科学研究費補助金 特定領域研究

新世代の計算限界
— その解明と打破 —

研究課題別成果報告書 分冊 A

平成20年 4月

領域代表者　岩間 一雄
京都大学 大学院 情報学研究科　教授

目 次

A01: オンライン予測の手法を用いた意思決定モデルに関する研究
瀧本 英二，酒井 義文，天野 一幸 1
A02: 計算理論的設計による知識抽出モデルに関する研究
徳山 豪，塩浦 昭義，河原林 健一，全 眞嬉 30
A03: 論理関数表現のモデルとシンボリックアルゴリズム
武永 康彦，湊 真一 63
A04: 組み合わせ最適化における指数サイズ・多項式時間近傍の設計
玉木 久夫，井口 幸洋 81
A05: ハードウェアアルゴリズムの性能評価に関する研究
高木 直史，高木 一義 105
A06: 幾何的配置問題の計算量限界と近似可能性の研究
加藤 直樹，田村 明久，大崎 純，瀧澤 重志 129
A07: 現実的な状況下での量子計算の能力に関する研究
山下 茂，中西 正樹 166
A08: 動的な構造をもつネットワーク上の資源割当て問題の研究
藤田 聡，中野 浩嗣，森本 康彦 214
A09: 非明示的表現に対するアルゴリズムの開発
篠原 歩，竹田 正幸，下薗 真一，坂本 比呂志，石野 明 248
A10: 変移する要素間の関係を条件とする組合せ最適化モデル
宮野 英次，朝廣 雄一，廣瀬 英雄 277

文部科学省科学研究費補助金 特定領域研究 (課題番号 16092201)

A01: オンライン予測の手法を用いた意思決定モデルに関する研究

本研究では，意思決定とデータの提示が交互に繰り返されるオンラインでの意思決定過程を統一して扱うことのできる普遍的なモデルの構築を目指し，幅広い見地から３つのアプローチを試みた．１つめは，計算学習理論の分野で開発されたオンライン予測の手法を手がかりにしようとするもので，従来モデルの拡張を行うとともに，オンライン予測の手法がいくつかのオンライン意思決定問題に適用できることを示した．２つめは，様々な現実の問題において求めるべき解の品質に関する基準を定式化し，本研究に還元しようとするもので，情報検索やゲノムのアラインメントなどの問題に対し，それぞれ有用な評価基準を与えるとともに，その基準を満たすアルゴリズムの開発を行った．３つめは，アルゴリズムが保持する知識表現の複雑さに基づいてオンライン意思決定の限界について考察しようとするもので，論理回路，決定木，しきい値回路などの様々な表現形式の複雑さの下界を与えた．

研究組織

研究代表者： 瀧本 英二　東北大学 大学院 情報学研究科
研究分担者： 酒井 義文　東北大学 大学院 農学研究科
天野 一幸　群馬大学 大学院 工学研究科

交付決定額 (配分額)

平成 16 年度	3,600,000 円
平成 17 年度	3,500,000 円
平成 18 年度	3,600,000 円
平成 19 年度	3,500,000 円
合　計	14,200,000 円

研究成果の概要

- 学術誌
  JASS (2004 年), JOIN (2005 年), SIAM J. Comput.(2005 年), TCS (2006 年), Algorithmica (2006 年), DAM (2007 年) など
- 国際会議
  ISAAC (Hong Kong, 2004 年), ALT (Padova, 2004 年)，ALT (Singapore, 2005 年), MFCS (Gdansk, 2005 年), EuroComb (Berlin, 2005 年), COCOON (Taipei, 2006 年), CCC (San Diego, 2007 年) など
- 受賞
  電子情報通信学会情報・システムソサイエティ活動功労賞 (瀧本，2005 年 9 月), 石田（實）記念財団研究奨励賞 (瀧本, 2005 年 10 月)
- 国際会議の開催
  ALT (瀧本, 大会委員長/実行委員として, 仙台, 2007 年)

1　はじめに

意思決定とデータの提示が交互に繰り返されるオンラインの状況での意思決定過程に関する問題は，機械学習，情報探索，情報圧縮，言語モデル，ゲーム理論，ファイナンス，ルーティング，ページングなど様々な領域に現れる普遍的なものである．また，インターネットやゲノム解析などの急速な発展に伴い，膨大なデータが逐次的に与えられる状況のもとで解決しなければならない新しい問題が次々と提起されている．これまで，このようなオンライン意思決定の問題について数多くの理論が展開されてきたが，それらは，それぞれの領域に固有の問題意識を動機付けとして，比較的独立に発展してきた．

本研究では，これらの問題を統一して扱うことのできる普遍的なモデルの構築，すなわち，様々なプロトコルや評価基準のもとで，理論的に保証された性能を持つ意思決定アルゴリズムの統一的な設計と解析方法を確立することを目指すものである．この目的を達成するため，我々は幅広い見地から次の３つのアプローチを試みた．

オンライン予測の手法の拡張

１つめのアプローチは，計算学習理論の分野で開発されたオンライン予測の手法が広い問題のクラスに対してほぼ最適な性能保証を与えることに注目し，オンライン予測の分野で確立されているアルゴリズムの設計と解析の理論を手がかりにしようとするものである．このアプローチに基づき，我々は主に以下の成果を得た（２章）．(2.1) 従来のオンライン予測モデルの１つである動的資源分散モデルを拡張し，投資の対象となるオプションの損失の範囲（リスク情報）がオプションごとに異なり，それらが予め知らされている場合についてアルゴリズムの設計と解析を行い，厳密な性能評価を与えた．(2.2) オンラインオークションと呼ばれる商品売買のモデルが，リスク情報つきの動的資源分配モデルに還元できることを示し，妥当な問題設定の下で，従来手法より優れた利得下界を持つアルゴリズムを与えた．(2.3) 複数の仮説を統合して予測性能の良い仮説を構築するブースティングの問題に対し，サンプルの分割と併合を繰り返すことによって決定ダイアグラムを構成する新しいスキームを得た．これは，従来異なる原理に基づくものと思われていた AdaBoost タイプと決定木生成タイプのブースティング手法を統一的に扱うための理論的枠組みを与えるものである．(2.4) 各ノードに否定のない単項式がラベルづけされた単調項決定リストと，複数の単調 DNF 式を排他的論理和で結んだ XOR-MDNF 式が等価となる条件を与え，その結果から，単調項決定リストに対する質問による学習アルゴリズムを導いた．(2.5) 非線形な幾何概念に対する学習の枠組みとして，ランダム写像に基づく新しい手法を提案した．(2.6) ブール関数に対する帯域フィルタの概念を定式化し，そのノイズ除去効果をフーリエ係数によって特徴付けることに成功した．

新しい評価基準に基づくモデル化

２つめのアプローチは，さまざまな現実の問題において，社会的あるいは学術的に求められている解の品質に関する基準を定式化し，モデル構築研究に還元しようとするものである．このアプローチに基づき，我々は主に以下の成果を得た（３章）．(3.1, 3.2) ネットワーク上の文書の新規公開及び改訂を即時に認識し，ユーザへの検索結果に直ちに反映させることのできる検索システムの構築を目指して，分散型の検索エンジンの開発を行った．特に，提示文書の順位付けとして，文書の新鮮さを加味した新しい文書採点法を提案し，実データによってその有効性を示した．(3.3) 複数個のアミノ酸配列の間に存在する共通構造のうち，特定の機能に関連するもののみを見出す問題を，最長共通部分配列問題を一般化した問題として定式化し，効率の良いアルゴリズムを与えた．(3.4) 複数本のゲノム同士のアラインメントを求める問題を，最長共通増加部分列問題として定式化し，線形領域アルゴリズムを与えた．(3.5) 複数の文字配列の多重アラインメントを求める問題を，閉路のない有向グラフにおけるトポロジカルソートが表す複数の文字配列と，別に与えられた１つの文字配列同士のすべての組み合わせにおいて，最も長い共通部分列を求める問題として定式化し，効率の良いアルゴリズムを与えた．(3.6) 航空旅客機内における映画の上映など，待ち時間においてユーザを退屈させないためのサービスの費

用対効果に関する問題を，オンライン意思決定問題として定式化し競合比解析を行った．

オンライン意思決定の計算論的限界の解明

３つめのアプローチは，意思決定アルゴリズムが保持すべき知識表現の複雑さに注目し，オンライン意思決定アルゴリズムの計算量的な限界について考察しようとするものである．このアプローチに基づき，我々は主に以下の成果を得た（４章）．(4.1) 否定素子数を限定した論理回路，２層しきい値回路，および論理式のサイズの下界を導出する新しい手法をそれぞれ与えた．(4.2) 否定素子数を限定した論理回路で，ソーティングや反転回路を構成する新しい手法を与えた．これらは，入力系列を k トニック列と呼ばれるものに限定すると，従来の構成法よりも優れている場合があることを示した．(4.3) 任意の二次関数を計算するサイズ最小の単調論理回路は，論理積素子を単層として含むという単層予想と呼ばれる未解決問題を，否定的に解決する重要な手がかりを与えた．(4.4) 乗算を計算するほぼ最適な OBDD の設計を与えた．(4.5) ある制約を満たす単調積和形論理式のうち，充足割り当ての個数を最小，あるいは最大にするものを特定することに成功した．(4.6) 論理回路のサイズの下界を求めるための，計算機支援による証明系の実現可能性について検討を行い，幾つかの知見を得た．(4.7) 方形描画を符号化に必要な符号長の限界値の導出を，超大規模行列の第一固有値を求める問題に帰着することにより，符号長の下界を与えた．(4.8) 回路計算量の下界を導出する従来手法は，k 乱雑性と呼ばれる論理関数の組み合わせ論的な性質に基づいていることを指摘し，その性質のみを利用するだけでは，従来の下界を改良できないことを示した．(4.9) 脳の計算機構をモデル化した閾値論理回路において，発火する閾素子の数の最大値をエネルギー消費量を表す新しい尺度として導入し，エネルギー消費量を制限した閾値論理回路の計算能力が著しく制約される場合があることを示した．

2　オンライン予測の手法の拡張

2.1　リスク情報つき動的資源配分問題

計算学習理論の分野で開発されたオンライン予測のモデルで最も適用範囲の広いものに，資源の動的配分問題があり，次のように定式化される．投資の対象（オプション）が K 個あるとし，投資の方法についてアドバイスを行う「エキスパート」が N 人いると仮定する．各時刻 $t = 1, 2, \ldots$ において，各エキスパート i は，投資ベクトル $x_{i,t} = (x_{i,t}(1), \ldots, x_{i,t}(K))$ を提示する．ただし，投資ベクトルは確率ベクトル，すなわち，各成分が非負で成分の総和が 1 であるとする．これらを受け，投資アルゴリズム（学習者と呼ぶ）は，自分の投資ベクトル $p_t = (p_t(1), \ldots, p_t(K))$ を定める．その後，各オプション j のコスト $l_t(j) \in [0, 1]$ が明らかとなる．このとき，学習者は，投資ベクトルとコストベクトルの内積 $p_t \cdot l_t = \sum_{j=1}^{K} p_t(j) l_t(j)$ の損失を被る．同様に，各エキスパート i も損失 $x_{i,t} \cdot l_t$ を被る．学習者の目標は，このプロセスをある時刻 T まで行った後の累積損失 $\sum_{t=1}^{T} p_t \cdot l_t$ を，最適なエキスパートの累積損失 $\min_i \sum_{t=1}^{T} x_{i,t} \cdot l_t$ に比べてそれほど大きくならないようにすることである．この問題に対し，ヘッジアルゴリズムと呼ばれる単純な重みつき平均アルゴリズムが，ほぼ最適な性能を与えることが知られている．

本研究では，各オプション j のコスト $l_t(j)$ がいずれも一様に 1 で上から抑えられているという仮定を緩め，その上界（リスク情報）b_j が予め与えられているというモデルを提案した．リスク情報は，競馬の投資におけるオッズ情報に対応するものであり，非常に自然な概念である．このモデルの下で，Vovk の統合アルゴリズム（Aggregating Algorithm）[5] を詳細に解析し，その性能の厳密な評価を与えた．その結果，リスク情報に大きな偏りがある場合，ある一定値以下のリスク情報は，アルゴリズムの性能に影響を与えないという意外な性質が明らかとなった．

また，リスク情報を用いないアルゴリズムとの性能の差は，リスクが一様のときに最大になることを示唆する結果を得た．これは，リスク情報が本質的に役に立たないかも知れないことを意味しており，投資戦略に対し，新たな知見を与えるものである．

2.2 オンラインオークション

Bar-Yossef らが提案したオンラインオークションモデルは，1 種類の商品を次々と来訪する入札者に売るための価格の決定過程をモデル化したものである [1]．このモデルでは，各入札者 $t = 1, 2, \ldots$ について，次のような取引が行なわれる．売り手が商品の価格 r_t を提示した後，入札者の入札金額 $m_t \in [1, h]$ が明かされる．ここでは簡単のため，入札金額の最低額を 1 に規格化し，入札金額の最高額を h と仮定する．このとき，$r_t \leq m_t$ ならば売り手は商品を売り利得 $g_t = r_t$ を得るが，$r_t < m_t$ ならば利得を得られない ($g_t = 0$)．目標は，時刻 T までの累積利得 $G_T = \sum_{t=1}^{T} g_t$ を，最適な価格 $r*$ で売り続けたときの売り上げ OPT に匹敵させることである．

この問題は次のように（利得型の）リスク情報つき動的資源配分問題へと還元される．まず，$[1, h]$ を離散化し，N 個の価格候補 $b(1) > b(2) > \cdots > b(N)$ を定め，これらをオプションと考える．また，各オプション（価格）$b(i)$ に対し，そのオプションにのみ全資源を投資するエキスパートを導入する．すなわち，エキスパート i（$1 \leq i \leq N$）の与える投資ベクトル $x_{i,t}$ は，常に，第 i 成分が 1 の単位ベクトルであるとする．これらの投資ベクトルを受け，学習者が時刻 t の投資ベクトル p_t を定めたとき，これをオプション集合上の確率分布とみなし，その分布に従って価格 r_t を選ぶ．すなわち，確率 $p_t(i)$ で $r_t = b(i)$ とする．そして，入札者 t の入札額 m_t が提示されると，各オプション i に対し，$b(i) \leq m_t$ ならば $l_t(i) = b(i)$，そうでないならば $l_t(i) = 0$ とおくことにより，コスト（利得）ベクトル l_t を定める．ここで，オプション i のリスク（利得の上界）が $b(i)$ となることに注意されたい．この変換のもとで，時刻 t の利得 g_t の期待値は，投資ベクトルとコストベクトルの内積 $p_t \cdot l_t$ で与えられ，エキスパート i（価格 $b(i)$）の利得も内積 $x_{i,t} \cdot l_t$ で与えられる．従って，価格候補集合を $\{b(1), \ldots, b(N)\}$ に限定したオンラインオークション問題は，この集合をオプション集合とするリスク情報つき動的資源分配問題とみなすことができ，前節で与えた投資アルゴリズムを適用することにより，価格候補集合の中で最適な価格 $b(i^*)$ で売り続けたときの累積利得 G_T^* に匹敵する累積利得を得ることができる．さらに，N を大きくして離散化レベルを十分細かくすることにより，真の最適価格 r^* による利得 OPT と $G_{i^*,T}$ の比を十分小さく抑えることができる．

このとき用いる動的資源配分問題の解析手法は，コストベクトル l_t が直積集合 $\{0, b(1)\} \times \cdots \times \{0, b(N)\}$ の任意の要素を取りうるとして最悪の場合の利得下界を求めるものである．しかし，オークションモデルは，入札金額 m_t 以下の価格を提示すれば必ずその価格で売れるという単調性を持つため，還元によって得られた動的資源配分問題においては，コストベクトル l_t は $\{(0, \ldots, 0, b(i-1), b(i), \ldots, b(N)) \mid 1 \leq i \leq N\}$ の N 通りしか取り得ない．本研究ではこのコストベクトル空間を用いて Vovk の統合戦略を厳密に適用することによって，アルゴリズムのさらなる改良を行い，次に示す期待利得の下界を得た．

便宜的に $b(0) = \infty$ とおき，確率ベクトル r と s を，それぞれ

$$
\begin{aligned}
r(j) &= \left(\frac{1}{b(j)} - \frac{1}{b(j-1)}\right) b(N), \\
s(j) &= \left(\frac{1}{\alpha^{b(j)} - 1} - \frac{1}{\alpha^{b(j-1)} - 1}\right) (\alpha^{b(N)} - 1)
\end{aligned}
$$

とおく．ここで，$\alpha > 1$ は適当な定数である．このとき，

$$
E[G_T] \geq \frac{b(N) \ln \alpha}{D(r||s) + b(N) \ln \alpha} \left(G_{i^*} - \frac{\ln N}{\ln \alpha}\right).
$$

ここに，$D(r||s)$ は，KL ダイバージェンスである．

価格候補集合を適当に定めると，この下界は，$h \leq 10^{100}$ 程度以下ならば，従来手法の下界よりも大きくなることを数値計算により確かめた．これは，現実的な問題設定のもとでの本手法の優位性を示すものである．

2.3 ブースティング

精度の低い複数の仮説を統合して予測精度の高い仮説を構築するブースティングの問題に対し，一つの新しい設計手法を与えた．ブースティングは，動的資源配分問題の双対問題という見方もできることが知られており，オンライン予測とは非常に関連が深い問題である．

ブースティングは，以下のような共通のスキームの元で論じることができる．サンプル

$\{(x_1, y_1), \ldots, (x_m, y_m)\} \subseteq X \times \{-1, 1\}$ が与えられていると仮定する．ここで，X は適当なインスタンス空間である．各時刻 $t = 1, 2, \ldots$ に対し，ブースターはサンプル上の分布 D_t を作り，分布 D_t のもとでのマージンの期待値が正の仮説（弱仮説）$h_t : X \to \{-1, 1\}$ を得る．すなわち，$\sum_{i=1}^{m} D_t(i) h_t(x_i) y_i > 0$．ある時刻 T において，ブースターは弱仮説 $h_1, \ldots, h_T$ を統合してマージンの最小値が正の仮説 F_T を構成する．すなわち，任意の i に対し，$F_T(x_i) y_i > 0$.

ブースティングには大きく 2 通りの設計手法が知られている．一つは，マージン空間上にポテンシャル関数 ϕ を導入し，$\sum_i \phi(F_T(x_i) y_i)$ が最小となることを目指すもので，AdaBoost，InfoBoost，LogitBoost などがこの設計手法に基づくものである [2]．もう一つは，ラベル空間の分布上にエントロピー関数を導入し，結合仮説の与える条件付エントロピー $H(Y|F_T)$ が最小となることを目指すものである．C4.5 などのトップダウン型決定木生成アルゴリズムや，これを改良して決定ダイアグラムを生成する MM アルゴリズムなどがこの設計手法に基づく [3].

従来，これらの 2 つの方法論はまったく異なるものだと思われていたが，本研究では InfoBoost がある意味で決定ダイアグラムを生成していることに注目し，これを一般化することによって，これら 2 つの方法論を統一的に論じることのできる新しい設計手法を開発した．すなわち，本研究で得られたブースティング手法は，AdaBoost が用いているポテンシャル関数 $\phi(x) = \exp(-x)$ に基づくタイプ 1 のブースティングアルゴリズムという見方もできるし，MM が用いているエントロピー関数 $G(p_1, p_2) = 2\sqrt{p_1 p_2}$ に基づくタイプ 2 のブースティングアルゴリズムという見方もできる．従って，本研究の成果は，ブースティングの設計手法に対し，統一的な理論的枠組みを与えたいうことができる．

2.4 単調項決定リストの学習

決定リストは，ブール関数の最も基本的な表現の 1 つであり，

$$f = ((f_1, a_1), \ldots, (f_d, a_d), (\mathrm{TRUE}, a_{d+1}))$$

のように表される．ここで，各 f_i は積和型論理式（DNF 式），a_i は 0 または 1 であり，一般性を失わず，$a_i \neq a_{i+1}$ とする．d を f の交代数と呼ぶ．決定リスト f は，ベクトル $v \in \{0, 1\}^n$ に対し，v が満たす最初の DNF 式が f_i のとき，a_i を値として取る．これまで，決定リストのさまざまな部分クラスの学習可能性が示されているが，本研究では，各項に否定リテラルを含まない単調項決定リストの学習可能性について考えた．単調項決定リストは，任意のブール関数を表現することができるため，決定リストの部分クラスとして，その学習可能性は特に重要である．

我々は，単調項決定リストのクラスと単調 DNF の排他的論理和式（⊕MDNF 式）との密接な関係に着目した．⊕MDNF 式は，ブール関数の単調関数への分解方法の 1 つとして提案されたもので，$f_1 > \cdots > f_d$ なる d (≥ 1) 個の単調 DNF 式を用いて，$f = f_1 \oplus \cdots \oplus f_d$ で表される．ただし，同じ項が式の中に 2 度以上現れることはないとする．ここで，d を式の深さと呼ぶ．⊕MDNF 式のクラスは，学習対象 f の積閉包複雑度と呼ばれる組み合わせ量に比例する時間で，所属質問と等価性質問を用いて学習可能であることが知られている [4]．f の積閉包複雑度 $\mathrm{cl}(f)$ とは，f を構成する各単調 DNF 式 f_i から高々 1 つずつ項を選び，それらを論理積で結ぶことにより得られる項全体からなる集合の要素数である．我々は，単調項決定リストに対する積閉包複雑度の概念を同様に定義し，任意の ⊕MNDF 式 f から，f と等価で積閉包複雑度が等しい単調項決定リスト f' が直ちに得られること，および，交代数が d の任意の単調項決定リスト f' に対し，f' と等価で $cl(f) \leq (cl(f')/d)^d$ を満たす ⊕MDNF 式 f が存在することを示した．さらに，これらの結果より，$O(\log n)$ 項単調項決定リストのクラス，および，交代数が定数の単調項決定リストのクラスが，所属質問と等価性質問を用いて多項式時間学習可能であることを導いた．

2.5 ランダムプロジェクションによる非線形概念の学習

ユークリッド空間中のベクトルとそれが正か負かを示すラベルの対の集合がサンプル $S = \{(x_1, y_1), \ldots, (x_m, y_m)\} \subseteq R^n \times \{-1, +1\}$ が与

えられており，正と負のベクトルを分離する仮説超曲面を求める問題について考える．いま，ある非線形写像 $\phi : R^n \to R^N$ が存在して，S の像 $\phi(S) = \{(\phi(x_i), y_i)\}$ が原点を通る超平面によって線形分離可能であると仮定する．ここで，各ベクトルの像 $z_i = \phi(x_i)$ を特徴ベクトル，R^N を特徴空間と呼ぶ．この仮定より，以降では，特徴空間において $\phi(S)$ を分離する超平面を求める問題について考える．

サンプルを分離する超平面 $h \in R^N$ に対し，その超平面から一番近い特徴ベクトルまでのユークリッド距離を h のマージンと呼ぶ．マージン理論より，マージンが大きいほど汎化誤差の上界が小さくなる．サポートベクトルマシン（SVM）は，特徴空間において，マージンが最大の分離超平面 h^* の双対表現を求める手法であり，その汎化誤差は，$\tilde{O}(1/(m\ell^2))$ となることが知られている．ここで，ℓ は h^* のマージンである．しかし，SVM では 2 次計画問題として知られる最適化問題を解く必要があり，その計算量を減らすことが，SVM の学習における中心的な課題のひとつとなっている．一方，2 次計画法の計算過程において，特徴ベクトルは内積（$K(x_i, x_j) = \phi(x_i)^\mathsf{T}\phi(x_j)$）の形でしか現れないため，カーネル関数 K が効率よく計算可能であれば，計算時間は特徴空間の次元数 N に依存しない．

本研究では，ランダム写像と呼ばれる技術を用いて，特徴空間（R^N）中のサンプルを低次元空間（R^d）に写像し，その低次元空間中で分離超平面を直接求める新しい学習方式について検討している．これまで知られているランダム写像は，各成分が独立なランダム $N \times d$ 行列に基づくもので，ユークリッド距離を高い確率で保存するという性質（Johnson-Lindenstrauss の補題）を持つ．この性質から，次元数を $d = O(1/\ell^2)$ とすると，サンプルの線形分離性（マージン）も高い確率で保存されることを導くことができる．VC 理論より，d 次元空間中でサンプルを線形分離する超平面の汎化誤差は $\tilde{O}(d/m) = \tilde{O}(1/(m\ell^2))$ なので，線形計画法によってそのような超平面を 1 つ求めることにより，SVM と同等の汎化能力を持つ仮説が，2 次計画問題を解くことなく効率よく得られることになる．しかし，従来のランダム写像を用いた場合，サンプルを d 次元空間に変換するのに N に依存する時間がかかってしまう．

本研究では，ランダム写像を与える行列 R を，n の多項式程度のビット数のランダムシード r を用いて構成する（$R = \psi(r)$）という学習の枠組みを提案する．この枠組みで効率のよい学習アルゴリズムを設計するためには，次の 2 つの条件が必要十分である．

条件 1 ランダム性の弱いランダム行列 $R = \psi(r)$ を用いて，サンプルのマージンを保存する性質を持つ低次元空間へのランダム写像を構築すること．

条件 2 各ベクトル x_i に対し，$\phi(x_i)$ と $\psi(r)$ を陽に計算することなく，x_i の像 $z'_i = K'(x_i, r) = R^\mathsf{T} z_i = \psi(r)^\mathsf{T}\phi(x_i)$ （SVM におけるカーネル関数に相当）が効率よく計算できること．

我々は，条件 1 を満たす 2 種類の新しいランダム写像の手法を与えることに成功した．以下に，そのうちの 1 つの構成法を述べる．

$\{-1, +1\}$ 上の k (≥ 4) 限定独立なランダム $N \times P$ 行列を Q 個 $R_1, \ldots, R_Q$ 選び，各 $1 \leq q \leq Q$ に対し，R^N から R^P への写像 $f_q(x) = R_q^\mathsf{T} x / \|R_q^\mathsf{T} x\|$ を定義する．ここに，各行列 R_q は k 限定独立であるが，$R_1, \ldots, R_Q$ は互いに独立となるように選ぶ．本研究で与えるランダム写像は，ベクトル $x \in R^N$ を，Q 個の P 次元ベクトル $f_1(x)$，…，$f_Q(x)$ を連結した $d = PQ$ 次元ベクトルへと写すものである．このとき，P と Q と適当に選ぶことにより，

$$d = \tilde{O}(1/\ell^{3+4/k})$$

で，このランダム写像は高い確率で線形分離性を保存する．また，このとき用いるランダムシードのビット数は，$\tilde{O}((\log N)/\ell)$ 程度であり，特徴空間の次元数 N が n の指数程度に大きい場合でも，ランダムシードの大きさは n の多項式で抑えられる．

2.6 ブール関数に対するフィルタのノイズ除去効果

実数上の関数は，さまざまな周波数を持つ単純な三角関数の組み合わせによって構成されているとみなすことができる．実際，電波や音声などの

時系列信号は，その信号を三角関数に分解することで信号源や伝送路などの性質を知ることができる．その分解をフーリエ変換と呼ぶ．フーリエ変換は，数学的には単なる座標変換であり，同じ関数を別の角度から眺める手段にすぎないが，信号処理の分野でフーリエ変換が有用なのは，現実に現れる時系列信号の多くが，単振動に基づく物理現象によって引き起こされているからである．一方，理論計算科学の分野で扱うブール関数には，現実の時系列信号のように，その関数の源になる単振動のような「自然な」基底関数は存在しない．したがって，フーリエ変換のような座標変換によって関数を表現するという試みは，1988 年に Kahn，Kalai，Linial の結果 [6] が発表されるまで，ほとんどなされてこなかった．Kahn らの結果は，Walsh 変換と呼ばれる有限体上の関数の直交変換をブール関数に適用し，関数の感受度をその “Walsh 係数” によって特徴付けたというものである．n 変数ブール関数 f の感受度 $s(f)$ とは，$x \in \{0,1\}^n$ と $i \in \{1,\ldots,n\}$ をそれぞれ一様ランダムに選んだとき，$f(x) \neq f(x^{(i)})$ となる確率として定義される．ここで，$x^{(i)}$ は x の第 i ビットを反転して得られるベクトルである．直感的には，$s(f)$ は，入力変数のどこかのビットが反転したときの f の値の変わりやすさを表している．Kahn らの結果は，Walsh 変換によって，f が n 次元ベクトル $\alpha \in \{0,1\}^n$ をパラメータとして持つ「基底関数」$\chi_\alpha : \{0,1\}^n \to \{-1,1\}$ の線形結合 $f(x) = \sum_{\alpha\in\{0,1\}^n} \hat{f}(\alpha)\chi_\alpha(x)$ として一意に表されることを指摘した上で，f の感受度と Walsh 係数 $\hat{f}$ との美しい関係 $s(f) = \frac{1}{n}\sum_{\alpha\in\{0,1\}^n} |\alpha|\hat{f}^2(\alpha)$ を見出したことである．ここで，$|\alpha|$ は α の L_1 ノルムであり，これを「周波数」α の高さとみなすとすると，この等式は，高周波成分の大きい関数ほど感受度が高いと解釈することができる．

この結果以降，この変換はブールドメイン上の「フーリエ変換」と呼ばれるようになり，さまざまな問題がフーリエ変換を通して解析されてきたが，それらの結果は，基底関数の正規直交性やフーリエ係数の組み合わせ論的な性質を解析することにより導かれたものであり，その手法は，信号処理の分野における従来のフーリエ解析とは，ほぼ独立に発展してきている．その理由として，ブールドメイン上のフーリエ変換は，信号処理の分野における周波数成分への分解というような物理的な意味を持っていないことが考えられる．一方，長い歴史を持つ信号処理の分野には，フーリエ解析について膨大な知識の蓄積があり，それらの結果に対応する概念をブールドメイン上で見出すことにより，計算機科学においても有用な新しい知見が得られる可能性がある．このような信号処理の分野からの類推研究はこれまでほとんどされておらず，わずかにサンプリングの定理や高速フーリエ変換に対応する概念が見出されているにすぎない．そこで，本研究では，信号処理の分野で大きな役割を果たしているフィルタを取り上げ，ブール関数に対するフィルタの概念の定式化を試み，その性質を調べた．

フィルタは，信号処理の分野では，入力信号の各成分を強調もしくは減衰し，適切な性質を持つ信号に変換する処理全般を指し，信号の平滑化，エッジ検出，復調，倍音除去等，あらゆる用途で用いられている．その中でも，本研究では，フィルタの最も基本的な役割の１つであるノイズ除去の機能に着目する．ここで，ブール関数 $f : \{0,1\}^n \to \{0,1\}$ に対するノイズモデルとして，各入力 $x \in \{0,1\}^n$ に対して独立に，ある確率 p で $f(x)$ の値を反転するものを考える．

我々は，フィルタを帯域フィルタと呼ばれるものに限定した場合のノイズ除去効果の大きさを，元の関数のフーリエ係数を用いて評価する公式を２つ与えた．１つめの評価式は，低域フィルタが，自然なブール関数のクラスに対するノイズ除去フィルタとして，極めて有効であることを示している．２つめの評価式は，フィルタと対象のブール関数から定まるマージンと呼ばれる幾何学的なパラメータによって，ノイズ除去効果の大きさが特徴づけられることを示している．これらの結果は，最適なフィルタの設計法の指針を与えると同時に，与えられたブール関数のノイズ耐性を評価する方法を与えている．

3　新しい評価基準に基づくモデル化

3.1　協調検索エンジンの開発について

ネットワーク上の新規文書の公開及び既公開文書の改訂を即時に認識し，ユーザへの検索結果に

直ちに反映させることのできる検索システムの設計及び構築を目指して，互いに協調する複数個のサーバによって構成される分散型の検索エンジンCSE(Cooperative Search Engine) の開発を進めており，アーキテクチャの観点からの対故障性の向上に関する技術を提案し，その有効性を検証した．また，検索結果の提示法として，公開や改訂の時刻に依存した文書の新しさを加味した文書採点法を提案した．

CSE の概要は以下の通りである．CSE は，ネットワーク上の各局所サーバに存在する局所検索エンジン LSE(Local Search Engine) と，ユーザからの検索要求に応じて適切な LSE に局所検索を依頼する位置サーバ LS(Location Server) からなる．

LSE は，周期的に自身が存在する局所サーバ L おいて新規に公開された文書及び改訂された既公開文書を検知し，L がもつ文書を局所検索する際に用いる索引を更新する．集中型の検索エンジンでは，他の局所サーバにおいて公開されている文書をロボットによって収集した後に索引を更新しなければならないため，新規文書や改訂された文書の内容を短い間隔で索引に反映させることは難しい．これに対して，LSE は L に存在する文書のみを対象とするため，短い周期で頻繁に索引を更新することが可能である．また，LSE は，索引の更新の後に，索引に現れるすべての語を，L 内においてその語が現れる任意の文書に対する最大のスコアとともに LS に送信する．これをメタ索引という．メタ索引は，ユーザからの検索要求の際に，LS が局所検索を依頼する LSE の個数を絞り込む目的で用いられる．更に，LSE は LS からの局所検索の依頼に応じて，L に存在する文書を検索し，結果を LS に返す．

LS は，キーワードをそれが文書中に現れるか否かを表す論理変数にみたてた論理式の形でユーザから指定される検索クエリに対して，各 LSE から送信された索引に存在する語に関する情報を利用することによって，局所検索をする LSE の個数を絞り込む．たとえば，一つのキーワードのみからなる検索クエリが与えられた場合に，そのキーワードが索引に現れない LSE には検索対象となる文書が局所サーバに存在しない．したがって，そのキーワードが索引に現れる LSE のみに局所検索の依頼をすれば十分である．検索クエリが一般の論理式 f である場合も，$f \leq g$ であるような単調な論理式 g を求め，g に現れるキーワードを，それが文書に現れるか否かを表す論理変数とみなすのではなく，それが LSE の索引に現れるか否かを表す論理変数とみなすことによって，$g = 0$ である LSE の局所サーバに $f = 1$ である文書が存在しないことが保証される．ただし，f から最小の g を求めるのは NP 困難であるため，実際には最小な g ではなく，これを近似する単調論理式を用いて，局所検索を依頼する LSE の個数を絞り込む．また，検索結果としてスコアの高い順に n 個の文書を検索するような場合に，OR 演算のみの論理式が検索クエリとして与えられるならば，キーワードを含む文書の最大のスコアに関する情報を用いることで，局所検索を依頼する LSE の個数を高々n に絞り込むことができる．LS は，ユーザからの検索クエリに対して，選ばれた LSE に局所検索を依頼し，各 LSE から返された結果を統合したものを検索結果としてユーザに出力する．

一般に，分散型の検索エンジンは，検索時にサーバ間の通信が生じるため，集中型の検索エンジンと比較して，検索結果を得るまでに時間がかかる．CSE では，上に述べた LS による局所検索を依頼する LSE の個数の絞り込みのほかに，様々なキャッシュ技術を用いて，検索結果を得るまでの時間の短縮化を実現している．ただし，現状ではインターネット規模のネットワークでの検索エンジンとして用いるのは難しく，たとえば社内イントラナットなどの，100 台以下の局所サーバからなる比較的小規模なネットワークでの利用を想定している．以上が，CSE の概要である．

CSE は複数のサーバからなる分散型の検索エンジンであるものの，1 台の LS が他のすべての LSE に接続し，ユーザからの検索要求に対して必ず LS を用いて検索結果を得る．したがって，LS が故障した場合に，CSE のすべての検索サービスが停止する．そこで，LS を複数台用いて冗長な構成に変更することにより，対故障性の向上を試みた．

複数台の LS による CSE の構成の概要は以下の通りである．各 LS は，LSE から送信されるメタ索引を共有しなければならない．しかし，各 LSE がすべての LS に同じ情報を送信するのは，ネッ

トワーク上のメッセージ数を増加させ，輻輳の原因となる．そこで，各 LSE は最寄の故障していない 1 台の LS にのみメタ索引を送信する．また，メタ索引についても，送信メッセージの大きさを小さく抑えるために，前回の更新時の索引からの差分のみを，前回の更新時刻と今回の更新時刻からなるバージョン情報と，送り先である LS のみからなる送信済み LS のリストとともに送信する．メタ索引の差分を受け取った LS は，送信済み LS のリストに隣接する LS をすべて加え，送信済みでない送信可能なすべての LS に，受け取ったメタ索引の差分を転送する．なお，メタ索引の差分を受け取った LS は，もし現在保持しているメタ索引のバージョンと，受け取った差分に付随するバージョンに不整合がある場合に，送り主に差分ではなくメタ索引そのものを送信するように要求する．この状況は，LS が故障から復帰した場合に起こり得る．以上により，故障していないすべての LS によって，LSE のメタ索引が等しく共有される．したがって，すべての LS が故障しない限り，CSE のサービスは停止することなく提供される．

各 LS の故障率に対するシステム全体の停止率をシミュレーションによって調べたところ，LS の台数よりも各 LS に隣接する LS の個数の影響が強いことが確認された．ただし，各 LS に重複なくランダムに選ばれた指定された個数の LS が隣接するネットワークトポロジーを用いた．16 台の LS に対して完全グラフをネットワークトポロジーとして用いた場合においては，各 LS の故障率が 90%であっても，システム全体の停止率は 20%程度であり，実用として許容されるレベルの対故障性を有するシステムを実現できることが確認された．

CSE では，検索における文書の順位付けに，TF·IDF の値に基づくスコアを用いている．ここで，TF は，その文書に現れるキーワードの頻度を表し，IDF は，そのキーワードが現れる文書の頻度の逆数の対数を表す．したがって，検索クエリに複数のキーワードが指定されている場合に，検索対象となる文書全体で希にしか現れないキーワードが数多く現れる文書ほど高いスコアが与えられる．このスコアに基づく順位付けを用いた場合に，公開からの時間経過による新鮮さという意味での文書の価値の劣化は，検索結果に反映されない．そこで，文書に各キーワードが現れてからの経過時間に応じて TF の値を減衰させることによって定義される FTF を用いた，FTF·IDF の値に基づくスコアを提案した．

提案した FTF の定義は以下の通りである．t_i を，文書が公開された後の LSE による i 回目の更新の時刻とする．時刻 t_i における TF，FTF の値をそれぞれ，TF_t，FTF_t とする．このとき，

$$\begin{aligned}\mathrm{FTF}_1 &= \mathrm{TF}_1 \\ \mathrm{FTF}_i &= \mathrm{FTF}_{i-1} \cdot \alpha e^{-(t_i - t_{i-1})} \\ &\quad + (\mathrm{TF}_i - \mathrm{TF}_{i-1})\end{aligned}$$

である．この定義のもとで，Web 上にて文書を公開している特定のサイトに対して，1 ヶ月にわたり各文書における FTF の値の推移を観察したところ，新しくかつキーワードに関連性の高い文書が上位に順位付けられる傾向が確認された．

CSE を設計する上での今後の重要な課題は，LSE 同士のメタ索引の交換により，各 LSE が LS の部分的な役割を担う仕組みを組み込むことである．時間の経過とともに変化する各局所サーバの文書集合に対して，どの LSE がどの LSE のメタ索引をもつことが検索の高速化につながるかの判断については，オンライン予測の手法が有用となる状況が存在するものと予想される．

3.2 新鮮な情報検索のスコアリングについて

ウェブ検索における新鮮さを重視する文書のスコアリングとしてこれまでに提案した FTF · IDF に関して，改良と実データを用いた性能の評価実験を行った．

文書集合 D における単語 w と文書 d に対する FTF · IDF は，以下のように定義される．TF は，単語 w が文書 d に現れる回数である．IDF は，単語 w が少なくとも 1 回以上現れる D 中の文書の個数の D 中のすべての文書の個数に対する割合の対数である．したがって，頻繁に用いられるありふれた単語ほど小さな値が設定される．多くの検索エンジンにおいて，TF · IDF をスコアとする文書のランク付けが用いられている．FTF は，

TF を用いて，

$$\mathrm{FTF} = \sum_{j=1}^{\mathrm{TF}} \exp\left(-\frac{t-t_j}{\alpha}\right) \cdot \mathrm{TF}$$

と定義される．ここで，t は現在の時刻，t_j は文書 d の先頭から j 番目に現れる単語 w が更新などによって文書 d に初めて生じた時刻である．α は，FTF が時刻の経過するにしたがってどの程度減衰するかを指定する因子である．この定義より，単語 w が生じてから時間がたつほど，FTF は小さな値が設定されることになる．したがって，FTF · IDF によるランク付けのもとでは，検索に用いた単語が最近に生じた文書ほど上位にランクされる．なお，上に定義した FTF を正確に計算するには，文書に現れるすべての単語に対してそれぞれタイムスタンプを保存し，また，更新の前後で各単語の対応関係を正確に判定しなければならない．これを検索システムに実装して用いるのはあまり現実的ではないため，実際には，正確ではないものの代わりの定義として，

$$\begin{aligned}
\mathrm{FTF}_0 &= \mathrm{TF}_0, \\
\mathrm{FTF}_i &= \mathrm{FTF}_{i-1} \cdot \exp\left(-\frac{t_i - t_{i-1}}{\alpha}\right) \\
&\quad + \mathrm{TF}_i - \mathrm{TF}_{i-1} \\
&\quad (i \geq 1 \text{ のとき})
\end{aligned}$$

を提案している．ただし，t_0 は文書 d が初めて公開された時刻，t_i は文書 d の i 回目の更新の時刻を表し，FTF_i，TF_i はそれぞれ，時刻 t_i における FTF，TF の値を表す．

上の定義より，単語 w に関する検索に対して FTF·IDF で文書のランク付けをした場合，$\alpha = \infty$ ならば，従来の TF · IDF によるランク付けと同様に，文書中に単語 w が現れる回数が多いものほど上位にランク付けされる．一方，α が 0 に近いならば，最近の更新において単語 w が現れた文書ほど上位にランク付けされるものの，その文書に更新以前から存在している単語 w の個数はランク付けにあまり考慮されない．したがって，単語 w を多く含むという意味で単語 w に対する関連性の高く，かつ，公開されてまだあまり時間が経っていない w に関する内容を含むという意味で新鮮である文書を上位にランク付けするには，α をどのように設定するかが重要である．

そこで，文書集合全体における TF の総和が大きく変化する平均的な周期と，平均的な変化の総和の大きさから適応的に α の値を設定する方法を考案し，更新の周期の異なる実在する 2 つのウェブサイトをそれぞれ文書集合全体とみなして，FTF · IDF のランク付けを試した．この結果，どちらのウェブサイトにおいても，最も新しい文書が最上位にランク付けされる検索単語の種類の個数が飽和する境界の α が，考案した方法によって効率よく得られることを確認した．

3.3 弱制約最長共通部分配列問題について

複数個の配列（sequence）の間の共通の部分配列（subsequence）の中で長さが最大のものを求める最長共通部分配列問題は，配列の間に存在する類似性や共通構造を見出す問題として，計算機科学や分子生物学など様々な分野での応用が考えられる．この問題に関する研究の一環として，最近，任意に指定された配列 P が部分列として含まれるという条件のもとで，配列 X，Y の最長の共通部分配列を求める制約最長共通部分配列問題が Tsai[12] によって提案され，この問題を解く $O(|X| \cdot |Y| \cdot |P|)$ 時間アルゴリズムが Chin ら [8] によって与えられた．

制約最長共通部分配列問題は，例えば，分子生物学において，アミノ酸配列 X，Y に部分配列として共通に含まれる配列 P が両者から得られるタンパク質に共通する特定の機能に関与していることが解っているときに，この前提のもとで更なる共通構造を見出すことに利用できるものと期待できる．ただし，制約として指定される P が冗長な配列であり，実際には P に含まれるより短い部分配列がタンパク質の機能に関与しているような場合には，制約が強すぎることにより短い共通配列しか得られず、有用な共通構造が見落とされる可能性がある．そこで，制約最長共通部分配列問題を更に一般化して，制約である配列 P の他に，パラメタ $0 \leq r \leq |P|$ を与えて，P の少なくとも長さ r の部分列を部分列として含む X と Y の最長の共通部分配列を求める状況を想定し，これを効率よく解くための手法について考えた．以降では，この問題を，弱制約最長共通部分配列問題ということにする．$r = 0$ のときの弱制約最長共通

部分配列問題は，従来の最長共通分配列問題に等しく，$r = |P|$ のときの弱制約最長共通部分配列問題は，制約最長共通分配列問題に等しい．

弱制約最長共通部分配列問題を解く一つの方法は，P の長さ r の部分配列それぞれに対して，これを制約配列としたときの制約最長共通部分配列問題を $O(|X| \cdot |Y| \cdot r)$ 時間で解き，得られた配列の中で最も長いものを選ぶことである．ただし，この方法では，$\binom{|P|}{r}$ 個存在する P の長さ r の部分配列の中から制約最長共通部分配列が最長となるものをしらみつぶしに探さなければならないため，実行時間に関して非常に効率が悪い．そこで，X，Y，P の各長さの接頭辞に対する 3 次元の表を用いる動的計画法に基づく Chin ら [8] の制約最長共通部分配列問題に対するアルゴリズムを，更に P の部分配列の長さ $0 \leq q \leq r$ に関する次元を加えた 4 次元の表を用いるアルゴリズムに変形することによって，$O(|X| \cdot |Y| \cdot |P| \cdot r)$ 時間で弱制約最長共通部分配列問題を解くアルゴリズムを提案した．

また，P の部分配列の長さ $0 \leq q \leq r$ に関する次元を加える代わりに，P の部分配列の長さ $l - r \leq q \leq l$ に関する次元を加えることによって，$O(|X| \cdot |Y| \cdot |P| \cdot (l - r))$ 時間で弱制約最長共通部分配列問題を解くアルゴリズムを提案した．ただし，l は，X，Y，P の最長共通部分配列の長さである．このアルゴリズムは，たとえ $l < |P|$ であっても，$r = l$ に対して $O(|X| \cdot |Y| \cdot |P|)$ 時間で動作するという意味で，Chin ら [8] の制約最長部分配列問題に対する $O(|X| \cdot |Y| \cdot |P|)$ 時間アルゴリズムの機能を，漸近的な時間計算量を変化させることなく拡張したものとなっている．

提案した後者のアルゴリズムは，例えば $l = |P|/2$ の場合に，$\binom{|P|}{|P|/2}$ 個存在する P の長さ $|P|/2$ のどの部分配列が弱制約最長共通部分配列に部分配列として含まれるかに関する事前の知識なしに，$O(|X| \cdot |Y| \cdot |P|)$ 時間で弱制約最長共通部分配列を求める．一方，制約最長共通部分配列問題は，事前にどの配列が含まれるかについての知識が事前に与えられる場合の弱制約最長共通部分配列問題とみなすことができるものの，この問題に対する $O(|X| \cdot |Y| \cdot |P|)$ 時間よりも真に高速なアルゴリズムは知られていない．この意味で，共通部分配列に含まれる配列そのものに関する事前の知識がアルゴリズムの漸近的な時間計算量に実質的にどのように影響しているのかを検討することは，今後の課題の一つである．

3.4 最長共通増加部分配列問題の線形領域解法について

最長共通増加部分配列問題は，与えられた複数個の整数配列において共通に含まれる部分配列の中で，各要素である整数の値が先頭から末尾に向かって真に増加する最長の部分配列を求める問題である．ただし，配列に含まれる部分配列とは，元の配列の任意の位置から任意の個数の要素を削除して得られる配列のことをいう．この問題は，古くから知られている単一の整数配列に対する最長増加部分配列問題を，複数個の配列を扱う形に一般化した問題である．また，同じく古くからよく知られている複数個の配列に対する最長共通部分配列問題とも密接な関係があり，この関係に基づく応用の一例として，複数本のゲノム同士のアラインメントを求める際の，最長共通増加部分配列問題の解の利用が考えられる．

2 つの整数配列 X，Y に対する最長共通増加部分配列問題に対しては，Yang らによって提案された $O(|X| \cdot |Y|)$ 時間 $O(|X| \cdot |Y|)$ 領域アルゴリズムが知られている [13]．そこで，本研究では，アルゴリズムが計算に用いる領域が，2 つの配列の長さの積に比例する大きさではなく，長さの和に比例する大きさで十分であるような，$O(|X| \cdot |Y|)$ 時間 $O(|X| + |Y|)$ 領域アルゴリズムを提案した．

配列 X，Y に対する最長共通部分配列に対しては，$O(|X|)$ 領域を用いて $O(|X| \cdot |Y|)$ 時間で解くアルゴリズムが古くから Hirschberg によって与えられている [9]．このアルゴリズムでは，配列 Y を半分の長さの接頭辞と接尾辞に分割し，Y の接頭辞に対しては，X の各長さの接頭辞との最長共通部分配列の長さを $O(|X| \cdot |Y|)$ 時間 $O(|X|)$ 領域で求め，同様に Y の接尾辞に対しては，X の各長さの接尾辞との最長共通部分配列の長さを求める．そして，得られた X の各接頭辞と Y の接頭辞の最長共通部分配列の長さと，X からその接頭辞を取り除いて得られる接尾辞と Y の接頭辞の最長共通部分配列の長さの和が最大になるように，X を接頭辞と接尾辞へ分割する．この結果と

して，X と Y の最長共通部分配列問題は，X と Y の接頭辞の最長共通部分配列問題と接尾辞の最長共通部分配列問題に分割されるため，再帰的にこの分割を繰り返すことによって解を求める．

この分割統治法に基づく手法は、整数配列 X, Y の最長共通増加部分配列問題を線形領域で解く際にも利用できる．ただし，X と Y それぞれの接頭辞の最長共通増加部分配列の長さと，それらの接頭辞を取り除くことによって得られる X と Y それぞれの接尾辞の最長共通増加部分配列の長さの和の最大値は，必ずしも X と Y の最長共通増加部分配列の長さに一致するわけではない．なぜなら，接頭辞に関する最長共通増加部分配列の末尾の整数が接尾辞に関する最長共通増加部分配列の先頭の整数よりも大きいか等しい場合には，これら二つの配列の連接は，各要素である整数が先頭から末尾に向かって真に単調に増加するという条件を満たさないからである．

上記の問題点は，X と Y のある最長共通増加部分配列が部分配列として現れる Y 上の位置の中で Y の長さが半分の接尾辞の最も先頭よりの位置に現れる整数を z とすると，この z が現れる位置で Y を接頭辞と接尾辞に分割し，また，同じ最長共通増加部分配列が部分配列として現れる X 上の位置における z が現れる位置を，X における接頭辞と接尾辞への分割位置とすることによって回避できる．この X と Y それぞれの接頭辞、接尾辞への分割のもとで，X と Y の接頭辞については末尾が z 未満である最長共通増加部分配列を求め，一方，接尾辞については先頭が z を超える最長共通増加部分配列を求めれば，前者に整数 z を連接し，これに後者を連接することによって，X と Y 全体の最長共通増加部分配列が得られるからである．X と Y の各長さの接頭辞の対に対して，上記の分割位置を長さが 1 だけ短い接頭辞に関する値から帰納的に求める漸化式を与えることによって，Hirschberg [9] の分割統治法に基づく $O(|X| \cdot |Y|)$ 時間 $O(|X| + |Y|)$ 領域アルゴリズムを得た．

3.5 トポロジカルソート配列と配列の最長共通部分配列アルゴリズムについて

同じ記号が同じ列に揃わなければならないという条件のもとで複数個の配列からなる最短の（列の個数が最も少ない）多重アラインメントを求めるのは NP 困難であることが知られている [10]．一方，多重アラインメントの各列をそれに現れる記号とみなすことによって 1 個の配列と扱う場合に，2 個の多重アラインメントからなる最短のアラインメントは多項式時間で求めることができる．このことから複数個の配列からなる短い多重アラインメントを求めるための手法として，対象となる配列の集合から 2 個を選んで最短のアラインメントを求め，これを選ばれた 2 個の代わりに集合に戻すことを繰り返すことによって全体の多重アラインメントを求める累進アルゴリズムが広く用いられている．

多重アラインメントのいくつかの列の位置を入れ換えたときに，アラインメントを構成する各配列に矛盾が生じないならば，特に理由がない限り，列の入れ換えによって得られる多重アラインメントをもとの多重アラインメントと同等に扱うのが自然であろう．このような同等な多重アラインメントは見取り図グラフ [11] のトポロジカルソートが表す配列として得られるため，一般に複数個（場合によっては，多重アラインメントの長さに関する指数関数個）存在する．例えば，3 個の配列 aacc，bacac，abbbc の 2 個の多重アラインメント

```
a-a-c--c     a--a-c-c
-ba-c-ac ,   -b-a-cac
ab-b-b-c     abb-b--c
```

において，右のアラインメントは左のアラインメントの各列を 1, 2, 4, 3, 6, 5, 7, 8 の順に並べ替えることによって得られるから，両者は同等な多重アラインメントであり，それぞれ配列 ababcbac, abbabcac と扱われる．この例における多重アラインメントの見取り図グラフを図 1に示す．グラフの各頂点 v_i は左のアラインメントの第 i 列に対応する．

一般に累進アルゴリズムにおいて，多重アラインメントは他の同等なアラインメントを考慮することなく，任意に選ばれた 1 個のアラインメント

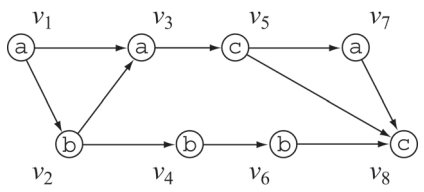


図 1: 見取り図グラフ

として扱われる．しかし，多重アラインメントを，その見取り図グラフのトポロジカルソートが表すすべての同等なアラインメントの中で，相手とのアラインメントを求めたときに最短のものが得られるものとして扱うならば，より短いアラインメントが最終的に得られる可能性がある．この方針に基づく累進アルゴリズムを実現するためには，与えられた 2 個の多重アラインメントの見取り図グラフ（各頂点に記号を付された有向非循環グラフ）に対して，それぞれのトポロジカルソートが表す配列同士のすべての組み合わせにおいて最も長い最長共通部分配列をもつものを求める効率のよい計算手法が必要となる．

本研究では，与えられる見取り図グラフには，頂点に付されている各記号に対して，そのグラフ上に経路が存在して，その記号を付された頂点はすべてその経路上に存在するものと仮定し，更に，与えられる 2 個の見取り図グラフの一方は 1 つの経路のみからなるグラフ，すなわち，1 個の配列そのものである仮定した場合に，最も長い最長共通部分配列をもつトポロジカルソートの組を求める $O(|G|^2 \cdot |B|^4)$ 時間アルゴリズムを提案した．ただし，$|B|$ は 1 つの経路のみからなる見取り図グラフの頂点の個数を表し，$|G|$ はもう一方の見取り図グラフの頂点の個数を表す．前者の仮定は，多重アラインメントのいくつかの列を一つにまとめることによってより短い多重アラインメントが得られることがないという意味で，非冗長なアラインメントのみを対象としていることを意味する．累進アルゴリズムにおいて，選ばれた 2 個のアラインメントから最短のアラインメントが求められるという観点から，この仮定をおくことは現実的である．

3.6 読捨てコンテンツ更新タイミングのオンライン意思決定について

待ち時間を伴うサービスにおける読捨てコンテンツの提供はユーザを退屈させないための手法の一つである．銀行や理髪店の待合室における週刊誌の提供や，航空旅客機内における映画の上映などはその一例である．通常，ユーザを退屈させないという目的のもとでは，過去に閲覧したことのある読捨てコンテンツは，再度の閲覧においてその目的を果たさない．一方，古いコンテンツを新しいものに更新する際には，一般にコストが生じる．したがって，退屈させないことがサービスを継続的に利用するユーザの獲得に影響する状況のもとでは，どのようなタイミングでコンテンツを更新するかが問題となる．

このような状況を，ユーザによる同一コンテンツの再閲覧によって生じるコストを 1，古いコンテンツを新しいコンテンツに更新する際に生じるコストを m としたときに，u 人のユーザによる閲覧要求列に対してコストの総和を小さく抑えるための更新のタイミングを求める問題としてモデル化し，この問題を解く決定性オンライン更新アルゴリズムの競合比に関して評価した．ただし，閲覧要求列において複数のユーザは同一時刻に要求をしないものと仮定し，また，更新アルゴリズムは最初の閲覧要求が起こる前に必ず 1 回コンテンツを更新し，コンテンツの更新と閲覧要求は同一時刻に起こらないものと仮定した．

はじめに，競合比の下界に関して，どのような決定性オンライン更新アルゴリズムであっても，最適なオフライン更新アルゴリズムのコストの総和と比較して，漸近的に少なくとも，

- $u = \infty$ ならば，2 倍，
- $4m+1 \leq u < \infty$ ならば，$\dfrac{2\left\lfloor \dfrac{u-1}{2m} \right\rfloor - 1}{\left\lfloor \dfrac{u-1}{2m} \right\rfloor + \dfrac{1}{2}}$ 倍，
- $2m+5 \leq u \leq 4m$ ならば $\dfrac{3}{2 + \dfrac{m}{2\left\lfloor \dfrac{u-1}{4} \right\rfloor}}$ 倍

のコストの総和を生じさせることを示した．

また，ユーザの人数に制限のない $u = \infty$ の場合に，漸近的に最適なオフライン更新アルゴリズムのコストの総和のたかだか 2 倍のコストの総和で更新をするという意味で最適な決定性オンライン更新アルゴリズムとして，以下の 2 つを提案した．

アルゴリズム 1：前回の更新の後に起こった再閲覧要求の回数がちょうど m になった直後に更新をする．

アルゴリズム 2：前回の更新の後に複数回閲覧要求をしたユーザの人数がちょうど m になった直後に更新をする．

特に，後者のアルゴリズムに関しては，漸近的に最適なオフライン更新アルゴリズムのコストの総和に対して，$m \leq u < 2m$ ならば，たかだか u/m 倍，$u < m$ ならば，たかだか 1 倍のコストの総和で更新をすることを示した．

更に，これらの二つのアルゴリズムを比較し，m が大きいときに，ある閲覧要求列に対しては前者のアルゴリズムのコストの総和が後者のアルゴリズムのコストの総和の 2 倍に近くなるのに対して，後者のアルゴリズムのコストの総和はどのような閲覧要求列に対しても前者のアルゴリズムのコストの総和のたかだか 1.5 倍に抑えられることを示した．また，m が大きくかつ $u \geq 2m$ の場合に，後者のアルゴリズムのコストの総和が実際に前者のアルゴリズムのコストの総和の 1.5 倍に近くなるような閲覧要求列が存在することを示した．

最後に，動的計画法に基づくオフライン更新アルゴリズムを提案し，このアルゴリズムが漸近的に閲覧要求列の長さの 2 乗に比例する時間で，任意に与えられた閲覧要求列に対して最小のコストの総和を実現する更新のタイミングを求めることを示した．

4 論理関数の複雑さに基づくオンライン意思決定の計算論的限界の解明

個々の論理関数を，種々の計算モデルを用いて表現する際に必要なサイズを評価する問題，特にそのサイズに対する下界を導出する問題は，計算量理論における中心的かつ挑戦的課題の一つである．これら表現のサイズは，様々な意思決定アルゴリズムが保持する知識表現の構築の困難さを特徴付ける尺度としても使用できることから，本研究においても基本的研究課題として位置づけている．本研究では，論理回路，論理式，決定木，順序付き二分決定グラフといった，理論上あるいは，実用上重要な各計算モデルにおいて，種々の論理関数を表現する際に必要なサイズの下界を導出する手法の開発を目指した．また，これらの表現を用いて，実用上有用な様々な論理関数に対するコンパクトな表現を求めることによって，既知の上界を改善する研究も合わせて行った．

4.1 回路計算量の下界導出手法の開発

単調論理回路と一般の論理回路との中間モデルにあたる，制限された個数の否定素子を用いることのできる，否定数限定論理回路を用いた場合の計算複雑さの下界導出手法の開発を行った．これについては，NP 完全に属するクリーク関数と呼ばれる論理関数を計算するためには，否定素子の使用を $O(\log\log n)$ 個程度許した場合にも，超多項式的なサイズの論理回路を要することを明らかにし，権威ある欧文雑誌の一つである SIAM Journal on Computing において発表した．否定素子の個数を $\log n$ 個にまで増加させた回路に対しても同様の下界を得ることができれば，直ちに P≠NP を意味することから，この結果は，前述した P≠NP 予想の解決へ向けた一つのステップと位置づけられる．

更には，回路中に現れる各素子の出次数を高々 1 に制限した論理回路に相当する，論理式を用いた際の論理関数の表現サイズの下界導出手法の開発にも取り組んだ．これに対しては，近年，量子回路において論理関数を評価する際に必要なオラクル呼び出しの回数の下界を導出する手法として開発された量子敵対者法が，論理式サイズの下界導出手法としても利用可能であるという最新の報告を受けて，本手法の詳細な解析と更なる拡張の可能性について研究を行った．本手法は，ある種の半正定値問題に対する実行可能解の存在が，ある一定値以下のサイズの論理式の不存在性を保証するという著しい性質を利用したものであり，これによって，論理式サイズの下界の証明を構成的に行うことが可能となる．本研究では，本手法を

適用する際に得られる半正定値問題とその双対問題における許容解空間の性質に対するいくつかの性質を明らかにし，本手法の更なる拡張への手掛かりを得た．

しきい値回路とは，線形しきい値素子を用いて構成される非周期的回路である．しきい値回路は，単純化した脳のモデルとして，その計算能力について，古くより盛んな研究が行われている．本研究では，2 層のしきい値回路に着目し，これを用いて論理関数を計算する際に必要なサイズの下界を導出する手法の開発を目指した．これに対して，与えられた論理関数を計算する 2 層しきい値回路のサイズの下界の導出を，ある法則によって与えられる線形計画問題の目的関数の最適解の導出に帰着する，新たな手法を提案発表した．また，提案手法を，2 層しきい値回路で計算が困難であることが強く予想されている内積関数に実際に適用し，得れらた大規模な線形計画問題を高い計算能力を持った計算機を用いて解くことによって，これまで同種の回路に対して得られていた最良の下界に匹敵する指数関数的下界を導出することに成功した．

4.2 論理関数の計算における否定素子の効果

論理回路モデルにおいては，計算の複雑さは通常，回路のサイズや深さで評価されるが，それらは，使用できる否定素子の個数を制限すると，制限しない場合に比べて大きいオーダーの量を必要とすることがある．本研究では，そのようなギャップのある問題として，入力ビット系列を昇順に（0*1* の形に）並び変えるソートと入力ビット系列の 1 の補数を求めるビット反転問題を取り上げ，否定素子数と回路のサイズとの間の関係を調べた．このとき，入力ビット系列 $x = (x_1, x_2, \ldots, x_n)$ において，値が 0 から 1 へ，あるいは 1 から 0 へ変化する回数 $A(x) = |x_2 - x_1| + \cdots + |x_n - x_{n-1}|$ をパラメータとして解析を行った．以下では，入力ビット系列 x を，$A(x) \leq k$ を満たす系列（k トニック列と呼ぶ）に限定する．

ソートについては，$k = O(1)$ で否定素子数を $\log\log n$ に制限したとき，サイズが，ある定数 $c \in (0,1)$ に対して $O(n(\log n)^{1-c})$ となる回路を構成した．マージは $k = 2$ におけるソート問題とみなせるので，この結果はマージに対する同様の結果 [14] の拡張となっている．

ビット反転問題については，否定素子数を $O(k \log n)$ に制限したとき，深さが $O(k \log n)$ でサイズが $O(kn)$ の回路を構成した．この結果は，$k = o(\log n)$ のときは，従来の構成法 [15] を改善したものとなっている．さらに，$k = O(1)$ のときは，我々の構成法は本質的にこれ以上改善できない最適なものとなっている．

4.3 単層予想の否定的解決

明示的に与えられた論理関数に対して，これを計算する論理回路のサイズに対する有効な下界を導出する問題は，理論計算機科学分野最大の未解決問題である P≠NP 予想の解決にも直結する重要な課題と位置づけられている．本研究では，この問題の解決の手掛かりを得るために，種々の制約を設けた論理回路における複雑さの下界の導出手法の開発を行うことを目指した．

単調論理回路とは，論理積素子と論理和素子のみからなる組み合わせ論理回路である．単調論理回路は，論理否定素子を用いることのできる一般の論理回路に比べ，表現できる関数が単調制の制約を満たすものに限定されるという制約を受けるものの，実用上現れる多くの関数は，単調性を有しており，それゆえ，汎用的モデルと言うことができる．本研究では，特に二次の論理関数，すなわち，丁度 2 個の変数の積の論理和として表される論理関数を取り上げ，論理回路の構造的性質と，関数の計算に要する回路サイズとの関係に焦点をあて研究を行った．その結果，任意の二次関数を計算するサイズ最小の単調論理回路は，常に単層回路—入力から出力へ至る任意の道上に論理積素子が高々 1 個しか出現しない回路—であることを予想した，いわゆる単層予想と呼ばれる十数年来の未解決問題に対し，多出力関数を考えた場合には，この予想が偽であることを示し，部分的に解決した．この予想は，その後，Jukna[17] によって，単出力関数の場合にも偽であることが証明され，完全な解決を見た．本研究における結果は，本予想解決への足がかりとなる有効なものであったと考えられる．

4.4 乗算を計算する OBDD の設計

順序付き二分決定グラフ (Ordered Binary Decision Diagram, 以下 OBDD) は，近年活発に研究されている論理関数の表現形式である．OBDD は，各節に論理変数が割り当てられた決定二分木であり，かつ変数の出現する順序が根から各葉に至る道上で一様であるという制約を満たすものをいう．OBDD は，変数の出現順序を固定した場合には，その表現の最小化や，等価性判定が効率良く行うことが出来，この性質を利用して，VLSI の正当性の検証や，学習に際しての知識表現等の場面で幅広く用いられている．

OBDD は，変数順序を上手く選択することによって，数多くの実用上重要な論理関数を，非常にコンパクトに表現することができる一方，基本的な幾つかの関数は，いかなる変数順序の元でも，その表現サイズが指数関数的となることが知られている．二進数で与えられた整数同士の掛け算は，この様な性質を持つ最も代表的な関数である．90 年代初頭に Bryant によって証明されたこの事実は，以来，計算機科学分野において最も多く引用される結果の一つとなったが，掛け算を実現する最小の OBDD のサイズの具体的な値，および，その実現方法については，明らかとなっていなかった．

本研究では，この状況を打破すべく，二進 n 桁同士の掛け算を OBDD で表現する手法について理論と実験の両面から評価した．その結果，ベンチマークとして最も良く用いられる，n 桁同士の二進整数の掛け算の中心の桁を計算する，サイズが $O(2^{6n/5})$ である OBDD の構成手法を与えることに成功した．これは，従来知られる最良の上界を改善するものである．この OBDD は，入力として与えられる二つの数を下位桁から交互にプローブすることにより得られる．すなわち，入力 $X = x_n x_{n-1} \cdots x_1$，$Y = y_n y_{n-1} \cdots y_1$ に対し，変数順序を $x_1, y_1, x_2, y_2, \ldots, x_n, y_n$ として構成した OBDD が上記の上界を達成する．

また，比較的小さな入力桁数に対して，掛け算を表現する最適な OBDD を実際に計算機実験により構成することで，上で示した上界の性能を評価した．その結果，n 桁の掛け算の中心桁を実現する最適な OBDD のサイズは $2^{6n/5}$ にほぼ比例すること．すなわち，本研究で得た構成手法は，ほぼ最適なサイズの OBDD を与えることを強く示唆する結果を得た．

4.5 単調 DNF 式の充足割り当て数の評価

ある制約を満たす単調積和型論理式 (MDNF 式) のうち，充足割り当ての個数を最大にする，あるいは最小にするものを特定する問題にも取り組んだ．これは，機械学習における知識表現の際にも広く用いられる MDNF 式に対して，極限的性質を持った式を特徴付けることを通じて，学習の効率化などに貢献しようという試みである．本研究では，極値集合論，あるいは極値グラフ理論といった分野において開発されてきた様々な手法が，この種の問題に適用できることを明らかにし，特に，変数の個数に比べて項の数が少ない，いわゆる疎な MDNF 式に対しては，上記の性質を持つ式を特定することに成功した．

4.6 計算機支援下界証明系の開発にむけて

明示的に与えられた関数を計算する論理回路のサイズの下界を与える，計算機支援による証明の実現可能性に対する検討を行った．四色問題の解決に端を発する計算機支援証明は，近年の著しい計算機性能の向上のもと，より一層の広がりを見せることが予想されている．本研究では，岩間らによる論理回路サイズに対する既知の最良の下界を与える証明 [16] を，計算機支援のもとで再構成することを試み，その実現可能性や，より強い下界導出手法へと発展させる際に必要となる技術に対する幾つかの知見を得た．

4.7 大規模計算に基づく方形描画のコード長の下界の導出

計算機支援による実際の下界証明の適用例として，C04 版の中野 (群馬大学) と共同で方形描画の最適コード長に関する研究を行った．方形描画とは，長方形をより小さな長方形に分割した描画を指し，この表現は VLSI のフロアプラン設計等に非常に幅広く用いられることから，コンパクトで扱いやすい符号化等について広く研究が行われている．大規模なフロアプランを扱う場合には，よ

り短いコードに符号化することが求められるが，その符号長の情報論的限界値については非常に基本的な問題であるにもかかわらず未解決問題となっている．

本研究では，方形描画の符号長の限界値の導出を，超大規模行列の第一固有値を求める問題に帰着する手法を提案した．より具体的には，まず，方形描画の家系木 T を構成する．家系木 T は，その k 層目の各頂点に，面数 k の描画が抜けも重複もなく割り当てられた木で，これは各描画に対して上手に親子関係を定義することで構成することができる．

次に，同じ構造を子木に持つ頂点を併合した多重辺グラフ G を考える．面数 n の描画の個数は，こうして得られたグラフ G の長さ $n-1$ の道の総数で与えられる．しかし，このグラフ G の頂点数も深さに対して指数関数的に増加するので，厳密な数え上げは依然困難である．そこで，本研究では G のある隣接する 2 層に着目し，これを 2 部グラフとしてみたとき，もとのグラフの道の総数の増加割合は，この 2 部グラフの隣接行列の第一固有値で下から押さえられることを示した Perron-Frobenius の定理を適用し，道の総数の下界の導出を行った．

実際，上記の手続きにより得られた約 1200 万次の平方行列の第一固有値を計算機実験により求めることで，面数 n の方形描画の個数は $\Omega(11.56^n)$ であることを示した．このことから直ちに，方形描画に対する任意の 2 進コードは $3.53n$ ビット以上の長さとなることが導かれる．これは，現在知られる最良のコード長である $5n$ ビットのコード化手法の改善余地が既に 30%未満であることを示している．

4.8 従来型回路計算量下界導出手法の限界の解明

1 または 2 入力論理関数のうち，排他的論理和とその否定を除く任意の演算が計算可能である素子からなる論理回路を U_2 回路と呼ぶ．明示的に定義された論理関数に対して，それを計算する U_2 回路のサイズに対するできるだけ大きな下界を導出しようという問題は，理論計算機科学の挑戦的課題として多くの研究者を引き付けてきた．1990 年代初頭の Zwick による $4n$ の下界の導出のあと，約 10 年間の停滞が見られたが，2001 年 Lachish と Raz がこれを $4.5n$ に改良し，その直後に岩間と森住 [16] が更に $5n$ に改良した．これら 2 点の改良は，Strongly-Two-Dependent 性や，更に強い性質である k-乱雑性という性質を持った論理関数の構造に着目することによって得られたものである．

論理関数 n が k-乱雑であるとは，任意に選ばれた k 変数に対する任意の異なる 2 つの部分割り当てが，異なる $n-k$ 変数論理関数を誘導することを言う．上で述べた直近 2 点の下界は，対象とする論理関数が $n-o(n)$-乱雑である場合には，全て適用可能であることが知られている．これまで，$n-o(n)$-乱雑な論理関数の実現については，$O(n\log n)$ サイズの論理回路によるものが最良であり，したがって，現在最良の下界 $5n$ との間には，まだ大きなギャップが存在していた．

本研究では，$n-o(n)$-乱雑性を満たす論理関数が，サイズ $5n+o(n)$ の U_2 論理回路で実現できることを構成的に証明した．以下に，その関数の具体的定義を示す．

入力変数の個数を n とする．簡単の為，n は $\log^2 n$ によって割り切れるものとする．$p(n)$ を n より大きな，かつ，最小の素数とする．自然数 s を入力とし，1 から n の整数を出力とする関数 w_n を次のように定義する．

$$w_n(s) \quad = \quad \begin{cases} t, & 1 \le t \le n \text{ のとき}, \\ 1, & \text{それ以外}. \end{cases}$$

ここで，$t = s \bmod p(n)$ とする．入力変数 $X = \{x_1, \ldots, x_n\}$ を各ブロック $\log^2 n$ 個の変数からなる $\ell = n/\log^2 n$ 個のブロックに分割し，各ブロックに含まれる変数のパリティを $p_1, \ldots, p_\ell$ とする．n 変数論理関数 f を以下の様に定義する．

$$f(x_1, \ldots, x_n) = x_z,$$

ただし，

$$z = w_n\left(\sum_{i=1}^{\ell} i \cdot p_i\right).$$

以上で定義した関数 f は $n-o(n)$-乱雑で，かつ，サイズ $5n+o(n)$ の U_2 回路で構成可能である．

本結果は，論理関数に対する新たな性質の発見無しには，上記の下界を改良することが不可能であることを意味し，同時に，論理関数の高い複雑性を担保する，新たな性質の発見に対する強い動機付けを与えるものである．本成果はB03班の垂井 (電通大) との共同研究によるものである．

4.9 エネルギー消費量に基づく脳の計算機構のモデル化

通常の半導体素子とは違い，ニューロンは，発火する（1を出力する）ときに消費するエネルギー量が，発火しない（0を出力した）ときに比べて大きいという性質がある．一方，脳全体の単位時間あたりのエネルギー代謝量は限られているので，脳におけるニューロンの発火活動は，極めて疎であると考えられている．にもかかわらず，実際に脳は複雑な計算を行っているように見える．このことは，脳の計算機構をモデル化して計算過程を論じる際，エネルギー消費量をパラメータの１つとして導入すべきであることを示唆している．本研究では，脳の計算機構モデルとして最も単純な閾値回路を考える．閾値回路とは，線形分離関数を計算する閾素子を単位素子とする組み合わせ論理回路である．さて，脳における上述の性質から，閾値回路の発火活動を制限したとき，その計算能力はどの程度の影響を被るのかという自然な疑問が生じる．閾素子や閾値回路は，意思決定アルゴリズムが保持する知識表現として広く用いられているため，この問は，本研究にとって基本的かつ重要なものである．Uchizawa らは，閾値回路が消費するエネルギー量を，その回路において同時に発火しうる閾素子の数の最大値と定義し，これをその回路のエネルギー複雑度と名付けた [7]．そして，n 変数関数の広いクラスが，$O(\log n)$ のエネルギー複雑度を持つ多項式サイズの閾値回路で計算できることを示した．そのクラスには，線形分離関数が分類規則として各内部頂点に割り当てられた決定木などが含まれる．この結果は，素子の発火活動が疎でも，回路は複雑な計算を行えることを意味している．

一方，我々は，計算量理論の見地から，エネルギー複雑度を制限することによる否定的な側面に着目し，閾値回路のエネルギー複雑度と計算能力の関係を調べた．その結果，それらの間にトレードオフがあることを意味する不等式の導出に成功した．特に，エネルギー複雑度を $n^{o(1)}$ に制限し，さらに深さを定数に制限すると，ある自然な条件を満たす論理関数を計算するには，指数的な素子数を必要とすることを示した．そのような関数の１つに，内積関数 $IP(x,y) = x_1y_1 \oplus x_2y_2 \oplus \cdots \oplus x_ny_n$ がある．内積関数は，エネルギー複雑度を $O(n)$ に緩めると，深さが３で線形サイズの閾値回路で計算できるので，この結果は，エネルギー消費量の制限が，閾値回路の計算能力に本質的な影響を与えていることを示している．

参考文献

[1] Z. Bar-Yossef, K. Hildrum and F. Wu, Incentive-Compatible Online Auctions for Digital Goods, *Proceedings of 13th SODA*, 964–970, 2002.

[2] J. Friedman, T. Hastie, and R. Tibshirani, Additive logistic regression: a statistical view of boosting, *Annals of Statistics*, 2, 337–374, 2000.

[3] Y. Mansour and D. A. McAllester, Boosting using branching programs, *Journal of Computer and System Sciences*, 64(1), 103–112, 2002.

[4] E. Takimoto, Y. Sakai, and A. Maruoka, The learnability of exclusive-or expansion based on monotone DNF formulas, *Theoretical Computer Science*, 241, 37–50, 2000.

[5] V. Vovk, A game of prediction with expert advice, *Journal of Computer and System Sciences*, 56(2), 153–173, 1998.

[6] J. Kahn, G. Kalai, N. Linial, The influence of variables on Boolean functions, *Proceedings of the 29th Annual IEEE Symposium on Foundations of Computer Science*, 66–80, 1988.

[7] K. Uchizawa, R. Douglas and W. Maass, Energy complexity and entropy for threshold

circuits. *Proceedings of 33rd ICALP*, 631–642, 2006.

[8] F.Y.L. Chin, A. De Santis, A.L. Ferrara, N.L. Ho, S.K. Kim, A simple algorithm for the constrained sequence problem, *Information Processing Letters*, 90, 175–179, 2004.

[9] D.S. Hirschberg, A linear space algorithm for computing maximal common subsequence, *Communicatoins of the ACM*, 18(6), 341–343, 1975.

[10] D. Maier, The complexity of some problems on subsequences and supersequences, *Journal of the ACM*, 25, 322–336, 1978.

[11] J. Meidanis, J.C. Setubal, Multiple alignment of biological sequences with gap flexibility, *Lecture Notes in Computer Science*, 911, 411–426, 1995.

[12] Y.T. Tsai, The constrained longest common subsequence problem, *Information Processing Letters*, 88, 173–176, 2003.

[13] I.H. Yang, C.P. Huang, K.M. Chao, A fast algorithm for computing a longest common increasing subsequence, *Information Processing Letters*, 93, 249–253, 2005.

[14] K. Amano, A. Maruoka and J. Tarui, On the Negation-Limited Circuit Complexity of Merging, *Discrete Applied Mathematics*, 126(1), 3–8, 2003.

[15] R. Beals, T. Nishino and K. Tanaka, More on the Complexity of Negation-Limited Circuits, *SIAM Journal on Computing*, 27(5), 1334–1347, 1998.

[16] K. Iwama and H. Morizumi, An Explicit Lower Bound of $5n - o(n)$ for Boolean Circuits, *Proceedings of 27th MFCS*, 353–364, 2002.

[17] S. Jukna, Disproving the Single Level Conjecture, *SIAM Journal on Computing*, 36(1), 83–98, 2006.

研究業績一覧

学術論文

1. Eiji Takimoto, Syuhei Koya, Akira Maruoka:
"Boosting Based on Divide and Merge", Proceedings of the 15th International Conference on Algorithmic Learning Theory (ALT 2004), Lecture Notes in Artificial Intelligence, 3244, 127–141, Oct. 2004.
概要: InfoBoost is a boosting algorithm that improves the performance of the master hypothesis whenever each weak hypothesis brings non-zero mutual information about the target. We give a somewhat surprising observation that InfoBoost can be viewed as an algorithm for growing a branching program that divides and merges the domain repeatedly. We generalize the merging process and propose a new class of boosting algorithms called BP.InfoBoost with various merging schema. BP.InfoBoost assigns to each node a weight as well as a weak hypothesis and the master hypothesis is a threshold function of the sum of the weights over the path induced by a given instance. InfoBoost is a BP.InfoBoost with an extreme scheme that merges all nodes in each round. The other extreme that merges no nodes yields an algorithm for growing a decision tree. We call this particular version DT.InfoBoost. We give an evidence that DT.InfoBoost improves the master hypothesis very efficiently, but it has a risk of overfitting because the size of the master hypothesis may grow exponentially. We propose a merging scheme between these extremes that improves the master hypothesis nearly as fast as the one without merge while keeping the branching program in a moderate size.

2. Nobuyoshi Sato, Minoru Udagawa, Minoru

Uehara, Yoshifumi Sakai, Hideki Mori:
"Redundant Architecture in Cooperative Search Engine", Journal of Applied Systems Studies, vol. 5, no. 2, pp. 157–168, Jul., 2004.
概要: We have developed a distributed search engine, called Cooperative Search Engine (CSE), in order to retrieve fresh information. In CSE, a local search engine located in each Web server makes an index of local pages, and a meta search server integrates these local search engines in order to realize a global search engine. This meta server is a single point of failure in CSE. Therefore, we propose the introduction of redundancy for the meta search servers in order to increase the availability of CSE. In this paper, we describe and evaluate a reliable architecture for CSE.

3. Nobuyoshi Sato, Minoru Uehara, Yoshifumi Sakai:
"A Case Study on Freshness based Scoring for Fresh Information Retrieval", Proceedings of IEEE International Symposium on Communications and Information Technologies 2004, 210–215, Oct. 27, 2004.
概要: For most businesses, fresh information retrieval is very important. However, it is difficult for conventional search engines based on centralized architecture to retrieve really fresh information, because they take a long time to collect documents via Web robots. In contrast to a centralized architecture, a search engine based on a distributed architecture does not need to collect documents, because each site independently makes an index. As this result, distributed search engines can retrieve really fresh information. However, fast indexing is not enough to easily retrieve fresh information. The value of information is determined by both freshness and relevance. Traditional ranking methods consider either freshness or relevance; so, we proposed FTF·IDF (Fresh Term Frequency multiplied by Inverse Document Frequency) as a scoring method that considers both freshness and relevance. In this paper, we will describe a verification of FTF·IDF on an actual web diary.

4. Kazuyuki Amano, Akira Maruoka:
"On the Monotone Circuit Complexity of Quadratic Boolean Functions", Proc. of the 15th International Symposium on Algorithms and Computation (ISAAC 04), Lecture Notes in Computer Science, 3341, 28–40, Dec. 2004.
概要: Several results on the monotone circuit complexity and the conjunctive complexity, i.e., the minimal number of AND gates in monotone circuits, of quadratic Boolean functions are proved. We focus on the comparison between single level circuits, which have only one level of AND gates, and arbitrary monotone circuits, and show that there is a huge gap between the conjunctive complexity of single level circuits and that of general monotone circuits for some explicit quadratic function. Almost tight upper bounds on the largest gap between the single level conjunctive complexity and the general conjunctive complexity over all quadratic functions are also proved. Moreover, we describe the way of lower bounding the single level circuit complexity, and give a set of quadratic functions whose monotone complexity is strictly smaller than its single level complexity.

5. Eiji Takimoto, Naomitsu Aizawa, Akira Maruoka:
"On Proper Learning for Monotone Term Decision Lists from Queries", Proc. Workshop on Learning with Logics and Logics for Learning, 45–51, Jun. 2005.
概要: A monotone term decision list is a decision list where each node is labeled with a conjunction of literals with no nega-

tion. We characterize the class of monotone term decision lists in terms of another class of Boolean formulas called the XOR expansions of monotone DNF formulas. The relationships we give imply the proper learnability for some subclasses of monotone term decision lists. In particular, monotone term decision lists with $O(\log n)$ terms and monotone term decision lists with a constant number of alternations are both polynomial time properly learnable from membership and equivalence queries.

6. 酒井義文:
"弱制約最長共通部分配列問題", 電子情報通信学会論文誌, D-I, Vol. J88-D-I, No. 8, pp. 1261–1263, 2005 年 8 月.
概要: 配列 X, Y の共通部分配列（subsequence）の中で，制約として与えられた配列 P の長さ r の部分配列を部分配列として含む，長さが最大の配列の一つを求める $O(|X| \cdot |Y| \cdot |P| \cdot r)$ 時間アルゴリズム及び $O(|X| \cdot |Y| \cdot |P| \cdot (l-r))$ 時間アルゴリズムを提案する．ただし，l は X, Y, P の最長共通部分配列の長さである．

7. Kazuyuki Amano, Akira Maruoka:
"On the Complexity of Depth-2 Circuits with Threshold Gates", Proc. of the 30th Mathematical Foundations of Computer Science (MFCS05), Lecture Notes in Computer Science 3618, 107-118, Aug. 2005.
概要: The paper investigates the complexity of depth-two circuits with threshold gates and consisting of two parts. First, we develop a method for deriving a lower bound on the size of depth two circuits with a threshold gate at the top and a certain type of gates at the bottom. We apply the method for circuits with symmetric gates at the bottom that compute the "inner product mod 2", and obtain a lower bound of 1.3638^n. Although our lower bound is slightly weaker than the best known lower bound of $\Omega(2^{n/2}/n)$, which was recently proved by Forster et al., our method has unique features: A lower bound is obtained by solving a certain linear program, and solving larger linear programs yield higher lower bounds. We also discuss the generalization of the proposed method. Second, we develop a simplified simulation of a depth-one threshold circuit with unbounded weights by a depth-two threshold circuit with small weights. Precisely, we give an explicit construction of depth-two circuits with small weights consist of $\tilde{O}(n^5)$ gates that compute an arbitrary threshold function. We also give the construction of such circuits with $O(n^3/\log n)$ gates computing the COMPARISON and CARRY functions, and that with $O(n^4/\log n)$ gates computing the ADDITION function. These improve the previously known constructions on its size and simplicity.

8. N. Sato, M. Uehara, Y. Sakai:
"Evaluation of Freshness Considering Scoring on Fresh Information Retrieval", Journal of Interconnection Networks, Vol. 5, No. 3, pp. 265–282, 2005 年 9 月.
概要: Fresh information is important for real business. In order to realize fresh information retrieval, we need not only to collect documents in a short time, but also to rank the results in the suitable order. However, conventional ranking methods are not suited for fresh information retrieval because they ignore temporal value of information. So, we have proposed the novel ranking method FTF·IDF for fresh information retrieval. FTF · IDF extends TF · IDF by means of using FTF (Fresh Term Frequency) instead of TF (Term Frequency). FTF differs from TF because FTF decreases as time goes. The speed of decreasing FTF is determined by the dumping factor. The dumping factor is sensitive against small changes of documents. So, we use a thresh-

old to ignore such small changes. In some papers we published, we detect the optimal threshold manually. In this paper, we proposed an adaptive calculating method in order to detect threshold automatically. In this method, the optimal value is determined by iterating to test generated thresholds. In this paper, we describe our method and its evaluation.

9. Kazuyuki Amano, Jun Tarui:
"Monotone Boolean Functions with s Zeroes Farthest from Threshold Functions", Proc. of the 3rd European Conf. on Combinatorics, Graph Theory and Applications (EuroComb05), Discrete Mathematics and Theoretical Computer Science, Vol. AE, 11-16, Sep. 2005.
概要: Let T_t denote the t-threshold function on the n-cube: $T_t(x) = 1$ if $|i : x_i = 1| \geq t$, and 0 otherwise. Define the distance between Boolean functions g and h, $d(g, h)$, to be the number of points on which g and h disagree. We consider the following extremal problem: Over a monotone Boolean function g on the n-cube with s zeros, what is the maximum of $d(g, T_t)$? We show that the following monotone function p_s maximizes the distance: For $x \in 0, 1^n$, $p_s(x) = 0$ if and only if $N(x) < s$, where $N(x)$ is the integer whose n-bit binary representation is x. Our result generalizes the previous work for the case $t = n/2$ and s=2^{n-1} by Blum, Burch, and Langford [FOCS98], who considered the problem to analyze the behavior of a learning algorithm for monotone Boolean functions, and the previous work for the same t and s by Amano and Maruoka [ALT02].

10. Shigeaki Harada, Eiji Takimoto, Akira Maruoka:
"Online Allocation with Risk Information", Proc. 16th International Conference on Algorithmic Learning Theory, Lecture Notes in Artificial Intelligence, Vol. 3734, 343–355, Oct. 2005.
概要: We consider the problem of dynamically apportioning resources among a set of options in a worst-case online framework. The model we investigate is a generalization of the well studied online learning model. In particular, we allow the learner to see as additional information how high the risk of each option is. This assumption is natural in many applications like horse-race betting, where gamblers know odds for all options before placing bets. We apply the Aggregating Algorithm to this problem and give a tight performance bound. Surprisingly, the Hedge algorithm without seeing risk information performs nearly as well as the Aggregating Algorithm. So the risk information does not help much. We discuss some interesting behaviors of the loss bound.

11. Kazuyuki Amano, Akira Maruoka:
"A Superpolynomial Lower Bound for a Circuit Computing the Clique Function with at Most $(1/6)\log\log n$ Negation Gates", SIAM Journal on Computing, 35(1), 201–216, Oct.2005.
概要: In this paper, we investigate the lower bound on the number of gates in a Boolean circuit that computes the clique function with a limited number of negation gates. To derive strong lower bounds on the size of such a circuit we develop a new approach by combining three approaches: the restriction applied to constant depth circuits due to Håstad, the approximation method applied to monotone circuits due to Razborov, and the boundary covering developed in the present paper. We prove that if a circuit C with at most $\lfloor(1/6)\log\log m\rfloor$ negation gates detects cliques of size $(\log m)^{3(\log m)^{1/2}}$ in a graph with m vertices, then C contains at least $2^{(1/5)(\log m)^{(\log m)^{1/2}}}$ gates. No non-

trivial lower bounds on the size of such circuits were previously known, even if we restrict the number of negation gates to be a constant. Moreover, it follows from a result of Fischer (Lecture Notes in Comput. Sci., 33 (1974), pp. 71–82) that if one can improve the number of negation gates from $\lfloor (1/6)\log\log m \rfloor$ to $\lfloor 2\log m \rfloor$ in the statement, then we have P $\neq$ NP. We also show that the problem of lower bounding the negation-limited circuit complexity can be reduced to the one of lower bounding the maximum of the monotone circuit complexity of the functions in a certain class of monotone functions.

12. Kazuyuki Amano, Akira Maruoka:
"On Learning Monotone Boolean Functions under the Uniform Distributions", Theoretical Computer Science, Vol. 350(1), pp. 3–12, Jan. 2006.
概要: In this paper, we prove two general theorems on monotone Boolean functions which are useful for constructing a learning algorithm for monotone Boolean functions under the uniform distribution. A monotone Boolean function is called fair if it takes the value 1 on exactly half of its inputs. The first result proved in this paper is that a single variable function $f(x) = x_i$ has the minimum correlation with the majority function among all fair monotone functions. This proves the conjecture by Blum, Burch and Langford (1998, Proc. 39th FOCS, pp.408–415) and improves the performance guarantee of the best known learning algorithm for monotone Boolean functions under the uniform distribution they proposed. Our second result is on the relationship between the influences and the average sensitivity of a monotone Boolean function. The influence of variable x_i on f is defined as the probability that $f(x)$ differs from $f(x \oplus e_i)$ where x is chosen uniformly from $\{0,1\}^n$ and $x \oplus e_i$ means x with its i-th bit flipped. The average sensitivity of f is defined as the sum of the influences over all variables x_i. We prove that a somewhat unintuitive result which says if the influence of every variable on a monotone Boolean function is small, i.e., $O(1/n^c)$ for some constant $c > 0$, then the average sensitivity of the function must be large, i.e., $\Omega(\log n)$. We also discuss how to apply this result to the construction of a new learning algorithm for monotone Boolean functions.

13. 瀧本英二:
"オンライン予測の理論に基づく意思決定", 電子情報通信学会論文誌 A, Vol. 89, No, 6, 405–418, 2006 年 6 月.
概要: 計算学習理論の枠組みの一つとして開発され発展してきたオンライン予測の方法論は，情報圧縮，資源配分，ゲームの戦略選択，線形回帰，ネットワークルーチングなど，様々な領域における意思決定問題に適用できることが明らかになってきた．本解説では，この普遍性を支えるオンライン予測アルゴリズムの設計と解析の原理を，いくつかの応用例とともに紹介する．

14. Shigeaki Harada, Eiji Takimoto and Akira Maruoka:
"Online Allocation with Risk Information", IEICE Transactions on Information and Systems, E89-D (8), 2340–2347, 2006.
概要: We consider the problem of dynamically apportioning resources among a set of options in a worst-case online framework. The model we investigate is a generalization of the well studied online learning model. In particular, we allow the learner to see as additional information how high the risk of each option is. This assumption is natural in many applications like horse-race betting, where gamblers know odds for all options before placing bets. We apply Vovk's Aggregating Algorithm to this problem and give a tight performance bound. The results support our intuition that it is safe to bet

more on low-risk options. Surprisingly, the loss bound of the algorithm does not depend on the values of relatively small risks.

15. Yoshifumi Sakai:
"A linear space algorithm for computing a longest common increasing subsequence", Information Processing Letters, Vol. 99, No. 5, pp. 203–207, 2006 年 7 月.
概要: Let X and Y be sequences of integers. A common increasing subsequence of X and Y is an increasing subsequence common to X and Y. In this note, we propose an $O(|X| \cdot |Y|)$-time and $O(|X| + |Y|)$-space algorithm for finding one of the longest common increasing subsequences of X and Y, which improves the space complexity of Yang et al. [Inform. Process. Lett. 93 (2005) 249–253]'s $O(|X| \cdot |Y|)$-time and $O(|X| \cdot |Y|)$-space algorithm, where $|X|$ and $|Y|$ denote the lengths of X and Y, respectively.

16. Shigeaki Harada, Eiji Takimoto and Akira Maruoka:
"Aggregating Strategy for Online Auctions", Proceedings of the 12th Computing and Combinatorics Conference (COCOON 2006), Lecture Notes in Artificial Intelligence, 4112, 33–41, Aug. 2006.
概要: We consider the online auction problem in which an auctioneer is selling an identical item each time when a new bidder arrives. It is known that results from online prediction can be applied and achieve a constant competitive ratio with respect to the best fixed price profit. These algorithms work on a predetermined set of price levels. We take into account the property that the rewards for the price levels are not independent and cast the problem as a more refined model of online prediction. We then use Vovk's Aggregating Strategy to derive a new algorithm. We give a general form of competitive ratio in terms of the price levels. The optimality of the Aggregating Strategy gives an evidence that our algorithm performs at least as well as the previously proposed ones.

17. Takayuki Sato, Kazuyuki Amano, Akira Maruoka:
"On the Negation-Limited Circuit Complexity of Sorting and Inverting k-tonic Sequences", Proceedings of the 12th Computing and Combinatorics Conference (COCOON 2006), Lecture Notes in Computer Science, 4112, 104–115, Aug. 2006.
概要: A binary sequence $x_1, \ldots, x_n$ is called k-tonic if it contains at most k changes between 0 and 1, i.e., there are at most k indices such that $x_i \neq x_{i+1}$. A sequence $\neg x_1, \ldots, \neg x_n$ is called an *inversion* of $x_1, \ldots, x_n$. In this paper, we investigate the size of a negation-limited circuit, which is a Boolean circuit with a limited number of NOT gates, that sorts or inverts k-tonic input sequences. We show that if $k = O(1)$ and $t = O(\log \log n)$, a k-tonic sequence of length n can be sorted by a circuit with t NOT gates whose size is $O((n \log n)/2^{ct})$ where $c > 0$ is some constant. This generalizes a similar upper bound for merging by Amano, Maruoka and Tarui (Disc. Appl. Math. 126(1) 3–8, 2003), which corresponds to the case $k = 2$. We also show that a k-tonic sequence of length n can be inverted by a circuit with $O(k \log n)$ NOT gates whose size is $O(kn)$ and depth is $O(k \log^2 n)$.

18. Kazuyuki Amano, Akira Maruoka:
"On the Monotone Circuit Complexity of Quadratic Boolean Functions", Algorithmica (Special Issue on ISAAC 2004), Vol. 46(1), 3–14, Sep. 2006.
概要: Several results on the monotone circuit complexity and the conjunctive complexity, i.e., the minimal number of AND gates in monotone circuits, of quadratic

Boolean functions are proved. We focus on the comparison between single level circuits, which have only one level of AND gates, and arbitrary monotone circuits, and show that there is an exponential gap between the conjunctive complexity of single level circuits and that of general monotone circuits for some explicit quadratic function. Nearly tight upper bounds on the largest gap between the single level conjunctive complexity and the general conjunctive complexity over all quadratic functions are also proved. Moreover, we describe the way of lower bounding the single level circuit complexity and give a set of quadratic functions whose monotone complexity is strictly smaller than its single level complexity.

19. 酒井義文, 上原稔, 佐藤永欣:
"読捨てコンテンツをいつ更新するべきか", 電子情報通信学会論文誌, D-I, Vol. J89-D, No. 9, pp. 1913–1922, 2006 年 9 月.
概要: ユーザが以前に閲覧したことのある読捨てコンテンツを再閲覧したときのコストを 1，読捨てコンテンツを新しいものに更新する際に生じるコストを m としたときの，u 人のユーザによる閲覧要求列に対する決定性オンライン更新アルゴリズムの競合比を，競合解析の手法により評価する.

20. Kazuyuki Amano, Akira Maruoka:
"Better Upper Bounds on the QOBDD Size of Interger Multiplication", Discrete Applied Mathematics, Vol. 155(10), 1224-1232, May. 2007.
概要: We show that the middle bit of the multiplication of two n-bit integers can be computed by an OBDD of size less than $2.8 \cdot 2^{6n/5}$. This improves the previously known upper bound of $(7/3) \cdot 2^{4n/3}$ by Woelfel (J. Comp. Sys. Sci. 71(4) pp. 520–534, 2005). The experimental results suggest that our exponent of $6n/5$ is optimal or at least very close to optimal. A general upper bound of $O(2^{3n/2})$ on the OBDD size of each output bit of the multiplication is also presented.

21. K. Amano, J. Tarui:
"A Well-Mixed Function with Circuit Complexity $5n \pm o(n)$: Tightness of the Lachish-Raz-type Bounds", Proc. of the 5th Annual Conference on Theory and Applications of Models of Computation (TAMC 08), (掲載予定).
概要: A Boolean function on n variables is called k-mixed if for any two different restrictions fixing the same set of k variables must induce different functions on the remaining $n-k$ variables. In this paper, we give an explicit construction of an $n-o(n)$-mixed Boolean function whose circuit complexity over the basis U_2 is $5n+o(n)$. This shows that a lower bound method on the size of a U_2 circuit that uses the property of k-mixed, which gives the current best lower bound of $5n-o(n)$ on a U_2-circuit size (Iwama, Lachish, Morizumi and Raz [STOC '01, MFCS '02]), has reached the limit.

研究会等

1. Kazuyuki Amano, Akira Maruoka:
"Better Simulation of Exponential Threshold Weights by Polynomial Weights", Electronic Colloquium on Computational Complexity, 10, 90-1–90-8, Nov. 2004.
概要: We give an explicit construction of depth two threshold circuit with polynomial weights and $\tilde{O}(n^5)$ gates that computes an arbitrary threshold function. We also give the construction of such circuits with $O(n^3/\log n)$ gates computing the COMPARISON and CARRY functions, and that with $O(n^4/\log n)$ gates computing the ADDITION function. These improve the previously known constructions on its size and simplicity.

2. 原田薫明，瀧本英二，丸岡章:
"リスク情報を用いたオンライン資源分配",

電子情報通信学会技術研究報告（コンピュテーション）, Vol. 104, No. 743, 105–110, 2005 年 3 月.

概要: 本論文では，動的資源分配問題について，オンライン予測の枠組みで考察を行う．我々は，広く研究されているオンライン学習のモデルを一般化したモデルを提案する．このモデルでは，学習者はそれぞれの投資先のリスクがどれほど高いかという情報を知ることができる．このような仮定はごく自然なもので，多くの例が考えられる．例えば競馬では，投資家は資金を賭ける前にリスク情報として馬の倍率を知ることができ，投資家はその倍率を考慮した投資が許される．本論文ではアグリゲーティングアルゴリズムをこのような問題に適用し，厳密な性能評価を行う．我々の結果は，投資先のリスクが一様である場合と比較して，それぞれの投資先でリスクが異なるほど，アルゴリズムはよりよい性能を持つことを示唆している．

3. Shigeaki Harada, Eiji Takimoto, Akira Maruoka:
"Online Allocation with Risk Information", Proceedings of the 4th Japanese-Hungarian Symposium on Discrete Mathematics and Its Applications, 84–91, Jun. 2005.

概要: We consider the problem of dynamically apportioning resources among a set of options in a worst-case online framework. The model we investigate is a generalization of the well studied online learning model. In particular, we allow the learner to see as additional information how high the risk of each option is. This assumption is natural in many applications like horse-race betting, where gamblers know odds for all options before placing bets. We apply the Aggregating Algorithm to this problem and give a tight performance bound. The bound we give intuitively implies that the algorithm performs better when faced with options of various risks than when faced with options of the same risk.

4. Kazuyuki Amano, Akira Maruoka:
"Tighter Bounds on the OBDD Size of Integer Multiplication", Proceedings of the 4th Japanese-Hangarian Symposium on Discrete Mathematics and Its Applications, 9–15, Jun. 2005.

概要: We show that the middle bit of the multiplication of two n-bit integers can be computed by an OBDD of size less than $2.8 \cdot 2^{6n/5}$. This improves the previously known upper bound of $(7/3) \cdot 2^{4n/3}$ by Woelfel (STACS, 2001). The experimental results suggest that our exponent of $6n/5$ is (at least very close to) optimal. A general upper bound of $O(2^{3n/2})$ on the OBDD size of each output bit of the multiplication and some conjecture on the structural properties of multipliers inspired by the experimental results are also presented.

5. Tatsuya Watanabe, Eiji Takimoto, Kazuyuki Amano, Akira Maruoka:
"Random Projection and Its Application to Learning", Proceedings of 2005 Workshop on Randomness and Computation, 3–4, Jul. 2005.

概要: 任意の 2 点間のユークリッド距離を保存するランダムプロジェクションを用いた線形分離関数の学習について述べる．また，ランダムシードが 4 限定独立のときは弱い距離保存の性質を持つが，3 限定独立のときは距離保存の性質をまったく持たないことを示す．

6. 渡邊辰也，瀧本英二，天野一幸，丸岡章:
"マージンを保存するランダム性を限定したプロジェクションとブール空間への埋め込み", 電子情報通信学会技術研究報告（コンピュテーション）, Vol. 105, No. 343, 21–28, 2005 年 10 月.

概要: マージンが最大となる仮説を求めることは機械学習の分野における最も重要な問題のひとつである．しかし，与えられたデータを分離する大きなマージンを持った超平面を見つけることは難しい．本稿では，高次元空

間でのマージンをほぼ保存したまま低次元空間へのランダムな写像を行うような，3つのランダムプロジェクションの手法を提案する．これらの手法により写像された低次元空間内でパーセプトロン等のアルゴリズムを用いることで，高速に適切な仮説を探すことができる．我々の手法は，Johnson-Lindenstrauss の補題に基づく既存の手法に比べ，マージンの保存という点では，より効果的である．JL 補題に基づく手法が任意の2点間の距離を保存するという強い性質を持つのに対し，我々の手法はマージンの保存のみを目的とするからである．具体的には，マージンの保存性能をほぼ維持したままランダムビットを削減する写像を与え，さらに，ブール空間へのマージンを保存した写像も与える．

7. 天野 一幸，佐藤 貴之，丸岡 章:
“充足割り当て数を最小化/最大化する単調 DNF 式について”，電子情報通信学会技術研究報告（コンピュテーション），Vol. 105, No. 344, 21–25, 2005 年 10 月.
概要: 変数の個数，項の数，項の長さ等が固定された中で，充足割り当ての個数が最も少ない，あるいは，最も多い単調 DNF 式を特定する問題について考える．まず，充足割り当ての個数を最小化する単調 DNF 式は，項を “辞書順” に並べた式であることが，極値集合論における Kruskal-Katona の定理より，容易に導くことができることを述べる．一方，充足割り当ての個数を最大化する単調 DNF 式は，各項の長さを 2 に限定し，かつ，項の個数にもある制約を設けた場合には，Bollobás による完全部分グラフの個数に関する定理から容易に求められるものの，項の長さや項数を一般化した問題は，非常に困難な問題となっている可能性のあることを述べる．

8. 高橋良介，瀧本英二，丸岡章:
“シャノンスイッチングゲームにおけるペアリング戦略の複雑さについて”，電子情報通信学会技術研究報告（コンピュテーション），Vol. 105, No. 680, 7–14, 2006 年 3 月.
概要: 与えられたグラフ上で，特定の 2 点間を結ぶ道を作ろうとするプレイヤーと，それを阻止しようとするプレイヤーが戦うシャノンスイッチングゲームにおいては，最適戦略を求めることが PSPACE 困難であることが知られている．本論文では，後手の戦略をペアリング戦略と呼ばれる単純な戦略に限定した場合，先手の戦略に対応する問題の計算量が NP 完全で，後手の戦略に対応する問題が Σ_2^p 完全となることを示す．

9. Kazuyuki Amano:
“A Procedure that Generates a Class of Optimal Boolean Formulas”, 情報処理学会アルゴリズム研究会, 2006-AL-106, 1–8, 2006 年 5 月.
概要: In this paper, we investigate the size of Boolean formulas for composite functions. For two Boolean functions $f : \{0,1\}^n \to \{0,1\}$ and $g : \{0,1\}^m \to \{0,1\}$, $f \otimes g$ denotes a Boolean function on nm variables defined by $(f \otimes g)(x_1, \ldots, x_{nm}) = f(g(\tilde{x}_1), \ldots, g(\tilde{x}_n))$ where $\tilde{x}_i = (x_{(i-1)m+1}, \ldots, x_{im})$. We give a class of base functions G such that for every function of the form $f = f_1 \otimes \cdots \otimes f_k$ with $f_1, \ldots, f_k \in G$, a “naive” construction yields an optimal formula for f. The class G contains every two-variable functions, 134 out of 256 three-variable functions, and more. This can be viewed as a generalization of Khrapchenko’s result that says that the formula size of the parity function on $n = 2^k$ variables is n^2, which corresponds to the case $f_i = x_1 \oplus x_2$ for every i. We also give a procedure that recursively generates Boolean formulas whose optimality is guaranteed. Our results are based on a careful inspection of a recently proposed complexity measure called SumPi, which originated from the quantum adversary method.

10. 内沢啓，瀧本英二:
“エネルギー計算量に制限のある定数段しきい値論理回路のサイズの指数下界について”，電子情報通信学会技術研究報告（コンピュ

テーション）, Vol. 106, No. 289, 71–76, 2006 年 10 月.

概要: エネルギー計算量とは，脳が計算過程で消費するエネルギーと計算能力の関係を調べるために，しきい値論理回路に対して提案されたパラメータである．脳内のニューロンが“ 発火する ”時に,“ 発火しない ”時と比べて非常に大きなエネルギーを必要とするという事実に基づき，エネルギー計算量は，計算過程で発火する（1 を出力する）素子の個数として定義される．実際の脳内の計算は，非常に少ない発火回数で行われることが知られているため，本稿ではこの事実に着目し，エネルギー計算量の制限がしきい値論理回路の計算能力にどのように影響を与えるかについて理論的に解析した．具体的には，しきい値論理回路のエネルギー計算量，素子数，深さの 3 つのパラメータに関する関係式を導き，この関係式から，あるクラスに属するブール関数を計算する，エネルギー計算量と深さに制限のあるしきい値論理回路は，素子数が入力数に対して必ず指数的であることを示す．

11. 天野 一幸:
“回路計算量の線形下界に対する計算機支援証明について”, 電子情報通信学会技術研究報告（コンピュテーション）, Vol. 106, No. 289, 65–70, 2006 年 10 月.

概要: 近年の計算機の著しい性能向上を受けて，組み合わせ論的問題に対する計算機支援証明の適用可能性が広がっている．本稿では，明示的論理関数に対する回路計算量の線形下界の導出問題を取り上げ，これに対する計算機支援証明を構成する試みについて報告する．まず，この問題に対する証明手法として広く用いられるゲート除去法を，四色定理の証明との類似性についても言及しつつ，より機械的取り扱いが容易となるよう定式化する．次いで，岩間ら (MFCS ’02) による現在最良の線形下界の証明を，計算機支援のもとで再構成する試みを通じて得られた知見について述べる．

12. 澤田 清，天野 一幸:
“完全 k 分木型組織構造の多階層関係追加モデル”, 情報処理学会 アルゴリズム研究会, 2007-AL-110, 47–54, 2007 年 1 月.

概要: 本研究では，高さ H の完全 K 分木型組織構造に関係を追加するモデルを提案する．ここでは，$L(L = 1, 2, \cdots, H)$ 個の階層それぞれのすべての頂点対に辺を追加する場合に，完全 K 分木の全頂点対の最短経路の短縮長さを合計した総頂点間短縮経路長を最大にする最適深さの組 $\{N_1, N_2, \ldots, N_L\}$ $(N_1 > N_2 > \cdots > N_L)$ を求める．その結果，$\{N_1, N_2, \ldots, N_L\} = \{H, H-1, \ldots, H-L+1\}$ が示される．

13. 唐崎 正史，瀧本 英二:
“ブール関数に対するフィルタのノイズ除去効果について”, 電子情報通信学会技術研究報告（コンピュテーション）, Vol. 107, No. 73, 7–12, 2007 年 5 月.

概要: 理論計算機科学の分野における様々な問題が，ブールドメイン $0, 1^n$ 上の関数のフーリエ変換を通して解析されているが，その手法は，基底関数の正規直交性やフーリエ係数の組み合わせ論的な性質を調べることを主としており，信号処理の分野における従来のフーリエ解析とはほぼ独立に発展している．一方，信号処理分野では，フーリエ解析について長い歴史があり，すでに膨大な知識が蓄えられている．我々は，計算機科学の立場からこれらの結果を見直し，類推することによって計算機科学においても有用な新たな知見が得られる可能性があると考える．そこで，本論文では，信号処理分野で基本的かつ重要な役割を果たしているフィルタを取り上げ，ブール関数に対するフィルタの概念を定式化し，その性質を調べる．特に，ブール関数に対するフィルタのノイズ除去効果の度合を表わす公式を 2 つ与える．

14. K. Amano, J. Tarui:
“A Well-Mixed Function with Circuit Complexity $5n \pm o(n)$: Tightness of the Lachish-Raz-type Bounds”, The 10th Korea-Japan Joint Workshop on Algorithms and Computation (WAAC 07), No. 2, Aug. 2007.

概要: A Boolean function on $X_n =$

$\{x_1, \ldots, x_n\}$ is called k-mixed if for every $V \subseteq X_n$ with $|V| = k$ and for any two distinct partial assignments with support V yield different subfunctions. In this paper, we give an explicit construction of an $n - o(n)$-mixed Boolean function whose circuit complexity over the basis U_2 is $5n + o(n)$. This shows that a lower bound method that uses the property of k-mixed, which gives the current best lower bound of $5n - o(n)$ on a U_2-circuit size (Iwama, Lachish, Morizumi and Raz [STOC '01, MFCS '02]), has reached the limit.

15. K. Amano, S. Nakano, and K. Yamanaka: "On the Number of Rectangular Drawings : Exact Counting and Lower and Upper Bounds", 情報処理学会アルゴリズム研究会, 2007-AL-115-5, 33–40, 2007 年 11 月.
概要: A rectangular drawing is a plane drawing of a graph in which every inner face is a rectangle. In this paper, we consider the problem of counting the number of rectangular drawings with n faces, denoted by $R(n)$, and show the following : (i) There is an algorithm for counting $R(n)$ in time $O(poly(n) \cdot 2^n)$ which enables us to determine exact values of $R(n)$ for $n \leq 30$, and (ii) There is a limit $c_R = \lim_{n\to\infty} R(n)^{1/n}$ such that $11.56 < c_R < 28.3$.

学会大会等

1. 天野 一幸:
"しきい値素子を用いた回路における計算量について (招待講演)", 日本オペレーションズ・リサーチ学会関西支部研究部会, 2005 年 1 月.

2. 瀧本 英二:
"オンライン予測の理論と応用（チュートリアル講演）", 第 18 回 回路とシステム軽井沢ワークショップ, 595–600, 2005 年 4 月.

3. 天野 一幸:
"論理関数の複雑さの下界導出問題に対する数理計画的アプローチ（チュートリアル講演）", 電子情報通信学会コンピュテーション研究会, 2005 年 10 月.

4. 渡邊 辰也，瀧本 英二，天野 一幸，丸岡 章:
"ランダム写像による非線形概念の学習の効率化に向けて", 電子情報通信学会総合大会 COMP-NHC 学生シンポジウム，DS-1-14, S35–S36, 2006 年 3 月.

5. Eiji Takimoto:
"Path Kernels and Finding the Optimal Strategy for a Routing Game (Tutorial Talk)", Tutorial workshop on game-theoretic probability and related topics, 2006 年 3 月.

6. H. Fukuhara, K. Amano, E. Takimoto:
"On formula size lower bounds for synthesis of Boolean functions over disjoint sets of variables", 電子情報通信学会総合大会 COMP-NHC 学生シンポジウム, DS-1-2, S3–S4, 2007 年 3 月.

7. 天野 一幸，垂井 淳:
"回路計算量の 5n の下界に対する 5n の上界", 2007 年度夏の LA シンポジウム, 2007 年 7 月.

文部科学省科学研究費補助金 特定領域研究 (課題番号 16092202)

A02: 計算理論的設計による知識抽出モデルに関する研究

本研究の目的はデータマイニングにおける現在の精度限界を打破するための知識抽出モデルの提案と、それに関する理論研究及び実際のシステムの構築である。特に、研究代表者の持つ先端技術である計算幾何学的手法を中心に、計算理論手法による限界突破を目指した。本研究により、既存の多くのデータマイニングシステムの重要なコンポーネントである意志決定システムに透明性と高い学習汎用性が与えることが出来た。更に、基本技術として用いるグラフ・計算幾何アルゴリズム理論を中心とした理論計算機科学での理論発展と共に、理論計算機科学技術の知識抽出への応用の新しい進展を試みた。

研究組織

研究代表者：	徳山 豪	東北大学 大学院 情報科学研究科
研究分担者：	塩浦 昭義	東北大学 大学院 情報科学研究科
	河原林 健一	国立情報学研究所 [平成 18 年度より] 東北大学 大学院 情報科学研究科 [平成 17 年度まで] (平成 17–18 年度)
	全 眞嬉	東北大学 大学院 情報科学研究科 (平成 18–19 年度)

交付決定額 (配分額)

平成 16 年度	3,500,000 円
平成 17 年度	3,500,000 円
平成 18 年度	3,500,000 円
平成 19 年度	3,600,000 円
合　計	14,100,000 円

研究成果の概要

- 学術誌
  Advances in Applied Mathematics (2004 年 8 月), Mathematical Programming (2004 年 12 月), Algorithmica (2005 年 4 月, 2006 年 2 月), International Journal of Computational Geometry and Applications (2006 年 6 月), SIAM Journal on Discrete Mathematics (2007 年 6 月)
- 国際会議
  SWAT (Humlebaek, Denmark, 2004 年 7 月), ISAAC (香港, 2004 年 12 月), FOCS (ピッツバーグ, 2005 年 10 月), ISAAC (中国海南島, 2005 年 12 月), STOC (シアトル, 2006 年 5 月) (2 件), SODA (ニューオリンズ, 2007 年 1 月) (2 件)
- 国際会議の開催: The 5th Hungarian-Japanese Symposium on Discrete Mathematics and Its Applications (宮城県仙台市, 2007 年 4 月)
- 受賞: Best Paper Award of ISAAC 2006, 平成 18 年度文部科学省若手科学者賞

1　はじめに

本研究の目的はデータマイニングにおける現在の精度限界を打破するための知識抽出モデルの提案と、それに関する理論研究及び実際のシステムの構築である。特に、研究代表者の持つ先端技術である計算幾何学的手法を中心に、計算理論手法による限界突破を目標とした。データマイニングは大規模なデータから情報をコンパクトな知識として抽出する技術であり、近年に情報科学から発信した技術であるにもかかわらず、経済学、経営学、統計学などの学問分野、及び経営コンサルティング、在庫管理や顧客サービスといった実用分野に幅広く応用が期待され、情報化社会における最重点技術の一つである。

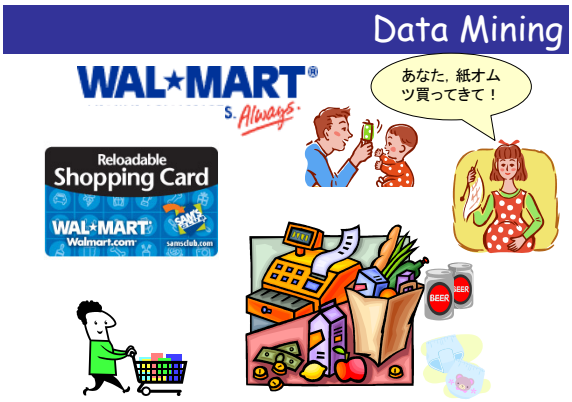


図 1: データマイニングのイメージ：紙おむつを買う人は高い確率でビールも同時に買う

自動的に抽出し、表示された知識形態は、ユーザにより意志決定等の補助として用いられる。重要な条件は、抽出した知識形態がシンプルであり（単純性)、正確にデータの特性を記述すること（正確性)、知識として（サンプルに依存しない) 汎用性を持つ事（学習汎用性) さらにユーザにとって説得力があり、検証が容易であること (透明性) である。単純性と透明性の観点から、結合ルール及びそれを用いた決定木は有力な手法である。これらの手法を用いたデータマイニングの研究は IBM 社アルマデン研究所や米国スタンフォード大学、カナダ国サイモンフレーザー大学、オーストラリア国シドニー大学等で 1980 年代半ばに始まり、主に二値データに対する研究が行われた。スタンフォード大学での研究成果の一例として、スーパーマーケットの売上情報のデータのマイニングでの「紙おむつを買う人はビールを買う」といったルールは広く知られている。これは、どちらかというとネガティブなイメージ（くだらないルールしか発見できない) に捉えられがちだが、「常識からは見出せないルールも発見できる」と見るべきである。現在では、広くデータマイニング技術は用いられており、例えば本のインターネット販売でも、データマイニングを用いた解析により、購買者の過去の履歴から、興味のありそうな書籍のリストを作成して提示するというような事が行われている。

図 2: 書籍販売でのデータマイニング

さて一方、実用データマイニングで必須である数値データベースの取り扱いを考えると、数値データの二値化誤差から生じる正確性と学習汎用性のトレードオフに関する精度限界があり、それを乗り越えるのは技術的に困難と思われていた。日本でのデータマイニングの研究はこれに数年遅れ、研究代表者を含む日本 IBM のチームが 1990 年代後半に行った研究が国際的に最初に認められた成果である。具体的には、結合ルール条件生成時のアダプティブな二値化のアイデアを導入し、二値化誤差最小化の理論的枠組みである最適結合ルールの定式化を与え、データマイニングに計算理論の手法を導入することに世界に先駆けて成功した（1996–1998 年)。これにより、上記のトレードオフは大きく改善された。世界的に独創的な研究として高く評価され、現在広く関連研究がなされている。しかしながら、最小誤差二値化を用い

ても誤差には理論的下界があり、上記限界の完全な克服には至っていない。

本研究では、透明性を保ちながら、この二値化誤差からくる限界を克服するデータマイニングシステムの設計とその理論的解析を行った。このようなデータマイニングシステムは要望されてはいるが、現実化は難しく、独創的な新しいアイデアが必要とされていた。本研究により、既存の多くのデータマイニングシステムの重要なコンポーネントである意志決定システムに透明性と高い学習汎用性が与えることが出来た。更に、基本技術として用いるグラフ・計算幾何アルゴリズム理論を中心とした理論計算機科学での理論発展と共に、理論計算機科学技術の知識抽出への応用の新しい進展を目指した。

2 各年度の計画と成果

2.1 平成 16 年度

平成 16、17 年度においては階層化ルールをエキスパートとして用いた柔軟な決定木の実現を目標とした。初年度である平成 16 年度は、そのコンポーネント技術の開発とプロトタイプ作成をターゲットとした。

平成 16 年度の具体的な成果としては、下記のものが挙げられる。

1. 柔軟な決定木の構築に用いる階層ルールの高次元におけるモデル化と、高速アルゴリズムの設計及び最適データ近似の研究
2. データマイニングでの主要な道具であるクラスタリングにおいて、クラスタ発見の高濃度部分グラフ問題としての定式化と、計算クラスに関する新しい発見
3. 階層化ルールをエキスパートとして用いた柔軟な決定木のプロトタイプ実装、および実験による評価の開始
4. インターネットにおける知識抽出の実用的なモデル化の側面から見た、クラスタリングのモデルに対する検討及び実用的定式化の研究と実装
5. データサンプリングに関する離散システム理論とアルゴリズム理論の研究

その他、経済予測でのモデル化とアルゴリズム理論の応用、また、最適化の基礎理論である離散凸解析の理論などにおいても成果をあげた。

2.2 平成 17 年度

前年度に続き、基礎理論に関する研究と階層化ルールをエキスパートとして用いた柔軟な決定木の実現を目指した。平成 17 年度は特に、グラフ理論に関する最適化技術、学習理論の先端技術の研究とその利用について研究を行った。

本年度の具体的な成果としては、下記のものがある。

1. 柔軟な決定木の構築に用いる階層ルールの高次元におけるモデル化と高速アルゴリズムの設計及び最適データ近似の研究
2. 上記目的のための計算幾何学の最適化理論の研究
3. データマイニングでの主要な道具であるクラスタリングにおいて、クラスタ発見の高濃度部分グラフ問題としての定式化と、計算クラスに関する新しい発見
4. ヨーロピアン・アジアンオプションの価格付けに対する近似アルゴリズムの設計

また、本年度より河原林が研究計画に加わり、グラフ理論に関する研究を開始した。

2.3 平成 18 年度

平成 18 年度の具体的な成果としては、柔軟かつ実用的なデータマイニング手法の実現と、関連する基礎理論とその利用についての研究を行ったことが挙げられる。また、本年度より研究計画に加わった全と、博士研究員として参加した Ashkan Sami 研究員は、Web データや医療データなどでの問題への応用研究を行った。さらに、データマイニングで広く使われるクラスタリングのための基礎理論である、ボロノイ図についても研究を行

い、 さまざまな拡張について成果を得た。特にゾーン図という新しいボロノイ図に関する成果は非常に高い国際評価を得た。また、モデル班のメンバーとして、近年提案された様々な計算モデルについて研究を行った。特に、河原林はグラフマイナーを考慮した最適化アルゴリズムの設計理論について研究を精力的に行い、その成果は世界的なブレークスルーであると高い評価を得た。河原林の研究の一部は、本特定領域の経費で招聘したBruce Reed 教授との共同研究であり、データマイニングで用いるグラフの視覚化に応用できる非常に優れた成果である。また、徳山と塩浦はパラメトリック複雑度モデルでの計算量の理論と離散凸解析モデルでの 計算理論での成果をあげた。そのほかに、徳山と塩浦は計算幾何学を用いた金融工学における新しいアルゴリズム設計モデルについていくつかの研究成果を得た。

本年度の具体的な主な成果としては、下記のものがある。

1. 柔軟な決定木の構築と実験
2. 階層ルールの一般化に対する計算量の探求と高速アルゴリズムの設計
3. 計算幾何学の最適化理論の研究
4. Web データや医療データにおけるデータマイニングの実用研究
5. パラメトリック複雑度やグラフマイナーを考慮した新しいモデルでのアルゴリズム設計理論と解析
6. 計算幾何学手法を用いた金融工学における近似アルゴリズムの設計

2.4 平成 19 年度

継続中の研究テーマについては研究をさらに進めると共に、過去 3 年間に行った研究の取りまとめを行った。新たに得られた研究成果については、国内外の会議や研究会にてを発表するとともに、論文に纏めて学術論文誌に投稿をした。

3 研究活動

研究活動予定に従い、徳山は基礎理論とモデル化及び計算限界の算定、塩浦は数理最適化とアルゴリズム理論に絞って研究を実施した。途中から研究計画に参加した河原林は新しいモデルでの最適化とグラフアルゴリズムの研究、全は階層ルールを用いた柔軟な決定システムの実装と実験及び改良を中心に研究を行った。システム実装と実験に関しては、全を中心に大学院学生のプロジェクトを構成し実施した。

研究代表者の徳山は研究成果に関連した論文誌・査読付き国際学会論文を 30 本以上、研究協力者の塩浦は同 11 本を発表した。また、河原林は権威ある STOC、FOCS、SODA などの国際会議での発表を行い、2006 年の国際会議 ISAAC では最優秀論文賞を獲得した。その他、国際学会招待講演、国内学会、研究会等での発表も数多く行った。総じて、研究グループにおいて FOCS の論文 1 件、STOC の論文 2 件、SODA の論文 2 件、ISAAC での最優秀論文を含む 20 件近くの質の高い研究論文の発表を行い、十分高い評価を受ける事が出来る活動成果であると判断する。

徳山は 2005 年 2、3 月に行われた本特定研究主催の NHC 国際会議において PC 委員長として活動し、多くの国際的な研究者を招聘すると共に、活発な国際交流を行った。また、徳山は 2005 年 7 月に仙台で開催した SMAPIP-NHC 共催の国際ワークショップ「Randomness and Computation」の開催責任者の一人として貢献をおこなった。2007 年 4 月に仙台にて行われた国際会議「The 5th Hungarian-Japanese Symposium on Discrete Mathematics and Its Applications」においては、徳山、塩浦、全の 3 名は実行委員として中心的な役割を果たした。

徳山は市民講座での講演など本研究テーマに関連した啓蒙活動も行った。

研究期間中、塩浦はドイツ、河原林は米国やデンマーク等に中長期の訪問を行うことにより、共同研究を行い、広く国際的に活動した。

経費について

理論的な研究が主体であるため、研究活動に必須である国際的活動と、仙台からの国内旅費の必

要性から、小規模な備品・消耗品を除き研究費のほとんどは旅費に当てた。徳山はSTOC、SODA国際会議などのアルゴリズム分野での最高レベルのシリーズ国際会議で研究成果発表経費、ISAAC国際会議の参加・成果発表経費、その他国内旅費を本研究費から拠出した。さらに塩浦はヨーロッパ各国の研究者との共同研究を行い、広く国際的に活動したが、それらの旅費の一部を本研究費から拠出した。全はデータベースでの一流国際会議であるVLDB国際会議への参加費用や国内旅費を本研究費から拠出した。また、博士研究員として採用 (採用経費は他の経費から拠出) した Ashkan Sami 研究員はデータマイニングと人工知能の 2 つの国際会議での研究発表を行い、その経費を本特定領域研究費から拠出した。更に、NHC プロジェクトの国際会議や春、秋の強化合宿には、本研究計画に参加する学生を参加させ、国際的な一流研究者との交流経験を積ませたが、その旅費及び国内研究会発表のための旅費も本研究費から拠出した。また、特定研究の研究費で O. Cheong、J. Matoušek、M. Halldórsson ら国際的に一流な研究者を東北大学に招聘し、講演・共同研究を依頼し、研究面での多大な成果をあげた。

4 研究成果の概要

4.1 階層ルールの高次元におけるモデル化とアルゴリズム設計

データベースからの知識抽出システムの構成において、数値属性の取り扱いは大きなトピックである。数値属性の 2 値化に関してはさまざまな研究があり、特に数値属性のペアについては領域ルールというアイデアが福田、森本、森下、徳山 (SIGMOD96) によって 1996 年に考案され、大きな効果を上げている。しかし高次元 (多属性) 領域ルールの実現にあたっては、直接的なモデル化を行うと、近似計算すら困難な問題に定式化される。この計算困難性を避けるために計算理論的な考察を行い、それに従ったモデル化とアルゴリズム設計を行う必要性がある。更に進んで、2 値化ではなく柔軟な多値化によって得られる階層化ルールは、後述のようにエキスパート付きの決定木構築に重要な役割を果たすことが知られている。しかし、1 次元の場合にのみ良い手法が考案され、1 次元手法の直接的な拡張は多項式時間アルゴリズムにならないという問題点があった。実際、高次元 (多属性) 階層化ルールの実現にあたっては、直接的なモデル化を行うと、近似計算すら困難な問題に定式化される。

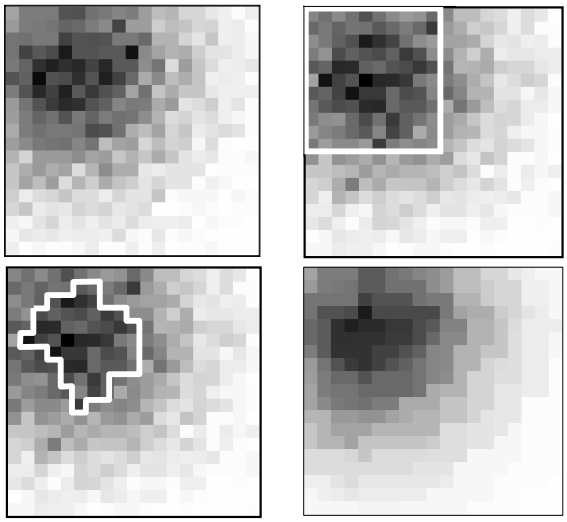


図 3: 2 種類 (長方形、ピクセル凸) の領域ルールと階層化ルール (右下図)

この計算困難性を避けるために計算理論的な考察を行い、それに従ったモデル化とアルゴリズム設計を行う必要性がある。この困難を克服するために、本研究ではピクセルグリッド上で定義された単峰曲面によるデータ分布の近似として問題を定式化した。さらに有向グラフを用いて任意の次元における高次元ルールをモデル化し、階層化された領域ルールの最適化問題を最大支配カットというネットワークフロー問題に帰着し、高速アルゴリズムの設計を実現した。これは、一般次元での数値属性のアソシエーションルールの設計としては世界初であり、大きな成果である。図3には2次元の問題例が示されており、ピクセルグリッドの濃度がデータの確信度分布を表す。左上と右下が領域ルール、右下図が階層化された領域ルールであり、濃度を高さと思うと曲面の等高線表示を示していることになる。正確にはデータマイニングで利用する際は、ピクセルグリッドにはサポートに対応した非均一な測度を導入し、この測度に対して最適化を行う。そのため、我々の手法は任

意の測度に対して有効であるように設計されている。図 4では対応するグラフと最大支配カットを示す。

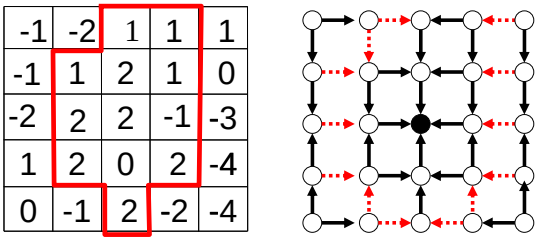


図 4: 領域の切り出し (左図) と対応するグラフの最大支配カット (右図)

さらに、2 次元の階層化ルールにおける計算量の改善、一般の L_p 基準による 1 次元の階層化ルールの最適アルゴリズム設計なども成果としてあげられる。更に、1 次元データに関しては、データの確信度分布のヒストグラムから最適な k ピーク近似を計算するアルゴリズムを開発した。これにより、データから幾つのルールを抽出すべきかの判定、隠れた比較的小さいルールの抽出が可能となった。

また、1 次元ルールの設計に関連して、ピークの数ではなく区分線形関数の折れ線の数に注目し、尺度を L_1 もしくは L_2 にした、データ集合の最適折れ線近似に対しても新しい理論成果をあげた。これは、従来 L_∞ 基準でしか考えられていなかった折れ線近似問題に新しい定式化を与えたものである。これらの成果は COCOON、ISAAC 等の国際会議で発表され、論文誌にも採択された。

4.2 高濃度部分グラフを用いたクラスタリング

知識抽出システムでは、重複を許した非排他的クラスタリングが重要なツールである。そこでは、データを分割するのではなく、クラスタを一つずつ発見していくというアプローチが採られる。代表的なアルゴリズムとして知られる Agrawal らのアプリオリアルゴリズムは、データと (カテゴリ) 属性の作る 2 部グラフを考え、データ部分の大きい 2 部クリーク (完全 2 部グラフ) を全て列挙するという方針をとっている。しかしながら、この手法を用いた場合、属性数が多くなるとクリークの数が爆発するため、全て列挙することは不可能である。改良案として極大 2 部クリーク列挙アルゴリズムも提案されているが、計算爆発を完全には克服できない。一方、枝数が最大のクリークを求める問題、もしくは指定した属性数に対する最大クリークを求める問題は NP 完全であり、近似する事も計算量理論的に困難である。したがって本研究では、クリークではなく、(サイズを指定した) 最大濃度部分グラフ問題を用いてクラスタ発見を行うという新しいアプローチを採った。最大濃度部分グラフ問題は古くから研究されている。従来の計算理論の研究においては、U. Feige によってランダム 3SAT 判定からの帰着が示されており、計算量は近似困難問題のクラスに入ることがわかっているため、多項式時間アルゴリズム設計は (計算階層が崩壊しない限り) 不可能であった。知られている最良の近似比は $O(n^{1/3})$ を少し改良したもの (Feige、Kortsarz、Peleg、Algorithmica 1993) であり、これを越えることは非常に難しいとされていた。多項式時間の定数近似アルゴリズム設計が可能な既知の場合は、Arora、Karger、Karpinski らによる部分グラフが $\Omega(n^2)$ の辺数を持つ場合 (STOC 1995) の PTAS があるが、これは実用的なアルゴリズムではなかった。また、朝廣、岩間、玉木、徳山によるグリーディーアルゴリズムの解析 (J. Algorithms 2000) では、定数近似を達成するのは頂点数が $\Omega(n)$ の場合に限られていた。一方、本研究では、最大濃度部分グラフの濃度に注目し、この濃度が高い (クリークに比べて定数倍の辺をもつ) 場合は、高速な近似アルゴリズムの設計が可能であることを示した。この高濃度条件は、データマイニングのようにクリークに近い部分グラフのみを必要とする場合には実用的に妥当な条件である。

成果としては、2 部グラフに対しては $O(n^\epsilon)$ 近似アルゴリズムおよび擬多項式時間近似スキームを与え、この近似スキームが多項式時間近似スキームになるための条件の明確な提示を行った。

これらは、新しい見地から最大濃度部分グラフ問題の計算量を追求したものであり、世界的にも新規性の高い成果である。更に、一般の重み付きグラフに対してもこの理論を拡張し、$O(n^{\epsilon})$ 近似の多項式時間アルゴリズム及び 2 近似を達成する擬多項式時間アルゴリズムを考案している。これらの成果は国際会議 ISAAC 等で発表された。また、一般の組合せ最適化問題に 対しての同様のアプローチの検討を行った。本研究の成果は、subdense なインスタンスでの計算困難性の追求という観点から、世界的にも大きな貢献であるといえる。

4.3 エキスパート付き決定木の構成と実装

アソシエーションルールを用いた決定木はデータマイニングにおける意思決定システムとして広く用いられている。しかしながら、数値属性を含むデータにおいては、データのカテゴリ属性への変換による情報損失があり、この損失と決定木における過学習のバランスを取る事が重要になる。そのために、数値属性の情報をエキスパートとして付加した決定木の構成を目指した。エキスパート付き決定木に関しては主に学習理論の見地から理論的な研究は行われているが、現実のエキスパートの作成を階層化ルールを用いて実現し、数値データベースでのデータマイニングへの応用を意識した実装実験を行ったのは世界でも初めてのものであった。

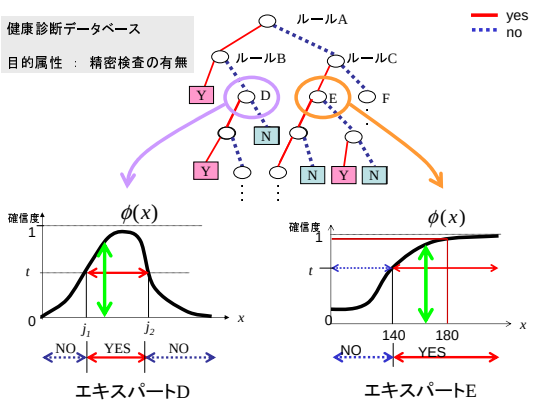


図 5: エキスパート付き決定木

最適ピラミッドを用いた決定木構成アルゴリズムは以下のように記述される。

ステップ 1 学習データ T を読み込む。

ステップ 2 T のすべての事例が同一クラスならば葉とする。
そうでなければ、各属性に対して次の 1〜4 を行う。

1. 各属性値における確信度を求める。
2. 求めた確信度を元に最適ピラミッドを構築する。
3. 全ての高さに対して分割を行い、相互情報量を計算する。
4. 得られた各属性の相互情報量を比較する。
   (a): 閾値 g_ℓ 以上かつ最も大きな相互情報量が得られた属性があるならば、それをノードとし、T を T_1, T_2 に分割する。同時に、T_1, T_2 の確信度を求めておく。
   (b): もし、どの属性の相互情報量も g_ℓ 未満であるならば、T に最も多く現れるクラスラベルの葉とする。
5. 分割頂点に最適ピラミッドから構成したエキスパートを配備する

ステップ 3 分割された事例集合に対し、全てが葉になるまでステップ 1、ステップ 2 を繰り返す。

生成された決定木の各ノードでは分割テストが行われる。例えば、根である属性 A テストは、全学習データ T の、属性 A について求められた確信度のヒストグラム f をもつ。また、f から求められた最適ピラミッド ϕ、そして、分割された T_1, T_2 の確信度 $\mathrm{conf}(T_1)$, $\mathrm{conf}(T_2)$ をそれぞれもっている。この確信度は、テストデータが分類され枝を通るときに、各テスト事例に対して与えられる。分割された事例集合 T_1, T_2 は、子ノードに入力され、子ノードも同様な構造を持つ。

また、決定木の末端である葉には、クラスラベルが与えられる。その葉に振り分けられた学習データの事例数 $|T_{11}|$, その中でクラスが C でない事例数 $|E_{11}|$ (エラー数) も記録してある。

エキスパートは、決定木構築アルゴリズムのステップ 2-4 の後に生成され、該当ノードにおける

学習データ数 $|T|$ 及び T より構築された最適ピラミッド関数 $\phi(T)$ をもつ。エキスパートは決定木の各ノードに配置される。入力されたデータ i に対し、ノード A のエキスパートは、学習データ数 $|T|$ 及び $\phi(T)$ におけるデータ i の確信度 $\alpha_T(i)$ を与える。そして、与えられた確信度と事例数のセットにより、データ i のクラスを決定できるかどうか判断する。あるノード A の情報の信頼性が大きいと判断したら、そこでデータ i にラベルを与え、テストを終了する判断を行う。即ち、ノード A, B 間で枝刈りを行うことになる。ここで、比較に用いた $\alpha_T(i)$ の値は各データによって値が異なる。つまり、エキスパートは、個々のデータに依存した、データことに柔軟かつ暗に (implicit に) 決定木の枝刈りを行う。

実験では、C4.5 が生成した枝刈り前の決定木の各ノードに最適ピラミッドを構築し、エキスパートを配置した。各ノードのエキスパートは、テストデータに対し、個々のデータに依存する確信度を与える。その確信度がユーザの指定した範囲にあれば、その時点でラベルを与え、分類を終了し、次のデータの分類へと進む。実験データは UCI で提供する Diabetes データセット (`http://www.ics.uci.edu/~mlearn/MLRepository.html`) で、糖尿病診断データを用いた。実験ではエキスパートのパラメタ と精度の相関を調べた。パラメタをチューンする事により、エキスパートなしの場合に比べて精度の大幅な向上を見ることが出来た。

更にスペシャリストモデルを用いたオンライン学習理論を適用し、従来よりも高い理論保証を持つシステム構築を行った。スペシャリストモデルでのオンライン学習は、A01 班の研究テーマと関連し、多くの決定システムを統合する理論的な枠組みである。スペシャリストモデルにおいては、スペシャリストは決定を行うことも、決定を保留することも出来、最適なスペシャリストグループの統合した決定と同等の判定を、(その最適なグループを知らずに) 自動的に行う学習システムである。理論的な保証の導出と、実験による検証を行い、従来のシステムに比べての優位性を見出した。

4.4 インターネットでの探索における知識抽出

図 6: 排他的ラベルクラスタによる検索結果表示

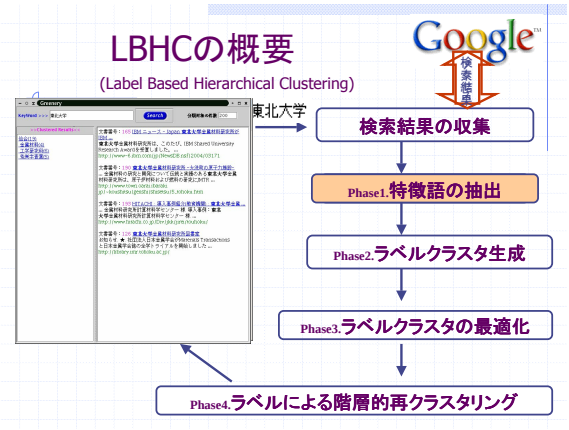


図 7: システム構成

インターネットでの検索エンジンは現代生活に不可欠なものとなっている。現在、検索エンジンにより得られた検索結果はページランクアルゴリズムなどにより整列され、リストとして表示される。しかしながら、WWW の巨大化と検索の高速化・大規模化に伴い、検索結果として得られた文書集合を知識抽出により判りやすい形で分類し、表示する技術が必要になってきている。さまざまなクラスタリング手法による分類が考えられているが、本研究では、latent semantics を用いた文書集合の非排他的クラスタリングを行う時に、クラスタに自動的に付随するラベルを知識として捉え

る。 これにより、ラベルに関して排他的であり文書に関して非排他的であるクラスタリングをネットワークフローなどを用いて考察し、従来よりも利用しやすい検索エンジン (LBHC：Label Based Hierachical Clustering 法) の設計を行った。プロトタイプの実験結果は研究会等で発表を行った。

4.5 データサンプリングと最適近似に関する理論研究

データ圧縮及びデータサンプリングで重要である丸めとディスクレパンシー基準による最適性評価において、計算理論的な成果をあげた。特に、グラフの最短路と幾何学的領域族に関するハイパーグラフに対する大域丸め集合の離散的構造を解明し、その列挙アルゴリズムの設計を行った。 特に領域族に対する丸め理論は、幾何学データの圧縮とサンプリングに応用があると期待される。これらの成果は本研究のテーマに間接的に影響を及ぼすものであり、海外の一流論文誌、一流国際会議で発表を行った。

4.6 階層ルールの一般化に対する計算量と高速アルゴリズムの設計

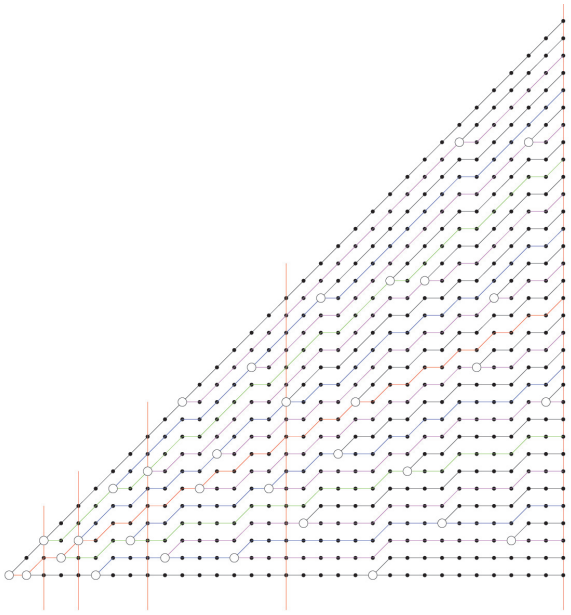


図 8: ディジタル星型領域を定義するグラフ

本研究では、数値属性用のエキスパート設計に用いられる階層化ルールに注目した。数値属性用のエキスパート設計に用いられる階層化ルールでは、ルールの基本形状が、これまでは「1 点を共有する長方形領域の和」に限られていたため、より広いクラスの形状への拡張が望まれていた。これまでの階層化ルール設計手法は、有向グラフの最大重み支配閉包が多項式時間で計算できる事を用いているが、これに関連する未解決問題として、2 つの異なったグラフの支配閉方の和集合で最大重みであるものの計算の複雑度の同定が求められていた。本研究における成果の一つとして、上記の未解決問題は、グラフが木である場合でも NP 完全であり、かつ近似アルゴリズムの設計も困難である事を証明した。さらに、有向グラフの最大重み支配集合の理論を用いた新しい階層化ルールの設計として、ディジタル平面やディジタル空間での (中心を固定した) 星型領域の集合を最大重み支配集合の集合とするグラフの設計に成功した (図 8参照)。これらの仕事は、全、徳山と、本研究グループに 1 ヶ月滞在したドイツ国カールスルーエ大学の Martin Nöllenburg(博士課程学生) の共同研究である。

4.7 ヨーロピアン・アジアンオプションの価格付けに対する近似アルゴリズムの設計

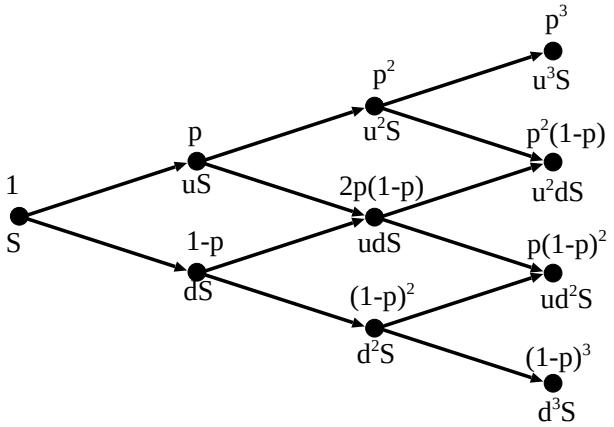


図 9: 二項モデル。各頂点の上には、根からその頂点に到達する確率、下にはその頂点における株価が書かれている。

オプションとは、金融市場で取引されている金融派生商品 (デリバティブ) の一つである。本研究では、パス依存型オプションの一種であるヨーロピアン・アジアンオプションの価格付けという問題に対し、高速かつ近似誤差が小さい近似アルゴリズムを提案した。

狭義には、オプションは株式などの原資産を、指定された時点もしくは期間に、決められた価格で売買する権利のことである。より広義には、オプションは原資産の価格の履歴に依存して決まる報酬 (ペイオフと呼ばれる) をオプションの所有者が得ることができる、という契約である。オプションは資産売買に伴うリスクのヘッジや、少ない投資で多額の利益を得るという目的のために使われている。様々なオプションが世界の金融市場で取引されている現在、オプションの適正な価格付けは重要な計算問題となっている。

オプションの購入日から満期までの平均株価を A、行使価格を X とおいたとき、ヨーロピアン・アジアンオプションのペイオフは $\max\{A-X,0\}$ と表される。二項木モデル (図 9参照) 上でパス依存型オプションの価格を厳密に求めることは、一般に#P 困難であることが知られている。ゆえに、効率的かつ理論的に近似精度が保証された近似アルゴリズムの設計が望まれており、これまでに様々な近似手法が開発された。

ヨーロピアン・アジアンオプションの厳密な価格を計算する単純な方法として、二項木のすべてのパスを列挙するという、全パス法と呼ばれる方法がある。しかし、最悪の場合、二項木のパスの数は指数個であり、ゆえに全パス法は指数時間を要する。

Aingworth、Motwani、Oldham (AMO) (2000) は、近似誤差を理論的に保証した最初の多項式時間近似アルゴリズムを提案した。この近似アルゴリズムでは、理論的な近似誤差はボラティリティに依存しないという利点をもつ。このアルゴリズムは動的計画法とバケッティングに基づいたものであり、高いペイオフをもつ指数個のパスを数式を使って枝刈りする、というアイディアを用いることで計算効率を高めている。このアルゴリズムの各ステップにおいて生じる誤差は X/k により押さえられる。ここで k は二項木の各頂点において使用されるバケットの数を表す。ゆえに、アルゴリズムの総ステップ数を n としたとき、AMO アルゴリズムの誤差は nX/k で押さえられる。またアルゴリズムの実行時間は $\mathrm{O}(kn^2)$ となることが示されている。のちに、AMO アルゴリズムの近似誤差は Dai, Huang, Lyuu (2002) および我々の研究グループ (2005) により改善された。AMO アルゴリズムでは、二項木の各頂点において同じ数のバケットが使用されているのに対し、Dai, Huang, Lyuu のアルゴリズムでは各頂点ごとに異なる数のバケットを使用する。近似アルゴリズムの計算時間を $\mathrm{O}(kn^2)$ に保ちつつ、各頂点でのバケットの数を適切に調整することにより、$\mathrm{O}(\sqrt{n}X/k)$ という近似誤差を実現している。ここで k は各頂点で用いられるバケットの数の平均値を表す。一方、我々が提案したアルゴリズムは、AMO アルゴリズムに「ランダム丸め」の技法を適用したものであり、$\mathrm{O}(n^{1/4}X/k)$ という近似誤差を実現した。

本研究では、計算時間がこれまでと同じ $\mathrm{O}(kn^2)$ であり、かつより小さい近似誤差 $\mathrm{O}(X/k)$ をもつランダム化アルゴリズムを提案した。AMO アルゴリズムおよび過去に提案された改良版アルゴリズムと異なり、このアルゴリズムの理論的な近似誤差は二項木の深さ n と独立である、という利点をもつ。提案するアルゴリズムでは、過去の 2 つのアルゴリズムで用いられたアイディアを同時に利用している。まず、我々のグループ (2005) が以前提案したアルゴリズムと同様に、今回提案したアルゴリズムでは、適切な確率変数を導入することにより、アルゴリズムの動きを $\mathrm{O}(n^2)$ 段階のマルチンゲール過程と見なす。そして、アルゴリズムの出力の期待値が厳密な値と等しいことを証明するとともに、各ステップで生じる近似誤差が二項木の頂点で用いられるバケット数の関数により抑えられることを示した。したがって、上記で述べたマルチンゲール過程に Azuma の不等式を適用することが可能となり、これにより近似誤差の理論的な解析が可能となった。ここで、二項木の各頂点でちょうど k 個のバケットを用いた場合、我々のグループ (2005) が以前提案した近似アルゴリズムと同一のものが得られる。このアルゴリズムの近似誤差を出来る限り小さくするために、二項木の各頂点で用いられるバケットの数を調整し、$\mathrm{O}(X/k)$ という近似誤差を得ることに成功した。ここで k は各頂点で用いられるバケット

の数の平均値を表す。X/k という値は、二項木の各頂点で生じる近似誤差の「平均値」と見なすことが出来るので、我々のアルゴリズムの近似誤差は、AMO アルゴリズムの枠組みの中で最良といえる。また、本研究では提案する近似アルゴリズムおよび過去に提案された近似アルゴリズムをコンピュータ上で実装し、実際の計算により得られるオプション価格の近似値の精度についても評価を行った。この結果は国際会議および国際的な論文誌において発表された。

4.8 アメリカン・アジアンオプションの価格付けに対する計算幾何学手法を用いた近似アルゴリズムの設計

本研究では、資産価格の変動を離散的に表現する確率モデルである二項モデルにおいて、アメリカン・アジアンオプションの価格を計算する、という問題を扱った。

オプションの価格は (利子を割り引いた) ペイオフの期待値として与えられるので、アメリカン・アジアンオプションのペイオフの期待値を計算する問題と言い換えることができる。この問題は非常に計算困難な問題であるが、その主な理由は次の 2 つである。

1 つ目の理由は、このオプションがアジアン型であり、ペイオフの値が資産価格の過去の履歴に依存するので、厳密値の計算のためには資産価格の履歴の全ての可能性を列挙しなければならない、ということである。したがって、アジアンオプションの価格を厳密に計算するには指数時間が必要となる。現実的な問題を解くためにはオプション価格の近似値を高速に求めるアルゴリズムが求められる。

2 つ目の理由は、アメリカンオプションでは任意の時期において権利行使が可能であるが、最適な行使時期の決定は計算困難である、ということである。例えば、モンテカルロ法はオプション価格の近似アルゴリズムとしてしばしば用いられ成功を収めているが、アメリカンオプションの場合にはその行使時期決定が難しいため、モンテカルロ法の適用には様々な工夫が必要とされる。

二項モデルにおけるアメリカン・アジアンオプションの近似アルゴリズムは Hull, White (1993)、Neave (1997)、Chalasani ほか (1999), Dai ほか (2002) によって提案されている。しかし、いずれのアルゴリズムについても近似精度は実験的に検証されているだけであり、理論的な近似精度保証は与えられていない。アメリカン・アジアンオプションの価格付けに対して近似精度を理論的に保証した近似アルゴリズムは、著者らの知る限り、これまでに提案されていなかった。

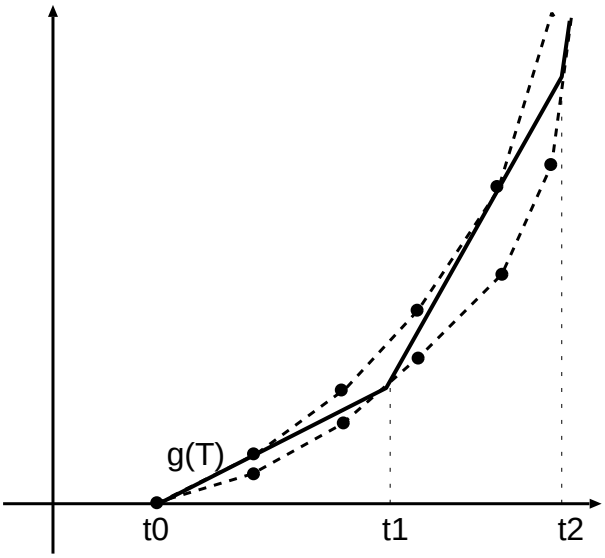


図 10: 計算幾何学手法を用いた近似アルゴリズムの動き

本論文では、アメリカン・アジアンオプションの価格付けに対し、近似精度を保証する多項式時間近似アルゴリズムを提案した。アメリカン・アジアンオプションの価格の厳密値は、各時点でのペイオフの期待値を表す区分線形関数を繰り返し計算することにより求めることが可能である。しかし、計算の過程で区分線形関数の複雑度は指数オーダーとなり、計算には指数時間が必要となる。これに対し、本論文では計算幾何学的な手法を用いて区分線形関数をより複雑度の小さい区分線形関数により近似する (図 10参照)。これにより、近似精度を保証しつつアルゴリズムの高速化を計った。また、本研究で提案した近似アルゴリズムをコンピュータ上で実装し、実際の計算により得られるオプション価格の近似値の精度についても評価を行った。

4.9 計算幾何学の最適化理論の研究

知識抽出において、特にイメージや幾何学データからの知識抽出にあたっては、幾何学的最適化技法は要素技術として必須である。本研究計画では幾何学的最適化技法に関する様々な研究を行ったが、ここでは特にゾーン図に対する研究について記述する。この成果は浅野哲夫教授 (JAIST)、Jiří Matoušek 教授 (チェコ国チャールス大学) との共同研究の成果であり、計算理論分野で最高レベルの国際会議 STOC(ACM Symposium on Theory of Computing)、SODA(ACM-SIAM Symposium on Discrete Algorithms) などで発表された。また、成果の一部は数学分野での一流専門誌である Advances in Mathematics に掲載された。

ボロノイ図は、平面上のデータに対応する勢力図として与えられる平面分割であり、知識抽出に必須であるクラスタリングやベクトル量子化を行うための基本技法である。本研究の大きな成果として、ボロノイ図の新しい変種であるゾーン図の性質の解明が挙げられる。

平面上の 2 点が与えられたとき、コンパスと定規を用いてそれらの垂直二等分線を作図するのは、小学校で習う初歩の作図法であるが、これは古代ギリシャでの「幾何算術」の基本であり、四則計算や根号計算などを作図を用いて行うための基本技法であった。垂直二等分線は 2 点からの等距離点の集合である。同様に、平面上の 3 点に対して、3 点からの等距離点は 3 点を通るユニークに定まる円の中心であり、三角形の外接円の中心 (外心) に等しい。三角形の外接円の中心が、3 本の垂直二等分線 (三角形の 3 つの辺の二等分線の集合) の交点として求める事ができることは中学校で学ぶ事実である。

3 点 $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3$ に対して、平面を、どの点に一番近いかで 3 つの領域 (支配領域) に分けることが出来る。これは外心から放射状にこれら 3 本の垂直二等分線を三角形の対応する辺に向けて書くことで実現できる。これを一般化して、平面上の n 点集合 $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$ に対して、

$$\begin{aligned}&\mathrm{Vor}(\mathbf{p}_i)\\&=\{\mathbf{z}\in\mathbb{R}^2 : d(\mathbf{z},\mathbf{p}_i)\le d(\mathbf{z},\mathbf{p}_j)\ \ (1\le j\le n)\}\end{aligned}$$

を点 $\mathbf{p}_i$ の支配領域とし、$\mathrm{Vor}(\mathbf{p}_i)$ によって平面分割を行ったものがボロノイ図である。ここで $d(\cdot,\cdot)$ はユークリッド距離である。ボロノイ図は垂直二等分線を用いて比較的容易に描くことができる。さてここで、ボノロイ図の変種を一つ考えてみよう。図 11の左図は通常のボロノイ図であり、右図は新たに導入した概念であるゾーン図 (zone diagram) である。

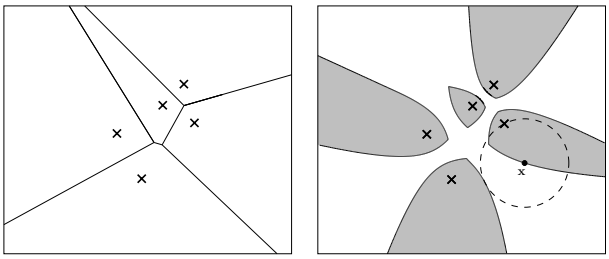


図 11: 5 つの母点が×印で描かれている。左は古典的ボロノイ図。右のゾーン図では母点 $\mathbf{p}_i$ は R_i を領域として持ち、R_i の境界上の点 $\mathbf{x}$ から $\mathbf{p}_i$ への距離は他の領域群への最短距離と等しい

平面上の点 $\mathbf{a}$ と集合 $X \subseteq \mathbb{R}^2$ に対して、$\mathbf{a}$ の X に対峙する支配領域 (*dominance region*) を

$$\mathrm{dom}(\mathbf{a}, X) = \{\mathbf{z}\in\mathbb{R}^2 : d(\mathbf{z},\mathbf{a}) \le d(\mathbf{z}, X)\}$$

で定義する。ここで $d(\mathbf{z}, X) = \inf_{\mathbf{x}\in X} d(\mathbf{z},\mathbf{x})$ である。

古典的なボロノイ図では、母点 $\mathbf{p}_i$ の支配領域は $\mathrm{dom}(\mathbf{p}_i, \{\mathbf{p}_j : j \ne i\})$ であり、これらの領域は平面を敷き詰める。ゾーン図では、$\mathbf{p}_i$ の支配領域は

$$R_i = \mathrm{dom}\Big(\mathbf{p}_i, \bigcup_{j\ne i} R_j\Big)\ \ (i = 1, 2, \ldots, n) \qquad (4.1)$$

で定まる。すなわち、支配領域は、他のどの支配領域よりも母点に近い点の集合である。支配領域 R_i の集合和は、中立地帯 (*neutral zone*) と呼ばれる空でない補集合をもつ。

ゾーン図の概念を「戦国時代」の均衡を例にして説明しよう。互いに敵対する n 国の領主たちがいるとしよう。領国 i はその居城を母点 $\mathbf{p}_i$ にもち、領国 R_i をその周りに持っている。これら n 個の領国は中立地帯で隔てられている。領国 R_i が他の領主から攻撃された場合、迎撃隊が居城 $\mathbf{p}_i$ から出発する。迎撃隊が R_i の境界にある攻撃地点に、敵軍よりも早く到着した場合に撃退に成功

する。しかしながら、攻撃側は自分の領内では秘密に部隊の移動を行うことが出来、迎撃隊の出発は攻撃側がその領国の境界を離れたときに始めて居城を出発できる。ゾーン図は上記の状況での均衡図であり、どの国も領土を守ることが出来、更にどの国も他の国からの侵略のリスクなしに領土の拡大が出来ない状態である。

ゾーン図の概念は非常に面白く、多くの数学的、アルゴリズム論的なチャレンジを提供してくれる。さらに、以下の理由により、ゾーン図やその変種は自然界の現象のモデル化に活用できるように思える。古典的なボロノイ図は基本的な幾何学構造であるが、様々な分野で利用され、中でも多体の成長過程 (細胞成長や結晶格子、地理的パターン、社会科学における領域均衡など) の数学的モデル化に利用される。今までのボロノイ図やその多くの一般化は、空間を母点たちに対応する領域に分割する。しかし、自然界における幾何学的構造においては、支配領域が全空間を覆っているとは限らない。例えば、成長過程において、支配領域群が接する前に成長が停止したような場合や、何らかの非接触 (遠隔) 作用によって領域成長が抑えられるような場合である。このような場合にはゾーン図のように中立地帯のあるモデルが適当である。

先に書いたように古典的なボロノイ図の定義は直接的であり、存在は自明である。しかしながら、ゾーン図の定義は陰的に与えられ、各々の領域の定義に他の領域が使用されている。従って、この定義を満たすような領域族の存在は非自明であり、また、一意に存在するかどうかも明らかでない。本研究ではこの 2 つの問題を肯定的に解決した。すなわち、下記の定理である。

定理 平面上の任意の n 点の集合 $\mathbf{p}_1, \ldots, \mathbf{p}_n \in \mathbb{R}^2$ に対して、領域族 $(R_1, \ldots, R_n)$ で (4.1) を満たすものが存在し、一意に定まる。

驚くべきことに、母点数が 2 である場合にすら上記定理は自明でない。この場合、2 つの領域は互いの鏡映像であり、境界線は distance trisector curve と呼ばれる面白い曲線になる。これは垂直二等分線の自然な拡張であるが、この曲線は代数的でなく、基本的な関数を複数個組み合わせても表示できないのではないかと予想されている。一方、曲線上の点は任意の精度までビット数の多項式の時間で計算することが出来る。

上記の定理の主張のうち存在証明は、解析学でよく知られた角谷–Shauder のバナッハ空間における不動点定理を利用することで与えられる。バナッハ空間の一点は領域の n 個の組に対応し、従ってこの空間は連続無限次元を持つ。証明は概念的にシンプルであり、d 次元空間におけるゾーン図において成立する。さらに、複数の幾何学物体に対するゾーン図、α-ゾーン図 (集合 R_i に含まれる点と $\mathbf{p}_i$ の距離が $\bigcup_{j \neq i} R_j$ への距離の α 倍以下。ここで $\alpha > 0$ は任意のパラメタ) など、ゾーン図の様々な自然な拡張に対して存在性の証明は適用できると思える。

一意性の証明では、ゾーン図を近似する自然な繰り返しプロシージャを考える。母点 $\mathbf{p}_1, \ldots, \mathbf{p}_n$ を固定し、$\mathbf{R} = (R_1, \ldots, R_n)$ を平面上の領域の n 組で $\mathbf{p}_1 \in R_1, \ldots, \mathbf{p}_n \in R_n$ を満たすものとする。これに対して、新しい領域の n 組 $\mathbf{Dom}(\mathbf{R}) = \mathbf{S} = (S_1, S_2, \ldots, S_n)$ を $S_i = \mathrm{dom}(\mathbf{p}_i, \bigcup_{j \neq i} R_j)$ として定義する。このようにすると、$\mathbf{Dom}$ は領域の n 組全体の集合に作用する作用素 (ドミナンス作用素) であり、ゾーン図の定義は下記のように書き直すことが出来る。

> 領域の n 組 $\mathbf{R} = (R_1, \ldots, R_n)$ が点集合 $\mathbf{p}_1, \ldots, \mathbf{p}_n$ のゾーン図であるとは、作用素 $\mathbf{Dom}$ の不動点である、すなわち、$\mathbf{R} = \mathbf{Dom}(\mathbf{R})$ であることである。

以下、二つの n 組 $\mathbf{R} = (R_1, \ldots, R_n)$ と $\mathbf{S} = (S_1, \ldots, S_n)$ に対して、$R_i \subseteq S_i$ $(i = 1, 2, \ldots, n)$ であるとき、$\mathbf{R} \preceq \mathbf{S}$ と表記する。定義から、$\mathbf{R} \preceq \mathbf{S}$ であれば $\mathbf{Dom}(\mathbf{R}) \succeq \mathbf{Dom}(\mathbf{S})$、すなわち $\mathbf{Dom}$ は $\preceq$ に関して反単調作用素である。初期状態 $\mathbf{I}^{(0)} = (\{\mathbf{p}_1\}, \ldots, \{\mathbf{p}_n\})$ を単一点集合領域からなるシステムとし、 $\mathbf{O}^{(0)} = (O_1^{(0)}, \ldots, O_n^{(0)}) = \mathbf{Dom}(\mathbf{I}^{(0)})$ を $\mathbf{p}_1, \ldots, \mathbf{p}_n$ たちの古典的なボロノイ領域からなるシステムとする。帰納的に $k = 1, 2, \ldots$ に対して $\mathbf{I}^{(k)} = \mathbf{Dom}(\mathbf{O}^{(k-1)})$, $\mathbf{O}^{(k)} = \mathbf{Dom}(\mathbf{I}^{(k)})$ と定義する。

作用素 $\mathbf{Dom}$ の反単調性と帰納的定義により $\mathbf{I}^{(0)} \preceq \mathbf{I}^{(1)} \preceq \mathbf{I}^{(2)} \preceq \cdots$ かつ $\mathbf{O}^{(0)} \succeq \mathbf{O}^{(1)} \succeq \mathbf{O}^{(2)} \succeq \cdots$ となる。さらに、もし R がゾーン図であれば、$\mathbf{R} = \mathbf{Dom}(\mathbf{R})$ かつ $\mathbf{I}^{(0)} \preceq \mathbf{R}$ なので、帰納法と反単調性から任意の k に対して $\mathbf{I}^{(k)} \preceq \mathbf{R} \preceq \mathbf{O}^{(k)}$ が成立する。すなわち図 12に示すように、$\mathbf{I}^{(k)}$ はゾーン図を内側から近似する増加列

であり、一方 $\mathbf{O}^{(k)}$ は外側から近似する減少列である。

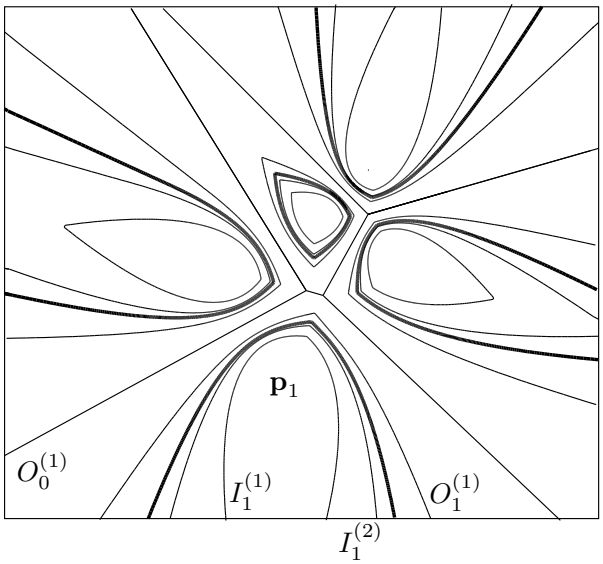


図 12: ゾーン図の内側近似と外側近似 $\mathbf{I}^{(k)}$ と $\mathbf{O}^{(k)}$.

内側近似と外側近似が同一の極限に収束することを示す事により、ゾーン図が一意に存在することが示される。この証明は同時にゾーン図を構成する実用的なアルゴリズムを与える。領域族 $\mathbf{I}^{(k)}$、$\mathbf{O}^{(k)}$ に属する領域を凸多角形を用いて近似することにより、証明に沿ったアルゴリズムの実行が可能になる。実際、図 12のゾーン図はこのアルゴリズムで描画したものである。実装に注意は要るが、これらの凸多角形が実際にゾーン図の領域の内側近似と外側近似になっているようにアルゴリズムを設計できる。小規模データでの実験結果では、このアルゴリズムの収束は非常に早いことが示された。

4.10 Web データや医療データにおけるデータマイニングの実用研究

Web データでのデータ切り分け

HTML で表現されたリスト形式のデータを切り分けて表形式のテーブルにする事は、知識抽出の道具として、さまざまの応用がある。これは、リストに対する詳細ページ情報を用いたデータ切り分けによって実現する事が出来、充足制約問題として定式化できる。全と修士学生の小沼らにより、従来の研究の実装と改良、書籍販売 Web データ等での実験を行い、更にグラフアルゴリズムを用いた高速計算法の研究を行った。研究成果の一部については研究会にて発表された。

医療データでのデータ抽出

データマイニング技術の実データへの適用の試みとして、博士研究員 Ashkan Sami、東北大学医学研究科永富教授との共同研究を行い、データマイニングにおけるデータ評価指標として、SDI(Shape Distribution Indicator) を提案し、医療の実データに適応し、いくつかの新しい知見を得た。成果はデータマイニングと AI 分野の 2 つの国際会議で発表を行った。

4.11 パラメトリック複雑度やグラフマイナーを考慮した新しいモデルでのアルゴリズム設計理論と解析

知識抽出に必要な最適化問題の多くはグラフ上の最適化問題として定式化されるが、それらの多くが NP 完全である事が知られている。しかしながら、グラフマイナー理論において、パーフェクトグラフ予想の解決や Hadwiger 予想 (4 色問題の拡張) での研究成果など、多くの理論的な道具立てがそろい、その結果、特定のマイナーを含まないグラフでは多くの最適化問題が効率的に解けるということが明らかにされてきている。河原林はこのテーマに網羅的に取り組み、多くの新しい成果をあげた。これらは最高レベルの国際会議 STOC 及び SODA で発表され、また、権威ある国際会議である ISAAC での最優秀論文賞を受けるなど理論的には非常に高い評価を上げている。但し、多項式時間解法ではあるが、計算時間は非常に掛かるので、知識抽出への利用はまだ先のことであり、実装等は行っていない。さらに、グラフマイナー理論から発展した特筆すべき成果は、交差数が定数であるグラフの最小交差数描画を求める線形時間アルゴリズムの開発である。これは河原林と、東北大学で本特定領域の研究費 (総括班) を用いて平成 18 年度に 2ヶ月間招聘したカナダ国マクギル大学の Bruce Reed 教授との共同研究で

ある。交差がなく平面に描画できるグラフ (平面グラフ) の描画を行うアルゴリズムは Hopcroft と Tarjan によって 1970 年代 (J. ACM, 1974) に提案され、アルゴリズム理論の一つの金字塔とされる。しかしながら、データマイニングの関係描画など多くの応用では、疎ではあるが必ずしも平面的でないグラフを視覚化させる必要があり、そのため、Hopcroft と Tarjan の手法を拡張し、与えられたグラフの最小交差数を計算し、かつ最小交差数を持つ描画の実現を行うアルゴリズム設計は広く知られていた未解決問題であった。河原林らは、この未解決問題を解決し、最小交差数が定数である場合は線形時間で作動するアルゴリズムの設計に成功した。数学的にも非常に深い結果であり、最高レベルの国際会議 STOC で 2007 年 6 月に発表された。

これに関連し、徳山は Halldórsson 教授 (アイスランド大学)、Knauer 教授 (ベルリン自由大学) Spillner 教授 (英国 East Anglia 大学) との共同研究で、グラフの非交差全域木の計算アルゴリズムに関して、入力グラフの交差数をパラメタとした時に従来より格段に高速なアルゴリズムの設計を与えた。

研究業績一覧

学術論文

1. K. Murota, A. Shioura:
   "Fundamental Properties of M-convex and L-convex Functions in Continuous Variables", IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, E87-A, 1042–1052, 2004 年 5 月.
   概要: The concepts of M-convexity and L-convexity, introduced by Murota (1996, 1998) for functions over the integer lattice, extract combinatorial structures in well-solved nonlinear combinatorial optimization problems. These concepts are extended to polyhedral convex functions and quadratic functions over the real space by Murota–Shioura (2000, 2004). In this paper, we consider a further extension to general convex functions. The main aim of this paper is to provide rigorous proofs for fundamental properties of general M-convex and L-convex functions.

2. S. Moriguchi, A. Shioura:
   "On Hochbaum's Scaling Algorithm for the General Resource Allocation Problem", Mathematics of Operations Research, 29, 394–397, 2004 年 5 月.
   概要: It is pointed out that the polynomial-time scaling algorithm by Hochbaum (1994) does not work correctly for the general resource allocation problem. We fix the algorithm so that it works correctly, and analyze the time complexity of the fixed algorithm.

3. J.Jansson, T. Tokuyama :
   "Semi-Balanced Coloring of Graphs: Generalized 2-Coloring Based on a Relaxed Discrepancy Condition", Graphs and Combinatorics, 20, 205–222, 2004 年 6 月.
   概要: Let G be an undirected, unweighted, connected graph with n vertices and m edges. We introduce the concept of a semi-balanced coloring of G, which is a 2-coloring of the shortest-paths hypergraph of G under a certain discrepancy condition yielding a class of independent sets of G. We prove that the number of different semi-balanced colorings of G is: (1) at most $n+1$ if G is bipartite; (2) at most m if G is non-bipartite and triangle-free; and (3) at most $m+1$ if G is non-bipartite. Based on the above combinatorial investigation, we design an algorithm that enumerates all semi-balanced colorings in $O(nm^2)$ time.

4. T. Asano, N. Katoh, H. Tamaki, T. Tokuyama :
   "On Geometric Structure of Global Roundings of Graphs and Range Spaces", Proceedings of the 9th Scandinavian Workshop on Algorithms, LNCS3111, 455–467, 2004

年7月.

概要: Given a hypergraph $H = (V, F)$ and a $[0,1]$-valued vector $\mathbf{a} \in [0,1]^V$, its global rounding is a binary (i.e.,$\{0,1\}$-valued) vector $\alpha \in \{0,1\}^V$ such that $|\sum_{v \in f}(\mathbf{a}(v) - \alpha(v))| < 1$ holds for each hyperedge $f \in F$. We study geometric (or combinatorial) structure of the set of global roundings of $\mathbf{a}$ using the notion of compatible set with respect to the discrepancy distance. We conjecture that the set of global roundings forms a simplex if the hypergraph satisfies "shortest-path" axioms, and prove it for some special cases including some geometric range spaces and the shortest path hypergraph of a series-parallel graph.

5. T. Ishikawa, K. Kawarabayashi, T. Tokuyama :
"On Properties of a Set of Global Roundings Associated with Clique Connection of Graphs", Interdisciplinary Information Sciences , 10(2), 159–163, 2004年9月.
概要: Given a connected weighted graph $G = (V, E)$, we consider a hypergraph $H(G) = (V, F(G))$ corresponding to the set of all shortest paths in G. For a given real assignment $\mathbf{a}$ on V satisfying $0 \leq \mathbf{a}(v) \leq 1$, a global rounding α with respect to $H(G)$ is a binary assignment satisfying that $|\sum_{v \in F} \mathbf{a}(v) - \alpha(v)| < 1$ for every $F \in F(G)$. Asano *et al.* conjectured that there are at most $|V| + 1$ global roundings for $H(G)$. In this paper, we present monotone properties on size and affine corank of a set of global roundings under a clique connection operation.

6. D. Chen, J. Chun, N. Katoh, T. Tokuyama :
"Efficient Algorithms for Approximating a Multi-Dimensional Voxel Terrain by a Unimodal Terrain", Proceedings of the 10th Computing and Combinatorics Conference, LNCS3126, 238–248, 2004年8月.
概要: We consider the problem of approximating a function on a d-dimensional voxel grid by a unimodal function to minimize the L_2 approximation error with respect to a given measure distribution on the grid. The output unimodal function gives a layered structure on the voxel grid, and we give efficient algorithms for computing the optimal approximation under a reasonable assumption on the shape of each horizontal layer. Our main technique is a dominating cut algorithm for a graph.

7. K. Murota, A. Shioura:
"Quadratic M-convex and L-convex Functions", Advances in Applied Mathematics, 33, 318–341, 2004年8月.
概要: The concepts of L-convexity and M-convex are introduced by Murota (1996, 1998) for functions defined over the integer lattice, and then extended to polyhedral convex functions by Murota–Shioura (2000). In this paper, we consider these concepts for quadratic functions and investigate the structure of the coefficient matrices of such quadratic functions. It is shown that quadratic L-convex and M-convex functions can be characterized by nice combinatorial properties of their coefficient matrices.

8. T. Asano, N. Katoh, H. Tamaki, T. Tokuyama :
"The Structure and Number of Global Roundings of a Graph", Theoretical Computer Science, 325, 425–437, 2004年10月.
概要: Given a connected weighted graph $G = (V, E)$, we consider a hypergraph $H_G = (V, P_G)$ corresponding to the set of all shortest paths in G. For a given real assignment $\mathbf{a}$ on V satisfying $0 \leq \mathbf{a}(v) \leq 1$, a global rounding α with respect to H_G is a binary assignment satisfying that $|\sum_{v \in F} \mathbf{a}(v) - \alpha(v)| < 1$ for every $F \in P_G$. We conjecture that there are at most $|V|+1$

global roundings for H_G, and also the set of global roundings is an affine independent set. We give several positive evidences for the conjecture.

9. B. Aronov, T. Asano, N. Katoh, K. Mehlhorn, T. Tokuyama :
"Polyline Fitting of Planar Point Set Under Min-Sum Criteria", Proceedings of 15th International Symposium on Algorithms and Computation, LNCS3341, 77-88, 2004年12月.
概要: Fitting a curve of a certain type to a given set of points in the plane is a basic problem in statistics and has numerous applications. We consider fitting a polyline with k joints under the min-sum criteria with respect to L_1- and L_2-metrics, which are more appropriate measures than uniform and Hausdorff metrics in statistical context. We present efficient algorithms for the 1-joint versions of the problem, and fully polynomial-time approximation schemes for the general k-joint versions.

10. K. Murota, A. Shioura:
"Conjugacy Relationship Between M-convex and L-convex Functions in Continuous Variables", Mathematical Programming, 101, 415–433, 2004年12月.
概要: By extracting combinatorial structures in well-solved nonlinear combinatorial optimization problems, Murota (1996, 1998) introduced the concepts of M-convexity and L-convexity to functions defined over the integer lattice. Recently, Murota–Shioura (2000, 2004) extended these concepts to polyhedral convex functions and quadratic functions in continuous variables. In this paper, we consider a further extension to more general convex functions defined over the real space, and provide a proof for the conjugacy relationship between general M-convex and L-convex functions.

11. R. Farooq, A. Shioura:
"A Note on the Equivalence Between Substitutability and M$^\natural$-convexity", Pacific Journal of Optimization, 1, 243–252, 2005年1月.
概要: The property of "substitutability" plays a key role in guaranteeing the existence of a stable solution in the stable marriage problem and its generalizations. On the other hand, the concept of M$^\natural$-convexity, introduced by Murota–Shioura (1999) for functions defined over the integer lattice, enjoys a number of nice properties that are expected of "discrete convexity" and provides with a natural model of utility functions. In this note, we show that M$^\natural$-convexity is characterized by two variants of substitutability.

12. K. Sadakane, N. Takki-Chebihi, T. Tokuyama :
"Combinatorics and Algorithms for Low-Discrepancy Roundings of a Real Sequence", Theoretical Computer Science, 331, 23–36, 2005年2月.
概要: We discuss the problem of computing all the integer sequences obtained by rounding an input sequence of n real numbers such that the discrepancy between the input sequence and each output binary sequence is less than one. The problem arises in the design of digital halftoning methods in computer graphics. We show that the number of such roundings is at most $n+1$ if we consider the discrepancy with respect to the set of all subintervals, and give an efficient algorithm to report all of them. Then, we give an optimal method to construct a compact graph to represent the set of global roundings satisfying a weaker discrepancy condition.

13. T. Tokuyama :
"Minmax Parametric Optimization Problems and Multidimensional Parametric

Searching", Interdisciplinary Information Sciences, 11-1, 1–9, 2005年3月.

概要: The parametric minimax problem, which finds the parameter value minimizing the weight of a solution of a combinatorial maximization problem, is a fundamental problem in sensitivity analysis. Moreover, several problems in computational geometry can be formulated as parametric minimax problems. The *parametric search* paradigm gives an efficient sequential algorithm for a convex parametric minimax problem with one parameter if the original non-parametric problem has an efficient parallel algorithm. We consider the parametric minimax problem with d parameters for a constant d, and solve it by using multidimensional version of the parametric search paradigm. As a new feature, we give a feasible region in the parameter space in which the parameter vector must be located. Typical results obtained as applications are: (1) Efficient solutions for some geometric problems, including theoretically efficient solutions for the minimum diameter bridging problem in d-dimensional space between convex polytopes. (2) Solutions for parametric polymatroid optimization problems, including an $O(n \log n)$ time algorithm to compute the parameter vector minimizing k-largest linear parametric elements with d dimensions.

14. K. Ohta, K. Sadakane, A. Shioura, T. Tokuyama:
"A Fast, Accurate and Simple Method for Pricing European-Asian and Saving-Asian Options", Algorithmica, 42, 141–158, 2005年4月.

概要: We propose an efficient and accurate randomized approximation algorithm for computing the price of European-Asian options. Our algorithm can be seen as a modification of the approximation algorithm developed by Aingworth et al. (2000) into a randomized algorithm, which improves the accuracy theoretically as well as practically. We also propose a new option named the Saving-Asian option which enjoys advantages of both the European-Asian and American-Asian options. It is shown that our approximation algorithm also works for pricing Saving-Asian options.

15. A. Shioura, T. Tokuyama:
"Efficiently Pricing European-Asian Options: Ultimate Implementation and Analysis of the AMO Algorithm", Proceedings of the 1st International Conference on Algorithmic Applications in Management (AAIM 2005), LNCS 3521, 291–300, 2005年6月.

概要: We propose an efficient and accurate randomized approximation algorithm for pricing a European-Asian option on the binomial tree model. For an option with the strike price X on an n-step binomial tree and any positive integer k, our algorithm runs in $\mathrm{O}(kn^2)$ time with the error bound $\mathrm{O}(X/k)$ which is independent of n. Our algorithm is a modification of the approximation algorithm developed by Aingworth, Motwani, and Oldham (2000) into a randomized algorithm, which improves the accuracy theoretically as well as practically.

16. J. Chun, K. Sadakane, T. Tokuyama:
"Construction of Expert Guided Decision Tree on Numeric Database", Proceedings of the 8th Japan-Korea Workshop on Algorithm and Computation, WAAC05, 27–34, 2005年8月.

概要: In this paper, we propose the technique for constructing a decision tree with high prediction accuracy. There is over fitting as a problem at the time of constituting a decision tree. If the optimal decision tree is used to training data, trade-off that the prediction accuracy over unknown data will fall will arise. In this paper, we propose a

new method for removing these drawbacks. Indeed, we give optimized numeric association rules with multi attributes, and give layered-structure numeric association rules. We use layered-structure numeric association rules as expert of decision tree.

17. J. Chun, K. Sadakane, T. Tokuyama, M. Yuki:
"Peak-Reducing Fitting of a Curve under the L_p Metric", Interdisciplinary Information Sciences, 11(2), 191–197, 2005 年 9 月.
概要: Given a function $y = f(x)$ in one variable, we consider the problem of computing a k-peaked curve $y = \phi(x)$ minimizing the L_p distance between them. In other words, $\phi(x)$ has at most k local peaks and minimizes the area bounded by the curves $f(x)$ and $\phi(x)$. This gives extension of the authors' previous work on the unimodal (i.e., single-peaked) approximation for the L_2 distance.

18. K. Murota, A. Shioura:
"Substitutes and Complements in Network Flows Viewed as Discrete Convexity", Discrete Optimization, 2, 256–268, 2005 年 9 月.
概要: We study combinatorial properties of the optimal value function of the network flow problem. It is shown by Gale–Politof [Substitutes and complements in networks flow problems, Discrete Appl. Math. 3 (1981) 175–186] that the optimal value function has submodularity and supermodularity w.r.t. problem parameters such as weights and capacities. In this paper we shed a new light on this result from the viewpoint of discrete convex analysis to point out that the submodularity and supermodularity are naturally implied by discrete convexity, called M-convexity and L-convexity, of the optimal value function.

19. E. Demaine, M. Hajiaghayi, K. Kawarabayashi:
"Algorithmic Graph Minor Theory: Decomposition, Approximation, and Coloring", Proceedings of the 46th IEEE Symposium on Foundation of Computer Science, FOCS 2005, 637–646, 2005 年 10 月.
概要: At the core of the seminal Graph Minor Theory of Robertson and Seymour is a powerful structural theorem capturing the structure of graphs excluding a fixed minor. This result is used throughout graph theory and graph algorithms, but is existential. We develop a polynomial-time algorithm using topological graph theory to decompose a graph into the structure guaranteed by the theorem: a clique-sum of pieces almost-embeddable into bounded-genus surfaces. This result has many applications. In particular, we show applications to developing many approximation algorithms, including a 2-approximation to graph coloring, constant-factor approximations to treewidth and the largest grid minor, combinatorial polylogarithmic-approximation to half-integral multicommodity flow, subexponential fixed-parameter algorithms, and PTASs for many minimization and maximization problems, on graphs excluding a fixed minor.

20. A. Suzuki, T. Tokuyama:
"Dense Subgraph Problems with Output-Density Conditions", Proceedings of the 16th Annual International Symposium on Algorithms and Computation, LNCS 3837, 266–276, 2005 年 12 月.
概要: We consider the dense subgraph problem that extracts a subgraph with a prescribed number of vertices that has the maximum number of edges (total edge weight in the weighted case) in a given graph. We give approximation algorithms with improved theoretical approximation ratios assuming that the density of the optimal output subgraph is high, where density is the ratio of number of edges (or sum

of edge weights) to the number of edges in the clique on the same number of vertices. Moreover, we investigate the case where the input graph is bipartite, and design a pseudo-polynomial time approximation scheme that can become a PTAS even if the size of the optimal output graph is comparatively small. This is a significant improvement in a theoretical sense, since no constant-ratio approximation algorithm was known previously if the output graph has $o(n)$ vertices.

21. J. Chun, K. Sadakane, T. Tokuyama: "Linear Time Algorithm for Approximating a Curve by a Single-peaked Curve", Algorithmica, 44(2), 103–115, 2006 年 2 月.
概要: Given a function $y = f(x)$ in one variable, we consider the problem of computing the single-peaked (unimodal) curve $y = \phi(x)$ minimizing the L_2-distance between them. If the input function f is a histogram with $O(n)$ steps or a piecewise linear function with $O(n)$ linear pieces, we design algorithms for computing ϕ in linear time. We also give an algorithm to approximate f with a function consisting of the minimum number of unimodal pieces under the condition that each unimodal piece is within a fixed L_2-distance from the corresponding portion of f.

22. J. Chun, K. Sadakane, T. Tokuyama: "Efficient Algorithms for Constructing a Pyramid from a Terrain", IEICE Transactions on Informations and Systems, E89-D(2), 783–788, 2006 年 2 月.
概要: Given nonnegative valued functions ρ and μ in d variables, we consider the optimal pyramid maximizing the total parametric gain of ρ against μ. The pyramid can be considered as the optimal unimodal approximation of ρ relative to μ, and can be applied to hierarchical data segmentation. In this paper, we give efficient algorithms for a couple of two-dimensional pyramid construction problems.

23. K. Kawarabayashi, B. Mohar: "Approximating the List-chromatic Number and the Chromatic Number in Minor-closed and Odd-minor-closed Classes of Graphs", Proceedings of the 38th ACM Symposium on Theory of Computing, STOC 2006, 401–406, 2006 年 5 月.
概要: It is well-known (Feige and Kilian, Hastad) that approximating the chromatic number within a factor of $n^{1-\epsilon}$ cannot be done in polynomial time for $\epsilon > 0$, unless coRP = NP. Also, it is known that computing the list-chromatic number is much harder than the chromatic number. In fact, it is known that the problem of deciding if the list-chromatic number is k for $k \geq 3$ is Π_2^p-complete. In this paper, we focus on minor-closed and odd-minor-closed families of graphs. In doing that, we may as well consider only graphs without K_k-minors and graphs without odd K_k-minors for a fixed value of k, respectively. Our main results are that there is a polynomial time approximation algorithm for the list-chromatic number of graphs without K_k-minors and there is a polynomial time approximation algorithm for the chromatic number of graphs without odd-K_k-minors. Their time complexity are $O(n^3)$ and $O(n^4)$, respectively. The algorithms have multiplicative error $O(\sqrt{\log k})$ and additive error $O(k)$, and the multiplicative error occurs only for graphs whose list-chromatic number and chromatic number are $\Theta(k)$, respectively. Let us recall that H has an odd complete minor of order at least l if there are l vertex disjoint trees in H such that every two of them are joined by an edge, and in addition, all the vertices of trees are two-colored in such a way that the edges within the trees are bichromatic, but the edges between trees are monochromatic.

Hence it is easy to see that odd minor is a generalization of minor. Let us observe that the complete bipartite graph $K_{n/2,n/2}$ certainly contains a K_k-minor for $k \leq n/2$, but on the other hand, it does not contain any odd K_k-minor for any $k \geq 3$. Odd K_5-minor-free graphs are closely related to one field of discrete optimization which is finding conditions under which a given polyhedron has integer vertices, so that integer optimization problems can be solved as linear programs. Also, the odd version of the well-known Hadwiger's conjecture has been paid attention by many researchers. This is our motivation to consider odd minor-closed class of graphs. Our main idea involves precoloring extension. This idea is used in many results; one example is Thomassen's proof on his celebrated theorem on planar graphs. The best previously known approximation for the first result is a simple $O(k\sqrt{\log k})$-approximation following algorithm that guarantees a list-coloring with $O(k\sqrt{\log k})$ colors for K_k-minor-free graphs. The best previous approximation for the second result comes from the recent result of Geelen et al. who gave an $O(k\sqrt{\log k})$-approximation algorithm. We also relate our algorithm to the well-known conjecture of Hadwiger and its odd version of the conjecture. In fact, we give an $O(n^3)$ algorithm to decide whether or not a weaker version of Hadwiger's conjecture is true. Here, by a weaker version of Hadwiger's conjecture, we mean a conjecture which says that any $27k$-chromatic graph contains a K_k-minor. Also, we shall give an $O(n^{2500k})$ algorithm for deciding whether or not any $2500k$-chromatic graph contains an odd-K_k-minor.

24. T. Asano, J. Matoušek, T. Tokuyama:
"The Distance Trisector Curve", Proceedings of the 38th ACM Symposium on Theory of Computing, STOC 2006, 336–343, 2006 年 5 月.
概要: Given points p and q in the plane, we are interested in separating them by two curves C_1 and C_2 such that every point of C_1 has equal distance to p and to C_2, and every point of C_2 has equal distance to C_1 and to q. We show by elementary geometric means that such C_1 and C_2 exist and are unique. Moreover, for $p = (0, 1)$ and $q = (0, -1)$, C_1 is the graph of a function $f : \mathcal{R} \to \mathcal{R}$, C_2 is the graph of $-f$, and f is convex and analytic (i.e., given by a convergent power series at a neighborhood of every point). We conjecture that f is not expressible by elementary functions and, in particular, not algebraic. We provide an algorithm that, given $x \in \mathcal{R}$ and $\epsilon > 0$, computes an approximation to $f(x)$ with error at most ϵ in time polynomial in $\log \frac{1+|x|}{\epsilon}$. The separation of two points by two "trisector" curves considered here is a special (two-point) case of a new kind of Voronoi diagram, which we call the *zone diagram* and which we investigate in a companion paper.

25. B. Aronov, T. Asano, N. Katoh, K. Mehlhorn, T. Tokuyama:
"Polyline Fitting of Planar Points under Min-sum Criteria", International Journal of Computational Geometry and Applications, 16(2-3), 97–116, 2006 年 6 月.
概要: Fitting a curve of a certain type to a given set of points in the plane is a basic problem in statistics and has numerous applications. We consider fitting a polyline with k joints under the min-sum criteria with respect to L_1- and L_2-metrics, which are more appropriate measures than uniform and Hausdorff metrics in statistical context. We present efficient algorithms for the 1-joint versions of the problem and fully polynomial-time approximation schemes for the general k-joint versions.

26. 徳山豪:
"関数近似における幾何学アルゴリズムの最近の進展：データ解析への応用に向けて（招待解説論文)", 電子情報通信学会論文誌 A, J89-A, 419–429, 2006 年 6 月.
概要: データの単純化は現代の情報社会では非常に重要なテーマである。生のデータは通常巨大なサイズを持ち，分布関数としてみるとサンプル誤差や例外データ，誤入力などによる揺らぎを持っている。したがって，データマイニングなどのデータ解析においては，入力データをシンプルな関数に近似して利用する事が重要である。関数の近似においてはさまざまなアプローチがあり，関数解析的な手法，学習による手法などとともに離散アルゴリズムを用いた計算幾何学的な最適近似のアプローチは広く研究されている。しかしながら，計算幾何学的な手法においては，従来の応用はパターンマッチなどであり，データ解析に用いる場合は定式化や最適化基準を適当なものに変更する必要がある。それにしたがってアルゴリズム理論上のさまざまな困難や計算限界が生じ，それらを打破あるいは回避する必要が生じる。本稿では，計算幾何学におけるデータ近似の説明と現状の上で，著者たちの最近の研究を中心にした動向を解説する。

27. A. Shioura, N. Sun, Z. Yang:
"Efficient Strategy Proof Fair Allocation Algorithms", Journal of the Operations Research Society of Japan, 42 (2), 144–150, 2006 年 6 月.
概要: We study a fair division problem with indivisible objects like jobs, houses, and one divisible good like money. Each individual is to be assigned with one object and a certain amount of money. The preferences of individuals over the objects are private information but individuals are assumed to have quasi-linear utilities in money. It is shown that there exist efficient algorithms for eliciting honest preferences and assigning the objects with money to individuals efficiently and fairly.

28. M. Halldórsson, T. Tokuyama:
"Minimizing Interference of a Wireless Ad-Hoc Network in a Plane", ALGOSENSORS 2006, LNCS 4240, 71–82, 2006 年 7 月.
概要: We consider the problem of topology control of a wireless ad-hoc network on a given set of points in the plane, where we aim to minimize the maximum interference by assigning a suitable transmission radius to each point. By using computational geometric ideas and ϵ-net theory, we attain an $O(\sqrt{\Delta})$ bound for the maximum interference where Δ is the interference of a uniform-radius ad-hoc network. This generalizes a result given by Rickenbach et al. for the special case of highway model (i.e., one-dimensional problem) to the two-dimensional case. We also give a method based on quad-tree decomposition and bucketing that has another provable interference bound in terms of the ratio of the minimum distance to the radius of a uniform-radius ad-hoc network.

29. T. Ito, Y. Okamoto, T. Tokuyama:
"Algorithms for the Full Steiner Tree Problem", Proceedings of iETA (improving Exponential Time Algorithms), 39–45, 2006 年 7 月.
概要: Assume that each edge of a graph $G = (V, E)$ is assigned a nonnegative real number cost, and that a set $K \subseteq V$ of vertices, called *terminals*, is given. A *full Steiner tree* (or a *terminal Steiner tree*) of G with respect to K is a subtree of G whose leaf set is K. The (minimum-cost) *full Steiner tree problem* is to find a full Steiner tree for K which has the minimum total cost of edges. In this paper we first give an exact algorithm to solve the problem for general graphs in time $O(3^k)$ where $k = |K|$, and hence it shows the fixed-parameter tractability of the problem. We then show that the problem can be solved in polynomial time for plane graphs if all

the terminals lie only on a fixed constant number of face boundaries.

30. M. Halldórsson, A. Wolff, and T. Tokuyama:
"Improved Fixed-Parameter Algorithms for Non-crossing Subgraphs", Proceedings of iETA (improving Exponential Time Algorithms), 31–37, 2006年7月.
概要: We consider the problem of computing non-crossing spanning trees in topological graphs. It is known that it is NP-hard to decide whether a topological graph has a non-crossing spanning tree, and that it is hard to approximate the minimum number of crossings in a spanning tree. We give improved fixed-parameter algorithms, where we use the number k of crossing edge pairs and the number μ of crossing edges in the given graph as parameters. The running times of our algorithms are $O^*(k^{O(\sqrt{k})})$ and $O^*(\mu^{O(\mu^{2/3})})$, where the O^*-notation neglects polynomial terms. Our method can be applied to several other non-crossing subgraph problems in topological graphs.

31. T. Asano, N. Katoh, H. Tamaki, T. Tokuyama:
"Angular Voronoi Diagram with Applications", Proceedings of the 3rd International Symposium on Voronoi Diagrams in Science and Engineering, ISVD 2006, 18–24, 2006年7月.
概要: Voronoi diagram for a set of geometric objects is a partition of the plane (or space in higher dimensions) into disjoint regions each dominated by some given object under a predetermined criterion. This paper defines a new Voronoi diagram called an angular Voronoi diagram for a set of line segments in the plane and investigate its combinatorial and structural properties. Given a set of line segments, an angular Voronoi diagram partitions the plane into regions so that a point belongs to a Voronoi region of a line segment if the visual angle of the line segment from the point is smallest (or largest) among all line segments. For simplicity we assume that given line segments are transparent when computing their visual angles. In other words, the visual angle of a line segment from a point is defined without considering other line segments.

32. A. Sami, R. Nagatomi, M. Takahashi, T. Tokuyama:
"OSDM: Optimized Shaper Distribution Method", ADMA2006, LNCS 4093, 1057–1064, 2006年8月.
概要: Comprehensibility is vital in results of medical data mining systems since doctors simply require it. Another important issue specific to some data sets, like Fitness, is their uniform distribution due to tile analysis that was performed on them. In this paper, we propose a novel data mining tool named OSDM (Optimized Shape Distribution Method) to give a comprehensive view of correlations of attributes in cases of uneven frequency distribution among different values of symptoms. We apply OSDM to explore the relationship of the Fitness data and symptoms in medical test dataset for which popular data mining methods fail to give an appropriate output to help doctors decisions. In our experiment, OSDM found several useful relationships.

33. A. Sami, R. Nagatomi, M. Takahashi, T. Tokuyama:
"SDI: Shape Distribution Indicator and its Application to Find Interrelationships between Physical Activity Tests and Other Medical Measures", Proceedings of Australian Joint Conference on Artificial Intelligence 2006, LNCS 4304, 383–392, 2006年12月.
概要: Comprehensibility is driving force in medical data mining results since doctors

utilize the outputs and give the final decision. Another important issue specific to some data sets, like physical activity, is their uniform distribution due to tile analysis that was performed on them. In this paper, we propose a novel data mining tool named SDI (Shape Distribution Indicator) to give a comprehensive view of co-relations of attributes together with an index named ISDI to show the robustness of SDI outputs. We apply SDI to explore the relationship of the Physical Activity data and symptoms in medical test dataset for which popular data mining methods fail to give an appropriate output to help doctors decisions. In our experiment, SDI found several useful relationships.

34. E.D. Demaine, M. Hajiaghayi, K. Kawarabayashi:
"Algorithmic Graph Minor Theory: Improved Grid Minor Bounds and Wagner's Contraction", Proceedings of the 17th International Symposium on Algorithms and Computation, LNCS 4288, 3–15, 2006 年 12 月.
概要: We explore the three main avenues of research still unsolved in the field of algorithmic graph-minor theory, which all stem from a key min-max relation between the treewidth of a graph and its largest grid minor. This min-max relation is a keystone of the Graph Minor Theory of Robertson and Seymour, which ultimately proves Wagner's Conjecture about the structure of minor-closed graph properties. First, we obtain the only known polynomial min-max relation for graphs that do not exclude any fixed minor, namely, map graphs and power graphs. Second, we obtain explicit (and improved) bounds on the min-max relation for an important class of graphs excluding a minor, namely, $K_{3,k}$-minor-free graphs, using new techniques that do not rely on Graph Minor Theory. Third, we disprove Wagner's Conjecture in the context of graph contractions for general graphs, and in a sense characterize which graphs satisfy the conjecture. Our results have several combinatorial and algorithmic implications.

35. A. Shioura, T. Tokuyama:
"Efficiently Pricing European-Asian Options: Ultimate Implementation and Analysis of the AMO Algorithm", Information Processing Letters, 100(6), 213–219, 2006 年 12 月.
概要: We propose an efficient and accurate randomized approximation algorithm for pricing a European-Asian option on the binomial tree model. For an option with the strike price X on an n-step binomial tree and any positive integer k, our algorithm runs in $\mathrm{O}(kn2)$ time with the error bound $\mathrm{O}(X/k)$ which is independent of n. Our algorithm is a modification of the approximation algorithm developed by Aingworth, Motwani, and Oldham (2000) into a randomized algorithm, which improves the accuracy theoretically as well as practically.

36. T. Asano, J. Matoušek, T. Tokuyama:
"Zone Diagrams, Existence, Uniqueness, and Algorithmic Challenge", Proceedings of the 18th Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2007, 756–765, 2007 年 1 月.
概要: A *zone diagram* is a new variation of the classical notion of Voronoi diagram. Given points (sites) $p_1, \ldots, p_n$ in the plane, each p_i is assigned a region R_i, but in contrast to the ordinary Voronoi diagrams, the union of the R_i has a nonempty complement, the *neutral zone*. The defining property is that each R_i consists of all $x \in \mathcal{R}^2$ that lie closer (non-strictly) to p_i than to the union of all the other R_j, $j \neq i$. Thus, the zone diagram is defined implicitly, by a "fixed-point property," and neither its exis-

tence nor its uniqueness seem obvious. We establish existence using a general fixed-point result (a consequence of Schauder's theorem or Kakutani's theorem); this proof should generalize easily to related settings, say higher dimensions. Then we prove uniqueness of the zone diagram, as well as convergence of a natural iterative algorithm for computing it, by a geometric argument, which also relies on a result for the case of two sites in an earlier paper. Many challenging questions remain open.

37. K. Kawarabayashi:
"Half-integral packing, Erdos-Posa Property and Graph Minors", Proceedings of the 18th ACM-SIAM Symposium on Discrete Algorithms, SODA 2007, 1187–1196, 2007 年 1 月.
概要: We consider the following problem: Given a graph G. Decide whether or not G has k half-integral packing of an H-minor for fixed k and H. Here, "k half-integral packing of an H-minor" is that there are k H-minors such that each vertex is used at most twice. Let us observe that if G contains H as a minor, then G has $|H|$ disjoint trees such that after contracting all the trees, we can get an H. So k half-integral packing of an H-minor means that there are $k|H|$ trees $T_{1,1}, \ldots, T_{1,|H|}, T_{2,1}, \ldots, T_{k,|H|}$ such that trees $T_{i,1}, \ldots, T_{i,|H|}$ are disjoint for $1 \leq i \leq k$, and after contracting all these trees $T_{i,1}, \ldots, T_{i,|H|}$, we can get an H for $1 \leq i \leq k$, but each vertex is used at most twice in these trees $T_{1,1}, \ldots, T_{1,|H|}, T_{2,1}, \ldots, T_{k,|H|}$. This problem is motivated by the study of the half disjoint paths problem and unspilittable flow problem, . Let us observe that when $k = 1$, this case is exactly testing a given minor. So our problem includes Robertson-Seymour's result. Also, it would not follow from the existence of kH-minors. Furthermore, the algorithm of the half disjoint paths problem by Kleinberg does not give our algorithm. Hence, our setting is not directly comparable to Robertson and Seymour's result, and Kleinberg's result. We prove that when $H = K_6$ or K_7, this problem is solvable in polynomial time. Our motivation for the K_6-minor is that the structure of graphs with no K_6-minors. Unfortunately, it is not known at all, although the structure of graphs without K_5-minors is well known from 1930's. Moreover, we prove that Erdős-Pósa property holds for k half-integral packing of a K_6-minor (a K_7-minor, respectively). More precisely, either G has k half-integral packing of a K_6-minor (a K_7-minor, respectively) or G has a vertex set T of order at most $f(k)$ for some function of k such that $G - T$ does not contain any K_6-minor (K_7-minor, respectively). This settles the conjecture of Thomas for the K_6-minor case and the K_7-minor case. Finally, we shall address why our approach works for the K_6-minor case and the K_7-minor case, but does not work for the K_8-minor case.

38. T. Tokuyama:
"Recent Progress on Geometric Algorithms for Approximating Functions：Towards Applications to Data Analysis", Electronics and Communications in Japan, 90(3), 1–12, 2007 年 3 月.
概要: Data simplification is an extremely important issue in our current information-oriented society. Normally, a real-world database contains there exists a massive amount of raw data, and when we an attempt is made to consider the that data as a distribution function, it has fluctuations due to sampling errors, outliersexception data, and/or invalid inputs. Therefore, for data analysis technology such as data mining, it is important to approximate the input data by a simplifiede function. There are various approaches to function approximation, and including functional analyti-

cal methods techniques, and learning-based techniques are quite popular. and computational geometric Apart from them, computational geometric approach based on optimizational approximation approaches using discrete algorithms, which have been is widely studied. However, the conventional application of computational geometrical techniques is pattern matching, and to applyuse them to for data analysis, their formulation and or optimization criteria must be changed accordingly. Therefore, various difficulties and or computational barriers arise, which must be eliminated or avoided. In this paper, we discuss the author explains data approximation in computational geometry and describes current trends centered on the author's his latest research.

39. A. Shioura, K. Tanaka:
"Polynomial-time Algorithms for Linear and Convex Optimization on Jump Systems", SIAM Journal on Discrete Mathematics, 21(2), 504–522, 2007 年 6 月.
概要: The concept of jump system, introduced by Bouchet and Cunningham (1995), is a set of integer points with a certain exchange property. In this paper, we discuss several linear and convex optimization problems on jump systems and show that these problems can be solved in polynomial time under the assumption that a membership oracle for a jump system is available. We firstly present a polynomial-time implementation of the greedy algorithm for the minimization of a linear function. We then consider the minimization of a separable-convex function on a jump system, and propose the first polynomial-time algorithm for this problem. The algorithm is based on the domain reduction approach developed in Shioura (1998). We finally consider the concept of M-convex functions on constant-parity jump systems which has been recently proposed by Murota (2006). It is shown that the minimization of an M-convex function can be solved in polynomial time by the domain reduction approach.

40. T. Asano, J. Matoušek, T. Tokuyama:
"The Distance Trisector Curve", Advances in Mathematics, 212 (1), 338–360, 2007 年 6 月.
概要: Given points p and q in the plane, we are interested in separating them by two curves C_1 and C_2 such that every point of C_1 has equal distance to p and to C_2, and every point of C_2 has equal distance to C_1 and to q. We show by elementary geometric means that such C_1 and C_2 exist and are unique. Moreover, for $p = (0, 1)$ and $q = (0, -1)$, C_1 is the graph of a function $f : \mathcal{R} \to \mathcal{R}$, C_2 is the graph of $-f$, and f is convex and analytic (i.e., given by a convergent power series at a neighborhood of every point). We conjecture that f is not expressible by elementary functions and, in particular, not algebraic. We provide an algorithm that, given $x \in \mathcal{R}$ and $\epsilon > 0$, computes an approximation to $f(x)$ with error at most ϵ in time polynomial in $\log \frac{1+|x|}{\epsilon}$. The separation of two points by two "trisector" curves considered here is a special (two-point) case of a new kind of Voronoi diagram, which we call the *zone diagram* and which we investigate in a companion paper.

41. J. Chun, Y. Okada, T. Tokuyama:
"Distance Trisector of Segments and Zone Diagram of Segments in a Plane", Proceedings of the 4th International Symposium on Voronoi Diagrams in Science and Engineering, ISVD 2007, 66–73, 2007 年 7 月.
概要: Motivated by the work of Asano et al., we consider the distance trisector problem and Zone diagram considering segments in the plane as the input geometric objects. As the most basic case, we first consider the pair of curves (distance trisector curves) tri-

secting the distance between a point and a line. This is a natural extension of the bisector curve (that is a parabola) of a point and a line. In this paper, we show that these trisector curves C_1 and C_2 exist and are unique. We then give a practical algorithm for computing the Zone diagram of a set of segments in a digital plane.

42. T. Asano, N. Katoh, H. Tamaki, T. Tokuyama:
"Voronoi Diagram with Respect to Criteria on Vision Information", Proceedings of the 4th International Symposium on Voronoi Diagrams in Science and Engineering, ISVD 2007, 25–32, 2007 年 7 月.
概要: Voronoi diagram for a set of geometric objects is a partition of the plane (or space in higher dimensions) into disjoint regions each dominated by some given object under a predetermined criterion. In this paper we are interested in various measures associated with criteria on goodness of an input line segment with respect to each point in the plane as the "point of view." These measures basically show how the segment or information displayed on the segment can be seen from the point. Mathematically, the measures are defined in terms of the shape of the triangle determined by the point and the line segment. Given any such measure, we can define a Voronoi diagram for a set of line segments. In this paper we are interested in investigating their common combinatorial and structural properties. We investigate conditions for those measures to define regular Voronoi diagrams and also conditions that local optima on the measures lie only on Voronoi edges, not in the proper interior of Voronoi regions.

43. T. Tokuyama:
"Recent Progress on Combinatorics and Algorithms for Low Discrepancy Roundings", Graphs and Combinatorics, 23, 359–378, 2007 年 7 月.
概要: Given a $[0,1]$-valued n-dimensional vector $a = (a_1, a_2, \ldots, a_n) \in [0,1]^V$ indexed by a set $V = \{v_1, v_2, \ldots, v_n\}$, we consider the problem of approximating a by a binary (i.e., $\{0,1\}$-valued) vector $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \{0,1\}^V$ under the discrepancy measure with respect to a hypergraph $\mathcal{H} = (V, \mathcal{F})$. We are interested in the properties of low-discrepancy roundings. Especially, we survey recent works on the combinatorial properties of a global rounding; that is, rounding whose discrepancy is less than 1.

44. M. M. Halldórsson, C. Knauer, A. Spillner, T. Tokuyama:
"Fixed-Parameter Tractability for Non-Crossing Spanning Trees", Proceedings of the 10th International Workshop on Algorithms and Data Structures, LNCS 4619, 410–421, 2007 年 8 月.
概要: We consider the problem of computing non-crossing spanning trees in topological graphs. It is known that it is NP-hard to decide whether a topological graph has a non-crossing spanning tree, and that it is hard to approximate the minimum number of crossings in a spanning tree. We consider the parametric complexities of the problem for the following natural input parameters: the number k of crossing edge pairs, the number μ of crossing edges in the given graph, and the number ι of vertices in the interior of the convex hull of the vertex set. We start with an improved strategy of the simple search-tree method to obtain an $O^*(1.93^k)$-time algorithm. We then give more sophisticated algorithms based on graph separators, with a novel technique to ensure connectivity. The time complexities of our algorithms are $O^*(2^{O(\sqrt{k})})$, $O^*(\mu^{O(\mu^{2/3})})$, and $O^*(2^{O(\sqrt{\iota})})$. By giving a reduction from 3-SAT, we show that the $O^*(2^{\sqrt{k}})$ complexity is hard to improve un-

der a hypothesis of the complexity of 3-SAT.

研究会等

1. 塩浦 昭義, 徳山 豪:
“ヨーロピアンアジアンオプションの効率的な価格付け手法: AMO アルゴリズムの実装と解析の改良”, 情報処理学会アルゴリズム研究会, 2004-AL-96, 43–50, 2004 年 7 月.
概要: 本稿では、二項木モデル上でのヨーロピアンアジアンオプションの価格付けに対する効率的かつ正確な乱択近似アルゴリズムを提案する。n ステップの二項木上での行使価格 X のオプションと任意の正整数 k が与えられたとき、我々のアルゴリズムは n に依存しない誤差 $O(X/k)$ の範囲の近似値を $O(kn^2)$ 時間で求める。我々のアルゴリズムは Aingworth, Motwani, および Oldham (2000) による近似アルゴリズムを乱択アルゴリズムへ修正したものであり、近似精度を理論的にも実用的にも改善している。

2. 鈴木晶子, 徳山豪:
“二部グラフの高濃度部分グラフ問題の近似アルゴリズムについて”, 情報処理学会アルゴリズム研究会, 2004-AL-97, 17–23, 2004 年 10 月.
概要: 二部グラフの高濃度部分グラフ問題と、そのデータマイニングにおける次元制限クラスタリングへの応用を論じる。過去の研究と異なり、出力される部分グラフが小さい場合にも良い理論的近似アルゴリズムを与える。

3. 全眞嬉, 金子法正, 定兼邦彦, 徳山豪:
“数値属性を用いたエキスパート付き決定木の構築”, 日本ソフトウェア科学会第 4 回データマイニングワークショップ, 2004 年 9 月.
概要: 数値属性に対する単峰近似を用いてエキスパートを作成し, データマイニングにおける C4.5 決定木をベースとしたエキスパート付き決定木を構築した。実験により性能の向上を示した。

4. 丸山謙志, 王冠超, 徳山豪:
“Web 検索解析によるクラスタリング手法の研究”, 日本ソフトウェア科学会第 4 回データマイニングワークショップ, 2004 年 9 月.
概要: Web 検索におけるクラスタリングで、 ラベルクラスタの精密化をネットワークフローを用いて行い、 従来より利用しやすいクラスタリングシステムを試作した。

5. 全眞嬉, 金子法正, 定兼邦彦, 徳山豪:
“数値データベースに対するエキスパート付き決定木の構築”, FIT2004 第 3 回情報科学技術フォーラム, F-50, 331–332, 2004 年 9 月.
概要: 予測精度の高い決定木を構築するための手法の提案を行う。決定木を構成する際の問題点として過学習がある。学習データに対する最適な決定木と未知データに対する予測精度にはトレードオフが生じるため，決定木の枝刈りが必要となる．本論文では分割時の誤差をより小さくするために，数値データを最適ピラミッド近似を利用して整理し，構築されたピラミッドをエキスパートとして用いた，個々のデータに依存する柔軟な枝刈り手法を提案した。

6. 丸山謙志, 王冠超, 徳山 豪:
“Web 検索解析によるクラスタリング手法の研究 ”, FIT2004 第 3 回情報科学技術フォーラム, D-43, 99–100, 2004 年 9 月.
概要: WEB 検索におけるクラスタリングで, ラベルクラスタの精密化をネットワークフローを用いて行い, 従来より利用しやすいクラスタリングシステムを試作した。

7. 廣川裕, 徳山豪:
“チェッカーボード丸めに関する考察と実装”, 電子情報通信学会コンピュテーション研究会, COMP2004-70, 91–96, 2005 年 1 月.
概要: 濃淡画像を白黒画像に変換するデジタルハーフトーニングへの応用を目的とし，実数値行列を二値行列に丸める時に，全ての 2×2 部分行列のディスクレパンシーの最大値を小さく方法を考える。本論文では，従来の方法に比べて簡単なアルゴリズムを考察し，実験結果を報告する。また，理論的な限界についても考察する。

8. 結城匡人, 全眞嬉, 徳山豪:
“曲線のピーク削減アルゴリズムの考察と実装”, 電子情報通信学会コンピュテーション研究会, COMP2004-71, 83–90, 2005 年 1 月.
概要: 複雑な曲線をよりシンプルな曲線で近似する問題は、幾何学データ処理の基本操作である。本論文では、出力する曲線に特徴をもたせることに重点を置く。関数 $y = f(x)$ で与えられる曲線を近似した関数表示を $y = \phi(x)$ としたとき、k 個のピーク値をもち、2 つの曲線の間での距離が最小になる $\phi(x)$ を考える。この $\phi(x)$ を得るためのアルゴリズムを L_∞ 距離および L_1 距離に関して提案する。

9. 丸山謙志, 王冠超, 徳山豪:
“Web 検索結果におけるクラスタリングアルゴリズムの研究”, 情報処理学会アルゴリズム研究会, 2004-AL-100, 17–24, 2005 年 3 月.
概要: 本論文では，ラベル付きクラスタリングによる WWW 検索エンジンの閲覧効率の向上手法を提案する。現存の WWW 検索結果クラスタリングアルゴリズムでは，クラスタラベルの重複及びラベルに無関係な文書の混在という欠点があり，ユーザにとって文書発見が困難になることがある。提案するアルゴリズムでは，従来法で得られた初期クラスタリングに対し，ネットワークフローを用いて最適なラベル配置を行う事によりラベル集合のクラスタリングを求める。上記の問題点は，ラベルクラスタリングを利用した文書の再クラスタリングにより解消される。また，アルゴリズムの実装を行い，本手法による WWW 検索エンジンのユーザビリティーの向上を示す実験結果を与える。

10. 浅野哲夫, J. Matoušek, 徳山豪:
“中立地帯を持ったボロノイ図に関する考察”, 情報処理学会アルゴリズム研究会, 2005-AL-101, 51–58, 200 年 5 月.
概要: 中立地帯のあるボロノイ図の概念の提案を与える。従来のボロノイ図が空間の分割だったのに対し、中立地帯のあるボロノイ図では、空間内の母点に対応する領域は互いに隣接せず、様々なパターンの生成を可能にする。数学的な定義を与え、その存在性と幾何学的性質、及び構築アルゴリズムに関して議論する。

11. 全眞嬉, 定兼邦彦, 徳山豪:
“データ分布の近似を用いたエキスパート付き決定木構築”, 日本応用数理学会 2005 年度年会, 2005 年 9 月.
概要: 本論文では，予測精度の高い決定木を構築するための手法の提案を行う。決定木を構成する際の問題点として過学習がある。学習データに対して最適な決定木を用いると，未知データに対する予測精度が落ちてしまうといったトレードオフが生じる。そのため決定木の枝刈りが必要となるが，どのように枝刈りを行えば最適であるかはわからない。予測精度の高い決定木を生成するためには，学習データの分割をできるだけ小さな誤差で行い，そして適切な枝刈りを行うことが必要である。従来の数値データの 1 点分割手法は，分割時にデータの誤差問題がある。また，最適サポートルールの区間を用いた分割手法では，誤差の少ない区間分割は可能であるが，適切な枝刈りは困難である。本論文では，分割時の誤差をより小さくするために，数値データを最適ピラミッド近似を利用して整理し，構築されたピラミッドをエキスパートとして用いた，個々のデータに依存する柔軟な枝刈り手法を提案する。

12. 関野剛永, 塩浦昭義, 徳山豪:
“ヨーロピアン・アジアンオプションの価格付けに関する近似的解法”, 電子情報通信学会 COMP 研究会, COMP2005-46, 13–18, 2005 年 10 月.
概要: アジアオプションの価格決定において、金利や株価のボラティリティーに変動があるモデルで理論保証のあるアルゴリズムを提案する。

13. 岡田有司，徳山 豪:
“点と直線間の 3 等分線と 2 点間の 6 等分線”, 2006 年度夏の LA シンポジウム, 2006 年 8 月.
概要: 平面上に点 p と直線 ℓ を与え, これらを次の条件を満たす曲線 C_1, C_2 で分割する.C_1 の全ての点から点 p への距離と C_2 への距離

が等しく, また C_2 の全ての点から直線 ℓ への距離と C_1 への距離が等しい。我々は C_1,C_2 が一意に存在することを示す。

14. 田代 太一，M. Halldórsson，徳山 豪:
“Wireless Ad-Hoc Network における干渉の軽減法”，2006 年度夏の LA シンポジウム，2006 年 8 月.
概要: 本研究では, 平面上に与えられた点集合を用いて Ad-Hoc ネットワークを構築する際, 各端末の電波送信半径を調整することによって, 最も干渉が起こる端末の干渉レベルを抑える手法を考える。本論分では, ϵ-ネットの理論と計算幾何のアイディアを用いることで, 干渉を $O(\sqrt{\Delta})$ に抑える手法を提案する. ここでの Δ は, 全端末を同じ電波送信半径としてネットワークを構築した場合の最大干渉レベルである。この手法は一次元モデルにおける結果を二次元に拡張したものである。さらに 4 分木を用いた手法も提案し, 最大干渉レベルの下界を考察する。

15. 小沼寛明，全眞嬉，徳山豪:
“Web データの自動抽出とデータ変換”，FIT2006, D-036, 2006 年 9 月.
概要: 複数の web ページからデータを自動的に抽出しデータ変換を行い XML にまとめる。本稿では，リストページとディテールページの構造を持った web ページ群から情報を自動抽出する Lerman らの手法の実装と改善方法を提案し実験を行った。Lerman らの手法ではタグに挟まれている文字列を extract とし，リストページとディテールページの完全一致する文字列のみを抽出し，CSP の制約式を立ててレコード分割に用いる。完全一致した文字列のみ抽出されるため，情報ロスが生じる。本稿では上記の問題を回避するための改善法を提案する。

16. T. Asano, J. Matoušek, T. Tokuyama:
“Zone Diagram, Existence, Uniqueness and Algorithmic Challenge”，電子情報通信学会コンピュテーション研究会, COMP2006-30, 39–46, 2006 年 9 月.
概要: A *zone diagram* is a new variation of the classical notion of Voronoi diagram. Given points (sites) $p_1, \ldots, p_n$ in the plane, each p_i is assigned a region R_i, but in contrast to the ordinary Voronoi diagrams, the union of the R_i has a nonempty complement, the *neutral zone*. The defining property is that each R_i consists of all $x \in \mathcal{R}^2$ that lie closer (non-strictly) to p_i than to the union of all the other R_j, $j \neq i$. Thus, the zone diagram is defined implicitly, by a “fixed-point property,” and neither its existence nor its uniqueness seem obvious. We establish existence using a general fixed-point result (a consequence of Schauder’s theorem or Kakutani’s theorem); this proof should generalize easily to related settings, say higher dimensions. Then we prove uniqueness of the zone diagram, as well as convergence of a natural iterative algorithm for computing it, by a geometric argument, which also relies on a result for the case of two sites in an earlier paper. Many challenging questions remain open.

17. 徳山豪:
“風変わりなボロノイ図と計算可能性”, 日本 OR 学会第 18 回 RAMP シンポジウム, 2006 年 10 月.
概要: 垂直二等分線とその一般化であるボロノイ図は計算幾何学ではもっともポピュラーな研究対象である。その定義を少し変更し，垂直二等分線ではなく三等分線を用いたボロノイ図であるゾーン図（zone diagram) を考察する。本公演では，このような変更により生じる，数学的な困難，及び計算可能性に対する新しい問題点について述べる。(浅野哲夫氏，Jiří Matoušek 氏との共同研究)。

18. 渋谷彰信，塩浦昭義，徳山豪:
“アメリカン・アジアンオプションの価格付けに対する計算幾何手法を用いた近似アルゴリズム”, 2006 年度冬の LA シンポジウム, 2007 年 1 月.
概要: 本論文は，金融市場で取引されている金融派生商品 (デリバティブ) の一つである，オプションの価格付け (価格の計算) という問

題を扱う。本論文では，アメリカン・アジアンオプションと呼ばれるオプションに対し，計算幾何学の手法を用いて、オプション価格の近似値を効率的に求めるアルゴリズムを提案する。

19. 田代太一，M. Halldórsson，徳山豪:
“Wireless Ad-Hoc ネットワークにおける干渉の軽減法に関する研究”, 電子情報通信学会コンピュテーション研究会, COMP2006-59, 67–74, 2007 年 3 月.
概要: 本研究では, 無線通信形態の一つである Ad-Hoc ネットワークに着目し, 平面上に与えられた点集合を用いて Ad-Hoc ネットワークを構築する際の電波干渉を抑える手法を議論する. 各端末の電波送信半径を調整することによって, 連結性を保ちながら, 最も干渉が大きい端末の干渉レベルを抑えることが可能となる. 本論分では, 計算幾何における ϵ-net のアイディアを用いることで, 干渉を $O(\sqrt{\Delta})$ に抑える手法を提案する. ここでの Δ は, 全端末を同じ電波送信半径としてネットワークを構築した場合の最大干渉レベルである. この手法は一次元モデルにおける Rickenbach らの結果を二次元に拡張したものである. さらに 4 分木を用いた手法も提案し, 最大干渉レベルの下界を考察する.

20. コルマン マティアス, 徳山豪:
“最適ハイウェイ配置問題”, 電子情報通信学会コンピュテーション研究会, COMP2007-14, 19–24, 2007 年 5 月.
概要: 平面上に 2 組の点集合が与えられたときに、2 つの集合の点の間のシティティスタンスに関する目的関数を最小化するように高速道路 h を設置することを考える. シティティスタンスとは、h 上では $v > 1$、その他の一般道路では 1 の速さを用いて測る距離である。Ahn らの結果を拡張し、我々はすでに道路の集合の一部 H が建設されているというオプションを考える。さらに、異なるオプションとして、高速道路の速さや長さを考えることができ、それぞれに異なるアルゴリズムが必要であると考えられる。我々のアルゴリズムは、低次の多項式計算時間を持ち、また、様々な最適化基準に対して統一的な構造を持つ。

21. M. Halldórsson, C. Knauer, A. Spillner, 徳山豪:
“Fixed-Parameter Tractability for Non-Crossing Spanning Trees”, 電子情報通信学会コンピュテーション研究会, COMP2007-15, 25–30, 2007 年 5 月.
概要: 平面上に辺を曲線として描画されたグラフ (トポロジカルグラフ) において，辺が交差しない全域部分木を計算する問題を考える。非交差な全域部分木を持つかどうかの判定は NP 困難問題であり，交差数を最小化する問題は近似困難である事が知られている。本論文では、以下のような自然なパラメタに対し、パラメトリック計算量 (パラメタへの依存度を考慮する計算量) を考察する。対象とするパラメタは，入力されるトポロジカルグラフにおける交差辺対の数 k，交差辺の数 μ，頂点集合の凸包の内点である頂点数 ι である。まず，シンプルなバックトラック探索における改良された戦略を与え，$O^*(1.93^k)$ 時間の計算量を与える。これは理論的には大きな改善ではないが、バックトラック探索における実装のガイドラインを与えるという価値を持つ。次に，グラフのセパレータを用いたより洗練されたアルゴリズムにより計算量を改善する。アルゴリズムの計算量はそれぞれのパラメタに対し，$O^*(2^{O(\sqrt{k})})$, $O^*(\mu^{O(\mu^{2/3})})$, $O^*(\iota^{O(\sqrt{\iota})})$ となる。また，3-SAT 問題への帰着により，この問題に対して $O^*(2^{\sqrt{k}})$ の計算量を改善するのは 3-SAT に対する広く信じられている計算量の仮定の下では困難である事を示す。

22. 渋谷彰信, 塩浦昭義, 徳山豪:
“アメリカン・アジアンオプションの価格の近似に対する計算幾何手法的アプローチ”, 情報処理学会アルゴリズム研究会, 2007-AL-112, 49–56, 2007 年 5 月.
概要: 本論文では二項モデルにおけるアメリカン・アジアンオプションの価格付けという問題に対し、精度保証付きの近似アルゴリズムを提案する。この問題に対してはこれまで

に様々な近似アルゴリズムが提案されたが、その計算困難性のため、いずれも近似精度の保証はされていなかった。アメリカン・アジアンオプションの価格の厳密値は、各時点でのペイオフの期待値を表す区分線形関数を繰り返し計算することにより求めることが可能である。提案するアルゴリズムでは、計算幾何学的な手法を用いて区分線形関数をより複雑度の小さい区分線形関数を用いて近似することにより、オプション価格の近似値を求める。

23. 全眞嬉, コルマン マティアス, M. Nöllenburg, 徳山豪:
“ディジタル星型領域とその応用”, 情報処理学会アルゴリズム研究会, 2007-AL-114, 81–88, 2007 年 9 月.
概要: 平面グリッド上の中心点 o が与えられた時, o を通るディジタル線分と o が中心点となる任意の $p \in P$ がディジタル線分 $dig(po)$ を含むディジタル星型図形 P について定義を行う。すべての $p \in P$ 上の $dig(po)$ の結合はグリッド上で木の形をし，ユークリディアン線分 po と $dig(po)$ との Hausdorff 距離は十分に小さい。高次元の場合も考慮する。最適星型図形と星型環形の抽出をイメージセグメンテーション問題に用いることによって有効なアルゴリズム設計が可能となる。さらに，山のような形をした図形の水平断面が星型領域である図形の最適近似を行うことができる。NP 困難性の結果を用いて 2 つのディジタル星型の結合を抽出することの難しいさについて述べる。

24. V. Kolmogorov, 塩浦昭義:
“凸費用ネットワークフロー問題の双対に対する効率的なアルゴリズムとそのコンピュータビジョンへの応用”, 情報処理学会アルゴリズム研究会, 2007-AL-115, 25–32, 2007 年 11 月.
概要: 凸費用ネットワークフロー問題の双対として表される整数凸最適化問題にはコンピュータビジョンへの応用が多数存在する. 本論文ではまず, この問題の最適解を求める新たな主算法を提案する. 提案する主算法では, 繰り返し最小カット問題を解くことによって主変数を更新する. 本論文での主な貢献は, この算法の反復回数に対するタイトな上界 $K \cdot T(n,m)$ を示すことである. ここで, K は主変数の動く範囲の大きさを表し, $T(n,m)$ は n 個の頂点および m 本の枝をもつグラフにおいて最小カットを求めるのに必要な時間を表す. 次に, 双対変数であるフローを利用することにより主算法を改良し, 主双対算法を提案する. 提案する主双対算法の理論的な計算時間は主算法と同じであるが, 実用的にはより良い性能をもつことが期待できる. 最後に, パノラマイメージ接合問題というコンピュータビジョンの問題に対する応用を考える. 提案する主双対算法に対していくつかの実装を考え, パノラマイメージ接合問題の例題に適用し, その実用上の性能を評価する.

学会大会等

1. 鈴木晶子，徳山豪:
“データマイニングにおけるクラスタリング”, 日本オペレーションズ・リサーチ学会アルゴリズム研究部会, 2004 年 5 月.

2. 全眞嬉，徳山豪:
“数値データに対するデータマイニング”, 日本オペレーションズ・リサーチ学会アルゴリズム研究部会, 2004 年 5 月.

3. J. Chun，K. Sadakane，T. Tokuyama:
“Construction of Expert Guided Decision Tree on Numeric Database”, The 2nd Student-Organizing Int’l Mini-Conference on Information Electronics System, 241-244, 2004 年 10 月.

4. 塩浦 昭義:
“二項モデル上でのヨーロピアン・アジアンオプションの価格付けに対する精度保証付き近似解法”, 日本オペレーションズ・リサーチ学会アルゴリズム研究部会, 2004 年 11 月.

5. A. Shioura, T. Tokuyama:
“Ultimate Implementation and Analysis of

the AMO Algorithm for Pricing European-Asian Options", Workshop on New Horizons in Computing, 2005 年 3 月.

6. 徳山豪:
"計算理論の先端的手法による実用分野開拓: データマイニングを例にして", 電子情報通信学会 2005 年総合大会, 情報システムソサイエティ企画 "計算限界の中心で解法を叫ぶ" チュートリアル講演, 2005 年 3 月.

7. Akiko Suzuki, Takeshi Tokuyama:
"Approximation Algorithms for Computing a Highly Dense Subgraph", The 4th Japanese-Hungarian Symposium on Discrete Mathematics and Its Applications (Invited talk), 2005 年 6 月.

8. Akiko Suzuki, Takeshi Tokuyama:
"Dense Subgraph Problem Revisited", Proceedings of 2005 Workshop on Randomness and Computation, 21-22, 2005 年 7 月.

9. Takeshi Tokuyama:
"Semi-Balanced Colorings of Graphs", Proceedings of 2005 Workshop on Randomness and Computation, 23-24, 2005 年 7 月.

10. Takeshi Tokuyama:
"Distance Trisector and Voronoi Diagram with Neutral Zone", NHC Workshop on Discrete Algorithms(Invited talk), 2006 年 3 月.

11. K. Murota, A. Shioura:
"Substitutes and Complements in Network Flows Viewed as Discrete Convexity", ECCO XIX – CO 2006 Joint Meeting, 2006 年 5 月.

12. 徳山豪:
"計算理論の現状と未来", 特定領域研究 NHC 全体会議, 2006 年 11 月.

13. 徳山豪:
"数学の使い方：コンピュータサイエンスでの利用", 茨城大学イブニングセミナー（公開市民セミナー), 2006 年 12 月.

14. 徳山豪:
"計算理論の現場での応用例：講演者の経験から", 情報処理学会全国大会（特別セッション基調講演), 2007 年 3 月.

15. Akiyoshi Shioura:
"Polynomial-Time Algorithms for Convex Optimization on Jump Systems", Workshop on Advances in Optimization, 2007 年 4 月.

16. 塩浦 昭義:
"ジャンプシステム上の最適化問題に対するアルゴリズム", 日本 OR 学会 計算と最適化研究部会, 2007 年 4 月.

17. 塩浦 昭義:
"離散凸解析入門：M 凸関数と L 凸関数", 組合せ論サマースクール 2007, 2007 年 9 月.

18. V. Kolmogorov, 塩浦昭義:
"最小費用テンション問題に対する算法と画像処理への応用", 日本オペレーションズリサーチ学会 2007 年秋季研究発表会, 2007 年 9 月.

文部科学省科学研究費補助金 特定領域研究 (課題番号 16092207)

A03: 論理関数表現のモデルとシンボリックアルゴリズム

本研究では、二分決定グラフ (OBDD) や二分木などによる論理関数のグラフ表現の性質や、これらを用いた種々のデータの表現法とシンボリックアルゴリズムに関する研究を行なった。OBDD 等を用いたデータや問題の効率的表現手法や、シンボリックアルゴリズムについては、種々のデータの効率的シンボリック表現とシンボリックアルゴリズムについては、主にグラフの表現を中心に研究を行ったほか、解の列挙のアルゴリズムの効率化に関する研究を行い、その有用性を示した。また、関連する一般的なグラフアルゴリズムの研究として、いくつかのグラフ族に辺の追加、削除を行なった、パラメータ化グラフに対する頂点彩色問題の計算量を明らかにした。グラフによる論理関数の表現の基本的性質他の表現との関連については、論理関数の二分木表現と主項が一対一に対応する tree-shelalble 論理関数族について、その性質や判定問題の計算量について研究を行ない、特に、判定が効率的に行なえる実用的なグラフ族を明らかにした。

研究組織

研究代表者： 武永 康彦　電気通信大学 電気通信学部
研究分担者： 湊 真一　北海道大学 大学院 情報科学研究科 (平成 19 年度)

交付決定額 (配分額)

平成 16 年度	2,100,000 円
平成 17 年度	3,000,000 円
平成 18 年度	2,700,000 円
平成 19 年度	3,400,000 円
合　計	11,200,000 円

研究成果の概要

- 学術誌
  電子情報通信学会論文誌（2006 年), Information Processing Letters（2007 年), IEICE Trans. Inf. and Syst.（2008 年）など
- 国際会議
  WG（Norway, 2006 年), IWOCA (Australia, 2007 年)

1　はじめに

分岐プログラムは有向グラフによる論理関数の表現法であり、古くから主に理論的研究が行われていた。近年になって、分岐プログラムに強い制限を加えた二分決定グラフ（OBDD：Ordered Binary Decision Diagram）と呼ばれる表現法が注目され、論理設計支援などの分野で論理関数の内部表現として多く用いられている。そのため、二分決定グラフや、それに類したいくつものモデルについて、論理関数の処理手法やその計算複雑さについて、理論・応用の両面から研究が盛んに行なわれるようになった。

二分決定グラフを用いたアルゴリズムは、基本的に、問題をすべて符号化して論理関数の形で表現し、二分決定グラフ上で論理関数処理を行うことによって問題を解くという形をとっている。このようなアルゴリズムはシンボリックアルゴリズムと呼ばれ、その実用性から主に論理設計支援の分野で利用されてきた。最近では、グラフアルゴリズムなどにおいても、アルゴリズム理論の立場から興味深いシンボリックアルゴリズムの研究が行われるようになりつつある。

本研究では、シンボリックアルゴリズムに関して、アルゴリズムの入力となる二分決定グラフやそれに類する論理関数表現による種々のデータや問題の表現方法に関する研究を行ない、その有効性を明らかにした。また、これらの研究成果をもとに、種々の問題に対するシンボリックグラフアルゴリズム、それに関連する一般的なグラフアルゴリズムの設計を行なった。

また、論理関数のグラフ表現の性質や他の表現との関連については、論理関数の主項と二分木表現が一対一に対応付けられる正論理関数のクラスである、(ordered) tree-shellable 論理関数について研究を行なった。

以下に、具体的な研究成果の概要をまとめる。

第一に、二分決定グラフやそれに類するグラフ表現を用いた種々のデータの表現法とその複雑さ、それらに基づくシンボリックアルゴリズムについては、以下の研究を行なった。

区間グラフの区間表現に基づく効率的な二分決定グラフ表現の方法を提案し、その上で隣接頂点集合を求める、頂点からある距離以内にある頂点集合を求めるなどのグラフ操作のシンボリックアルゴリズムの設計・評価を行なった。特定のグラフクラスを対象にすることにより、グラフの一般的な OBDD 表現を用いた場合より効率的な表現や操作が可能となる場合があることを示した。

二分決定グラフに類するグラフ表現によるデータの効率的表現については、二分決定グラフの非明示的表現を用いることにより、広い応用範囲を持つ多変量閾値関数を極めて小さなサイズで表現できることを示した。また、重みつきグラフの二分モーメントグラフを用いた表現方法について研究を行い、規則的な重みを持つグラフについて有効な場合があることを示した。

また、OBDD を用いた組合せ問題の解の列挙については、与えられたグラフの全ての可能なトポロジカルソートの列挙を行なうアルゴリズムの提案、実装を行なった。OBDD を用いると多数の解を比較的小さなサイズで保持できるため、全ての解を求め、記憶する処理に向いていると考えられる。特に規則的なグラフに対しては、極めて多数の解を小さなサイズの OBDD で表現できることを明らかにし、また列挙の効率を大幅に改善する手法を示した。また、ナンバーリンクと呼ばれるパズルの OBDD による解法を提案、実装し、問題の解の列挙、唯一解のチェックを可能にした。

他に、画像データの OBDD による表現法を提案し、その上での閾値処理、平行移動・回転、フィルタ処理など基本的な画像処理アルゴリズムの設計、実装を行なった。本手法により、OBDD で表されたまま全画素に対して並列に処理が実行でき、特に小さな OBDD サイズで表現できる場合には非常に効率的な演算が可能となることを示した。

また湊は、A07 班の山下准教授と連携し、量子アルゴリズムを実現する「量子論理回路」の設計自動化の基盤技術として、二分決定グラフを活用する手法の研究を進めた。具体的には、量子状態を含む量子論理関数を効率よく表現するための二分決定グラフの変種データ構造をいくつか比較検討し、その中で山下が以前に提案した「Decision Diagrams for Matrix Functions (DDMFs)」と呼ばれるデータ構造の有効性を理論的に考察した。そして、このデータ構造が、量子論理回路に特有の制約条件をうまく利用しており、それによって、単純な二分決定グラフよりも一層、データを圧縮

し、処理効率を高めているということを明らかにした。

第二に、基本的グラフ問題のアルゴリズムに関する研究としては、特定のグラフ族に辺の追加や削除を行なったパラメータ化グラフに対する、頂点彩色問題のアルゴリズムや計算量について研究を行ない、パラメータ化計算量の観点から、より広範なグラフに対する効率的な解法を示した。

比較可能グラフに k 本の辺を加えた、比較可能$+ke$ グラフの頂点彩色問題については、k が 1 の場合のみ多項式時間アルゴリズムが存在し、2 以上の場合は NP 完全となることを示した。また、比較可能グラフから k 本の辺を除いた、比較可能-ke グラフに対しては、k が 1 の場合には多項式時間アルゴリズムが存在することを示した。

Chordal グラフに辺の追加と削除を同時に行なった、chordal$+k_1$e $-$ k_2e グラフに対する頂点彩色問題のアルゴリズムを設計し、パラメータ k_1、k_2 に関して fixed parameter tractable であることを示した。これは、辺の追加と削除をはじめて同時に扱った、パラメータ付きグラフに対するアルゴリズムである。

第三に、グラフによる論理関数の表現については、非冗長な積和形論理式で与えられた正論理関数が (ordered) tree-shellable 論理関数について研究を行ない、主に、(ordered) tree-shellability の判定が効率的に実行できる実際的な論理関数のクラスを解明した。各リテラルの出現回数が制限された場合の tree-shellable 関数の性質を明らかにし、出現回数が定数の場合には多項式時間で判定可能であることを示した。また、各積項のリテラル数に制限がある場合、tree-shellability の判定は多項式時間で実行可能であるが、ordered tree-shellable 論理関数に対しては、リテラル数が最大 4 の場合に NP 完全となることを示した。

これらの他に、本特定領域研究のミニ研究集会などを通じ、組合せゲーム・パズルについても研究を行った。

以下の各章で、これらの研究成果について説明を行なう。

2 二分決定グラフ等を用いたデータの表現とシンボリックアルゴリズム

2.1 二分決定グラフ

二分決定グラフ（Ordered Binary Decision Diagram : OBDD) [2] とは論理関数を表現する有向非巡回グラフである。変数でラベル付けされた内部節点と、論理値でラベル付けされた終端節点を持ち、内部節点は 0 枝、1 枝の 2 つの出枝を持つ。変数の全順序関係が存在し、根から終端節点への全てのパスにおいて現れる変数の順序がその全順序に従う。変数に値を割り振り、根から変数の値に従って 0 枝または 1 枝を辿り、到達した終端節点のラベルが関数の値となる。OBDD から冗長、等価な節点をすべて削除したものを既約な OBDD と呼ぶ。既約な OBDD は、変数の全順序を定めると論理関数の表現が一意に定まる、多くの実用的な関数が比較的小さなサイズで表現できる、論理演算・等価性判定・充足可能性判定など、論理関数に対する基本的処理が高速に行なえる、といった特長を持つ。そのため、論理関数や各種データの表現法として広く用いられている。

2.2 区間グラフの二分決定グラフ表現

本研究では、区間グラフの区間表現に基づく OBDD 表現を提案し、その表現の上での基本的なグラフ操作のシンボリックアルゴリズムを示した。通常のグラフの OBDD 表現を用いた場合より小さなサイズでグラフを表現でき、主なグラフの操作もより効率的に実行できることを明らかにした。これにより、特定のグラフ族に対しての効率的なグラフ OBDD 表現やシンボリックアルゴリズムが可能となる一例を示した。

Definition 1 区間グラフ *[5]* とは、数直線上の区間の集合を $I = \{K_1, K_2, \cdots, K_n\}$ としたとき、

1. 各区間 $K_1, K_2, \cdots, K_n$ に対応させて頂点を $V = \{K_1, K_2, \cdots, K_n\}$ とする、
2. K_i と K_j が共通する区間をもつ時、$(K_i, K_j) \in E$ とする、

という形で表せるグラフである。区間 K_i の最小値を a、最大値を b とした時、$K_i = [a, b]$ と表す。また、区間集合 I を区間グラフの区間表現という。

区間グラフは実際的な応用範囲も広く、重要なグラフクラスのひとつとして広く研究が行なわれている。

基本的グラフ操作としては、頂点に隣接する頂点集合を求める、頂点集合に隣接する頂点集合を求める、頂点からある距離以内にある頂点集合を求める、という操作を考える。後の 2 つの操作は入力の OBDD サイズに対し、最悪で指数時間がかかることが知られている。しかし、一般的な OBDD 表現を用いた場合との比較のため、またどの程度の大きさのグラフを扱えるかを把握するため、区間グラフの頂点数を基準にした計算量の評価を行っている。

区間表現に基づく OBDD 表現

一般にグラフを OBDD で表現する場合、頂点を 2 進数で符号化し、頂点対が辺を持つ場合のみ 1 となる関数を OBDD で表す。それに対し、区間グラフではその区間表現に基づいて OBDD で表現することも可能である。本章では、区間表現に基づく OBDD の表現方法を示す。

- 区間の両端の値を s ビットの 2 進数で表し、辺を $2s$ ビットで符号化する。
- 区間の左端、右端を符号化したものをそれぞれ x、y とし、その特性関数 $\chi_I(x_1, x_2, \cdots, x_s, y_1, y_2, \cdots, y_s)$ を OBDD で表す。変数順序は $x_1, x_2, \cdots, x_s, y_1, y_2, \cdots, y_s$ の順とする。つまり、$[x, y]$ が区間集合の一要素と一致していれば 1 を返す OBDD を作成する。

区間表現において、グラフに変化を与えない範囲で区間を短くする事が出来る。したがって、区間表現では扱う区間の端点を $0 \sim 2n - 1$ の整数で表せる。このように表した場合、χ_I の OBDD の高さは $2\lceil \log 2n \rceil$、幅は最大で n になる。よってサイズについては $O(n \log n)$ となる。一般的な表現法を用いた場合のサイズは、辺の数を m とすると、$O(m \log n)$ となり、区間表現に基づいた表現の方がサイズが小さくなる。

また、区間同士に包含関係がない区間表現から出来る区間グラフである proper 区間グラフについてもその表現法を提案し、そのサイズは $O(n \log n)$ となることを示した。

区間グラフ上でのアルゴリズム

[区間に隣接する区間集合を求めるアルゴリズム]

χ_I の OBDD から区間 $[a, b]$ に隣接しない区間を取り除くことにより求める区間集合を得る。区間 $[a, b]$ に隣接しない区間 $[x, y]$ の条件とは $x < b$ または $y > a$ を満たす事である。以下にアルゴリズムを示す。

1. x, y が $x < b$ または $y > a$ をみたすときに 0 を返す OBDD を作成する。
2. 1 で得られた OBDD と χ_I の OBDD の AND をとる。
3. 2 で得られた OBDD と区間 $[a, b]$ の OBDD との排他的論理和をとり，得られた OBDD が求める区間集合を表す。

このアルゴリズムの計算量は $O(n \log^2 n)$ となる。

これをもとに、区間集合に隣接する区間集合を求める操作は区間集合に含まれる区間数を k とすると、$O(kn^2 \log^2 n)$ 時間で実行できる。

[頂点から距離 l 以内にある頂点集合を求めるアルゴリズム]

ある区間 $[a, b]$ から距離 1 以内の区間集合 I_1 に含まれる区間 $[x_i, y_i]$ のうち、端点の最小値 x_{min} と最大値 y_{max} の間にある点は必ず区間集合 I_1 に含まれる。よって、区間集合 I_1 に隣接する区間集合 I_2 を求める事は、区間 $[x_{min}, y_{max}]$ に隣接する区間集合を求める事と同じである。この操作を繰り返すことにより、距離 l 以内にある頂点集合を求めることができる。この計算量は $O(ln \log^2 n)$ となり、一般的なグラフの OBDD 表現を用いた場合より効率的に実行できる。

2.3 多変量閾値関数の非明示的 OBDD 表現

実際に用いられる論理関数の中には、その規則性により、Leveled OBDD(LOBDD) で表現した際にレベル間の接続関係が類似した形になることがある。そのような場合に、OBDD のレベル間の接続関係を OBDD を用いて非明示的に表現することにより、さらに効率的な表現が可能になる場合がある [10]。

本研究では、種々の規則性を持つデータの論理関数表現に現れる、多変量閾値関数の OBDD の非明示的表現について研究を行なった。多変量閾値関数とは、k 個の n ビット整数 $x^{(0)}, x^{(1)}, ..., x^{(k-1)} (\in \{0,1\}^n)$ を入力変数とし、

$$f(x) = \begin{cases} 0 & (\sum_{i=0}^{k-1} w_i x^{(i)} < t) \\ 1 & (\sum_{i=0}^{k-1} w_i x^{(i)} \geq t) \end{cases}$$

の形で表せる nk 変数論理関数である。非明示的 OBDD 表現を用いた場合、多変量閾値関数の重みの和を w とすると、$O(k \log w)$ のサイズで表現可能であることを示した。

OBDD の非明示的表現

それぞれのレベルの節点に対してレベルごとに独立に 2 進符号を与える。そして、レベル間の接続関係を、節点の 2 進符号 y とそこから出る枝のラベル x を入力とし、枝の先にある節点の符号 z を出力する多出力関数 $\delta(y, x) = z$ で表す。定数節点にも同様に 2 進符号を与え、節点の 2 進符号を入力とし、その節点にラベル付けされている論理値 z を出力する関数 $\lambda(y) = z$ を作る。

一般に、n 変数 m 出力論理関数を表現する LOBDD L のレベル j の節点を w_j ビットで二進符号化すると、この LOBDD を表す iLOBDD I は以下の 3 つ組として表される。ここで $\sigma_j : N_j \to B^{w_j}$ を LOBDD のレベル j の節点の符号を表す関数とする。

定義 n 変数 m 出力の LOBDD $L = (\boldsymbol{N}, \boldsymbol{e}, c, \boldsymbol{i})$ を、符号化 σ のもとで表す iLOBDD I は 3 つ組 $I = (\boldsymbol{\delta}, \lambda, \boldsymbol{s})$ である。ただし、

- $\boldsymbol{\delta} = (\delta_0, ..., \delta_{n-1})$：写像 $(\delta_j : B^{w_j} \times B \to B^{w_{j+1}})$ は j レベルの節点の枝が接続する節点の符号を表し、$\delta_j(\sigma_j(v), x) = \sigma_{j+1}(\boldsymbol{e}_j(v, x))$ を満たす。
- $\lambda : B^{w_n} \to B$ は定数節点に割り当てられている論理値を表し、$\lambda(\sigma_n(v)) = c(v)$ を満たす。
- $\boldsymbol{s} = (s_0, ..., s_{m-1})$ は初期節点を表し、$s_j = \sigma_0(i_j)$ を満たす。

多変量閾値関数の非明示的な表現法

OBDD のあるパス上で、1 枝をたどる節点に付けられた重みの和をパスの重みと定義する。初期節点から節点 v までのあるパスの重みを、そのパスに対する節点 v の中間和 (temporary sum) と呼び $S(v)$ と書く。初期節点から節点 v までの全てのパスに対する中間和の集合を $TS(v)$ と書く。

k 個の n ビット変数を入力とする多変量閾値関数で、変数順序を $x_0^{(0)}, x_0^{(1)}, ..., x_0^{(k-1)}, x_1^{(0)}, ..., x_{n-1}^{(k-1)}$ とした時について考える。

多変量閾値関数を OBDD で表すとき、各レベルに掛かる重さは、$w_0, w_1, ..., w_{k-1}, w_0 * 2, w_1 * 2, ..., w_{k-1} * 2, w_0 2^2, w_1 2^2, ..., w_{k-1} * 2^2, w_0 * 2^{(n-1)}, w_1 * 2^{(n-1)}, ..., w_{k-1} * 2^{(n-1)}$ となり、k レベル毎に $w_0, w_1, ..., w_{k-1}$ が繰り返され入力変数の桁が上がる度に重みが倍になる。

入力変数の l 桁目でラベル付けされたレベルの集合をブロック l と書く。ブロックが変わる時に各頂点の中間和、閾値、各レベルに掛かる重みを更新する。閾値を $t = t_{m-1} t_{m-2} ... t_0$ とする各レベルの節点 v を $S(v)$ を二進符号化したもので符号化する。ブロック 0 では $S_0(v) = S(v)$ とする。レベル $k-1$ でブロック 0 からブロック 1 に変わる際、$S_1(v) = \lfloor \frac{S_0(v) - t_0}{2} \rfloor$ とし、閾値を最下位ビットを除いた $t_{m-1} t_{m-2} ... t_1$ に更新し、以降のレベルに掛かる重みを全て 2 で割る。これによって、ブロック 1 の各レベルに掛かる重みは $w_0, w_1, ..., w_{k-1}$ となり、各レベルの節点 v を $S_1(v)$ を二進符号化したもので符号化することでブロック 0 と同じ関数でレベル間の接続関係を表すことが出来る。以降のブロックに関しても同様に、$S_{i+1}(v) = \lfloor \frac{S_i(v) - t_i}{2} \rfloor$ として閾値や重みも同様に更新していく。ブロッ

ク i の各節点 v に対して $S'(v) = S_i(v)$ とし、全ての節点 v を $S'(v)$ を用いて符号化することですべてのブロックのレベル間の接続関係が同じ関数で表現できる。これらの関数は、加算の形で表せることから、そのサイズが $O(k \log w)$ となることが示される。

2.4 二分モーメントグラフを用いた重み付きグラフの表現

OBDD を拡張した表現法として、Binary Moment Diagram (BMD) および multiplicative BMD (*BMD) が考案されている [3]。ここではこれらをまとめて二分モーメントグラフと呼ぶ。二分モーメントグラフは論理変数に対し整数値を返す関数を表現するためのデータ構造であり、特に乗算などの OBDD では効率的に表現できない算術演算を効率的に表現することができるという特長を持つ。

本研究では、この二分モーメントグラフを用いて重みつきのグラフを表現する方法について研究を行い、効率的な表現が可能なグラフの例を示した。

二分モーメントグラフ

二分モーメントグラフとは、$f : \{0,1\}^n \rightarrow Z$ となる関数を表す有向非巡回グラフである。基本的構造は OBDD と類似しているが、以下の点で異なる。

二分モーメントグラフは次のような関数の展開に基づいている。$f = f|_{x_i=0} + x_i(f|_{x_i=1} - f|_{x_i=0})$ この式の $f|_{x_i=0}$ を定数モーメントと呼び、$f|_{x_i=1} - f|_{x_i=0}$ を線形モーメントと呼ぶ。x_i でラベル付けされた頂点の 0 枝は定数モーメントを指し、1 枝は線形モーメントを指す。そして、定数頂点は整数でラベル付けされている。二分モーメントグラフで与えられた関数の関数値は、開始頂点から変数の値に従って、値が 0 なら 0 枝、値が 1 なら 0 枝と 1 枝両方を辿り、到達した終端頂点のラベルの値の総和で与えられる。*BMD では各枝に重みを与えることができ、その枝を通る際には、枝の重みがその枝の指す頂点の値に掛けあわされる。

辺に重みの付いたグラフの表現

辺重み付きグラフを表現する方法として、まず頂点の組 (x, y) を入力とし、その 2 頂点間に辺があればその辺の重みを返す関数 $w(x, y)$ の二分モーメントグラフによる表現を提案した。このとき、辺の有無の区別がつけられるように、$(\forall(x, y) \in E)(w(x, y) < w_{up})$ となる定数 w_{up} を定義し、2 頂点間に辺があればその重みを返し、無ければ w_{up} を返す。頂点の重みが k ビットで表現可能な場合、OBDD を用いた通常の表現のサイズは $O(m \log n + mk)$ となるのに対し、二分モーメントグラフで表したときは、辺の数にかかわらず $O(n^2)$ となる。

次に、二分モーメントグラフと OBDD を併用し、2 頂点間に辺があれば 1 を返す OBDD と、辺を符号化したものからその辺の重みを求める *BMD で表現する方法を提案した。この表現の OBDD のサイズは $O(m \log n)$ となる。二分モーメントグラフのサイズは、一般的には $O(n^2)$ となり効率よく表現できないが、w が $|x|+|y|, |x|*|y|, 2^{|x|}$ 等の算術演算を用いて表される場合は $O(\log n)$, $|x|^2$ などを用いて表される場合は $O(\log^2 n)$ となる。

二分モーメントグラフと OBDD を用いた表現が有効なグラフの例として、2 次元メッシュの各辺に、規則的な重みを付加したグラフを考える。2 次元メッシュの辺 $((x, y), (x', y'))$ の重みを $w(x, y, x', y') = |x'||y'|$ と定義した場合、辺の重みだけでなく、頂点 (x, y) を始点とする長さ k の経路の重みを表す関数も効率的に表現できることを示した。

2.5 OBDD によるトポロジカルソートの列挙

解の列挙とは、問題の条件を満たす解を全て求めることである。問題の解を 2 進符号化して表現すると、解集合は x が解の符号化となるときそのときに限り、$f(x) = 1$ となる論理関数で表現できる。この論理関数を OBDD を用いて表現することにより、多数の解が存在する場合においても、部分グラフの共有が行なわれるため、比較的小さなサイズで表現可能になると期待できる。指数的な個数の解を多項式サイズで表現できる場合も存

在する。特に解の集合は問題の条件を満たすために、共通の性質を持つ部分が存在することから、適切な符号化や変数順序の選択により、効率的なOBDD 表現が可能となる場合が多いと考えられる。また、解集合を OBDD で表現することにより、多数の解の中から、与えられた条件を満たす解を容易に抽出することができる利点がある。

本研究では、与えられた有向非巡回グラフに対し、その全てのトポロジカルソートの結果を列挙するアルゴリズムの提案、実装評価を行なった。入力のグラフはOBDD の形で与えられるものとし、論理演算などのシンボリック操作のみにより、トポロジカルソートの列挙を行なう。複数の列挙アルゴリズムを実装し、それらの比較を行なった。その結果、非常に多くの解を小さなサイズの OBDD で表現でき、特に規則的な構造を持つグラフに対してはその効果が顕著であることが明らかになった。しかし、演算結果は小さく表現できるが、計算途中の OBDD のサイズが非常に大きくなる場合がある。しかし、この問題点は各頂点の開始頂点からの距離を求めるアルゴリズム [9] の適用、到達可能性行列の利用により大幅に改善可能であることが示された。

トポロジカルソートの列挙アルゴリズム

有向非巡回グラフ $G = (V, E)$ のトポロジカルソートとは、$(v_1, v_2) \in E \Leftrightarrow \pi^{-1}(v_1) < \pi^{-1}(v_2)$ の条件を満たす頂点順序 $v_{\pi(1)}, \ldots, v_{\pi(n)}$ を求めることである。グラフ G の頂点は、$0, \ldots, |V| - 1$ の数字でラベル付けされているものとする。

グラフを OBDD で表現する際には、頂点のラベルを n ビットで 2 進符号化し、2 頂点間に辺が存在するときのみ 1 を返す関数で表現する。トポロジカルソートは $n \times |V|$ 変数の入力 x をとる論理関数 f で表す。頂点の並び順 $(x_0 \ldots x_{n-1}), (x_n \ldots x_{2n-1}), \ldots, (x_{n(|V|-1)} \ldots x_{n|V|-1})$ がトポロジカルソートになっているとき、f は 1 になる。

$S(i, v)$ はソートの i 番目に頂点 v が現れることを表す関数である。つまり、v の符号を変数 $x = x_{n\times(i-1)} \ldots x_{n\times i-1}$ としたとき、1 になる関数である。

本研究では、3 通りの方法でトポロジカルソートの列挙を行った。

削除法では、全ての頂点の組合せから条件に当てはまらないものを削除していく。

削除法＋最大距離では、それぞれの頂点の、開始頂点（終端頂点）からの最大距離を求めるアルゴリズム [9] を用い、あらかじめ不要な順列を削除することで、計算途中の OBDD のサイズを小さくする。

頂点 v の開始頂点からの最大距離 d は、頂点 v のソート順が d 以上であることを保証する。これにより条件を満たさない順列を除去する。同様に、終端頂点からの最大距離 d は、頂点 v のソート順が $|V| - 1 - d$ 以下であることを保証する。

削除法＋到達可能性では、グラフの到達可能性行列を利用する。到達可能性行列は入力の OBDD より簡単に計算することができ、その行ごと、列ごとの 1 の数の和を求めることにより、各頂点より先、あるいは後に現れなければならない頂点数が得られる。これにより、最大距離を用いた場合よりも、各頂点がソート結果の中で現れ得る位置を絞り込むことができ、より多くの条件を満たさない順列を除去することができる。

実験結果と考察

OBDD ライブラリは CMU の bddlib を利用し、Opteron265*2CPU、8GB メモリの計算機上で実験を行なった。入力のグラフとして、$k \times k$ のグリッドグラフ、これから指定した確率で辺を除去したグラフ、2 個の頂点を長さ $m + 1$ のパス l 本で結んだ l 行 m 列グラフをなどを用いた。

図 1に、グリッドグラフにおける、全ての解を保持する OBDD の頂点数を示す。

表 1: 解の個数と OBDD サイズ

頂点数	解の個数	OBDD サイズ
4	2	13
9	42	87
16	24024	431
25	–	1925
36	–	9591
49	–	41911

解の個数の増加に対し、非常に小さいサイズのOBDDで全ての解を保持できることがわかる。

しかし、計算途中において、OBDDサイズが非常に大きくなる場合がある。表2は5×5グリッドグラフでの実行結果である。

削除法、削除法+最大距離、削除法＋到達可能性の順にOBDDの最大サイズが小さくなり、実行時間も速くなる。削除法＋到達可能性では、他の方法で計算できない6×6、7×7について列挙を実行できた。

表 2: 5×5 グリッドグラフでの実行結果

アルゴリズム	実行時間 (s)	最大サイズ
削除法	1587.12	56993208
削除法+最大距離	212.70	16211270
削除法+到達可能性	170.90	965026

2.6 OBDDを用いたナンバーリンクの解法

本研究では、ナンバーリンクと呼ばれるパズルをOBDDを用いて解く手法、および与えられた問題の正当性をチェックする手法の提案・実装を行なった。問題の解の条件を論理式で表現することにより、全ての解を生成することができるが、問題の性質より、局所的な条件のみで解の制約条件を表現するのが困難である。しかし、解の性質を検討することにより、正当性のチェックを実現できることを示した。

ナンバーリンク

ナンバーリンクは以下のように定義される。

入力: 盤面は大きさ$m \times n$の正方格子であり、i行j列目のマスを(i, j)で表す。盤面上には1からsの数字がそれぞれ2個ずつのマスに与えられており、他のマスは空白である。入力において数字の与えられたマスを基点という。

問題: 以下のような条件を満たすように、基点以外のマスに線を記入する。

- 基点以外の全てのマスに線が記入される。
- 1個のマスには、上下左右のマスのうちの2個を結ぶ直線または折れ線が一本だけ記入される。
- 同じ数字の基点のマス同士が線で結ばれる。
- 全ての線が基点を結ぶために使われる。

この条件を満たす解と正当解と呼ぶ。また、上記の条件のうち最初の条件以外の条件をすべて満たすものを短絡解という。唯一の正当解を持ち、短絡解を持たない問題を正当な問題と呼ぶ。

OBDDによるナンバーリンクの解法

数iの基点同士を結ぶ線が通るマスに数iを割り当てることにより解を表現する。各マスに割り当てられる数は$\lceil \log(s+1) \rceil$ビットの2進数で符号化する。短絡解において、マスに数字を割り当てないとき、そのマスには0を割り当てる。

論理変数$x_{i,j,k}$を以下のように定義する。

$$x_{i,j,k} = \begin{cases} 1 & (\text{マス } (i,j) \text{ の数 } k \text{ ビット目が } 1) \\ 0 & (\text{マス } (i,j) \text{ の数 } k \text{ ビット目が } 0) \end{cases}$$

$$(0 \le i \le m-1, 0 \le j \le n-1, 0 \le k \le r-1)$$

この変数を用いて、ナンバーリンクの解となるための制約条件を論理関数の形で記述する。たとえば、基点(i, j, t)における接続条件は以下のように表される。

$C_1(i,j,t) = (S_1(i,j,t)(\wedge\overline{S_2(i,j,t)} \wedge \overline{S_3(i,j,t)} \wedge \overline{S_4(i,j,t)}) \vee (\overline{S_1(i,j,t)} \wedge S_2(i,j,t) \wedge \overline{S_3(i,j,t)} \wedge \overline{S_4(i,j,t)}) \vee (\overline{S_1(i,j,t)} \wedge \overline{S_2(i,j,t)} \wedge S_3(i,j,t) \wedge \overline{S_4(i,j,t)}) \vee (\overline{S_1(i,j,t)} \wedge \overline{S_2(i,j,t)} \wedge \overline{S_3(i,j,t)} \wedge S_4(i,j,t))$

ここで、関数S_1, S_2, S_3, S_4はそれぞれ基点であるマス(i, j)に対して左、上、右、下に対する接続条件を表す論理関数であり、$S_1(i,j,t) = (i \neq 0) \wedge SET(i-1,j,t)$, $SET(i,j,a) = \bigwedge_{0 \le k \le r-1}((a_k = 0) \wedge \overline{x_{i,j,k}}) \vee (a_k = 1) \wedge x_{i,j,k}))$ などと表される。

解の列挙と問題の正当性チェック

各マスでの接続条件全ての論理積をとることにより、基点以外の全てのマスに数が記入され、各マスでの接続条件を満たしたものが全て得られる。しかし、これらが全て問題の正当な解となるとは限らない。なぜなら、解の最後の条件がチェックされていないためである。しかしこのような場合には、必ず短絡解が発生する。したがって、入

力が正当な問題であることは、上記の条件を満たす割り当てが唯一存在し、かつ短絡解を許しても条件を満たす割り当てが唯一存在することによりチェックすることができる。

2.7 OBDDを用いた画像処理アルゴリズム

本研究では、画像データをOBDDで表現することを考える。画像の性質によってはかなり効率的に画像データを表現できることが期待できる。さらに、OBDD表現を用いる事により、OBDD上の操作によって種々の画像の処理をより効率的に実行できる可能性がある。特に画像処理では、画像の全画素に対して同じ処理を施すことが非常に多いため、OBDD上の演算により全画素に対して同時に操作を実行できる点でOBDD表現の利用が有効であると考えられる。しかも、これらの処理はOBDD表現により圧縮されたデータのまま実行することができる。

本研究では、多値画像データのOBDDによる表現方法として、1個のOBDDで表現する方法、複数のOBDDを用いてSBDD(Shared OBDD)で表現する方法の2種類を考えた。そして、この2種類の表現方法の上での、閾値処理、平行移動・回転、エッジ抽出などのフィルタ処理等の画像処理アルゴリズムを提案、実装した。

OBDDを用いた画像の表現

画像をOBDDで表現する際には、x座標、y座標、画素の濃度zをそれぞれ$x = x_{n-1}x_{n-2}\cdots x_0$などと2進表現により符号化する。また論理関数$f(x_{n-1}x_{n-2}\cdots x_0, y_{n-1}y_{n-2}\cdots y_0, z_{n-1}z_{n-2}\cdots z_0)$を$f(\mathrm{x,y,z})$のように表す。OBDDにより画像を表現する方法として、x座標、y座標を表す節点の後にzを表す節点を持つ1つのOBDDで表す方法と、複数のOBDDでzの各ビットをそれぞれ個別に表し、その複数のOBDDをSBDDで保持する方法の2通りが考えられる。なお以下では、簡単のため画像のサイズは$2^n \times 2^n$とする。

・1個のOBDDによる表現

この方法ではx座標、y座標、画素の濃度zを表す$x_0, \cdots, x_{n-1}$, $y_0, \cdots, y_{n-1}$, $z_0, \cdots, z_{m-1}$を変数として持ち、1つのOBDDで画像を表現する。座標(x, y)の画素の濃度がzのとき、$f(\mathrm{x,y,z}) = 1$となるOBDDにより画像を表現する。

・SBDDによる表現

この方法では、OBDDの変数はx, yのみで各OBDDがzの各ビットを表し、それら全てのOBDDをSBDDとして保持する。座標(x, y)の画素の濃度zのz_iが1のとき、$f_i(\mathrm{x,y}) = 1$ $(i = 0, \cdots, m-1)$となるOBDDにより画像を表現する。

OBDDを用いた画像処理

OBDD上での画像処理を行うシンボリックアルゴリズムのうち、平行移動、およびフィルタ処理について説明する。

平行移動

平行移動は、x軸方向またはy軸方向に距離dだけずらす操作である。OBDD表現に対するアルゴリズムを以下に示す。

1. $g(\mathrm{x,u}) = 1 \Leftrightarrow u - x = d$となるような$g$を作る。
2. $f'(\mathrm{u,y,z}) = \exists \mathrm{x}[f(\mathrm{x,y,z}) \cdot g(\mathrm{x,u})]$

gは多変量閾値関数のクラスに属している[9]。そのためgはx, uの変数順序を$x_0, u_0, x_1, u_1, \cdots, x_{n-1}, u_{n-1}$のように入れ子にすることによって定数幅で表すことができ、元画像のOBDDサイズの線形時間で処理を行なえる。

SBDD表現によるフィルタ処理多値のままフィルタをかける処理は、OBDD表現での効率的実現は困難であると考えられる。しかしSBDD表現では、これらを含めた様々なフィルタ処理を行うことが出来る。SBDD表現によるフィルタ処理のアルゴリズムを以下に示す。

1. 周囲8方向への平行移動処理を行う。
2. 1.で作った8個の画像のSBDDと元の画像のSBDDに対し、フィルタの対応する要素の絶対値との乗算を行う。
3. フィルタの対応する要素の符号がプラスのもの同士を加算したものを$g(\mathrm{x,y})$、マイナスのもの同士を加算したものを$h(\mathrm{x,y})$とする。

4. $g(\mathrm{x},\mathrm{y}) - h(\mathrm{x},\mathrm{y})$ の減算を行う。結果が濃度の最大値を超えた場合は最大値に、マイナスになった場合は 0 にする。

我々の提案したシンボリック画像処理アルゴリズムは、最悪の場合入力 OBDD サイズに対して指数時間かかるものもあるが、画像の画素数に対しては最悪でも多項式時間となる。

3 パラメータ化グラフに対する頂点彩色問題

3.1 概論

一般の入力に対して NP 完全である問題への取り組みの 1 つとして, 問題の入力の一部をパラメータとして特定の値 k に固定して, 特定の k の値の元では容易に解けるかどうかを考える方法がある.

本研究では、グラフの頂点彩色問題を扱う。頂点彩色問題は次で定義される.

Input: グラフ $G = (V, E)$, 正の整数 $t \leq |V|$

Question: G は t 彩色可能か. つまり, すべての G の辺 (u, v) に対して, $f(u) \neq f(v)$ を満たす関数 $f : V \to \{1, 2, \ldots, t\}$ が存在するか.

t 彩色可能である t のうち最小のものを彩色数 $\chi(G)$ という. 頂点彩色問題は彩色数を 3 に固定したところで NP 完全である.

しかし、特定の性質を持つグラフだけを対象とした場合には、頂点彩色問題が多項式時間で解けることがある。したがって、このようなグラフ族に含まれるグラフに何らかの意味で「近い」グラフに対しても効率的なアルゴリズムが存在する可能性がある。このように、あるグラフ族に近いグラフとして、ここでは以下のようなグラフ族を考える。

$\mathcal{F}$ をあるグラフ族とすると, $\mathcal{F}$ グラフに辺を k 本加えたグラフからなるグラフ族を $\mathcal{F}$+ke グラフと呼ぶ. また同様に, $\mathcal{F}$ グラフから辺を k 本取り除いたグラフからなるグラフ族を $\mathcal{F}$-ke グラフと呼ぶ. $\mathcal{F} + ke$ グラフ $G = (V, E)$ のモジュレータとは, $G - E_k \in \mathcal{F}, |E_k| \leq k$ となるような $E_K (\subset E)$ のことである. つまり, $\mathcal{F}$+ke グラフからモジュレータに含まれる辺を取り除くと $\mathcal{F}$ グラフが得られる.

3.2 比較可能+ke, -ke グラフの彩色問題

比較可能グラフ [5] に対しては、頂点彩色問題を多項式時間で解くことができる [6]. グラフ G が比較可能であるとは, G の各無向辺を適当な方向の有向辺にしたときに, そのグラフが推移的グラフになるものが存在するときをいう. ただし, 推移的グラフとは辺 (u, v), (v, w) が存在するならば辺 (u, w) も存在するような有向グラフのことである.

本研究では, k を入力のパラメータとし, k の値を固定したときに比較可能+ke グラフおよび比較可能-ke グラフの頂点彩色問題の計算量がどのように変化するかを調べる. 比較可能+ke グラフに対しては、$k = 1$ のとき $O(\gamma|E|)$ 時間で解くことができ、$k \geq 2$ の場合は NP 完全となることを示した。また、比較可能-ke グラフに対しては、$k = 1$ のとき多項式時間で解くことができることを示した。

比較可能+1e グラフの彩色問題

グラフ G のクリーク数 $\omega(G)$ とは, G の完全部分グラフのうち最大のものの位数である. 比較可能グラフは理想グラフであり、 クリーク数と彩色数が等しい. 比較可能グラフのクリーク数および彩色数は多項式時間で求められる. 推移的グラフ $G = (V, E)$ の頂点 u, v について, u が v へ隣接しているとき $f(u) < f(v)$ を満たすような関数 $f : V \to \{1, 2, \ldots, \omega(G)\}$ を, G のレベル付けといい, このとき $f(u)$ を u のレベルという. 同じレベルの頂点を同じ色で彩色すれば、$\omega(G)$ 色での頂点彩色が得られる。

比較可能+1e グラフを G, G のモジュレータを $E_1((a, b) \in E_1)$ とすると, $G' = G - E_1$ は比較可能グラフとなる. なお, ここではモジュレータはあらかじめ与えられているものとする. もし仮にモジュレータが未知であったとしても, 比較可能グラフの判定は多項式時間でできるので, モジュレータを探すことは多項式時間でできる.

比較可能+1e グラフ G の彩色数 $\chi(G)$ は少なくとも $\chi(G')$, また高々 $\chi(G')+1$ である. G' における彩色のうち頂点 a,b が互いに異なる色で彩色されるものはそのまま G の彩色となり, 頂点 a,b がともに同じ色で彩色されるものはそのままでは G の彩色として不適切である. そこで, G における $\chi(G)$ 彩色において, 頂点 a,b を互いに異なる色で彩色できる条件として, 以下の補題を示す.

Lemma 1 *レベル l の 2つの頂点 a,b を異なる色で彩色できる必要十分条件は, レベル $l-1$ と l の頂点から誘導される部分グラフ, およびレベル l と $l+1$ の頂点から誘導される部分グラフの少なくともどちらか一方において a,b が連結でないことである.*

この補題の条件を満たすかどうかを判定するアルゴリズムの時間計算量は、比較可能グラフ G' の推移的グラフを求める時間に依存し、$O(\gamma|E|)$ となる。ただし, γ は頂点の次数の最大値である.

Theorem 1 *比較可能+1e グラフの彩色問題は $O(\gamma|E|)$ 時間で解くことができる. ただし γ は頂点の次数の最大値である.*

比較可能+2e グラフの彩色問題

比較可能+2e グラフを G, G のモジュレータを E_2 としたとき, G の彩色数 $\chi(G)$ は少なくとも $\chi(G-E_2)$ であり, また高々 $\chi(G-E_2)+2$ である.

比較可能+2e グラフ G の頂点彩色問題の 2 つのパターンとして $\chi(G-E_2)$ 色彩色問題と $\chi(G-E_2)+1$ 色彩色問題の NP 完全性を示す. 以下の 2 つの定理は、ともに 3SAT からの帰着により証明される。

Theorem 2 *比較可能+2e グラフ G の $\chi(G-E_2)$ 色彩色問題は NP 完全である.*

Theorem 3 *比較可能+2e グラフ G の $\chi(G-E_2)+1$ 色彩色問題は NP 完全である.*

また, 定理 2および定理 3のどちらからでも, 次の系が導かれる.

Corollary 1 *比較可能+ke グラフの頂点彩色問題は $k \geq 2$ のとき NP 完全である.*

比較可能-1e グラフの彩色問題

以下では、与えられた比較可能-1e グラフを G、G にモジュレータの辺 (a,b) を追加して得られる比較可能グラフを G_c とする。G においては、モジュレータの両端の頂点を同じ色で彩色することにより、G_c より少ない色数で彩色できる可能性がある。

モジュレータの有無により最大クリークサイズが異なる場合、まず彩色数を変えないグラフの変形を行なう。ハッセ図表現において頂点 w,x,y,z において辺 $(w,x),(y,x),(w,z)$ を持ち、w と x が同じ $(\omega(G_c)-1)$-クリークに含まれるとき、ハッセ図において辺 (y,z) を付加する。この操作によって彩色数が増えないことから、上の条件を満たす頂点の組がなくなるまでこの操作を繰り返し、得られたグラフを G_+ とする。

ハッセ図においてモジュレータ (a,b) が a から b への有向辺となるとき、V_a, V_b をそれぞれ、G_+ のハッセ図において頂点 a からのパスを持つ頂点の集合、頂点 b へのパスを持つ頂点の集合とする。このとき、以下の補題が成り立つ。

Lemma 2 *モジュレータの両端の頂点を同時に含まず、また全ての頂点が $V_a \cup V_b$ に含まれるような $(\omega(G_c)-1)$-クリークが存在するとき、G_+ は $(\omega(G_c)-1)$ 色で彩色可能である。*

この補題の条件をチェックすることにより多項式時間で頂点彩色問題を解くことができる。

3.3 $\mathcal{F}+k_1\mathrm{e}-k_2\mathrm{e}$ グラフの頂点彩色問題

本研究では、$\mathcal{F}$ グラフに辺の付加、削除の両方を行って得られるグラフを扱う。頂点彩色問題を多項式時間で解くことが出来るグラフ族 $\mathcal{F}$ に含まれるグラフに高々k_1 本の辺を追加し、高々k_2 本の辺を削除して出来る $\mathcal{F}+k_1\mathrm{e}-k_2\mathrm{e}$ グラフを考える。追加された高々k_1 本の辺の集合と、削除された高々k_2 本の辺の集合をそれぞれプラスモジュレータ、マイナスモジュレータと呼ぶ。またプラスモジュレータの端点を正特殊頂点、マイナスモジュレータの端点を負特殊頂点と呼び、合わせて特殊頂点と呼ぶ。

$\mathcal{F}+k\mathrm{e}$、$\mathcal{F}-k\mathrm{e}$ グラフの頂点彩色問題に関して Cai[4] は、グラフ族 $\mathcal{F}$ が非隣接頂点の同一化、辺

の縮退に関して閉じているという性質を持つ場合 fixed parameter tractable となることを示している。非隣接頂点の同一化とは、非隣接である 2 頂点を、それらに接続している辺全てと接続する新しい 1 つの頂点に置き換える操作であり、辺の縮退とは、隣接している 2 頂点間の辺を削除し、それら 2 頂点に対して上の同一化を行う操作である。

本研究では、第一に、$\mathcal{F}$ が辺の縮退・非隣接頂点の同一化という 2 つの操作に対して閉じているなら、そのグラフの頂点彩色問題はパラメータ k_1、k_2 に関して fixed parameter tractable、すなわち $f(k_1, k_2)p(n)$ 時間で解けることを示した。

第二に、辺の縮退に関しては閉じているが非隣接頂点の同一化に関しては閉じていないグラフとして、chordal グラフを取り上げ、chordal$+k_1$e$-k_2$e グラフに対する頂点彩色問題を扱った。本研究では chordal$+k_1$e$-k_2$e グラフに対する頂点彩色問題のアルゴリズムを設計し、fixed parameter tractable となることを示した。以下ではこのアルゴリズムを説明する。

Chordal$+k_1$e$-k_2$e グラフの頂点彩色問題

本研究のアルゴリズムは、グラフの tree decomposition を用いた、Marx の chordal$+k$e グラフの頂点彩色問題に対するアルゴリズム [7] をもとに、マイナスモジュレータに対する処理を可能にしたものである。

グラフ G に対し、以下を満たす木 T を G の tree decomposition と呼ぶ [5]。「木 $T(U, F)$ と、各 $v \in V$ に対するその部分木 $T_v \subseteq V$ が存在し、$u, v \in V$ がグラフ $G(V, E)$ 内で隣接している必要十分条件は、$T_u \cap T_v \neq \phi$ である。」G が chordal グラフの場合、G の tree decomposition は多項式時間で求められる。

ノード $x \in U$ に対して、V_x により x または x の子孫を含む部分木を構成する頂点の集合を表す。また、V_x によって誘導される G の部分グラフを $G_x = G[V_x]$ で表す。T のノード $x \in U$ に対して、K_x で x に含まれる頂点集合を表す。

Nice tree decomposition とは、tree decomposition のうち、さらに以下の条件を満たすものを言う。

・全てのノード $x \in U$ が高々2 つの子を持つ。
・もし $x \in U$ が 2 つの子 y、$z \in U$ を持つならば、$K_x = K_y = K_z$ である (x は join ノード)。
・もし $x \in U$ が 1 つの子 $y \in U$ しか持たないならば、ある $v \in V$ に対して $K_x = K_y \cup \{v\}$(introduce ノード) か、$K_x = K_y \backslash \{v\}$(forget ノード) のどちらかである。
・もし $x \in U$ が子を持たないならば、K_x はちょうど 1 つの頂点を含む (leaf ノード)。

[7] のアルゴリズムの概要は以下の通りである。色集合 C による彩色が可能かどうかを判定する。G_x にモジュレータを加えて出来るグラフを H_x と表し、$|C| - |K_x|$ 個の頂点 $u_1, u_2, \ldots, u_{|C|-|K_x|}$ からなるクリークを H_x に加え、K_x の各頂点と繋げて出来るグラフを H_x^* とし、その時出来る $|C|$ クリークは K_x^* で表す。アルゴリズム中では、部分グラフ H_x^* の彩色は「同じ色で彩色する頂点の組」を要素とする集合を S_x で表す。Set system $\mathcal{S}_x$ とは、部分グラフ H_x^* の彩色の方法 S_x の集合となる。Set system の各集合の要素数 n が指数的に増加することを抑えるために、representative system を用いる。木 T の葉ノードから順に、ノードの種類ごとの操作を行なうことにより、H_x^* に対する set system を求める。根ノードに達した時に set system が空でなければ、グラフ G 全体に対する C 彩色が存在する。モジュレータの両端が H_x^* に含まれた時点で、両端が同じ色になる彩色は排除される。

本研究では、マイナスモジュレータに対する処理を行なうため、次の概念を導入した。マイナスモジュレータの端点のうち後で introduce されるものを後入負特殊頂点、先に forget されるものを先忘負特殊頂点と呼ぶ。Chordal$+k$e グラフの頂点彩色問題アルゴリズムとの大きな違いは、後入負特殊頂点によって K_x 内にマイナスモジュレータが現れ、その両端が同色の彩色パターンを考えることが必要になり、先忘負特殊頂点によって K_x 内のペアが無くなったマイナスモジュレータの組を別の同色の頂点で置き換える操作が必要になることである。

K_x 内で構成されているマイナスモジュレータの両端が同色になっている場合、$|K_x|'$ を K_x で使用している色数とし、頂点 $u_1, u_2, \ldots, u_{|C|-|K_x|'}$ からなるクリークを付け加える。

以上の方針のアルゴリズムにより、次の定理が

得られる。

Theorem 4 *Chordal*$+k_1\mathrm{e}-k_2\mathrm{e}$ グラフの頂点彩色問題は、パラメータ k_1、k_2 に関して *fixed parameter tractable* である。

4 Tree-Shellable 論理関数の性質と判定複雑さ

4.1 概論

論理関数の特性を理解し、明らかにすることは、計算機科学の様々な分野において重要なことである。積和形論理式 (Disjunctive Normal Form Boolean Formula：DNF) は、論理関数の代表的な表現法の 1 つである。それぞれの積項は論理関数の内項に相当し、特に主項は、論理関数の最小形を表すなど、非常に重要な概念である。特に、正論理関数は非冗長な積和形正論理式 (Positive DNF：PDNF) により一意に表現される。一方、二分決定木も論理関数のグラフによる表現法として広く用いられている。

Tree-shellable 論理関数 [8] とは、正論理関数において関数の主項とその関数を表す二分決定木の 1 ノードへたどるパスの数が等しい論理関数である。また、Ordered tree-shellable 論理関数は tree-shellable 論理関数の一種であり、関数を表す二分決定木が順序付き二分決定木となっているものである。論理関数 f が (oedered) tree-shellable であれば、f の双対関数 f^d を求めるのは容易であり、f の DNF 表現から f^d の非冗長な DNF 表現を多項式時間で求めることが可能である。また、各変数が 1 となる確率が与えられたとき、$f=1$ となる確率を求める問題はある種のシステムの信頼性計算に対応しており、f が shellable であれば、この問題も容易に解くことができる。

Tree-shellable 論理関数の判定複雑さを明らかにし、また効率的に tree-shellability の判定が行える論理関数の条件を示すことが重要である。

本研究では、各変数の出現回数に制限を加えた場合と積項の長さに制限を加えた場合の、tree-shellable 論理関数の性質や判定複雑さについて考察し、効率的な tree-shellability の判定が行なえる実際的な論理関数のクラスを明らかにした。第一に、積和形論理関数で表した場合の各リテラルの出現回数と関数の tree-shellability の関係について考察した。同一リテラルの出現回数が高々i 回である tree-shellable 論理関数 f が tree-shellable であるためには、積項数が高々$2i-1$ でなければならないことを示した。また、f が tree-shellable であることを示す二分木の分岐の深さについても明らかにした。この結果、同一リテラルの出現回数が定数回で抑えられる場合、tree-shellability の判定が多項式時間で行なえることを示した。第二に、積項の長さが制限された場合の (ordered) tree-shellable 論理関数の判定複雑さについて、tree-shellable 関数の判定は k が定数の場合は多項式時間で実行できるが、[1] の結果を改良し、積項の長さが高々4 の場合 ordered tree-shellable 論理関数判定問題が NP 完全であることを示した。以下では、前者について説明する。

他に、quadratic 関数にリテラル数の多い積項を少数加えた場合の判定複雑さや、(oedered) tree-shellable 関数の定義の正論理関数に限らない一般の論理関数への拡張についても研究を行なった。

4.2 Tree-Shellable 論理関数

二分決定木は論理関数を表す、ラベル付けされた木である。二分決定木の根ノードから値が 1 である葉ノードまでのパスを 1 パスと呼ぶ。特に、変数の全順序 π が存在し、二分木の全てのパスにおいて、現れる変数の順序が π に矛盾なく従っている場合、その二分木を順序付き二分決定木という。π を順序付き二分決定木の変数順序と呼ぶ。

Definition 2 [8] f は *PDNF* $f=\bigvee_{k=1}^{m}\bigwedge_{i\in I_k} x_i$ で表されているとする。f を表す *(*順序付き*)* 二分決定木で m 個の *1* パスを持つものが存在するとき、f は *(ordered) tree-shellable* であるという。

Theorem 5 [8] $f=\bigvee_{k=1}^{m}\bigwedge_{i\in I_k} x_i$ が *tree-shellable* であるとき、以下を満たす二分決定木 T を構成できる。

1. T は $P_1,\cdots,P_m$ の m 個の *1* パスを持つ。

2. 各々の $P_k(1\le k\le m)$ は $i\in I_k \leftrightarrow x_i\in P_k$ を満たす項 I_k に対応する。

4.3 リテラルの出現回数を制限した関数の Tree-Shellability

非冗長な積和形論理式において、各リテラルの出現回数に制限を加えた tree-shellable 論理関数の判定複雑さについて考える。中心となる結果として、次の定理が示すように、リテラルの出現回数に制限がある場合、tree-shellable となる関数の積項数にも上限があることを示した。

Theorem 6 各リテラルの出現回数が高々i回の *PDNF* で表される論理関数 f で、項を i^2+1 個以上持つものは *tree-shellable* ではない。

リテラルの出現回数を高々i回に制限すると、f が tree-shellable であることを示す f の二分決定木は根のラベル x_s を含む 1 パスを高々i 個しか持たない。また x_s を含む項 1 つに対して定理 5の条件を満たす x_s を含まない項は高々$i-1$ 個であることがわかる。よってこの場合、tree-shellable 論理関数の積項数は高々i^2 個となる。

この結果をもとに、tree-shellable となる二分木の根に選べる変数の条件や、i 回以上出現するリテラルが少数存在するときの tree-shellable 論理関数の積項数の上限についても明らかにした。

また、実際に各リテラルの出現回数が高々i回かつ積項数が i^2 の PDNF で表される tree-shellable 論理関数が無数に存在することも明らかにした。つまり、i^2 は最適な値である。具体的には、以下にそのような関数の例を示す。A, B をそれぞれ変数の積とする。ただし A、B は共通の変数を含まず、また変数 $x_1, \cdots, x_{2i-2}$ も含まないものとする。$f = \sum_{j=2}^{i}(x_1x_j + x_jB + \sum_{k=i+1}^{2i-2} x_jx_k) + \sum_{k=i+1}^{2i-2} Ax_k + x_1A + AB$ とする（ただし $i > 1$）。

Tree-shellability の判定は、一般には、与えられた論理関数のすべての可能な木を調べることになる。その際、判定アルゴリズムの計算時間は木のどの深さまでパスの分岐があるかに依存する。i^2 個の項を持てば、分岐の深さは最大 i^2-1 となり得るが、関数が tree-shellable である場合には、このように大きな分岐の深さを持たないことがわかる。二分木の根ノードから、根ノードを含めて 0 枝のみで辿れるノードの列を street とすると、以下の結果が成り立つことを示した。

Theorem 7 各リテラルの出現回数が高々i 回の *PDNF* で表される *tree-shellable* 論理関数を f とすると、f が *tree-shellable* であることを示す二分決定木 T の *street* のノード数は $2i-1$ 以下である。

長さ $2i-1$ の street を持つ関数も存在するため、この値も最適であることがわかる。

この結果を判定アルゴリズムに適用すると、関数が tree-shellable 論理関数であるかどうかは $O(n^{4i-1}i^5)$、また関数が ordered tree-shellable 論理関数であるかどうかは $O(n^{3i}i^5)$ で判定可能であることがわかる。したがって、各リテラルの出現回数が定数であれば、多項式時間で (ordered) tree-shellable であるかどうかが判定可能となる。特に、多項式の次数も i に関して線形となり、効率的に判定が可能であることがわかる。

参考文献

[1] E. Boros, Y. Crama, O. Ekin, P. L. Hammer, T. Ibaraki and A. Kogan, Boolean Normal Forms, Shellability and Reliability Computations, SIAM J. Discrete Mathematics, vol.13, no.2, 212-226, 2000.

[2] R.E. Bryant, “Graph-Based Algorithms for Boolean Function Manipulation”, IEEE Trans. Comput. Vol.C-35, No.8 (1986).

[3] R.E. Bryant and Y.A. Chen, “Verification of arithmetic circuits with binary moment diagram” In Proc. 32nd ACM/IEEE Design Automation Conference, 535-541, 1995.

[4] Leizhen Cai, “Parameterized Complexity of Vertex Colouring”, Discrete Applied Mathematics, 127, No.3, 415-429, 2003.

[5] M. C. Golumbic, Algorithmic Graph Theory and Perfect Graphs 2nd Ed., Elsevier, Amsterdam, 2004.

[6] Grotschel, M., Lovasz, L. and Schrijver, A., “Polynomial Algorithms for Perfect Graphs.”, Annals of Discrete Mathematics, 21, 325-356, 1984.

[7] D. Marx,“Parameterized coloring problems on chordal graphs”, Theoret. Comput. Sci., 351, 3, 407-424, 2006.

[8] Y. Takenaga, K. Nakajima and S. Yajima, “Tree-Shellability of Boolean Functions”, Theoret. Comput. Sci. 262, 633-647, 2001.

[9] P. Woelfel, “Symbolic Topological Sorting with OBDDs”, Mathematical Foundations of Computer Science 2003, LNCS 2747, 671-680, 2003.

[10] 山内 仁, 石浦 菜岐佐, 高橋 浩光, “二分決定グラフの非明示的表現法とその操作法”, Technical Report of IEICE. CAS94-126,VLD94-142,ICD94-250, 1995.

研究業績一覧

学術論文

1. 松金輝久、武永康彦:
“組合せ最適化問題としてのぷよぷよの連鎖数判定問題”, 電子情報通信学会論文誌, Vol.J89-D, No.3, 405–413, 2006 年 3 月.
概要: ゲームやパズルの計算量や解法に関する研究は古くから行われている。特に最近ではテトリスのようなゲームが注目を集めている。本論文では、対戦型ゲームとして広く知られるぷよぷよを、入力として初期盤面と落下してくるピース列が与えられ、ピースを落下させることにより特定の目的を達成するパズルゲームとして定式化し、その連鎖数判定問題を考える。連鎖はぷよぷよにおける特徴的な性質であり、最大の連鎖を発生させる連鎖数判定問題が NP 完全であることを証明する。

2. Yasuhiko Takenaga and Kenichi Higashide:
“Vertex Coloring of Comparability+ke and -ke Graphs”, 32nd International Workshop on Graph-Theoretic Concepts in Computer Science, Lecture Note in Computer Science 4271, pp.102-112, June 2006.
概要: $\mathcal{F}+k$e and $\mathcal{F}-k$e graphs are classes of graphs obtained by adding or deleting at most k edges from a graph in $\mathcal{F}$. In this paper, we consider vertex coloring of comparability+ke and comparability$-k$e graphs. As vertex coloring can be solved in polynomial time for comparability graphs, it may also be easy for such graph classes. For comparability+ke graphs, we show that vertex coloring can be solved in polynomial time for $k = 1$, and is NP-complete for $k \geq 2$. For comparability-ke graphs, we show that vertex coloring can be solved in polynomial time for $k = 1$.

3. Yasuhiko Takenaga and Toby Walsh:
“TETRAVEX is NP-complete”, Information Processing Letters, Vol.99, No.5, pp.171-174, Sept. 2006.
概要: Tetravex is a widely played one person game in which you are given n^2 unit tiles, each edge of which is labelled with a number. The objective is to place each tile within a n by n square such that all neighbouring edges are labelled with an identical number. In this paper, we prove that deciding if there exists a tiling for the given unit tiles is NP-complete.

4. Yasuhiko Takenaga and Yusuke Miura:
“Vertex Coloring of Chordal+k_1e$-k_2$e Graphs”, 18th International Workshop on Combinatorial Algorithms, pp.192-201, Nov. 2007.
概要: $\mathcal{F}+k$e ($\mathcal{F}-k$e resp.) graph is the classes of graphs which can be obtained by adding (deleting resp.) at most k edges to a graph in graph class $\mathcal{F}$. Complexity of the problems on these graphs can be measured using k which is a parameter to represent the closeness to graph class $\mathcal{F}$. In this paper, we consider chordal+k_1e$-k_2$e graphs which are the classes of graphs obtained by adding and deleting edges at the same time from chordal graphs. We show an algorithm for vertex coloring of

chordal+k_1e−k_2e graphs and prove that it is fixed parameter tractable.

5. Yasuhiko Takenaga and Nao Katougi: "Tree-Shellability of Restricted DNFs", IEICE Trans. Inf. and Syst., Vol.E91-D, No.4, April 2008.
概要: A tree-shellable function is a positive Boolean function which can be represented by a binary decision tree whose number of paths from the root to a leaf labeled 1 equals the number of prime implicants. In this paper, we consider the tree-shellability of DNFs with restrictions. We show that, for read-k DNFs, the number of terms in a tree-shellable function is at most k^2. We also show that, for k-DNFs, recognition of ordered tree-shellable functions is NP-complete for $k = 4$ and tree-shellable functions can be recognized in polynomial time for constant k.

研究会等

1. 東海林貴司, 武永康彦:
"積項の長さに制限を付けた論理関数の Ordered Tree-Shellability", 京都大学数理解析研究所講究録, 1375, 130–136, 2004 年 5 月.
概要: Ordered tree-shellable 論理関数とは、1 へのパスの個数が主項の数と同じになるような順序付き二分木表現を持つ正論理関数である。本論文では、積和形論理式で与えられた関数が ordered tree-shellable であるか否かを判定する問題を考える。各積項の長さが 2 以下の場合は多項式時間で判定できることが知られているが、長さ 3 以上の積項を 1 個含む場合、長さ 3 の積項を 2 個含む場合も多項式時間で判定できることを示す。

2. 東出賢一, 武永康彦:
"比較可能+ke グラフの彩色問題の計算量", 電子情報通信学会技術研究報告, COMP2004-82, 71–77, 2005 年 3 月.
概要: 比較可能 $+ke$ グラフとは、比較可能グラフに辺を k 本加えたグラフである。本論文では、k を定数としたときの比較可能 $+ke$ グラフの頂点彩色問題の計算量を考える。まず、比較可能 $+1e$ グラフの彩色問題が多項式時間で解けることを示す。次に、3SAT 問題から比較可能 $+2e$ グラフの彩色問題への多項式時間帰着を示し、比較可能 $+2e$ グラフの彩色問題が NP 完全であることを示す。

3. 東出賢一, 武永康彦:
"比較可能+ke グラフの彩色問題", 京都大学数理解析研究所講究録, 1426, 153–158, 2005 年 4 月.
概要: 一般には NP 完全である頂点彩色問題は、比較可能グラフに対しては多項式時間で解くことができる。そのため、比較可能グラフに近いグラフについても多項式時間で解ける可能性がある。本論文では、比較可能グラフに辺を k 本加えた比較可能 $+ke$ グラフを考える。まず、比較可能 $+1e$ グラフの彩色問題が多項式時間で解けることを示す。次に、3SAT 問題からの帰着により、比較可能 $+2e$ グラフの彩色問題が NP 完全であることを示す。

4. Yasuhiko Takenaga:
"Coloring Comparability-ke Graphs", 京都大学数理解析研究所講究録, 200?年?月.
概要: $\mathcal{F}$+ke and $\mathcal{F}$-ke graphs are classes of graphs close to graphs in a graph class $\mathcal{F}$. They are the classes of graphs obtained by adding or deleting at most k edges from a graph in F. In this paper, we consider vertex coloring of comparability−ke graphs. We prove that vertex coloring of comparability-1e graphs is solved in polynomial time. We also show that vertex coloring of comparability−ke graphs is closely related with vertex coloring of comparability graphs with restrictions that some pairs of vertices must have the same color.

5. 加藤木直, 武永康彦:
"リテラルの出現回数に制限を加えた関数の tree-shellablity 判定問題", 電子情報通信学会技術研究報告, COMP2005-12, 25–30, 2005 年 5 月.

概要: Tree-shellable 論理関数とは、根のノードから１にラベル付けされた葉ノードへのパスの数が、主項の数と同じになるような二分木表現を持つ正論理関数である。本稿では、積和形論理式で与えられた関数中の同一変数の出現回数と tree-shellability に関する性質を明らかにする。またそれを用い、リテラルの出現回数に制限を加えた場合の、(ordered) tree-shellable であるか否かを判定する問題の複雑さ明らかにする。本稿のアルゴリズムにより、同一リテラルの出現回数が高々定数の場合、tree-shellable 論理関数、ordered tree-shellable 論理関数の判定がともに多項式時間で実行できる。

6. 加藤木直, 武永康彦:
"種々の制限を加えた ree-shellable 論理関数判定問題の複雑さ", 電子情報通信学会技術研究報告, COMP2005-67, 31–37, 2006 年 3 月.
概要: Tree-shellable 論理関数とは、根のノードから 1 にラベル付けされた葉ノードへのパスの数が、主項の数と同じようになるような二分決定木表現をもつ正論理関数である。また二分決定木が順序付きであるならば、ordered tree-shellable 論理関数であるという。本稿では、関数内の積項に含まれるリテラルの最大数が 4 以上の場合、ordered tree-shellable 論理関数判定問題は NP 完全問題であることを明らかにする。また同一変数の出現回数に制限を加えた場合の (oredered) tree-shellable 論理関数判定問題の性質を明らかにし、その結果より (oredered) tree-shellable 論理関数判定問題の計算複雑さを明らかにする。

7. Yasuhiko Takenaga:
"Coloring Comparability-ke Graphs", 京都大学数理解析研究所講究録, 1489, pp.239-244, May 2006.
概要: $\mathcal{F}$+ke and $\mathcal{F}$-ke graphs are classes of graphs close to graphs in a graph class $\mathcal{F}$. They are the classes of graphs obtained by adding or deleting at most k edges from a graph in $\mathcal{F}$. In this paper, we consider vertex coloring of comparability$-k$e graphs. We prove that vertex coloring of comparability-1e graphs is solved in polynomial time. We also show that vertex coloring of comparability$-k$e graphs is closely related with vertex coloring of comparability graphs with restrictions that some pairs of vertices must have the same color.

8. 三浦勇介，武永康彦:
"Chordal+k_1e$-k_2$e グラフの頂点彩色問題", 電子情報通信学会コンピュテーション研究会, COMP2006-53, pp.21-27, 2007 年 3 月.
概要: $\mathcal{F}+k_1\mathrm{e}-k_2\mathrm{e}$ グラフとは、グラフ族 $\mathcal{F}$ に含まれるグラフに高々k_1 本の辺を追加し、高々k_2 本の辺を削除して得られるグラフ族である。本稿では、$\mathcal{F}+k_1\mathrm{e}-k_2\mathrm{e}$ グラフの頂点彩色問題を考える。$\mathcal{F}$ が辺の縮退、非隣接頂点の同一化の両方に関して閉じているならば、そのグラフに対する頂点彩色問題は fixed parameter tractable であることを示す。また、chordal+$k_1\mathrm{e}-k_2\mathrm{e}$ グラフの頂点彩色問題のアルゴリズムを提案し、パラメータ k_1,k_2 に関して fixed parameter tractable であることを示す。

9. S. Yamashita, S. Minato, and D. M. Miller:
"An EfficientDecision Diagram Structure for Design Verification of Quantum Circuits under a Practical Restriction", Hokkaido University, Division of Computer Science, TCS Technical Reports, TCS-TR-A-07-33, Dec. 2007.
概要: Recently much attentions have been paid to quantum circuit design to prepare for the future "quantum computation era." Like the conventional logic synthesis, it should be important to verify and analyze the functionalities of generated quantum circuits. Therefore, we propose an efficient verification method for quantum circuits under a practical restriction. Thanks to the restriction, we can introduce an efficient verification scheme based on decision diagrams called Decision Diagrams for Matrix

Functions (DDMFs). Then, we show analytically what are the advantages of our approach based on DDMFs over the previous techniques. In order to introduce DDMFs, we also introduce new concepts, quantum functions and matrix functions, which may also be interesting and useful on their own for designing quantum circuits.

10. 中山昌光, 武永康彦:
"多変量閾値関数の非明示的OBDD表現", 電子情報通信学会コンピュテーション研究会, 2008年3月.
概要: 規則的な構造を持つ OBDD では、OBDD のレベル間の接続関係を OBDD を用いて非明示的に表現することにより、さらに効率的な表現が可能になる場合がある。本研究では、種々の規則性を持つデータの論理関数表現に現れる、多変量閾値関数の OBDD の非明示的表現について研究を行なった。非明示的 OBDD 表現を用いた場合、多変量閾値関数の重みの和を w とすると、$O(k \log w)$ のサイズで表現可能となり、通常の OBDD 表現を用いた場合より大幅に小さなサイズで表現可能であることを示した。

学会大会等

1. 松川弘明, 武永康彦:
"区間グラフの OBDD 表現について", 2005年電子情報通信学会総合大会, D-1-9, 2005年3月.

2. 新井滋, 武永康彦:
"Rogue 脱出判定問題", 特定領域研究「新世代の計算限界」組合せゲーム・パズル第1回ミニ研究集会, 2005年9月.

3. 新井滋, 武永康彦:
"Rogue 脱出判定問題の PSPACE 完全性", 特定領域研究「新世代の計算限界」組合せゲーム・パズル第2回ミニ研究集会, 2007年3月.

4. 番能孝生, 武永康彦:
"OBDD を用いた画像処理アルゴリズム", 2008年電子情報通信学会総合大会, 2008年3月.

5. 古妻浩一, 武永康彦:
"OBDD によるナンバーリンクの解法", 特定領域研究「新世代の計算限界」組合せゲーム・パズル第3回ミニ研究集会, 2008年3月.

文部科学省科学研究費補助金 特定領域研究 (課題番号 16092226)

A04: 組み合わせ最適化における指数サイズ・多項式時間近傍の設計

組み合わせ最適化問題に対する指数サイズ近傍のアプローチに対して、基礎的な技法とその理論、具体的な問題に対する近傍と解法の設計と実装、解法のハードウェア化ないし準ハードウェア化の各側面から研究を行った。特に、技法の基礎としてグラフの分枝分割アルゴリズムについての研究に努力を傾注し、平面グラフの場合に理論的および実験的に良い結果を得ることができた。具体的な問題としては巡回セールスマン問題、集合被覆問題、クリーク被覆問題、グラフ描画問題などを対象として研究を行い、特にグラフ描画問題に対してこのアプローチの効果を確認した。解法のハードウェア化については、組み合わせ最適化問題の範囲にとどまらずより広いクラスの問題に対するアルゴリズムのハードウェア化を研究対象とし、多くの成果を挙げることができた。

研究組織

研究代表者： 玉木 久夫　明治大学 理工学部
研究分担者： 井口 幸洋　明治大学 理工学部

交付決定額 (配分額)

平成 16 年度	2,100,000 円
平成 17 年度	3,000,000 円
平成 18 年度	3,100,000 円
平成 19 年度	3,200,000 円
合　計	11,400,000 円

研究成果の概要

- 学術誌
  Theoretical Computer Science (2004 年), IEICE Transactions on Fundamentals of Electronics(2004 年 2 件), IEICE Transactions on Infromation Systems(2004 年、2006 年、2007 年), Random Structures and Algorithms (2005 年), Theory of Computing Systems (2006 年), Japanese Journal of Applied Physics(2006 年) Journal of Multiple-Valued Logic (2007 年) ACM Transaction on Algorithms (2007 年　採録決定),
- 国際会議
  SWAT (Humlebaek, 2004 年), STACS (Montpellier, 2004 年), ISMVL (Toronto, 2004年, Calgary, 2005年, Singapore 2006年, Oslo 2007年), MWSCAS (Hiroshima, 2004年), GLSVLSI (Chicago, 2005年), Cool Chips VIII(Yokohama, 2005年), DSD (Porto, 2005年), RM (Tokyo, 2005年), ICALP (Lisbon, 2005年), ALENEX (San Francisco, 2008年)

1 はじめに

本課題では、組み合わせ最適化問題に対する指数サイズの近傍のアプローチをテーマにしている。おもに、次の３つの方向から研究を進めて来た。

(A) 一般的、基礎的な技法およびそのための理論

(B) 具体的な問題に対する近傍と解法の設計と実装

(C) 解法のハードウェア化ないし準ハードウェア化

(A) において、技法の核にあるのは、グラフやハイパーグラフの再帰的分割であり、幅の小さい木分割 [3, 4] や分枝分割 [2] を効率的に求めることが重要になる。特にこの研究では、実装のしやすい動的計画法の基礎となる分枝分割を多用している。一般のグラフの場合には、最適の木分割や分枝分割を求める問題はＮＰ困難であるが、平面グラフの場合には、分枝分割を求める $O(n^4)$ 時間のアルゴリズムが知られていた [5]。本研究では、海外研究協力者である Qiangping Gu と共にこの計算時間を $O(n^3)$ に改良することに成功した。この結果は、ICALP2005 で報告し、また ACM Transactions on Algorithms に掲載決定している。また、同じく Qiangping Gu らと共に平面グラフの分子幅決定アルゴリズムの実装の研究を行い、時間およびメモリ量に関して効率の良い実装法を提案した。この結果は、ALENEX2008 （Workshop on Algorithm Engineering and Experiments) で報告した。これらの技術的な詳細については 2節で述べる。

(B) においては、巡回セールスマン問題、集合被覆問題、クリーク被覆問題、グラフ描画問題などを対象として研究を行った。巡回セールスマン問題については、文献にあるさまざまな指数サイズ近傍の実験的評価を行った他に、グラフ分割に基いた指数サイズ近傍を効率よく解くための動的計画法の表の効率よい構成法を提案した。グラフ描画問題については、平面グラフのコンパクト描画問題、平面グラフの低縮退問題などについて指数サイズ近傍アプローチのアルゴリズムの実装を行い、その効果を確認した。これらについての技術的概要については 3節で述べる。

(C) については、予備的な考察と使用する技法についてＦＰＧＡを中心にさまざまな研究を行った。その技術的概要については 4節で述べる。

2 平面グラフの分枝分割

2.1 平面グラフ最適分枝分割の $O(n^3)$ 時間アルゴリズム

G を無向グラフとする。G の分枝分割とは、G の辺集合を葉の集合として持つ、内点の次数がいずれも３であるような木のことを言う。T が G の分枝分割であるとき、その各辺 $\{p, q\}$ に対して $L_T(p, q) \subseteq E(G)$ によって p から q を通らない T の道によって到達できる T のすべての葉からなる集合を表す。T における G に関する $\{p, q\}$ の幅を、$L_T(p, q))$ の辺と $L_T(q, p))$ の辺の両方に接続するような G の頂点の個数と定義する。G の分枝分割 T の幅は、その辺の G に関する幅の最大値と定義される。G の分枝幅は、G の分枝分割の幅の最小値と定義される。

分枝幅は Robertson と Seymour [2] によってよりよく知られた木幅の概念 [3, 4] との関連において導入された。分枝幅や木幅の小さいグラフにおいては、動的計画法によりさまざまな組み合わせ問題を効率よく解くことができるため、グラフの分枝幅・木幅を決定したり、最小幅の分割を求めたりする問題は極めて重要である。分枝幅や木幅が定数で抑えられる場合には、これらの問題はいずれも線型時間で解くことができる [20]。しかしながら、与えられたグラフの幅が k を超えないかどうか決定する問題は、k が入力の一分として与えられた場合ＮＰ完全である [18, 5]。グラフを平面グラフに制限した場合には、木幅に対しては依然として多項式時間アルゴリズムが知られていないものの、分枝幅を求める $O(n^2)$ 時間のアルゴリズムが Seymour と Thomas [5] によって与えられている。同じ論文には、この分枝幅アルゴリズムをサブルーチンとして、平面グラフの最適な分枝分割を求める $O(n^4)$ 時間のアルゴリズムが与えられている。我々は、このアルゴリズムの実行時間を $O(n^3)$ に改良した。以下では、その技術的な概要を述べる。

準備

多重グラフ G に対して、$V(G)$ によってその頂点集合、$E(G)$ によってその辺集合を表す。各頂点集合 $X \subseteq V(G)$ に対して、$\delta_G(X)$ によって、X のなかにひとつの端点をもち、その外にもう一方の端点を持つ辺の集合を表す。互いに素な頂点集合 $U, V \subseteq V(G)$ に対して、$E_G(U,V)$ によって、U のなかにひとつの端点をもち、V の中にもう一方の端点を持つ辺の集合を表す。Seymour と Thomas の平面グラフに対する分枝幅アルゴリズムおよび分枝分割アルゴリズムの双方は共に関連した多重グラフに対する刻み幅・分割への帰着に基づいている。

多重グラフ G の刻み分割は G の頂点集合を葉の集合として持つ、内点の次数がいずれも 3 であるような木のことを言う。T が G の刻み分割であるとき、その各辺 $\{p,q\}$ に対して $L'_T(p,q) \subseteq V(G)$ によって p から q を通らない T の道によって到達できる T のすべての葉からなる集合を表す。T における G に関する $\{p,q\}$ の幅を、$L'_T(p,q))$ の頂点と $L'_T(q,p))$ の頂点を結ぶような G の辺の個数と定義する。G の刻み分割 T の幅は、その辺の G に関する幅の最大値と定義される。G の刻み幅は、G の刻み分割の幅の最小値と定義される。

G を、描画の指定された平面グラフとする。G のメディアルグラフ $M(G)$ [5] とは、G の平面双対グラフの、頂点集合と面集合の接続 2 部グラフのことを言う。G が単純グラフであっても一般には $M(G)$ は多重グラフになることに注意する。Seymour と Thomas [5] は、$M(G)$ の刻み幅は G の分枝幅のちょうど 2 倍であることを示した。さらに、$M(G)$ の最適刻み分割から G の最適分枝分割への変換は線型時間で行うことができる。そこで、我々は、平面多重分割の最適刻み分割を求めるアルゴリズムを考えることにする。

定理 1 G を n 個の頂点と $O(n)$ 本の辺を持つ平面多重グラフとするとき、G の最小幅刻み分割は $O(n^3)$ 時間で構成することができる。

この結果を証明するために、Seyour と Thomas のアルゴリズムを幾分異なった枠組みで記述する。まず、刻み分割を根付き 2 分木とみなすことから始める。$\sigma(V)$ によって、集合 V の一元部分集合全体からなる集合を表す。集合 V における 2 進併合を次のように帰納的に定義する。

1. $\sigma(V)$ は V における 2 進併合である。
2. もし $\mathcal{M}$ が V における 2 進併合であり、X, Y が $\mathcal{M}$ における V の相異なる極大部分集合であるならば、$\mathcal{M} \cup \{X \cup Y\}$ は V における 2 進併合である。

V における 2 進併合が V 自身を要素として持つとき、その 2 進併合は完全であるという。G をグラフとするとき、$V(G)$ の完全 2 進併合には自明な形で G の刻み分割が対応する。反対に、G の刻み分割が与えられたとき、その分割の辺のひとつを根として選ぶことにより、その刻み分割に対応する $V(G)$ における完全 2 進併合が決定される。$V(G)$ における 2 進併合 $\mathcal{M}$ の幅を $X \in \mathcal{M}$ のすべてを考えたときの $|\delta_G(X)|$ の最大値と定義する。$\mathcal{M}$ が完全であるとき、この定義は $\mathcal{M}$ に対応する刻み分割の幅と一致する。$\mathcal{M}$ と $\mathcal{M}'$ が共に $V(G)$ における 2 進併合であるとき、$\mathcal{M}' \supseteq \mathcal{M}$ であるならば $\mathcal{M}'$ は $\mathcal{M}'$ を拡張すると言う。$\mathcal{M}$ を $V(G)$ における 2 進併合とし、その幅は G の刻み幅 k 以下であるとする。X と Y を $\mathcal{M}$ の相異なる極大要素とするとき、もし $\mathcal{M} \cup \{X \cup Y\}$ を拡張する完全 2 進併合で幅が k であるものがあるならば、X と Y は $\mathcal{M}$ において併合可能と言うことにする。

我々は、Seymour と Thomas の平面グラフに対する分枝分割アルゴリズムを、2 進併合の概念を用いて次のように定式化する。これ以降 G は常にアルゴリズムに与えられた多重平面グラフを表し、k はその刻み幅を表すものとする。このアルゴリズムは初期 2 進併合 $\sigma(V(G))$ から出発して、$V(G)$ の完全 2 進併合が得られるまで 2 進併合を拡張して行く。その仮定で、「保持している 2 進併合は、幅 k の 2 進併合に拡張可能である」という不変条件を保持する。ここで、k は入力グラフ G の刻み幅を表す。

Algorithm ST
入力: 連結な n 頂点、$O(n)$ 辺平面多重グラフ G
出力: $V(G)$ の刻み分割で 幅が最小のもの
(1) G の刻み幅 k を決定する。
(2) $\mathcal{M}^0 = \sigma(V(G))$ とおき、$i = 1$ とおく。
(3) 以下を $i < n$ である間繰り返す。

(3.1) $\mathcal{M}^{i-1}$ の対 $\{X,Y\}$ で、$E_G(X,Y) \neq \emptyset$ であるようなものを、併合可能な対がみつかるまで走査する。併合可能な対 $\{X,Y\}$ がみつかったら、$X^i = X \cup Y$. とおく。
(3.2) $\mathcal{M}^i = \mathcal{M}^{i-1} \cup \{X^i\}$ とおいて、i を $i+1$ におきなおす。
(4) 完全２進併合 $\mathcal{M}^{n-1}$ の表す刻み分割を返す。

このアルゴリズムの正しさは、i についての数学的帰納法によって各 $\mathcal{M}^i$ が定義されかつ幅 k の完全２進併合に拡張可能であることを示すことによって証明される。証明はここでは省略する。

X と Y が $\mathcal{M}$ において併合可能かどうか調べるには、G において $\mathcal{M}' = \mathcal{M} \cup \{X \cup Y\}$ の各要素である頂点集合をそれぞれひとつの頂点に縮約して得られる多重グラフ $G_{\mathcal{M}'}$ に対して Seymour と Thomas の刻み幅アルゴリズムを呼び、その幅が k 以下であるかどうかを調べる。このテストの呼び出し回数は $O(n^2)$ 回であり、刻み幅アルゴリズムの実行時間は $O(n^2)$ であるため、アルゴリズムＳＴの全体の実行時間は $O(n^4)$ となる。

この、刻み幅アルゴリズムを用いた併合可能性のテストを高価なテストと呼ぶ。次のサブセクションでは、高価なテストの回数を減らすための補題を与える。

バリア補題

補題 1 G を刻み幅 k の多重グラフとする。$\mathcal{M}$ を $V(G)$ における２進併合とし、X, Y を $\mathcal{M}$ の相異なる極大要素とする。X, Y が $X', Y', Z, W \in \mathcal{M}$ を用いて $X = X' \cup Z$、$Y = Y' \cup W$ と表され、かつ $E_G(Z,W) = \emptyset$ であると仮定する。このとき、もし X と Y が $\mathcal{M}$ において併合可能であり $|\delta_G(X' \cup Y')| \leq k$ であるならば X' と Y' は $\mathcal{M} \setminus \{X,Y\}$ において併合可能である。

証明: $\mathcal{M}' = \mathcal{M} \setminus \{X,Y\}$ とおき、また $\mathcal{M}^*$ によって $\mathcal{M} \cup \{X \cup Y\}$ を拡張する幅 k の完全２進併合を表す。$\mathcal{M}_1 = \mathcal{M}' \cup \{X' \cup Y', X' \cup Y' \cup Z\} \cup (\mathcal{M}^* \setminus \mathcal{M})$、$\mathcal{M}_2 = \mathcal{M}' \cup \{X' \cup Y', X' \cup Y' \cup W\} \cup (\mathcal{M}^* \setminus \mathcal{M})$ とおくと、$\mathcal{M}_1$ と $\mathcal{M}_2$ はともに $\mathcal{M}' \cup \{X' \cup Y'\}$ を拡張する $V(G)$ における完全２進併合である。したがって X' と Y' が $\mathcal{M}'$ において併合可能であることを示すためには、, $\mathcal{M}_1$ または $\mathcal{M}_2$ の幅が k であることを示せば良い。$\mathcal{M}_1 \setminus \mathcal{M}^* = \{X' \cup Y', X' \cup Y' \cup Z\}$、$\mathcal{M}_2 \setminus \mathcal{M}^* = \{X' \cup Y', X' \cup Y' \cup W\}$、でありまた仮定より $|\delta_G(X' \cup Y')| \leq k$ であるから、$|\delta_G(X' \cup Y' \cup Z)| \leq k$ または $|\delta_G(X' \cup Y' \cup W)| \leq k$ であることを示せば目的は達せられる。. $A = X' \cup Y'$ および $B = V(G) \setminus (X \cup Y)$ とおこう。仮定より $E_G(Z,W) = \emptyset$ であるから、

$$\begin{aligned} &\delta_G(A \cup Z) \\ &\quad = \quad E_G(A,B) \cup E_G(Z,B) \cup E_G(A,W), \end{aligned}$$

が成り立ち、右辺の集合は互いに素であるから

$$\begin{aligned} &|\delta_G(A \cup Z)| \\ &\quad = \quad |E_G(A,B)| + |E_G(Z,B)| + |E_G(A,W)|, \end{aligned}$$

が成り立つ。同様にして

$$\begin{aligned} &|\delta_G(A \cup W)| \\ &\quad = \quad |E_G(A,B)| + |E_G(W,B)| + |E_G(A,Z)| \\ &|\delta_G(A)| \\ &\quad = \quad |E_G(A,B)| + |E_G(A,Z)| + |E_G(A,W)| \\ &|\delta_G(B)| \\ &\quad = \quad |E_G(A,B)| + |E_G(Z,B)| + |E_G(W,B)|. \end{aligned}$$

が成り立つ。したがって

$$\begin{aligned} &|\delta_G(A \cup Z)| + |\delta_G(A \cup W)| \\ &\quad = \quad 2|E_G(A,B)| + |E_G(Z,B)| + \\ &\qquad\quad |E_G(W,B)| + |E_G(A,Z)| + |E_G(A,W)| \\ &\quad = \quad |\delta_G(A)| + |\delta_G(B)|. \end{aligned}$$

が成り立つ。さらに、仮定より $|\delta_G(A)| \leq k$ でありまた X と Y は $\mathcal{M}$ において併合可能であるから $|\delta_G(B)| = |\delta_G(X \cup Y)| \leq k$ である。したがって、

$$|\delta_G(A \cup Z)| + |\delta_G(A \cup W)| \leq 2k$$

である、$|\delta_G(A \cup Z)| \leq k$ か $|\delta_G(A \cup W)| \leq k$ の少なくとも一方が成り立つ。[]

$\mathcal{M}$ を $V(G)$ における２進併合とする。M の要素の列 $X_0 \supset X_1 \supset \ldots \supset X_m$ は、各 X_i（$1 \leq i \leq m$）が X_{i-1} の $\mathcal{M}$ における極大真部分集合であるとき、$\mathcal{M}$ における連鎖と呼ばれる。C によって

上の連鎖を表すとき、X_0 と X_m はそれぞれ C の極大元、極小元と呼ばれ、$\top C$ および $\bot C$ と表される。連鎖 C が条件 $|\delta_G(\top C \setminus \bot C)| > k$ を満たすとき C はバリアと呼ばれる。ここで k は約束どおり、G の刻み幅を表す。

補題 2 *(バリア補題)* $\mathcal{M}$ を $V(G)$ における幅 k 以下の２進併合とし、X, Y を $\mathcal{M}$ において併合可能な極大元の対とする。２進併合 $\mathcal{M}' \subseteq \mathcal{M}$ とその相異なる極大元 $X' \subseteq X$ と $Y' \subseteq Y$ で、次の条件を満たすものがあったとする。

1. $|\delta(X' \cup Y')| \leq k$
2. $E_G(X \setminus X', Y \setminus Y') = \emptyset$
3. $\top B = X$ かつ $\bot B \supseteq X'$ であるようなバリア B は $\mathcal{M}$ のなかに存在しない
4. $\top B = Y$ かつ $\bot B \supseteq Y'$ であるようなバリア B は $\mathcal{M}$ のなかに存在しない

このとき、X' と Y' は $\mathcal{M}'$ において併合可能である。

証明: $Z = X \setminus X'$、$W = Y \setminus Y'$ とおく。$X_0 \supset X_1 \supset \ldots \supset X_h$ を極大元 $X_0 = X$ と極小元 $X_h = X'$ を持つ $\mathcal{M}$ の連鎖とする。同様に、$Y_0 \supset Y_1 \supset \ldots \supset Y_j$ を極大元 $Y_0 = Y$ と極小元 $Y_j = Y'$ を持つ $\mathcal{M}$ の連鎖とする。$\mathcal{M}_1$ を $\mathcal{M}$ から X_i（$0 \leq i < h$）と Y_i（$0 \leq i < j$）をすべて除いて得られる $V(G)$ の２進併合とする。すると $\mathcal{M}_1$ は $\mathcal{M}'$ を拡張し、$X_i \setminus X_{i+1}$（$0 \leq i < h$）のそれぞれと $Y_i \setminus Y_{i+1}$ （$0 \leq i < j$）のそれぞれは $\mathcal{M}_1$ の極大要素である。この $\mathcal{M}_1$ を次のように拡張して、極大元 X, Y を持ち X の極大真部分集合が X' と Z と、Y の極大真部分集合が Y' と W であるような２進併合 $\mathcal{M}_2$ をつくる。まぅ、$\mathcal{M}_1$ に $i = 2, \ldots, h$ に対して順次 $X_0 \setminus X_i$ を加えて極小元が $X_0 \setminus X_1$ で、極大元が $X_0 \setminus X_h = Z$ であるような連鎖を導入する。同様にして、極小元が $Y_0 \setminus Y_1$ で、極大元が W であるような連鎖を導入し、その結果を $\mathcal{M}_2$ とする。もし $\delta_G(X_0 \setminus X_i) > k$ であるなら極大元 X_0 と極小元 X_i を持つ $\mathcal{M}$ の連鎖は仮定に反してバリアであるから、したがって $i = 2, \ldots, h$ に対して $\delta_G(X_0 \setminus X_i) \leq k$ が成り立つ。同様にして、$i = 2, \ldots, j$ に対して $\delta_G(Y_0 \setminus Y_i) \leq k$ が成り立つ。したがって、$\mathcal{M}_2$ の幅は k 以下である。$\mathcal{M}_2$ の極大要素の集合は $\mathcal{M}$ のそれと等しいので、X と Y が $\mathcal{M}$ において併合可能であるという仮定からそれらが $\mathcal{M}_2$ においても併合可能であることがわかる。

こうして、$X', Y', Z, W \in \mathcal{M}_2$ が成り立ち、$X = X' \cup Z$ と $Y = Y' \cup W$ は $\mathcal{M}_2$ の極大元で、X と Y は $\mathcal{M}_2$ において併合可能である。しかも、仮定より $E_G(Z, W) = \emptyset$ と $|\delta_G(X' \cup Y')| \leq k.$ が成り立つ。したがって補題1を $\mathcal{M}_2$ に適用することができ、X' と Y' は $\mathcal{M}_2 \setminus \{X, Y\}$ において併合可能であることが結論される。$\mathcal{M}_2 \setminus \{X, Y\}$ は $\mathcal{M}'$ を拡張するので X' と Y' は $\mathcal{M}'$ において併合可能である。[]

高価なテストの節約

この節では $\mathcal{M}^i$, $0 \leq i \leq n-1$ によってアルゴリズム ST の G に対する実行において順次構築される $V(G)$ の２進併合を表す。我々の併合可能性テスト法を記述するために、次の記法を用いる。T^i によって、実行開始から $\mathcal{M}^i$ が構築される直前までに、併合可能性テストが適用された $\mathcal{M}^i$ の要素対すべてからなる集合を現す。T^i は、アルゴリズム ST における、要素対の走査順を指定すれば、入力グラフ G に対して一意に決まる。T^i の部分集合で、以下に述べる我々の方法において、高価なテストが適用される要素対の集合を P^i で表す。さらに、Q^i によて、P^i に属する要素対のうち、テストの結果が「併合不能」であったようなものの集合を現す。X, Y を $\mathcal{M}^{n-1}$ の相異なる２要素で $E_G(X, Y) \neq \emptyset$ であるようなものとするとき、$\mu_Y(X)$ によって X の $\mathcal{M}^{n-1}$ に属する部分集合で $E_G(X, Y) = E_G(X', Y)$ を満たすののうち、極小なものを表す。さらに、$Q(X, Y)$ によって、$X \supseteq X' \supseteq \mu_Y(X)$ かつ $Y \supseteq Y' \supseteq \mu_X(Y)$ を満たすような $\{X', Y'\} \in Q^{n-1}$ のすべてからなる集合を表す。$\mu_Y(X)$ と $Q(X, Y)$ は i を $X, Y \in \mathcal{M}^i$ なる最小のインデックスとするとき、$\mathcal{M}^i$ と Q^i から決定可能であるので、X と Y の併合可能性をテストするときにアルゴリズムが使用することが可能である。また、対要素ごとの集合包含関係は集合 $Q(X, Y)$ の全順序になっていることに注意する。

倹約併合可能性テスト
入力： $\mathcal{M}^{i-1}$ の相異なる極大元 X と Y
目的：X と Y が $\mathcal{M}^{i-1}$ において併合可能かどうか答える
(1) もし $|\delta(X \cup Y)| > k$ ならばNOと答える。
(2) そうでなく、もし $Q(X,Y) = \emptyset$ であれば高価なテストを実行しその結果を答える。
(3) そうでなければ、$\{X', Y'\}$ を $Q(X,Y)$ の極大対とする。. もし $\top B = X$ かつ $\perp B \supseteq X'$ であるか、$\top B = Y$ かつ $\perp B \supseteq Y'$ であるようなバリア B が $\mathcal{M}^{i-1}$ のなかに存在すればNO答える。そうでなければ, 高価なテストを実行してその結果を答える。

補題 3 上の併合可能性テストは正しい結果を答える。

証明: 答えがNOであるときに、X と Y が実際に併合不可能であることを示せばよい。テストがステップ(1)で終わるときには、これは自明である。テストがステップ(3)で高価なテストをスキップしてNOと答える場合を考える。$\{X', Y'\}$ を上の手続きで定義されたものとし、$Z = X \setminus X'$ および $W = Y \setminus Y'$ とおく。$Q(X,Y)$ の定義から、$\delta_G(Z,W) = \emptyset$ が導かれる。したがってバリア補題を $\mathcal{M} = \mathcal{M}^{i-1}$ および、$\mathcal{M}'$ を X' と Y' を要素として持つ $\mathcal{M}_j$ で j が最小のものに置き換えて適用することができ、X, Y が $\mathcal{M}^{i-1}$ において併合可能でないことが結論される。[]

補題 4 上の併合可能性テスト法を用いてアルゴリズム ST を実行したとき高価なテストの呼ばれる回数は $O(n)$ である。

証明: 省略。[]

併合可能性テスト以外の部分の計算時間は素朴なデータ表現とアルゴリズムによって $O(n^3)$ に抑えることができる。したがって、定理1が成り立つ。

2.2 分枝幅決定アルゴリズムの実装

SeymourとThomasによる平面グラフの分枝幅決定アルゴリズムは、分枝幅の決定を下で定義するネズミ捕りゲームの勝敗判定に帰着するため、ネズミ捕りアルゴリズムと呼ばれる。ネズミ捕りアルゴリズムを実装するためには、$O(n^2)$ のメモリが必要であるため、これがこのアルゴリズムを実用に供する際の大きな制約となっている。Hicksは[7]で、メモリを節約するために、基礎情報を格納せずに再計算する手法を提案している。我々は、再計算によらずにメモリを節約するためにアルゴリズムの背後にあるネズミ捕りゲームの性質を用い、できるだけ局所的に計算を完結させる方針で実装上の改良を行った。これらの手法により、素朴な実装ではメモリ不足のために解く事ができない例を扱うことができるようになった。以下でこれらの手法の詳細について述べる。

諸定義

SeymourとThomasのアルゴリズムは、平面グラフの分枝分割の問題を次に述べる刻み分割の問題に帰着する事によって解く[5]。分枝分割がグラフの辺の集合の階層的分割を表すのに対し、刻み分割は同様な表現でグラフの頂点の集合の階層的分割を表わす。

グラフ G の刻み分割 C は、C の葉頂点の集合が $V(G)$ であり、C の任意の内部頂点の次数が3であるような木である。G の刻み分割 C の各辺 c は、$V(G)$ の2分割 $(V1, V2)$ と自然に対応する。ここで、$V1$ と $V2$ は C を c によって切断して得られる2つの部分木における葉の集合を表す。$V1$ と $V2$ を結ぶ辺の本数を C における辺 c の幅と呼ぶ。刻み分割 C の幅とは、C の辺の中で一番大きい辺の幅のことであり、グラフ G の刻み幅とは、G の刻み分割を全て考えたときのそれらの幅の最小値のことである。

平面グラフの刻み幅決定アルゴリズム

平面グラフ G の分枝分割問題は実際は G を変換して得られる多重平面グラフの刻み分割に帰着される。以下では記述の簡明さのために単純平面グラフについての刻み幅決定について述べるが、多重平面グラフへの一般化は容易である。また、平面グラフ G に対しては固定した平面描画が与えられていると仮定する。刻み幅決定をするために、ネズミ捕りゲーム $RC(G,k)$ を考える。

ネズミ捕りゲーム $RC(G,k)$

プレイヤーはネズミとネズミ捕りの二人である。ゲームは G の平面描画上で行われる。ネズミは頂点に位置し、辺を通って移動する。ネズミ捕りは面に位置し、辺を横切って移動する。移動の途中でいったん辺上に停止する。ネズミとネズミ捕りはグラフ全体の構造を知っており、また自分と相手のグラフ上での位置を知っている。ネズミ捕りは、ネズミの移動を妨害するためにノイズを発生させる。ネズミ捕りの位置が f （f は G の辺また面）であるとき、G の各辺 e は、「G の平面双対グラフ上で f と e をとおる長さ k 以内の閉じた歩道が存在する」という条件のもとで通行不能となる。ここで、k は G とともに与えられたゲームのパラメータである。ゲームは次の手順で進行する。

1. ネズミ捕りが最初に自分が位置する面を選ぶ。
2. それに応じてネズミは自分の位置する頂点を選ぶ。
3. もしネズミが次数が k 以下の頂点にいて、ネズミ捕りがネズミのいる頂点と接する面にいるならばネズミ捕りの勝ちでゲームを終える。
4. ネズミ捕りは自分がいる面と接する面の中から、移動先の面を選び、現在面と移動先の面の共有する辺上に移動する。
5. ネズミは、通行不能でない辺を０個以上任意個通って移動する。
6. ネズミ捕りは予定していた面に移動し、３に戻る。

ネズミが永久に負けを免れて逃げ続けるとき、ネズミをゲームの勝者とする。

定理 2 (Seymour and Thomas[5]) 平面グラフ G が幅 k 以下で刻み分割可能であることの必要十分条件は $RC(G,k)$ においてネズミ捕りが必勝であることである。

ゲームのグラフ化

ネズミ捕りアルゴリズムに対する我々の実装手法を説明しやすくするために、アルゴリズムを Seymout と Thomas のオリジナルとは多少異なる形で定式化する。$RC(G,k)$ のゲームの状態を頂点、状態遷移を辺とする二部グラフ $H(G,k)$ を次のように定義する。ネズミ捕りが面 r にいるときにノイズが通過する辺を G から取り除くことによってできる G の部分グラフを G_r とする。同様に、ネズミ捕りが辺 e にいるときにノイズが通過する辺を取り除くことによってできる G の部分グラフを G_e とする。G_r、G_e の連結成分の集合をそれぞれ、C_e、C_r と表す。G の辺 e と面 r が接しているとき、$E(G_r) \subseteq E(G_e)$ であり、従って、$V(G)$ の分割 C_r は C_e の細分である。よってネズミ捕りが面 r から辺 e に移動したいときに、ネズミが移動できる頂点集合 $c \in C_e$ は、ネズミがどの $c' \in C_r$ の中にいたかだけに依存し、c' のどの頂点にいたかには依存しない。つまり、ネズミ捕りが面 r にいるとき、ゲームの状態は (r,c') によって表すことができる。

$H(G,k)$ の頂点集合は S,T であり、$R(G)$ を G の面の集合とすると、$S = \{(r,c)|r \in R(G), c \in C_r\}$ はネズミ捕りが面にいるときの状態の集合である。$T = \{(e,c)|e \in R(G), c \in C_e\}$ はネズミ捕りが辺にいるときの状態の集合である。そして、$(r,c) \in S$ と $(e,c') \in T$ の間に $H(G,k)$ の辺があるのは、これらの状態間で遷移が可能であるときである。すなわち、r と e が接していて、かつ、$c \subseteq c'$ であるときである。S の頂点 (r,c) からは、ネズミ捕りが次に向かう面 r' と r の間の辺 e を選択することによって、T の頂点 (e,c') へ遷移する。一方、T の頂点 (e,c) からの遷移はネズミの選択した頂点 v とネズミ捕りが移動する先である面 r' によって、$v \in c''$ である S の頂点 (r',c'') と決まる。

ネズミ捕りゲームの勝敗判定

もし、グラフ G の最大次数が k より大きい場合、ネズミは k よりも大きい次数を持つ頂点にいれば捕まることはないので、ネズミ必勝が自明であう。よってこれ以後はグラフ G の最大次数が k 以下である場合の判定法を述べる。$RC(G,k)$ における、ネズミ捕り必勝の状態を表す集合を以下のように帰納的に定義する。

1. $\hat{T}_0 = \phi$

2. $\hat{S}_0 = \{(r, \{v\}) | G$ の面 r と頂点 v は接している $\}$

3. $\hat{T}_{i+1} = \{(e, c) | e$ と接しているある面 r と $c' \subseteq c$ なるすべての $c' \in C_r$ に対して $(r, c') \in \hat{S}_i\} \cup \hat{T}_i$

4. $\hat{S}_{i+1} = \{(r, c) | r$ と接しているある辺 e と $c' \supseteq c$ なるある $c' \in C_e$ に対して $(e, c') \in \hat{T}_{i+1}\} \cup \hat{S}_i$

$\hat{S} = \bigcup_i S_i, \hat{T} = \bigcup_i T_i$ とおく。各 $\hat{S}_i$ に属する状態がネズミ捕り必勝状態であることと $S - \hat{S}$ に属す状態はネズミ必勝であるはともに数学的帰納法によって証明することができる。

勝敗判定アルゴリズムの実装においては、$H(G, k)$ から頂点を削除することにより $\hat{S}$、$\hat{T}$ を表現する。実装を工夫することにより、この計算は $O(n^2)$ 時間で実行することができる [5]。また、次の観察により、$\hat{S}$ を全て計算しなくても判定できる場合がある。

観察 1 ある面 r に対して、すべての $(r, c), c \in C_r$ が $\hat{S}$ に属するならば、$S = \hat{S}$ である。つまり、ゲームはネズミ捕り必勝である。

実装上の改良

改良を述べるために以下の定義を行う。R を G の平面双対グラフにおいて連結な面の集合とし、E_R を R の面と接する辺の集合とする。S, T の部分集合 S_R, T_R を $S_R = \{(r, c) | r \in R, c \in C_r\}, T_R = \{(e, c) | e \in E_R, c \in C_e\}$ と定義し、H_R を $S_R \cup T_R$ によって誘導される $H(G, k)$ の部分グラフとする。$\hat{S}, \hat{T}$ の構成規則を H_R の中だけで、適用して得られる $\hat{S}, \hat{T}$ の部分集合を $\hat{S}_R, \hat{T}_R$ で表す。$\hat{S}_R$ はネズミ捕りが R, E_R を動いて勝つことのできる状態の集合を表している。

面 $r \in R$ に対して $(r, c) \in S_R - \hat{S}_R$ なる $c \in C_r$ が唯一であるとき、c を R における面 r のサバイバル成分と呼ぶ。サバイバル成分が $\hat{S}$ に属さないことは、ネズミが勝つために必要である。$\hat{S}$ の帰納的定義において、$\hat{S}_0$ を $\hat{S}_0 \cup X \subseteq S$ で、$\hat{T}_0$ を $Y \subseteq T$ で置き換えて得られる集合を、それぞれ $\hat{S}(X, Y), \hat{T}(X, Y)$ で表す。$\hat{S}_R(X, Y), \hat{T}_R(X, Y)$ も同様に定義する。

改良 1．局所的な計算により、ネズミ捕り必勝であることが判る状態を見つける

ネズミ捕りが面 r にいて、ネズミが頂点 v から一歩も動けずにネズミ捕りに近づかれて、捕まってしまうための十分条件の真偽を調べる。次の定理を用いる。

定理 3 次数が k 以下である平面グラフ G において、r を G の面、v を G の頂点とする。以下の 2 条件を同時に満たす、ある二つの面 s, t があるならば、 $(r, v) \in \hat{S}$ である。

1. G の平面双対グラフにおいて r から最短距離で s まで行き、v の周りを時計回りして t まで行き、 r へ最短距離で戻ってくる閉じた歩道 C_1 の長さが k 以下である

2. 同じく G の平面双対グラフにおいて r から最短距離で s まで行き、v の周りを反時計回りして t まで行き、r へ最短距離で戻ってくる閉じた歩道 C_2 の長さが k 以下である

改良 2．$H(G, R)$, および、$\hat{S}, \hat{T}$ の漸増的な構成

G のある面 r_0 から、幅優先探索で面の集合 R を次第に増大させながら、$S_R, T_R, \hat{S}_R, \hat{T}_R$ を計算していく。

前に述べたように, $\hat{S}, \hat{T}$ は S, T から要素を削除することによって表現するので、使用メモリ量の節約が期待できる。実験により、改良 1 と組み合わせたときに効果が著しいことが判った。観察 1 から、$H(G, k)$ 全体を構成せずともアルゴリズムを停止することができる場合があるので、この方法は計算時間の改良にもなる可能性がある。

改良 3. サバイバル成分の特別扱い

以下の定理を用いる。

定理 4 R を G の平面双対グラフにおいて連結な面の集合とし、各 $r \in R$ がサバイバル成分 C_r を持つと仮定する。$\tilde{E}_R$ によって R に属する面と R に属さない面を隔てる辺の集合を表す。R を G の平面双対グラフから除いて作られる各連結成分 Q に対して $Y_Q = \{(e, c) | e \in \tilde{E}_R \cap \tilde{E}_Q, (e, c) \in \tilde{T}_R\}$ と定義する。$\hat{S} = S$ と次の条件は同値である。

ある辺 $e \in \tilde{E}_R$ と e に接する面 $r \in R$ があって、C_r を包含する C_e の要素を c_e とするとき、$(e, c_e) \in \hat{T}_Q(\phi, Y_Q)$ である。但し Q は G の平面双対グラフから R を取り除いたときにできる連結成分のうち、e を境界に持つようなものの集合である。

上の定理より、サバイバル成分を持つ連結な面の集合 R と E_R については $\tilde{E}_R$ に属す辺 e についてのみの状態集合だけを保持すればよく、メモリ節約効果が期待できる。

実験

素朴な実装と以上の改良を組み込んだ実装を比較する実験結果を表1に示す。実験に使用したグラフは巡回セールスマン問題のベンチマークライブラリ TSPLIB にある点集合を delaunay 三角形分割をしたものであり、Hicks[7] が使用したグラフと同じものである。例題名中の数値は点の個数を示している。Pentium4 3.2GHz の CPU を使用した。使用言語は java である。また、使用ヒープメモリ上限は６００ＭＢバイトとした。

表 1: 例題ごとの分枝幅決定計算時間

例題名	辺数	分枝幅	計算時間 (sec)	
			素朴	改良
eil51	140	8	0.937	0.265
lin105	292	8	4.44	0.625
ch130	377	10	23.1	0.735
pr144	393	9	4.50	0.500
kroB150	436	10	11.7	0.828
tsp225	622	12	37.1	1.16
pr226	586	7	7.64	1.13
rd400	1183	17	×	3.58
u574	1708	17	×	11.2
p654	1806	10	×	11.3
d657	1958	22	×	7.28
pr1002	2972	21	×	27.9
rl1323	3950	22	×	75.9
d1655	4890	29	×	57.8
rl1889	5631	22	×	188
u2152	6312	31	×	76.2
pr2392	7125	29	×	158
fl3795	11326	25	×	502
fnl4461	13359	48	×	622
rl5934	17770	41	×	1690
pla7397	21865	33	×	2370
usa13509	40503	63	×	9230
brd14051	42128	68	×	6680

3 具体的な問題

3.1 巡回セールスマン問題

平面巡回セールスマン問題における無効辺の除去

巡回セールスマン問題（TSP) のベンチマークインスタンスの多くは、平面上のユークリッド距離に基づくものである。これを平面 TSP と呼ぶことにする。一方我々の技法は、グラフ上の TSP に対して特化している。平面 TSP インスタンスからグラフ上の TSP インスタンスを得るために、これまでは主に、他の手法で良い巡回路を多数得て、それらを併合したグラフを用いて来た。我々は他の手法に頼らずに、もとの平面 TSP インスタンスから、最終的な解に至るまでのすべての過程を巨大近傍アプローチによって構成することを目指している。そのための最初のステップとして、最適性を損なわずに平面インスタンスから平面グラフインスタンスに変換する方法を開発して来た。ふたつの頂点を結ぶ辺のうち、直観的に決

して最適な巡回路に含まれることがないと思われるものが多数あるが、この直観を裏付けるための十分な理論は文献に見当たらない。かなり多くの場合に当てはまる、下の定理を証明した。2点 p, q の距離を $d(p,q)$ で表す。

定理 5 平面上の有限点集合 S が与えられたとする。S 中の相異なる 4 点 p, q, t_1, t_2 が次の条件を満たすならば、辺 pq は、S に対する最適巡回路に決して含まれることはない。この条件の記述において、c は十分大きな普遍的定数である。

1. 線分 pq と線分 t_1t_2 は交差する。
2. 線分 pq と線分 t_1t_2 の交点を r とおくと、$d(p,r)) \geq cd(t_1,t_2)$ かつ $d(q,r)) \geq cd(t_1,t_2)$
3. 線分 t_1t_2 を線分 pq に射影した結果の線分の長さを h, t_i、$i=1,2$、と線分 pq との距離を v_i とおく。このとき、直線 pq に関して t_1 と同じ側にあるどの点 $s \neq t_1$ に対しても $d(t_1,s) > \max\{h, v_1\}$ が成り立つ。また、直線 pq に関して t_2 と同じ側にあるどの点 $s \neq t_2$ に対しても $d(t_2,s) > \max\{h, v_2\}$ が成り立つ。

証明: （概略）最適な巡回路 T を固定する。T が辺 pq を含むと仮定して矛盾を導く。T は、点 p の直後に q を訪れると仮定する。一般に p_1、p_2、q_1、q_2 が相異なる S の点であり、T が p_1 の直後に p_2 を訪れ、q_1 の直後に q_2 を訪れるとき、「p_1 の直後に q_1 を訪れ、その後 T と逆の順に点をたどって p_2 にいたり、次に q_2 から T と同じ順に点をたどって p_1 に至る」という巡回路を、$\mathrm{flip}(T, p_1, p_2, q_1, q_2)$ と呼ぶことにする。

t_i から pq に降ろした垂線の足を f_i とするとき、一般性を失うことなく、f_1 の方が f_2 よりも p に近いか、または f_1 と f_2 は等しいと仮定する。T において t_1 の直前の点を s_1、t_2 の直後の点を s_2 とおく。T の最適性より、線分 pq は線分 s_1t_1 とも線分 t_2s_2 とも交差しない。また、有向直線 s_1t_1 と、有向直線 pq のなす角を 0 と π の間で表した値は $\frac{2}{3}\pi$ 以上である。もしそうでなければ、$\mathrm{flip}(T, p, q, s_1, t_1)$ の長さは T より短く、T の最適性に矛盾する。（ここで、c が十分に大きいことと、条件 $d(t_1,s) > v_1$ を用いる。）同様に、有向直線 t_2s_2 と、有向直線 pq のなす角を 0 と π の間で表した値は $\frac{2}{3}\pi$ 以上である。T から 3 辺 pq、s_1t_1、t_2s_2 を取り除いたときにできる 3 つの経路のうち、q を始点とするものを P_1、p を終点とするものを P_2、残りのひとつを P_3 と呼ぶことにする。P_1、P_2、および P_3 の逆転をつなげてできる巡回路が T よりも短い。（ここでは、c が十分に大きいことと、条件 $d(t_i,s) > h$、および上で示した角度の条件を用いる。）[]

平面グラフ TSP に対する動的計画法の表管理

平面グラフ TSP に対するグラフ分割に基づいた動的計画法アルゴリズムでは、動的計画法の表をどのように管理するかが重要なポインタになる。本研究では、以下に述べるような改良を発見した。

平面グラフ TSP の、グラフ分割に基づいた動的計画法に現れる部分問題は、次のように定義される、非交差マッチングによって表現される。

定義 1 有限集合 V 上の マッチングとは、V 上の非順序対（無向辺）の集合 M でグラフ (V,M) における各頂点の次数がちょうど 1 であるようなもののことを言う。全順序集合 $V = \{v_1 < v_2 < v_3 < \cdots < v_n\}$ 上の マッチング M が非交差であるとは、V のどのような 4 要素 v_i 、v_j 、v_k 、v_h、$i < j < k < h$ に対しても M が $\{v_i, v_k\}$ と $\{v_j, v_h\}$ の両方を持つことはないことを言う。

定義 2 M_1 を $U = \{u_1 < u_2 < \cdots, u_n\}$ 上の非交差マッチング M_2 を $V = \{v_1 < v2 < \cdots, v_m\}$ 上の非交差マッチングとするとき、$0 \leq k \leq \min(n,m)$ なる整数 k に対して、グラフ $G(M_1, M_2, k)$ を、頂点集合が $U \cup V \setminus \{v_1, v_2, \ldots v_k\}$ であり、辺集合が $M_1 \cup \{\{f(v_i), f(v_j)\} \mid (v_i, v_j) \in M_2\}$ であるようなグラフとして定義する。ただし、関数 f は、$1 \leq i \leq k$ に対しては $f(v_i) = u_i$, $k < i \leq m$ に対しては $f(v_i) = v_i$ によって定義される。

$G(M_1, M_2, k)$ の各連結成分は経路か閉路であることに注意する。

定義 3 M_1 を $U = \{u_1 < u_2 < \cdots, u_n\}$ 上の非交差マッチング M_2 を $V = \{v_1 < v2 < \cdots, v_m\}$ 上の非交差マッチングとするとき、M_1 と M_2 が k-貼り付け可能であるとは、$G(M_1, M_2, k)$ が閉路を持たないか、あるいはただひとつの閉路だけからなることを言う。後者は $m = n = k$ のときにのみ起こりうる。M_1 と M_2 が k-貼り付け可能であ

るとき、$\mathrm{paste}_k(M_1, M_2)$ によって $G(M_1, M_2, k)$ の経路を辺とみなして得られる $\{u_n < u_{n-1} < \ldots < u_{k+1} < v_{k+1} < \ldots < v_{m-1} < v_m\}$ 上のマッチングを表す。

$2n$ 個の頂点集合上の非交差マッチングは、n 組の括弧からなる wfp （well-formed parenthesis) によって表現される。W_n によって n 組の括弧からなる wfp 全体からなる集合を表す。$p \in W_m$ が非交差マッチング M_p を表し、$q \in W_n$ が非交差マッチング M_q を表を表すとき、$\mathrm{paste}_k(M_p, M_q)$ を表現する W_{m+n-k} の要素を $\mathrm{paste}_k(p, q)$ と表す。

定義 4 $2n$ 次の *DP* 表とは、W_n から非負実数への写像を言う。t_1 が $2m$ 次の *DP* 表、t_2 が $2n$ 次の *DP* 表であるとき、t_1 と t_2 の k-合成とは、$2(m+n-k)$ 次の *DP* 表 t でありそのエントリは、各 $p \in W_{m+n-k}$ に対して

$$\begin{aligned} t[p] &= \min\{t_1[q] + t_2[r] \\ &\quad \mid q \in W_m, r \in W_n, p = \mathrm{paste}_k(q, r))\} \end{aligned}$$

によって定義される。

DP 表の合成を素朴に行った場合、W_m と W_n に関する２重ループを実行することになり、時間は C_n をカタラン数として、$O(C_m C_n)$ を要することになる。この合成操作を、２段階に分けて効率的に実行する次のような手法を開発した。

定義 5 p、q を *well-formed* とは限らない任意の括弧列とするとき、$p+q$ で p と q の連結を表し、$\mathrm{pref}_k(p)$ によって p の長さ k の接頭辞を、$\mathrm{suf}_k(p)$ によって p の長さ k の接尾辞を表す。$\mathrm{rv}(p)$ によって、p のなかの括弧の順番と同時に括弧の向きを反転してできる括弧列を表す。また、$\mathrm{lp}(p)$ によって、p の中にある開き括弧の個数、$\mathrm{rp}(p)$ によって、p の中にある閉じ括弧の個数を表す。また、$\mathrm{open}(k)$ によって k 個の開き括弧からなる括弧列を、$\mathrm{close}(k)$ によって k 個の閉じ括弧からなる括弧列を表す。最後に、$\mathrm{trim}_k(p)$ によって、$\mathrm{pref}_k(p) + \mathrm{close}(\mathrm{lp}(\mathrm{pref}_k(p)) - \mathrm{rp}(\mathrm{pref}_k(p)))$ を表す。

定義 6 $p + \mathrm{close}(i)$ が W_m に属す括弧列であり、q が W_n に属す括弧列であると仮定する。このとき、$\mathrm{lftPlug}(p, q)$ によって、p の開き括弧のなかで対応する閉じ括弧のないものを $\mathrm{pref}_i(q)$ の要素で置き換えてできる列の後に $\mathrm{suf}_{n-i}(q)$ をつなげてできる括弧列を表す。また。$\mathrm{rgtPlug}(p, q)$ によって $\mathrm{rv}(\mathrm{lftPlug}(reverse(q), reverse(p))$ を表す。

定理 6

$$\begin{aligned} &\mathrm{paste}_k(p, q) \\ &= \mathrm{lftPlug}(\mathrm{pref}_{m-k}(\mathrm{rv}(p)), \mathrm{paste}_k(\mathrm{trim}_k(p), q)) \\ &= \mathrm{lftPlug}(\mathrm{pref}_{m-k}(\mathrm{rv}(p)), \\ &\quad \mathrm{rgtPlug}(\mathrm{paste}_k(\mathrm{trim}_k(p), \mathrm{trim}_k(q))), \\ &\quad \mathrm{suf}_{n-k}(q))) \end{aligned}$$

この定理を用いて、DP 表の合成を２段階で行うことができる。各 $p \in W_{m+n-k}$ に対して、$t[p]$ の定義式の右辺は、$\min\{t_1[q]] + t_2[q] \mid q \in W_m, r \in W_n, p = \mathrm{lftPlug}(\mathrm{pref}_{m-k}(\mathrm{rv}(q)), \mathrm{paste}_k(\mathrm{trim}_k(q), r))\}$ と表すことができる。ここで、$\mathrm{trim}_k(q) = s$ と $\mathrm{paste}_k(trim_k(q), r)) = u$ を固定すると t_2 の表エントリのみに基づいて、一部の最適化が可能であることに注目する。各 $s \in \{\mathrm{trim}_k(q) \mid q \in W_m\}$ と $u \in W_h$ に対して $a[s, u]$ を $\min\{t_2[r] \mid r \in W_n, u = \mathrm{paste}_k(s, r)\}$ によって定義すると、$t[p]$ の値は、$\min\{t_q[q] + a[s, u] \mid q \in W_m, s = \mathrm{trim}_k(q), p = \mathrm{lftPlug}(\mathrm{pref}_{m-k}(\mathrm{rv}(q)), u\}$ によって表すことができる。$a[s, u]$ の計算は表 t_2 のみに基づいているため、全体の表アクセス回数の削減が期待できる。実際、予備的に行った実験においては、m, n の値を２０程度まで大きくした場合に、１５パーセント程度の高速化が得られることを確認した。

3.2 三角形メッシュのコンパクト描画問題

以下で述べるアプローチは、三角形メッシュ以外にも適用可能であるがここでは簡単のため対象は三角形メッシュであると過程する。メッシュの頂点は整数座標の点（格子点）におかれるものとし、メッシュ全体をできるだけ小さい面積のなかに描くことが目的である。

与えられたメッシュに対する我々の近傍は、そのメッシュにおいて頂点を隣の格子点に動かすことにより得られる。移動しないことも含めると、

移動先の候補は５通りとなる。すべての頂点を独立に動かすことを許す場合、頂点数を n とすると近傍の大きさは n^5 となる。メッシュをグラフと見たときの分枝幅が十分小さければ（例えば６以内）であれば、この近傍内の最適解は動的計画法により高速に求めることができる。（ただし、目的関数は部分の最適値から全体の最適値を合成できるための一定の性質を満たすことが条件である。）

分枝幅が大きすぎる場合には、グラフの縮約の技法を用いる。いくつかの頂点集合を選び、そのそれぞれをひとつの頂点に縮約する。縮約された頂点集合においては、各頂点が同一の動きをするという制約を設けた近傍のなかでの最適解は、縮約結果のグラフの分割に基づいた動的計画法により求めることができる。実験によって、この技法による解の質が小近傍アプローチの結果を上回ることを確認した。

3.3 平面グラフの低縮退描画問題

平面グラフ G が与えられたとき、すべての辺が直線分となるように G を平面描画し、隣接する２辺のなす角の正弦の絶対値の最小値ができるだけ大きくなるようにしたい。この問題は、二次元凸包の分割統治法アルゴリズムのトレース例題を学生の自習用にランダムに生成する必要から生まれた。このトレースの課題においては、学習者は平面上の２０点程度の点集合を図で与えられて、分割統治法アルゴリズムを紙上で実行することを要求される。コピー解答の誘惑から学習者を解放するために、例題は乱数によって学習者ごとに生成される。与えられた点集合に対して、分割統治法アルゴリズムが再帰の各レベルで生成するすべての凸包の辺の和集合からなるグラフ G を考える。このグラフ G のなかで、隣り合う２辺のなす角が小さいかまたは π に近いと、アルゴリズムの実行に必要な３点の位置関係に対する判断を正しくできない可能性がある。そのために、G は上に述べたような基準によって描画されていることが望ましい。

この問題に対する厳密アルゴリズムはおそらく難しいので、次のような指数サイズ近傍を設計した。まず、頂点は整数座標の点のみに配置することにする。現在の描画を D とし、d を $(1,0)$、$(1,1)$、$(0,1)$、$(-1,1)$、$(-1,0)$、$(-1,-1)$、$(0,-1)$、$(1,-1)$ の８つのベクトルのうちのひとつとする。このとき、D の d-近傍 $N_d(D)$ は $N_d(D) = \{D_i \mid 0 \leq i < 2^n\}$ で定義される。ここで、n は G の頂点数であり、D_i は「D において頂点 j を、もし i の第 j ビットが１ならば d によって動かし、０ならば動かさない」という規則により得られる描画である。$N_d(D)$ 内の最適化には、D の平面双対グラフの分枝分割に基づいた動的計画法を用いる。アルゴリズム全体では、d を８通りに変化させながら近傍内最適化を行うフェーズを改良ができなくなるまで繰り返す。また、目的関数は隣接２辺の角の正弦の絶対値の辞書式列として、最小値が変化しない場合も改良が起きる可能性があるようにした。

この指数サイズ近傍を実装し、小近傍の局所探索との比較を行った。小近傍の探索では、各頂点に対して上の８通りのベクトルを足して改良が起きる場合にはその移動を施すということを変化がなくなるまで繰り返した。実験対象は、実際の教育で使用する問題のサイズと同様の２０頂点のランダムな例題１００題を用いた。指数サイズ近傍の結果の質が一貫して小近傍の結果を上回った。特に小近傍の場合には、まれに 0.04 などの非常に小さい値に収束するのに対し、指数サイズ近傍では、もっとも小さい場合でも収束値が 0.2 を下回ることはなかった。小近傍で他の工夫をすることにより同等の効果をあげることができないか、さらなる検討が必要だが、指数サイズ近傍によって素朴な小近傍の欠陥を補うことができることは明らかになった。

4 ハードウェア化の検討と技法の研究

大規模な組合せ問題を解く場合問題を専用のハードウェアで解かせる場合，そのハードウェアの構成は大きく分けて２つの形式が考えられる（並列コンピュータなどは除く）：

1. 組合せ問題を解く大規模な専用ハードウェア＋汎用コンピュータ
   汎用コンピュータは入出力などに使い，大規模計算には一切かかわらない

2. コプロセッサ（専用ハードウェア）＋汎用コンピュータ
特定の時間がかかる部分のみを計算する専用ハードウェアで行わせる．この部分は実際はＰＣの拡張バスなどにカードとして挿す形態となる。

(１）の方法は，高性能が見込まれるが高価である．(２）の方法は，安価ではあるが，バスの通信スピードがボトルネックとなりスピードアップの高価が薄くなることが問題となる。

我々は，実際にシステムを構築することを考えているので費用の点で (２）の方式をとらざるをえない．そこで，大規模問題をコンパクトに格納すること，高速な演算を行うことを目標にLUTカスケードというメモリを直列多段に接続する論理関数の格納法･計算法をFPGA上で実現し，一回の通信でなるべく多くの計算を専用ハードウェアで高速に計算させることを目標にこのアーキテクチャの有効性を検討した。専用ハードウェアとしてはFPGA (Field Programmable Gate Array) を利用するのが現実的である。この研究では、FPGA技術の広い知識と深い理解が必要であるために、組み合わせ最適化問題に限らずより広い範囲の問題についてFPGAによる解法実装の研究を行った。特にLUT (Look-Up Table) を直列多段に接続することで所望の論理回路がLUTカスケード法として提案されているが [12]，これをFPGA内部の組込みメモリを用いて実装するツールの開発と実際の応用例としてFIRフィルタの実装に用いて性能評価を行うなど、関連研究を積極的に行った。[13, 14, 15, 16, 17].

これらのハードウェアの記述には，Verilog-HDLを用いている．これを直に記述する手間とバグが混入する可能性が高くなるので，仕様から直にVerilog-HDL記述を生成するようなソフトウェアツールの開発も行い実際に使用している．この場合，仕様をメニューなどで与えて，開発したツールを用いてVerilog-HDL記述を生成し，最後にFPGA開発システムで論理合成・配置配線を行い，実装することになる．これと並行して，一度，Verilog-HDLを生成するのではなく，C言語で動作を記述することでEDIFと呼ばれるネットリストに直接論理合成するシステムを購入し，性能評価を行った。実装の準備を進めている．

4.1 平面グラフ巡回セールスマン問題解法の準ハードウェア化の準備

平面グラフ巡回セールスマン問題を動的計画法を用いて専用ハードウェアで解く準備について本節で述べる．

専用ハードウェアで, マイクロプロセッサとソフトウェアとの組合せより高速に解を得るためには, 次の2つの工夫が必要である．

1. 並列度を上げること,
2. 解を予め計算する．これを圧縮して格納し，使用時に効率良く検索できること．

並列度を上げるためには, グラフを再帰的に分割し, 個々のグラフで解を求めればよい．個々に求めた解を併合するときに, それがまたTSPとしての解の性質を満たしているかの判定を行う．また, 併合後のグラフの内部接続を高速に求めたい．グラフをその度に辿って解を求めると時間がかかりすぎるので, 予め計算しておきそれを表に蓄える．

図1にグラフ G_1 と G_2 とからなるグラフを示す．G_1 は, 下から反時計回りに $0, 1, 2, \ldots, 7$ とラベル付けされた8本の辺によって外部と接続している．内部では, 0と7, 1と6, 2と3, 4と5とが内部の太い実線で表されたツアーで接続されている． これを ((()())) で表現する．

一方, G_2 は時計回りに $0, 1, 2, \ldots, 5$ とラベル付けされた6本の辺がグラフから出ている．内部では, 0と5, 1と2, 3と4とが接続されている．これは (()()) と表現する．

この2つのグラフを図1にあるように5本の共通辺で接続し, 一つのグラフ G_{12} に併合する．このグラフから出る辺A,B,C,Dは, 図から明らかなようにAとD, BとCとでグラフ内部で接続されている．これは (()) で表せる．

ここで，入力として解の型 ((()())), および, (()()) を与えると, 合成結果の型 (()) を高速に出力する方法を考える．一番単純な方法は，予めすべての解をグラフを辿ることで求めておき表に蓄え, 実際に値が必要なときは表引きによる求めることである．この括弧を用いた記法は, 左括弧を0, 右括弧を1とすれば2進数で表せる．本例では, 14ビット入力, 4ビット出力の論理関数で表現できる．

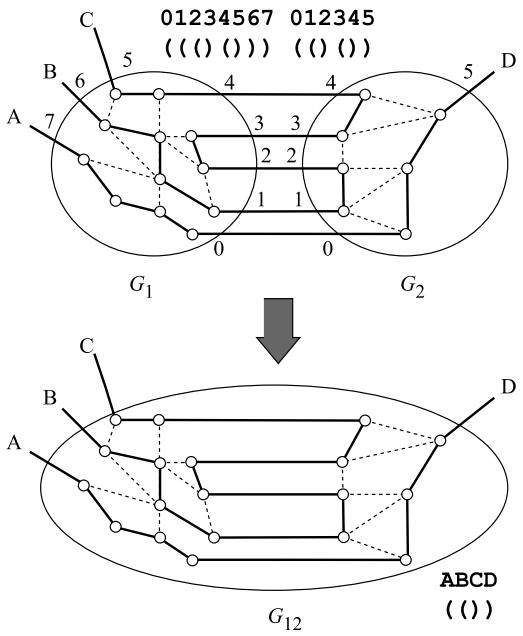


図 1: 解の合成

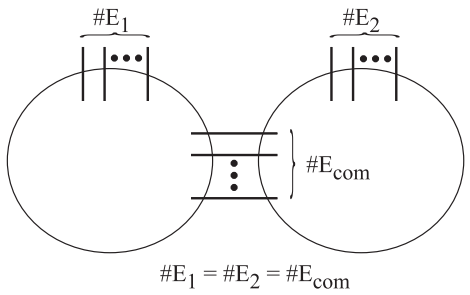


図 2: 取り扱うグラフ

ここでは, 問題を簡単にするために図 2のように, 2 つのグラフ G_1 と G_2 の共通辺の個数 $\#E_{com}$ とその他のグラフから出る辺の個数 $\#E_1$, および, $\#E_2$ とが全て等しい場合についての結果を示す.

無効な入力（それぞれの部分グラフにおいて括弧の対応が正しくない入力）が多いから, すべての有効な解 (TSP として成立する解) は, 入力の全ての組合せに比べるとわずかである. 無効な入力は, 計算時（表引き時）には, 与えないので出力は don't care としてよい.

また, 図 3のように併合したグラフ内に内部ループが存在する場合は, TSP の解として不適当である. この場合は, TSP の解としては成立しないという出力を出す.

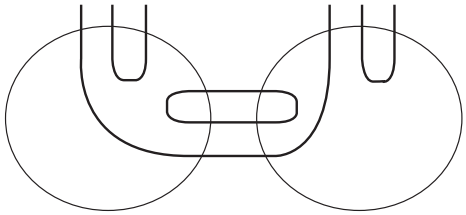


図 3: 内部ループが存在するグラフ

例 1 図 *2*において, $\#E_1 = \#E_2 = \#E_{com} = 3$ の場合を考えてみよう. この場合, 入力は *12* ビット, 出力は *6* ビットとなる. 単純にメモリで記憶させると $2^{12} \times 6 = 4096 \times 6 = 24576$ ビットの記憶容量を必要とする. 一方, *TSP* として意味のある接続は表 *2*に示すように *17*通りのみである. *(*例終り*)*

例 1では, 意味のある答えは 17 通りだけであった. これを記憶する場合は, 入力と出力のペアで記憶させればよいので $(12 + 6) \times 17 = 306$ ビットあれば十分である. なお, このように格納した場合は, 表引きにはハッシュ関数や連想メモリ, および, 連想メモリを模擬する専用ハードウェアを用いれば高速に検索が可能となる [9, 10].

このように, 有効な解の個数は, 入力の全ての組合せに比べて少ない. ソフトウェアで TSP の解を求めるとき, この計算を表引きに変えれば高速化の可能性がある. また, ハードウェアで計算する場合も, この方法は利用できる.

そこで, 実際にメモリに蓄えられる大きさがどのくらいか, 解を圧縮して記憶させたいので実際の解のパターンを全て求めてみた. 計算結果を表 3に示す. 計算に用いた計算機は, Intel Xeon CPU 3.2GHz×2, 3.5GB RAM, Windows XP Professional V.2002 Service Pack 2 の PC (Personal Computer) を用いた. 使用言語は, C 言語であり, コンパイラには gcc を用いた. 表 3に入力ビット数の増加に対して, 有効な解の個数とその解全てを求めるのに要した計算時間, 有効解の割合を示す. 例 1の例は, 表 3では, $in = 12$ の場合に相当する. この場合, 意味のある解は 17 個であり, 17 個の解を求めるのに $6.8\mu s$ 要している. 有効な組合せは 4096 個中 17 個であるので, これを $17 \div 2^{12} \times 100 = 0.41\%$ と示した.

表より, 意味のある組合せは割合としては, わ

表 2: 12 入力 6 出力の論理関数で表したグラフの意味のある接続

in	*out*
((())))((()))	((()))
((())))(()())	(()())
((())))(())()	(())()
((())))()(())	()(())
((())))()()()	()()()
(()())((()))	(()())
(()())()(())	(()())
(()())()()()	(())()
(())()((()))	()(())
(())()()(())	()(())
(())()()()()	()()()
()(())((()))	(())()
()(())(()())	(()())
()(())(())()	(())()
()()()((()))	()()()
()()()(()())	()(())
()()()(())()	()()()

ずかであることがわかる. しかし, 48 入力のときの全ての解の個数は 10G 個を超えているこれをそのままメモリなどに記憶するのは難しい

今後, 2 つのグラフの併合で入力数が異なる場合, また, 効率的な圧縮法の考案, もしくは, 一部を表引きに変えるなどの検討が必要であることがわかった.

表 3: 意味のある解の種類

in	有効な解の個数	計算時間	有効解の割合 [%]
4	1	93ns	6.2000
8	3	580ns	1.2000
12	17	6.8μs	0.4100
16	117	100μs	0.1800
20	955	2ms	0.0910
24	8,389	36ms	0.0500
28	79,889	630ms	0.0300
32	791,905	11s	0.0180
36	8,228,777	210sec	0.0120
40	87,675,915	1hours	0.0080
44	963,501,291	17hours	0.0055
48	10,771,003,505	12days	0.0038

参考文献

[1] D.L. Applegate and R.E. Bixby and V. Chvátal and W.J. Cook, The Traveling Salesman Problem - A Computational Study, Princeton University Press, 2006.

[2] N. Robertson and P.D. Seymour, Graph minors X. Obstructions to tree-decomposition, Journal of Combonatorial Theory, Series B, 153-190, 1991.

[3] N. Robertson and P.D. Seymour, Graph minors I. Excluding a forest, Journal of Combinatorial Theory, Series B, 35: 39-61, 1983

[4] N. Robertson and P.D. Seymour, Graph minors. II. Algorithmic aspects of tree-width, Journal of Algorithms, 7: 309-322, 1986

[5] P.D. Seymour and R. Thomas, Call Routing and the Ratcatcher. Combinatorica, 14(3), 217-241,1994.

[6] Q.-P. Gu and H. Tamaki, Optimal Branch-decomposition of planar graphs in $O(n^3)$time. ICALP 2005, 373-384.

[7] I.V. Hicks,Planar Branch Decomposition I:The Ratcatcher,INFORMS Journal on Computing. Vol.17, No.4, Fall2005, pp.402-412.

[8] H. Tamaki, A liner time heuristic for the branch-decomposition of planar graphs. ESA2003, 765-775.

[9] T. Sasao, “A Design method of address generators using hash memories,” *International Workshop on Logic and Synthesis(IWLS06),* Vail, Colorado, U.S.A, June 7-9, 2006, pp.102-109.

[10] H. Nakahara, T. Sasao and M. Matsuura, "A CAM emulator using look-up table cascades," *14th Reconfigurable Architectures Workshop RAW 2007*, March 2007, Long Beach California, USA. CD-ROM RAW-9-paper-2.

[11] T. Sasao, M. Matsuura, and Y. Iguchi, "A cascade realization of multiple-output function for reconfigurable hardware," *International Workshop on Logic and Synthesis(IWLS01),* Lake Tahoe, CA, June 12-15, 2001, pp. 225-230.

[12] Y. Iguchi, T. Sasao, and M. Matsuura , "Realization of multiple-output functions by reconfigurable cascades, " International Conference on Computer Design :VLSI in Computers & Processors (ICCD-2001), Austin, TX, Sept. 23-26, 2001. pp. 388-393.

[13] T. Sasao, Y. Iguchi, and T. Suzuki, "On LUT cascade realizations of FIR filters," *DSD2005, 8th Euromicro Conference on Digital System Design: Architectures, Methods and Tools*, Porto, Portugal, Aug. 30 - Sept. 3, 2005, pp.467-474.

[14] K. Nakamura, T. Sasao, M. Matsuura, K. Tanaka, K. Yoshizumi, H. Nakahara and Y. Iguchi, "A memory-based programmable logic device using look-up table cascade with synchronous static random access memories," Japanese Journal of Applied Physics, Vol. 45, No. 4B, pp. 3295-3300, April 2006.

[15] Y. Iguchi, T. Sasao, and M. Matsuura, "On designs of radix converters using arithmetic decompositions," IEEE Proc. of ISMVL-2006 (CD-ROM), May 17-20, 2006.

[16] Y. Iguchi, T. Sasao, and M. Matsuura, "On designs of radix converters using arithmetic decompositions, — Binary to Decimal Converters —," IEEE Proc. of ISMVL-2007, (CD-ROM), May 14-15, 2007.

[17] Y. Iguchi, T. Sasao, and M. Matsuura, "Design methods of radix converters using arithmetic decompositions," IEICE Trans. on Information and Systems, Vol. E90-D, No. 6, pp. 905-914, 2007.

[18] S. Arnborg and D.G. Corneil and A. Proskurowski, Complexity of finding embeddings in a k-tree, SIAM Journal on Algebraic and Discrete Methods, 8: 277-284, 1987

[19] S. Arnborg and J. Lagergren and D. Seese, Easy problems for tree-decomposable graphs, Journal of Algorithms, 12: 308-340, 1991

[20] H.L. Bodlaender, A Linear-Time Algorithm for Finding Tree-Decompositions of Small Treewidth, SIAM Journal on Computing, 25: 1305-1317, 1996.

[21] E.D. Demaine and M.T. Hajiaghayi and N. Nishimura and P. Ragde and D.M. Thilikos, Approximation algorithms for classes of graphs excluding single-crossing graphs as minors, Journal of Computer and System Sciences, 69-2: 166-195, 2004.

研究業績一覧

学術論文

1. Holger Bast, Kurt Mehlhorn, Guido Schäfer, Hisao Tamaki:
"Matching Algorithms Are Fast in Sparse Random Graphs", Proc. of the 21st Symposium on Theoretical Aspects of Computer Science (STACS04), LNCS 2996, 81-92, 2004 年 3 月.
概要: c をある程度大きい正の定数とするとき、頂点数が n で辺の本数が cn のランダムグラフ（またはランダム 2 部グラフ）において、任意の非最大マッチングは長さ $O(\log n)$ の増加路を高確率で持つことを示した。このことは、代表的なマッチングア

ルゴリズムがそのようなランダムグラフ上で高速であることを意味する

2. Tetsuo Asano, Naoki Katoh, Hisao Tamaki, Takeshi Tokuyama:
"On Geometric Structure of Global Roundings for Graphs and Range Spaces", Proc. of the 9th Scandinavian Workshop on Algorithm Theory (SWAT04), LNCS 3111, 455-467, 2004 年 10 月.
概要: ハイパーグラフ H と H の頂点をインデックスとする [0,1]-値のベクトルが与えられたとき、そのベクトルの大域丸めとは、２進値ベクトルで、H のどの辺においても、もとのベクトルの要素の総和丸めベクトルの要素の総和の差が１未満であるようなものを言う。我々は両立集合の概念を用いて、大域丸めの集合のなす構造を調べる。ハイパーグラフが最短路公理と呼ばれる公理集合を満たすときには、大域丸めの個数が $n+1$ 個以下である（n はハイパーグラフの頂点数）と予想し、この予想がいくつかの幾何的レンジスペースや、直並列グラフの最短路ハイパーグラフに対しては成り立つことを示した。

3. Tetsuo Asano, Naoki Katoh, Hisao Tamaki, Takeshi Tokuyama:
"The structure and number of global roundings of a graph", Theoretical Computer Science, 325-3, 425-437, 2004 年１０月.
概要: グラフの各頂点に実数値が割り当てられているときに、それらの値を整数値に丸めたい。このとき、あらかじめ定められた頂点集合族に属する集合のそれぞれにおいて、もとの値の合計と丸め結果の合計の差が１未満であるという大域的な条件をつける。頂点集合族として、グラフの最短路上の頂点集合をとった場合について、制約を満たす丸めの集合の構造を調べ、木、サイクルや、それらを拡張したグラフに対して、丸めの個数が常に $n+1$ 以下であることを示した。

4. Qiangping Gu, Hisao Tamaki:
"Optimal branch-decomposition of planar graphs in $O(n^3)$ time", Simon Fraser University School of Computing Technical Report, SFU-CMPT-TR 2004-14, 1-12, 2004 年１１月.
概要: n 頂点の平面グラフ G が与えられたとき、G の分枝分割で幅が最小であるようなものを求める $O(n^3)$ 時間のアルゴリズムを与える。これはこれまでの最良の Seymour と Thomas による $O(n^4)$ 時間アルゴリズムを改良するものである。

5. Marin Dietzfelbinger, Hisao Tamaki:
"On the probability of rendezvous in graphs", Random Structures and Algorithms, 26-3, 266 - 288, 2005 年 3 月.
概要: グラフにおいて、各頂点がその隣接点のひとつを当確率で選ぶという実験を考える。この実験において、お互いを選びあうような頂点対が生ずればランデブーが起きたという。この論文では、頂点数を一定にしたとき、ランデブー起きる確率が最も高いのは完全グラフであることを証明する。また、与えられたグラフに対してランデブーの起きる確率を計算する問題は#P 完全であることを示す。

6. S. Nagayama, T. Sasao, Y. Iguchi, and M. Matsuura:
"Area-time complexities of multi-valued decision diagrams", IEICE Transactions on Fundamentals of Electronics, Vol.e87-A, No.5, pp.1020-1028, May 2004.
概要: This paper considers Quasi-Reduced ordered Multi-valued Decision Diagrams with k bits (QRMDD(k)s) to represent binary logic functions. Experimental results show relations between the values of k and the numbers of nodes, the memory sizes, the numbers of memory accesses, and area-time complexity for QRMDD(k). For many benchmark functions, the numbers of nodes and memory accesses for QRMDD(k)s are nearly equal to $1/k$ of the corresponding Quasi-Reduced ordered Binary Deci-

sion Diagrams (QRBDDs), and the memory sizes and the area-time complexities for QRMDD(k)s are minimum when $k = 2$ and $k = 3 - 6$, respectively.

7. H. Qin, T. Sasao, M. Matsuura, K. Nakamura S. Nagayama, and Y. Iguchi:
"A realization of multiple-output functions by a look-up table ring", IEICE Transactions on Fundamentals of Electronics, Vol.e87-A, 3141-3150, Dec. 2004.
概要: A look-up table (LUT) cascade is a new type of a programmable logic devices (PLDs), which provides an alternative way to realize multiple-output functions. Two types of LUT cascades exist: A combinational LUT cascade, and a sequential LUT cascade. Compared with a combinational LUT cascade, the sequential LUT cascade is more flexible. Unlike FPGA realization of a logic function, accurate prediction of the delay time is relatively easy in a sequential-LUT cascade realization. A prototype of a sequential LUT cascade has been custom-designed with 0.35

8. Y. Iguchi, T. Sasao, and M. Matsuura:
"A method to evaluate logic functions in the presence of unknown inputs using LUT cascades", Proc. of the 34th International Symposium on Multiple-Valued Logic (ISMVL-2004), 302-308, May 19-22, 2004.
概要: In logic simulation, we often need to evaluate two-valued logic functions in the presence of unknown inputs. However, a naive method often produces imprecise results. In these case, we can obtain precise values by evaluating the regular ternary logic function for the given two-valued logic function. This paper shows a hardware realization of regular ternary logic functions. We use look-up table (LUT) cascades to implement double-rail logic representation. The evaluation time for n-input logic function is O(n). They require much smaller amount of hardware than naive memory implementations.

9. T. Sasao, H. Nakahara, M. Matsuura, and Y. Iguchi:
"Realization of sequential circuits by look-up table ring", Proc. of The 2004 IEEE International Midwest Symposium on Circuits and Systems (MWSCAS 2004), I-517-I-520, July 25-28, 2004.
概要: The LUT ring is a new type of memory-based realization of a sequential circuit that requires much smaller memory than conventional methods. In this paper, a method to realize a sequential circuit by a look-up table (LUT) ring is presented. The sequential circuit consists of a combinational part and feedback flip-flops. The combinational part is represented by a set of LUT cascades, and they are sequentially emulated by the LUT ring. The LUT ring uses two types of clocks: One evaluates the combinational part, and the other is the clock for state transitions of the sequential circuit. We present a method to reduce the clock period for state transitions, given limited memory size.

10. H. Qin, T. Sasao, and Y. Iguchi:
"An FPGA design of AES encryption circuit with 128-bit keys", Proc. of GLSVLSI 2005, pp.147–151, April 17, 2005.
概要: This paper addresses a pipelined partial rolling (PPR) architecture for the AES encryption. The key technique is the PPR architecture, which is suitable for FPGA implementation. Using the proposed architecture on the Altera Stratix EP1S20F780C5 FPGA, the AES-4SM achieves a throughput of 5.61 Gbps by using 20 M4Ks, and the AES-8SM achieves a throughput of 10.49 Gbps by using 40 M4Ks. Compared with the unrolling implementation that achieves a throughput of

20.48 Gbps by using 80 M4Ks on the same FPGA, implementations with the PPR architecture reduce the amount of memory up to 75(i.e., throughput divided by the size of memory for core) up to 9.6and rolling architectures, and fits on less expensive FPGAs.

11. K. Nakamura, T. Sasao, M. Matsuura, K. Tanaka, K. Yoshizumi, H. Qin, and Y. Iguchi:
"Programmable logic device with an 8-stage cascade of 64K-bit asynchronous SRAMs", Proc. of Cool Chips VIII, IEEE Symposium on Low-Power and High-Speed Chips, April 20, 2005.
概要: The first implementation of a new programmable logic device using LUT(Look-Up Table) cascade architecture is developed in 0.35um CMOS logic process. Eight 64Kb asynchronous SRAMs are simply connected to form an LUT cascade with a few additional circuits. Benchmark results show that it has a competitive performance to FPGAs.

12. Y. Iguchi and T. Sasao:
"Hardware to compute Walsh coefficients", Proc.of ISMVL-2005, pp. 75–81, May 19, 2005.
概要: This paper presents a method to compute a fragment of the Walsh coefficients of logic functions using hardware. First, it introduces the Walsh transformation tree, and shows a method to compute Walsh coefficients using the Walsh transformation tree. Next, it shows the hardware realization for the Walsh tree. The amount of hardware to compute a coefficient and the entire coefficients are O(2n) and O(n2・2n), respectively. FPGA implementations show their feasibility up to n = 14. The FPGA realization is at least 1253 times faster than a software implementation on a microprocessor for n = 14.

13. T. Sasao, Y. Iguchi, and T. Suzuki:
"On LUT cascade realizations of FIR filters", Proc. of 8th Euromicro Conference on Digital System Design: Architectures, Methods and Tools, pp. 467–474, Aug. 30, 2005.
概要: This paper first defines the n-input q-outputWS function, as a mathematical model of the combinational part of the distributed arithmetic of a finite impulse response (FIR) filter. Then, it shows a method to realize theWS function by an LUT cascade with k-input q-output cells. Furthermore, it 1) shows that LUT cascade realizations require much smaller memory than the single ROM realizations; 2) presents new design method for a WS function by arithmetic decomposition, and 3) shows design results of FIR filters using FPGAs with embedded memories.

14. T. Sasao, Y. Iguchi, and M. Matsuura:
"LUT cascades and emulators for realizations of logic functions", RM2005, pp. 63–70, Sept. 5, 2005.
概要: Two types of programmable logic devices using LUTs (Look-Up Tables) are presented. An LUT cascade directly implements logic functions by a series connection of LUTs, while an emulator emulates an LUT cascade by sequentially accessing LUTs. The LUT cascade is faster, but has a limited logic capability, while the emulator is slower, but has a higher logic capability. LUT cascades and emulators can be directly generated from the BDDs of target functions. Their performances are easy to estimate. The C-measure that show the complexity of LUT cascades are also presented. Functions with small C-measures have efficient LUT cascade and emulator realizations. Classes of functions that are suitable for LUT cascade and emulator realizations are also presented.

15. Qiangping Gu, Hisao Tamaki:

"Optimal branch-decomposition of planar graphs in $O(n^3)$ time", Proc. 32nd International Conference on Automata, Languages, and Programming, LNCS 3580, 373-384, 2005 年 7 月.
概要: n 頂点の平面グラフ G が与えられたとき、G の分枝分割で幅が最小であるようなものを求める $O(n^3)$ 時間のアルゴリズムを与える。これはこれまでの最良の Seymour と Thomas による $O(n^4)$ 時間アルゴリズムを改良するものである。

16. Holger Bast, Kurt Mehlhorn, Guido Schäfer, Hisao Tamaki:
"Matching Algorithms Are Fast in Sparse Random Graphs", Theory of Computing Systems, Vol. 39, No. 1, 3-14, 2006 年 2 月.
概要: c をある程度大きい正の定数とするとき、頂点数が n で辺の本数が cn のランダムグラフ（またはランダム 2 部グラフ）において、任意の非最大マッチングは長さ $O(\log n)$ の増加路を高確率で持つことを示した。このことは、代表的なマッチングアルゴリズムがそのようなランダムグラフ上で高速であることを意味する

17. K. Nakamura, T. Sasao, M. Matsuura, K. Tanaka, K. Yoshizumi, H. Nakahara and Y. Iguchi:
"A memory-based programmable logic device using look-up table cascade with synchronous static random access memories", Japanese Journal of Applied Physics, Vol. 45, No. 4B, pp. 3295-3300, April, 2006.
概要: A large-scale memory-technology-based programmable logic device (PLD) using a look-up table (LUT) cascade is developed in the 0.35-mm standard complementary metal oxide semiconductor (CMOS) logic process. Eight 64 K-bit synchronous SRAMs are connected to form an LUT cascade with a few additional circuits. The features of the LUT cascade include: 1) a flexible cascade connection structure, 2) multi phase pseudo asynchronous operations with synchronous static random access memory (SRAM) cores, and 3) LUT-bypass redundancy. This chip operates at 33MHz in 8-LUT cascades at 122mW. Benchmark results show that it achieves a comparable performance to field programmable gate array (FPGAs).

18. Y. Iguchi, T. Sasao, and M. Matsuura:
"On designs of radix converters using arithmetic decompositions", IEEE Proc. of ISMVL-2006, May 17-20, 2006.
概要: In arithmetic circuits for digital signal processing, radixes other than two are often used to make circuits faster. In such cases, radix converters are necessary. However, in general, radix converters tend to be complex. This paper considers design methods for p-nary to binary converters. It introduces a new design technique called arithmetic decomposition. It also compares the amount of hardware and performance of radix converters implemented on FPGAs.

19. Zhengbing Bian, Qian-Ping Gu, Marjan Marzban, Hisao Tamaki, Yumi Yoshitake:
"Empirical study on branchwidth and branch decomposition of planar graphs", ALENEX08 (Workshop on Algorithm Engineering and Experiments)), 2008 年 1 月 19 日.
概要: We propose efficient implementations of Seymour and Thomas algorithm which, given a planar graph and an integer β, decides whether the graph has the branchwidth at least β. The computational results of our implementations show that the branchwidth of a planar graph can be computed in a practical time and memory space for some instances of size about one hundred thousand edges. Previous studies report that a straightforward implementation of the algorithm is memory consuming, which could be a bottleneck for solving instances with more than a few thousands edges. Our

results suggest that with efficient implementations, the memory space required by the algorithm may not be a bottleneck in practice. Applying our implementations, an optimal branch decomposition of a planar graph of practical size can be computed in a reasonable time. Branch-decomposition based algorithms have been explored as an approach for solving many NP-hard problems on graphs. The results of this paper suggest that the approach could be practical.

20. Qian-Ping Gu, Hisao Tamaki:
"Optimal branch-decomposition of planar graphs in $O(n^3)$ time", ACM Transaction on Algorithms, （掲載予定）.
概要: n 頂点の平面グラフ G が与えられたとき、G の分枝分割で幅が最小であるようなものを求める $O(n^3)$ 時間のアルゴリズムを与える。これはこれまでの最良の Seymour と Thomas による $O(n^4)$ 時間アルゴリズムを改良するものである。

21. Yukihiro Iguchi, Tsutomu Sasao, and Munehiro Matsuura:
"On designs of radix converters using arithmetic decompositions", IEEE Proc. of 37th Int'l Symposium on Multiple-Valued Logic, CD ROM のためページ数の記載なし, 2007 年 5 月.
概要: In digital signal processing, radixes other than two are often used for high-speed computation. In the computation for finance, decimal numbers are used instead of binary numbers. In such cases, radix converters are necessary. In general, radix converters are complex. This paper considers design methods for binary to decimal converters. It introduces a new design technique based on weighted-sum (WS) functions. The method computes a WS function for each digit by an LUT cascade and a binary adder, then adds adjacent digits with decimal adders.

22. Yukihiro Iguchi, Tsutomu Sasao, and Munehiro Matsuura:
"Design methods of radix converters using arithmetic decompositions", IEICE Trans. on Information and Systems, Vol. E90-D, No. 6, pp. 905-914, 2007 年 6 月.
概要: In arithmetic circuits for digital signal processing, radixes other than two are often used to make circuits faster. In such cases, radix converters are necessary. However, in general, radix converters tend to be complex. This paper considers design methods for p-nary to binary converters. First, it considers Look-Up Table (LUT) cascade realizations. Then, it introduces a new design technique called arithmetic decomposition by using LUTs and adders. Finally, it compares the amount of hardware and performance of radix converters implemented by FPGAs. 12-digit ternary to binary converters on Cyclone II FPGAs designed by the proposed method are faster than ones by conventional methods.

23. Yukihiro Iguchi, Tsutomu Sasao, and Munehiro Matsuura:
"On designs of radix converters using arithmetic decompositions", Journal of Multiple-Valued Logic, Vol. 13, No. 4-6, pp. 503-520, 2007 年 11 月.
概要: In digital signal processing, radixes other than two are often used for high-speed computation. In the computation for finance, decimal numbers are used instead of binary numbers. In such cases, radix converters are necessary. This paper considers design methods for binary to q-nary converters. It introduces a new design technique based on weighted-sum (WS) functions. The method computes a WS function for each digit by an LUT cascade and a binary adder, then adds adjacent digits with q-nary adders. A 16-bit binary to decimal converter is designed to show the method.

研究会等

1. 井口幸洋, 鈴木隆広, 笹尾 勤:
"LUT リングを用いたプログラマブル・コントローラの構成法", 電子情報通信学会 第 2 種研究会 リコンフィギャラブルシステム研究会 (RECONF), pp.41-49, 2004 年 9 月 15 日.
概要: LUT(Look-Up Table) の出力の一部を入力にリング状に接続した LUT リングを用いてプログラマブル・ロジック・コントローラ (PLC: Programmable Logic Controller) を構成する方法を提案する．PLC で実行する命令の多くは，論理演算である．同時に実行可能な論理演算をまとめ, 全ての組合せに対する出力値を予め計算して一つの LUT に蓄える．これらの LUT を個別に実装するかわりに, 一個のメモリを繰り返しアクセスすることにより複数の LUT として実現し, 出力の一部は外部出力レジスタに, 残りは状態変数レジスタに記憶し処理を進める．PLC のプログラミング入力としては，ラダー図を用いる．LUT 内部に格納するデータ, および, 出力を外部出力や内部記憶用に振り分けるための制御回路データを生成するツールも開発した．いくつかのベンチマーク関数に対する性能評価を示し, 従来の PLC よりも高速であることを示す.

2. 田中克征, 吉住謙一, 中村和之, 笹尾勤, 松浦宗寛, Qin Hui, 井口 幸洋:
"LUT カスケードアーキテクチャによるプログラム可能 LSI の開発,", 電子情報通信学会 第 2 種研究会・第 8 回システム LSI ワークショップ, pp.315-318, 2004 年 11 月 30 日.
概要: LUT カスケードアーキテクチャによるプログラム可能 LSI の開発を行い，チップ試作を行った. 動作速度，チップ面積，消費電力について FPGA との比較を行い, LUT カスケードアーキテクチャの有効性を示す

3. Hui QIN, 笹尾勤, 井口幸洋:
"A design of AES encryption circuit with 128-bit keys using look-up table ring", 電子情報通信学会 VLSI 設計技術研究会, VLD2004-122, 2005 年 01 月 26 日.
概要: This paper presents a partial-rolling architecture for an AES encryption processor. By using a look-up table ring and pipeline techniques, the proposed architecture reduces the amount of memory by 75% while maintaining the memory efficiency (i.e., throughput divided by the size of memory for core). We implemented two architectures: AES-4SM and AES-8SM on Altera Stratix FPGA. The AES-4SM achieved 5.61 G-bit/s throughput by using 20 memory blocks (M4Ks), and the AES-8SM achieved 10.49 G-bit/s throughput by using 40 M4Ks. We also implemented the unrolling architecture that achieves 20.48 G-bit/s throughput by using 80 M4Ks on the same FPGA. Compared with the unrolling implementation, the AES-4SM and the AES-8SM improved the memory efficiency by 10.9% and 18.2%, respectively, and reduced the amount of the memory blocks by 75% and 50%, respectively. The proposed implementation AES-8SM has an higher memory efficiency than both the fastest unrolling implementation and the fastest rolling implementation designed by other group. The proposed implementation fills gap between the rolling and unrolling implementations, and fits on less expensive FPGA.

4. 笹尾勤, 井口幸洋, 鈴木隆広:
"FIR フィルタの LUT カスケードによる実現について", 電子情報通信学会 VLSI 設計技術研究会, VLD2004-126, 2005 年 03 月 10 日.
概要: 本論文では，FIR フィルタのビット分散演算を表す関数として, n 入力 q 出力 DA 関数を定義する．次に, DA 関数を k 入力 q 出力のセルを用いた LUT カスケードで実現する場合，単一の ROM で実現する場合に比べ必要メモリ量を大幅に削減でき, セル数，メモリの総ビット数が量子化ビット数, タップ数, セルの入力数が決まれば上限が見積もれることを示す.

5. 井口幸洋, 笹尾勤:
“Walsh 係数を計算するハードウェアについて”, 電子情報通信学会 VLSI 設計技術研究会, VLD2004-127, 2005 年 03 月 10 日.
概要: 論理関数の Walsh スペクトラムの係数の一部をハードウェアを用いて計算する方法を示す. まず, Walsh 変換木を定義し, Walsh 変換木から Walsh 係数を計算する方法を示す. 次に, この操作をハードウェアで実現する方法を示す. 1 個の係数を計算するハードウェア量は, $O(2^n)$ であり, 全ての係数を求めるハードウェア量は, $O(n^2 \cdot 2^n)$ である. FPGA デバイス対して回路設計を行った結果, 市販 FPGA 上で計算可能な論理関数の変数の個数 14 であった. また, マイクロプロセッサに比べて FPGA 実現は, $n = 14$ のとき 1253 倍高速である.

6. 笹尾勤, 井口幸洋, 鈴木隆広:
“FIR フィルタの算術分解を用いた LUT カスケードによる実現”, 電子情報通信学会 リコンフィギャラブルシステム研究会, 2005 年 9 月 15 日.
概要: FIR フィルタの分散演算の組合せ論理回路の数学モデルとして n 入力 m 出力 WS 関数を定義する. 次に WS 関数を k 入力 q 出力のセルを用いた LUT カスケードで実現する方法を示す. 更に, 次の 3 点を明らかにする：(1) 単一の ROM で実現する場合に比べ, LUT カスケードを用いた実現は, 必要メモリ量を大幅に削減できる. (2) 算術分解を用いて WS 関数を実現する新しい方法を提案する. (3) FIR フィルタを組込みメモリを用いて FPGA 上に実現した結果を示す.

7. 井口幸洋, 笹尾勤, 松浦宗寛:
“算術分解を用いた基数変換回路の構成法”, 電子情報通信学会, ディペンダブルコンピューティング研究会, DC2006-23, pp. 61-66, 2006 年 8 月 1 日.
概要: デジタル信号処理用の算術演算回路などでは, 基数が 2 以外の論理回路を用いて, 高速かつコンパクトな回路を実現する場合がある. この場合, 基数変換回路が必要であるが, 一般には複雑な回路となる. 本論文では, p 進数を 2 進数に変換する基数変換回路の構成法を提案する. 本構成法では, 算術分解と呼ぶ新しい方法を用いる. また, FPGA(Field Programmable Gate Array) 上に基数変改回路を実現したときのハードウェア量と性能の比較を行う.

8. 井口幸洋、笹尾勤、松浦宗寛:
“算術分解を用いた基数変換回路の構成法 (2)”, 電子情報通信学会、リコンフィギャラブルシステム研究会, RECONF2006-47, pp. 19-24, 2006-11-30.
概要: デジタル信号処理では, 高速演算のために 2 以外の基数がよく用いられる. また, 金融計算では, 10 進数が 2 進数の代わりに用いられる. このような場合, 基数変換回路が必要である. 本論文では, 2 進数を q 進数に変換する新しい基数変換回路の構成法を提案する. これは, weighted-wum (WS) 関数の概念に基づく新しい方法である. 各桁の WS 関数を LUT カスケードと 2 進加算器とで計算し, それらを q 進加算器で足し合わせることで基数変換器を構成する. 16 ビットの 2 進 10 進変換器を例にその構成法を説明する.

9. 井口幸洋、笹尾勤、松浦宗寛:
“算術分解を用いた基数変換回路の構成法 (3)”, 電子情報通信学会 VLSI 設計技術研究会, VLD2006-135, 2007 年 03 月 8 日.
概要: デジタル信号処理では, 高速演算のために 2 以外の基数がよく用いられる. また, 金融計算では, 10 進数が 2 進数の代わりに用いられる. このような場合, 基数変換回路が必要である. 我々は, 2 進数を q 進数に変換する新しい基数変換回路の構成法を提案している. これは, weighted-wum (WS) 関数の概念に基づく新しい方法である. 各桁の WS 関数を LUT カスケードと 2 進加算器とで計算し, それらを q 進加算器で足し合わせることで基数変換器を構成する. 本稿では, 基数変換の合成ツールを開発したので報告する. これは, FPGA 上の組込みメモリのデータ・パターンと Verilog-HDL での回路記述を生成する. 16 ビットの 2 進 10 進変換器を例にその結果を示す.

10. 玉木久夫、吉武由実:
“平面グラフの分枝幅決定アルゴリズムの効率的実装”, 情報処理学会アルゴリズム研究会, SIGAL-109-4, 19-26, 2006 年１１月.
概要: 平面グラフの分枝幅を決定するための、Seymour と Thomas によるネズミ捕りアルゴリズムについて主にメモリを節約するための工夫を施した上で実装し、素朴な実装では実行できない規模の例題に対して実行することを可能とした。

11. 玉木久夫、土屋裕希:
“平面ユークリッドＴＳＰの分割統治法ヒューリスティック”, 情報処理学会アルゴリズム研究会, SIGAL-111-17, 125-132, 2007 年３月.
概要: 平面ユークリッドＴＳＰの初期解構築法として、分割統治法によるアプローチを提案した。この方法は、部分解（巡回路）の統合のステップにおいてふたつの部分巡回路にさらに多数の辺を付加した統合グラフを作成して、その上での最適巡回路を動的計画法によって求めることに特徴がある。

12. 井口幸洋、笹尾勤、松浦宗寛,:
“基数変換回路の構成法 (4)”, 電子情報通信学会 リコンフィギャラブルシステム研究会, RECONF2007-48, pp.31-36, 2007年11月22日.
概要: デジタル信号処理を高速化するために RNS (residue number system: 剰余数系) を用いることがある. このような場合, RNS と 2 進数との間の基数変換回路が必要である. 本稿では, 剰余数 (RNS) から 2 進数への高速な変換回路の構成法を示す. 基数を指定することにより, RNS 二進変換回路を自動的に合成するシステムを試作し, 多数の回路を FPGA 上に実現した結果を示す.

13. 井口幸洋、笹尾勤、清水大和:
“基数変換回路の構成法 (5)”, 第 58 回 FTC 研究会, 2008 年 1 月 12 日.
概要: RNS (Residue Number System) を使用したシステムでは, 誤り検出や誤り訂正を RNS の性質を用いて行うことが知られている. 本発表では, 1 個の冗長桁を付加したシステムで, 誤り検出回路を LUT (Look-Up Table) カスケードで構成する方法を提案し, そのハードウェア量を RNS から 2 進数への基数変換回路のハードウェア量とほぼ同等であることをいくつかの例で示した.

14. 八木沢圭、山崎浩二、細川利典、玉木久夫:
“テストパターンの静的圧縮における厳密解と貪欲解の比較”, 電子情報通信学会ディペンダブルコンピューティング研究会, 2008 年 2 月 8 日.
概要: ドントケアに基く静的テストパターン圧縮をクリーク被覆問題や頂点彩色問題として定式化し、厳密解法と貪欲解法による解法実験を行った。その結果、今回実験の対象としたほとんどのテストパターンについて圧縮されたテストパターン数の貪欲解と厳密解が一致することがわかった。

文部科学省科学研究費補助金 特定領域研究 (課題番号 16092210)

A05: ハードウェアアルゴリズムの性能評価に関する研究

集積回路技術の進展に伴い、専用回路を構成することにより、処理時間に関わる計算限界を打破することが可能となってきている。このような専用回路を開発する場合、その基礎となるハードウェアアルゴリズムの設計が重要である。本研究では、近年の集積回路技術の進展を踏まえ、より現実に則した評価基準に基づくハードウェアアルゴリズムの性能評価と優れたハードウェアアルゴリズムの設計について研究を行い、次の成果を得た。１．配線遅延を考慮したハードウェアアルゴリズムの計算時間の評価基準を定めるとともに、配線遅延を考慮した高速な整数除算用および整数開平用ハードウェアアルゴリズムを開発した。２．多層配線に適した小面積な乗算用ハードウェアアルゴリズムを開発するとともに、配線層数の回路面積への影響を調べた。３．算術演算のための減算シフト型ハードウェアアルゴリズムの自動合成システムを開発した。４．剰余系演算及びガロア体上の演算に対するハードウェアアルゴリズムを開発した。

研究組織

研究代表者： 高木 直史　名古屋大学 大学院 情報科学研究科
研究分担者： 高木 一義　名古屋大学 大学院 情報科学研究科

交付決定額 (配分額)

平成 16 年度	2,000,000 円
平成 17 年度	3,500,000 円
平成 18 年度	3,700,000 円
平成 19 年度	2,000,000 円
合　計	11,200,000 円

研究成果の概要

- 学術誌
  IEEE Trans. Computers（2005 年），IEICE Trans. Fundamentals（2005 年），IEICE Trans. Fundamentals（2006 年），IEICE Trans. Fundamentals（2006 年），IEEE Trans. Computers（2008 年）
- 国際会議
  ARITH 17（Cape Cod, MA, USA, 2005 年），CHES 2005（Edinburgh, UK, 2005 年），ISPACS 2006（Tottori, Japan, 2006 年），ARITH 18（Montpellier, France, 2007 年），SASIMI 2007（Sapporo, Japan, 2007 年）
- 受賞
  Best Paper Award of CHES 2005（2005 年）
- 特許出願
  特願 2005-242956（2005 年），特願 2006-040859（2006 年），米国 20070050442（2007 年）

1　はじめに

集積回路技術の進展に伴い、これまでソフトウェアでは実現できなかった高度の実時間処理を、専用回路を構成して実現することが可能になってきている。すなわち、専用回路を構成することにより、処理時間に関わる計算限界を打破することが可能となってきている。このような専用回路を開発する場合、並列処理による高速計算や集積回路化に適した規則正しい回路構造等を意識した、ハードウェア実現向きのアルゴリズム、すなわち、「ハードウェアアルゴリズム」を設計することが重要である。

ハードウェアアルゴリズムの性能は、従来、アルゴリズムに基づく回路の、組合せ回路モデル上での段数や素子数、VLSI モデル上での面積等で評価してきた。しかし、集積回路技術の進展に伴い、これらのモデルおよび評価基準は、現実に則さなくなってきている。たとえば、回路の微細化に伴い、論理素子間を結ぶ配線における遅延が論理素子における遅延に対して無視できなくなってきており、回路の段数のみで計算時間を見積もることができなくなってきている。

本研究では、近年の集積回路技術の進展を踏まえ、より現実に則した評価基準に基づくハードウェアアルゴリズムの性能評価と優れたハードウェアアルゴリズムの設計について研究を行い、次の成果を得た。

1. 配線遅延を考慮したハードウェアアルゴリズムの計算時間の評価基準を提案した。
2. 配線遅延を考慮した整数除算用高速ハードウェアアルゴリズム、および、整数開平用高速ハードウェアアルゴリズムを開発した。
3. 多層配線に適した小面積な乗算用ハードウェアアルゴリズムを開発した。
4. 配線層数の回路面積への影響を調べた。
5. 算術演算のための減算シフト型ハードウェアアルゴリズムの自動合成システムを開発した。
6. 剰余系演算およびガロア体上の演算に対するハードウェアアルゴリズムを三つ開発した。

以下、これらの詳細を述べる。

2　配線遅延を考慮したハードウェアアルゴリズムの評価

ハードウェアアルゴリズムの性能評価は、回路を数学的にモデル化し、評価基準を定義した上で、その回路モデルに基づいて行う。評価基準として、従来は組合せ回路モデルにおける回路の段数や素子数、VLSI モデルにおける面積が用いられてきた。各論理素子の遅延を定数値とし、配線遅延を無視すれば、回路の計算時間は回路の段数に比例する。しかし、実際の CMOS 論理回路においては、遅延は素子のファンアウトや出力に接続される配線の長さに依存する。集積回路技術の進展に伴い、回路の微細化が進み、論理素子の遅延に対して配線遅延が相対的に大きくなってきている。

本研究では、より現実に則したハードウェアアルゴリズムの性能評価を行うために、配線長に依存する配線遅延を考慮した回路モデルを提案する。その上で、種々の並列乗算のハードウェアアルゴリズムについて回路全体の配線遅延を見積もり、配線による遅延を考慮した場合の回路の計算時間を評価する。

2.1　従来の回路モデルと評価基準

組合せ回路モデルでは、回路はある与えられた種類の論理関数（AND，OR など）を実現する入次数（ファンイン）制限のある論理素子（２入力 AND ゲート、３入力 OR ゲートなど）で構成される、フィードバックループをもたない論理回路である。各論理素子の出次数（ファンアウト）は制限しない。組合せ回路の段数とは、その回路の入力から出力に至る全ての経路の内、最も多くの論理素子を含む経路上の論理素子の個数である。組合せ回路の素子数とは、その回路を構成する論理素子の個数である。各論理素子での遅延が同一の一定値であり、配線上での遅延を無視するという仮定の下で、回路での計算時間は回路の段数に比例する。また、回路の素子数は回路のハードウェア量の一つの評価基準となる。

VLSI モデルでは、正方格子を考え、以下を仮定をする [1]。

1. 配線は、格子線に沿って、縦方向または横方向のみに引かれる。配線はｋ層である。（一

つの格子点で交わる配線は最大k本である。）二つの層の配線を接続するためには、コンタクトが必要である。

2. 入出力パッド、論理素子、コンタクトなどの回路要素は、一つの格子点上に位置する。一つの格子点を占める回路要素は一つである。配線が、回路要素が占めている格子点に行き当っているとき、その配線はその回路要素に接続している。（配線は回路要素が占めている格子点を通過できない。）

3. 配線は、一方向の信号のみを伝える。

4. 組合せ回路モデルで述べたように、論理素子は、入次数はある定数以下で、（論理回路レベルで）出次数は制限がない。（論理素子が占めている格子点から出る配線は1本で、これが分岐する。）

5. 回路の領域は、回路を含む、各辺が格子線に沿った、最小の長方形である。入出力パッドは領域の外周線上にある。

VLSI 上での回路の面積とは、5. で定義した長方形の領域の大きさである。

2.2 配線遅延を考慮した回路モデル

配線長に依存する配線遅延を考慮に入れて回路の計算時間を評価するために、回路の論理素子間を接続する配線を規定しなければならない。本研究では、前述の VLSI モデルに従って回路のレイアウトを行い、論理素子の配置及び配線を行った上で、論理素子間の配線に対して配線長に依存する配線遅延を定義する。

配線を容量のみでモデル化すると、配線遅延は、配線長 l に対して $O(l)$ となる。ここで、論理素子の出力に、ドライブ能力の異なるいくつかのバッファを直列に挿入し、順次ドライブ能力を高めていくと、遅延時間は $O(\log l)$ となる。

配線を配線抵抗と配線容量でモデル化し、配線遅延の評価を Elmore 遅延モデル [2] により行うと、遅延時間は $O(l^2)$ となる。ここで、長い配線を分割してバッファをリピータとして挿入すると配線遅延は $O(l)$ となる。

上述の議論より、配線遅延は、配線長 l に対して $O(\log l)$ から $O(l)$ の範囲で定義することが妥当であると考えられる。本研究では、配線遅延が $O(\log l)$、$O(\sqrt{l})$、$O(l)$ の三つの場合を考える。

2.3 並列乗算のハードウェアアルゴリズムの評価

前節で定義した配線遅延を考慮した回路モデルにより、ハードウェアアルゴリズムの計算時間を評価する。ここでは、並列乗算の種々のハードウェアアルゴリズムを対象とする。

並列乗算は、部分積生成部、生成した部分積を桁上げ保存形で加え合わせる部分積累算部、桁上げ伝播加算により積を求める最終加算部により構成される。計算時間は部分積累算部が支配的となるため、ここでは、部分積累算部の評価を行う。

部分積の累算法により種々のハードウェアアルゴリズムが存在する。ここでは、配列型乗算法 [3]、バランス木を用いた乗算法 [4]、Wallace 木を用いた乗算法 [5]、4–2 加算木を用いた乗算法 [6] について、評価を行う。従来の評価基準である回路の素子数、面積、段数を表 2.3に示す。

表 1: 並列乗算の従来の評価基準による評価

	素子数	面積	段数
配列型	$O(n^2)$	$O(n^2)$	$O(n)$
バランス木	$O(n^2)$	$O(n^2)$	$O(\sqrt{n})$
Wallace 木	$O(n^2)$	$O(n^2 \log n)$	$O(\log n)$
4–2 加算木	$O(n^2)$	$O(n^2 \log n)$	$O(\log n)$

配列型乗算

配列型乗算では、部分積を桁上げ保存加算により順次累算する。配列型乗算器の VLSI モデル上でのレイアウトの概略を図 2.3に示す。図中の●は部分積、□は桁上げ保存加算器（CSA）、CPA は桁上げ伝播加算器である。各段の桁上げ保存加算器の出力が次の段の入力となっているため、桁上げ保存加算器間を接続する配線の長さは $O(1)$ となる。したがって、いずれの配線遅延モデルを用いても、回路の合計配線遅延は $O(n)$ となる。

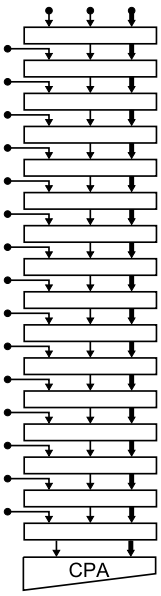
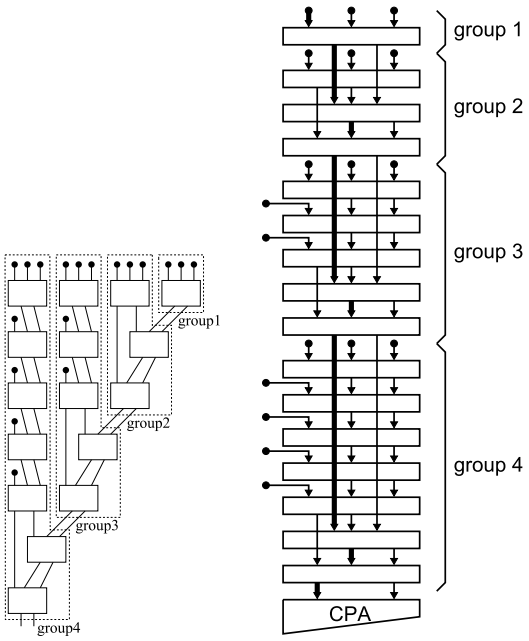


図 1: 配列型乗算器のレイアウト



(a) (b)

図 2: バランス木 (a) とそのレイアウト (b)

バランス木を用いた乗算

バランス木を用いた乗算では、部分積をいくつかのグループに分け、図 2.3(a) に示すように累算を並列に行う。バランス木を用いた乗算器の VLSI モデル上でのレイアウトの概略を図 2.3(b) に示す。図中の太線の経路が、最も配線遅延の大きな経路である。各桁上げ保存加算器の出力の配線の長さは、飛び越す加算器の個数に比例する。よって、太線の経路上の k 段目の加算器の出力の配線長は、k が奇数のとき $O(k)$、偶数のとき $O(1)$ となる。したがって、合計配線遅延は、配線遅延が $O(\log l)$、$O(\sqrt{l})$、$O(l)$ であるモデルに対して、それぞれ、$O(\sqrt{n}\log n)$、$O(n^{3/4})$、$O(n)$ となる。

Wallace 木を用いた乗算

Wallace 木を用いた乗算では、桁上げ保存加算を基本計算として、部分積を図 2.3(a) に示すように累算する。Wallace 木を用いた乗算器の VLSI モデル上でのレイアウトの概略を図 2.3(b) に示す。図中の太線の経路が、最も配線遅延の大きな経路である。この経路上の k 段目の加算器の出力の配線長は、$O((3/2)^k)$ となる。したがって、合計配線遅延は、配線遅延が $O(\log l)$、$O(\sqrt{l})$、$O(l)$ であるモデルに対して、それぞれ、$O((\log n)^2)$、$O(\sqrt{n})$、$O(n)$ となる。

4–2 加算木を用いた乗算器

4–2 加算木を用いた乗算では、桁上げ保存加算二段で実現される 4–2 加算を基本計算として、部分積を二分木状に累算する。4–2 加算木を用いた乗算器の VLSI モデル上でのレイアウトの概略を図 2.3に示す。図中の□は 4–2 加算器である。太線の経路が、最も配線遅延の大きな経路である。この経路上の k 段目の 4–2 加算器の出力の配線長は、$O(2^k)$ となる。したがって、合計配線遅延は、配線遅延が $O(\log l)$、$O(\sqrt{l})$、$O(l)$ であるモデルに対して、それぞれ、$O((\log n)^2)$、$O(\sqrt{n})$、$O(n)$ となる。

以上をまとめると、表 2.3のようになる。

表 2: 並列乗算の配線遅延を考慮した計算時間

	配線遅延モデル		
	$O(\log l)$	$O(\sqrt{l})$	$O(l)$
配列型	$O(n)$	$O(n)$	$O(n)$
バランス木	$O(\sqrt{n}\log n)$	$O(n^{\frac{3}{4}})$	$O(n)$
Wallace 木	$O((\log n)^2)$	$O(\sqrt{n})$	$O(n)$
4–2 加算木	$O((\log n)^2)$	$O(\sqrt{n})$	$O(n)$

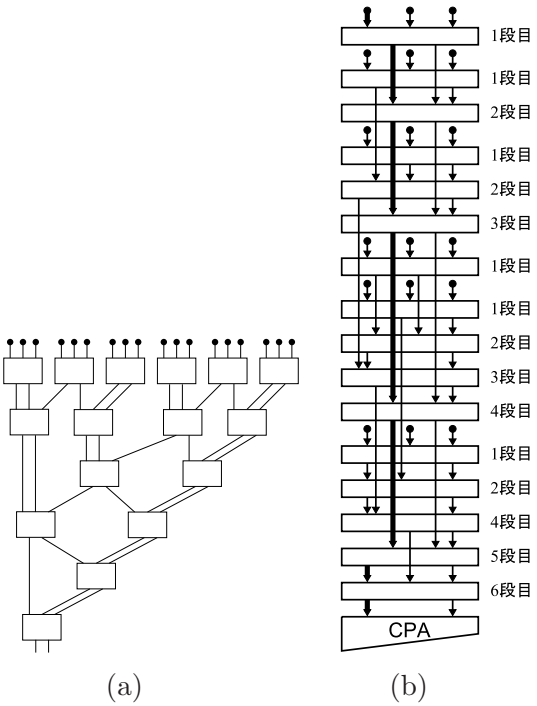


図 3: Wallace 木 (a) とそのレイアウト (b)

3 配線遅延を考慮した整数除算の高速ハードウェアアルゴリズム

整数除算は、基本算術演算の一つであり、いくつかの整数除算器が市販のＩＰライブラリに登録されている [7]。整数除算は、プロセッサにおいて多くのクロックサイクルを要する演算であり、サイクル数は演算数の値によって可変となる。

本研究では、整数除算のハードウェアアルゴリズムを提案する。アルゴリズムは減算シフト型の非回復型除算法に基づいている。専用回路によって計算を高速化するには、部分剰余の表現に桁上げ保存形や符号付きディジット表現 [8] を導入し、漸化式の計算を桁上げの伝搬なしに高速に行う手法がよく用いられる [9]。しかし、そのためには、除数を正規化する必要があり、したがって、大きな面積を要する先頭１検出器やバレルシフタが必要となる。

本研究で提案するハードウェアアルゴリズムでは、部分剰余を基数２の符号付きディジット表現（２進ＳＤ表現）で表し、漸化式の計算を桁上げの伝搬なしに高速に行い、かつ、除数の正規化を必要としない。部分剰余を符号と絶対値の組で保持し、絶対値を２進ＳＤ表現で表す。商の各桁は部分剰余の符号から決定する。漸化式の計算において、符号の計算と絶対値の計算を並行に行う。絶対値の計算は、前の部分剰余の符号が未知でも開始できる。

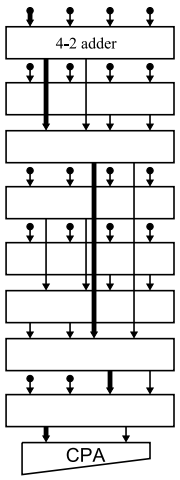


図 4: 4–2 加算木のレイアウト

アルゴリズムに基づく除算器の組合せ回路モデル上での段数は演算数の長さに比例し、素子数は演算数の長さの平方に比例する。また、ＶＬＳＩモデル上での面積は演算数の長さの平方に比例する。従来の減算シフト型アルゴリズムでは、商の桁選択回路の出力のファンアウトが演算数の長さに比例し、長い配線が存在したが、提案回路では、素子のファンアウトおよび配線長は、すべて、演算数の長さに依存しない定数以下である。

3.1 非回復型整数除算

ここでは、n ビット２の補数表示２進整数である被除数 X と除数 Y $(\neq 0)$ から、$X = Y \times Z + R$ を満す、n ビット２の補数表示２進整数の商 Z と剰余 R を求めるものとする。ただし、R は X と同じ符号になるようにする。

提案アルゴリズムは、以下に示す基数２の非回復型整数除算アルゴリズムに基づく。

[Algorithm IDIV]

$D := |Y|;\ \ R_n := X;$

for $j := n-1$ downto 0 do

  if $R_{j+1} = 0$ then do

    $Q := [q_{n-1}q_{n-2}\cdots q_{j+1}0\cdots 0];\ \ R := 0;$

    go to label;

endif;
if $R_{j+1} < 0$ then $q_j := -1$ else $q_j := 1$ endif;
$R_j := R_{j+1} - q_j \cdot 2^j \cdot D$;
endfor;
$Q := [q_{n-1}q_{n-2}\cdots q_0]$;
if $X > 0$ and $R_0 < 0$
then $R := R_0 + D$; $Q := Q - 1$;
elseif $X < 0$ and $R_0 > 0$
then $R := R_0 - D$; $Q := Q + 1$;
else $R := R_0$;
endif;
label: if $Y < 0$ then $Z := -Q$ else $Z := Q$ endif;
□

Q は、$R_{j+1} = 0$ となるまで、-1 と $+1$ の系列として求まる。Q の２の補数表示２進表現 $P = [p_{n-1}p_{n-2}\cdots p_0]$ $(p_j \in \{0,1\})$ は、$R_{j+1} = 0$ となる j を j_0 とする（そのような j がない場合は $j_0 = -1$ とする）と、j_0+2 以上の j については q_{j-1} が -1 なら p_j を 0、$+1$ なら 1 とし、p_{j_0+1} を 1 とし、j_0 以下の j については p_j を 0 とすることにより求まる。

3.2 ２進ＳＤ整数の符号と絶対値の計算

２進ＳＤ表現は桁集合が $\{-1,0,1\}$ の２進表現である。提案する除算アルゴリズムでは、２進ＳＤ整数の符号と絶対値の計算が必要である。２進ＳＤ整数の符号は、その中の最上位の非零桁の符号である。また、絶対値（の２進ＳＤ表現）は、その数が負である場合、各桁の符号を反転することにより求まる。

以下に、２進ＳＤ整数 $A = [a_{n-1}a_{n-2}\cdots a_0]$ の符号と絶対値を計算するアルゴリズム [10] を示す。得られる $sgn(A)$ は、A が正なら 1、零なら 0、負なら -1 である。$abs(A)$ は非負の２進ＳＤ整数である。

[Procedure SGNABS]
$s_n := 0$;
for $i := n-1$ downto 0 do
if $s_{i+1} = 0$ and $a_i = -1$ then $s_i := -1$;
elseif $s_{i+1} = 0$ and $a_i = 1$ then $s_i := 1$;
else $s_i := s_{i+1}$;
endif;
if $s_i = -1$ and $a_i \neq 0$ then $b_i := -a_i$;
else $b_i := a_i$ endif;
endfor;
$sgn(A) := s_0$;
$abs(A) := [b_{n-1}b_{n-2}\cdots b_0]$; □

3.3 整数除算のハードウェアアルゴリズム

提案アルゴリズムは、[Algorithm IDIV] に基づく。構造の規則性を増すために、途中で $R_{j+1} = 0$ となっても $q_j = 1$ として計算を続ける。[10] と同様に、数列 $\hat{R}_j = sign(R_{j+1}) \cdot R_j$ を導入し、２進ＳＤ表現で表す。定義より、$|\hat{R}_j| = |R_j|$、
$sign(\hat{R}_j) = sign(R_{j+1}) \cdot sign(R_j)$ $(R_j \neq 0)$、
$sign(\hat{R}_j) = +1$ $(R_j = 0)$
が成り立つ。よって、元の R_j の符号は、
$sign(R_j) = sign(R_{j+1}) \cdot sign(\hat{R}_j)$ $(R_j \neq 0)$、
$sign(R_j) = +1$ $(R_j = 0)$
となる。$sign(R_{j+1}) = -1$ のとき $R_j = R_{j+1} + 2^j \cdot D$ 他のとき $R_j = R_{j+1} - 2^j \cdot D$ であるから、$\hat{R}_j = |\hat{R}_{j+1}| - 2^j \cdot D$ である。$|\hat{R}_{j+1}|$ は、$\hat{R}_{j+1}$ から、[Procedure SGNABS] により求まる。したがって、$\hat{R}_j$ は $sign(R_{j+1})$ と独立に計算できる。すなわち、$\hat{R}_j$ と $|\hat{R}_j|$ の計算を $R_{j+1}(\hat{R}_{j+1})$ の符号がわかる前に開始できる。

右にハードウェアアルゴリズムを示す。D は Y が正なら Y、負なら $\bar{Y}+1$ である。ここに、$\bar{Y}$ は Y のビット毎の反転である。補正の $+1$ ビットは $\hat{R}_j$ の計算における最下位への桁借り入力として扱う。P の -1 補正は、最下位ビットを 1 から 0 にすることにより容易に行える。一方、$+1$ 補正は、$p_j = 0$ である最小の j を j_1 とすると、j_1 以下の j について p_j をビット反転することにより行える。２進表現の剰余 R を得るために、２進ＳＤ／２進変換が必要である。

[Hardware Algorithm IDIV-HA]
if $y_{n-1} = 0$ then $D := Y$ else $D := \bar{Y} + 1$ endif;
$\hat{R}_n := X$;
if $x_{n-1} = 1$ then $sign(R_n) := -1$
else $sign(R_n) := +1$ endif;
for $j := n-1$ downto 0 do
if $sign(R_{j+1}) = -1$ then do
if $y_{n-1} = 0$ then $p_{j+1} := 0$

```
      else p_{j+1} := 1 endif;
   else do
    if y_{n-1} = 0 then p_{j+1} := 1
      else p_{j+1} := 0 endif;
  endif;
  R̂_j := |R̂_{j+1}| - 2^j · D;
  |R̂_j| := abs(R̂_j);
  if sgn(R̂'_j) = -1
      or (sgn(R̂'_j) = 0 and sign(R_{j+1}) = -1)
   then sign(R_j) := -sign(R_{j+1})
   else sign(R_j) := sign(R_{j+1}) endif;
endfor;
P := [p_{n-1}p_{n-2}···p_1 1];
if x_{n-1} = 0 then do
    if sign(R_0) = -1 then do
        R := |R̂*_0|;
        if y_{n-1} = 0 then Z := P - 1
          else Z := P + 1 endif;
      else R := |R̂_0|;  Z := P;
    endif;
else do
    if sgn(R̂*_0) = 0 then do
        R := -|R̂*_0|;
        if y_{n-1} = 0 then Z := P - 1
          else Z := P + 1 endif;
     elseif sgn(R̂_0) ≠ 0 and sign(R_0) = +1
        then do
        R := -|R̂*_0|;
        if y_{n-1} = 0 then Z := P + 1
          else Z := P - 1 endif;
      else R := -|R̂_0|;  Z := P;
    endif;
endif;                                              □
```

3.4 アルゴリズムの評価

前節で提案したハードウェアアルゴリズムに基づく整数除算器を設計し、回路設計支援ソフトウェアを用いて論理合成と配置配線を行い、スタンダードセル方式のＬＳＩ上に実現した場合の計算時間および面積を見積ることにより、アルゴリズムの評価を行った。

アルゴリズムに基づく整数除算器をいくつか設計した。これらを Verilog-HDL で記述し、Synopsys Design Compiler により論理合成を行った。また、基数 2 の回復型アルゴリズムに基づく整数除算器を同様に合成した。スタンダードセルのライブラリとして、東京大学大規模集積システム設計教育センター（ＶＤＥＣ）提供のローム 0.35μm CMOS の京大 3.3V 高速版を用いた。

表 3(a) と (b) に、32 ビット整数除算器のセル面積、遅延、クロックサイクル数、計算時間（遅延×サイクル数）の見積りを示す。表 (a) 中の「$k = 2$」および「$k = 4$」はそれぞれ、内部でクロックサイクル毎に商の 2 および 4 ビット分ずつ計算を行う順序回路実現を示し、「comb」は組合せ回路実現を示している。表 (b) 中の「div_seq」は順序回路の除算器で、「div (rpl)」と「div (cla)」はそれぞれ加算器として順次桁上げ加算器と桁上げ先見加算器を用いた組合せ回路の除算器である。

表 3: 32 ビット整数除算器の見積り

(a) 提案除算器

divider	area $[mm^2]$	delay $[ns]$	#cycle	time $[ns]$
$k = 2$	0.69	4.67	21	98.07
$k = 4$	0.91	7.74	11	85.14
comb.	4.33	77.54	1	77.54

(b) 基数 2 の回復型アルゴリズムに基づく除算器

divider	area $[mm^2]$	delay $[ns]$	#cycle	time $[ns]$
div_seq	0.28	8.20	16	131.20
	0.39	13.40	8	107.20
div (rpl)	0.65	246.20	1	246.20
div (cla)	1.49	65.42	1	65.42

4 配線遅延を考慮した整数開平の高速ハードウェアアルゴリズム

前述した整数除算のハードウェアアルゴリズムと同様の手法に基づき、整数開平のハードウェアアルゴリズムを提案する。提案アルゴリズムに基づく回路においては、論理素子のファンアウトおよび配線長は、すべて、演算数の長さに依存しない定数以下である。

4.1 非回復型整数開平

ここでは、n ビット符号なし２進整数である演算数 X から、$Z = \lfloor X \rfloor$ を満す、$\frac{n}{2}$ ビット符号なし２進整数の平方根 Z を求めるものとする。

提案アルゴリズムは、以下に示す基数２の非回復型整数開平アルゴリズムに基づく。

[Algorithm ISQRT]

$Z_{\frac{n}{2}} := 0;$
$R_{\frac{n}{2}-1} := X - 2^{n-2};$
for $j := \frac{n}{2} - 2$ downto 0 do
  if $R_{j+1} < 0$ then
    $z_{j+1} := 0;$
    $R_j := R_{j+1} + 2^{j+1} \cdot (Z_{j+2} + 2^{j+1}) - 2^{2j};$
  else $z_{j+1} := 1;$
    $R_j := R_{j+1} - 2^{j+1} \cdot (Z_{j+2} + 2^{j+1}) - 2^{2j};$
  endif;
  $Z_{j+1} := Z_{j+2} + 2^{j+1} z_{j+1};$
endfor;
if $R_0 < 0$ then $z_0 := 0$ else $z_0 := 1;$
$Z_0 := Z_1 + z_0;$
$Z := Z_0$ □

4.2 整数開平のハードウェアアルゴリズム

提案アルゴリズムは [Algorithm ISQRT] に基づく。数列 $\hat{R}_j = sign(R_{j+1}) \cdot R_j$ を導入し、$\hat{R}_j$ を冗長２進表現で表す。冗長２進表現は桁集合が $\{-1, 0, 1\}$ の２進表現である。ここに、$sign(R_{j+1})$ は、$R_{j+1} \geq 0$ なら $+1$、< 0 なら -1 である。定義より、$|\hat{R}_j| = |R_j|$、

$sign(\hat{R}_j) = sign(R_{j+1}) \cdot sign(R_j)$ $(R_j \neq 0)$、
$sign(\hat{R}_j) = +1$ $(R_j = 0)$

が成り立つ。よって、元の R_j の符号は、

$sign(R_j) = sign(R_{j+1}) \cdot sign(\hat{R}_j)$ $(R_j \neq 0)$、
$sign(R_j) = +1$ $(R_j = 0)$

となる。$sign(R_{j+1}) = -1$ のとき $R_j := R_{j+1} + 2^{j+1} \cdot (Z_{j+2} + 2^{j+1}) - 2^{2j}$、他のとき $R_j := R_{j+1} - 2^{j+1} \cdot (Z_{j+2} + 2^{j+1}) - 2^{2j}$ であるから、

$\hat{R}_j = |\hat{R}_{j+1}| - 2^j \cdot (2Z_{j+1} + 2^{j+1}\bar{z}_{j+1} + 2^j)$

である。この計算は、桁上げの伝搬なしに高速に行う。$|\hat{R}_{j+1}|$ は、$\hat{R}_{j+1}$ を最上位桁から順に調べることにより、上位桁から順に求めることができる。したがって、$\hat{R}_j$ は $sign(R_{j+1})$ と独立に計算できる。すなわち、$\hat{R}_j$ と $|\hat{R}_j|$ の計算を $R_{j+1}(\hat{R}_{j+1})$ の符号がわかる前に開始でき、漸化式の繰返し計算を高速に行うことができる。

以下に提案アルゴリズムを示す。

[VLSI Algorithm ISQRT-VA]

$\hat{R}_{\frac{n}{2}-1} := X - 2^{n-2};$
if $sgn(\hat{R}'_{\frac{n}{2}-1}) = -1$ then $z_{\frac{n}{2}-1} := 0$
  else $z_{\frac{n}{2}-1} := 1;$
$Z_{\frac{n}{2}-1} := 2^{\frac{n}{2}-1} z_{\frac{n}{2}-1};$
for $j := \frac{n}{2} - 2$ downto 0 do
  $\hat{R}_j := abs(\hat{R}_{j+1}) - 2^j \cdot (2Z_{j+1} + 2^{j+1}\bar{z}_{j+1} + 2^j);$
  if $sgn(\hat{R}'_j) = -1$ or
    $(sgn(\hat{R}'_j) = 0$ and $z_{j+1} = 0)$
   then $z_j := \bar{z}_{j+1}$ else $z_j := z_{j+1};$
  $Z_j := Z_{j+1} + 2^j z_j;$
endfor;
$Z := Z_0$ □

アルゴリズムに基づく開平回路の組合せ回路モデル上での段数は演算数の長さに比例し、素子数は演算数の長さの平方に比例する。また、VLSI モデル上での面積は演算数の長さの平方に比例する。

4.3 アルゴリズムの評価

表 4に、VDEC、ローム 0.35μm CMOS、京大 3.3V 高速版セルライブラリによる、32 ビット整数開平回路の遅延とセル面積の見積りを示す。

表 4: 32 ビット整数開平回路の見積り

	遅延 $[ns]$	面積 $[mm^2]$
提案法	28.29	0.487
従来法	46.78	0.261

5 多層配線に適した小面積な乗算のハードウェアアルゴリズム

スタンダードセル方式の VLSI においては、配線層数の増加により配線の複雑化による面積の増加量が小さくなり、セルの総面積が回路面積に直接反映されるようになってきた。本研究では、多倍長乗算のソフトウェア実装においてしばしば用いられる Karatsuba の方法 [11] に基づく、配線が複雑であるが素子数の少ない乗算器を設計し、小面積の並列乗算器が実現可能であることを示す。評価の結果、本研究の乗算器は配列型乗算器よりも高速かつ小面積であった。

5.1 Karatsuba の方法に基づく並列乗算器

N ビット符号なし 2 進整数の乗算について考える。被乗数 X と乗数 Y を上位 $k_H = \lceil N/2 \rceil$ ビットと下位 $k_L = \lfloor N/2 \rfloor$ ビットにわけ、それぞれ X_H, X_L, Y_H, Y_L と表す（$X = 2^{k_L} X_H + X_L, Y = 2^{k_L} Y_H + Y_L$）。これらを用いて積 XY を表すと、$XY = 2^{2k_L} X_H Y_H + 2^{k_L}(X_H Y_L + X_L Y_H) + X_L Y_L = 2^{2k_L} P_1 + 2^{k_L}(P_3 - P_1 - P_2) + P_2$ となる。ここに $P_1 = X_H Y_H$、$P_2 = X_L Y_L$、$P_3 = (X_H + X_L)(Y_H + Y_L)$ である。

Karatsuba の方法では、上式に従い P_1, P_2, P_3 を計算する約 $N/2$ ビットの乗算 3 回と、加算、減算、シフトで N ビットの乗算を行う。本研究ではアルゴリズムを再帰的に適用せず、一度だけ適用した乗算器を設計する。一般的な並列乗算器はおよそ N^2 個の全加算器が必要であるが、Karatsuba 乗算器では $N/2$ ビットの乗算 3 回におよそ $\frac{3}{4}N^2$ 個の FA ですむため、素子数が少なくなる。

本研究で設計する Karatsuba 乗算器は、P_1, P_2, P_3 の計算に配列型乗算器や Wallace 木による乗算器を用いる。ここで P_3 を求める際に桁上げ伝搬加算器を使用せず、P_3 の部分積と P_1, P_2 をともに足しあわせる。乗算器の中で利用する桁上げ伝搬加算器として、順次桁上げ加算器や桁上げ先見加算器を用いる。これらの乗算器や加算器の組み合わせによって、さまざまな特徴をもつ乗算器が構成できる。

5.2 Karatsuba 乗算器の高速化

本研究では、次の二つの高速化手法を導入する。

一つは、P_1 と P_2 の各ビットと、P_3 の部分積の各ビットの生成時間を考慮し、早く求まるビットから順に、全加算器や半加算器で足し合わせていくものである。全体の遅延がバランスされ、計算が高速化される。計算に必要な全加算器や半加算器の数は変らないので、素子数はほとんど増加しない。

もう一つは、P_1, P_2 を求める乗算器において、最終段の桁上げ伝搬加算器を除去し、P_1, P_2 を桁上げ保存形で出力するものである。これにより乗算器内での桁上げ伝搬がなくなり、高速化できる。ただし、P_1, P_2 は、二カ所で用いるので、後段の加算および減算のためのハードウェア量が増加する。

5.3 乗算器の評価

上述の方法を組み合わせたさまざまな 32 ビット符号なし整数 Karatsuba 乗算器を、Verilog HDL により設計し、ライブラリとして VDEC ROHM0.35μm を用いて論理合成を行った。また配列型乗算器や Wallace 木による乗算器を同様の条件で設計し、遅延、面積の比較を行った。結果を図 5に示す。Wallace 木の乗算器と比較すると遅延で劣るものの、配列型乗算器に対しては高速かつ小面積となっていることが確認できる。最も小さな面積の配列型乗算器と比較して、85% にまで小型化することができた。

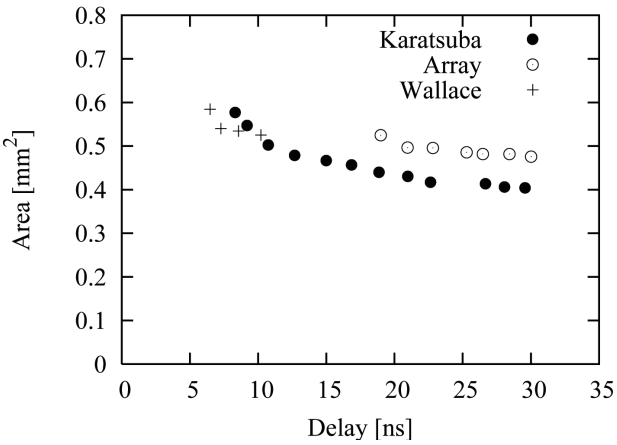


図 5: 32 ビット符号なし整数乗算器の遅延と面積

表 5: 乗算器のレイアウト結果

			3 層			5 層		
		セル面積 (μm^2)	回路面積 (μm^2)	総配線長 (μm)	増加率 (%)	回路面積 (μm^2)	総配線長 (μm)	増加率 (%)
32 bit	Array	99813	101333	108958	1.5	101333	116749	1.5
	Wallace	101245	102787	167065	1.5	102787	158746	1.5
	Karatsuba	84336	85623	102435	1.5	85623	100927	1.5
	Array+Booth	77447	81525	150266	5.3	78627	161320	1.5
	Wallace+Booth	78472	104632	207152	33.3	79668	186336	1.5
	Karatsuba+Booth	76011	89425	160781	17.6	77169	139270	1.5
64 bit	Array	404413	410573	423456	1.5	410573	446090	1.5
	Wallace	409054	415284	736326	1.5	415284	684370	1.5
	Karatsuba	320140	325019	411950	1.5	325019	390066	1.5
	Array+Booth	290543	363178	711159	25.0	294969	724458	1.5
	Wallace+Booth	294023	490043	1064182	66.7	298505	845112	1.5
	Karatsuba+Booth	253342	506686	1021491	100.0	257201	691701	1.5

6 配線層数の回路面積への影響

スタンダードセル方式の VLSI においては、配線層数の増加により配線の複雑化による面積の増加量が小さくなり、セルの総面積が回路面積に直接反映されるようになってきた。

本研究では、6 種類の乗算器について、配線層数の回路面積への影響を調べた。

6.1 乗算器の設計

前述の Karatsuba アルゴリズムに基づく乗算器（Karatsuba 乗算器）と配列型乗算器（Array 乗算器）、Wallace 木を用いた乗算器（Wallace 乗算器）、および、これらのそれぞれに 2 ビット Booth の方法を適用した乗算器（Karatsuba+Booth 乗算器、Array+Booth 乗算器、Wallace+Booth 乗算器）の計 6 種類の乗算器を、用いる配線層数を変えてレイアウトを行った。

Karatsuba 乗算器の部分積加算では、配列型乗算器の構成を用いた。Karatsuba+Booth 乗算器では、内部で用いる三つの乗算器のそれぞれに 2 ビット Booth の方法を適用し、部分積加算では [12] で提案した遅延を考慮して部分積をビット単位で加算していく手法を用いた。両乗算器とも、内部の加算器には順次桁上げ加算器を用いた。

6.2 回路面積の評価

乗算器は Verilog HDL を用いて記述し、ROHM 0.18μm プロセスを用いてレイアウトを行った。セルライブラリは東京大学大規模集積システム設計教育研究センター（ＶＤＥＣ）の ROHM 0.18μm プロセス用ライブラリを用い、レイアウトツールは Cadence 社の SOC Encounter を用いた。このプロセスでは配線層を 5 層利用できる。ただし、１層目には電源用の配線があるため、１層目は信号配線にはほとんど利用できない。

レイアウトは配線が可能な範囲で回路面積が最小になるように行った。レイアウトが可能かどうかは SOC Encouter による DRC(Design Rule Check) を行って判定した。DRC では配線の最小間隔や短絡、接続性などの幾何的なバイオレーションのみを考慮し、これらのバイオレーションがなければレイアウト可能と判断した。

配列型乗算器、Wallace 乗算器、Karatsuba 乗算器はほとんどがセル面積に対し 1.5%の面積増加率でレイアウトできている。1.5%の面積増加率はセルの配置効率によるものであり、配線量によ

るものではない。Wallace 乗算器は配列型乗算器や Karatsuba 乗算器よりも配線量が多い。一方、２ビット Booth の方法を用いた乗算器では、用いない乗算器と比較して面積増加率が大きくなっている。そのため、回路面積はセル面積に比べて大きく増加している。Wallace+Booth 乗算器と Karatsuba+Booth 乗算器の回路面積は Wallace 乗算器、Karatsuba 乗算器よりも大きくなっている。

配線に５層すべてを用いるレイアウトと、３層のみを用いるレイアウトを行った。３層のみの配線では、５層の配線層のうち、４層目と５層目の回路領域全体に配線禁止の領域を設定し、配線が行われないようにした。

乗算器のセル面積および配線層３層、５層の場合それぞれの回路面積を表 5.3に示す。面積増加率はセル面積に対する回路面積の増加率を示す。

配線層を５層用いた場合には、すべて面積増加率 1.5%でレイアウトができている。そのため、ほとんどの乗算器ではセル面積の大きさが回路面積に強く反映されており、セル面積が最も小さい Karatsuba+Booth 乗算器の回路面積が最も小さくなっている。

３層を用いたレイアウトに比べ、５層を用いたレイアウトでは回路面積の面積増加率が低下している。この結果より、配線層が増加することにより、配線が回路面積に及ぼす影響が小さくなっているといえる。また、どちらのビット数においても、配線層を５層用いた Karatsuba+Booth 乗算器の回路面積が最小となっている。Karatsuba+Booth 乗算器は素子数が少なくなることを重視した乗算器である。そのため、配線層数が多い場合には、素子数を重視した構成をとることにより、回路面積の小さな乗算器を構成することができるといえる。

7 算術演算のための減算シフト型ハードウェアアルゴリズムの自動合成

本研究では、算術演算のための減算シフト型ハードウェアアルゴリズムの自動合成の手法を提案する。まず、与えられた演算の算術式から、減算シフト型アルゴリズムのための漸化式を生成する。次に、漸化式および基数、桁集合等から桁選択関数を求める。

7.1 漸化式の生成

計算したい演算が算術式 f で与えられたときに、これに対する減算シフト型アルゴリズムのための漸化式を生成する手順について考える。f は、オペランドおよび加算、減算、乗算、除算（逆数計算も含む)、m 乗根（m は正の整数）の組合せで表されるとする。オペランドの個数は任意とする。f に対し、計算が加減算、シフト、一桁の乗算のみで行えるような漸化式を生成する。手順は以下の通りである。

まず、f に対し、残余 $W[j]$ の式を求める。f の算術式の各項について、$\frac{A}{B}$ があれば、式を $\frac{A}{B} = M$ として両辺に B を掛け、$A(B)^{\frac{1}{m}}$ があれば、式を $A(B)^{\frac{1}{m}} = M$ として両辺を m 乗じ、式を展開する。この操作を、式の各項に除算、m 乗根が存在しなくなるまで繰り返し、最後に、式を $g = 0$ の形に変形する。g には除算や m 乗根が存在しない。得られた $g = 0$ から、残余 $W[j]$ の式を、g を r^j 倍することにより得る。すなわち、

$$W[j] = r^j g \tag{7.1}$$

である。

次に、残余の漸化式を求める。漸化式は、式 (7.1) の j を $j+1$ に置き換え、中間結果の漸化式

$$S[j+1] = S[j] + s_{j+1} r^{-j-1} \tag{7.2}$$

を代入して展開することにより得られる。

$$W[j+1] = rW[j] + \cdots \tag{7.3}$$

の式で表される。ここで、得られた $W[j+1]$ の漸化式中に、オペランドと中間結果との乗算が存在するかを調べる。存在した場合、新しい中間変数で置き換える。さらに、新たに導入した変数に対する漸化式を導出する。乗算が漸化式中に存在しなくなるまでこの処理を繰り返す。

7.2 桁選択関数の生成

漸化式と基数 r、桁集合 $\{-a, \cdots, -1, 0, 1, \cdots, a\}$ 等から桁選択関数を生成する手順につい

て考える。$\frac{r}{2} \leq a \leq r-1$ とする。手順は以下の通りである。

まず、与えられた漸化式と基数、桁集合から、s_{j+1} として k を選んでもよい $rW[j]$ の範囲 $L_k[j] \leq rW[j] \leq U_k[j]$ を求める。ここで、残余 $W[j]$ の満たすべき範囲を $\underline{B}[j] \leq W[j] \leq \overline{B}[j]$ とする。$W[j+1]$ が $\underline{B}[j+1] \leq W[j+1] \leq \overline{B}[j+1]$ を満たすように s_{j+1} を $\{-a, \cdots, -1, 0, 1, \cdots, a\}$ の中から選ぶ必要がある。$s_{j+1} = k$ としたときの残余の漸化式を $W[j+1] = rW[j] + H_k[j]$ と表すと、

$$L_k[j] = \underline{B}[j+1] - H_k[j] \quad (7.4)$$

$$U_k[j] = \overline{B}[j+1] - H_k[j] \quad (7.5)$$

となる。$k = -a$ に対する下限及び $k = a$ に対する上限がそれぞれ $rW[j]$ の下限及び上限となるので、

$$r\underline{B}[j] = \underline{B}[j+1] - H_{-a}[j] \quad (7.6)$$

$$r\overline{B}[j] = \overline{B}[j+1] - H_a[j] \quad (7.7)$$

を得る。これらの漸化式から $\underline{B}[j]$、$\overline{B}[j]$ を求め、式 (7.4)、(7.5) より $L_k[j]$、$U_k[j]$ を定める。

次に、得られた $L_k[j]$ 及び $U_k[j]$ から、桁選択関数を決定する。桁選択関数は、$rW[j]$ と、$L_k[j]$ および $U_k[j]$ に含まれる変数に依存する。しかし、r^{-j} が含まれる項は、j が増えると急速に小さくなるので、桁選択関数を、あるとしてもごく少数の小さな j についてを除き、その項の中に存在する変数に依存しないようにすることができる。

桁選択関数への各入力の見積り値から s_{j+1} の値を決定する。桁選択関数を、選択の境界となる値の集合 $\{m_k | k \in \{-a+1, \cdots, -1, 0, 1, \cdots, a\}\}$ で表す。$m_k \leq \widehat{rW}[j] < m_{k+1}$ ならば、s_{j+1} として k が選択される。ここに、$\widehat{rW}[j]$ は $rW[j]$ の見積り値である。また、m_k は $rW[j]$ 以外の桁選択関数への入力に依存する。m_k は、s_{j+1} として $k-1$ が選ばれる $rW[j]$ の範囲と k が選ばれる範囲の重なりの範囲から、桁選択関数への入力となる変数の表現方法や範囲等を考慮して定めればよい。

桁選択関数の実現方法としては、比較器による実現、テーブルによる実現が考えられる。テーブルによる実現ではさらに、テーブルを ROM で作るのか論理回路で作るのかの選択がある。ここでは、桁選択関数を比較器で実現する場合に適した生成について述べてきたが、テーブルで実現する場合には、エントリに自由度が存在する。そのため、どのエントリを選ぶべきかという問題がある。また、桁選択関数の入力に必要なビット数のパラメータ（本研究では t と d）にも自由度があり、いくつかの可能な組合せの中からどれを選択すればよいかという問題がある。選択の基準としては、入力ビット数の和、有効なエントリ数、積項数等がある。また、桁選択関数の実現方法により、最適な基数や桁集合の選び方も異ってくると考えられる。

以上のことから、よりよい回路を実現するためには、桁選択関数の生成では、ある程度の自由度がある状態で出力し、次の段階で、回路実現方法を考慮して最適な桁選択関数を選ぶ、という二段階に分けて考えることが必要である。

7.3 桁選択関数生成ツール

以降、$L_k[j]$ 及び $U_k[j]$ に含まれる各変数（オペランド、中間結果、中間変数）の個数を l とし、各変数を $P_1[j], P_2[j], \cdots P_l[j]$ と表す。

本ツールでは、入力として、$L_k[j]$ 及び $U_k[j]$ の式、$W[j]$ の満たすべき範囲の下限 $\underline{B}[j]$ 及び上限 $\overline{B}[j]$ の式、$W[j]$ の表現方法、各 $P_i[j] (i = 1, 2, \cdots, l)$ の範囲の下限 $\underline{P_i}[j]$ 及び上限 $\overline{P_i}[j]$、各 $P_i[j]$ の表現方法、基数 r の値、桁集合の最大の要素 a の値、反復回数 j の初期値、$l+2$ 個の出力パラメータ（後述）の探索範囲の上限を与える。

$L_k[j]$、$U_k[j]$、$\underline{B}[j]$、$\overline{B}[j]$ の式は、基数 r の一般形で与える。$W[j]$ 及び各 $P_i[j]$ の表現方法については、非冗長表現か桁上げ保存形のどちらかを与える。$P_i[j]$ は全て符号なしで表現され、正の範囲をとるものとする。$P_i[j]$ の範囲が j や他の変数などに依存する場合も考えられるが、今回のツールではこれらに依存しない実数として $\underline{P_i}[j]$ 及び $\overline{P_i}[j]$ を与えるものとする。r は 2 のべきとし、a は $\frac{r}{2} \leq a \leq r-1$ を満たす正の整数であるとする。

このとき、桁選択関数の設定に必要な、「共通の桁選択関数が設定可能な反復回数の下限 J」と「その関数の入力に必要な $rW[j]$ の小数点以下のビット数 t、及び、$P_i[j]$ の小数点以下のビット数 $d_i (i = 1, 2, \cdots, l)$」を決定する。これらはトレードオフの関係にあり、いくつかの取り得る値の組

合せが存在する。ここでは、J の値と桁選択関数の入力ビット数の和の関係に着目し、「取り得る最小の J に対し、桁選択関数への入力ビット数の和が最小となる組合せ」、「取り得る最小の入力ビット数の和に対し、J の値が最小となる組合せ」を選ぶ。また、これらの組合せに対する、選択の境界となる値 $m_k(k=-a+1,\cdots,-1,0,1,\cdots,a)$ の下限及び上限を求める。

まず、$l+2$ 個のパラメータ J, t, d_i（$i=1,2,\cdots,l$）に対し、探索範囲の下限 $\underline{J}$、$\underline{t}$、$\underline{d_i}$ を設定する。$\underline{J}$ の値は反復回数 j の初期値である。$\underline{d_i}$ の値は $\underline{P_i}[j]$ 及び $\overline{P_i}[j]$ から適切に定める。$\underline{t}$ の値は $\underline{B}[j]$ 及び $\overline{B}[j]$ から、r 及び a の値などを考慮して適切に定める。

この探索範囲の中から、各パラメータにそれぞれある値を設定し、その値で桁選択関数が設定可能かどうか判定する処理 [ANALYSIS] を行う。この処理をパラメータの値を変化させて行う。その後、設定可能なパラメータの組合せの中から、前述の条件を満たす組合せを選ぶ。選ばれた各組合せに対し、パラメータの値及び選択の境界となる値 m_k の下限及び上限を出力する。

以降、処理 [ANALYSIS] について説明する。パラメータの値 J, t, d_i $(i=1,\cdots,l)$ を与え、以下の手順により桁選択関数が設定可能かどうか判定する。

[Step 1]

各変数 $P_i[j]$ について、$\underline{P_i}[j]$ 及び $\overline{P_i}[j]$ から、取り得る全ての見積り値 $\widehat{P_i}[j]$（2^{-d_i} の倍数）を求める。ここで、取り得る見積り値がただ一つとなる場合は、その変数が桁選択関数の入力として不要であるということを表す。

[Step 2]

各変数について、取り得る $\widehat{P_i}[j]$ の中から一つを選び（c_i とする）、その値が見積り値となる $P_i[j]$ の最小値 $\underline{P_i}[j]_{c_i}$ 及び最大値 $\overline{P_i}[j]_{c_i}$ を定める。$P_i[j]$ が非冗長表現ならば、

$$\begin{aligned}\underline{P_i}[j]_{c_i} &= \max\{c_i, \underline{P_i}[j]\}\\ \overline{P_i}[j]_{c_i} &= \min\{c_i+2^{-d_i}, \overline{P_i}[j]\}\end{aligned}$$

であり、$P_i[j]$ が桁上げ保存形ならば、

$$\begin{aligned}\underline{P_i}[j]_{c_i} &= \max\{c_i, \underline{P_i}[j]\}\\ \overline{P_i}[j]_{c_i} &= \min\{c_i+2^{-d_i+1}, \overline{P_i}[j]\}\end{aligned}$$

である。

[Step 3]

各 k（$k=-a+1,\cdots,a$）について、$\underline{P_i}[j]_{c_i} \le P_i[j] \le \overline{P_i}[j]_{c_i}$ 及び $j \ge J$ としたときの、$L_k[j]$ の最大値 $\max(L_k[j])$ 及び $U_{k-1}[j]$ の最小値 $\min(U_{k-1}[j])$ を求める。

[Step 4]

各 k について、選択の境界 $m_k(\widehat{P_i}[j])$ の下限 $\underline{m_k}(\widehat{P_i}[j])$ 及び上限 $\overline{m_k}(\widehat{P_i}[j])$ を求め、値を比較する。$m_k(\widehat{P_i}[j])$ は 2^{-t} の倍数となる必要があるので、残余の表現方法が非冗長表現ならば、

$$\begin{aligned}\underline{m_k} &= 2^{-t}\lceil 2^t \max(L_k[j])\rceil\\ \overline{m_k} &= 2^{-t}\lfloor 2^t \min(U_{k-1}[j])\rfloor\end{aligned}$$

であり、桁上げ保存形ならば、

$$\begin{aligned}\underline{m_k} &= 2^{-t}\lceil 2^t \max(L_k[j])\rceil\\ \overline{m_k} &= 2^{-t}\lfloor 2^t (\min(U_{k-1}[j]) - 2^{-t})\rfloor\end{aligned}$$

である。$\underline{m_k} > \overline{m_k}$ となる場合が一つでもあれば、与えたパラメータの値では桁選択関数を設定することができない。

[Step 5]

取り得る見積り値の全ての組合せに対して、[Step 2] から [Step 4] を行う。

[Step 6]

以上のステップにおいて、設定不可とならなければ、与えたパラメータの値で桁選択関数を設定できる。

上述のツールをＣ言語により実装した。

8 乗算剰余算のハードウェアアルゴリズム

乗算剰余算は、暗号処理に用いられる基本的な演算である。ＲＳＡ等の公開鍵暗号システムで必要な剰余系指数計算は、非常に大きな整数を法とする乗算剰余算の繰返しにより実現されている。

本研究では、乗算剰余算の並列処理を可能にした、二分乗算剰余算法と呼ぶ、新しい計算法を提案する。

8.1 ＫＴ剰余系

提案する計算法では、剰余系の乗算をＫＴ剰余という体系で行う。法 M は基数 r で n 桁の整数とし、$r^{n-1} < M < r^n$ および $\gcd(M, r) = 1$ を満たすものとする。U を $[0, M-1]$ の範囲の整数とするとき、U のイメージ（ＫＴ剰余）X を $X = U \cdot r^{\alpha n} \bmod M$ と定義する。ここに α は、$0 < \alpha < 1$ で、αn が整数となるような有理数である。

これらイメージの集合はＫＴ剰余系と呼ぶ剰余系を形成する。二つのイメージ X と Y が与えられたとき、ＫＴ剰余系における M を法とする乗算を、

$X * Y = X \cdot Y \cdot r^{-\alpha n} \bmod M$

と定義する。$r^{\alpha n}$ と M が互いに素であれば、$r^{-\alpha n} \bmod M$ は必ず存在する。暗号においては、M は奇数であるから、$r = 2^k$ とすれば、互いに素であるという条件が満たされる。

整数の元の表現から新しい剰余系への変換は、その整数と定数 $r^{\alpha n}$ との通常の乗算剰余算により行える。新しい剰余系から元の表現への逆変換は、イメージと定数 $r^{-\alpha n}$ との乗算剰余算で行え、モンゴメリアルゴリズムにより計算できる。

元の整数集合 Z_M とその上で定義された通常の乗算剰余算と、ＫＴ剰余系 Z'_M とその上で定義された演算 $*$ との間には、同型の関係が成り立つ。

8.2 二分乗算剰余算法

提案する計算法では、乗数を二つの部分に分け、並列に処理する。すなわち、ＫＴ剰余系におけるイメージ X と Y の M を法とする乗算は、乗数 Y を $Y = Y_H \cdot r^{\alpha n} + Y_L$ となるように Y_H と Y_L の二つの部分に分けると、以下の式により、効率よく計算できる。

$$X * Y = (X \cdot Y_H \bmod M + X \cdot Y_L \cdot r^{-\alpha n} \bmod M) \bmod M$$

$X \cdot Y_H \bmod M$ の項はインターリーブ法等、通常の乗算剰余算アルゴリズムにより計算でき、$X \cdot Y_L \cdot r^{-\alpha n} \bmod M$ の項はモンゴメリアルゴリズム [13] により計算できる。分割後の演算数 Y_H と Y_L は Y より短いので、$X \cdot Y_H \bmod M$ および $X \cdot Y_L \cdot r^{-\alpha n} \bmod M$ の計算は、それぞれ、分割しない演算数で通常の乗算剰余算あるいはモンゴメリ乗算を行う場合に比べ、高速に行える。

この二つの計算がほぼ同じ計算速度で行え、乗数を同じ大きさの二つの部分に分割すれば、分割しない演算数に対してそれぞれのアルゴリズムを単独で用いる場合に比べ、理論的に最大二倍の高速化が可能である。さらに、モンゴメリ法に比べ、元の剰余系とＫＴ剰余系との間の変換が最大二倍高速になり、変換に必要な定数の事前計算も不要になる。

提案手法では、中で行う通常の乗算剰余算およびモンゴメリ乗算の計算がどのようなアルゴリズムで実現されても構わない。さまざまなアルゴリズムを組合せることが可能であり、また、それぞれを高速化する方法が、独立に適用可能である。二つの計算の速度が異なる場合は、分割の割合 α を調節すればよい。例えば、後者が前者より２倍高速であれば、$\alpha \approx \frac{2}{3}$ とすればよい。

本手法は、ハードウェアばかりでなく、マルチプロッセ環境でソフトウェアで実現することもできる。

さらに、部分積の生成は独立に並列に行い、累算は統合して行うことも可能である。これをパイプライン化したハードウェアアルゴリズムも提案している [14]。

9 $GF(2^m)$ 上の乗算・逆元計算の統合ハードウェアアルゴリズム

ガロア体 GF(2^m) は、公開鍵暗号の一種である楕円曲線暗号において重要な役割を担っており、その上の乗算と逆元計算の高速化が望まれている。本研究では乗算と逆元計算のための統合ハードウェアアルゴリズムを提案する。

9.1 $GF(2^m)$ 上の乗算および逆元計算

GF(2) 上の m 次既約多項式 $G(x)$ によって定義される $GF(2^m)$ 上の元 $A(x)$ は、$m-1$ 次以下の GF(2) 上の多項式として表現される。GF(2^m) 上の乗算 “$\cdot$” は、$A(x) \cdot B(x) = A(x) \times B(x) \bmod G(x)$ と定義される。ここで、演算 “$\times$” は GF(2) 上の多項式乗算を表す。乗算アルゴリズムとして、MSB-first アルゴリズム等が知られている [15]。

GF(2^m) 上の零元を除く任意の元 $B(x)$ の乗法逆元 $B^{-1}(x)$ は、GF(2^m) 上の乗算を用いて $B(x) \cdot B^{-1}(x) = 1$ と定義される。逆元計算アルゴリズムとして、拡張ユークリッド法に基づくアルゴリズム等が知られている [16]。

9.2 統合ハードウェアアルゴリズム

MSB-first の乗算アルゴリズムと拡張ユークリッド法に基づく逆元計算アルゴリズムを統合した。二つのアルゴリズムで出現する計算の類似性に着目し、演算を区別する mode 信号を導入して、相違点には初期値の設定と mode 信号による簡単な制御で対処している。以下にアルゴリズムを示す。

[Algorithm GF-MULINV]

Input:

$mode \in \{0, 1\}$

$A(x), B(x)$: $m-1$ 次以下多項式

$G(x)$: m 次既約多項式

Output:

$B^{-1}(x)$ (if $mode = 0$)

$A(x) \cdot B(x)$ (if $mode = 1$)

Algorithm:

if $mode = 0$ then
  $R(x) := x \times B(x)$; $U(x) := x^3$; $V(x) := 0$;
  $\delta := -1$; $n := 2m - 1$;
else
  $R(x) := x \times B(x) + 1$;
  $U(x) := 0$; $V(x) := A(x)$;
  $\delta := m - 1$; $n := m + 1$;
endif
$S(x) := G(x)$;
for $i = 1$ to n
  if $r_m = 0$ then
    $R(x) := R(x) \times x$; $U(x) := U(x) \times x$;
  else
    if $\delta \geq 0$ then
      $R(x) := (R(x) + \overline{mode} \cdot S(x)) \times x$;
      $U(x) := (U(x) + V(x)) \times x$;
    else
      $\left\{ \begin{array}{l} R(x) := (R(x) + S(x)) \times x, \\ S(x) := mode \cdot S(x) + \overline{mode} \cdot R(x) \end{array} \right\}$;
      $\{U(x) := (U(x)+V(x)) \times x, V(x) := U(x)\}$;
      $\delta := -\delta$;
    endif
  endif
  if $u_{m+1} = 1$ then
    $U(x) := U(x) + x^{m+1}$
      $+ \sum_{k=1}^{m-1} mode \cdot s_t x^{k+1} + \overline{mode}$;
  endif
  $\delta := \delta - 1$;
endfor
output $V(x)/x$ as the result; □

9.3 ハードウェアアルゴリズムの評価

統合アルゴリズムに基づく複合回路を、$m = 163, 233, 283, 409, 571$ について、ＶＤＥＣ提供のローム $0.18\mu m$ CMOS 用セルライブラリを用いて合成した。また、単体の乗算回路と逆元計算回路を同様に合成した。表 6に $m = 571$ の場合の遅延とセル面積の見積りを示す。提案回路の面積は、乗算回路及び逆元計算回路を別々にもつ場合と比較して４割程度小さい。計算時間は同等である。

表 6: GF(2^m) 上の演算回路の見積り

	遅延 [ns]	面積 [mm^2]
乗算回路	1.1397	0.3489
逆元計算回路	1.2554	0.3754
複合回路 (提案回路)	1.3409	0.4343

10 $GF(2)$ 上の多項式乗算命令を用いた $GF(2^m)$ 上の逆元計算アルゴリズム

本研究では、二つの１ワードオペランドから、２ワードの積を求める $GF(2)$ 上の多項式乗算命令を用いた実装に適した $GF(2^m)$ 上の乗法逆元の高速計算アルゴリズムを提案する。

10.1 拡張ユークリッド法に基づく逆元計算アルゴリズム

提案アルゴリズムは、以下に示す Brunner らによる拡張ユークリッド法に基づく VLSI 実装向けアルゴリズム [17] に基づいている。

ここで、$\{O1, O2\}$ は二つの演算 $O1$ と $O2$ が並列に実行されることを示し、r_m と s_m はそれぞれ $R(x)$ と $S(x)$ の m 次の係数であるとする。なお、このアルゴリズムでにおいて $R(x)$ と $S(x)$ の値は、本来の値に x の冪乗が乗じた値になっており、$\deg^*(\cdot)$ は多項式の本来の次数の上限を表す。

[Algorithm Brunner]

Input:

$A(x)$: $m-1$ 次以下多項式

$G(x)$: m 次既約多項式

Output:

$A^{-1}(x)$

Algorithm:

```
S(x) := G(x); V(x) := 0;
R(x) := A(x); U(x) := 1;
δ := 0;
for i := 1 to 2m do
  if r_m = 0 then
    R(x) := x × R(x); U(x) := x × U(x);
    δ := δ + 1;
  else
    if s_m = 1 then
     S(x) := S(x) − R(x); V(x) := V(x) − U(x);
    endif
    S(x) := x × S(x);
    if δ = 0 then
      {R(x) := S(x), S(x) := R(x)};
      {U(x) := x × V(x), V(x) := U(x)};
      δ := 1;
     else
      U(x) := U/x; δ := δ − 1;
    endif
  endif
endfor
output U(x) as the result                □
```

10.2 提案アルゴリズム

提案アルゴリズムは、Brunner らのアルゴリズムにおける連続した反復での演算を行列で表現する。以下では、この行列の決定法を説明する。Brunner らのアルゴリズムでは、$R(x), S(x), U(x), V(x)$ は $R(x)$ と $S(x)$ の m 次の係数に従って計算される。同様に、連続する k 反復で $R(x)$, $S(x)$, $U(x), V(x)$ に対して実行される演算は、$R(x)$ と $S(x)$ の m 次から $m-k+1$ 次までの係数によって決定する。この k 反復における計算を、以下で説明する行列 $H(x)$ を用いて以下の式のように一挙に計算することが可能である。

$$\begin{pmatrix} R(x) & U(x) \\ S(x) & V(x) \end{pmatrix} := H(x) \times \begin{pmatrix} R(x) & U(v) \\ S(x) & V(x) \end{pmatrix};$$

ここで、$H(x)$ は (2×2) 行列であり、各要素は次数が k 以下の GF(2) 上の多項式である。

例を用いて、どのように行列 $H(x)$ を得るかを説明する。例として、$m = 7$, $G(x) = x^7 + x^6 + x^3 + x + 1$, $A(x) = x^6 + x^4$, $k = 2$ の場合を考える。この場合、各変数の初期値は $R(x) = x^6 + x^4$, $S(x) = x^7 + x^6 + x^3 + x + 1$, $U(x) = x$, $V(x) = 0$ となる。Brunner らのアルゴリズムの最初の反復では、以下の計算が実行される。

$$R(x) := x \times R(x); \quad U(x) := x \times U(x);$$

この計算は、以下のように行列の形で表現できる。

$$\begin{pmatrix} R(x) & U(x) \\ S(x) & V(x) \end{pmatrix} := \begin{pmatrix} x & 0 \\ 0 & 1 \end{pmatrix} \times \begin{pmatrix} R(x) & U(x) \\ S(x) & V(x) \end{pmatrix};$$

同様に、2 回目の反復では

$$S(x) := x \times (S(x) + R(x));$$
$$V(x) := x \times (U(x) + V(x));$$

が実行され、この計算は以下の式で表現できる。

$$\begin{pmatrix} R(x) & U(x) \\ S(x) & V(x) \end{pmatrix} := \begin{pmatrix} 1 & 0 \\ x & x \end{pmatrix} \times \begin{pmatrix} R(x) & U(x) \\ S(x) & V(x) \end{pmatrix};$$

したがって、以上の 2 反復分の計算は、次の行列 $H(x)$ を用いて一挙に計算が可能である。

$$H(x) = \begin{pmatrix} x & 0 \\ x^2 & x \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ x & x \end{pmatrix} \times \begin{pmatrix} x & 0 \\ 0 & 1 \end{pmatrix}$$

以下に、提案アルゴリズムを示す。ここで、w はプロセッサのワード長であり、M は $\lceil (m+1)/w \rceil$ とする。$H(x)$ は $R(x)$ と $S(x)$ の上位ワードからシングルワードの演算のみで求める。アルゴリズム中で $U(x)$ と $V(x)$ は $2M$ ワード必要であるよ

表 7: MULGF2 命令を用いた実装における命令数比較

			従来アルゴリズム			提案アルゴリズム			提案/従来 [%]
m	w	M	#MULGF2	#XOR	Total	#MULGF2	#XOR	Total	
163	8	21	2551.990	5751.183	8303.173	4429.690	7435.794	11865.484	142.90
	16	11	1341.678	2900.299	4241.977	1776.748	2374.208	4150.956	99.45
	32	6	733.744	1475.552	2209.296	1092.291	939.702	2031.993	91.97
	64	3	437.356	781.916	1219.272	861.820	498.054	1359.874	111.53
233	8	30	5171.630	11782.952	16954.582	8498.470	14869.154	23367.624	137.82
	16	15	2686.084	5938.826	8624.910	2967.834	4339.090	7306.924	84.72
	32	8	1433.177	3005.143	4438.320	1608.673	1563.986	3172.659	71.48
	64	4	811.507	1550.121	2361.628	1189.397	772.566	1961.963	83.08
283	8	36	7606.275	17408.760	25015.035	12114.088	21526.934	33641.022	134.45
	16	18	3929.496	8770.623	12700.119	4046.623	6190.736	10237.359	80.61
	32	9	2069.663	4420.616	6490.279	1903.184	2005.679	3908.863	60.23
	64	5	1150.530	2265.673	3416.203	1519.226	1002.331	2521.557	73.81
409	8	52	15796.999	36401.858	52198.857	24358.389	44466.958	68825.347	131.85
	16	26	8090.381	18302.942	26393.323	7594.821	12405.278	20000.099	75.78
	32	13	4196.562	9195.477	13392.039	3167.822	3729.663	6897.485	51.50
	64	7	2262.231	4673.211	6935.442	2246.086	1664.397	3910.483	56.38
571	8	72	30726.338	71092.647	101818.985	45937.205	85235.844	131173.049	128.83
	16	36	15624.255	35652.145	51276.400	13611.722	23396.826	37008.548	72.17
	32	18	8019.095	17888.957	25908.052	5130.520	6652.412	11782.932	45.48
	64	9	4220.452	9012.505	13232.957	3055.222	2579.882	5635.104	42.58

うにみえるが、x の冪乗が乗じた値になっているため、実際には $M+1$ ワードで充分である。

(提案アルゴリズム)

$S(x) := G(x)$; $V(x) := 0$;
$R(x) := A(x)$; $U(x) := x$;
while $\deg(R(x)) > 0$ do
  if $\{R(x)$ の最上位ワードが $0\}$ then
    $R(x) := x^w \times R(x)$; $U(x) := x^w \times U(x)$;
  else
    $R(x), S(x)$ の最上位ワードから $H(x)$ を計算
$$\begin{pmatrix} R(x) & U(x) \\ S(x) & V(x) \end{pmatrix} := H(x) \times \begin{pmatrix} R(x) & U(v) \\ S(x) & V(x) \end{pmatrix};$$
  endif
endwhile
$$\text{output} \begin{cases} U(x)/x^{Mw} & if\ \deg(R(x)) = 0 \\ V(x)/x^{Mw} & otherwise \end{cases} \quad \square$$

10.3 評価

提案アルゴリズムを、MULGF2 命令及び XOR 命令の実行数によって従来アルゴリズムと比較した。ここで、MULGF2 命令のレイテンシはシングルサイクルであると仮定し、評価にはランダムな 1000 個の元の逆元を計算した場合の平均値を用いた。また、既約多項式には NIST 推奨の多項式 [18] を用いた。結果を表 7に示す。

プロセッサのワード長が 32 と 16 の場合、提案アルゴリズムはほぼ全ての m で高速であった。特に、プロセッサのワード長と m の値がそれぞれ 571 と 32 である場合、提案アルゴリズムは従来のアルゴリズム [19] の約半数の命令数で逆元計算が可能であった。

参考文献

[1] J. D. Ullman: "Computational Aspects of VLSI", Computer Science Press (1984)

[2] W. C. Elmore: 'The Transient Response of Damped Linear Networks with Particular Regard to Wideband Amplifiers,' Journal of Applied Physics, vol. 19, p. 55, 1947.

[3] C. R. Baugh and B. A. Wooley: 'A Two's Complement Parallel Array Multiplication Algorithm,' IEEE Trans. Computers, vol. C-22, no. 12, pp. 1045–1047, Dec. 1973.

[4] S. D. Pezaris: 'A 40-ns 17-Bit by 17-Bit Array Multiplier,' IEEE Trans. Computers, vol. C-20, no. 4, pp. 442–447, Apr. 1971.

[5] C. S. Wallace: 'A Suggestion for a Fast Multiplier,' IEEE Trans. Elec. Comput., vol. EC-13, no. 1, pp. 14–17, Feb. 1964.

[6] J. Vuillemin: 'A very fast multiplication algorithm for VLSI implementation,' INTEGRATION, the VLSI Journal, vol. 1, pp. 39–52, 1983.

[7] Synopsys Inc.: 'DesignWare Building Block IP,' 2003.

[8] A. Avizienis: 'Signed-Digit number representations for fast parallel arithmetic,' IRE Trans. Elec. Comput., vol. EC-10, no. 3, pp. 389-400, Sept. 1961.

[9] M. D. Ercegovac and T. Lang: "Division and Square Root – Digit-Recurrence Algorithms and Implementations," Kluwer Academic Publishers, 1994.

[10] H. Dawid and H. Meyr: 'The differential CORDIC algorithm: constant scale factor redundant implementation without correcting iterations,' IEEE Trans. Comput., vol. 45, no. 3, March 1996.

[11] A. Karatsuba and Yu. Ofman: 'Multiplication of multidigit numbers on automata,' Sov. Phys. Dokl., 7, pp.595-596, 1963.

[12] 柴岡雅之，高木直史，高木一義：「Karatsuba アルゴリズムに基づく小面積並列乗算器」，2005 年電子情報通信学会総合大会，A-3-1，2005 年３月.

[13] P. L. Montgomery: 'Modular multiplication without trial division,' Mathematics of Computation, vol. 44, no. 170, pp. 519–521, April 1985.

[14] M. E. Kaihara, N. Takagi: 'Pipelined bipartite modular multiplication,' 信学会技報，VLD2005-67，2005 年 12 月

[15] S. K. Jain, L. Song and K. K. Parhi: 'Efficient Semisystolic Architectures for Finite-Field Arithmetic,' IEEE Trans. VLSI Systems, vol. 6, no. 1, pp. 101–113, March 1998.

[16] Z. Yan and D. V. Sarwate: 'New Systolic Architectures for Inversion and Division in GF(2^m),' IEEE Trans. Computers, vol. 52, no. 11, pp. 1514–1519, Nov. 2003.

[17] H. Brunner, A. Curiger and M. Hofstetter: On Computing Multiplicative Inverses in $GF(2^m)$,' IEEE Trans. Comput., vol. 42, no. 8, pp. 1010–1015, Aug. 1993.

[18] National Institute of Standards and Technology 'FIPS PUB 186-2: Digital Signature Standard (DSS),' pub-NIST, Jan. 2000. (URL http://www.itl.nist.gov/fipspubs/fip186-2.pdf)

[19] D. Hankerson, J. L. Hernandez and A. J. Menezes: 'Software Implementation of Elliptic Curve Cryptography over Binary Fields,' Proc. Cryptographic Hardware and Embedded Systems – CHES 2000, pp. 1–24, Aug. 2000.

研究業績一覧

著書

1. 高木直史:
"算術演算のＶＬＳＩアルゴリズム（並列処理シリーズ５）", コロナ社, 2005 年３月.
概要: さまざまな信号処理や画像処理、画像圧縮／伸長、グラフィクス、暗号処理などのための専用回路がＶＬＳＩ上に実現されるようになり、「ＶＬＳＩアルゴリズム」の設計が重要になってきている。「ＶＬＳＩアルゴリズム」は、ＶＬＳＩ上での実現を前提として、並列処理や規則正しい回路構造を意識して設計されたアルゴリズムである。本書では、さまざまな専用回路において中心的な役割を果

たす算術演算を主な対象として、種々のＶＬＳＩアルゴリズムを示し、その中で用いられている手法を紹介する。これらの手法は、さまざまな専用回路のためのＶＬＳＩアルゴリズムの設計に適用できる。また、算術演算の種々のＶＬＳＩアルゴリズムを理解することは、専用回路のためのＶＬＳＩアルゴリズムの設計において、最適な算術演算回路を選択するために必須である。

学術論文

1. M. .E. Kaihara, N. Takagi:
   "A Hardware Algorithm for Modular Multiplication/Division", IEEE Transactions on Computers, Vol. 54, No. 1, pp. 12–21, January, 2005.
   概要: A mixed radix-4/2 algorithm for modular multiplication/division suitable for VLSI implementation is proposed. The algorithm is based on Montgomery modular multiplication and on the extended Binary GCD algorithm for modular division. Both algorithms are modified and combined into the proposed algorithm so that almost all the hardware components are shared. The new algorithm carries out both calculations using simple operations such as shifts, additions, and subtractions. The radix-2 signed-digit representation is used to avoid carry propagation in all additions and subtractions. A modular multiplier/divider based on the algorithm performs an n-bit modular multiplication/division in $O(n)$ clock cycles where the length of the clock cycle is constant and independent of n. The modular multiplier/divider has a linear array structure with a bit-slice feature and can be implemented with much smaller hardware than that necessary to implement both multiplier and divider separately

2. N. Takagi, S. Kadowaki, K. Takagi:
   "A hardware algorithm for integer division", Proceedings of the 17th IEEE Symposium on Computer Arithmetic, pp. 140–146, June, 2005.
   概要: A hardware algorithm for integer division is proposed. It is based on the digit-recurrence, non-restoring division algorithm. Fast computation is achieved by the use of the radix-2 signed-digit representation. The algorithm does not require normalization of the divisor, and hence, does not require area-consuming leading one (or zero) detection nor shifts of variable-amount. Combinational (unfolded) implementation of the algorithm yields a regularly structured array divider, where pipelining is possible for increasing the throughput. Sequential implementation yields a compact divider.

3. M. E. Kaihara, N. Takagi:
   "Bipartite modular multiplication", Springer LNCS3659: Proceedings of 7th International Workshop on Cryptographic Hardware and Embedded Systems - CHES 2005, pp. 201–210, August, 2005.
   概要: This paper proposes a new fast method for calculating modular multiplication. The calculation is performed using a new representation of residue classes modulo M that enables the splitting of the multiplier into two parts. These two parts are then processed separately, in parallel, potentially doubling the calculation speed. The upper part and the lower part of the multiplier are processed using the interleaved modular multiplication algorithm and the Montgomery algorithm, respectively. Conversions back and forth between the original integer set and the new residue system can be performed at speeds up to twice that of the Montgomery method without the need for precomputed constants. This new method is suitable for both hardware implementation and software implementationin a multiprocessor environment. Although this paper focusing

on the application of the new method in the integer field, the technique used to speed up the calculation can also easily be adapted for operation in the binary extended field $GF(2^m)$.

4. M. E. Kaihara, N. Takagi:
"A hardware algorithm for modular multiplication/division based on the extended Euclidean algorithm", IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, Vol. E88-A, No. 12, pp. 3610–3617, December, 2005.
概要: A hardware algorithm for modular multiplication/division which performs modular division, Montgomery multiplication, and ordinary modular multiplication is proposed. The modular division in our algorithm is based on the extended Euclidean algorithm. We employ our newly proposed computation method that consists of processing the multiplier from the most significant digit first to calculate Montgomery multiplication. Finally, the ordinary modular multiplication is based on shift-and-add multiplication. Each of these three operations is carried out through the iteration of simple operations such as shifts and additions/subtractions. To avoid carry propagation in all additions and subtractions, the radix-2 signed-digit representation is employed. A modular multiplier/divider based on the algorithm has a linear array structure with a bit-slice feature and carries out n-bit modular multiplication/division in $O(n)$ clock cycles, where the length of the clock cycle is constant and independent of n. This multiplier/divider can be implemented using a hardware amount only slightly larger than that of the modular divider.

5. F. Kumazawa, N. Takagi:
"Hardware algorithm for computing reciprocal of Euclidean norm of a 3-D vector", IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, Vol. E89-A, No. 6, pp. 1799–1806, June, 2006.
概要: A hardware algorithm for computing the reciprocal of the Euclidean norm of a 3-dimensional (3-D) vector which appears frequently in 3-D computer graphics is proposed. It is based on a digit-recurrence algorithm for computing the Euclidean norm and an on-line division (on-line reciprocal computation) algorithm. These algorithms are modified, so that the reciprocal of the Euclidean norm is computed by performing on-line division where the divisor is the partial result of Euclidean norm computation. Division, square-rooting, and reciprocal square-root computation, which are important operations in 3-D graphics, can also be performed using a circuit based on the proposed algorithm.

6. N. Takagi, S. Kadowaki, K. Takagi:
"A hardware algorithm for integer division using the SD2 representation", IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, Vol. E89-A, No. 10, pp. 2874–2881, October, 2006.
概要: A hardware algorithm for integer division is proposed. It is based on the radix-2 non-restoring division algorithm. Fast computation is achieved by the use of the radix-2 signed-digit (SD2) representation. The algorithm does not require normalization of the divisor, and hence, does not require area-consuming leading one (or zero) detection nor shifts of variable-amount. Combinational (unfolded) implementation of the algorithm yields a regularly structured array divider, and sequential implementation yields compact dividers.

7. N. Takagi, K. Takagi:
"A VLSI algorithm for integer square-rooting", Proc. of 2006 International Sym-

posium on Intelligent Signal Processing and Communication Systems (ISPACS2006), pp. 626–629, December, 2006.
概要: A VLSI algorithm for integer square-rooting is proposed. It is based on the radix-2 non-restoring square-rooting algorithm. Fast computation is achieved by the use of the radix-2 signed-digit representation. Nonetheless, the algorithm does not require normalization of the operand. Combinational (unfolded) implementation of the algorithm yields a regularly structured array square-rooter. Its delay is proportional to n, the bit length of the operand, while that of conventional ones is at least proportional to $n \log n$.

8. K. Kobayashi, N. Takagi, K. Takagi: "An algorithm for inversion in GF(2^m) suitable for implementation using a polynomial multiply instruction on GF(2)", Proceedings of the 18th IEEE Symposium on Computer Arithmetic, pp. 105–112, June, 2007.
概要: An algorithm for inversion in GF(2^m) suitable for implementation using a polynomial multiply instruction on GF(2) is proposed. In the algorithm, operations corresponding to several contiguous iterations of the VLSI algorithm proposed by Brunner et al. is represented as a matrix. They are calculated as once through the matrix efficiently by means of a polynomial multiply instruction on GF(2). For example, in the case whre the word size of a propcessor and m are 32 and 571, respectively, the algorithm calculate inversion with about the half number of instructions of the conventional algorithm on the average.

9. K. Kobayashi, N. Takagi: "Design of a combined circuit for multiplication and inversion in GF(2^m)", Proceedings of the Workshop on Synthesis and System Integration of Mixed Information Technologies – SASIMI 2007, pp. 15–20, October, 2007.
概要: A combined circuit for multiplication and inversion in GF(2^m) is proposed. We combine the inversion algorithm proposed by Yan et al. that is based on the extended Euclid's algorithm and the MSB-first multiplication algorithm by focusing on the similarities between them so that multiplication and inversion can share almost all hardware components of the circuit. The area of the circuit is estimated to be approximately 40% smaller than the total area of an ordinary multiplication circuit and an ordinary inversion circuit.

10. M. E. Kaihara, N. Takagi: "Bipartite modular multiplication method", IEEE Transactions on Computers, Vol. 57, No. 2, pp. 157–164, February, 2008.
概要: This paper proposes a new fast method for calculating modular multiplication. The calculation is performed using a new representation of residues modulo M called KT-residue. This representation enables the multiplier to be split into two parts which can then be processed separately, in parallel, potentially doubling the calculation speed. The upper part of the split multiplier can be processed by calculating a multiplication modulo M between the multiplicand and this part of the split multiplier. The lower part of the split multiplier can be processed by calculating a multiplication modulo M between the multiplicand, this part of the split multiplier and the inverse of a constant R, where R $<$ M. Two different implementations based on this method are proposed. One uses an interleaved modular multiplier and an interleaved Montgomery multiplier. The other one generates the partial products for each part of the split multiplier separately, which are added and accumulated in a single

pipelined unit. A radix-4 version of this hardware implementation is presented as an example. The proposed method is also suitable for software implementation in a multiprocessor environment.

研究会等

1. カイハラ マルセロ，高木直史:
“拡張ユークリッド法に基づく剰余系乗除算回路”，電子情報通信学会技術研究報告, VLD2004-1, 1–6, 2004 年５月.
概要: ＶＬＳＩ実現向きの剰余系乗除算回路を提案する。除算は拡張ユークリッド法、乗算は通常の乗算アルゴリズムとモンゴメリ乗算に基づく。通常の乗算は二倍の操作と加減算の繰り返しで行う。モンゴメリ乗算は新しく提案した左シフト型の手法に基づく。これらの演算はビットシフトや加減算などの単純な操作の繰り返しで行う。基数 2 の符号付きディジット表現を用いることにより全ての加減算を桁上げ伝播なしに行う。提案する剰余系乗除算回路は、規則正しいビットスライス構造を持ち、n ビットの剰余系乗算、モンゴメリ乗算及び剰余系除算を $O(n)$ クロックサイクルで実行する。クロックサイクルの長さは n に依存しない。提案する剰余系乗除算回路は、モンゴメリ法に基づく乗算回路、通常の剰余系乗算回路と剰余系除算回路を別々に設計したものより少ないハードウェア量で実装できる。

2. 小林克希，高木直史，高木一義:
“GF(2^m) 上の逆元算出のための拡張ユークリッド法に基づくテーブルを用いたアルゴリズム”，電子情報通信学会技術研究報告, VLD2004-2, 7–12, 2004 年５月.
概要: $\mathbf{GF}(2^m)$ 上の逆元算出のための拡張ユークリッド法に基づくテーブルを用いたアルゴリズムを提案する。提案アルゴリズムでは、従来の拡張ユークリッド法に基づく逆元算出アルゴリズムの数回の反復における演算を行列で表現し、事前に計算してテーブルに保持しておく。これにより、提案アルゴリズムをハードウェア実装した際に、従来よりも XOR の段数を削減することが可能となる。また、$w \times (w+1)$ ビットの $\mathbf{GF}(2^m)$ 上の乗算器を用いれば、ソフトウェア実装した際に多倍長演算の処理時間を削減し、高速化することが可能である。

3. 門脇俊介，高木直史，高木一義:
“冗長２進数の絶対値計算を用いた整数除算回路”，電子情報通信学会技術研究報告, VLD2004-3, 13–18, 2004 年５月.
概要: 非回復型の除算アルゴリズムに基づいた新しい整数除算回路を提案する。各部分剰余を冗長２進数で表し、その絶対値を計算するとともに、その部分剰余の符号も求めていくものである。その符号を用いて商の各桁を最上行桁から順に決定していく。n ビットの被除数、除数から n ビットの商と剰余を $O(n)$ の時間で計算できる。32 ビットの被除数、除数から 32 ビットの商と剰余を出力する整数除算器として、提案するアルゴリズムに基づく回路を組合せ回路として実現した。順次桁上げ加算器を用いて実現した非回復型整数除算器と比較すると、遅延時間は約 65%の減少であった。また、桁上げ先見加算器を用いて実現した非回復型整数除算器と比較すると、遅延時間は約 25%の減少であった。

4. 熊澤文雄，高木直史:
“算術演算のための減算シフト型ハードウェアアルゴリズムの自動合成”，電子情報通信学会技術研究報告, VLD2004-90，ICD2004-176，DC2004-76, 175–178, 2004 年 12 月.
概要: プロセッサ等の開発において、より高い演算性能を実現するためには、頻出する算術演算のための専用回路を開発することが有効であると考えられる。本報告では、算術演算の専用回路のための減算シフト型ハードウェアアルゴリズムの自動合成について考察する。計算したい演算が算術式として与えられたとき、加減算、シフト、一桁の乗算の繰り返しで計算が行えるような漸化式を生成する。漸化式、基数、桁集合、残余の表現方法などが与えられたとき、桁選択関数を生成する。

5. 熊澤文雄，高木直史:

“算術演算のための減算シフト型ハードウェアアルゴリズムの自動合成ツールの構築”, 情報処理学会 DA シンポジウム 2005 論文集, pp. 231–236, 2005 年 8 月.
概要: プロセッサ等の開発において、より高い演算性能を実現するためには、頻出する算術演算のための専用回路を開発することが有効であると考えられる。本報告では、以前示した考察に基づく、算術演算の専用回路のための減算シフト型ハードウェアアルゴリズムの自動合成ツールについて述べる。本報告で説明するツールは、減算シフト型アルゴリズムのための桁選択関数の生成を支援する。漸化式から導出される選択領域、基数、桁集合などを与え、アルゴリズムの性能の見積りを基に桁選択関数の生成に必要なさまざまなパラメータを決定する。

6. カイハラ マルセロ，高木直史:
“Pipelined bipartite modular multiplication”, 電子情報通信学会技術研究報告, VLD2005-67, pp. 37–42, 2005 年 12 月.
概要: We propose a fast hardware algorithm for calculating modular multiplication. The algorithm is based on the recently proposed bipartite modular multiplication method. The calculation of modular multiplication is performed in the KT-residue system which enables the splitting of the multiplier into two parts which can then be processed in parallel. In the hardware algorithm that we propose, the addition of the partial products to the intermediate accumulated product is pipelined in order to reduce the critical path delay. A radix-4 version of the hardware algorithm is given and its hardware implementation is discussed.

7. 川島裕崇，高木直史，高木一義:
“配線層数の乗算器の回路面積への影響について”, 電子情報通信学会技術研究報告, VLD2006-141, pp. 7–12, 2007 年 3 月.
概要: 近年の VLSI 製造技術の進歩により、配線に用いることができるメタル層数が増加している。使用できる配線層数が増えると、セルの上を通過することができる配線の量も増加する。そのため、配線のための面積は小さくなり、より小さい回路面積で回路が構成できる。乗算は基本的な算術演算のひとつであり、多数の特定用途向けの集積回路 (ASIC, Application Specific Integrated Circuit) で乗算器が搭載されている。ASIC において乗算器は大きな面積を占めるため、乗算器の面積は ASIC の面積や製造コストに大きな影響を与える。本稿では複数の構成法を用いて乗算器を構成し、配線層数が回路面積に及ぼす影響を確認した。

8. 小林克希，高木直史:
“拡張ユークリッド法に基づく GF(2^m) 上の乗算・逆元計算のための複合回路”, 電子情報通信学会技術研究報告, VLD2006-142, pp. 13–18, 2007 年 3 月.
概要: GF(2^m) 上の乗算及び逆元計算のための複合回路を提案する。提案回路は、これまでに提案されている複合回路と異なり、構成が GF(2^m) を定義する既約多項式に依存せず、また、入出力される多項式の係数の順序を反転する必要がない。提案回路において、逆元計算は拡張ユークリッド法に、乗算は MSB-first アルゴリズムに基づいており、それぞれのアルゴリズムの類似性に着目して回路の大部分を共用できるように複合した。複合には乗算と逆元計算で剰余多項式の次数が同一である必要があるため、乗算の場合の剰余多項式を変形して次数を揃え、他の変数もその変形に合わせた。提案回路を論理合成して回路の規模を見積もったところ、面積はそれぞれの回路を別々に持つよりも 4 割程度小さかった。

学会大会等

1. 高木直史，高木一義:
“冗長ＳＤ／非冗長２進のオンザフライ変換の新手法”, 情報処理学会第 67 回全国大会, 1, 13–14(6B-1), 2005 年 3 月.

2. 柴岡雅之，高木直史，高木一義:
“Karatsuba アルゴリズムに基づく小面積並

列乗算器”, 2005 年電子情報通信学会総合大会, 基礎・境界, 66(A-3-1), 2005 年 3 月.

3. 小林克希，高木直史，高木一義:
“反復回数を削減した拡張ユークリッド法に基づく GF(2^m) 上の逆元算出アルゴリズム”, 2005 年電子情報通信学会総合大会, 基礎・境界, 67(A-3-2), 2005 年 3 月.

4. 高木直史:
“テーブルサイズを削減した逆数の区分二次近似法”, 2005 年電子情報通信学会総合大会, 基礎・境界, 68(A-3-3), 2005 年 3 月.

5. 小林克希，高木直史，高木一義:
“GF(2^m) 上の乗除算のためのハードウェアアルゴリズム”, 2006 年電子情報通信学会基礎・境界ソサイエティ大会, 51(A-3-7), 2006 年 9 月.

6. 川島裕崇，高木直史，高木一義:
“Karatsuba 乗算器の設計と評価”, 2006 年電子情報通信学会基礎・境界ソサイエティ大会, 52(A-3-8), 2006 年 9 月.

文部科学省科学研究費補助金 特定領域研究 (課題番号 16092214)

A06: 幾何的配置問題の計算量限界と近似可能性の研究

建築におけるトラス構造物の最適設計問題や、都市工学における通路、道路、鉄道などの配置問題および道路網などにおける避難問題、配送問題をユークリッド幾何構造の上での数理的な問題として定式化し、個々の問題の計算量限界を明らかにするとともに、実用的に解ける近似アルゴリズムの開発をおこなった。

研究組織

研究代表者： 加藤 直樹　京都大学 大学院 工学研究科
研究分担者： 田村 明久　慶應義塾大学 理工学部 (平成 18–19 年度)
大崎 純　京都大学 大学院 工学研究科
瀧澤 重志　京都大学 大学院 工学研究科

交付決定額 (配分額)

平成 16 年度	2,700,000 円
平成 17 年度	3,600,000 円
平成 18 年度	3,600,000 円
平成 19 年度	3,100,000 円
合　計	13,000,000 円

研究成果の概要

- 学術誌
  J. of Comb. Opt. (2005 年), Structural and Multidisciplinary Opt. (2005 年), Theoretical Comp. Science (2004 年, 2007 年), Graphs and Combinatorics (2007 年), Math. of Operations Res. (2007 年), Int. J. Solids and Struct. (2006 年), Int. J. of Comp. Geometry (2006 年), IEICE Trans. on Fundamentals (2006 年), Discrete Applied Math. (2006 年), J. Int. Assoc. for Shell and Spatial Struct.(2006 年), Comput. Geometry: Theory and Applications (2005 年) など

- 国際会議
  AAIM (香港 (2006 年), Portland(2007 年)), COCOON (Jeju Island, 2004 年), ISAAC (香港 (2004 年), 仙台 (2007 年)), CCCG (Windsor, 2005 年), CJK-OSM3 (金沢, 2004 年), CJK-OSM4 (昆明, 2006 年), IASS (Bucharest, 2005 年), KyotoCGGT2007 (京都, 2007 年), ACM/SIAM SODA (San Francisco, 2008 年) など

- 本研究班が主体となって開催した国際会議
  2007 年 6 月 11 日から 15 日まで京都大学百周年時計台記念館において、秋山仁氏および Vasek Chvatal 氏の還暦を記念して離散・計算幾何学・グラフ理論の国際会議 (Kyoto Conference on Computational Geometry and Graph Theory) を開催した.

1　はじめに

建築におけるトラス構造物の最適設計問題や、都市工学における通路、道路、鉄道などの配置問題および道路網などにおける避難問題、配送問題は、数理的にはユークリッド幾何構造の上での最適化問題として特徴付けられる。また、これらの問題の多くは、実現可能な解集合を考慮すると本質的に離散構造を有する。たとえば、建築構造物の設計では、実現できる部材長や部材断面積は、あらかじめ与えられたサイズ集合から選択されるのが望ましい。通路、道路の配置問題、配送問題、避難問題でも対象となる幾何構造が限定されるケースが多い。また、現実に現れる問題は、最適化基準が明示的ではないばあいは、最適解を一つ求めるより、実行可能解を列挙する方が望ましい。このような観点から、本研究では、これらの問題を幾何的配置問題、幾何的最適化問題、ネットワーク最適化問題としてモデル化し、個々の問題の計算量限界を明らかにするとともに、離散組合せ手法や計算幾何学を用いて実用的に解ける近似アルゴリズムの開発をおこなった。平成 16 年度から 19 年度において得られた主要な成果の概略は以下の通りである。

(a) 建築構造物などの設計問題に関して得られた成果

1. 存在することのできる節点と部材の位置が与えられたトラスに対し，静的外力の下での応力と，局所的な安定性の制約の下で最適な部材トポロジーを求める問題を，混合整数計画法として定式化し，分枝限定法で大域最適解を求めた（学術論文 [26]）。

2. 静定構造物の位相的な特徴はラーマングラフによって特徴付けられる。構造最適化への応用を考慮すると、辺が交差しないことが重要である。そこで、与えられた点集合に対して辺が交差しないラーマングラフをすべて列挙する問題を考察した。本問題で考えるグラフのクラスは pseudo-triangulation を含み、より難しい問題である。この問題に対して、逆探索法を基礎にした効率のよい列挙アルゴリズムを開発した。また、あらかじめ指定された辺集合を必ず含むような無交差ラーマングラフをすべて列挙する問題も考察し、効率のよいアルゴリズムを提案した（学術論文 [26], 研究会等 [11,16], 学会等 [18,19,40]）。

3. 棒材とケーブルで構成されるテンセグリティ構造に対して，釣合い条件と非退化条件を満たす軸力密度ベクトルを，釣合い行列のスペクトル分解を用いた繰返し計算によって求める方法を提案した。さらに，釣合い状態の安定性を判別するための新しい手法を提案した。

4. スナップスルーを用いたバイステーブル・コンンプライアントメカニズムの形状とトポロジーを，非線形最適化によって求めるアルゴリズムを提案した。さらに，ラーマングラフ列挙法を用いることにより，平面グラフ構造を有するコンンプライアントメカニズムを効率よく求めることが可能であることを示した。

5. 異なる部材長の数を最小にするトラスの設計問題を、異なる辺長の数を最小化する二次元平面上の三角形メッシュ生成問題としてモデル化し、計算量限界、近似性能比を明らかにした。また、異なる節点のタイプの数を最小化するトラスの設計問題を、隣り合う辺間の角度の種類数を最小にする問題としてモデル化し、計算量限界、近似性能比を明らかにした（学術論文 [17,24], 研究会等 [6], 学会等 [6,7,8,17]）。

(b) 配送計画や避難計画に関して得られた成果

1. 配送問題では、積込み、配送の 2 種類の荷物を有する場合について、とくに取り扱うネットワークが木構造を有する場合について新しい近似アルゴリズムを開発した。車両の起点、終点は単一のデポに限定されている場合を扱う。効率のよい 2– 近似アルゴリズムを提案した（学術論文 [13]）。

2. 災害発生時における建物内部から安全な外部への最適避難経路や、津波や大地 震発生時の安全な建物・地域への最適避難経路を求める問題を考察した。実用的な観点から、格子状ネットワークに限定し、通路幅も 一定という制約や、避難経路が途中で枝分かれすることを許さない、または交差しないという制約下

での最適経路群を求める問題を考察した。とくに、格子状ネットワークにおける最適避難計画問題を動的なネットワークフロー問題として定式化し，以下の研究成果を得た。枝容量，枝移動時間が一定の格子状ネットワーク(サイズ $\sqrt{n} \times \sqrt{n}$) において避難完了時間最小化問題を解く $O(n \log n)$ 時間アルゴリズムを開発した（学術論文 [18], 研究会等 [13,14], 学会等 [21,22,23]）。

3. 上記の格子状ネットワークにおける最適避難計画問題をより広いクラスのネットワークに拡張した。辺容量が一定なグリッドネットワークを一般化した，単一パス長に加え，それぞれの点 v に対して最小 v-s カットが s に接続している辺のうちのその始点が v から到達可能なものによって構成されているネットワークに対する避難計画問題の避難完了時間とその避難を実現するフローを求める $O((k^2n^2 + kmn) \log n)$ 時間アルゴリズムを開発した．ただし，m は入力のネットワークの辺数，k はネットワークにおけるシンクまでの異なったパス長の種類数であるとする（学術論文 [27], 研究会等 [20], 学会等 [32,33,35]）。

4. 上記の問題を考察した際の副産物として、Edmonds による「有向グラフにおける arborescence のパッキング定理」を拡張することに成功した（学術論文 [31], 研究会等 [18], 学会等 [37,38]）。

5. これまでの避難計画問題に対して，現実の避難の際に考えるべき問題を制約を加えた問題に対する研究も行った．具体的には、避難時の混乱を避けるため，避難の途中の点における滞留を禁止する．避難時の指示を容易にするため，各点からシンクまでの避難経路を一つに定めるという制約である。これらの制約が問題をどのくらい複雑にしているかを考察した（研究会等 [19], 学会等 [36]）。

6. 平面上に存在する利用者間の交通量が与えられている時、総和が一定の長さの線分集合で表現される道路（または鉄道）を配置することにより増大する利用者の利便性を最大化する交通網を設計する問題を幾何的最適化問題としてモデル化し、計算量限界、近似性能比を明らかにした（研究会等 [22], 学会等 [41,42]）。

(c) その他、幾何的最適化問題に関する成果

1. 与えられた平面上の点集合に対して、それらを k 個の折れ点のある区分線形関数で近似する問題（誤差評価は L_1 および L_2-距離）を取り扱い、近似解法を提案した（学術論文 [20], 研究会等 [3]）。

2. 平面上に与えられた長方形集合を結ぶ橋のなかで迂回率を最小にするものを求める問題を取り扱った。ある条件の下ではその問題が多項式時間で得られることを示した（学術論文 [8]）。

以上の成果のうちのいくつかについて，詳しく紹介する．

2 テンセグリティの形状設計

図 1に示すような，互いに接触しない棒材（ストラット）と，張力の導入される連続なケーブルで構成されるピン接合構造物を，テンセグリティ構造物という。ここで，図 1において，太線がストラット，細線がケーブルである。張力導入によって安定化される構造物（ケーブルネット，膜構造物，テンセグリティ構造など）を総称して張力構造物という。張力構造物に対して，自重を無視し，外力の作用しない自己釣合い状態での形状を求めることを，釣合い形状解析という。このような構造は，張力が存在しない状態で不安定であり，釣合い状態において節点位置（形状）と張力が連成するため，設計者が期待するような釣合い形状を求めることは極めて困難である。本研究では，釣合い行列のスペクトル分解を用いて，釣合い状態での軸力と形状を効率よく求める方法を提案した（学術論文 [11], 研究会等 [8], 学会等 [14,16]）。その概要は以下のとおりである。

ケーブルとストラットを合わせた部材数を m, 節点数を n とする。部材 i の軸力を N_i, 長さを L_i としたとき，$q_i = N_i/L_i$ を軸力密度（force density）といい，そのベクトルを $\mathbf{q} = (q_1, \ldots, q_m)^\top$ とす

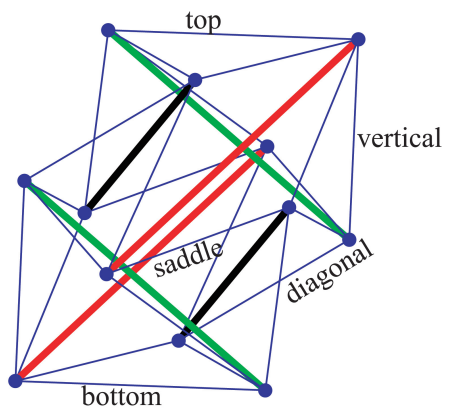


図 1: 2 層テンセグリティ構造物

る。各節点の x, y, z 座標をまとめたベクトルを，それぞれ $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$ とし，それらに関する釣合い行列を $\mathbf{E} \in \Re^{n \times n}$ で定義すると，構造物の自己釣合方程式は次のように書ける。

$$\mathbf{EX} = \mathbf{EY} = \mathbf{EZ} = \mathbf{0} \tag{2.1}$$

ここで，$\mathbf{E}$ は $\mathbf{q}$ の関数であり，実対称行列である。3 次元空間に存在するテンセグリティの釣合い形状解析において，立体的な構造を期待したにもかかわらず，平面，直線あるいは点に縮退した構造が得られることを,「退化」という。形状が退化しないためには，条件

$$\operatorname{rank} \mathbf{E} = n - 4 \tag{2.2}$$

を満たす必要がある。

行列 $\mathbf{E}$ の固有値解析を行い，固有ベクトルを並べた行列 $\mathbf{U}$ と，固有値を対角成分に持つ対角行列 $\mathbf{S}$ を用いて，$\mathbf{E}$ は次のようにスペクトル分解できる。

$$\mathbf{E} = \mathbf{USU}^{\top} \tag{2.3}$$

$\mathbf{E}$ の非零の固有値の数が $\mathbf{E}$ のランクと一致する。したがって，$\mathbf{S}$ の最小の h^* 個の固有値をゼロにした行列 $\bar{\mathbf{S}}$ を用いて $\bar{\mathbf{E}}$ を次のように構成する。

$$\bar{\mathbf{E}} = \mathbf{U}\bar{\mathbf{S}}\mathbf{U}^{\top} \tag{2.4}$$

以上より，次のような繰返し計算で $\mathbf{q}$ を求めることができる。

Step 0: トポロジー（部材と節点の接続関係）を設定し，$\mathbf{q}$ の初期値を与える。

Step 1: $\mathbf{q}$ から $\mathbf{E}$ 作成し，ランクを計算する。条件 (12.2) を満足すれば，アルゴリズムを終了する。

Step 2: $\mathbf{E}$ の固有値解析を行い，最小の h^* 個の固有値をゼロにし，新しい釣合い行列 $\mathbf{E}$ を (2.4) を用いて構成する。

Step 3: $\mathbf{E}$ を近似的に構成する $\mathbf{q}$ を最小 2 乗法で求めて Step 1 に戻る。

ところで，上記の方法で得られるテンセグリティは，釣合い条件は満たすが，安定であるとは限らない。本研究では，接線剛性を構成する線形剛性行列を幾何剛性行列の固有値などの特性量を用いて，安定性を判別するための理論と手法を提案した（学術論文 [22], 学会大会等 [16]）。

テンセグリティの接線剛性行列 $\mathbf{K} \in \Re^{n \times n}$ は，線形剛性行列 $\mathbf{K}^{\mathrm{E}}$ と幾何剛性行列 $\mathbf{K}^{\mathrm{G}}$ の和として $\mathbf{K} = \mathbf{K}^{\mathrm{E}} + \mathbf{K}^{\mathrm{G}}$ のように定義できる。$\mathbf{K}^{\mathrm{E}}$ と $\mathbf{K}^{\mathrm{G}}$ の r 次固有値，固有ベクトルを，それぞれ λ_r^{E}, $\mathbf{\Psi}_r^{\mathrm{E}}$ 及び λ_r^{G}, $\mathbf{\Psi}_r^{\mathrm{G}}$ で表す。λ_i^{E} 及び λ_i^{G} のなかで，$\lambda_i^{\mathrm{E}} = 0$ 及び $\lambda_i^{\mathrm{G}} < 0$ がそれぞれ成立する添字（次数）の集合を J^{E} 及び J^{G} で表す。また，テンセグリティは，一般に張力が存在しないとき剛体モード以外の不安定な変形モードが存在するので，J^{E} は空でないものとする。このとき，安定性に関して，次のような条件が成立する。

1. $\lambda_r > 0$, $(r = 1, \ldots, 3n)$ のとき安定である。
2. $\mathbf{\Psi}_i^{\mathrm{E}}$ $(i \in J^{\mathrm{E}})$ が $\mathbf{\Psi}_i^{\mathrm{G}}$ $(i \in J^{\mathrm{G}})$ の線形結合で表されるとき，不安定である。
3. 安定な構造において張力の符号を全て入れ替えると，不安定になる。
4. $\lambda_{\min}^{\mathrm{E}} = \min_{i \notin J^{\mathrm{E}}} \lambda_i^{\mathrm{E}}$, $\lambda_{\min}^{\mathrm{G}} = \min_{i \notin J^{\mathrm{E}}} \lambda_i^{\mathrm{G}}$ とする。次の 2 つの条件が成立するとき安定である。
   (a) 全てのベクトル $\mathbf{\Psi}_i^{\mathrm{E}}$ $(i \in J^{\mathrm{E}})$ が $\mathbf{\Psi}_i^{\mathrm{G}}$ $(i \notin J^{\mathrm{G}})$ の線形結合で表され，少なくとも 1 つの正の固有値 λ_j^{G} に関する固有モードの成分が 0 でない。
   (b) 張力は十分に小さく，$\lambda_{\min}^{\mathrm{E}} + \lambda_{\min}^{\mathrm{G}} > 0$ が成立する。

3 バイステーブル・コンプライアントメカニズムの設計

コンプライアントメカニズムとは，構造物の適切な位置に柔軟性を付加することにより構造一体で機能を実現するメカニズムである。医療部品やマイクロメカニクスの分野で積極的に利用されている。バイステーブル構造物は，2つの自己釣合い状態を持つ構造である。

本研究では，座屈現象を積極的に利用したバイステーブル・コンプライアントメカニズムを，最適化手法によって生成する方法を提案した（学術論文 [10]）。本構造は，建築構造や宇宙構造物での展開構造などで利用できる。

以下のような条件の下で，最適化問題を定式化する。

1. スナップスルーを利用して，小さい外力で大変形を発生させる。
2. 最終状態での外力は0であり，強制変位を与える点に微小な外力を与えて拘束すれば安定化できる。
3. 全部材体積を最小化。
4. 初期剛性が指定値以上。
5. 最終状態で拘束した構造物の剛性が指定値以上。
6. 最終状態で拘束を解除して微小な外乱を与えれば初期状態が回復される。

例えば，図2に示すような平面トラスを初期解（グランドストラクチャ）として，節点位置と断面積を変数として最適化し，断面が0となる部材を除去して最適な節点位置とトポロジーを求める。最適化アルゴリズムは以下のようにまとめられる。

Step 1 断面積ベクトル $\mathbf{A}$ 及び節点位置ベクトル $\mathbf{X}$ に初期値を与える。

Step 2 入力節点に強制変位を与えて釣合い経路を解析し，最終状態での変形を求める。

Step 3 目的関数（全部材体積）V 及び制約関数の設計変数ベクトル $(\mathbf{A}, \mathbf{X})$ に関する感度を求める。

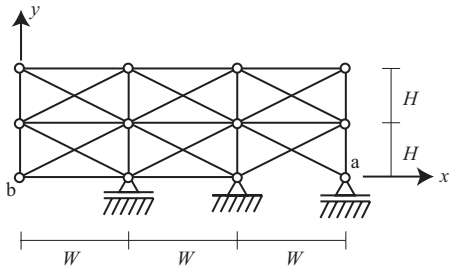
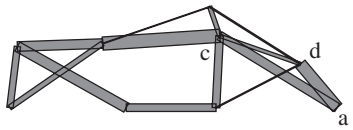
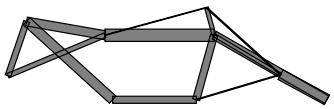


図 2: 平面トラスモデル



(a) 初期状態



(b) 最終状態

図 3: 最適解 Type-1 ($V = 1.2858 \times 10^{-4}$ m^3)

Step 4 最適化アルゴリズムにしたがって設計変数 $(\mathbf{A}, \mathbf{X})$ を変更する。

Step 5 収束条件を満たしていなければ Step 2 へ。

図2のトラスに対して，節点 a の x 軸方向に強制変位を与えたとき，節点 b が y 軸方向に指定量変位するようなコンプライアントメカニズムを求める。最適化によって得られた形状を図3に示す。この解での入力節点の変位と外力の関係は図4にようであり，スナップスルーが実現されている。最終状態は不安定であるが，入力節点の変位が最終状態を少し超えたところで止まるような障害物を与えることにより，障害物との接触によって最終状態は安定化される。また，初期解を変更して最適化すると，図5のような解が得られる。

上記のような手法によって，さまざまな形状のメカニズムを生成することができる。しかし，上記の Step 1–5 で解かれる最適化問題は極めて非

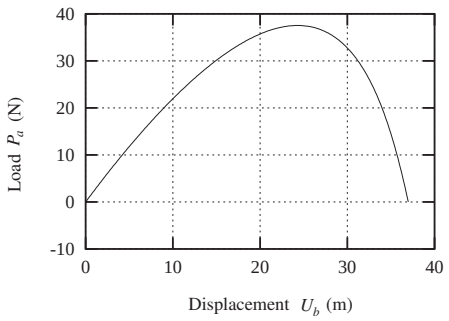
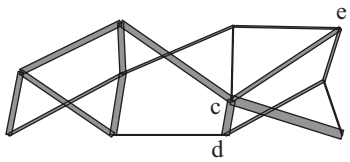
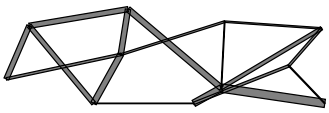


図 4: 最適解 Type-1 の入力変位 U_a と入力荷重 P_a の関係



(a) 初期状態



(b) 最終状態

図 5: 最適解 Type-2 ($V = 8.7875 \times 10^{-5}$ m^3)

線形性が強いので，実際的に意味のあるメカニズムを得るためには最適化問題を多数回解かなければならない。

ところで，最適化によって得られるコンプライアントメカニズムは，静定構造であり，ラーマングラフの性質を満たすことが知られている。そこで，ラーマングラフ列挙手法にによって得られた種々のトポロジーを初期解として，節点位置のみを変数として最適化を行う方法を提案した。本手法によると，最適化問題が極めて単純な形式となり，収束性が向上するので，さまざまなトポロジーを持つコンプライアントメカニズムを容易に求めることができる（研究会等 [12,24], 学会大会等 [29,44,46]）。また，平面グラフ構造を持つメカニズムが得られるので，実際に製作する際に極めて有効である。

例として，図 6のような節点・支点配置を持つ構造において，節点 B を下方に移動させたとき

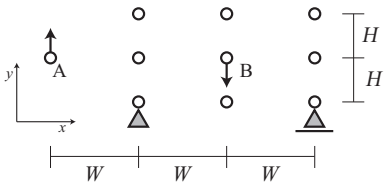
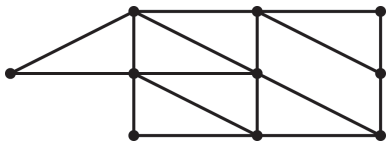
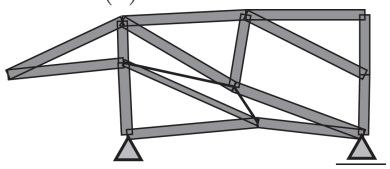


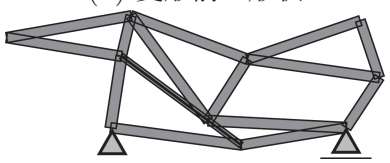
図 6: 節点，支点と荷重の配置



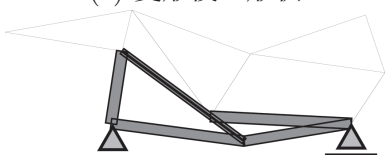
(a) トポロジー



(b) 変形前の形状



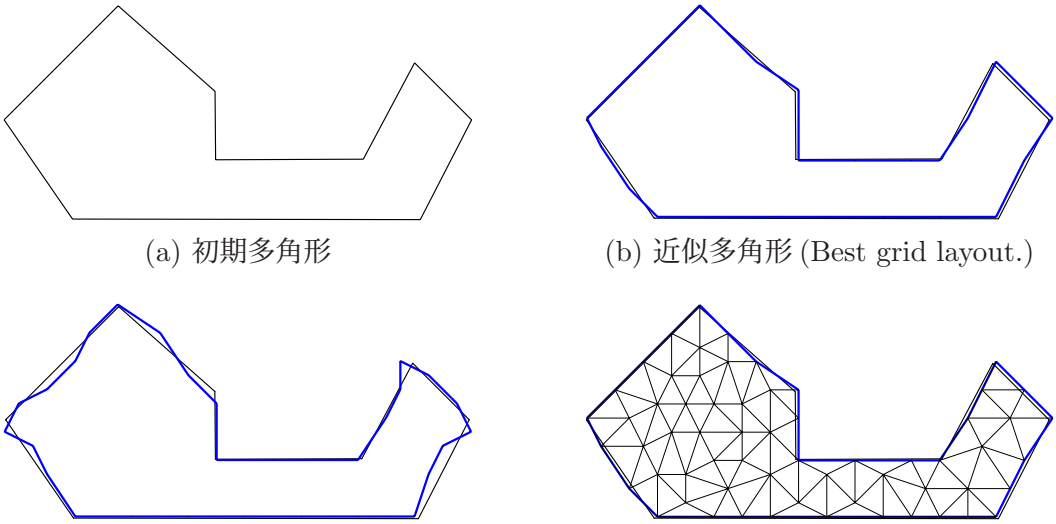
(c) 変形後の形状



(d) 変形後の形状（弾性変形部材のみ）

図 7: 節点，支点と荷重の配置

に，節点 A が上方に所定の量変位するメカニズムを求めた。詳細は省略するが，ラーマングラフの列挙によって，図 6から 7793 個のトポロジーが得られ，それらを初期値として，1148 個の異なるトポロジーを持つメカニズムを得ることが出来た。得られたトポロジーの 1 例を図 7(a) に示す。図 7(b) は変形後の形状であり，弾性変形している部材のみを図 7(c) に示す。



(a) 初期多角形　(b) 近似多角形 (Best grid layout.)

(c) 近似多角形 (Worst grid layout.)　(d) 三角形分割

図 8: 近似多角形とその三角形分割．異なる辺長種類数は 7．($\underline{l} = 50, \alpha = 0.5$ and $\beta = 2.0$)

4 異なる辺長種類数制約下における平面三角形分割

利用可能な異なる部材長数が制限されている条件下において建築三角形トラス構造物を設計する問題は，異なる辺長の種類数制約下における二次元平面上の三角形メッシュ生成問題としてモデル化することが可能である．一般に，任意に与えられた境界多角形内部を定数種類の辺長のみを用いて三角形分割を行うことは，非常に難しい問題と考えられる．そこで本研究では，与えられた境界多角形を僅かにずらす (別の近似多角形を考える) ことによって，異なる辺長種類数が一定である三角形分割を求めるアルゴリズムを提案した．具体的には，以下のように問題を考える．

入力: 各辺長の長さが $\underline{l}$ 以上，$2\underline{l}$ 以下で与えられた頂点数 n の x-単調多角形 P.

出力: 各辺長が $\underline{l}$ 以上，$\beta\underline{l}$ 以下の P の近似多角形 Q と，Q 内部の異なる辺長種類数が一定である三角形分割.

アルゴリズムのアイデアは，近似多角形 Q の頂点及び，メッシュ生成のために Q 内部に追加する Steiner 点を幅 $\alpha\underline{l}$ で与えられた正方形グリッド G の頂点上に配置することである．そのことによって，得られる三角形分割の辺長の上限値が定数ならば，異なる辺長種類数も定数となる．

まず P を G の頂点のみを用いた多角形 Q へと近似を行う問題を考え，Hausdorff distance 最小となる近似多角形を $O((n/\alpha^4)\log(n/\alpha))$ 時間で求めるアルゴリズムを提案した．次に，Q の内部に逐次的に Steiner 点を G の頂点上にのみ追加していき，Q の頂点と Steiner 点に対する Delaunay 三角形分割を求めることで得られる三角形分割が，辺長及び，角度に関して以下を満たしていることを示した．

- 得られた三角形分割の辺長は $\underline{l}$ 以上，$(2+\sqrt{2}\alpha)\underline{l}$ 以下である．
- 得られた三角形分割の各三角形の内角は $\min\{\arccos(\beta/2), \arcsin(1/(2+\sqrt{2}\alpha))\}$ 以上である．

このことにより，アルゴリズムによって得られる三角形分割の異なる辺長種類数は $3\pi/8\alpha^2 + o(1/\alpha^2)$ となる．

図 8にこのアルゴリズムを実装した結果を示す．図 8(a) を初期多角形，$\alpha = 0.5$ とした時，図 8(b),(c) が多角形近似の結果，図 8(d) は図 8(b) の三角形分割である．三角形分割の辺長種類数は 7 種類という結果が得られた．

また図 8(b) と (c) は，同じグリッド幅であっても，グリッドの位置によって得られる近似多角形の近似精度が大きく異なる例である．そこで新たな問題設定としてグリッドの最適配置問題を考え，グリッドの最適配置と近似多角形を同時に求める

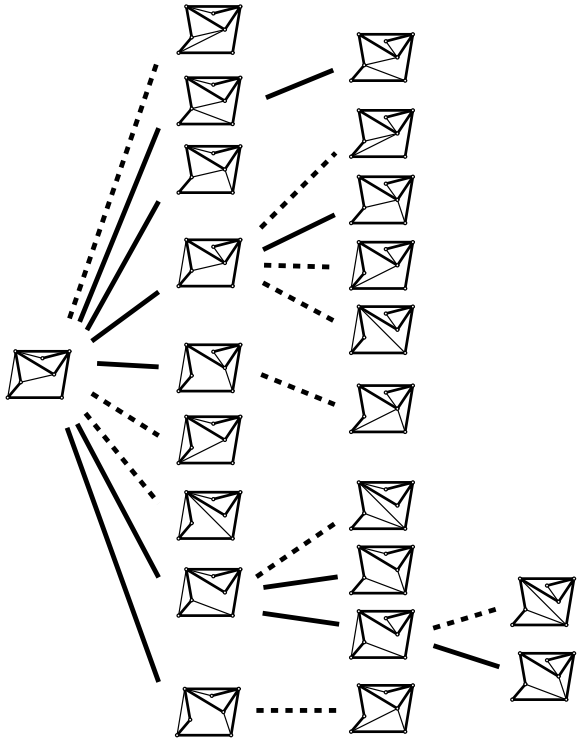


図 9: 頂点集合上の辺制約付き無交差ラーマングラフ列挙の様子．太線が辺制約．

$O((n^3/\alpha^{12})\log(n/\alpha))$ 時間のアルゴリズムを提案した．

5 無交差ラーマンフレームワークの列挙

頂点数 n の頂点集合 $V = \{1, \ldots, n\}$ と辺数 m の辺集合 E から成るグラフ $G = (V, E)$ について，$m = 2n - 3$ でかつ任意の n' 個の頂点部分集合においてその誘導部分グラフの辺数が $2n' - 3$ 以下である時，G をラーマングラフという．平面上の点集合 $P = \{p_1, \ldots, p_n\} \subset \mathbf{R}^2$ が与えられた際，頂点 $i \in V$ を点 $p_i \in P$ へ，辺 $ij \in E$ を伸び縮みのできない線分 (p_i, p_j) へと埋め込んで得られる幾何グラフを棒とジョイントからなるフレームワークと呼び，$G(P)$ と表される．組合せ剛性理論とは，平面や空間内に棒とジョイントを用いて構成されたフレームワークの剛性，安定性を組合せ的に特徴付ける理論である．フレームワークの安定性に関する研究は多数の工学分野の研究者によって行われてきたが，近年，ロボットアームやリンク機構の設計，立体抽出等，その応用の多様さからますます注目を集めることとなった．特に G がラーマングラフである時，G を一般の点配置上に埋め込む事で得られるフレームワークをラーマンフレームワークと呼び，ラーマンフレームワークは P 上の最小数の棒から構成される安定したフレームワークである事が知られている．また一般の点集合上のラーマンフレームワークの集合は，辺集合を台集合をする剛性マトロイドの基集合を構成するなど，優れた組合せ的特徴を有している．

$G(P)$ のすべての直線分対 $(p_i, p_j), (p_k, p_l)$ に対して，(p_i, p_j) と (p_k, p_l) がその端点以外で共有点を持たないとき $G(P)$ は無交差であるという．ここでは P 上のすべての無交差ラーマンフレームワークの列挙を行うアルゴリズムを考える．

幾何グラフの列挙問題は，多数のグラフクラスに対して効率的なアルゴリズムの構築がなされている．列挙を行うオブジェクト (幾何グラフ) の集合を $\mathcal{O}$ とした時，あるオブジェクト $o_1 \in \mathcal{O}$ から辺を一本取り除き新たに一本加えることによって新たなオブジェクト $o_2 \in \mathcal{O}$ を得る局所探索操作を 1-フリップと呼ぶ．2 つのオブジェクトが一回の 1-フリップによって相互に変換可能である時，それらは隣接しているものと捉え，各オブジェクトを頂点とし，隣接しているオブジェクトの頂点間を辺で結ぶことでフリップグラフ $\mathcal{G}_{\mathcal{O}}$ が得られる．列挙アルゴリズムの基本的な考えはこの $\mathcal{G}_{\mathcal{O}}$ 上を効率よく辿ることにあるが，列挙アルゴリズムのみならず近傍探索に基づく最適化アルゴリズムやヒューリスティックの多くは $\mathcal{G}_{\mathcal{O}}$ 上の探索を行う事により実現されている．そのため $\mathcal{G}_{\mathcal{O}}$ の性質を解明することは，各種アルゴリズムの有効性や性能を評価する際の重要な理論を与える．

平面上に固定された点集合上の (無交差とは限らない) ラーマンフレームワークの集合 $\mathcal{L}$ は，剛性マトロイドの基集合を構成することから，1-フリップ (基底変換) を用いて $\mathcal{G}_{\mathcal{L}}$ は連結であり，その列挙も容易である．しかしその部分集合である無交差ラーマンフレームワークの集合 $\mathcal{NL}$ は必ずしも 1-フリップを用いて連結であるとは限らず，$\mathcal{G}_{\mathcal{NL}}$ が連結である事は自明ではない．我々は，Delaunay 三角形分割の性質と剛性マトロイドの性質を用い，以下の定理を新たに示した．

定理 1 全ての無交差ラーマンフレームワークは

$O(n^2)$ 回の *1*-フリップで互いに変換可能である．

また我々の手法は，指定された辺集合 F を必ず含む無交差ラーマンフレームワーク，F 制約付き無交差ラーマンフレームワークの集合に対しても適用可能である．

定理 2 平面上の頂点集合 P と P 上の剛性マトロイド独立辺集合 F に対して，全ての F 制約付き無交差ラーマンフレームワークは $O(n^2)$ 回の *1*-フリップを用いて互いに変換可能である．

三角形分割 T に着目し，T の部分グラフとなっているラーマンフレームワークの集合を $\mathcal{L}_T$ とする．このとき三角形分割の辺集合は無交差であるため，$L \in \mathcal{L}_T$ は必然的に無交差ラーマングラフとなる．さらに $\mathcal{L}_T$ は T の辺集合を台集合とするマトロイドの基集合となるため 1-フリップを用いて互いに変換可能である．よって任意の三角形分割は $O(n^2)$ 回の Delaunay フリップを用いて Delaunay 三角形分割へ変換可能であるという事実と剛性マトロイドの公理を用いることによって，Delauany フリップで変化可能な 2 つの三角形分割 T と T' 上において互いに 1-フリップを用いて変換可能であるような $L \in \mathcal{L}_T$ と $L' \in \mathcal{L}_{T'}$ が必ず存在する事を示すことができる．この事により，全ての無交差ラーマングラフが連結である事が導かれる．

Avis and Fukuda による逆探索手法に基づき，新たに定義したフリップを用いて列挙アルゴリズムの設計を行った．逆探索手法は $\mathcal{G}_{\mathcal{NL}}$ 上を深さ優先探索の要領で辿ることで全てのオブジェクトの列挙を行う．ある無交差ラーマンフレームワーク L 上において，$e_1 \in L$ を取り除き $e_2 \in K_n - L$ を加えるフリップの可能な e_1, e_2 の組合せは $O(n^3)$ である事から，$\mathcal{G}_{\mathcal{NL}}$ 上で L は $O(n^3)$ 個のオブジェクトと隣接している．また与えられたグラフがラーマングラフであるかを判定するのに $O(n^2)$ の計算時間が必要である事から，アルゴリズムが L に隣接している未探索なオブジェクトを探索するのに必要な計算時間は自明な方法で $O(n^5)$ となる

剛性マトロイドの剛体成分を保持するデータ構造等を用い，局所探索を行う辺のペアを適切に特徴づけることにより与えられた頂点集合上の F 制約付き無交差ラーマンフレームワークの集合を $O(n^2)$ のメモリを用いてオブジェクト 1 つあたり $O(n^3)$ の計算時間で求める方法を示した．図 9に平面頂点集合上の列挙の様子を例示する．

6 辞書式順序最大三角形分割と辺制約付き無交差木の列挙

5節で述べた Delaunay 三角形分割を用いて無交差ラーマンフレームワークの列挙を行うアルゴリズムの考えは他のグラフクラスへも適用可能であり，さらに既存の結果を制約付き問題へと拡張することが可能である．

本研究では、さらに新たに辺制約付き辞書式順序最大三角形分割を定義し，この辞書式順序最大三角形分割を Delaunay 三角形分割の変わりに用いることで，辺制約付き三角形分割列挙問題，辺制約付き無交差全域木列挙問題の高速化を実現した．

点集合 P の各要素は x 座標の小さい順に $p_1, p_2, \ldots, p_n$ と記す．P の 2 点を結ぶ線分 (p_i, p_j) を辺と呼ぶ．2 辺 (p_i, p_j) と (p_k, p_l)，但し $i < j, k < l$，に対して，(i)$i < k$ 又は (ii)$i = k$ かつ p_i 周りの時計回り順で p_l が p_j より大きい時，$e \prec e'$ と順序を定める．さらに辺集合 $E = \{e_1 \prec e_2 \prec \cdots \prec e_n\}$ と $E' = \{e'_1 \prec e'_2 \prec \cdots \prec e'_n\}$ に対して，$e_i \neq e'_i$ となる最小添字 i において，$e_i \prec e'_i$ が成り立つ時，辞書式順序で E は E' より小さいとする．

F 制約付き辞書式順序最大三角形分割とは，全ての F 制約付き三角形分割の中で辞書式順序が最大な三角形分割と定義する．極大無交差幾何グラフが三角形分割となることは良く知られた事実であり，このことから F に無交差な辺を貪欲的に加えていく事で必ず F 制約付き三角形分割を得ることができる．さらにこの辺追加の作業を辺の順序の大きい順に行う事で得られる F 制約付き三角形分割が F 制約付き辞書式順序最大三角形となることを示すことができる．

辺制約を考えない場合，つまり通常の辞書式順序三角形分割に対しては，幾つかの優れた性質が知られており，Bespamiyatnikh はそれらの性質を用い高速な三角形分割列挙アルゴリズムを開発している．我々は，辺制約制約を課した場合，つまり F 制約付き三角形分割の集合上に対して，以

下のように辞書式順序最大三角形分割の性質の拡張を行った.

定理 3 頂点集合 P 上の任意の F 制約付き三角形分割は，辞書式順序を増加させる局所操作 *(1-フリップ)* を $O(n^2)$ 回行うことによって F 制約付き辞書式順序最大三角形分割へ変換する事ができる.

さらに Bespamiyatnikh による高速な三角形分列挙アルゴルズムを拡張し，辺制約付き三角形分割を 1 つ当たり $O(\log\log n)$ 時間で列挙を行うアルゴリズムを得た.

我々はさらにこの辺集合の辞書式順序に関する単調性を辺制約付き無交差全域木へ拡張することにも成功した．証明の考えは，各 F 制約付き無交差全域木 ST の辺集合を制約とする ST 制約付き三角形分割を考えることである.

定理 4 頂点集合 P 上の任意の F 制約付き無交差全域木は，辞書式順序を増加させる局所操作 *(1-フリップ)* を $O(n^2)$ 回行うことにより F 制約付き辞書式順序最大無交差全域木へ変換することができる.

また逆探索手法を用いることで，F 制約付き無交差全域木を 1 つ当たり $O(n^2)$ 時間で列挙を行う高速なアルゴリズムを得た.

7 無交差幾何グラフ列挙手法の開発

5節及び 6節の三角形分割を用いた列挙手法を拡張させ，グラフクラスに依存しない一般的な無交差幾何グラフ列挙手法の開発を行った．ある無交差幾何グラフのグラフクラス (無交差全域木，無交差マッチング等) を $\mathcal{NGC}$ を用い表した際，無交差幾何グラフ列挙問題とは平面上の (固定された) 点集合 P，P 上の $\mathcal{NGC}$ に属する全てのグラフを漏れなく列挙する問題である.

新たに得られた手法を適用することで，無交差連結グラフ, 無交差全域木, 無交差ラーマングラフ列挙問題に対して，既存のアルゴリズムより高速な列挙アルゴリズムを得ることができる．また提案する手法は非常に単純であるため，無交差マッチング，無交差 k(辺) 連結グラフ，無交差有向グラフ等のこれまで列挙アルゴリズムが知られていない多くのグラフクラスに対しても適用可能である．表 1 に得られた結果を記す．ここで以下の記号を用いて P 上の各グラフの総数を記す; $\mathsf{pg}(P)$: 平面グラフ, $\mathsf{cg}(P)$:無交差連結グラフ，$\mathsf{sp}(P)$:無交差全域木，$\mathsf{mrf}(P)$:無交差ラーマングラフ，$\mathsf{tri}(P)$: 三角形分割，$\mathsf{pm}(P)$:無交差完全マッチング.

ここで，無交差幾何グラフの列挙問題と無交差とは限らない (通常の) グラフの列挙問題の違いについて言及しておきたい．無交差幾何グラフの (多項式時間) 列挙が自明ではない大きな要因として，分割法と呼ばれる枠組みを適用することができない事が挙げられる．例題として，K_n 内の全域木の列挙問題を考えてみよう．無交差制約がない場合においては，分割法を適用することにより出力グラフ 1 つ当たり $O(n^3)$ 時間の全域木の列挙アルゴリズムを容易に得ることができる．具体的には，まずある辺 e に着目し，e を含む全域木を列挙する問題と e を含まない全域木を列挙する問題の 2 つの子問題に分割する．e を含む全域木を列挙する子問題においては e を縮約することができ，一方 e を含まない全域木を列挙する子問題においては e を問題から取り除くことができる．つまりどちらの子問題においても問題のサイズは小さくなる．さらにグラフが少なくとも 1 つの全域木を含んでいるか，つまりグラフが連結であるかは $O(n)$ 時間で判定することができるため，得られた子問題に対してさらに探索を行うべきかを $O(n)$ 時間で判定することができる．よって分割木を $O(n^2)$ ステップ辿る事で必ず新たな全域木を得ることができ，1 つ当たり $O(n^3)$ 時間で全域木を列挙することができる.

つまり多くの無交差制約のないグラフクラス $\mathcal{C}$ の列挙問題においては，与えられたグラフ内に部分グラフとして $\mathcal{C}$ に属するグラフが存在するかを多項式時間で判定することができれば，上述の分割法を用いて多項式遅延の列挙アルゴリズムを設計することができる．一方，無交差幾何グラフクラス $\mathcal{NGC}$ の場合，与えられた幾何グラフ内に部分グラフとして $\mathcal{NGC}$ に属するグラフが存在するかを問う問題が，無交差全域木や無交差完全マッチング等の基本的なグラフクラスにおいても NP 困難であることが知られている．そのため上述した無交差幾何グラフ列挙の既往手法の何れもが，分割法ではなく 5節と 6節で示した辺フリップに基づくものである．しかし我々は，この無交差幾

表 1: 得られた結果と既存の結果.

	得られた結果	既存の結果
平面グラフ	$O(\mathsf{pg}(P))$	$O(n \cdot \mathsf{pg}(P))$
無交差連結グラフ	$O(\mathsf{cg}(P))$	$O(n \cdot \mathsf{cg}(P))$
無交差全域木	$O(n \cdot \mathsf{tri}(P) + \mathsf{sp}(P))$	$O(n \cdot \mathsf{sp}(P))$
無交差ラーマングラフ	$O(n^2 \cdot \mathsf{mrf}(P))$	—
無交差完全マッチング	$O(n^{3/2} \cdot \mathsf{tri}(P) + n^{5/2}\mathsf{pm}(P))$	—

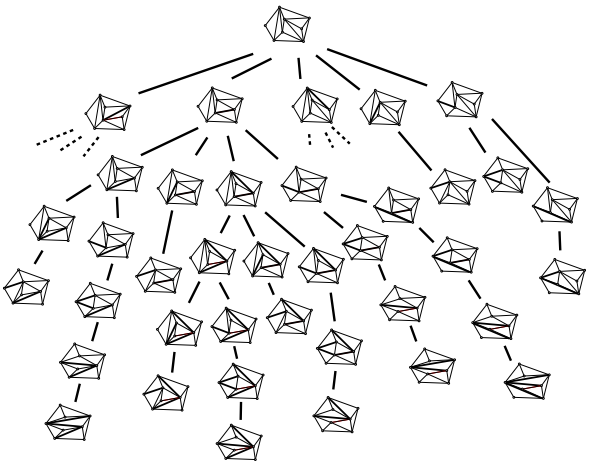


図 10: Bespamiytnikh による三角形分割列挙アルゴリズムの列挙木．太線が各三角形分割の極小元.

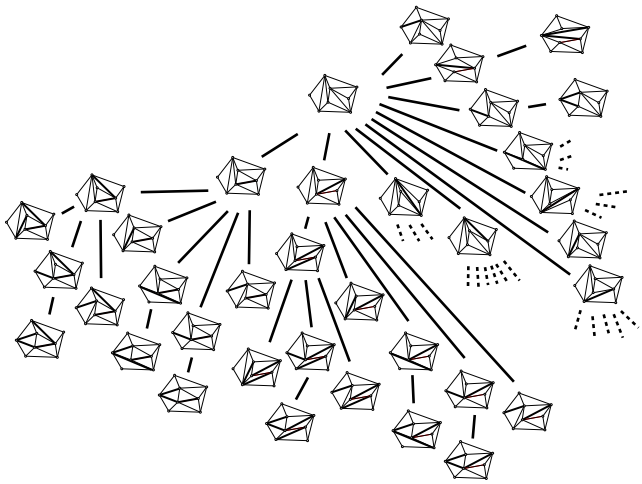


図 11: 新たに開発した三角形分割列挙アルゴリズムの列挙木．太線が各三角形分割の極小元.

何グラフ列挙問題に生じる困難性をいかに回避するかに焦点を絞り，問題を無交差幾何グラフの列挙から通常のグラフの列挙問題へ変換する方法を開発した.

入力頂点集合 P 上の三角形分割の集合を $\mathcal{T}$，列挙を行う無交差幾何グラフの集合を $\mathcal{NGC}$ と記す．提案する手法は，まず P 上の三角形分割 $\mathcal{T}$ の列挙を試み，新たな三角形分割 $T \in \mathcal{T}$ が得られる度に T 上で T の部分グラフとなっている無交差幾何グラフ $G \in \mathcal{NGC}$ の列挙を行う．全ての三角形分割は 1 つあたり $O(\log\log n)$ 時間で効率的に列挙が可能である．また任意の無交差幾何グラフは必ず幾つかの三角形分割の部分グラフであるため，上述の手法で $\mathcal{NGC}$ の全てのグラフを列挙することができる．さらに重要な事として，三角形分割の各辺は互いに無交差であるため三角形分割上で無交差幾何グラフを列挙する際，アルゴリズムは無交差制約を無視することができる．例えば，無交差全域木を列挙する際，列挙の対象をある三角形分割の部分グラフに限ってしまえば，通常の高速な全域木列挙アルゴリズムを用いて無交差全域木を列挙することができる．しかし，任意の無交差幾何グラフ対し，それを含む三角形分割は複数存在する事に注意しなければならない．そのため上述の方法では 1 つの無交差幾何グラフを複数回出力してしまう．我々は以下のように無交差辺集合族の分割を導入することで，漏れや重複なく $\mathcal{NGC}$ を列挙する手法を確立した.

$\mathcal{F}$ を P 上を結ぶ無交差辺集合の集合族とする．まず $\mathcal{F}$ から $\mathcal{T}$ への写像 T^* を定義する．具体的には，$F \in \mathcal{F}$ に対して，$T^*(F)$ は 6節で定義した F 制約付き辞書式順序最大三角形分割と定める．T^* を用いて無交差辺集合族上の同値関係 $\sim$ を定義することができる．つまり任意の無交差辺集合 $F, F' \in \mathcal{F}$ に対して，$T^*(F) = T^*(F')$ のとき $F \sim F'$ と記す．これにより各三角形分割 $T \in \mathcal{T}$ の辺集合 $E(T)$ を代表元とする同値類 $[E(T)] = \{F \in \mathcal{F} \mid F \sim E(T)\}$ が自然に定まる．我々は各同値類 $[E(T)]$ に対して，包含関係に対して極小な元 $R(T)$ が存在することを示した: 各 $[E(T)]$ に対して，ある $R(T) \in [E(T)]$ が存在し，任意の $F \in \mathcal{F}$ が $F \in [E(T)]$ であるための必

要十分条件は $R(T) \subseteq F \subseteq E(T)$ が成り立つことである．この $R(T)$ を T の極小代表元と呼ぶ．この極小代表元を用いて我々のアルゴリズムは以下のように記述される (図 10):

Phase 1: P 上の三角形分割の列挙を行う．

Phase 2: 新たな三角形分割 T が得られる度に，T の部分グラフでありかつ T の極小代表元 $R(T)$ を含む $G \in \mathcal{NGC}$ を全て列挙する．

このアルゴリズムが $\mathcal{NGC}$ を漏れや重複なく列挙することは，$\{[E(T)] \mid T \in \mathcal{T}\}$ が $\mathcal{F}$ の分割である事と，任意の $F \in \mathcal{F}$ が $F \in [E(T)]$ であるための必要十分条件が $R(T) \subseteq F \subseteq E(T)$ である事によって簡単に示すことができる．つまり任意の $G \in \mathcal{NGC}$ に対して，その辺集合を $E(G)$ と記した際，$E(G) \in [E(T)]$ となる三角形分割 T が唯一存在し，この時 $R(T) \subseteq E(G) \subseteq E(T)$ が成り立つ．よって G はこの T に対する Phase 2 において出力される．また逆に $R(T) \subseteq E(G) \subseteq E(T)$ が成り立つならば $E(G) \in [E(T)]$ であり，そのような T は唯一である．よって G はこの T に対する Phase 2 においてのみ列挙される．よって上述の手法で $\mathcal{NGC}$ の全てのグラフを重複なく列挙することができる．

しかし上述したアルゴリズムは列挙の際，全ての三角形分割の列挙するため，列挙を行うグラフの総数が三角形分割の総数に比べ極端に小さい場合は非常に非効率と考えられる．例えば，無交差マッチングの列挙問題に対して上述のアルゴリズムを適用した場合，ある三角形分割 T において，もし極小代表元 $R(T)$ が誘導するグラフが頂点次数 2 以上の頂点を持つならば，$R(T)$ を含み T に含まれるような無交差マッチングは存在しない．つまり T を列挙することは無駄な作業である．しかし Bespamiyatnikh の三角形分割列挙アルゴリズムは図 10に示した探索木を辿ることにより列挙を行うため，ある特定の三角形分割 T の列挙を回避する事は探索木上でそれ以降の T の孫となっている三角形分割 T' の列挙も行わないこととなる．そのため T' に含まれており $R(T')$ を含むマッチング M が存在した場合，M を列挙し忘れることとなる．

我々はこのような問題を回避し，かつ全ての三角形分割を列挙することなく目的のグラフクラスに応じて必要な三角形分割のみを列挙する発展手法を開発した．具体的には上記の問題を回避するため，極小代表元 $R(T)$ の包含関係に対して単調性を有する三角形分割探索木が存在することを示し，その探索木上を辿る三角形分割列挙アルゴリズムを新たに開発した．つまり新たに得られた探索木上では，三角形分割 T 及びその孫 T' に対して，$R(T) \subset R(T')$ が成り立つ (図 11)．よって無交差マッチングの列挙の場合，ある三角形分割 T において極小元 $R(T)$ が誘導するグラフが頂点次数 2 以上の頂点を持つならば，その孫 T' においても同様であり，結果そのような T に遭遇した場合，その孫全ての列挙を回避することができる．つまり必要な三角形分割のみの列挙が可能である．この手法はより一般的に記述することができ，無交差ラーマングラフの高速な列挙にも応用可能である．

8 避難計画問題に対する効率的なアルゴリズム

近年，地震のみならず数多くの災害が発生している．そのような背景から，大地震や火事，そして津波といった大規模災害に対する危機管理システムの構築の必要性が叫ばれている．本論文は，そのような危機管理システム構築のための理論的基礎研究である．最も効率的な避難計画を求める問題は，動的ネットワークフローを用いた避難計画問題として定式化される．避難計画問題においては，n 個の要素を持つ点集合 V と m 個の要素を持つ辺集合 A を持つ有向グラフ $D = (V, A)$，容量関数 $c\colon A \to \mathbb{R}_+$，移動時間関数 $\tau\colon A \to \mathbb{Z}_+$，サプライ関数 $b\colon V \to \mathbb{R}_+$，そしてシンク $s \in V$ で構成される動的ネットワーク $\mathcal{N} = (D = (V, A), c, \tau, b, s)$ が与えられる（図 12参照）．ただし，$\mathbb{R}_+$ と $\mathbb{Z}_+$ は，それぞれ非負の実数と非負の整数の集合である．

都市地域における避難計画を想定しているならば，点は建物や部屋，出口をモデル化し，辺は道路や通路をモデル化している．辺 $e \in A$ に対して，容量 $c(e)$ は単位時間当たり e を通過することのできる人数を，移動時間 $\tau(e)$ は e を通過するのに必要な時間を表している．点 $v \in V$ に対して，サプライ $b(v)$ は v に存在する人の数を表している．この時，避難計画問題は全てのサプライをシ

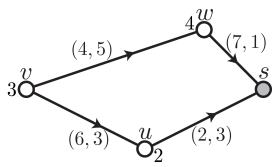


図 12: 動的ネットワークの例．ただし，点に付されている数字はサプライを，辺に付されている数字の組は容量と移動時間を表わしている．

ンクまで流すことが可能な最小時間を求める問題である．

形式的には，動的ネットワーク $\mathcal{N}$ が与えられたとき，避難計画問題 EP($\mathcal{N}$) は以下のように定義される．

$$\mathrm{EP}(\mathcal{N})\colon \mathrm{minimize}\{\Theta(f)\colon f \text{ は } \mathcal{N} \text{ で実行可能}\}.$$

ただし，動的フロー $f\colon A \times \mathbb{Z}_+ \to \mathbb{R}_+$ が以下の三つの条件を満たすとき実行可能であるとする [3].

容量制約： 任意の $e \in A$ と $\theta \in \mathbb{Z}_+$ に対して

$$0 \le f(e,\theta) \le c(e).$$

フロー保存則： 任意の $v \in V$ と $\Theta \in \mathbb{Z}_+$ に対して

$$\sum_{e\in\delta^+(v)} \sum_{\theta=0}^{\Theta} f(e,\theta) - \sum_{e\in\delta^-(v)} \sum_{\theta=0}^{\Theta-\tau(e)} f(e,\theta) \le b(v).$$

要求制約： 以下を満たす $\Theta \in \mathbb{Z}_+$ が存在する．

$$\sum_{\delta^-(s)} \sum_{\theta=0}^{\Theta-\tau(e)} f(e,\theta) = \sum_{v\in V} b(v). \tag{8.1}$$

ただし，$\delta^+(v)$ と $\delta^-(v)$ は，それぞれ A の辺の中で始点と終点が v の辺の集合，$\Theta(f)$ は (8.1) を満たす最小の Θ である．

避難計画問題に対しては，Ford と Fulkerson によって，時間拡大ネットワークを用いた擬多項式時間アルゴリズムが提案された [1]．ただし，動的ネットワーク $\mathcal{N}$ に対する時間拡大ネットワークとは，時間制限 T が与えられたとき，時刻 0 から T までの $\mathcal{N}$ における動的フローと等価な通常のフローを表現することができる，$\mathcal{N}$ の $O(T)$ 倍のサイズを持つ通常の（つまり移動時間関数を持たない）ネットワークである（図 13参照）．

多項式時間アルゴリズムとしては，唯一のものが Hoppe と Tardos [2] によって提案されている

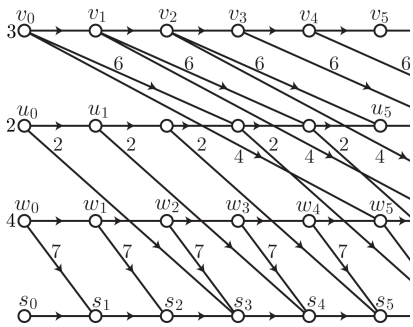


図 13: 図 12に対する $T = 7$ の場合の時間拡大ネットワーク．点に付されている数字はサプライを表わし，数字の付されていない点のサプライは 0 であるとする．また辺に付されている数字は容量を表わし，数字の付されていない辺の容量は $+\infty$ であるとする．

が，このアルゴリズムはサブルーチンとして劣モジュラ関数最小化を必要としており，計算時間が高次元の多項式となっている．ネットワークの構造を制限した研究としては，Mamada 等 [3] が木構造に対する $O(n \log^2 n)$ 時間アルゴリズムを提案しており，また Hall 等 [4] は単一パス長という性質を持つネットワークに対する避難計画問題に対する研究を行っている．

現実に計算機で避難計画問題を解く際には，Hoppe と Tardos のアルゴリズムは複雑すぎるため，Ford と Fulkerson によって提案された時間拡大ネットワークが必要となることがある．しかし，我々の実験によると，そのネットワークのサイズの巨大性より，必要な精度を保つためには，一辺の点数が高々20 のグリッドネットワークまでしか，通常の計算機では解くことができず，やはり特別なネットワークのクラスに対する，高速で簡素なアルゴリズムの必要性が確認された．

本論文では，避難計画問題に対して実際の応用に有益であると考えられるネットワークのクラスに対する効率的なアルゴリズムを提案し，さらに通常の避難計画問題に加え，現実の避難の際に考えるべき問題を制約に加えた避難計画問題に対する研究も行った．

8.1 容量が一定のグリッドネットワークに対する効率的なアルゴリズム

まず，グリッド構造（図 14参照）を持ち，全ての辺の容量と移動時間が等しいネットワークに対する避難計画問題の最速避難完了時間を求める効率的なアルゴリズムを開発した．この様なグリッドネットワークは都市のモデル化の際，頻繁に現れる構造であり，この構造を持つネットワークに対する高速なアルゴリズムは大変有益であるといえる．

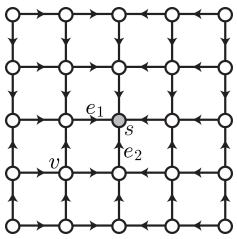


図 14: グリッドネットワーク

グリッドネットワークに対する効率的なアルゴリズムを開発するために必要な幾つかの観察を説明する．まず，一つめの観察はグリッドネットワークが図 15のように，四つの辺素な部分ネットワークに分解することができることである．

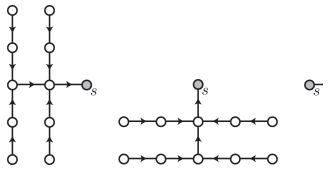


図 15: グリッドネットワークの分解

次に図 14の点 v に注目しよう．v のサプライは，必ず辺 e_1 か e_2 のいずれかを通って s に到達する．同様にして，任意の点に対して，その点のサプライが s に到達するために必ず通過する二本または一本の辺が決まる．ここで，全ての点に関して，それぞの辺に対するサプライの分配が固定されている問題を考えよう．本論文では，この問題に対して分配されたサプライが上記のように分解されたネットワークを流れるフローが最速避難完了時間を達成し，その最速避難完了時間は容易に求めることができることを示した．よって，残された問題はどのようにして最適なサプライの分配を求めるかであるが，本論文ではこの最適なサプライの分配を求める問題がネットワーク制約下の資源配分問題に帰着することができることを証明し，最適なサプライと最小避難完了時間を $O(n \log n)$ 時間で求めることの出来るアルゴリズムを提案した．また，その最小避難完了時間を実現するフローを具体的に求めるためには，[3] のアルゴリズムを用いれば，$O(n \log^2 n)$ 時間で求めることができる．

8.2 完全連結単一パス長ネットワークに対する効率的なアルゴリズム

続いて辺の容量と移動時間関数が一定なグリッドネットワークを一般化したネットワークに対する効率的なアルゴリズムの説明をする．本論文では，以下のような三つの条件を満たすネットワークを考える： (i) 全ての辺の容量が 1，(ii) 各点からシンクまでの任意のパスの辺の移動時間の合計が一定，(iii) 各点 v とシンクとの最小カットが，シンクに入る辺のうち始点が v から到達可能なもので構成される（図 16参照）．

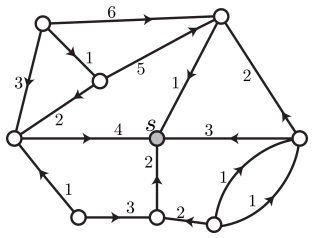


図 16: 単一パス長ネットワーク．辺に付されている数字は移動時間を表わす．また全ての辺の容量は 1 とする．

先ほどのグリッドネットワークにおける避難計画問題が効率的に解けたのは，上記の条件 (iii) の代わりに，四つの辺素な部分グラフに分解することが出来たからであった．本論文では，まず以下の定理を証明することにより，辺素な部分ネットワークへの分解と条件 (iii) が等価であることを証明した．ただし，以下の定理を避難計画問題に適用する場合には，集合 S をシンク s の子供の集合と設定する．

定理 5 特別な点の集合 $S = \{s_1, \ldots, s_d\}$ を持つ有向グラフ $D = (V, A)$ と関数 $f\colon S \to \mathbb{N}$ が与えられているとする．各 $v \in V$ に対して $R(v)$ は S の点のうち v から到達可能なものを表わす．各 $i = 1, \ldots, d$ に対して，V の点うち s_i に到達可能なものの集合を V_i で表すとする．さらに，D に新しい点 s^* を加え，各 s_i から s^* に $f(s_i)$ 本の並行辺を加えて得られる有向グラフを D^* とする．このとき，各 $T_{i,1}, \ldots, T_{i,f(s_i)}$ が s_i を根とし，各 $T_{i,j}$ が V_i を張るような，$i = 1, \ldots, d$ に対して $T_{i,1}, T_{i,2}, \ldots, T_{i,f(s_i)}$ と表される $\sum_{s_i \in S} f(s_i)$ 個の辺素な木をパッキングできる必要十分条件は，全ての点 $v \in V$ に対して，D^* における v から s^* への局所連結度が $\sum_{s_i \in R(v)} f(s_i)$ 以上であることである．

この定理は，Edmonds [5] によって証明された，有向グラフにおける内向木のパッキングに関する有名な定理の一般化になっており，様々な応用があり，この結果単独でもグラフ理論的に興味深い結果となっている．

定理 5は，全ての $v \in V$ に対して $R(v) = S$ が成り立つときは，[5] の結果より容易に導くことができるが，その結果から定理 5が直接導くことができないことを示す．図 17(a) の有向グラフを考えよう．だたし $f(s_1) = f(s_2) = 1$ とする．そして，全ての $v \in V$ に対して $R(v) = S$ が成り立つように，図 17(b) のように新しく二本の辺を加えたグラフを考えよう．すると，先程の事実より s_1 と s_2 を根とし V を張る二つの内向木が得られる．図 18の二つの内向木が得られたとしよう．すると，先ほど加えた辺（図の中では点線で表される辺）を取り除くことによって，定理 5の内向木，つまりそれぞれが s_1 と s_2 を根とし s_1 と s_2 に到達可能な点を張る内向木を得ることができない．これによって定理 5の非自明性が示される．

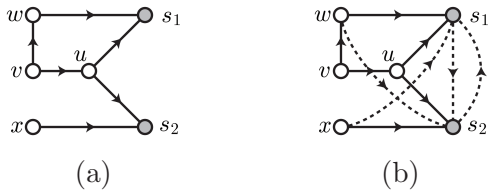


図 17: (a) 入力のグラフ．(b) 変形後のグラフ．

では定理 5の証明の概要を説明しよう．証明は $\sum_{i=1}^{d} f(s_i)$ の値に関する帰納法で証明する．つまり，ある $i \in \{1, \ldots, d\}$ を固定したとき D^* 上の s^* を根とし，点集合が $V_i \cup \{s^*\}$ である内向木以下の条件を満たす T が常に存在することを証明すればよい．(i) T の辺のうち始点が s^* のものは一本だけである．(ii) 各点 $v \in V$ に対して，もし v か s_i に到達可能であるならば，D^* から T の全ての辺を取り除いたグラフ $D^* \backslash T$ における v から s^* への局所連結度が $\sum_{s_i \in R(v)} f(s_i) - 1$ 以上であり，もし v が s_i に到達可能でなければ，$D^* \setminus T$ における v から s^* への局所連結度が $\sum_{s_i \in R(v)} f(s_i)$ 以上である．本論文では，このような T を辺を一本ずつ加えることによって構成できることを証明した．また，避難計画問題とは直接関係はないが，辺素な内向木のパッキングに対する以下のような定理 5に関してとは別の特徴付けの証明も行った．ただし，有向グラフの連結部分グラフが辺の向きを無視したとき閉路を持たないとき木であると言う．

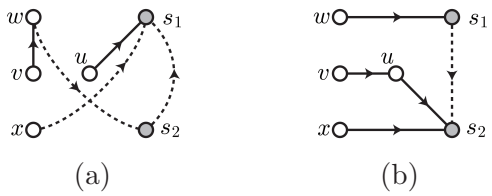


図 18: (a) s_1 を根とする内向木．(b) s_2 を根とする内向木．

定理 6 特別な点の集合 $S = \{s_1, \ldots, s_d\}$ を持つ有向グラフ $D = (V, A)$ と関数 $f\colon S \to \mathbb{N}$ が与えられているとする．$R(v)$，V_i と D^* は定理 *5*と同様に定義する．このとき，定理 *5*で定義されている内向木 $\{T_{i,j}\colon i = 1, \ldots, d, j = 1, \ldots, f(s_i)\}$ がパッキングできる必要十分条件は，以下の条件を満たす D^* の部分グラフの集合 $\{T^*_{i,j}\colon i = 1, \ldots, d, j = 1, \ldots, f(s_i)\}$ が存在することである：*(i)* $T^*_{i,j}$ は $V_i \cup \{s^*\}$ を張る木である．*(ii)* 全ての $T^*_{i,j}$ の合併を F と表したとき，各 $v \in V$ に対して F の辺のうち始点が v のものの本数は $f(R(v))$ で，s^* が始点のものは 0 本である．

この定理は有向グラフの被覆などに応用がある．

では，避難計画問題に戻ろう．定理 5により，入力のネットワークが辺素な部分ネットワークに分解することができることが証明されたので，条件 (ii) を用いることにより，グリッドネットワークの場合と同様にして，最適なサプライの分配を他の

問題に帰着することができる．ただし，グリッドネットワークの場合はネットワーク制約下の資源配分問題に帰着したのに対して，今回はさらに効率的に解くことができる限定パラメトリックフロー問題 [7] に帰着できることを証明した．以上の結果より，この問題に対する $O(\{k^2n^2 + kmn\}\log n)$ で最小避難完了時間と，最小避難時間完了時間を実現するフローを求めるアルゴリズムを提案した．また，このアルゴリズムは，容量が有理数である場合に容易に拡張することができる．

8.3 現実的な制約を加えた避難計画問題

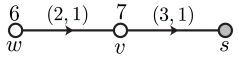
また，本論文では通常の避難計画問題に加え，現実の避難の際に考えるべき問題を制約に加えた避難計画問題に対する研究も行った．具体的には以下の二つの問題を扱った

1. 避難時の困難を避けるため，避難の途中の点における滞留を禁止する．
2. 避難時の指示を容易にするため，それぞれの点からシンクまでの避難経路を一つに定める．

本論文では，1 の問題を滞留制約避難計画問題，2 の問題をパス制約避難計画問題と呼ぶ．

滞留制約避難計画問題

まず，滞留を禁止したフローの定義を行う．図 19の動的ネットワークを考えよう．一つめのフローとして各点のサプライが目の前の辺に可能なだけ流れるものを考えよう．表 (8.2) はこの場合の，各辺に入るフローの量を表わしている．便宜上，v のサプライと w のサプライは区別できるとしよう．f_v と f_w は v と w のサプライのフローの量を表わしている．この場合，w のサプライは v で待っている．よってこのフローは好ましくない．次に，表 (8.3) で表されるフローを考えよう．この場合は，どのサプライも途中の点において待っていはいないが，v のサプライが w のサプライに道を譲っている．これは不自然である．最後に表 (8.4) のフローを考えよう．このフローにおいては，w のサプライは時刻 0 において，辺 wv が空であるのに待たなければならくなっている．これも不自然である．



(a)

図 19: Input dynamic network.

time		0	1	2	3	4
vs	f_v	3	3	1	0	0
	f_w	0	0	2	3	1
wv	f_w	2	2	2	0	0

(8.2)

time		0	1	2	3	4
vs	f_v	3	1	1	1	1
	f_w	0	2	2	2	0
wv	f_w	2	2	2	0	0

(8.3)

time		0	1	2	3	4
vs	f_v	3	3	1	0	0
	f_w	0	0	2	2	2
wv	f_w	0	2	2	2	0

(8.4)

以上の議論より本論文では，滞留制約フローを以下の二条件を満たすフローとして定義する．

1. 途中の点における滞留は禁止する．
2. 各点 v と時刻 θ において, v のサプライのうち v からでるサプライの量は，$\sum_{e\in\delta^+(v)} c(e)$ と v の残っているサプライの量の最小値となる．

条件 2 は後方のサプライに道を譲らないことを意味している．本論文では，単一パス長ネットワークに対する，滞留制約フローの存在判定問題に対する効率的なアルゴリズムを提案した．

パス制約避難計画問題

パス制約避難計画問題とは，通常の避難計画問題に，各点のサプライは全て同じ道を通ってシンクまで流れる制約を加えたものである．本論文では，この問題が単一パス長で全ての辺の容量が一のネットワークに対しても $\mathcal{NP}$ 困難であることを証明した．具体的な証明は，$\mathcal{NP}$ 完全問題であるパーティションの任意のインスタンスをこの問題に帰着できることを証明した．パーティションのインスタンス $\{a_1, \ldots, a_p\}$ に対する具体的なネットワークの構造は図 20のようになる．

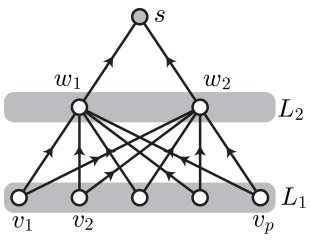


図 20: 帰着の図．ただし，各 v_i のサプライは a_i, w_1, w_2 のサプライは 0 であるとする．また，全ての辺の容量は 1，移動時間は 0 とする．

9 線的施設配置問題に関する研究

今や自動車は生活になくてはならないものであるが, 自動車交通による混雑は深刻な問題となっている．そうした中で車社会を見直す動きが高まってきており, LRT(Light Rail Transit) などの公共交通機関が注目されている．このことからも, 公共交通の根幹となる都市内鉄道網の形状を決定することは, 重要な意味を持つ問題であると言える．こうした背景を踏まえ, 本研究では文献 [8] の内容を発展させた手法を開発した。

9.1 線的施設配置問題

・都市モデル

対象領域を 1 辺の長さが 1 の正方形領域 A（簡易的な都市を表しているとする）とし, ここに住民が住んでいると想定する．正方形領域 A 上に, 都市交通ネットワークに対応するグラフを配置する問題を考え, そのグラフを $G = (Q, E)$ で表す．ただし, G の頂点集合 Q は駅を表し, 枝集合 E は鉄道網を表す．また, 徒歩速度を v_w, 列車速度を v_r とし, 移動距離はユークリッド距離を用いる．

・需要点

対象領域 A を辺長 Δx と Δy の方形の微小地区 $a_i (i = 1, 2, \ldots, n)$ に分割する．微小地区 a_i の中心点を p_i とし, この微小地区内の住民は p_i に集中しているものと考える．

・ＯＤ (Origin Destination) **交通量**

地区 a_i 内から地区 a_j 内への交通量は, すべてその中心点 p_i から p_j への移動とみなす．その 2 点間のＯＤ交通量 w_{ij} は地区 a_i と地区 a_j の人口の積に比例し, そのユークリッド距離 $d(p_i, p_j)$ の二乗に反比例すると仮定して, $\alpha_{ij} = s_i s_j / \{d(p_i, p_j)\}^2$ とする．ただし s_i, s_j は, 地区 a_i, a_j の人口を表す．

・線的施設配置問題の定式化

ある領域内に存在する交通ネットワークの利便性を示す評価指標として, 移動の起点から終点までに要する移動時間を用いて, 全利用者の平均移動時間 T_{avg} を最小化する交通ネットワークの形状を求める問題を考える．任意の 2 点,p_i, p_j の最短移動時間 $t_{\min}(p_i, p_j)$ は, 徒歩のみで移動する場合に出発地から目的地に到達するまでにかかる時間を t_w, 鉄道を経由して移動する場合にかかる時間を t_r とすると, $t_{\min}(p_i, p_j) = \min\{t_w(p_i, p_j), t_r(p_i, p_j)\}$ で与えられる．よって対象領域内における全移動の平均移動時間は, 全利用者の移動時間の総和を対象領域内の総ＯＤ交通量 α で除すればよいから, Minimize $T_{\mathrm{avg}} = \frac{1}{\alpha} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \alpha_{ij} t_{\min}(p_i, p_j)$ と定式化できる．

9.2 ネットワーク最適化の手法

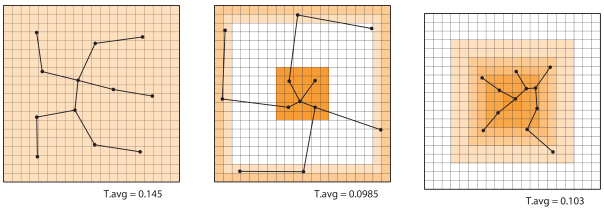
本研究では, 前節で定式化した問題を 2 段階にわけて最適化する手法を提案する．まず 1 段階としては、駅の位置を固定した状態でトポロジーの最適化を行い, 次に 2 段階としてトポロジーを固定した状態で駅の配置の最適化を行う．この 2 つの操作を何度も繰り返し行うことで, 最適解を求めようというものである．本研究で取り扱う問題は, 問題をネットワークのトポロジーを求めることに限定した場合, Minimum Communication Spanning Tree(MCST) と呼ばれる最適化問題に帰着される．しかし, MCST は一般に困難な問題であることが知られている（文献 [9]）．そこで, 本研究では以下の 3 つの方法を用いてトポロジーを与え, 得られた最良解について比較考察を行うこととする．(i) 最小全域木（Minimum Spanning Tree : MST）によってトポロジーを与えることで最適化を行う．(ii)MCST を整数計画問題 (Integer Program : IP) として定式化することで MCST の解を厳密に求め, MCST によってトポロジーを与えて最適化を行う．(iii) MCST の解を文献 [10] で提案されている進化的アルゴリズム (Evolutionary Algorithm : EA) を用いることで発見的に求め, MCST によってトポロジーを与えて最適化を行う．駅の配置最適化については, 最急降下法を用いて行う．

9.3 数値実験

実際に数値を代入し, ネットワークのトポロジーを前述した 3 つの手法で求め, 得られた解の比較を行う. ここでは例題として, $n = 400, m = 12, v_w = 1, v_r = 10$ とした. さらに, 各需要点の密度を図 21のような 0〜5 の 6 段階の重みで表し, 3 種類の異なる需要点密度分布モデルで実験を行う.

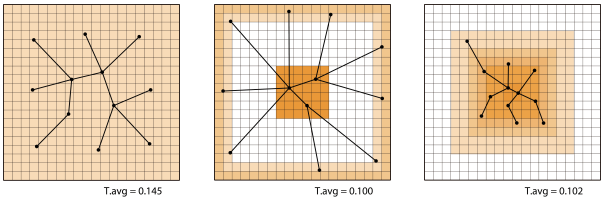
それぞれの手法を用いて得られた, 各モデルにおける最良解を図 22-24に示す.

図 21: 各需要点における密度の重み.



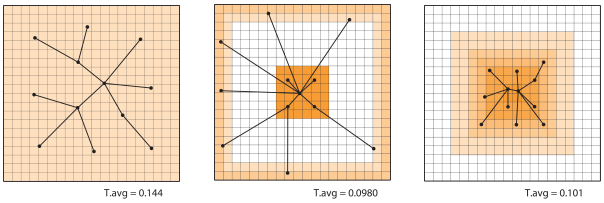
Model1 Model2 Model3

図 22: トポロジーを MST で与えた場合の最良解のネットワークの形状.



Model1 Model2 Model3

図 23: トポロジーを MCST で与えた場合（IP）の最適解のネットワークの形状.



Model1 Model2 Model3

図 24: トポロジーを MCST で与えた場合（EA）の最適解のネットワークの形状.

参考文献

[1] L.R. Ford and D.R. Fulkerson. *Flows in Networks*. Princeton University Press, Princeton, NJ, 1962.

[2] B. Hoppe and É. Tardos, The quickest transshipment problem. *Mathematics of Operations Research*, 25(1):36–62, 2000.

[3] S. Mamada, T. Uno, K. Makino, and S. Fujishige, An $O(n \log^2 n)$ algorithm for the optimal sink location problem in dynamic tree networks. *Discrete Applied Mathematics*, 154(16):2387–2401, 2006.

[4] A. Hall, S. Hippler, and M. Skutella, Multicommodity flows over time: Efficient algorithms and complexity, *Theor. Comput. Sci.*, 379(3):387-404, 2007.

[5] J. Edmonds. Edge-disjoint branchings. In R. Rustin, editor, *Combinatorial Algorithms*, pp.91-96. Academic Press, New York, 1973.

[6] T. Ibaraki, N. Katoh, Resource allocation problems under submodular constraints, In Resource Allocation Problems : Algorithmic Approaches, MIT Press, Cambridge, MA, pp.144–176, 1988.

[7] G. Gallo, M. D. Grigoriadis, and R. E. Tarjan, A fast parametric maximum flow algorithm and applications, *SIAM J. Comput.*, 18(1):30–55, 1989.

[8] 塚本智久, 加藤直樹, 瀧澤重志. 『都市における線的施設配置問題に関する研究』. 学術講演梗概集 A-2, pp.557-558, 2005 年度大会（近畿）.

[9] Michael R. Garey and David S. Johnson. COMPUTERS AND INTRACTABILITY, A Guide to the Theory of NP-Completeness. *W. H. FREEMAN AND COMPANY*, 1991.

[10] Gunther R. Raidl and Bryant A. Julstrom. Edge Sets: An Effective Evolutionary Coding of Spanning Trees. *IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION*, 7(3):225-239, JUNE 2003.

研究業績一覧

著書

1. 日本建築学会編:
"建築最適化への招待", 日本建築学会, 2005.4.
概要: 本書は主に建築を学ぶ学生を対象に最適化技術の魅力を理解してもらうための啓蒙書として，建築に関する平易な事例を通して最適化理論の初歩が学べるように工夫されている。また，取り扱う建築分野も都市計画，建築計画，構造形状，構造設計，熱環境，内外装計画の多岐にわたっている。

2. 加藤直樹:
"数理計画法", コロナ社, 2007.12.
概要: 数理計画はシステム最適化の理論的基礎をなす学問分野であり、本書は線形計画，ネットワーク計画，整数計画，組合せ最適化，非線形計画の各最適化問題の理論的基礎と解法について直観的な理解が得られるよう具体的な例題を用いながら詳述している。近似アルゴリズム、動的計画法、施設配置、マトロイド理論についても解説をおこなっている。

学術論文

1. Danny Z. Chen, Jinhee Chun, Naoki Katoh, and Takeshi Tokuyama:
"Efficient Algorithms for Approximating a Multi-Dimensional Voxel Terrain by a Unimodal Terrain", Proc. of 10th International Computing and Combinatorics Conference(COCOON 2004), LNCS 3106, Springer, 238-248, 2004.8.
概要: Optimized region rules developed by Fukuda et al. are effective tools in data mining for a database with numeric data. However, there are two defects in the previous methods: (1) each rule can contain at most two numeric conditional attributes, and (2) the decision is only done by whether a given data is inside of a region R or outside of it, irrelevant to its exact position. In this paper, we propose a method for resolving these defects. Indeed, by applying graph algorithms, we give optimized numeric association rules with more than two attributes, and also give layered-structured numeric association rules. The method is also applicable to removal of exceptional data and also clustering.

2. 天沼はるか、加藤直樹:
"2 値化エッジ画像からの高速長方形成分検出手法の開発", 日本建築学会環境系論文集, 586, 83-90, 2004.12.
概要: 本研究では、カラー画像やグレースケール画像に対して、2 値化処理が予め施された 2 値化エッジ画像を入力画像とし、1 枚の画像から実用的な計算時間で長方形集合を検出する方法を提案する。また、画像に対して主軸方向が同一で、スライシング構造を持つ複数の長方形とする。本論文では、実用的な計算時間で高精度の長方形検出をおこなうため、次のようなさまざまな工夫をおこなっている。(1) 長方形群の主軸 2 方向に沿ったスライス処理を施すことによって、長方形が存在する可能性の高い領域を限定する処理をおこなっている。そこで得られた領域ごとに、長方形検出のためのしきい値を適応的に設定することによって、欠損率の異なる長方形を一度に検出できるようにしている。(2) ノイズカットや濃度値データの計算といった前処理を施すことによって、計算効率と検出精度の向上を図っている。(3) 線形時間の動的計画法を利用して長方形の辺候補が存在する部分を絞り込むことによって、辺候補 4 本の組合せで得られる長方形候補の数を劇的に減らし、結果として計算時間の大幅な短縮を達成している。

3. 大崎 純，西脇眞二:

“スナップスルーを利用したバイステーブルコンプライアントメカニズムの形状設計法”, 日本機械学会論文集 A, 70(700), 23-28, 2004.12.

概要: A general approach is presented for generating compliant mechanisms from a highly connected ground structure. The structure is modeled as a pin-jointed truss, and geometrical nonlinearity is considered. An optimization problem is formulated for minimizing the total structural volume under constraints on the displacement at the specified node, and stiffnesses at initial and final states. The design variables are cross-sectional areas and the nodal coordinates. It is shown in the numerical examples that several mechanisms can be naturally found as a result of optimization starting from randomly selected initial solutions.

4. D.Z. Chen, O. Daescu, Y. Dai, N. Katoh, X. Wu:
“Efficient Algorithms and Implementations for Optimizing the Sum of Linear Fractional Functions, with Applications”, Journal of Combinatorial Optimization, 9(1), 69-90, 2005.1.

概要: The problem of optimizing the sum of m linear fractional functions (SOLF) in a fixed dimension d, subject to n linear constraints, arises in a number of theoretical and applied areas. This paper presents an improved algorithm for solving the SOLF problem in 2-D. A key subproblem to our solution is the *off-line ratio query* (OLRQ) problem, which computes the optimal values of a sequence of m linear fractional functions (called ratios), with the ratios subject to a dynamically changing feasible domain defined by $O(n)$ linear constraints. Based on useful geometric properties and the parametric linear programming technique, we develop an algorithm that solves the 2-D OLRQ problem in $O((m+n)\log(m+n))$ time. Our OLRQ algorithm can be easily implemented and is robust. More importantly, it enables us to speed up every iteration of a known iterative SOLF algorithm in 2-D, from $O(m(m+n))$ time to $O((m+n)\log(m+n))$. Implementation results of our improved SOLF algorithm have shown that in most cases our algorithm outperforms the commonly-used approaches for the SOLF problem.

5. 宮高泰昌，加藤直樹，瀧澤重志:
“建築画像データベースの自動索引付けを目指した建築構成要素の認識手法の開発 畳と格子を例として”, 日本建築学会環境系論文集, 588, 63-70, 2005.2.

概要: 建築画像データベースを構築することを最終目標として，建築画像の自動索引付けには，どのような検索キーを付与することが有用であるか，という点を考慮することが重要である。本研究では，建築画像データベースにおける有効な検索システムの構築を目指し，画像への自動索引付けを目的として,「画像中に特定の建築構成要素を含むかどうか」を判別する方法を「畳」と「格子」を例として，二つ提案する。
提案する一つ目の方法は，特定の建築構成要素を含む画像は，画像内の大まかな色分布と構図も概ね似通っているという考え方を基礎に，パターン認識と類似画像検索の技術を応用することによって，画像内の色分布と構図情報の類似度を計算して判別を行おうとするものである。類似画像検索の研究において画像間の類似度として，建築画像を探す際に重要な観点や要素である,「色」と「構図」を考慮した研究も存在する。しかし，それらの研究では構図情報の抽出に複雑な領域分割処理を行うものが多く，処理に手間がかかる上，必ずしも建築的な構図を反映したものではなかった。そこで本研究では，画像内の画素の位置情報が建築画像の構図にとって重要であると考え，粗く切った画像において「色と構図の情報」を考慮した類似度を，最小費用流問題によって計算する方法を提案する。
二つ目は，画像認識技術を用いて特定の構成要素に対して，その判別に重要な役割を果た

す特徴量を抽出することにより，構成要素の存在の有無に関する学習モデルを構築し，判別を行う方法である。
判別精度は 7,8 割とまだ実用レベルには達していないが，既に実用化されている文字認識などと比較した場合に，建築構成要素のもつ構図や対象の多様性を考慮すると，研究初期段階の目的は達したものと考えられる。

6. M. Ohsaki and N. Katoh:
"Topology optimization of trusses with stress and local constraints", Structural and Multidisciplinary Optimization, 29, 190-197, 2005.3.
概要: A truss topology optimization problem under stress constraints is formulated as a Mixed Integer Programming (MIP) problem with variables indicating existence of nodes and members. The local constraints on nodal stability and intersection of members are considered, and a moderately large lower bound is given for the cross-sectional area of an existing member. A lower-bound objective value is found by neglecting the compatibility conditions, where linear programming problems are successively solved based on a branch-and-bound method. An upper-bound solution is obtained as a solution of NonLinear Programming (NLP) problem for the topology satisfying the local constraints. It is shown in the examples that upper- and lower-bound solutions with small gap in the objective value can be found by the branch-and-bound method, and the computational cost can be reduced by using the local constraints for obtaining a practically optimal solution.

7. Tetsuo Asano, Naoki Katoh, Hisao Tamaki and Takeshi Tokuyama:
"The structure and number of global roundings of a graph", Theoretical Computer Science, 325 (5) , 425-437, 2004.
概要: Given a connected weighted graph $G = (V, E)$, we consider a hypergraph $H_G = (V, P_G)$ corresponding to the set of all shortest paths in G. For a given real assignment $\mathbf{a}$ on V satisfying $0 \leq \mathbf{a}(v) \leq 1$, a global rounding α with respect to H_G is a binary assignment satisfying that $|\sum_{v \in F} \mathbf{a}(v) - \alpha(v)| < 1$ for every $F \in P_G$. We conjecture that there are at most $|V|+1$ global roundings for H_G, and also the set of global roundings is an affine independent set. We give several positive evidences for the conjecture.

8. Tetsuo Asano, Mark de Berg, Otfried, Cheong, Hazel Everett, Herman Haverkort, Naoki Katoh, and Alexander Wolff:
"Optimal Spanners for Axis-Aligned Buildings", Computational Geometry: Theory and Applications, 30(1), 59-77, 2005.
概要: We consider the following problem: We wish to connect buildings using bridges or tunnels, such that people can pass between the buildings without having to step outside. We want to place the bridges such that the dilation of the resulting network is minimum, where the dilation is the maximum ratio between distance within the network and L_1 distance over all pairs of points in the given buildings. Given n axis-parallel rectangular buildings and a set of pairs of buildings that are to be connected by axisparallel bridges, we can find a minimum dilation bridge placement in $O(n^2)$ time if the building-bridge graph is a path. For the case that the graph is a tree, we give a pseudo-polynomial solution based on an LP-formulation of the problem. We show that the problem is NP-hard if the graph has cycles.

9. M. Ohsaki and N. Katoh:
"Topology optimization of trusses with stress and local constraints", Structural and Multidisciplinary Optimization, 29, 190-197, 2005.
概要: A truss topology optimization prob-

lem under stress constraints is formulated as a Mixed Integer Programming (MIP) problem with variables indicating existence of nodes and members. The local constraints on nodal stability and intersection of members are considered, and a moderately large lower bound is given for the cross-sectional area of an existing member. An upper-bound solution is obtained as a solution of NonLinear Programming (NLP) problem for the topology satisfying the local constraints. It is shown in the examples that upper- and lower-bound solutions with small gap in the objective value can be found by the branch-and-bound method, and the computational cost can be reduced by using the local constraints for obtaining a practically optimal solution.

10. M. Ohsaki and S. Nishiwaki:
"Shape design of pin-jointed multi-stable compliant mechanisms by utilizing snapthrough behavior", Structural and Multidisciplinary Optimization, 30, 327-334, 2005.
概要: A general approach is presented for generating pin-jointed multi-stable compliant mechanisms utilizing snapthrough behavior. An optimization problem is formulated for minimizing the total structural volume under constraints on the displacements at the specified nodes, stiffnesses at initial and final states, and load factors to lead to snapthrough behavior. The design variables are cross-sectional areas and the nodal coordinates. It is shown in the numerical examples that several mechanisms can be naturally found as a result of optimization starting from randomly selected initial solutions. It is also shown that no local bifurcation point exists along the equilibrium path, and the obtained mechanism is not sensitive to initial imperfections.

11. J. Zhang and M. Ohsaki:
"Form-finding of self-stressed structures by an extended force density method", J. Int. Assoc. for Shell and Spatial Struct., 46(3), 159-166, 2005.
概要: A numerical method for the form-finding problem of self-stressed structures is presented as an extension of the basic idea and formulations of the force density method, which is initially developed for cable nets. Singular value decomposition of the equilibrium matrix with respect to nodal coordinates is utilized to find the feasible set of force densities for satisfying the non-degeneracy condition of the structure. A unique configuration of the structure can be obtained by specifying an independent set of nodal coordinates. The proposed method can have some, although not exact or direct, controls over the geometrical and mechanical properties of the structure. Its strong ability of searching new configurations and the good convergence properties are demonstrated by several numerical examples.

12. X. Gandibleux, H. Morita, N. Katoh:
"Population-based heuristic for solving assignment problems with two objectives", Journal of Mathematical Modelling and Algorithms, 掲載予定.
概要: A population-based heuristic for computing an approximation of the efficient solution set for the assignment problem with two objectives is presented. It is based on an intensive use of three operators performed on a population composed only of elite solutions. The operators are a local search, a crossover and a path-relinking. The initial population is a subset of supported solutions. A genetic information is derived from elite solutions. A lower bound set defined in the objective space provides an acceptable limit for performing a local search. Results of extensive numerical ex-

periments show that the heuristic is able to find quickly a very good approximation of the efficient frontier and outperforms the previous results published.

13. Naoki Katoh and Taihei Yano:
"An Approximation Algorithm for the Pickup and Delivery Vehicle Routing Problem on Trees", Discrete Applied Mathematics, 154(16), 2235-2349, 2006.
概要: This paper presents an approximation algorithm for a vehicle routing problem on a tree-shaped network with a single depot where there are two types of demands, pickup demand and delivery demand. Each customer has a positive demand of pickup and/or delivery. Demands of customers are served by a fleet of identical vehicles with unit capacity. Each vehicle can serve pickup and delivery demands. It is assumed that the demand of a customer is splittable, i.e., it can be served by more than one vehicle. The problem asks to find a set of tours of the vehicles with minimum total lengths. In each tour, a vehicle begins at the depot with certain amount of goods for delivery, visits a subset of the customers in order to deliver and pick up goods and returns to the depot. At any time during the tour, a vehicle must always satisfy the capacity constraint. We propose a 2-approximation algorithm for the problem.

14. 具 源龍，横田 隆志，瀧澤 重志，加藤 直樹:
"建築画像からの建築構成要素の認識手法に関する研究 ベイズ判別によるガラス開口部の認識", 日本建築学会総合論文誌, 第4号 情報化の視点からみた建築・都市のフロンティア, 123-127, 2006.2.
概要: 本研究では，デジタルカメラで撮影した住宅のファサード画像から，ガラス開口部を判別するモデルを構築し，その判別精度を検証するものである。ガラス開口部は，建物火災の延焼シミュレーションや，環境防犯設計において重要な因子となるものであるが，大量に存在するので, その位置を自動的に認識する必要がある。ガラスは透明性や反射性などの素材の持つ特徴から，他の建築部位と比較して色分布が安定しておらず，既往研究で対象とした畳や格子と比較して判別が難しいと考えられる。そこで，色情報だけでなく開口部が出現しやすい位置情報を加味するとともに，Naive Bayes を非独立な確率変数を扱えるように拡張した手法を新たに開発して判別モデルを構築した。

15. 神山直之，瀧澤重志，加藤直樹:
"フロアプランの列挙に基づく 2 階建て住宅の室配置アルゴリズム", 日本建築学会環境系論文集, 601, 65-72, 2006.3.
概要: 本研究は，2 階建て住宅の LDK 形式で表現される室配置を対象として，列挙アルゴリズムと最適化アルゴリズムを併用して，その配置を決定するアルゴリズムを提案するものである。概要を説明すると，まず，中野によって提案された高速なフロアプラン列挙アルゴリズムを用いてフロアプランを生成し，それらから隣接関係と方位条件，階を貫く室（階段室）の整合性を満たす室配置を抽出する。そして，タブー探索法を用いて室の大きさと形状を最適化し，要求に沿う室配置を複数求めようとするものである。2 階建て住宅を対象とし，吹き抜けなどの複数階で共有する空間の制約を利用して，列挙解の組み合わせを減少させる工夫をしている点にアルゴリズム面の特徴がある。

16. J. Zhang, M. Ohsaki and Y. Kanno:
"A direct approach to design of geometry and forces of tensegrity structures", Int. J. Solids and Struct., 43(7), 2260-2278, 2006.
概要: In the process of designing a tensegrity system, some constraints are usually introduced for geometry and/or forces to ensure uniqueness of the solution, because the tensegrity systems are underdetermined in most cases. In this paper, a new approach is presented to enable designers to specify independent sets of axial forces and nodal coordinates consecutively, under the

equilibrium conditions and the given constraints, to satisfy the distinctly different requirements of architects and structural engineers. The proposed method can be used very efficiently for practical applications because only linear algebraic equations are to be solved, and no equation of kinematics or material property is needed. Some numerical examples are given to show not only efficiency of the proposed method but also its ability of searching new configurations.

17. Shin-ichi Tanigawa and Naoki Katoh:
"Finding a Triangular Mesh with a Constant Number of Different Edge Lengths", IEICE TRANSACTIONS on Information and Systems, E89-D(8), 2364-2371, 2006.
概要: We consider the problem of triangulating an x-monotone polygon with a small number of different edge lengths using Steiner points. Given a parameter α, where $0 < \alpha < 1$, we shall present an algorithm for finding an almost uniform triangular mesh with $3\pi/8\alpha^2+o(1/\alpha^2)$ different edge lengths such that every edge length is between $\underline{l}$ and $(2+\sqrt{2}\alpha)\underline{l}$. Experiments demonstrate the effectiveness of this algorithm.

18. Naoyuki Kamiyama, Katoh Naoki, Astushi Takizawa:
"An Efficient algorithm for Evacuation Problem in Dynamic Network Flows with Uniform Arc Capacity", IEICE Transaction on Fundamentals, E89-D(8), 2372-2379, 2006.
概要: In this paper, we consider the evacuation problem in a network which consists of a directed graph with capacities and transit times on its arcs. We present an $O(n \log n)$ time algorithm for the evacuation problem in a network of grid structure with uniform arc capacity which has a single sink where n is the number of vertices in the network.

19. Sachio Teramoto, Tetsuo Asano, Benjamin Doerr, Naoki Katoh:
"Inserting Points Uniformly at Every Instance, The Institute of Electronics", IEICE Transaction on Fundamentals, E89-D(8), 2348-2356, 2006.
概要: A problem of arranging n points as uniformly as possible, which is equivalent to that of packing n equal and non-overlapping circles in a unit square, is frequently asked. In this paper we generalize this problem in such a way that points be inserted one by one with uniformity preserved at every instance. Our criteria on uniformity is to minimize the gap ratio (which is the maximum gap over the minimum gap) at every point insertion.We present a linear time algorithm for finding an optimal n-point sequence with the maximum gap ratio bounded by in the 1-dimensional case. We describe how hard the same problem is for a point set in the plane and propose a local search heuristics for finding a good solution.

20. Boris Aronov, Tetsuo Asano, Naoki Katoh, Kurt Mehlhorn:
"Polyline Fitting of Planar Points under Min-Sum Criteria", International Journal of Computational Geometry & Applications, Vol.16, No.2&3, 97-116, 2006.
概要: Fitting a curve of a certain type to a given set of points in the plane is a basic problem in statistics and has numerous applications. We consider fitting a polyline with k joints under the min-sum criteria with respect to L_1- and L_2-metrics, which are more appropriate measures than uniform and Hausdorff metrics in statistical context. We present efficient algorithms for the 1-joint versions of the problem and fully polynomial-time approximation schemes for the general k-joint versions.

21. J. Y. Zhang and M. Ohsaki:
"Adaptive force density method for form-finding problem of tensegrity structures",

Int. J. Solids and Struct., 43, 5658-5673, 2006.

概要: A numerical method is presented for form-finding of tensegrity structures. Eigenvalue analysis and spectral decomposition are carried out iteratively to find the feasible set of force densities that satisfies the requirement on rank deficiency of the equilibrium matrix with respect to the nodal coordinates. The equilibrium matrix is shown to correspond to the geometrical stiffness matrix in the conventional finite element formulation. A unique and non-degenerate configuration of the structure can then be obtained by specifying an independent set of nodal coordinates. A simple explanation is given for the required rank deficiency of the equilibrium matrix that leads to a non-degenerate structure. Several numerical examples are presented to illustrate the robustness as well as the strong ability of searching new configurations of the proposed method.

22. M. Ohsaki and J. Y. Zhang:
"Stability conditions of prestressed pin-jointed structures", Int. J. Non-Linear Mech., 41, 1109–1117, 2006.

概要: Pin-jointed structures are first classified to trusses, tensile structures, and tensegrity structures in view of their respective stability properties. A sufficient condition for stability of an equilibrium state is derived for tensegrity structures. The condition is based on the bilinear forms of the linear and geometrical stiffness matrices considering the flexibility of members. The stability is defined by the positive definiteness of the tangent stiffness matrix, whereas the definition of prestress-stability is based on the geometrical stiffness matrix and the infinitesimal mechanisms. Numerical examples verify that the so-called super-stability condition might not be satisfied by a stable tensegrity structure, and that a prestress-stable structure can be unstable if the prestresses are moderately large.

23. S. Fujishige and A. Tamura:
"A general two-sided matching market with discrete concave utility functions", Discrete Applied Mathematics, 154, 950-970, 2006.3.

概要: In the theory of two-sided matching markets there are two standard models: (i) the marriage model due to Gale and Shapley and (ii) the assignment model due to Shapley and Shubik. Recently, Eriksson and Karlander introduced a hybrid model, which was further generalized by Sotomayor. In this paper, we propose a common generalization of these models by utilizing the framework of discrete convex analysis introduced by Murota, and verify the existence of a pairwise-stable outcome in our general model.

24. Shin-ichi Tanigawa and Naoki Katoh:
"Polygonal Curve Approximation Using Grid Points with Application to a Triangular Mesh Generation with Small Number of Different Edge Lengths", Proc. 2nd International Conference on Algorithmic Aspects in Information and Management (AAIM 2006), Hong Kong, LNCS 4041, 161-172, 2006.

概要: For a given x-monotone polygonal curve each of whose edge lengths is between $\underline{l}$ and $2\underline{l}$, we consider the problem of approximating it by another x-monotone polygonal curve using points of a square grid so that there exists a small number of different edge lengths and every edge length is between $\underline{l}$ and $\beta\underline{l}$, where β is a given parameter satisfying $1 \leq \beta \leq 2$. Our first algorithm computes an approximate polygonal curve using fixed square grid points in $O((n/\alpha^4)\log(n/\alpha))$ time. Based on this, our second algorithm finds an approximate polygonal curve as well as an optimal grid placement simultaneously in

$O((n^3/\alpha^{12})\log^2(n/\alpha))$ time, where α is a parameter that controls the closeness of approximation. Based on the approximate polygonal curve, we shall give an algorithm for finding a uniform triangular mesh for an x-monotone polygon with a constant number of different edge lengths.

25. S. Fujishige and A. Tamura:
"A two-sided discrete-concave market with possibly bounded side payments: An approach by discrete convex analysis", Mathematics of Operations Research, 32, 2007.3.
概要: The marriage model due to Gale and Shapley and the assignment model due to Shapley and Shubik are standard in the theory of two-sided matching markets. We give a common generalization of these models by utilizing discrete concave functions and considering possibly bounded side payments. We show the existence of a pairwise stable outcome in our model. Our present model is a further natural extension of the model examined in our previous paper (Fujishige and Tamura (2006)), and the proof of the existence of a pairwise stable outcome is even simpler than the previous one.

26. David Avis, Naoki Katoh, Makoto Ohsaki, Ileana Streinu and Shin-ichi Tanigawa:
"Enumerating Non-crossing Minimally Rigid Frameworks", Graphs and Combinatorics, 23[sup], 117-134, 2007.
概要: In this paper we present an algorithm for enumerating without repetitions all the non-crossing generically minimally rigid bar-and-joint frameworks (simply called non-crossing Laman frameworks) on a given generic set of n points. Our algorithm is based on the reverse search paradigm of Avis and Fukuda. It generates each output graph in $O(n^4)$ time and $O(n)$ space, or, with a slightly different implementation, in $O(n^3)$ time and $O(n^2)$ space. In particular, we obtain that the set of all non-crossing Laman frameworks on a given point set is connected by flips which remove an edge and then restore the Laman property with the addition of a non-crossing edge.

27. Yinfeng Xu, Wenqiang Dai, Naoki Katoh, Makoto Ohsaki:
"Trinagulating a conve polygon with fewer number of non-standard bars", Theoretical Computer Science, 389, 143-151, 2007.
概要: For a given convex polygon with inner angle no less than $\frac{2}{3}\pi$ and boundary edge bounded by $[l, \alpha l]$ for $1 \leq \alpha \leq 1.4$, where l is a given standard bar's length, we investigate the problem of triangulating the polygon using some Steiner points such that (i) the length of each edge in triangulation is bounded by $[\beta l, 2l]$, where β is a given constant and meets $0 < \beta \leq \frac{1}{2}$, and (ii) the number of non-standard bars in the triangulation is minimum. This problem is motivated by practical applications and has not been studied previously. In this paper, we present a heuristic to solve the above problem, which is based on the heuristic to generate a triangular mesh with less number of non-standard bars and shorter maximal edge length, and a process to make the length of each edge lower bounded. Our procedure is simple and easily implemented for this problem, and we prove that it has good performance guaranteed.

28. Naoyuki Kamiyama, Naoki Katoh, Atsushi Takizawa:
"An Efficient Algorithm for the Evacuation Problem in a Certain Class of a Network with Uniform Path-Lengths", Proceedings of the third International Conference on Algorithmic Aspect in Information and Management (AAIM2007), Lecture Notes in Computer Science 4508, pp.178-190, 2007.
概要: In this paper, we consider the evacuation problem for a network which consists of a directed graph with capacities and

transit times on its arcs. This problem can be solved by the algorithm of Hoppe and Tardos in polynomial time. However their running time is high-order polynomial, and hence is not practical in general. Thus it is necessary to devise a faster algorithm for a tractable and practically useful subclass of this problem. In this paper, we consider a dynamic network with a single sink s such that (i) for each vertex v the sum of transit times of arcs on any path from v to s takes the same value, and (ii) for each vertex v the minimum v-s cut is determined by the arcs incident to s whose tails are reachable from v. We propose an efficient algorithm for this network problem. This class of networks is a generalization of the grid network studied in our previous paper.

29. J. Y. Zhang and M. Ohsaki:
"Stability conditions for tensegrity structures", Int. J. Solids and Struct., 44(11–12), 3875–3886, 2007.
概要: Stability conditions for tensegrity structures are derived based on positive definiteness of the tangent stiffness matrix, which is the sum of the linear and geometrical stiffness matrices. A necessary stability condition is presented by considering the affine motions that lie in the null-space of the geometrical stiffness matrix. The condition is demonstrated to be equivalent to that derived from the mathematical rigidity theory so as to resolve the discrepancy between the stability theories in the fields of engineering and mathematics. Furthermore, it is shown that the structure is guaranteed to be stable, if the structure satisfies the necessary stability condition and the geometrical stiffness matrix is positive semidefinite with the minimum rank deficiency for non-degeneracy.

30. 大崎 純, 西脇眞二:
"幾何学的非線形性を考慮したトラスの形状・トポロジー最適化によるリンク機構の生成", 日本機械学会論文集 A, 73(729), 659–665, 2007.
概要: A two-stage general optimization approach is presented for generating link mechanisms from a highly connected ground structure. The structure is modeled as a pin-jointed truss, and is to be optimized so that a large displacement is generated in the specified direction at the output node. The design variables are the cross-sectional areas of the members and the nodal locations. The equilibrium path of an unstable mechanism is traced by the displacement control method. In the first step, the unnecessary members are removed by solving the optimization problem for minimizing the total structural volume under constraints on the maximum load, the displacement at the specified node, and the stiffnesses at the initial and final states. In the second step, the deviation of the displacement of the output node from the specified direction is minimized. It is shown in the numerical examples that several mechanisms can be naturally found as a result of the two-stage optimization starting from randomly selected initial solutions.

31. Naoyuki Kamiyama, Naoki Katoh, Atsushi Takizawa:
"Arc-disjoint In-trees in Directed Graphs", Proceedings of the nineteenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA2008), pp.518-526, 2008.
概要: Given a directed graph $D = (V, A)$ and a set of specified vertices $S = \{s_1, \ldots, s_d\} \subseteq V$ with $|S| = d$ and a function $f\colon S \to \mathbb{N}$ where $\mathbb{N}$ denotes the set of natural numbers, we present a necessary and sufficient condition that there exist $\sum_{s_i \in S} f(s_i)$ arc-disjoint in-trees denoted by $T_{i,1}, T_{i,2}, \ldots, T_{i,f(s_i)}$ for every $i = 1, \ldots, d$ such that $T_{i,1}, \ldots, T_{i,f(s_i)}$ are rooted at s_i and each $T_{i,j}$ spans vertices from which s_i is

reachable. This generalizes the result of Edmonds, i.e., the necessary and sufficient condition that for a directed graph $D = (V, A)$ with a specified vertex $s \in V$, there are k arc-disjoint in-trees rooted at s each of which spans V. Furthermore, we extend another characterization of packing in-trees of Edmonds to the one in our case.

32. David Avis, Naoki Katoh, Makoto Ohsaki, Ileana Streinu and Shin-ichi Tanigawa:
"Enumerating Constrained Non-crossing Minimally Rigid Frameworks", Discrete & Computational Geometry, to appear, 2008.
概要: In this paper we present an algorithm for enumerating without repetitions all the non-crossing generically minimally rigid bar-and-joint frameworks under edge constraints (also called constrained non-crossing Laman frameworks) on a given generic set of n points. Our algorithm is based on the reverse search paradigm of Avis and Fukuda. It generates each output graph in $O(n^4)$ time and $O(n)$ space, or, slightly different implementation, in $O(n^3)$ time and $O(n^2)$ space. In particular, we obtain that the set of all the constrained non-crossing Laman frameworks on a given point set is connected by flips which restore the Laman property.

研究会等

1. Tetsuo Asano, Naoki Katoh, Hisao Tamaki, Takeshi Tokuyama:
"On Geometric Structure of Global Roundings for Graphs and Range Spaces", Proc. of 9th Scandinavian Workshop on Algorithm Theory, LNCS 3111, Springer, 455-467, 2004.7.
概要: Given a connected weighted graph $G = (V, E)$, we consider a hypergraph $H_G = (V, P_G)$ corresponding to the set of all shortest paths in G. For a given real assignment $\mathbf{a}$ on V satisfying $0 \leq \mathbf{a}(v) \leq 1$, a global rounding α with respect to H_G is a binary assignment satisfying that $|\sum_{v \in F} \mathbf{a}(v) - \alpha(v)| < 1$ for every $F \in P_G$. We conjecture that there are at most $|V|+1$ global roundings for H_G, and also the set of global roundings is an affine independent set. We give several positive evidences for the conjecture.

2. M. Ohsaki and S. Nishiwaki:
"Shape design of bistable compliant mechanisms by utilizong snapthrough behavior", Proc. Third China-Japan-Korea Joint Symposium on Optimization of Structural and Mechanical Systems (CJK-OSM3), 359-364, 2004.10.
概要: A general approach is presented for generating compliant mechanisms from a highly connected ground structure. The structure is modeled as a pin-jointed truss, and geometrical nonlinearity is considered. An optimization problem is formulated for minimizing the total structural volume under constraints on the displacement at the specified node, and stiffnesses at initial and final states, where the design variables are cross-sectional areas and the nodal coordinates. It is shown in the numerical examples that several mechanisms can be naturally found as a result of optimization starting from randomly selected initial solutions.

3. N. Katoh, M. Ohsaki and Y. Xu:
"A uniform triangle mesh generation of curved surfaces", Proc. Third China-Japan-Korea Joint Symposium on Optimization of Structural and Mechanical Systems (CJK-OSM3), 759-764, 2004.11.
概要: The purpose of this study is to maximize the number of edges with almost equal lengths for a triangular mesh of a curved surface, so that the distortion of the triangulation used for finite element model or architectural latticed roofs is to be mini-

mized. Two approximation algorithms are presented for triangulation of a polygon in a plane under upper and lower bounds on the edge lengths; (a) minimize the number of internal node, and (b) minimize the number of different edge lengths. For a parametric surface, the domain on a parametric surface is first mapped to a plane to obtain a polygon. The algorithm (b) is applied to the polygon to find a triangulation, which is remapped to the surface. The performance of the proposed method is demonstrated by using a tensor product Bézier surface.

4. Boris Aronov, Tetsuo Asano, Naoki Katoh, Kurt Mehlhorn, Takeshi Tokuyama:
"Polyline Fitting of Planar Points under Min-Sum Criteria", Proc. of 15th Symposium on Algorithms and Computations, LNCS 3341, Springer Verlag, 77-88, 2004.12.
概要: Fitting a curve of a certain type to a given set of points in the plane is a basic problem in statistics and has numerous applications. We consider fitting a polyline with k joints under the min-sum criteria with respect to L_1- and L_2-metrics, which are more appropriate measures than uniform and Hausdorff metrics in statistical context. We present efficient algorithms for the 1-joint versions of the problem, and fully polynomial-time approximation schemes for the general k-joint versions.

5. M. Ohsaki and N. Katoh:
"Topology optimization of trusses with stress and local constraints", Proc. 5th Int. Conf. on Computation of Shell and Spatial Structures, Salzburg, IASS-IACM, 2005.6.
概要: A mixed integer programming formulation is presented for the truss topology optimization with stress and local constraints. Linear programming problems are successively solved based on a branch-and-bound method, where an upper-bound solution is obtained by solving nonlinear programming problems. It is shown in the examples that upper- and lower-bound solutions with small objective gap can be found, and the computational cost can be reduced by utilizing the local constraints.

6. Shin-ichi Tanigawa and Naoki Katoh:
"Finding a triangular mesh with a constant number of different edge lengths", Proc. of 17th Canadian Conference of Computational Geometry, 271-274, 2005.8.
概要: The problem of triangulating a specified domain is a basic problem in computational geometry and has numerous applications. We consider the problem of finding an almost uniform triangular mesh for a given x-monotone polygon with a constant number of different edge lengths. This problem is motivated by constructing structure truss such as a dome with a limited number of different elements to reduce its construction cost. Though it is difficult in general to find a triangular mesh for any polygon with a constant number of different edge lengths, we give an algorithm for finding such a triangular mesh by approximating a given boundary using square grid points.

7. M. Ohsaki, J. Zhang and S. Kimura:
"An Optimization Approach to Design of Geometry and Forces of Tensegrities", Proc. IASS Symposium 2005, Bucharest, Romania, Int. Assoc. Shell and Spatial Struct., 603-610, 2005.9.
概要: Ageneral method is presented for direct design of member directions, force vectors and nodal locations of tensegrities with given topology, which is defined as a directed graph. The system of equilibrium equations is written in terms of the components of the member force vectors by using the incidence matrix. Since the number of the total components of the force vectors

is usually larger than the rank of the equilibrium matrix, some force components can be arbitrary specified. The target values are given here to formulate an optimization problem to obtain a set of force components that satisfies equilibrium conditions. By using the proposed method, the designers can directly control the axial forces and the configuration simultaneously, which is considered a major advantage of the method. A new configuration can also be obtained by changing the target values of the forces and geometrical constraints.

8. J. Zhang and M. Ohsaki:
"Form-finding of Self-stressed Structures by an Extended Force Density Method", Proc. IASS Symposium 2005, Bucharest, Romania, Int. Assoc. Shell and Spatial Struct., 93-100, 2005.9.
概要: The force density method is extended to find the feasible set of force densities, based on singular value decomposition of equilibrium matrix, and an algorithm for specifying an independent set of nodal coordinates is proposed. The force densities is first formulated in terms of elements of the equilibrium matrix, where constraints on force densities with desired values can be included at the same time. The least square solution of the force densities is then given based on this formulation. Singular value decomposition is next applied to the equilibrium matrix and its smallest elements with the specified number are set to be zero in order to attain the required rank. The equilibrium matrix is reconstructed for calculation of force densities by the equations formulated previously. By iteratively carrying out the above processes of calculating force density vector and equilibrium matrix with the updated ones, a feasible set of force densities is derived. The last step of the method is to specify an independent set of nodal coordinates to achieve a unique geometrical configuration of the structure.

9. T. Kinoshita and M. Ohsaki:
"Multistable compliant mechanisms for spatial structures using snapthrough behavior", Proc. IASS Symposium 2006, Beijing, China, Int. Assoc. Shell and Spatial Struct., .
概要: Contrary to conventional mechanisms, which are usually achieved by the rotations around hinges and/or sliding supports, a compliant mechanism uses the elastic deformation of structural parts to realize the mechanism for shape transformation of the entire structure. We investigate the applicability and effectiveness of compliant mechanism to realize shape transformation of spatial structures. Conventional mechanism achieves stability at undeformed and deformed states by applying additional forces or constraints. However, our compliant mechanism can keep stability at undeformed and deformed states through its own stiffness, because it has multiple stable states without any additional constraint. To achieve practically useful multistable compliant mechanism, structural optimization technique is used. The design variables of the optimization problem are the cross-sectional areas of structural members and the nodal coordinates. Design conditions are given on the magnitude of deformation, the stability before and after deformation. The total structural volume is minimized as the objective function. Optimization problem is formulated so that snapthrough buckling is naturally generated. Numerical tests are carried out to several types of spatial structures. It is shown that the mechanism which achieves the expected deformation can be obtained by solving the optimization problem.

10. David Avis, Naoki Katoh, Makoto Ohsaki, Ileana Streinu and Shin-ichi Tanigawa:

"Enumerating Non-crossing Minimally Rigid Frameworks", Proc. 9th Japan-Korea Joint Workshop on Algorithms and Computation, Sapporo, 73-78, 2006.
概要: In this paper we present an algorithm for enumerating without repetitions all the non-crossing generically minimally rigid bar-and-joint frameworks (simply called non-crossing Laman frameworks) on a given generic set of n points. Our algorithm is based on the reverse search paradigm of Avis and Fukuda. It generates each output graph in $O(n^4)$ time and $O(n)$ space, or, with a slightly different implementation, in $O(n^3)$ time and $O(n^2)$ space. In particular, we obtain that the set of all non-crossing Laman frameworks xson a given point set is connected by flips which remove an edge and then restore the Laman property with the addition of a non-crossing edge.

11. David Avis, Naoki Katoh, Makoto Ohsaki, Ileana Streinu and Shin-ichi Tanigawa:
"Enumerating Non-crossing Minimally Rigid Frameworks", Proc. 12th International Computing and Combinatorics Conference (COCOON 2006), Taipei, LNCS 4112, 205-215, 2006.
概要: In this paper we present an algorithm for enumerating without repetitions all the non-crossing generically minimally rigid bar-and-joint frameworks (simply called non-crossing Laman frameworks) on a given generic set of n points. Our algorithm is based on the reverse search paradigm of Avis and Fukuda. It generates each output graph in $O(n^4)$ time and $O(n)$ space, or, with a slightly different implementation, in $O(n^3)$ time and $O(n^2)$ space. In particular, we obtain that the set of all non-crossing Laman frameworks xson a given point set is connected by flips which remove an edge and then restore the Laman property with the addition of a non-crossing edge.

12. Naoki Katoh, Makoto Ohsaki, Takuya Kinoshita, Shin-ichi Tanigawa, David Avis and Ileana Streinu:
"Enumeration of Optimal Pin-Jointed Bistable Compliant Mechanisms", Proc. 4th China-Japan-Korea Joint Symposium on Optimization of Structural and Mechanical Systems (CJK-OSM4), Kunming, China, 2006.
概要: In this paper we present an algorithm for enumerating without repetitions all the non-crossing generically minimally rigid bar-and-joint frameworks (simply called non-crossing Laman frameworks) on a given generic set of n points. Our algorithm is based on the reverse search paradigm of Avis and Fukuda. It generates each output graph in $O(n^4)$ time and $O(n)$ space, or, with a slightly different implementation, in $O(n^3)$ time and $O(n^2)$ space. In particular, we obtain that the set of all non-crossing Laman frameworks on a given point set is connected by flips which remove an edge and then restore the Laman property with the addition of a non-crossing edge.

13. Naoyuki Kamiyama, Katoh Naoki, Astushi Takizawa:
"An Efficient algorithm for Evacuation Problem in Dynamic Network Flows with Uniform Arc Capacity", The Second International Conference on Algorithmic Aspect in Information and Management (AAIM2006), Hong Kong, China, LNCS4041, 231-242, 2006.
概要: In this paper, we consider the evacuation problem in a network which consists of a directed graph with capacities and transit times on its arcs. We present an $O(n \log n)$ time algorithm for the evacuation problem in a network of grid structure with uniform arc capacity which has a single sink where n is the number of vertices in the network.

14. Naoyuki Kamiyama , Katoh Naoki , Atsushi

Takizawa:
"An Efficient algorithm for Evacuation Problem in a Certain Class of Layered Networks", The 9th Japan-Korea Joint Workshop on Algorithms and Computation (WAAC2006), pp.119-126, 2006.
概要: In our previous paper, we proposed an $O(n \log n)$ time algorithm for the evacuation problem in a grid network with uniform arc capacity. In this paper, we extend the classes of networks to which we can apply the algorithm of our previous paper.

15. Tetsuo Asano, Hisao Tamaki, Naoki Katoh, Takeshi Tokuyama:
"Angular Voronoi Diagram with Applications", The 3rd International Conference on Voronoi Diagrams in Science and Engineering (ISVD 2006), Banff, Canada, IEEE Computer Society, 18-24, 2006.
概要: we define a new Voronoi diagram for a set of line segments in the plane, which is called an aspect-ratio Voronoi diagram. We show several interesting properties of this diagram. it looks quite different from an ordinary Voronoi diagram defined for a set of points. Voronoi edges consist of plane curves of degree-4 polynomial in x and y. A Voronoi cell associated with a line segment is not always connected. It may be divided into a number of connected regions. This fact causes high complexity of the diagram.

16. 加藤直樹:
"三角形分割とラーマングラフ：建築への応用", Proceedings of Eighteenth Research Association Mathematical Programming Symposium, 135-152, 2006 年 10 月.
概要: 部材がジョイントでつながっているトラス構造物はグラフ構造で表現される．トラス構造物が安定であるための必要最低限の部材から成る場合, 静定構造物と呼ばれている．静定構造物に対応するグラフは静定グラフとして知られている．本論文では平面上のトラス構造物を対象として, あたえられた平面上の点集合 P に対して, P 上に埋め込まれた, 辺が交差しない静定グラフをすべて列挙するアルゴリズムについて述べる．静定グラフはラーマングラフとも呼ばれ，その組合せ的性質については多くの研究者によって研究されており，そのようなグラフ全体がマトロイドの基を作ることが知られている．したがって，逆探索法によりすべてのラーマングラフを列挙することができる．一方，平面上に埋め込まれた無交差ラーマングラフ全体はマトロイドとはならない．最近，無交差ラーマングラフに対して，グラフ構造や幾何的構造を明らかにし、無交差ラーマングラフ同士に対して隣接関係を上手に定めることにより逆探索法によって，無交差ラーマングラフをすべて列挙するアルゴリズムを開発した．あらかじめ用いる枝が一部指定されている場合も取り扱うことができる．本講演ではその概要と構造最適化への応用について述べる．

17. J. Y. Zhang and M. Ohsaki:
"Multiobjective optimization for force design of tensegrity structures", Proc. Fourth China-Japan-Korea Joint Symposium on Optimization of Structural and Mechanical Systems (CJK-OSM4), 485–490, 2006.11.
概要: In this paper, we present a multiobjective optimization approach for force design of tensegrity structures, where their geometries are assumed to be determined a priori; we search for the optimal distribution of member forces that leads to the maximum stiffness, and that is closest to the target values assigned by the designers as well. The Pareto optimal solutions of the problem are presented for a tensegrity grid as an example to demonstrate the methodology

18. Naoyuki Kamiyama, Naoki Katoh and Atsushi Takizawa:
"Generalization of Theorem of Edmonds", Proceedings of Kyoto International Conference on Computational Geometry and

Graph Theory (KyotoCGGT2007), pp.195-196, 2007.

概要: Given a directed graph $D = (V, A)$ and a set of specified vertices $S = \{s_1, \ldots, s_d\} \subseteq V$ with $|S| = d$ and a function $f\colon S \to \mathbb{N}$ where $\mathbb{N}$ denotes the set of natural numbers, we present a necessary and sufficient condition that there exist $\sum_{s_i \in S} f(s_i)$ arc-disjoint in-trees denoted by $T_{i,1}, T_{i,2}, \ldots, T_{i,f(s_i)}$ for every $i = 1, \ldots, d$ such that $T_{i,1}, \ldots, T_{i,f(s_i)}$ are rooted at s_i and each $T_{i,j}$ spans vertices from which s_i is reachable. This generalizes the result of Edmonds, i.e., the necessary and sufficient condition that for a directed graph $D = (V, A)$ with a specified vertex $s \in V$, there are k arc-disjoint in-trees rooted at s each of which spans V. Furthermore, we extend another characterization of packing in-trees of Edmonds to the one in our case.

19. Naoyuki Kamiyama, Katoh Naoki, Atsushi Takizawa:
"Theoretical and Practical Issues of Evacuation Planning in Urban Areas", The 8th Hellenic European Research on Computer Mathematics & Conference its Applications (HERCMA 2007), pp.49-50, 2007.

概要: In the evacuation problem by using dynamic network flow, we need to consider several other features from the viewpoint of actual evacuation planning. (a) All the supplies at the same vertex flow through a single path from each vertex to a sink. Moreover, these evacuation paths for different vertices should not cross each other. (b)Theoretically, it is know that storage at intermediate vertices is unnecessary for an optimal flow of the evacuation problem. However, people at sources can not wait if the outgoing arc is not full. (c) We take into consideration the difference of the transition speed of people. In the model taking the difference of the speed of people, there exist multi-commodities in the network, and arcs have to be given several transit times for each commodity. In our paper, we present the well-define theoretical models in order to solve the above problems, and the results concerning these models.

20. Naoyuki Kamiyama, Naoki Katoh and Atsushi Takizawa:
"An Efficient Algorithm for the Evacuation Problem in a Certain Class of Networks with Uniform Path-Lengths", Proceedings of the sixth international Colloquium on Graphs and Optimisation (GO VI), pp.21, 2007.

概要: In this paper, we consider the evacuation problem for a network which consists of a directed graph with capacities and transit times on its arcs. This problem can be solved by the algorithm of Hoppe and Tardos in polynomial time. However their running time is high-order polynomial, and hence is not practical in general. Thus it is necessary to devise a faster algorithm for a tractable and practically useful subclass of this problem. In this paper, we consider a dynamic network with a single sink s such that (i) for each vertex v the sum of transit times of arcs on any path from v to s takes the same value, and (ii) for each vertex v the minimum v-s cut is determined by the arcs incident to s whose tails are reachable from v. We propose an efficient algorithm for this network problem. This class of networks is a generalization of the grid network studied in our previous paper.

21. Naoki Katoh and Shin-ichi Tanigawa:
"Enumerating Constrained Non-crossing Geometric Spanning Trees", Proc. of 13th International Computing and Combinatorics Conference, LNCS 4598, 243-253, 2007.

概要: In this paper we present an algorithm for enumerating without repetitions all non-crossing geometric spanning trees on a given

set of n points in the plane under edge inclusion constraints (i.e., some edges are required to be included in spanning trees). We will first prove that a set of all edge-constrained non-crossing spanning trees is connected via remove-add flips, based on the constrained smallest indexed triangulation which is obtained by extending the lexicographically ordered triangulation introduced by Bespamyatnikh. More specifically, we prove that all edge-constrained triangulations can be transformed to the smallest indexed triangulation among them by $O(n^2)$ times of greedy flips. Our enumeration algorithm generates each output graph in $O(n^2)$ time and $O(n)$ space based on reverse search technique. This result improves the previous $O(n^3)$ bound for the unconstrained case by factor of $O(n)$. Our algorithm can deal with the edge-constrained case in the same running time.

22. Takenao Taji, Shin-ichi Tanigawa, Naoyuki Kamiyama, Naoki Katoh and Atsushi Takizawa:
"Finding an optimal location of line facility using evolutionary algorithm", Third Asian-Pacific Congress on Computational Mechanics(APCOM2007), pp.129-135, 2007.
概要: In this paper, we consider the problem for determining an optimal location of line facility in a city such as railway system and so on. Line facility is modeled by a spanning tree embedded on the plane whose vertices represent stations and edges represent the rails connecting two stations. Hence, our problem is to find the optimal location of vertices as well as the optimal connection of vertices. A heuristic algorithm for this problem was proposed by Tsukamoto, Katoh and Takizawa. In this paper, we propose a new heuristic method and show the effectiveness of our proposed algorithm through the numerical examples.

23. J. Y. Zhang and M. Ohsaki:
"Optimization methods for force and shape design of tensegrity structures", Proc. 7th World Congress of Structural and Multidisciplinary Optimization (WCSMO7), Seoul, 40–49, 2007.
概要: This paper presents three optimization approaches for force and shape design of tensegrity structures. The first approach is to find the self-equilibrated configuration of a tensegrity structure, by minimizing the difference of strain energy between cables and struts; the second method is to find self-equilibrated configurations for the structures modeled as directed graphs, where the deviation of member forces from target values is to be minimized; and the third method is to find the optimal distribution of member forces so as to let the structure have maximum stiffness as well as uniform member forces.

24. N. Katoh, M. Ohsaki, T. Kinoshita, S. Tanigawa, D. Avis and I. Steinu:
"Enumeration of optimal pin-jointed bistable compliant mechanism", Proc. 7th World Congress of Structural and Multidisciplinary Optimization (WCSMO7), Seoul, 1041–1050, 2007.
概要: An optimization approach is presented for design of compliant mechanisms. In the first stage, the statically determinate trusses containing a given set of nodes are enumerated without repetitions by the graph enumeration algorithm. A statically determinate truss with non-crossing members is regarded as a planar embedded minimally rigid graph, which is also called a non-crossing minimally rigid framework or a non-crossing Laman framework. Constraints are given in the process of graph enumeration such that the existence of necessary members is ensured. In the

second stage, the nodal locations are optimized under mechanical constraints, where the snapthrough behavior that is usually avoided in structural design is extensively utilized to produce bistable compliant mechanisms. In the numerical examples, many bistable compliant mechanisms are generated to show the effectiveness of the proposed method.

学会大会等

1. 加藤直樹，宮高泰匡，瀧澤重志:
“建築画像のデータベースに対する自動索引付けのための建築構成要素の認識手法の開発　畳と格子を例として”, 日本建築学会大会学術講演梗概集, A-2, 539-540, 2004.

2. 神山直行，瀧澤重志，加藤直樹:
“フロアプラン列挙に基づく二層構造物の室配置最適化手法”, 日本建築学会大会学術講演梗概集, A-2, 559-560, 2004.

3. 大崎 純，西脇眞二:
“スナップスルーを利用したバイステーブルコンプライアントメカニズムの形状設計法”, 計算工学講演会論文集, 9, 399-402, 2004.

4. M. Ohsaki:
“Multiple eigenvalues in topology optimization”, International Workshops on Advances in Computational Mechanics, 114, 2004.

5. 大崎 純，西脇眞二:
“バイステーブルコンプライアントメカニズムの形状設計と不整感度解析”, 第 6 回最適化シンポジウム講演論文集，日本機械学会, 257-262, 2004.

6. 谷川眞一，加藤直樹:
“異なる辺長種類数制約下における一様三角形メッシュ生成アルゴリズム”, 形の科学会誌, 20(1), 97-98, 2005.

7. 谷川眞一，加藤直樹:
“必要部材種類数制約下における平面三角形分割手法に関する研究”, 日本建築学会大会学術講演梗概集, A-2, 483-484, 2005.

8. 谷川眞一，加藤直樹:
“グリッド点を用いた多角形近似と三角形メッシュ生成”, 日本応用数理学会 2005 年度会講演予稿集, 318-319, 2005.

9. 川口史恵，塚本智久，加藤直樹，瀧澤重志:
“都市内鉄道の最適ネットワークの形”, 形の科学会誌, 20(1), 81-82, 2005.

10. 塚本智久，加藤直樹，瀧澤重志:
“都市における線的施設配置問題に関する研究”, 日本建築学会大会学術講演梗概集, A-2, 557-558, 2005.

11. 加藤直樹，大原有祐，瀧澤重志:
“幾何学的迂回率最小化問題に関する研究”, 日本建築学会大会学術講演梗概集, A-2, 479-480, 2005.

12. 大原有祐，加藤直樹，瀧澤重志，川口史恵:
“幾何学的迂回率最小化問題に関する研究-迂回率の小さな街路設計への応用-”, 日本応用数理学会 2005 年度会講演予稿集, 316-317, 2005.

13. 木村 駿，大崎 純，張景耀:
“最適化手法によるテンセグリティ構造の軸力・形状決定”, 日本建築学会大会学術講演梗概集 (近畿), B-1, 941-942, 2005.

14. 張 景耀，大崎 純:
“釣合行列の固有値解析によるテンセグリティ構造の形状決定法”, 日本建築学会大会学術講演梗概集 (近畿), B-1, 943-944, 2005.

15. 大崎 純，張 景耀:
“テンセグリティの形状決定と安定性条件”, 日本応用数理学会 2005 年度年会講演予稿集, 78-79, 2005.

16. 大崎 純，張 景耀:
“固有値解析によるテンセグリティの形状設計と安定性条件”, 第 55 回理論応用力学講演会講演論文集, 121-122, 2006.

17. 谷川眞一, 加藤直樹:
“グリッドを用いた多角形近似”, 電子情報通信学会コンピュテーション研究会, comp2006-3, 25-32, 2006.

18. 谷川眞一, 加藤直樹, 大崎純:
“平面無交差静定構造列挙アルゴリズム”, 日本建築学会大会学術講演集梗概集, A-2, 457-458, 2006.

19. David Avis, 加藤直樹, 大崎純, Ileana Streinu, 谷川眞一:
“辺制約付き非交差静定グラフ列挙アルゴリズム”, 情報処理学会アルゴリズム研究会, 2006-AL-108, 33-40, 2006.

20. David Avis, 加藤直樹, 大崎純, Ileana Streinu, 谷川眞一:
“メカニズム最適化のための平面無交差静定構造列挙アルゴリズム”, 第 7 回最適化シンポジウム講演論文集, 209-214, 2006.

21. Naoyuki Kamiyama, Katoh Naoki, Astushi Takizawa:
“An Efficient algorithm for Evacuation Problem in Dynamic Network Flows with Uniform Arc Capacity”, 情報処理学会 COMP 研究会, 京都大学吉田キャンパス, 41-48, 2006.

22. 神山直之, 加藤直樹, 瀧澤重志:
“辺の容量が一定のネットワークにおける動的なフローを用いた避難計画問題に対する効率的なアルゴリズム”, RIMS 研究集会「モデリングと最適化の理論」, 232-240, 2006.

23. 神山直之, 加藤直樹, 瀧澤重志:
“辺の容量が一定のネットワークにおける動的なフローを用いた避難計画問題に対する効率的なアルゴリズム”, 日本建築大会, 465-466, 2006.

24. Naoyuki Kamiyama, Katoh Naoki, Astushi Takizawa:
“An Efficient algorithm for Evacuation Problem in a Certain Class of Layered Networks”, 日本オペレーションズ・リサーチ学会秋季研究発表会, 230-231, 2006.

25. M. Ohsaki:
“M. Ohsaki, Design of tensegrity and flexible structures”, Proc. Conf. Computational Eng. Sci., Japan Soc. Comp. Eng. Sci., 11, 855-858, 2006.

26. 木下拓也, 大崎 純:
“コンプライアントメカニズムを利用した可動建築構造物”, 日本建築学会大会学術講演梗概集, B-1, 725-726, 2006.

27. 張 景耀, 大崎 純, S. D. Guest:
“Multi-stable テンセグリティ構造物”, コロキウム構造形態の解析と創生 2006, 日本建築学会シェル空間構造運営委員会, 105–110, 2006.

28. 大崎 純, 西脇眞二:
“幾何学的非線形性を考慮したトラスの形状・トポロジー最適化によるリンク機構の生成”, 第 7 回最適化シンポジウム講演論文集, 日本機械学会, 173-178, 2006.

29. 木下拓也, 大崎 純, 加藤直樹, 谷川眞一, D. Avis:
“静定トラス列挙アルゴリズムを利用した柔構造物の最適設計”, 第 7 回最適化シンポジウム講演論文集, 日本機械学会, 203-208, 2006.

30. 田村明久:
“離散凸解析を利用したマッチング市場モデル”, 日本応用数理学会年会講演予稿集, 178-179, 2006.

31. 神山直之, 加藤直樹, 瀧澤重志:
“避難計画問題”, 京都大学 アルゴリズム・計算量・数理計画・OR・etc. 合同研究会 (KIDS), 2007.

32. Naoyuki Kamiyama, Katoh Naoki, Astushi Takizawa:
“An Efficient Algorithm for the Evacuation Problem in a Certain Class of a Network with Uniform Path-Lengths”, 電子情報通信学会 2007 年総合大会, pp.”S-11”-”S-12”, 2007 年 3 月.

33. Naoyuki Kamiyama, Katoh Naoki, Astushi Takizawa:
"An Efficient algorithm for the Evacuation Problem in a Certain Class of a Network with Uniform Path-Lengths", 研究集会「最適化：モデリングとアルゴリズム」, 2007年3月.

34. 谷川眞一, 加藤直樹:
"辺制約付きの無交差全域木の列挙", 研究集会「最適化：モデリングとアルゴリズム」, 2007年3月.

35. 神山直之, 加藤直樹, 瀧澤重志:
"単一パス長ネットワークにおける避難計画問題に対する効率的なアルゴリズム", 日本建築大会, pp.1167-1168, 2007年8月.

36. Naoyuki Kamiyama, Katoh Naoki, Astushi Takizawa:
"Theoretical and Practical Issues of Evacuation Planning in Urban Areas", スケジューリング・シンポジウム2007, pp.103-108, 2007年9月.

37. Naoyuki Kamiyama, Katoh Naoki, Astushi Takizawa:
"Arc-disjoint In-trees in Directed Graphs", 日本オペレーションズ・リサーチ学会 研究部会「コンピュテーション」, 2007年10月.

38. Naoyuki Kamiyama, Katoh Naoki, Astushi Takizawa:
"Arc-disjoint In-trees in Directed Graphs", 情報処理学会アルゴリズム研究会, pp.1-8, 2007年11月.

39. 加藤直樹，谷川眞一:
"辺制約付き無交差全域木列挙アルゴリズム", 情報処理学会アルゴリズム研究会, 2006-AL-111, 2007.

40. David Avis, 加藤直樹, 大崎純, Ileana Streinu, 谷川眞一:
"無交差ラーマンフレームワーク列挙アルゴリズム", 電子情報通信学会2007年総合大会, 2007.

41. 田治剛有, 谷川眞一, 加藤直樹, 瀧澤重志:
"線的施設配置問題に関する基礎研究", 日本建築学会大会, pp.435-436, 2007年8月.

42. 田治剛有, 谷川眞一, 加藤直樹, 瀧澤重志:
"線的施設配置問題に関する研究", 日本応用数理学会年会, pp.374-375, 2007年9月.

43. M. Ohsaki and J. Y. Zhang:
"Design of tensegrity structures by optimization methods", Proc. Int. Conf. on Computational Methods, Hiroshima, 255, 2007.

44. 木下拓也，大崎 純，加藤直樹，谷川眞一:
"静定トラス列挙アルゴリズムを利用したコンプライアンとメカニズムの生成", 日本建築学会大会学術講演梗概集 (九州), B-1, 333–334, 2007.

45. 張 景耀，大崎 純:
"幾何学制約付きのテンセグリティ構造形状決定法", 日本建築学会大会学術講演梗概集 (九州), B-1, 891–8924, 2007.

46. T. Kinoshita and M. Ohsaki:
"Topology synthesis of bistable compliant mechanism using truss elements", Abstracts of The 7th Int. Conf. on Optimization: Techniques and Applications (ICOTA7), Kobe, Japan, 2007.

47. 神山直之, 加藤直樹:
"内向木による有向グラフの被覆", 情報処理学会アルゴリズム研究会, 発表予定, 2008年3月.

48. Naoyuki Kamiyama, Katoh Naoki, Astushi Takizawa:
"Arc-disjoint In-trees in Directed Graphs", 日本オペレーションズ・リサーチ学会春季研究発表会, 発表予定, 2008年3月.

49. 田治剛有, 谷川眞一, 神山直之, 加藤直樹, 瀧澤重志:
"線的施設配置問題に関する研究", 日本オペレーションズ・リサーチ学会春季研究発表会, 発表予定, 2008年3月.

文部科学省科学研究費補助金 特定領域研究 (課題番号 16092218)

A07: 現実的な状況下での量子計算の能力に関する研究

量子計算および量子通信は，現在の計算・通信方式より本質的に勝る点があると広く信じられている一方で，その能力や具体的な利用方法に関してはまだ明らかになっていない点が多い．そのため，完全にその物理的な実現ができていない今日からでも，計算機科学の立場から多くの研究をしておく必要がある．本研究では，今まで計算機科学の分野で行われてきた量子計算や量子通信に関する研究に加えて，将来の量子計算や量子通信が実際に使われるであろう状況を考慮して，それまではあまり考えられてこなかった，量子計算の一部分が古典的な計算資源で置き換えられたり，量子計算の一部分がエラーを含む場合に対して，量子計算や量子通信の能力の解析を行なった．また，そのような状況でどのように量子計算や量子通信を具体的に利用するかの方法論や，逆にどのような状況ならば量子計算や量子通信が能力を発揮するかに関して，オートマトンの計算能力，オラクル計算における計算量，分散計算における通信量，暗号計算における安全性，などの様々な項目で多くの新たな知見を得た．

研究組織

研究代表者： 山下 茂　奈良先端科学技術大学院大学 情報科学研究科

研究分担者： 中西 正樹　奈良先端科学技術大学院大学 情報科学研究科

交付決定額 (配分額)

平成 16 年度	1,700,000 円
平成 17 年度	1,600,000 円
平成 18 年度	1,700,000 円
平成 19 年度	1,700,000 円
合　計	6,700,000 円

研究成果の概要

- 学術誌
  TCS (2007 年), IPL(2006 年), New J. Phys.(2006 年), IEICE(2005 年, 2007 年) など
- 国際会議
  STACS (Aachen, 2007 年), ICALP (Wroclaw, 2007 年), SWAT (Riga, 2006 年), ISAAC (Sendai, 2007 年), COCOON (Jeju Island, 2004 年, Taipei, 2006 年) など
- Nara Advanced Interdisciplinary Workshop on Quantum Information Science を山下と中西が中心となって，2007 年 1 月 25-26 日に開催した．

1　はじめに

量子計算および量子通信は，現在の計算・通信方式より本質的に勝る点があると広く信じられている一方で，その能力や具体的な利用方法に関してはまだ明らかになっていない点が多い．そのため，完全にその物理的な実現ができていない今日からでも，計算科学の立場から多くの研究をしておく必要がある．それに関して，量子計算の計算機科学からの研究は，量子チューリングマシンと呼ばれる非常に理想的な量子計算機のモデルの上で，計算ステップ数の漸近的な評価をしているものがほとんどである．そこで，本研究では，上記のような研究に加えて，より現実的な状況を考慮して，量子計算の一部分が古典的な計算資源で置き換えられたり，量子計算の一部分がエラーを含む場合に，量子計算や量子通信の能力の解析を行なうことを目的とした．また，そのような状況でどのように量子計算を具体的に利用するかの　方法論や，逆にどのような状況ならば量子計算や量子通信が能力を発揮するかについての知見を得ることも目標とした．

上述した目標達成のために，量子計算の一部分が古典的な計算資源で置き換えられたモデル，量子計算の一部分がエラーを含むモデル，量子計算が分散で計算するモデルなどの様々なモデルにおいて，量子計算や量子通信の能力の解析などを多角的に進めた．また，計算の大部分を古典計算機で行なわなければならないような現実的な状況において，量子計算機を利用すれば，古典計算だけの場合よりもより高速に解ける問題やその状況，およびその具体的な利用方法についての検討も行った．さらに，現実的な量子計算の利用方法を考えたて，計算に限定せずに古典通信では実現が困難な性質を持つ通信プロトコルへの応用がないかを検討し，様々な暗号やマジックプロトコルの考案とその性能解析を行った．また，量子エラー訂正回路を削減する手法や量子回路設計などに関しても多角的に研究を進めた．まず，主要な研究成果の概要を具体的に以下に示す．

古典スタック付き量子プッシュダウンオートマトンの計算能力

量子デバイスと古典デバイスが協調して動作する計算機のモデルとして古典スタック付き量子プッシュダウンオートマトンを提案し，片側誤りのもとで純粋な古典モデルより真に能力が高いことを示した．

エラーのある量子オラクルの分類を行う量子アルゴリズム

ある決められたオラクルの集合が与えられ，そのうちの一つがブラックボックスとして与えられた時に，何回そのブラックボックスのオラクルに質問すれば，オラクルを同定できるかという問題をオラクル同定問題という．オラクルが量子的なサブルーチンとして Grover Search に利用できる時に，たとえオラクルが定数確率で間違った値を返す場合でも，エラーがない場合の問い合わせ回数の下限と一致するオラクルへの問い合わせ回数でオラクル同定問題を解く量子アルゴリズムを設計した．

N 個の関数の値の OR を分散環境で効率よく計算する量子プロトコル

Alice と Bob が l 個の入力データ x と y をそれぞれ持っている状況で，l 個の $f(x,y)$ の出力の OR を計算するときに Alice と Bob の間で必要な通信量（ビット数）を考える．特に，f が $x \neq y$ かどうかをチェックする関数のとき，この問題は list-nonequality function:$LNE(l,k)$ と呼ばれ，分散環境での通信量に関して基本的な問題である．この問題に対して，量子のプロトコルに古典のランダムコインの概念をうまく取り入れることによって，古典の場合よりも効率的な計算手法を設計した．

古典計算機と量子計算機を協調させてプログラミングする枠組み

通常の C++で書かれたプログラムからほぼ全自動で，Grover Search で置き換えていい部分を

判定し，その量子アルゴリズムに相当する量子回路を生成する手法を実現した．今後は Grover Search 以外の量子アルゴリズムについても検討する課題が残っている．

量子情報を複数に秘密を分散して管理する手法

安全な情報の管理のために，パスワードなどの重要な秘密情報を複数に分散し，そのうちのいくつかが集まれば秘密が復元でき，しかも集まってきた情報のうち不正なものが何かを識別することができる手法を閾値秘密分散法と言う．秘密が量子情報の時に，直交配列とスタビライザ符号を用いて秘密分散を安全に行う手法を考案した．

公開する情報に「封」をする手法

封筒にいれた手紙などに紋章などの封をすることのように，データを途中で誰も見ていないことを保証する「封」の役割を実現する手法を考案した．考案した手法は，既に提案されている手法よりも不正がよりしにくい性質を持つと期待される．

リング状のネットワークでの量子分散アルゴリズム

量子計算機がリング状のネットワークでつながれた状況で，Distinctness と呼ばれる関数を計算する場合の通信量の下限を示した．また下限にほぼ一致する上限を持つ 2 つの量子プロトコルも設計した．得られた上下限は，各計算機に与えられる入力の可能性の数が計算機の数と同じオーダーの時には等しくなるので，設計したプロトコルはほぼ最適であると考えられる．

量子情報を送信可能な量子秘密通信プロトコル

量子暗号プロトコルとしては量子鍵配布が有名であるが，本手法では，鍵の配布だけでなく任意の量子ビットを安全に送信できるプロトコルを考案した．

エラーのある量子オラクルによるロバストな計算

量子バイアス付オラクルとは，各質問に対し確率 $1/2+\epsilon$ 以上で正しい答えを返すような量子オラクルである．そのようなオラクルが与えられたとき，そのオラクルへの $O(1/\epsilon)$ 回の質問を用いて，確率 2/3 以上で正しい答えを返すようなオラクルへの 1 回の質問をシミュレートするアルゴリズムを考案した．考案したアルゴリズムと既知の結果を用いると，オラクルがエラーを持つ場合の探索問題に関して既知の下限と一致する最適なアルゴリズムを得ることができた．

量子ネットワーク上での効率的な情報の伝送

バタフライネットワークと呼ばれるネットワークでは，s_1 から t_1 及び s_2 から t_2 という 2 つのフローパスが容量 1 のチャンネルを共有している．古典の場合，このボトルネックとなるチャンネルの存在にも拘らず，2 つのパスに沿って 2 ビットを同時に送ることができる．量子の場合，量子状態の非直交性がこのボトルネックをはるかに困難なものにする．しかし，工夫したプロトコルにより，以下の結果を得た．(i) 任意の量子状態 $|\psi_1\rangle$ を s_1 から t_1 に，$|\psi_2\rangle$ を s_2 から t_2 に，それぞれフィデリティが真に 1/2 より大きくなるように同時に送る事ができる．(ii) $|\psi_1\rangle$ か $|\psi_2\rangle$ の一方が古典情報なら，フィデリティが 2/3 となるように同時に送ることができる．(iii) もし，$|\psi_1\rangle$ と $|\psi_2\rangle$ がいくつかの既知の有限個の状態であれば，フィデリティを更に改善することができる．これにより，s_1 から t_1 への古典的な情報を 2 ビット（同時に s_2 から t_2 への 2 ビット）のうちの任意に（通信の後で）選ばれた 1 ビットを復号するようなランダムアクセスコーディングが可能であることを示した．

量子文字列封印における情報流出の上下界

量子封印とは，乱数と量子ビットを用いてメッセージを符号化することにより，第三者がそのメッセージを復号すると量子状態が壊れ，メッセージが読まれたことが確認できるというものである．

1 ビットの値を封印する量子ビット封印に関しては，安全なプロトコルが存在しないことがすでに知られている．しかし，量子ビット封印が安全でないからと言って量子文字列封印も安全でないとは限らない．そこで，量子文字列封印の可能性について様々な議論がなされている．しかし，今まで盗聴者が盗聴できる情報量と検出率の関係については知られていなかった．本研究では，盗聴される情報量の上界が自明な下界に一致するプロトコルを提案し，盗聴者が盗聴できる情報量と検出率のトレードオフの関係を明らかにした．

量子回路を二分決定グラフ構造から合成する手法

ブール関数を計算する量子回路の組織的な設計手法を提案した．提案手法は，既存の手法が入力数に関して指数的な精度のゲートを必要とするのに比べ，３種類のゲートしか用いないという点でより実用的であると言える．また，提案手法の導入のために，量子回路の動作を記述する行列関数という概念を新たに導入し，行列関数を二分決定グラフで表現する *Decision Diagram for a Matrix Function (DDMF)* というデータ構造も提案した．これらの概念も理論的に興味深いと考えられる．

多準位系の量子ゲートによる量子回路設計のためのユニタリ行列近似

近年，量子計算において多準位系を用いることが提案されている．本研究では，多準位系の量子ゲートによる量子回路設計のために，与えられた任意の n 次元のユニタリ行列を決められたいくつかの n 次元のユニタリ行列（基本ゲートに相当）の積で近似する手法を提案した．提案手法は，与えられた行列を n 次元だが「2 準位にしか作用しない行列」に分解し，その行列を基本ゲートの積に分解するという手法をとる．これにより，既存の手法（Solovay-Kitaev (SK) theorem およびそれに基づいた Dawson らによる手法）に比べて，合成される回路の規模はほぼ同じままで，アルゴリズムの前段階で要する処理時間を大幅に短縮することができるようになった．

Pure Dephasing を考慮した量子オラクル計算モデル

本研究では，量子オラクルと呼ばれる量子のブラックボックスがデコヒーレンスのエラーを受けた状況を計算科学の視点から考察を行った．まず，計算量とデコヒーレンスのエラーの度合いとの関係を解析するためにオラクルの出力に pure dephasing のエラーが起こってもそのようなオラクルを複数用いて pure dephasing のエラーを軽減する手法を提案した．これは，古典計算においてエラーのあるオラクルに対処するための最も基本的な手法である多数決による成功確率の増幅に対応する．我々の知る限りでは，我々の問題設定はオラクル計算の枠組みでデコヒーレンスの影響を考察する最初のものである．

ランダムアクセス符号のタイトな存在性の条件

1999 年に Ambainis, Nayak, Ta-shma, Vazirani によって導入された (n, m, p)-量子ランダムアクセス符号とは，次のような通信システムである．n ビットの情報を持つ送信者はその情報を m キュビットに符号化してそれを受信者に送る．受信者は正作用素値測度によって定義される量子的復号プロセスにより p 以上の確率で正しく元の n ビットの任意の 1 ビットを復号することができる．Ambainis らは (2, 1, 0.85)-量子ランダムアクセス符号の存在を示し，同時に古典的にはそのような符号は不可能であることを示した．Chaung は Ambainis らの符号を (3, 1, 0:79)-量子ランダムアクセス符号に拡張したが，それ以来 $p > 1 = 2$ をみたす $(4, 1, p)$-量子ランダムアクセス符号の存在は未解決であった．本研究では，そのような量子ランダムアクセス符号が存在しないことを示した．また，一般に量子および古典のランダムアクセス符号に関して，タイトな存在条件として以下のことを示した．(1) $(2^{2n}-1, n, > 1/2)$-量子ランダムアクセス符号は存在する．(2) $(2^{2n}, n, > 1/2)$-量子ランダムアクセス符号は存在しない．(3) $(2^n-1, n, > 1/2)$-古典ランダムアクセス符号は存在する．(4) $(2^n, n, > 1/2)$-古典ランダムアクセス符号は存在しない．

耐故障性量子計算におけるエラー訂正回数の削減手法

量子状態への量子的な操作は，外界からの相互作用などにより理想的な動作を 100 パーセントするとは考えられない．そのため, Shor により各量子ゲートが完全に動作しない状況でもフォールトトレラントに計算を実行する耐故障性量子計算の手法 [21] を提案されている．その手法は，演算に用いる量子ビットを量子誤り訂正符号（具体的には Steane 符号）で符号化して，各基本演算の前後で，量子誤り訂正により逐一エラーを訂正する手法である．しかし，量子誤り訂正は計算ステップ数が大きいのに加え，誤り訂正自体で誤りを生じる可能　性がある．そのため，全体の計算結果の信頼性を保ったまま，量子誤り訂正回路を挿入する場所を極力少なくする種々の手法を検討した．そのために，量子回路内での計算結果の信頼性（エラー訂正符号の中でのエラービット数がエラー訂正可能な数以下である確率）を効率よく見積もる手法を考案した.

関数の真理値表の 1 の数とそれを評価するための量子質問量との関係

量子計算により N 変数ブール関数を計算するときに必要な量子質問量に関しては，AND, OR（および AND-OR tree や NAND tree など）および，Parity や閾値関数などの典型的な関数にしか知られていなかった．本研究では，関数の真理値表の 1 の数というパラメーターだけで量子質問量の上限および下限ががほぼ決定できることを明らかにした．特に，1 の数がブール関数の変数の数 N の多項式で抑えられるときは，量子質問量が $\Omega(N)$ であることを示した.

以下では，これらのうち代表的な研究項目についてその詳細を述べる.

2　古典スタック付き量子プッシュダウンオートマトンの計算能力

シンプルな量子計算機のモデルとして，量子有限オートマトンや，量子プッシュダウンオートマトンが提案されている．特に量子有限オートマトンは量子チューリング機械や量子回路と異なり，必要な量子資源が有限であるため，現実的に実現できる量子計算機の規模を考えると，非常に重要なモデルである.

一方，計算機のすべてを量子デバイスを用いて実装するのは現実的ではない．量子計算機が実現した場合は，少ないメモリしか持たない量子計算機と多くのメモリを持つ古典計算機を協調して動作させることが必要となる．シンプルな量子計算機と古典計算機が協調して動作した場合，量子計算機単独で動作するよりも能力が高くなる場合がある.

そこで，本研究ではこのような量子デバイスと古典デバイスが協調して動作するような計算機の能力の解析を行った．量子デバイスと古典デバイスが協調して動作するモデルを考えることは，上述の理由の他に，可逆性を緩和できるという利点がある．量子計算モデルは量子力学から来る制約上，可逆な計算モデルでなければならない．この制約は場合によっては量子計算モデルの能力を古典計算モデルより弱くすることにもなりうる．しかしながら，古典デバイスはこのような可逆性という制約を受けないため，自由に操作ができる.

本研究では，計算モデルとしてプッシュダウンオートマトンを考え，その有限状態制御部とテープヘッドに量子デバイスを，スタックに古典デバイスを用いた量子-古典協調計算モデルを提案し，その能力の解析を行った．結果として，片側誤りのもとで，量子計算モデルの方が古典計算モデルよりも認識できる言語のクラスが大きくなる，つまり，能力が高くなることを示した.

以下に，本研究で提案した古典スタック付き量子プッシュダウンオートマトンの定義を与える.

Definition 1 古典スタック付き量子プッシュダウンオートマトンは次の *8* 項組で定義される.

$$M = (Q, \Sigma, \Gamma, \delta, q_0, \sigma, Q_{acc}, Q_{rej}),$$

ここで Q は状態の有限集合, Σ は左 *(右)* 終端記号 ¢ *(\$)* を含む入力アルファベット，Γ は底記号 Z を含むスタックアルファベット，$\delta : (Q \times \Sigma \times \Gamma \times Q \times \{0,1\}) \longrightarrow \mathbb{C}$ は状態遷移関数，q_0 は初期状態，$\sigma : Q \setminus (Q_{acc} \cup Q_{rej}) \longrightarrow \Gamma^+ \cup \{-, pop\}$ はスタック操作を決める関数，Q_{acc} $(\subseteq Q)$ は受

理状態集合，Q_{rej} $(\subseteq Q)$ は非受理状態集合である．ただし，$Q_{acc} \cap Q_{rej} = \emptyset$ とする．

□

このモデルは，量子有限オートマトンに古典のスタックを付加したものと考えることができ，量子有限オートマトン部の動作は，スタックトップの記号に基づいて決められるユニタリ変換によって規定される．また，スタックの操作は量子有限オートマトン部の観測結果に基づいて積むべき記号 (あるいはスタックトップの削除) が決定される．具体的には次のように動作を行う．

$\delta(q, b, a, q', D) = \alpha$ は状態 q で入力記号 b，スタック記号 a を読んだときに状態 q' に遷移して，ヘッドを D の方向に動かす確率振幅が α であることを示している．状態 q，ヘッド位置 k に対する時点表示を $|q, k\rangle$ と表すと，古典スタック付き量子プッシュダウンオートマトンの (量子有限状態制御部の) 1 ステップの動作は次のようにユニタリ変換 $U_a^{\boldsymbol{x}}$ を適用することに相当する．ここで，a はスタックトップの記号，$\boldsymbol{x}$ は入力文字列を表すものとする．

$$U_a^{\boldsymbol{x}}(|q, k\rangle) = \sum_{q' \in Q, D \in \{0,1\}} \delta(q, x(k), a, q', D) \left|q', k + D\right\rangle,$$

ここで，$x(k)$ は 入力文字列 $\boldsymbol{x}$ の k 番目の入力記号を表すものとする．

量子有限状態制御部に対して，ユニタリ変換 $U_a^{\boldsymbol{x}}$ が適用された後，次に示す観測量を用いて観測が行われる．

$$\begin{aligned} E_w &= span\{|q, k\rangle \,|\sigma(q) = w\}, \\ E_{\mathrm{acc}} &= span\{|q, k\rangle \,|q \in Q_{\mathrm{acc}}\}, \\ E_{\mathrm{rej}} &= span\{|q, k\rangle \,|q \in Q_{\mathrm{rej}}\}. \end{aligned}$$

と定め，観測量を $\mathcal{O} = \oplus_j E_j$ と定義する．

観測の結果が acc (rej) であれば，入力を受理 (非受理) する．観測の結果が pop であれば，スタックトップ記号をポップする．'$-$' であればスタックを変化させない．それ以外の場合 (つまり，j が Γ^+ の文字列の場合) は j をスタックにプッシュする．以上を繰り返すことにより，計算を進める．

本研究では，古典スタック付き量子プッシュダウンオートマトンで片側誤りで認識できるが，(古典の) 確率プッシュダウンオートマトンでは片側誤りで認識できない言語が存在することを示した．

Theorem 1 次に示す言語 L_1 は古典スタック付き量子プッシュダウンオートマトンで片側誤りで認識できる．

$$L_1 = \left\{ u\#v\natural w\flat x\%y \;\middle|\; \begin{array}{l} u, v, w, x, y \in \{a, b\}^*, \\ \neg \left(\begin{array}{l} |u| = |v| = |w| = |x| \\ and\ v = w \end{array} \right) \\ and\ \neg \left(\begin{array}{l} |u| = |x|\ \ and \\ |v| = |w| = 0 \end{array} \right) \\ and\ (y = v^R\ \ or\ y = w^R) \end{array} \right\}$$

□

Theorem 2 L_1 は文脈自由言語ではない．

□

定理 2 より，L_1 は非決定性プッシュダウンオートマトンで認識できないことがわかる．片側誤り確率プッシュダウンオートマトンが認識できる言語のクラスは，非決定性のそれより真に小さいことに注意すると，L_1 は片側誤り確率プッシュダウンオートマトンで認識できないことがわかる．

著者らは今までに任意の確率プッシュダウンオートマトンは同じ確率で入力を受理する可逆確率プッシュダウンオートマトンに変換できることを示している．

Theorem 3 可逆確率プッシュダウンオートマトンは任意の確率プッシュダウンオートマトンを同じ確率で入力を受理するようにシミュレートすることができる．

□

定理 3より，任意の確率プッシュダウンオートマトンは量子プッシュダウンオートマトンでシミュレート可能であることが言える．このことと定理 1，2 を合わせると，古典スタック付き量子プッシュダウンオートマトンは片側誤りのもとで，確率プッシュダウンオートマトンより真に能力が高いことが言える．

3 エラーのある量子オラクルの分類を行う量子アルゴリズム

Grover の探索アルゴリズムなどのオラクルを用いて解く問題の一般化として，以下のようなオラクル同定問題というものが考えられている．

オラクル同定問題

入力： M 個のオラクル集合 $S = \{f_1, ..., f_M\}$ とその中のオラクル $f_i : \{0,1\}^n \to \{0,1\}$，$f_i \in S$.

出力： オラクル f_i のインデックス i.

解が一つの場合の探索問題（つまりオラクル $f_i(i) = 1, f_i(j) = 0$ $(j \neq i)$ から i を求める Grover の探索問題），およびビット毎の積の論理和（内積）$f_i(j) = i \cdot j$ から i を求める問題もオラクル同定問題の一例である．言い替えると，オラクル同定問題は，M 個のオラクルの動作がわかっているという条件で，ブラックボックスにその内の一つが隠されている時に，オラクルへのクエリーを用いてどのオラクルがブラックボックスに隠されているかを決定する問題である．これは，$M \times N$ 行列 $(N = 2^n, N \leq M \leq 2^N)$ が与えられ，現在のオラクル（ブラックスボックスに隠されたオラクル）がどの行であるかを同定する問題と見ることもできる．そのため，以下では,「$M \times N$ 行列に対するオラクル同定問題」と呼ぶことにする．

オラクル同定問題において，オラクルがエラーを持つ場合に，エラー除去のために増加するクエリー回数が定数倍程度で済む量子アルゴリズムを以下で述べる．ここでは，$M = N$，すなわちオラクル候補の数がオラクル入力の長さと一致する場合を考える．(Grover の探索アルゴリズムの問題もこの場合に相当する.)

エラー付きオラクル同定問題

入力： M 個のオラクル集合 $S = \{f_1, ..., f_M\}$ とその中のある一つのオラクル $f_i \in S$ を計算するユニタリ変換 U_{f_i}:

$$U_{f_i} |x\rangle |0\rangle |0\rangle =$$

$$\sqrt{p_x} |x\rangle |\phi_i\rangle |f_i(x)\rangle + \sqrt{1-p_x} |i\rangle |\psi_i\rangle |\neg f_i(x)\rangle,$$

但しある定数 $0 < \varepsilon < 1/2$ に対して，$1/2 + \varepsilon \leq p_x^2 \leq 1$ である．

出力： オラクル f_i のインデックス i.

エラー無しの場合は，$O(\sqrt{N})$ 回のオラクルクエリーにより解けることが既に知られている．しかし，このアルゴリズムは古典的なオラクルクエリーを数多く利用しているために単純に頑健な量子探索アルゴリズムを利用するするだけでは，$O(\log N)$ 倍のオラクルクエリーが避けられない．そのため，エラーがあるオラクルの同定問題をエラーのない場合に対して定数倍程度のクエリー回数の増加で解く量子アルゴリズムは自明でない．しかし，実際には以下の定理が成り立つ．

Theorem 4 $N \times N$ 行列に対するエラー付きオラクル同定問題を $O(\sqrt{N})$ 回のオラクルクエリーにより少なくとも定数確率で解くアルゴリズムが構成可能である． □

ここでは，アルゴリズムを簡単に述べる．

1. 以下の 2.–6. をオラクル候補 Z が一つになるまで繰り返す．最初 は Z を与えられた N 個のオラクル候補全てとする．

2. 以下の操作では，Z の各列に対して**列反転** (column flip) を行う．ここで列反転とは Z の各列において，もし 1 が 0 の数より少なければ，その列の 0, 1 を反転し，そうでなければそのままにしておく，という操作である．この操作で各列での 1 の数が 0 の数よりも少ない行列として常に扱うようにする．(この操作にはクエリーを全く用いないことに注意する.)

3. 以下の 4. を選択された列によって**被覆**された行が $|Z|/3$ を超えるまで，あるいは選択する列がなくなるまで繰り返す．ここで列 j が行 i を被覆するとは $f_i(j) = 1$ を満たす，つまり i 行 j 列が 1 であることを意味する．最初，選択された列の集合 T は空集合とする．

4. T に属する列で被覆されていない行において 1 の数が最大の列を選び，T で被覆されてない行のうち $\sqrt{|Z|}/\log N$ 行を被覆しているならば，それを T に含める．

5. T に属する列に対して頑健な量子探索アルゴリズムを $O(\log N)$ 回適用する．1 が発見できれば T が被覆する行を新たに Z とおいて 2. へ戻る．

6. 5. で 1 が発見できなかった場合，以下のいずれかを行なう．

6.a T で被覆する行数が $|Z|/3$ 以上だった場合，T が被覆する行を除外したものを新たに Z として 2. に戻る．

6.b T で被覆する行数が $|Z|/3$ 未満だった場合，T が被覆する行を除外した行列に対して 7.-9. を実行する．

7. T に属さない列に対して頑健な量子探索アルゴリズムを適用する．

8. T に被覆されていない列の中で，7. で得られた列に 1 を持つ行を選ぶ．選ばれた行の集合を S' とする．

9. T に被覆されていない列の中で，S' 中に 0 と 1 両方含んだ列を探し，$\log N$ 回の古典的クエリーでその列の値を決定し，その結果に矛盾する行を Z から削除する．7.–9. を $|Z| = 1$ になるまで繰り返す．

上記のアルゴリズムに対して，オラクルのクエリー回数を評価する．5. で必要なクエリー回数は $O(\sqrt{|T|}\log N)$ であるが，$|T| \leq |Z|/(\sqrt{|Z|}/\log N) = \sqrt{|Z|}\log N$ より，5. で 1 が発見できた場合にはクエリー回数 $O(|Z|^{1/4}\log^{3/2} N)$ で確率 $1 - O(1/N)$ でオラクル候補の総数を 1/2 以下に絞り込むことができることになる．一方，5. で 1 が発見できなかった場合には，確率 $1 - O(1/N)$ 以上でオラクルは T に属さないことになる．従って，T が被覆する行数が $|Z|/3$ 以上だった場合には，オラクル候補の総数を 2/3 以下に絞り込むことができる．また，T が被覆する行数が $|Z|/3$ 未満だった場合には T で被覆された行を除外した後の各列に含まれる 1 の数が $\sqrt{N}/\log N$ という事実に注目すると 7.–9. で $\log N$ 回のクエリーをせいぜい $\sqrt{N}/\log N$ 回繰り返せばよい．よってこのときのクエリー回数は $O(\sqrt{N})$ 回である．従って全体でせいぜい $O(N^{1/4}\log^{5/2} N + \sqrt{N}) = O(\sqrt{N})$ のクエリー回数でオラクル同定が定数確率で可能である．

このアルゴリズムを改良して，さらに以下のことを示した．アルゴリズムが複雑なためここでは，結果のみを示す．

Theorem 5 $M \times N$ 行列に対するエラー付きオラクル同定問題を $M = poly(N)$ の時には，$O(\sqrt{N})$ 回のオラクルクエリーにより，少なくとも定数確率で解くアルゴリズムが構成可能である． □

Theorem 6 $M \times N$ 行列に対するエラー付きオラクル同定問題を $poly(N) \leq M \leq 2^{N^d}$ (d は $0 < d < 1$ の定数) の時には，$O(\sqrt{N\frac{\log M}{\log N}})$ 回のオラクルクエリーにより，少なくとも定数確率で解くアルゴリズムが構成可能である． □

オラクルにエラーがない場合でも，$M \times N$ 行列に対する同定問題のオラクルクエリーの下限が，$\Omega(\sqrt{N\frac{\log M}{\log N}})$ であることが既に示されているので，本研究の結果はほぼ最適であることがわかる．

4 N 個の関数の値のＯＲを分散環境で効率よく計算する量子プロトコル

Alice と Bob がネットワークでつながっていて，Alice が k ビットの入力 x を持ち，Bob が k ビットの入力 y を持っているとする．この状況で，x と y からある関数 $f(x, y)$ を計算するのに Alice と Bob の間で何ビットのデータ (通信量) をやり取りしなければならないかという問題は，分散環境において行う計算の効率を考える上での基本的な問題である．この応用で，Alice と Bob が l 個の入力データを持っている状況で，l 個の f の出力のＯＲ ($f^{\vee l}$ と表記する) を計算する問題に必要な通信量に関しても（古典）計算量の分野では多く研究されている．

本研究では，量子計算機が使える状況，すなわち，ネットワークと計算機が量子情報を扱える場合に，上記の問題を解く効率的な量子分散アルゴリズムが構築できることを示した．その手法は，古典のテクニックとして用いられる共通のコイン（乱数）を用いるモデルを拡張し，古典のコイン

を用いる量子計算のモデルを構築することを基本としている．特に f が $x \neq y$ かどうかをチェックする関数のとき，上記の問題は list-nonequality function: $LNE(l,k)$ と呼ばれる．古典計算の状況では LNE を解くのに必要な通信量は，$\Theta(l+\log k)$ であることがわかっているが，本研究で構築した量子アルゴリズムでは $O((\sqrt{l}\log l+\log k)$ で解けるため，古典計算より効率が良いことが分かる．

本研究の成果を述べる前に，Høyer らが示した再帰的に Amplitude Amplification を適用して INT_n を解く量子プロトコルを単純に応用した，量子プロトコルについて述べる．

Theorem 7 f を $f:\{0,1\}^k\times\{0,1\}^k \longrightarrow \{0,1\}$ であるような任意の関数，$\mathcal{P}$ を f を有界誤りで計算する量子通信プロトコルであるとする．このとき，ある定数 c に対して $Q_2(f^{\vee l}) = O(c^{\log^* l}\sqrt{l}\cdot C(\mathcal{P}))$ となる．ここで $C(\mathcal{P})$ はプロトコル $\mathcal{P}$ のコストを表すものとする． □

これより，$\overline{\mathrm{LNE}(l,k)} = \bigvee_{i=1}^{l} EQ_k$ であり，かつ，$Q_2(EQ_k) = \Theta(\log k)$ であることを考慮して，以下の Corollary を得る．

Corollary 1 $Q_2(LNE(l,k)) = O(c^{\log^* l}\sqrt{l}\cdot\log k)$. □

古典通信プロトコルの場合，共通コインを用いたプロトコルを基に，共通コインを用いないプロトコルを構築することができる．したがって，基となる共通コインを用いるプロトコルとして，効率が良いものを考えることにより，効率のよい (共通コインを用いない) プロトコルを構築することが可能である．具体的には，共通コインを用いた f を計算するプロトコルを基に共通コインを用いた $f^{\vee l}$ を計算するプロトコルを構築し，次に，このプロトコルを $O(\log lk)$ ビットの通信を付加することで共通コインを用いないプロトコルに変換する．ここで k は f の入力長，つまり，Alice と Bob の持つ入力のサイズは $O(lk)$ である．ここで，使用した共通コインの個数に関わらず，必要となる付加的な通信量は $O(\log lk)$ であることに注意されたい．この手法を用いることにより，$\mathrm{LNE}(l,k) = \overline{\mathrm{EQ}_k^{\vee l}}$ の古典有界誤りプロトコルの通信量は，$\Theta(l+\log k)$ となる．したがって，Corollay 1 で示した量子プロトコルは必ずしも古典のプロトコルより効率的であるとは限らない．

このことを踏まえ，本研究では上述の手法を量子プロトコルに適用し，定理 7の結果を改良する方法を示した．つまり，再帰の基底プロトコルとして EQ_k を計算する通信量 $O(1)$ の (古典) 共通コインを用いるプロトコルを用いた $\mathrm{LNE}(l,k)$ のプロトコルを考え，上述の手法により $O(\log lk)$ ビットの通信を付加して，共通コインを用いないプロトコルに変換する．これにより，古典の通信量 $\Theta(l+\log k)$ より常に少ない通信量を実現する量子プロトコルを構築することができる．しかしながら，量子プロトコル中で古典の共通コインを用いるため，上述の手法が単純に量子プロトコルに適用できるかは自明ではない．実際，古典コインを用いた量子プロトコルを構築するにはいくつかの問題点に注意を払う必要がある．この問題点を取り扱うために以下にいくつかの定義を行う．

Definition 2 任意の個数の古典共通コインを使用できる量子プロトコルを古典共通コイン量子プロトコルと呼ぶ．$Q_2^{pub}(f)$ を f を計算する有界誤り古典共通コイン量子プロトコルのコストのうち最小のものとする． □

上述の定義では，重ね合わせ中の基底それぞれについて，異なる古典コインを参照することができる．したがって，十分な個数の古典コインをあらかじめ用意しておく必要がある．ここで，古典共通コインプロトコルは古典共通コイン量子プロトコルに変換できることに注意されたい．しかしながら，たとえ定理 7 のプロトコルに古典共通コインを導入したとしても上述の変換手法に基づいて，ある定数 c_1，c_2 に対して，$Q_2(f^{\vee l}) = O(c_1^{\log^* l}\sqrt{l}R_2^{pub}(f)+\log k)$ となるような古典のプロトコルよりも常に効率の良い古典共通コイン量子プロトコルを構築できるとは限らない．これは以下のような理由による．量子振幅増幅アルゴリズムは純粋状態に対して適用可能なアルゴリズムである．つまり，次のような変換をおこなう．$\mathcal{A}\,|0^m\rangle = \alpha\,|\psi_{good}\rangle + \sqrt{1-\alpha^2}\,|\psi_{bad}\rangle$．ここで $\mathcal{A}$ は成功確率を増幅したいアルゴリズム (プロトコル) を表す．しかしながら，古典共通コイン量子プロトコルは (古典の) 確率にしたがうプロトコルであるため，量子状態は純粋状態でなく混合状態として記述される．したがって，単純に量子振幅増幅アルゴリズムを適用することができなくなる．

そこで，混合状態に量子振幅増幅アルゴリズムを適用するための手法が必要になる．以下では，このためにいくつかの準備を行う．

Definition 3 *(*古典共通コイン*)* 量子プロトコル $\mathcal{P}$ はプロトコル実行後の状態が確率 p で $|\psi\rangle = \alpha|\psi_{good}\rangle + \beta|\psi_{bad}\rangle$ と書けるとき， 確率 p で関数 f を純粋状態において有界誤りで解くという．ここで， $|\alpha|^2 > \frac{2}{3}$ かつ，$|\psi_{good}\rangle$ *(*$|\psi_{bad}\rangle$*)* は正しい *(*間違いの*)* 答えが解答ビットに書かれているような状態で張られる部分空間の要素である． □

これらの定義に基づいて，本研究では以下のことを示した．

Lemma 1 $\mathcal{P}$ を f を計算する有界誤り古典共通コイン量子プロトコルとする．また，$C(\mathcal{P})$ でプロトコル $\mathcal{P}$ のコストを表すとする．このとき，少なくとも $1-\frac{\epsilon}{l}$ の確率で f を純粋状態において有界誤りで計算する量子プロトコル $\mathcal{P}_f^{maj}$ を構築ことができる．ただし，$C(\mathcal{P}_f^{maj}) = O(\log l \cdot C(\mathcal{P}_f))$ とする． □

次に，古典通信量のモデルでの議論と同じように，古典共通コイン量子プロトコルにおいても，わずかな通信量を付加することにより共通コインを取り除くことができることを説明する．量子モデルに古典のコインを導入しているため，古典モデルにおける手法をそのまま適用できるかどうかは自明ではない．しかし，実際には古典の場合とほぼ同様の以下の結果を示すことができる．

Lemma 2 f を $f : \{0,1\}^n \times \{0,1\}^n \longrightarrow \{0,1\}$ なる関数とする．任意の δ，ϵ について，f を計算する任意の ϵ-誤り古典共通コイン量子プロトコルは $O(\log n + \log 1/\delta)$ ビットの通信を付加することにより，$(\epsilon+\delta)$-誤り量子プロトコルに変換できる． □

以上より，本研究における以下の主定理が示される．

Theorem 8 f を $f : \{0,1\}^k \times \{0,1\}^k \longrightarrow \{0,1\}$ なる任意の関数とする．このとき $Q_2(f^{\vee l}) = O(\sqrt{l} \cdot \log l \cdot Q_2^{pub}(f) + \log k)$.

Proof. 関数 $f^{\vee l}$ を計算する通信量が $O(\sqrt{l} \cdot \log l \cdot Q_2^{pub}(f) + \log k)$ の有界誤り量子プロトコルを以下に示す．

以下に示すプロトコルは，Høyer らの頑健な量子探索アルゴリズムに基づくものである．頑健な量子探索アルゴリズムにより，次のことが言える．「もし，f を純粋状態において有界誤りで計算する量子プロトコル $\mathcal{P}$ が存在すれば，$\mathcal{P}$ を $O(\sqrt{l})$ 適用することにより $f^{\vee l}$ を有界誤りで計算することができる．」

$\mathcal{P}_f$ を f を計算する古典共通コイン量子プロトコルとする．Lemma 1より $\mathcal{P}_f$ を $O(\log l)$ 回適用することで，少なくとも $1-\frac{\epsilon}{l}$ の確率で純粋状態において有界誤りで f を計算する量子プロトコル $\mathcal{P}_f^{maj}$ を構築することができる．よって，「l 個の入力に対する $\mathcal{P}_f^{maj}$ の試行がすべて純粋状態において有界誤りで関数 f を計算する」ということを少なくとも定数確率でできる．したがって，頑健な量子探索アルゴリズムを l 個の結果に対して適用することにより，全体としての成功確率が少なくとも定数になる．

l 個の入力に対して，プロトコル $\mathcal{P}_f^{maj}$ を量子重ね合わせの状態で適用する際，何番目の入力に対してプロトコルを実行すべきかという情報を Alice と Bob で交換する必要がある．この際 $O(\log l)$ ビットの通信が必要になる．したがって，全体としての通信量は $O(\sqrt{l} \cdot (\log l \cdot Q_2^{pub}(f) + \log l))(=O(\sqrt{l} \cdot \log l \cdot Q_2^{pub}(f)))$ となる．

頑健な量子探索アルゴリズム中で， S_0 と呼ばれる $|00\cdots0\rangle$ に対する位相を反転する操作を行う必要がある．この操作には定数個のビットの通信で十分であることに注意されたい．

上記のプロトコルは古典共通コインを用いているが，Lemma 2より，共通コインを使用しないプロトコルに変換可能である．この変換により，定理中に示した通信量のプロトコルが得られる．

□

最後に，$R_2^{pub}(EQ_k) = O(1)$ であることを利用して，LNE(l,k) に対する量子プロトコルとして，常に古典プロトコル (通信量 $\Theta(l+\log k)$) よりも効率の良いものを与える．

Corollary 2 $Q_2(LNE(l,k)) = O(\sqrt{l} \cdot \log l + \log k)$. □

以上より，本研究は，非常に効率が良い古典プロトコルが知られている LNE に対して，更により効率的な量子プロトコルを示したことになる．

5 古典計算機と量子計算機を協調させてプログラミングする枠組み

現在知られている汎用的な量子アルゴリズムは，データベースサーチ（ある条件を満たすものを探すこと）を高速に行う Grover Search と呼ばれるものしかない．そのため，Grover Search を使って通常のプログラミングをどのように行うのか（すなわちどのようなプログラミング言語を用いるのか）が実用的には非常に重要となる．そこで，C＋＋の通常の記述から Grover Search にすれば効率が良い部分をどのようにして切り出すべきかの検討を行った．既存の研究としては，量子計算向けのプログラミング言語を独自に提案しているものがあるが，本研究で提案する方法は，以下で述べるとおり，より現実的であると考えられる．その概要は以下の通りである．

1. プログラムはC＋＋プログラミング言語とまったく互換の言語で記述する．
2. プログラマが，if 文でチェックをして何かを探したい場合に，それを確率的におこなってもいいと考える場合には，その直前にC＋＋のコメント文で，//Quantum Search と記述する．また，最終的に達成してほしいエラー確率をその後に記述する．(図 1の例を参照)
3. システム側はその部分を量子計算機を用いた場合のオーバーヘッドを考慮して Grove Search で行うべきかを自動で判定し，指定されたエラー率を達成するだけの繰り返しの分も含めて量子回路を生成する．また，古典計算機からの呼び出しインターフェースも自動で生成する．
4. Grover Search をC＋＋でシミュレーションするサブルーチン関数も用意しておき，その関数への引数を自動生成する.（そのため実際の量子計算機を用いなくても，時間はかかるが通常のC＋＋の環境でも実行結果を検証可能となる.）

```
for(int i = 0; i<10000;i++){
   //Quantum Search 0.01
  if(Data1[i] == 100) {
     return i;
  }
}
```

図 1: C＋＋での記述例

ある探索が確率的でいいかどうかはプログラマにしか分からないため，確率的に動作する Grover Search を使用可能かどうかを自動で判定するのは現実的には不可能である．そのため，上述した 2. でのディレクティブを用いた指定はプログラマへの必要最小限の負担であり，現実的な解決方法である考えている．それ以外の作業は，すべてシステム側で判断する．具体的には，該当部分を量子計算機で行うと計算時間が早くなるかどうかを，その量子回路のサイズを見積もることにより自動で判断する．考案した枠組みは，特殊な量子プログラミング言語を用いる手法に比べて以下のような利点がある．

- プログラマは通常のC＋＋で記述でき，量子計算で行ってもよいと考えるところをディレクティブで指定するだけで，あとはシステムが自動で量子計算でやらすべきかを判断してくれる．つまり，プログラマが量子計算の詳細を知らなくても，プログラミング環境が自動的に量子計算機で実行すべき部分を切り出してくれる．
- 全体がC＋＋で記述され，また，量子計算で実行される Grover Search をC＋＋でシミュレーションするライブラリもあるため，プログラム全体を既存のC＋＋コンパイラの環境でシミュレーション可能である．量子計算と古典計算を協調させてプログラムを実行させる様子を見たい場合，従来のような特殊な量子回路シミュレータしかない場合には，そのシミュレーション結果と通常のC＋＋プログラムの出力を組み合わせてシミュレーションするしか方法がない．考案した枠組みでは，特別な量子回路シミュレータを使うことなく，全体の計算を簡単にシミュレーション可能で

ある．

以上の利点から，考案した手法は，将来の量子計算の利用の仕方として有効な手法となりうると考えられる．

上述した枠組みでは，Grover Search をすべき部分を切り出した後，それを基本ゲートによる量子回路として実現しなければいけない．従来，量子アルゴリズムを表現するユニタリ行列を行列分解により基本ゲートに分解する手法が多く提案されているが，大規模な Grover Search を表現するユニタリ行列はそれ自身が探索問題のサイズに対して指数的に大きくなるので，行列を表現して分解する手法は実用的なサイズの Grover Search には適用できない．そこで，本研究では，Grover Search の量子特有の操作で問題ごとに変化しない部分はテンプレートを用いることとし，問題ごとに変わる if の中の制御論理のみを行列分解とは違う手法で設計することを特徴としている．if(Data1[i] == 100) というような単純な場合にはこの制御論理の設計はほぼ自明だが，この部分が i の複雑な算術式で書かれていた場合それを量子回路で実現するのは自明ではない．今年度は，i の複雑な算術式で書かれている条件を，ブール関数で表現する効率的な手法を考案し，それをプログラムとして実装した．作成したプログラムを用いれば，今まで全く知られていなかった，大規模な Grover Search を実際に量子計算で行ったときの詳細な計算ステップを見積もることができるようになるため，多くの研究者に有益なツールとなると考えられる．

6 量子情報を複数に秘密を分散して管理する手法

6.1 はじめに

秘密情報を分散して管理する方法として，(k,n)-閾値秘密分散法 [20] が知られている．これは，秘密 S を n 個の分散情報に符号化し，k 個以上の分散情報を集めると，元の秘密 S を復元できるが，k 個未満の分散情報からは，S に関する情報は全く得られないというものである．

近年，量子情報処理を用いて秘密分散法を実現する量子秘密分散法に関する研究が盛んに行われており [7, 8, 9, 12, 14, 15]，盗聴者の検出や量子通信を行うことによる通信効率の改善が議論されている．

本研究では，さらに進んだ機能として，不正者識別機能を量子秘密分散法に付加する．つまり，不正者識別が可能な (k,n)-閾値量子秘密分散法を提案する．古典情報処理においては，黒沢らにより不正者識別が可能な秘密分散法が提案されている [16]．本研究で提案する手法は，この黒沢らの手法を量子秘密分散法に拡張したものである．黒沢らの手法においては，メッセージ認証を用いて分散情報の改竄を検出しているが，この認証法は量子メッセージには適用できない．そのため，我々の手法では，無条件安全であることが知られているスタビライザーコードを用いた量子メッセージ認証を利用する．我々が提案する手法においては，$t(k \geq 3t+1)$ 人までの不正者を識別可能である．

6.2 不正者識別可能な量子秘密分散法

諸定義

以下に，提案プロトコルを説明するために必要な諸定義を行う．

Definition 4 **t-不正者識別可能 (k,n)-閾値量子秘密分散法**

S を量子メッセージとする．分配者 D は S から n 個の分散情報 $\{v_1,\ldots,v_n\}$ を生成し，それぞれを n 人の参加者 $P_1,\ldots,P_n$ に配布する．ここで，*(i)* 分配者は不正を行わない．*(ii)* 不正者は高々 t 人しかいない． という二点を仮定する．このとき，以下の性質を持つ場合，そのプロトコルを t-不正者識別可能 (k,n)-閾値量子秘密分散法と呼ぶ．

完全性 不正者を除いて k 人以上の参加者が集まれば，元の秘密を高確率で復元できる．

健全性 k 人以下の参加者が集まっても，元の秘密に関する情報は全く得られない．

不正者識別性 k 人以上の参加者が集まれば，その中に存在する不正者を高確率で識別するチューリング機械 M が存在する．

□

Definition 5 不正者識別性

一般性を失うことなく，$P_1, \ldots, P_t$ を不正者とし，それぞれ，$\hat{v_1}, \ldots, \hat{v_t}$ を開示するとする．次のいずれかの条件を満たすとき，不正が成功したと見なす．

Case 1 不正者の存在が検出されたにもかかわらず，M が不正者を識別できない，かつ，秘密の復元に失敗する．

Case 2 参加者が分散情報 $\{\hat{v_1}, \ldots, \hat{v_t}, v_{t+1}, \ldots, v_m\}$ から誤った秘密を復元してしまう．ただし，$m \geq k$.

□

提案手法

提案手法の概略は以下の通りである．

元となる (k, n)-閾値量子秘密分散法として，クリーブらのスタビライザーコードを用いた量子秘密分散法 [7] を用いる．分配者はランダムに量子秘密分散法 Q を選択する．つまり，使用するスタビライザーコードをランダムに選択することになる．Q を用いて量子秘密 S を n 個の分散情報に符号化する．改竄を検出するため，各分散情報を量子メッセージ認証 [4] を用いて符号化する． これらの操作は，黒沢らの手法 [16] の量子メッセージへの拡張となっている．参加者はさらに古典情報である量子メッセージ認証の鍵と,「Q としてどの符号化 (スタビライザーコード) が使われたか」という情報を共有する．このとき，量子メッセージ認証の鍵は不正者を高々 t 人含んでいたとしても，k 人以上の参加者が集まった時，かつその時のみ復元できるように符号化する．これは，線形符号を用いて可能である．また,「Q としてどの符号化が使われたか」という情報は，古典 (k, n)-閾値秘密分散法を用いて共有する．これにより，k 人以上の参加者が集まれば，量子メッセージ認証の鍵を復元でき，各分散情報が改竄されていないかどうかを確認できる．改竄されていないことが確認できれば，クリーブらの手法に従って，元の量子秘密を復元できる．

Theorem 9 上記の手法は t-不正者識別可能 (k, n)-閾値量子秘密分散法である．ただし，$k \geq 3t + 1$ である．

□

6.3 むすび

本研究では，t-不正者識別可能 (k, n)-閾値量子秘密分散法を提案した．提案した手法において，不正者の人数の上界を得ることができれば，それによりスタビライザーコード $[n, 1, K]$ が存在するような n と K の関係について明らかになる可能性がある．また，不正者による攻撃の成功確率の詳細な解析等が今後の課題として挙げられる．

7 リング状のネットワークでの量子分散アルゴリズム

量子分散計算における通信量を解析することは，将来の量子ネットワークの利用の仕方を考える上で重要な理論研究であり，現在さまざまな角度から研究されいる．現在までのほとんどの研究は 2 者間の通信のモデルを扱っているが，インターネットなどのネットワークの形状を考えると，必ずしも現実的なネットワークのモデルを扱っているとはいえない．そこで，本研究では，より現実的な状況に近いネットワーク形状における通信量の研究のために，ネットワークの構造を考慮した通信量の解析を行った．具体的には，n 個の量子プロセッサがリング状の量子ネットワーク上に配置された状況を考え，以下の **Distinctness** と呼ばれる問題を解く量子分散プロトコルの通信量に関する下限と上限を解析した．以下で，プロトコルはすべて定数エラーを持つものを考える．

Definition 6 (Distinctness) リング状の量子ネットワーク上に配置された各量子プロセッサ P_i $(0 \leq i \leq n-1)$ が，整数 $x_i \in \{0, \ldots, L-1\}$ $(n \leq L)$ を入力として持つ．この時，$i \neq j \longrightarrow x_i \neq x_j$ が成り立つかどうかを決定する問題.

以下では，最終的に各プロセッサが **Distinctness** の問題の 1 ビットの結果を得るために各リンクで通信される量子ビットの合計を解析する．

下限の導出のために，以下を用いる．

Lemma 3 長さ n のビット列 x と y が，任意のネットワークでつながっているプロセッサ P_i と

P_j に与えられるとする．この時，s を P_i と P_j の間の最短パスの長さ，w を与えられたネットワークのグラフの P_i と P_j の間の最大のカット数とする．この時，ブール関数 $f(x,y)$ を計算するためのすべてのリンクにおける合計通信量は，$\Omega(s(Q_2(f(x,y)) - \log n)/\log w)$ となる．ただし，$Q_2(f(x,y))$ は，$f(x,y)$ を計算する *2* 者間プロトコルで必要となる通信量とする．

証明は，量子プロトコルにおける puclic coin プロトコルの private coin プロトコルへの変換 [22] を利用して，古典の場合 [23] と同様に示すことができる．

下限を示す定理は以下の通りである．

Theorem 10 (Lower Bound) Distinctness の通信量は，$\Omega(n^{\frac{3}{2}})$ である．

証明：$(n-2(cn-1))$（c は $c\ (\leq 1/4)$ であるような定数）個のプロセッサによるリング状のネットワークにおいて，プロセッサ P_A と P_C がちょうど対極に位置しているとする．P_A と P_C が長さ cn のビット列 x と y を入力として持つとする．この時，x と y の関数 $\mathbf{DISJ} = \bigwedge_{i=1}^{cn} \overline{x_i y_i}$ を計算する問題を $\mathbf{DISJ}_{\text{ring}}$ と呼ぶことにする． Lemma 3において，$w = 2$ および $s = (n-2(cn-1))/2 = O(n)$ とすれば，$\mathbf{DISJ}_{\text{ring}}$ を解くプロトコルがすべてのリンクにおいて通信する通信量の合計は $\Omega(n\sqrt{n})$ であることがわかる．そのため，$\mathbf{DISJ}_{\text{ring}}$ から **Distinctness** へのリダクションが示せれば定理を証明できることになる．

リダクションは以下の通りである．まず，**Distinctness** におけるリング状の n プロセッサを，サイズが cn である A と C，サイズが $(n/2 - cn)$ である B と D の 4 つのグループに分ける．ここで A と C は対極に位置するとする．次に，$L = (1+c)n$ とし，3 つの整数の集合 $I_1 = \{0, 1, \ldots, cn-1\}$，$I_2 = \{cn, \ldots, 3cn-1\}$ および $I_3 = \{3cn, \ldots, L-1\}$ を考える．$\mathbf{DISJ}_{\text{ring}}$ の任意のインスタンスから以下の手順により **Distinctness** のインスタンスを生成すれば，両者の解が一致することは容易にわかる．

1. P_A（P_C）への入力の i（$i = 1, \ldots, cn$）ビット目が 1 ならば，グループ A（C）の i 番目のプロセッサの入力を $(i-1) \in I_1$ とする．そうでないならば，I_2 の中から重なりが起こらないように選んだ整数をグループ A（C）の i 番目のプロセッサの入力とする．

2. C と D のグループの各プロセッサの入力は，I_3 の中から重なりが起こらないように選んだ整数とする．

以上のリダクションは $L \geq (1+c)n$ の時に有効であり，$c \to 0$ とすれば，$L > n$ において定理が成立することがわかる．

次に上限の定理を示す．

Theorem 11 (Upper Bound) Distinctness の通信量は，

$$\min\{O(n\sqrt{n}(\log\log L + \log n), O(n\sqrt{L})\}$$

証明：以下では，通信量が $O(n\sqrt{n}(\log\log L + \log n)$ および $O(n\sqrt{L})$ であるアルゴリズムをそれぞれ示す．

Algorithm 1:

解くべき問題は，以下の探索問題と本質的に同じである．「P_i と P_j $(j \neq i)$ が同じ入力を持っているような i は存在するか？」

もし，入力 i に対して，P_i と同じ入力を持つ $P_j (\neq P_i)$ が存在するときのみ 1 を返すオラクルがあれば，上記の探索問題は，Grover Search により $O(\sqrt{n})$ 回のオラクル呼び出しで解く事が可能である．そこで，ある決められたプロセッサ（便宜上，Alice と呼ぶ）が上記の Grover Search をするようにすれば問題が解ける．そのため，上記のオラクルを Alice がどのように模倣できるかを示せば良い．i に対するオラクルの出力を，Alice は以下のプロトコルで模倣できる．

1. Alice は，最初の L^t 個の素数の中からランダムに一つの素数 p を選ぶ．(t は定数とする．そのため素数定理より，p のビット長は $O(\log\log L)$ となることに注意.)

2. Alice は，$(i,\ p)$ が書かれたトークンをリング上に一周させる．この時，各プロセッサは受け取ったトークンの中の番号が自分のプロセッサ番号と一致する時， 自分の持っている値 x_i に対して $(x_i \bmod p)$ をトークンに書き込む．

3. Aliceは，$(x_i \bmod p)$ を知った後，新たなトークン $(p, (x_i \bmod p), c)$ をリング上に一周させる．この時，各プロセッサはトークンを受け取ったら，自分の入力 x に対して $(x \bmod p)$ を計算し，受け取ったトークンの中の $(x_i \bmod p)$ と比較して等しい場合には，c を1増やす．c の初期値は0であり，c が2以上の場合は値を増やさないようにする．（補助ビットを使うことにより可逆計算でこの操作は可能である.）そのため，c は2ビットあれば十分である．

このプロトコルの最後で $c = 2$ となっていれば，オラクルの出力は1で $c = 1$ ならばオラクルの出力は0とすればよい．上記のプロトコルでは，各プロセッサは補助ビットを用いて計算するので，最後にトークンを逆周りに回して逆計算を行い，プロトコル中に使用した全ての補助ビットのクリアを行う．上記において，p と $x_i \bmod p$ のビット長は $O(\log\log L)$，i のビット長は $O(\log n)$ なので，プロトコル全体では $O(n(\log\log L + \log n))$ の通信量を必要とする．また，上記のプロトコルではオラクルの出力にエラーがあるが，オラクルが定数エラーの場合には Robust Grover Searh[33] を用いることにより，Grover Searh と同様の $O(\sqrt{n})$ 回のオラクル呼び出しで探索問題が解ける．以上より，全体の通信量は $O(n\sqrt{n}(\log\log L + \log n)$ となる．

Algorithm 2:

解くべき問題は，以下の探索問題と本質的に同じである．「少なくとも2つ以上のプロセッサがその値を持っているような $x \in \{0, \ldots, L-1\}$ が存在するか?」

入力 x に対して，x を入力値としてもつプロセッサが2つ以上あるときに1を返すオラクルを考える．Aaronson と Ambainis による Disjointness を計算する量子プロトコル [24] 同様に，このオラクルを $O(\sqrt{L})$ 回呼ぶことによって上記の問題を解く事が可能である．Aaronson と Ambainis による Disjointness を計算する量子プロトコル [24] では，上記のオラクルを実際に呼ぶときには（再帰的に分割して行う手法で，x の可能性が定数個になった時点で初めてオラクルを呼ぶため），x の値を各プロセッサに伝えるのに定数ビットしか必要としない．そのためオラクルのシミュレートは $O(n)$ ビットの通信ですむ．以上より，アルゴリズム全体の通信量は $O(n\sqrt{L})$ となる．

上記の通信量は $L = O(n)$ のときは下限と一致するのでほぼ最適な手法であると考えられる．

今後の課題としては，他のネットワーク形状や他の問題に対して量子プロトコルの通信量の上下限を求めることが考えられる．

8 量子情報を送信可能な量子秘密通信プロトコル

8.1 はじめに

1994年にショアによって多項式時間因数分解アルゴリズムが発表されたことにより，現在の公開鍵暗号は量子計算機によって解読可能であることが知られている．このため，量子計算機による攻撃に対しても安全なものとして，1984年にベネットらによって提案された量子鍵配布プロトコル (BB84)[26] が注目されている．このプロトコルは，秘密通信を行いたい二者間で無条件安全に乱数を共有する手法である．このプロトコルの発表以降，ここ20年の間，さまざまな量子暗号プロトコルに関する研究が行われてきた．B92[25], E91[28], GV95[29] は BB84[26] と同様量子情報技術を応用して安全に鍵配布を行うプロトコルである．また，岡本らにより量子公開鍵暗号 [31] も提案されており，量子計算機を使っても解くのが困難と予想される問題に安全性を帰着している．さらに，河内らによって量子状態の判別の困難さとそれを応用した計算量的に安全な量子秘密通信プロトコルの提案が行われている [30]．さらに近年，鍵配布ではなく，二者間で直接秘密情報を通信するプロトコルが研究されている [27]．本研究では，新しい量子直接秘密通信プロトコルを提案する．提案プロトコルは，古典情報だけでなく量子情報も送信可能であり，また，光子数分割攻撃に強い，エンタングルメントを使用しないといった特徴を持つ．

8.2 提案プロトコル

以下に，提案プロトコルを示す．

$U_\theta = \begin{pmatrix} cos\theta & sin\theta \\ -sin\theta & cos\theta \end{pmatrix}$ とする．

プロトコル

1. アリスは秘密量子ビット $|\psi_s\rangle$ と $k-1$ 個の囮量子ビット $|\phi_1\rangle, \ldots, |\phi_{k-1}\rangle$ を用意し，ランダムな順に並べる．ここで，囮ビットの状態はランダムに決める．アリスが用意した状態を一般性を失うことなく，$|\psi_0\rangle = |\psi_s \phi_1 \ldots \phi_{k-1}\rangle$ と書くことにする．

2. アリスはユニタリ変換 $U_{\hat{\theta_A}} = U_{\theta_A} \otimes U_{\alpha_1} \otimes \cdots \otimes U_{\alpha_{k-1}}$ をランダムに用意する．ここで，θ_a，$\alpha_1, \ldots, \alpha_{k-1}$ はランダムに選んだ回転角である．$d_A = \alpha_1 \ldots \alpha_{k-1}$ とし， また，$U_{d_A} = U_{\alpha_1} \otimes \cdots U_{\alpha_{k-1}}$ とする．アリスは $|\Psi_1\rangle = U_{\hat{\theta_A}} |\Psi_0\rangle$ をボブに送る．

3. ボブはアリスと同様にユニタリ変換 $U_{\hat{\theta_B}} = U_{\theta_B} \otimes U_{\beta_1} \otimes \cdots \otimes U_{\beta_{k-1}}$ をランダムに用意する．ここで，d_B および U_B を d_A，U_A と同様に定める．$|\Psi_2\rangle = U_{\hat{\theta_B}} |\Psi_1\rangle$ をアリスに送る．

4. アリスはボブに何番目のビットが秘密ビットであったかを教える．これに対してボブは，囮ビットに適用した操作，つまり d_B をアリスに教える．アリスは囮ビットに対して $U^\dagger_{d_A}$ および $U^\dagger_{d_B}$ を適用した後，囮ビットが初期状態 $|\phi_1\rangle, \ldots, |\phi_{k-1}\rangle$ に戻っているかどうかを調べる．この時点での量子状態を $|\Psi'_2\rangle$ とする．アリスは $|\Psi_3\rangle = (U^\dagger_{\theta_A} \otimes I^{\otimes k-1}) |\Psi'_2\rangle$ という状態を用意し，ボブに送る．このとき，ビットの順序をランダムに入れ替えて送信する．

5. ボブは $|\Psi_3\rangle$ を受け取った後，アリスにビットの順序および各囮ビットの初期状態を問い合わせる．この情報を元に，ボブは囮ビットに改竄がないことを確かめる．改竄がなければ，$U^\dagger_{\theta_B}$ を秘密ビットに適用し，元の秘密情報 $|\psi_s\rangle$ を得る．

ここで，U_θ は可換であることに注意されたい．そのため，$U_{\theta^\dagger_B} U_{\theta^\dagger_A} U_{\theta_B} U_{\theta_A} = I$ となり，ボブは正しく秘密情報を復号できる．また，U_{θ_A}, U_{θ_B} は秘密鍵の役割をなしており，このため，盗聴者が通信路を盗聴しても情報を得ることができない．さらに，これら秘密鍵をアリスとボブで交換する必要がないことも提案手法の特徴である．

以下に，盗聴者 (イブ) が存在した場合の提案プロトコルの振る舞いについて説明する．今，イブは受動的攻撃者であるとする．イブは通信路を流れる量子ビットのうち，どれが囮ビットであるかを知らない．そのためイブができることは，通信路を流れる量子ビットをランダムに抽出して観測するだけである．しかし，仮に秘密ビットを抽出して観測したとしても，秘密ビットの状態はイブにとっては最大混合状態であるため，元の秘密に関する情報は全く得られない．また，仮に囮ビットを観測したとすると，そのことが 1/2 の確率で検出される．

次にイブが能動的攻撃者である場合を考える．このときイブはステップ 2 において，$|\Psi_1\rangle$ を奪い，ステップ 3 でのボブからアリスへの通信時に，奪った $|\Psi_1\rangle$ を送り返すことができる．しかし，この場合，ステップ 4 でアリスは送られてきた量子ビットがボブからのものであるかどうかを確認するため，イブがこの確認をすり抜けるためには，ステップ 2 において，秘密ビットのみを奪う必要がある．これが成功する確率は $1/k$ である．この確率をさらに下げるためにはボブが $|\Psi_2\rangle$ を送信する際にビットの並び順を置換すればよい．これにより，イブの攻撃の成功確率は $1/k^2$ になる．さらに，アリスとボブの間で提案プロトコルを繰り返すことにより，イブの攻撃の成功確率を指数的に下げることが可能である．

8.3 結び

本研究では，量子状態を送信できる量子直接秘密通信プロトコルを提案した．安全性についての証明や，多量子ビットの通信への拡張などが今後の課題として挙げられる．

9 エラーのある量子オラクルによるロバストな計算

現在までに知られている量子アルゴリズムは，エラー（物理的なノイズのエラーのことではない）を含むものが多い．そのような量子アルゴリズム

をサブルーチンとして呼び出しながら全体の計算を設計することを考えた時には，以下のような量子オラクル O_f^ε を考えてそれへの質問量を議論することが自然である．

Definition 7 ε バイアスでブール関数 f 計算する量子オラクルは，O_f^ε と表記され，以下のユニタリー変換として定義される．

$O_f^\varepsilon |x\rangle |0^{m-1}\rangle |0\rangle$
$= |x\rangle (\alpha_x |w_x\rangle |f(x)\rangle + \beta_x |w'_x\rangle \left|\overline{f(x)}\right\rangle)$ ここで，$|\alpha_x|^2 = 1/2+\varepsilon_x \geq 1/2+\varepsilon$ $(x \in [N])$ かつ $\varepsilon_{min} = \min_x \varepsilon_x$ とする．

なお，すべての x に対して $0 < \varepsilon \leq \varepsilon_{min} \leq \varepsilon_x \leq 1/2$ に注意する．通常，ε は明示的に与えられてアルゴリズムを設計するが，ε_{min} や ε_x は与えられないと仮定するのが一般である．

入力値が上記のオラクルのモデルで与えられた時に，N-bit OR の関数を $O(\sqrt{N}/\varepsilon)$ の質問量で計算するアルゴリズムを以下に示す．このアルゴリズムの質問量は，$\varepsilon = \varepsilon_{min}$ の場合には，知られている下限と一致するので最適である．

以下では，オラクル O_f^ε を以下の Lemma に示すオラクル $\tilde{O}_f^\varepsilon$ に変換して考える．なお，変換の構成方法は，文献 [32] の Lemma 1 と同様である．

Lemma 4 O_f^ε と $O_f^{\varepsilon\dagger}$ を *1* 回ずつ用いて，以下の変換*:* $\tilde{O}_f^\varepsilon$ を実現可能である．

$$\tilde{O}_f^\varepsilon |x, 0^m, 0\rangle = (-1)^{f(x)} 2\varepsilon_x |x, 0^m, 0\rangle + |x, \psi_x\rangle . \tag{9.1}$$

ここで，$|x, \psi_x\rangle$ はノルムが $\sqrt{1-4{\varepsilon_x}^2}$ で，$|x, 0^m, 0\rangle$ と直交する．

以下の Lemma は，Amplitude Estimation が量子並列に動作可能であることを意味する．

Lemma 5 $\mathcal{O}$ を観測を用いない以下のように表される量子オラクルとする．

$$\begin{aligned} \mathcal{O} |x\rangle |\mathbf{0}\rangle &= |x\rangle \mathcal{O}_x |\mathbf{0}\rangle = |x\rangle |\Psi_x\rangle \\ &= |x\rangle (\left|\Psi_x^1\right\rangle + \left|\Psi_x^0\right\rangle). \end{aligned}$$

ここで，状態 $|\Psi_x\rangle$ は，ブール関数 χ により望む状態 $\left|\Psi_x^1\right\rangle$ か望まない状態 $\left|\Psi_x^0\right\rangle$ かを判別できるとする．また，$\sin^2(\theta_x) = \langle\Psi_x^1|\Psi_x^1\rangle$ $(0 \leq \theta_x \leq \pi/2)$ を $\mathcal{O}_x |\mathbf{0}\rangle$ の成功確率と考える．（望む状態が観測で得られれば正しい答えが得られたことに相当すると考える．）この時，以下のような状態の変換を行う量子 $Par_Est_Phase(\mathcal{O}, \chi, M)$ を $\mathcal{O}$ とその逆変換を $O(M)$ 回利用して構成することができる．

$$|x\rangle |\mathbf{0}\rangle |\mathbf{0}\rangle \longmapsto |x\rangle \otimes \sum_{j=0}^{M-1} |\delta_{x,j}\rangle |v_{x,j}\rangle \left|\tilde{\theta}\ Par_Est_Phase(\mathcal{O}, \chi, M)_{x,j}\right\rangle .$$

ここで，任意の x に関して $\sum_{j:|\theta_x-\tilde{\theta}_{x,j}|\leq\frac{\pi}{M}} |\delta_{x,j}|^2 \geq \frac{8}{\pi^2}$ が成立し，任意の i, j に対して $|v_{x,i}\rangle$ と $|v_{x,j}\rangle$ は互いに直交する．

以上の準備により，以下の Lemma を示すことができる．

Lemma 6 $O_f^{1/6}$ の変換を $O(1/\varepsilon)$ 回の O_f^ε への質問によりシミュレートする量子アルゴリズムが存在する．

証明： O_f^ε を $O(1/\varepsilon)$ 回用いて，以下のような状態の変換を行う量子アルゴリズムが構築可能であることを示す．

$|x\rangle |\mathbf{0}\rangle |0\rangle \longmapsto$
$|x\rangle (\alpha_x |w_x\rangle |f(x)\rangle + \beta_x |w'_x\rangle \left|\overline{f(x)}\right\rangle)$.

（すべての x に対して $|\alpha_x|^2 \geq 2/3$ が成立するとする．）

アルゴリズムは以下の通りである．

以下では，$\sin\theta = 2\varepsilon$ かつ $\sin\theta_x = 2\varepsilon_x$ $(0 < \theta, \theta_x \leq \pi/2)$ とする． この時，$0 \leq \theta \leq \pi/2$ の時 $\sin\theta \leq \theta \leq \frac{\pi}{2}\sin\theta$ のため，$\Theta(\theta) = \Theta(\varepsilon)$ であることに注意する．また，$M_1 = \left\lceil \frac{3\pi(\pi+1)}{\theta} \right\rceil$ とし，χ を式 (9.1) にあらわれる状態の中で，望ましい状態 $(-1)^{f(x)} 2\varepsilon_x \left|0^{m+1}\right\rangle$ と望ましくない状態 $|\psi_x\rangle$ を区別するブール関数であるとする．χ は単に状態が $\left|0^{m+1}\right\rangle$ であるかないかをチェックするものなので，実装は容易である．Lemma 5より，$Par_Est_Phase(\tilde{O}_f^\varepsilon, \chi, M_1)$ により以下の変換ができる．

$$|x\rangle |\mathbf{0}\rangle |\mathbf{0}\rangle |\mathbf{0}\rangle \longmapsto |x\rangle \otimes \sum_{j=0}^{M-1} \delta_{x,j} |v_{x,j}\rangle \left|\tilde{\theta}_{x,j}\right\rangle |\mathbf{0}\rangle .$$

ここで，すべての x に対して $\sum_{j:|\theta_x-\tilde{\theta}_{x,j}|\le\frac{\theta}{3(\pi+1)}}|\delta_{x,j}|^2 \ge \frac{8}{\pi^2}$ であり，$|v_{x,i}\rangle$ と $|v_{x,j}\rangle$ は任意の i,j に関して互いに直交する．上記の状態の第 3 レジスタには，高い確率で良い精度の θ_x の近似値が得られている．第 4 レジスタ $|\mathbf{0}\rangle$ は以降の手順に使われるのに十分な大きさがあるとする．

以降の手順では，基本的には推定した $\tilde{\theta}_{x,j}$ の値を用いて Amplitude Amplification を行う． 文献 [32] と同様の手法で，まず $m^*_{x,j} = \left\lceil \frac{1}{2}\left(\frac{\pi}{2\tilde{\theta}_{x,j}}-1\right)\right\rceil$, $\theta^*_{x,j} = \frac{\pi}{4m^*_{x,j}+2}$, $p^*_{x,j} = \sin^2(\theta^*_{x,j})$ および $\tilde{p}_{x,j} = \sin^2(\tilde{\theta}_{x,j})$ を量子並列で計算し，最後の量子ビットに Hadmard 変換を行う．以上により以下の状態を得る．

$$|x\rangle\left(\sum_{j=0}^{M-1}\delta_{x,j}|v_{x,j}\rangle\left|\tilde{\theta}_{x,j}\right\rangle|m^*_{x,j}\rangle|\theta^*_{x,j}\rangle|p^*_{x,j}\rangle|\tilde{p}_{x,j}\rangle \otimes\left|0^{m+1}\right\rangle|0\rangle\otimes\frac{1}{\sqrt{2}}(|0\rangle+|1\rangle)\right).$$

次に，回転 $\mathbf{R}$: $\left|p^*_{x,j}\right\rangle|\tilde{p}_{x,j}\rangle|0\rangle \to \left|p^*_{x,j}\right\rangle|\tilde{p}_{x,j}\rangle\left(\sqrt{\frac{p^*_{x,j}}{\tilde{p}_{x,j}}}|0\rangle+\sqrt{1-\frac{p^*_{x,j}}{\tilde{p}_{x,j}}}|1\rangle\right)$ およびオラクル $\mathbf{O}=\tilde{O}^{\varepsilon}_f\otimes\mathbf{R}$ を考える．

$M_2 = \left\lceil\frac{1}{2}\left(\frac{3\pi(\pi+1)}{2(3\pi+2)\theta}+1\right)\right\rceil$, $\mathbf{Q} = -\mathbf{O}(\mathbf{I}\otimes\mathbf{S}_0)\mathbf{O}^{-1}(\mathbf{I}\otimes\mathbf{S}_\chi)$ とし，$\mathbf{O}$ と $\Lambda_{M_2}(\mathbf{Q})$ をこの順番で適用する．ここで，Λ_{M_2} はレジスタ $\left|m^*_{x,j}\right\rangle$ で制御される変換であり，最後の量子ビットが $|1\rangle$ だと $\mathbf{Q}$ がレジスタ $|x\rangle$ と $\left|0^{m+1}\right\rangle|0\rangle$ に適用される．次に $\mathbf{O}_x$ を $\mathbf{O}|x\rangle\left|0^{m+1}\right\rangle|0\rangle = |x\rangle\mathbf{O}_x\left|0^{m+1}\right\rangle|0\rangle$ という変換を行うユニタリ変換とする．$\mathbf{Q}_x = -\mathbf{O}_x\mathbf{S}_0\mathbf{O}_x^{-1}\mathbf{S}_\chi$ かつ任意の x,j に対して $m_{x,j} = \min(m^*_{x,j},M_2)$ とすると，以下のように状態を表現できる.（以下では，簡単のため最後の 3 レジスタのみを表記する．

$$\sum_{j=0}^{M-1}\frac{\delta_{x,j}}{\sqrt{2}}\times(\left|0^{m+1}\right\rangle|0\rangle|0\rangle + \mathbf{Q}_x^{m_{x,j}}\mathbf{O}_x\left|0^{m+1}\right\rangle|0\rangle|1\rangle), \tag{9.2}$$

式 (9.2) は以下のように書き直せる．$\sum_{j=0}^{M-1}\frac{\delta_{x,j}}{\sqrt{2}}\Big(\left|0^{m+1},0\right\rangle|0\rangle + \left((-1)^{f(x)}\gamma_{x,j}\left|0^{m+1},0\right\rangle+|\varphi_{x,j}\rangle\right)|1\rangle\Big)$. ここで，$|\varphi_{x,j}\rangle$ は $\left|0^{m+1},0\right\rangle$ と直交しノルムが $\sqrt{1-\gamma^2_{x,j}}$ である．

ここで，第 3 レジスタ $\left|\tilde{\theta}_{x,j}\right\rangle$ が $|\theta_x-\tilde{\theta}_{x,j}|\le\frac{\theta_x}{3(\pi+1)}$ を満たしているとする．この時，もし $|\theta_x-\tilde{\theta}_{x,j}|\le\frac{\theta_x}{3(\pi+1)}$ ならば，$m_{x,j}\le M_2$ であることは容易にわかる．そのため，$\mathbf{Q}_x$ は $m^*_{x,j}$ 回，すなわち第 4 レジスタで指定された回数だけ適用される．この時，文献 [32] の Lemma 2 と同様に，$\gamma_{x,j}\ge\sqrt{1-\frac{1}{9}}$ であることが示せる．

最後の量子ビットに Hadmard 変換を行うと，状態は以下のようになる．

$$\sum_{j=0}^{M-1}\frac{\delta_{x,j}}{2}((1+(-1)^{f(x)}\gamma_{x,j})\left|0^{m+2}\right\rangle|0\rangle + (1-(-1)^{f(x)}\gamma_{x,j})\left|0^{m+2}\right\rangle|1\rangle+|\varphi_{x,j}\rangle(|0\rangle-|1\rangle))\quad.$$

ここで，最後の量子ビットを観測して $|f(x)\rangle$ を得る確率は，

$$\sum_{j=0}^{M-1}\left(\left|\frac{\delta_{x,j}(1+\gamma_{x,j})}{2}\right|^2+\left|\frac{\delta_{x,j}\sqrt{1-\gamma^2_{x,j}}}{2}\right|^2\right) \ge \frac{1}{2}\sum_{j:|\theta_x-\tilde{\theta}_{x,j}|\le\frac{\theta}{3(\pi+1)}}|\delta_{x,j}|^2(1+\gamma_{x,j}) \ge \frac{2}{3}$$

となる．よって，最終の量子状態を

$$|x\rangle(\alpha_x|w_x\rangle|f(x)\rangle+\beta_x\left|w'_x\right\rangle\left|\overline{f(x)}\right\rangle$$

と表現したとき，任意の x に対して，$|\alpha_x|^2\ge 2/3$ が成立する．

以上の量子アルゴリズムが用いるオラクルへの質問量は，Amplitude Estimation M_1 と Amplitude Amplification M_2 に関するものである．よって，オラクル質問量は $O(\frac{1}{\theta})=O(\frac{1}{\varepsilon})$ である．

Lemma 6と文献 [33] により，$O(\sqrt{N}/\varepsilon)$ 回 O^{ε}_f を用いて，N-bit OR の関数を成功確率 2/3 以上で計算することが可能である．

同じ問題に対して，下限 $\Omega(\sqrt{N}/\varepsilon_{min})$ が文献 [32] の Theorem 6 より示せるため，上記のアルゴリズムは最適であることがわかる．

10 量子文字列封印における情報流出の上下界

10.1 量子文字列封印のモデル

本研究で取り扱う量子文字列封印のモデルを以下のように定める．

Alice は量子レジスタを二つ用意し，入力文字列を符号化して量子レジスタに書き込む．

$$\text{入力文字列 } i\text{:} \quad |0\rangle|0\rangle \longrightarrow \Sigma_j \alpha_j |\phi_j^i\rangle |\psi_j^i\rangle .$$

その後，Alice は二つ目のレジスタを公開する．つまり，

$$\rho_i = \Sigma_j |\alpha_j|^2 |\psi_j^i\rangle \langle\psi_j^i|$$

が公開された量子状態である．Bob は ρ_i に対して POVM を行い，文字列を復号する．Alice は ρ_i と手元にある一つ目のレジスタの両方に対して POVM を行い，開封確認をする．

上記のような量子文字列封印のモデルに対して，He も指摘している通り [13]，復元可能なビット数に関して，以下に示す自明な下界が存在する．

Lemma 7 復号化率 $1-\varepsilon$ の任意の n ビット量子文字列封印プロトコルに対して，開封確認の検出率が δ 以下で，$\delta\cdot(1-\varepsilon)\cdot n$ ビット復元する手法が存在する．

(証明)

確率 δ で正規の手順に従って復号し，確率 $1-\delta$ で何もしないという復号方法を用いればよい．

10.2 盗聴される情報量の上界

本節では，新しい量子文字列封印プロトコルを示し，復元できるビット数の上界が，補題 7 に一致することを示す．以下にそのプロトコルを示す．

入力 i $(0 \le i \le 2^n-1)$ をランダムに選んだ $X \in \{X_1, \ldots, X_{2^n}\}$ を用いて $\frac{1}{\sqrt{2}}(|i\rangle+|X\rangle)$ と符号化する．ここで，$\{|0\rangle \ldots, |2^n-1\rangle, |X_1\rangle, \ldots, |X_{2^n}\rangle\}$ は互いに直交するものとする． これを全て Bob に対して公開する．

Bob は $\{|0\rangle, \ldots, |2^n-1\rangle, |X_1\rangle \ldots, |X_{2^n}\rangle\}$ という基底を用いて観測を行う．よって復号化率は 1/2 である．

このプロトコルに対して，次の定理が成り立つ．

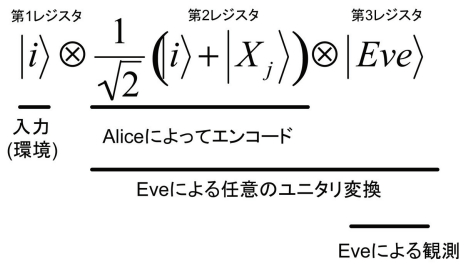


図 2: 攻撃のモデル

Theorem 12 上記のプロトコルに対して開封確認の検出率が δ 以下であるようなどのような攻撃についても，高々 $O(\delta n)$ ビットしか盗聴されない．

□

定理 12 の証明のために次のような盗聴のモデルを考える．(図 2 参照)

入力 $|i\rangle$ は $|0\rangle, \ldots, |2^n-1\rangle$ の一様な重ね合わせとし，これが第 1 レジスタに入っている．Alice は第 2 レジスタ，第 3 レジスタを使って入力を符号化し，第 3 レジスタを公開する．第 4 レジスタは盗聴者 Eve が使うアンシラとする．Eve は第 3 レジスタ及び第 4 レジスタに対してユニタリ変換 U_E をかけた後，第 4 レジスタを観測する．Alice は第 2, 第 3 レジスタを観測することにより Eve の存在を検出する．

上述のプロトコルでは，Alice は第 2 レジスタを使用しないため，以降第 2 レジスタは $|0\rangle$ と表記する．また，Alice は事前にランダムに $X \in \{X_1, \ldots, X_{2^n}\}$ を選択しておき，符号化の際に全ての i に対して同じ X を使用するものとする．このモデルでは第 1 レジスタが入力の量子的な重ね合わせになっているため実際の状況とは異なる．しかし，第 1 レジスタをトレースアウトすると，実際の状況と等価になり，また，Alice も Bob も第 1 レジスタにアクセスしないため，上記のモデルで安全性を議論して問題ないことに注意されたい．(図 3 参照)

上記のモデルでは Eve が観測できるレジスタを第 4 レジスタに制限しているが，この制限をおいても一般性を失わないことを示したのが以下の補題である．

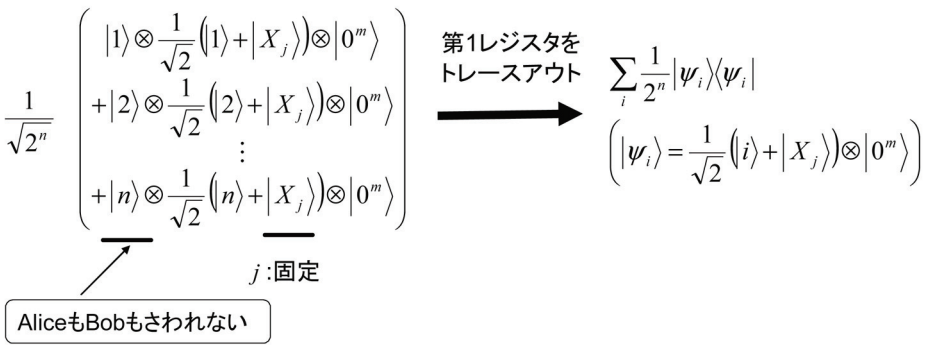


図 3: 攻撃モデルと実際の状況

Lemma 8 *Eve* は第 *4* レジスタのみを観測するとして一般性を失わない．

(証明)

Eve は十分多くのアンシラを持っているため，攻撃時に PVM のみを用いるとして一般性を失わない．ここで，Eve は攻撃時に第 3 レジスタを含めて観測を行うとする．この観測に使う射影作用素を $\{P_i\}$ とする．また，観測の直前の第 1 - 第 4 レジスタを $|\Phi\rangle = \Sigma_i \alpha_i |\phi_i\rangle$ と表記することにする．ただし $I \otimes I \otimes P_i |\phi_i\rangle = |\phi_i\rangle$ とする．このとき，アンシラを用意し，$I \otimes I \otimes U |\Phi\rangle |0\rangle = \Sigma_i \alpha_i |\phi_i\rangle |i\rangle$ となるようなユニタリ変換 U (U は第 3，第 4 レジスタにのみ作用するとする．) を施し，アンシラを観測することにより，元の観測をシミュレートできる．シミュレート後の状態からアンシラをトレースアウトしたものは，元の観測によって得られる密度行列に等しい．

上記のモデルにおいて，第 1 レジスタを観測した結果を表す確率変数を X，Eve が観測した結果を表す確率変数を Y とすると，相互情報量 $I(X:Y)$ が Eve が盗聴できる情報量を表している．$I(X:Y)$ の上界を与える補題として，Lo と Chau によって示された以下の補題 [19] がある．

Lemma 9 部分系 A，B からなる任意の純粋状態 $|\phi_{AB}\rangle$ が与えられたとする．A，B に対する任意の観測をそれぞれ X，Y とすると，各部分系のエントロピー $S(\rho_A)$ $(\rho_A = Tr_B |\phi_{AB}\rangle \langle\phi_{AB}|)$ は X と Y の相互情報量の上界を与える．

□

このことから直ちに，以下の系が得られる．

Corollary 3 観測直前の第 *1 - 3* レジスタのみに着目した量子状態を ρ_A，第 *4* レジスタのみに着目した量子状態を ρ_B とする．このとき *Eve* が得られる情報量は $S(\rho_A)$ 以下である．

□

さらに，定理 12 の証明のために，以下の補題を導入する．

Lemma 10

$$U_E |i\rangle |0\rangle = \Sigma_j a_j |X_j\rangle \left|w_{X_j}\right\rangle + \Sigma_j b_j |j\rangle |w_j\rangle$$

と表す．このとき，$|a_1| = |a_2| = \ldots = |a_{2^n}| (\leq \frac{1}{\sqrt{2^n}})$ として一般性を失わない．

(証明)

符号化時に，$X \in \{X_1, \ldots, X_{2^n}\}$ が一様ランダムに選ばれるため，

$$U_E |i\rangle |0\rangle = \Sigma_j a_j |X_j\rangle \left|w_{X_j}\right\rangle + \Sigma_j b_j |j\rangle |w_j\rangle$$

という攻撃の代わりに，j の置換 π をランダムに用意し，

$$U_E^{\pi} |i\rangle |0\rangle = \Sigma_j a_j \left|X_{\pi(j)}\right\rangle \left|w_{X_{\pi(j)}}\right\rangle + \Sigma_j b_j |j\rangle |w_j\rangle$$

という攻撃を行ったとしても，Eve にとって，検出率，盗める情報量に変化はない．π をランダムに選ぶため，攻撃後の状態は混合状態になるが，Eve に十分なアンシラを用意してこれを純粋化すると，

$$U_E |i\rangle |0\rangle = \Sigma_j a_j |X_j\rangle \left|w_{X_j}\right\rangle + \Sigma_j b_j |j\rangle |w_j\rangle$$

かつ，$|a_1| = |a_2| = \ldots = |a_{2^n}|(\leq \frac{1}{\sqrt{2^n}})$ という攻撃と等価になる．よって，$|a_1| = |a_2| = \ldots = |a_{2^n}|(\leq \frac{1}{\sqrt{2^n}})$ として一般性を失わない．

□

以下に定理 12 の証明を示す．

$$U_E |X\rangle |0\rangle = \alpha_1 |X\rangle |w_\alpha\rangle + \alpha_1' |\psi_{\alpha_1}\rangle |w_\alpha\rangle + \beta_1 |\psi_{\beta_1}\rangle$$

(ここで，任意の $|w\rangle$ に対して $|X\rangle |w\rangle \perp |\psi_{\beta_1}\rangle$，任意の $|\psi\rangle$ に対して $|\psi\rangle |w_\alpha\rangle \perp |\psi_{\beta_1}\rangle$，さらに，$|X\rangle \perp |\psi_{\alpha_1}\rangle$．) と表す．また，

$$U_E \left(\frac{1}{\sqrt{2}} |i\rangle + \frac{1}{\sqrt{2}} |X\rangle\right) |0\rangle = \alpha |\psi_\alpha\rangle |w_\alpha\rangle + \beta |\psi_\beta\rangle$$

と表す．(ここで，任意の $|\psi\rangle$ に対して，$|\psi\rangle |w_\alpha\rangle \perp |\psi_\beta\rangle$) 今，補題 10 より，

$$U_E |i\rangle |0\rangle = \Sigma_j \frac{e^{i\theta_{X_j}} a'}{\sqrt{2^n}} |X_j\rangle \left|w_{X_j}\right\rangle + \Sigma_j b_j |j\rangle |w_j\rangle$$

と書ける．ただし，$a' \leq 1$ である．ここで，

$$\frac{e^{i\theta_X} a'}{\sqrt{2^n}} |X\rangle |w_X\rangle = \frac{a''}{\sqrt{2^n}} |X\rangle |w_\alpha\rangle + \frac{a}{\sqrt{2^n}} |X\rangle \left|w_\alpha^\perp\right\rangle$$

とおくと，$|\psi_\beta\rangle$ は次のように書ける．

$$|\psi_\beta\rangle = \frac{a}{\sqrt{2}\sqrt{2^n}} |X\rangle \left|w_\alpha^\perp\right\rangle + b |i\rangle |w_i\rangle + c \left|\psi_\beta'\right\rangle$$

(ここで，任意の $|w\rangle$ について $|X\rangle |w\rangle \perp \left|\psi_\beta'\right\rangle$，$|i\rangle |w\rangle \perp \left|\psi_\beta'\right\rangle$，$|a| \leq a' \leq 1, |\frac{a}{\sqrt{2^{n+1}}}|^2 + |b|^2 + |c|^2 = 1$)

ここで，$|\psi_\beta\rangle$ を 第 3 レジスタが $(\frac{1}{\sqrt{2}} |i\rangle + \frac{1}{\sqrt{2}} |X\rangle)^\perp$ である空間に射影し，その大きさの 2 乗 (p とおく) を考える．これに $|\beta|^2$ をかけたものを d とすると，検出率は少なくとも d である．

以下，p の下界を求める．$c = 0$ かつ $\left|w_\alpha^\perp\right\rangle = |w_i\rangle (= |W\rangle)$ としたときが p が最も小さくなる．このとき

$$|\psi_\beta\rangle = \frac{a}{\sqrt{2^{n+1}}} |X\rangle |W\rangle + e^{i\theta} \sqrt{1 - \frac{a^2}{2^{n+1}}} |i\rangle |W\rangle$$

と書ける．このベクトルと，

$$\frac{1}{\sqrt{2}} |X\rangle |W\rangle - \frac{1}{\sqrt{2}} |i\rangle |W\rangle$$

というベクトルの内積の絶対値の 2 乗が p である．よって，

$$\begin{aligned}
p &= \left| \frac{a}{\sqrt{2^{n+2}}} - e^{i\theta} \sqrt{\frac{2^{n+1} - a^2}{2^{n+2}}} \right|^2 \\
&\geq \left(\frac{a}{\sqrt{2^{n+2}}} - \sqrt{\frac{2^{n+1} - a^2}{2^{n+2}}} \right)^2 \\
&\geq \left(\frac{1}{\sqrt{2^{n+2}}} - \sqrt{\frac{2^{n+1} - 1}{2^{n+2}}} \right)^2 \\
&= \frac{2^{n+1}}{2^{n+2}} - \frac{2\sqrt{2^{n+1} - 1}}{2^{n+2}} \\
&\geq \frac{1}{2} - \frac{1}{\sqrt{2^{n+1}}}
\end{aligned}$$

検出率が δ なので，$\delta \geq (\frac{1}{2} - \frac{1}{\sqrt{2^{n+1}}}) \cdot |\beta|^2$．よって $|\beta| \leq O(\sqrt{\delta})$．

以上より $|\alpha| \geq \sqrt{1 - O(\delta)}$ である．ここで，

$$U_E \left(\frac{1}{\sqrt{2}} |i\rangle + \frac{1}{\sqrt{2}} |X\rangle\right) |0\rangle = \alpha^i \left|\psi_\alpha^i\right\rangle |w_\alpha\rangle + \beta^i \left|\psi_\beta^i\right\rangle$$

と置き直し，前述の Eve の攻撃モデルを考えると，第 1 レジスタ - 第 4 レジスタまで含めた全体の状態 $|\Psi\rangle$ は，

$$\begin{aligned}
|\Psi\rangle &= \frac{1}{\sqrt{2^n}} \Sigma_i |i\rangle |0\rangle \left(\alpha^i \left|\psi_\alpha^i\right\rangle |w_\alpha\rangle + \beta^i \left|\psi_\beta^i\right\rangle\right) \\
&= A |\Psi_A\rangle |w_\alpha\rangle + B |\Psi_B\rangle
\end{aligned}$$

と書ける．ここで，$|A| \geq \sqrt{1 - O(\delta)}$ である．よって，第 4 レジスタをトレースアウトすると，

$$\rho_A = |A|^2 |\Psi_A\rangle \langle\Psi_A| + |B_1|^2 |\Psi_{B_1}\rangle \langle\Psi_{B_1}| + |B_2|^2 |\Psi_{B_2}\rangle \langle\Psi_{B_2}| + \ldots$$

と書ける．よって，

$$\begin{aligned}
S(\rho_A) &\leq -|A|^2 \log |A|^2 + \Sigma_j - B_j \log B_j \\
&\leq -|A|^2 \log |A|^2 \\
&\quad + \Sigma_1^{2^{n+1}} - \frac{(1 - |A|^2)}{2^{n+1}} \log \frac{(1 - |A|^2)}{2^{n+1}} \\
&= H(|A|^2) + (1 - |A|^2) \log 2^{n+1} \\
&\leq 1 + O(\delta)(n + 1)
\end{aligned}$$

以上より，$S(\rho_A) \leq O(\delta n)$ である．系 3 より，Eve が盗聴できる情報量は高々 $O(\delta n)$ である．

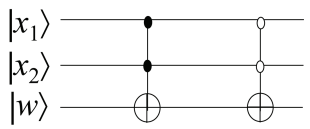


図 4: 量子回路 (1)

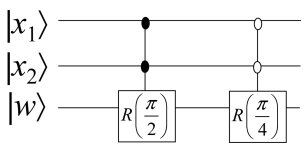


図 5: 量子回路 (2)

10.3 むすび

本研究では，量子文字列封印プロトコルを提案し，盗聴できる情報量の上界が自明な下界に一致することを示した．この意味で，本プロトコルは最適な量子文字列封印プロトコルと言える．本結果は，開封確認の検出率と盗聴される情報量のトレードオフを与えるものであり，提案プロトコルにおいて，検出率を δ とすると，盗聴される情報量は $O(\delta n)$ ビットであることが示された．

提案プロトコルの復号化率は 1/2 であるため，正規の復号手順を用いても誤って復号してしまう確率が高い．復号化率を高めた上で，盗聴される情報量の上界が，補題 7 に一致するプロトコルの開発等が今後の課題として挙げられる．

11 量子回路を二分決定グラフ構造から合成する手法

11.1 量子関数と行列関数

組織的な量子回路設計に役立つ概念として，量子関数と行列関数という概念を導入する．

図 4の回路は，3 ビット目の量子状態を $|w\rangle$ から $|w \oplus f(x_1, x_2)\rangle$ へ変換する．ここで，$f(x_1, x_2) = x_1 \cdot x_2 + \overline{x_1} \cdot \overline{x_2}$ である．そのため，この回路はブール関数 $f(x_1, x_2) = x_1 \cdot x_2 + \overline{x_1} \cdot \overline{x_2}$ を実現していると考えることができる．この種の量子回路は，有名な Grover Search [11] のオラクルとして用いることもでき，量子アルゴリズムの一部分としてよく利用される．

一方，図 5の量子回路では，3 ビット目の量子状態の出力をブール関数 $f(x_1, x_2)$ を用いて，$|w \oplus f(x_1, x_2)\rangle$ という形で表現することができない．たとえブール関数を実現する量子回路を設計対象としていても，部分的にこのような回路を利用することにより全体の回路規模が小さくなる場合があることが知られている．そのため，そのような量子回路も扱えるように新たに以下の定義を導入する．

表 1: 量子関数の真理値表

x_1, x_2	qf_2
$0, 0$	$R(\frac{1}{4}\pi)\,\|0\rangle$
$0, 1$	$\|0\rangle$
$1, 0$	$\|0\rangle$
$1, 1$	$R(\frac{1}{2}\pi)\,\|0\rangle$

Definition 8 n 変数 $x_1, x_2, \cdots, x_n$ 量子関数とは，$\{0, 1\}^n$ から一ビットの量子状態へのマッピングである．

図 5の回路の第 3 ビットを用いて説明する．もし，$|w\rangle$ の初期状態が $|0\rangle$ であれば，このビットの最終状態は，表 1に示される量子関数 $qf_2(x_1, x_2)$ となる．例えば，$x_1 = 0, x_2 = 0$ の時，最終量子状態は $R(\frac{1}{4}\pi)\,|0\rangle$ となるため，$qf_2(0, 0)$ が $R(\frac{1}{4}\pi)\,|0\rangle$ となっている．

表 2に図 4の第 3 ビット $|w\rangle$ の初期状態が $|0\rangle$ の時の最終状態を示す．これは，$|0\rangle$ と $|1\rangle$ をそれぞれブール値 0 および 1 と考えれば，ブール関数に対応していることがわかる．そのため，ブール関数は量子関数の特別なケースであることがわかる．

量子関数 $q(x_1, x_2, \cdots, x_n)$ の出力値は，$mf(x_1, x_2, \cdots, x_n)$ を $\{0, 1\}^n$ から二行二列の

表 2: ブール関数の真理値表

x_1, x_2	qf_1
$0, 0$	$\|1\rangle$
$0, 1$	$\|0\rangle$
$1, 0$	$\|0\rangle$
$1, 1$	$\|1\rangle$

表 3: 行列関数の真理値表

x_1, x_2	mf_2
$0,0$	$R(\frac{1}{4}\pi)$
$0,1$	I
$1,0$	I
$1,1$	$R(\frac{1}{2}\pi)$

表 4: 定数行列関数の真理値表

x_1, x_2	$CM(I)$	$CM(R(\frac{1}{2}\pi))$
$0,0$	I	$R(\frac{1}{2}\pi)$
$0,1$	I	$R(\frac{1}{2}\pi)$
$1,0$	I	$R(\frac{1}{2}\pi)$
$1,1$	I	$R(\frac{1}{2}\pi)$

ユニタリ行列へのマッピングとすると，常に $mf(x_1, x_2, \cdots, x_n)\,|0\rangle$ の形で表すことができるため，議論を容易にするため以下の定義を導入する．

Definition 9 n 個のブール変数 $x_1, x_2, \cdots, x_n$ に関する行列関数 mf とは，$\{0,1\}^n$ から二行二列のユニタリ行列へのマッピングである．

表 1の量子関数 qf_2 に対応する行列関数 $mf(x_1, x_2)$ を表 3に示す．

以下の定義も導入する

Definition 10 行列関数 $mf(x_1, x_2, \cdots, x_n)$ は，全ての変数割り当てに対してその出力が同じ行列 M となる場合，定数行列関数と呼び，$CM(M)$ で表す．

表 4に，$CM(I)$ と $CM(R(\frac{1}{2}\pi))$ を例として示す．

表 3の行列関数 mf_2 を用いれば，図 5の回路は，第 3 ビットを $|w\rangle$ から $mf_2(x_1, x_2)\,|w\rangle$ へ変換すると見ることができる．そのため，以降では回路の動作は行列関数で表現することとする．

行列関数に対し，演算子 “$\oplus$” を以下の通り定義する．

Definition 11 $mf_1(x_1, x_2, \cdots, x_n)$ と $mf_2(x_1, x_2, \cdots, x_n)$ を行列関数とすると $mf_1(x_1, x_2, \cdots, x_n) \oplus mf_2(x_1, x_2, \cdots, x_n)$ は，$\cdot$ を通常の行列の乗算演算として，$mf(x_1, x_2, \cdots, x_n) = mf_1(x_1, x_2, \cdots, x_n) \cdot mf_2(x_1, x_2, \cdots, x_n)$ となる行列関数 $mf(x_1, x_2, \cdots, x_n)$ と定義する．

表 5: 演算子 $\oplus$

x_1, x_2	mf_1	mf_2	$mf_1 \oplus mf_2$
$0,0$	$R(\frac{1}{2}\pi)$	$R(\frac{1}{2}\pi)$	$R(\pi)$
$0,1$	I	I	I
$1,0$	I	$R(\frac{1}{4}\pi)$	$R(\frac{1}{4}\pi)$
$1,1$	$R(\frac{1}{2}\pi)$	$R(\frac{1}{4}\pi)$	$R(\frac{3}{4}\pi)$

表 6: 演算子 $*$

x_1, x_2	mf_1	f_1	$f_1 * mf_1$
$0,0$	$R(\frac{1}{2}\pi)$	1	$R(\frac{1}{2}\pi)$
$0,1$	I	0	I
$1,0$	$R(\pi)$	1	$R(\pi)$
$1,1$	$R(\pi)$	0	I

例を表 5に示す．

次に演算子 “$*$” を以下の通り定義する．

Definition 12 $mf(x_1, x_2, \cdots, x_n)$ 行列関数，$f(x_1, x_2, \cdots, x_n)$ をブール関数とする．$f(x_1, x_2, \cdots, x_n) * mf(x_1, x_2, \cdots, x_n)$ は，$f(x_1, x_2, \cdots, x_n) = 1$ の時 $mf(x_1, x_2, \cdots, x_n)$，$f(x_1, x_2, \cdots, x_n) = 0$ の時 $CM(I)$ となる行列関数とする．

例を表 6に示す．

11.2 行列関数を表現する二分決定グラフ

通常のブール関数と同様 [6] に量子関数も二分決定グラフにより効率よく表現可能である．そのため，我々は以下のような *Decision Diagram for a Matrix Function (DDMF)* と呼ぶ二分決定グラフを導入する．

Definition 13 行列関数を表現する二分決定グラフ（*DDMF: Decision Diagram for a Matrix Function*）は，以下の *3*種類のノードからなる有向非巡回グラフである．

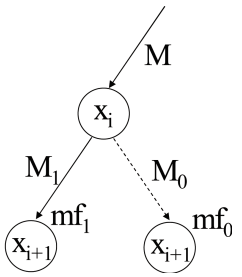


図 6: DDMF の非終端ノード

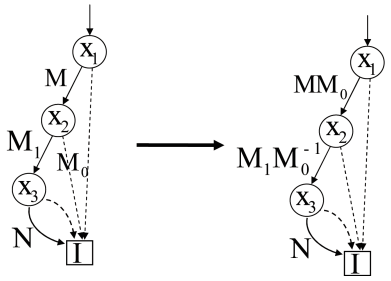


図 7: 標準形への変形

- 単位行列 I を表現する一つの終端ノード.
- 重み行列 M が付加された入力枝を持つ一つのルートノード.
- 複数の非終端ノード.

非終端ノードおよびルートノードは一つのブール変数 x_i と関連付けられ，別のノード *(1-child node)* につながっている出力枝：*1*-枝 *(*実線）および別のノード *(0-child node)* につながっている出力枝：*0*-枝 *(*点線）を持つ．各枝には重み行列が付加されている．

各ノードは以下の再帰的なルールにより行列関数を表現する．

- 終端ノードが表現する行列関数は $CM(I)$ である．
- 変数 x_i と関連付けられた非終端ノードおよびルートノードが表現する行列関数は，mf_1 と mf_0 を *1-child node* および *0-child node* のあらわす行列関数とし，M_1 と M_0 を *1*-枝および *0*-枝に付加された行列としたときに，$x_i*(CM(M_1)\oplus mf_1)\oplus\overline{x_i}*(CM(M_0)\oplus mf_0)$ となる．図 *6*に例を示す.
- ルートノードで表現される行列関数を mf，ルートノードへの入力枝に付加された行列を M とすると，*DDMF* により表現される行列関数は $CM(M)\oplus mf$ となる．

通常の BDD のように，以下の条件により DDMF の標準形を得ることができる．

Definition 14 変数の順序を固定した *DDMF* は以下の条件を満たすとき，同一の行列関数を表現するものは唯一となり，そのためそれを標準形とする．

- *0*-枝に付加される行列は I.
- *0*-枝と *1*-枝が同一のノードをさし，かつ *1*-枝に付加された行列が I であるようなノードは存在しない.（削除される.）
- 同じ部分グラフは存在しない.（共有される.）

終端ノードからルートノードに向けて順にノードを選択して，そのノードに対して以下の変形を行なえば，任意の DDMF は標準形となる．

変形の対象とするノードへの入力枝，0-枝および 1-枝に付加された行列を M, M_0 および M_1 とする．その時，もし M_0 が I でない場合，以下のように各枝に付加された行列を修正する．

- 入力枝に付加された行列を MM_0 とする．
- 1-枝に付加された行列を $M_1M_0^{-1}$ とする．
- 0-枝に付加された行列を I とする．

例を図 7に示す．例では行列 I は省略している．

便宜上 DDMF に対しても行列関数と同じ意味で演算子を次のように用いることにする．

- $mf = mf_1 \oplus mf_2$ の時，(mf を表す DDMF) = (mf_1 を表す DDMF) $\oplus$ (mf_2 を表す DDMF) と表現する．
- $mf = mf_1 * mf_2$ の時，(mf を表す DDMF) = (mf_1 を表す DDMF) $*$ (mf_2 を表す DDMF) と表現する．

11.3 行列関数からの量子回路設計

通常の回路設計と同様に，既に行列関数 mf_1 および mf_2 を実現する量子回路があれば，$mf_1 \oplus mf_2$ および $mf_1 * mf_2$ を実現する量子回路を実現できる．そのため，行列関数 mf を実現したい時には，mf よりも量子回路として実現しやすいような mf_1 および mf_2 により $mf_1 \oplus mf_2$ または $mf_1 * mf_2$ という形に mf を分解すればよいことになる．以下では，この 2 種類の分解に対して量子回路をどのように設計できるかを示す．

分解手法 1

$mf_1 \oplus mf_2$ という形に分解される行列関数を実現したい場合には，単純に以下により回路を合成可能である．

Proposition 1 行列関数 mf_1 および mf_2 で表現された変換を行なう量子回路を QC_1 および QC_2 とする．この時，QC_1 と QC_2 を連接して得られる量子回路は，$mf_1 \oplus mf_2$ で表現された変換を行なう量子回路である．

分解手法 2

既に行列関数 mf_1 と mf_2 で表現された変換を行なう量子回路が設計済みの時，$mf_1 * mf_2$ で表現される変換を行なう量子回路は以下の手順で合成可能である．

Step. 1 QC_1 を用意する．QC_1 が量子関数を計算するビットを w_1 とする．

Step. 2 QC_2 の各ゲートのコントロールに w_1 を加えた回路を用意し，Step. 1 で用意した回路の後に連接する．

Step. 3 QC_1 のゲートを逆順に並べた回路を用意し，Step. 2 で用意した回路の後に連接する．

mf_1 を実現する QC_1 および mf_2 を実現する QC_2 がそれぞれ図 4および図 5の時，$mf_1 * mf_2$ で表現される変換を行なう量子回路は図 8 に示すように合成される．

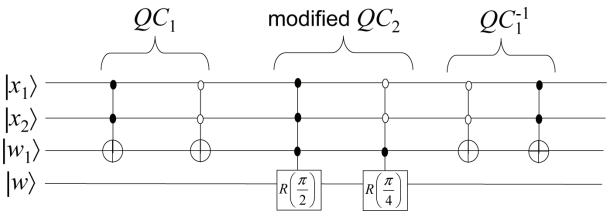


図 8: $mf_1 * mf_2$ の量子回路

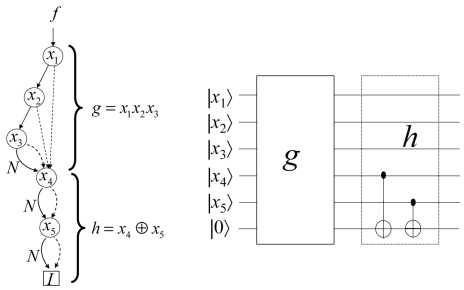


図 9: DDMF の分解: 簡単なケース

11.4 DDMF を用いた量子回路設計

DDMF の表現を利用して，制御-NOT ゲート，制御-V ゲートおよび制御-V^{-1} ゲートにより回路を組織的に合成する手法を以下に示す．

簡単なケース 1

(f を表す DDMF) = (g を表す DDMF) $\oplus$ (h を表す DDMF) を満たす f に対応する量子回路を設計するには，単純に f と g に対応する量子回路を連接すれば良い．図 9に，(DDMF for f) = (DDMF for g) $\oplus$ (DDMF for h) の時に f に対応する量子回路の設計例を示す．

簡単なケース 2

Definition 15 *DDMF の全てのノードの 1-枝と 0-枝が同じノードに向かっている時，その DDMF は chain と呼ばれる．*

もし，設計対象の DDMF が *chain* 構造を持つ場合，以下のように対応する量子回路を設計できる．

- 関連付けられた変数が x_i で 1-枝に付加された行列が M_i である各ノードに対して，x_i でコントロールされる制御-M_i ゲートを用意する．

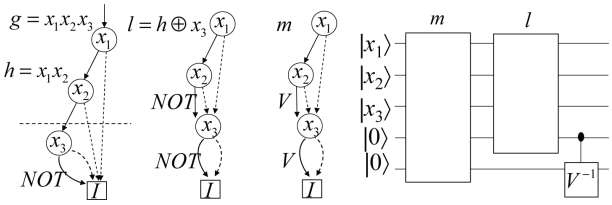


図 10: DDMF の分解: 複雑なケース

- 上記のゲート全てを連接すれば所望の回路を得る.

図 9の h を表す DDMF は chain 構造を持つので，図に示すように量子回路が設計できる.

複雑なケース

もし設計対象の DDMF の下部に chain 構造がなければ，上述した手法を用いることはできない．そのような場合も以下の手法により設計対象の DDMF をより簡単な DDMF に分解することが可能である.

今，設計対象の関数 g を表す DDMF の下部に chain 構造がないとする．そして，その部分の一番上に位置するノードを v_i とそれに関連付けられた変数を x_i とする．この時，以下のように設計対象の DDMF を分解できる.

- ルートノードから v_i に向かう全てのパスに対応するブール関数を h とする.
- $h \oplus x_i$ となるブール関数を l とする.
- l を表す DDMF の全ての枝に付加された行列 NOT を V に変更した DDMF の表現する量子関数を m とする.
- この m を実現する量子回路を用いて，図 10 のように所望の回路を設計することができる.

この設計手法では，l と m を表す DDMF は chain 構造をその下部に必ず持つことが保証できる．そのためこれらはより簡単な量子回路として設計できる．従って，上述した手法を再帰的に適用することにより量子回路を組織的に設計することが可能である.

既存の手法 [1] は設計すべき関数の入力数に対して指数的な精度のゲートを利用するので，提案手法のほうがより実用的であると考えられる.

12 関数の真理値表の 1 の数とそれを評価するための量子質問量との関係

12.1 はじめに

ブール関数を評価するための量子質問計算量（以下，単に質問量とする）は，量子計算の研究分野では最も基本的な研究トピックであり，今日まで多くの研究が行われている．そして，いくつかの OR, AND, AND-OR tree や NAND tree などの関数については，古典計算に対して，量子計算の質問量は平方倍になる（すなわち，N をブール関数の変数の数として，量子では $O(\sqrt{N})$ だが，古典では $\Omega(N)$）ことが知られている．しかし，一方で，Parity 関数を始めとして，ほとんど全ての関数に対しては，量子計算でも高々定数倍程度しか良くならない（具体的には、$\Omega(N)$）ことも知られている.

ブール関数を計算するための量子計算の質問量は，その両極端な場合である $O(\sqrt{N})$ と $O(N)$ については，上述した関数が知られているが，その間の質問量については良く知られていない．そこで，我々は，この 2 つの間の質問量に関して，ブール関数の on-set のサイズ（真理値表上での 1 の数）により質問量が特徴付けられることを示した.

以下では，特に断りがなければ N 変数で真理値表上の 1 の数が M であるようなブール関数 f を考える.

12.2 一般の関数に関する質問量

Theorem 13 *(General Upper Bound)* N 変数ブール関数 f の真理値表上の 1 の数が M であるとする．この時，ブール関数 f を定数確率で計算する質問量は，$O(\sqrt{N\frac{\log M}{\log N}})$ である．ただし，d を $(0 < d < 1)$ である任意の定数として，$poly(N) \leq M \leq 2^{N^d}$ が成り立つとする.

Remark. 文献 [5] の定理 4.10 より，閾値 $T = \log M/\log N$ の閾値関数に対しては，上記の定理の質問量はタイトである．

我々のキーアイデアは、オラクル同定問題（以下, OIP (*oracle identifying problem*) [3] のアルゴリズムを用いることである．OIP は, オラクルの候補が与えられた時に，質問によりブラックボックスのオラクルを１つに同定する問題である．オラクルは N ビットの値を持つとして，M 個のオラクル候補の問題は，$M \times N$ の行列で表現でき，$M \times N$ OIP と呼ぶ.

Lemma 11 (Ambainis et al. [3]) *$M \times N$ OIP は，質問量 $O(\sqrt{N\frac{\log M}{\log N}})$ で解くことができる．ただし，d を $(0 < d < 1)$ である任意の定数として，$poly(N) \leq M \leq 2^{N^d}$ が成り立つとする.*

Theorem 13の証明の概略

OIP は，M 個のオラクルの候補が与えられた時に，質問によりブラックボックスのオラクルを１つに同定する問題である．これをブール関数の計算に使うために，まず現在のオラクルの N ビットの入力がブール関数の on-set であると仮定して，$M \times N$ OIP を解く．これには，質問量 $O(\sqrt{N\frac{\log M}{\log N}})$ を要する.

OIP のアルゴリズムにより得られた解（N ビットの列）を z とする．また実際のオラクルの入力 N ビットを x とする．もし，$f(x) = 1$ ならば，上記の仮定が成り立っているので，OIP のアルゴリズムは正しく動き，$O(\sqrt{N\frac{\log M}{\log N}})$ の質問量で得られた z は x と等しくなり，$f = 1$ とすればいい.

しかし，もし $f(x) = 0$ ならばそもそも上記の OIP の仮定が成り立たないため，上述したアルゴリズムでは不完全である．そのため，我々のアルゴリズムでは以下のように対処する.

もし，OIP の仮定が成り立たないオラクルに対して OIP のアルゴリズムを走らせた場合，(1) OIP のアルゴリズムが途中で想定していない状況に陥るか，(2) 実際のオラクルとは異なる関数の on-set のうちの一つのオラクルを解として出力するかのかのいずれかとなる．(１) の場合は，OIP の仮定が成り立っていないと分るので，$f(x) = 0$ と結論付ければ問題ないが，(2) の場合は実際の入力が OIP のアルゴリズムの返した解と同じかどうかを調べる必要がある.

そこで，我々のアルゴリズムでは，OIP のアルゴリズムの解を z とすると現在のオラクル入力が z と等しいかを調べる．これは，Grover サーチにより $O(\sqrt{N})$ の質問量でできるため，$O(\sqrt{N\frac{\log M}{\log N}})$ に比べて無視できる．これにより，もし $x = z$ なら $f = 1$，そうでないなら $f = 0$ と結論付ければ，定数確率で正しいことは明らかである.

以上より, $f = 1$ か 0 かを質問量 $O(\sqrt{N\frac{\log M}{\log N}})$ で判定できる.

□

Corollary 4 *ブール関数 f の真理値表上の 1 の数が $O(N^{O(1)})$ なら，f の計算する質問量は $O(\sqrt{N})$ である.*

以下の定理より，この結果はタイトである.

Theorem 14 *(General Lower Bound) 真理値表上の 1 の数が，高々 $M = 2^{\frac{N}{2+\epsilon}}$ 個しかない任意の関数を計算するための質問量は，$\Omega(\sqrt{N})$ である.（ϵ は任意の正の定数とする.）*

この下限を証明するために，以下を用いる.

Lemma 12 (Beals et al. [5]) *T 回の質問をする量子アルゴリズム A があるとする．この時，このアルゴリズムの出力する解が 1 である確率は，次数が高々 $2T$ のマルチリニア多項式 $p_A(x_1, \ldots, x_N)$ で表現できる.*

Lemma 13 (Ambainis [2]) *$D(l, m) = \sum_{i=0}^{m} \binom{l}{i}$ とすると，高々 $2^{O(D(N,d)dN^2)}$ 個の関数が次数が d のマルチリニア多項式で近似できる.*

以下は，[17] および [18] の結果から得られる一般的な下限を証明するためのツールである.

Lemma 14 (Khrapchenko bound) *$Ham(x, y)$ を x と y のハミング距離とし，関数 f について 3 つの集合を $A \subseteq f^{-1}(1)$，$B \subseteq f^{-1}(0)$，$C = \{(x, y) \in A \times B \mid Ham(x, y) = 1\}$ として選ぶと，$Q(f) \geq \Omega(\sqrt{\frac{|C|^2}{|A||B|}})$ となる.*

Theorem 14の証明の概略

以下では，関数の値を1（0）とする入力（N ビットのオラクル）のことを 1-*point*（0-*point*）と呼ぶ．また，ハミング距離が1の入力を ”周り ”の入力と呼ぶことにする．

Lemma 14から（block-sensitivity による下限 $Q(f) \geq \Omega(\sqrt{bs(f)})$ [5] でも同様に），B を一つの 1-*point* の周りの 0-*point* の集合とすれば，下限が $\Omega(\sqrt{|B|})$ となることがわかる．

そのため，もし，その周りに $\Omega(N)$ 個の 0-*point* がある 1-*point* が存在すれば，下限は $\Omega(\sqrt{N})$ となる．

全ての 1-*point* の周りに，$o(N)$ 個の 0-*point* があると仮定し，以下でその矛盾を示す.（つまり，$N - o(N)$ 個の 1-*point* が周りにある.）一般性を失うことなく，入力 $(0, 0, .., 0)$ を 1-*point* とする．以下のように再帰的に定義される集合 S_k を考える．$S_0 = \{(0, 0, .., 0)\}$．S_k をハミング重みが k で，S_{k-1} の全ての 1-*point* の周りの全ての 1-*point* の集合とする．

仮定より，S_{k-1} の各 1-*point* の周りには，α を十分小さい定数として，$N - o(N) = N(1-\alpha)$ 個の 1-*point* が存在する．また，S_{k-1} の各 1-*point* の周りには，高々$(k-1)$ 個の S_{k-2} に含まれる 1-*point* しか存在しない．そのため，S_{k-1} の各 1-*point* の周りには，少なくとも $(N(1-\alpha) - (k-1))$ 個の S_k に含まれる 1-*point* が存在する．同様に，S_k の各 1-*point* の周りには，高々k 個の S_k に含まれる 1-*point* しか存在しない．

以上より，$|S_k| \geq |S_{k-1}|(N(1-\alpha) - (k-1))/k$ が成立する．

$|S_0| = 1$ であるので，上記より，

$|S_k| \geq (N(1-\alpha))(N(1-\alpha) - 1)(N(1-\alpha) - 2) \cdots (N(1-\alpha) - (k-1))/k!$

となる．

$f(x) = 1$ とする入力 x のうち，ハミング重みが k を超えないものを $T(k) = |S_1| + \ldots + |S_k|$ とする．

もし，$T(k) > M$ がある k で成立すれば，真理値表上の1の数が M 個であるという事実に矛盾するため，$\Omega(\sqrt{N})$ の下限を示せることになる．

$k \leq N/2$ として，$T(k)$ は以下のように抑えることができる．

$$\begin{aligned} T(k) &> |S_k| \\ &\geq (N(1-\alpha))(N(1-\alpha) - 1) \cdots \\ &\quad (N(1-\alpha) - (k-1))/k! \\ &> \left(\frac{N(1-\alpha)}{k}\right)^k \end{aligned}$$

以上より，$k = \frac{N}{2+\epsilon}$（ϵ は任意の正の定数）に対して，$T(k) \geq 2^{\frac{N}{2+\epsilon}} > M$ が成り立つ．従って，質問量は $\Omega(\sqrt{N})$ となる．

□

以上より，$M = O(2^{N^c})$ $(0 < c < 1)$ である全てのブール関数に対して，質問量の下限が $\Omega(\sqrt{N}) = \Omega(N^{1/2})$ で，上限が $O(\sqrt{N \frac{\log M}{\log N}}) = O(\frac{N^{1/2+c/2}}{\sqrt{\log N}})$ となる．これは，書き換えると下限は $\Omega(N^d)$ $(d = 1/2)$ で，上限が $O(N^d)(d = 1/2 + c/2)$ となる．この両者の開きは，以下の定理より許容できるものである．

Theorem 15 $M = 2^{N^c} + o(2^{N^c})$ が成り立つとき $(0 < c < 1)$，$1/2 \leq d < 1/2 + 1/2 \log_N(N^c/\log N)$ を満たす任意の定数 d に対して，質問量が $\Theta(N^d)$ となるブール関数が存在する．

Theorem 15の証明の概略

$f(x) = 1$ となる条件が，(i) x の最初の N^c ビットが 0^{N^c} でなく，残りの $N - N^c$ ビットが 1^{N-N^c} である場合か，(ii)x のハミング重みが N^{2d-1} より小さいか，のどちらかであるような関数を考える.（それ以外の場合は $f(x) = 0$ となる.）ここで，$0 \leq 2d - 1 < \log_N(N^c/\log N)$ $(0 < c < 1)$ に注意する．

すると，$M = |\{x \mid f(x) = 1\}|$ は，条件 (i) による $2^{N^c} - 1$ と条件 (ii) による $O(\binom{N}{N^{2d-1}}) = O(N^{N^{2d-1}}) = o(2^{N^c})$ の和となる．従って，$M = 2^{N^c} + o(2^{N^c})$ となる．

以下で，f を計算するための質問量を評価する．

条件 (i) により $f(x) = 1$ になるかどうかは，最初の N^c ビットと残りの $N - N^c$ ビットそれぞれに別々に Grover サーチを行えばわかる．この質問量は $O(N^{1/2})$ である．条件 (ii) に

より $f(x) = 1$ になるかどうかは，閾値関数を評価する量子アルゴリズムを用いて，$k \leq N/2$ を閾値として，$O(\sqrt{k(N-k)})$ の質問量で判定可能である．$k = N^{2d-1}$ とすると，質問量は $O(\sqrt{N^{2d-1}(N-N^{2d-1})}) = O(N^d)$ となる．以上より，f を計算するための質問量は $O(N^{1/2}) + O(N^d) = O(N^d)$ となる．

一方，以下の理由により，f を計算するための質問量の下限も $\Omega(N^d)$ となる．ハミング重みが高々 N^{2d-1} の入力で 1 となる閾値関数の評価は，上記の関数 f の評価と Grover サーチを組み合わせれば可能である．そのため，上記の関数の質問量は，ハミング重みが高々 N^{2d-1} の入力で 1 となる閾値関数の評価に必要な質問量の下限 $\Omega(N^d)$ で下から抑えることができる．

□

12.3 ほぼすべての論理関数に関する質問量

本節では，ほぼすべての論理関数についてその真理値表における 1 の個数によって質問量の特徴づけを行う．

Theorem 16 *(Randomized Upper Bound)* 真理値表に M 個の *1* を持つほぼすべての論理関数について，その質問量は $O(\log M + \sqrt{N})$ である．ただし，$M < 2^{\frac{N}{2+d}}$ とし，d は任意の正の定数とする．

Theorem 16 の証明の概略

定理を証明するために，次の考察を導入する．(証明は後述)

考察. 真理値表に M 個の 1 をもつ論理関数をランダムに生成したとする．このとき，ほぼすべての場合において，$f = 1$ を与える M 個の入力のうち，任意の 2 つの入力ついて，最初の k ビットが互いに異なる．ここで，$k = (2+d)\log M$ とし，d は正の定数とする． 以下 $k = (2+d)\log M$ とする．

この考察に基づいて，次のようなアルゴリズムを考える．オラクルに k 回問い合わせを行うことにより，入力の最初の k ビットを得る．このとき，場合によっては残りのビット値によらず，$f(x) = 0$ と結論付けられる場合がある．もし，そうでなければ，次のように残りの $N-k$ ビットに対して問い合わせを行う．上記の考察により，ほぼすべての関数について，$f = 1$ を与える残りの $N-k$ ビットの割り当てはひとつしか存在しない．よって残りのビット値がそのようなビット値になっているかどうかをグローバー探索アルゴリズムによって確認する．以上より，ほぼすべての関数について質問量は $O(k + \sqrt{N}) = O(\log M + \sqrt{N})$ となる．

以下に上記の考察の証明を与える．

真理値表に M 個の 1 を持つ関数の個数は $\binom{2^N}{M}$ である．そのような関数の中で，所望の関数，つまり，$f = 1$ を与える M 個の入力の任意の二つについて，最初の k ビットが異なる関数の個数を数える．まず最初に，k ビットの値に対する M 個の入力を割り当て方の個数を考えると，$\binom{2^k}{M}$ となる．次に，残りの $N-k$ ビットを任意に割り当てると，M 個の各入力に対して，$2^{(N-k)}$ 通りの割り当て方が存在する．よって，M 個の入力すべてに対して残りの $N-k$ ビットへの値の割り勝て方は，$(2^{(N-k)})^M = 2^{M(N-k)}$ 通り存在する．よって，所望の関数の個数は $\binom{2^k}{M}2^{M(N-k)}$ である．

よって，所望の関数の全体に対する割合は，以下のようになる．

$$
\begin{aligned}
& \frac{\binom{2^k}{M}2^{M(N-k)}}{\binom{2^N}{M}} \\
> \; & \frac{M!\binom{2^k}{M}2^{M(N-k)}}{2^{NM}} \\
= \; & \frac{(2^k)^{\underline{M}}2^{M(N-k)}}{2^{NM}} \\
= \; & \frac{(2^k)^{\underline{M}} \times 2^{MN-Mk}}{2^{Mk} \times 2^{MN-Mk}} \\
= \; & \frac{(2^k)^{\underline{M}}}{(2^k)^M} \\
= \; & (1-\frac{1}{2^k})(1-\frac{2}{2^k})\cdots(1-\frac{M-1}{2^k}) \\
\geq \; & \exp(\sum_{i=1}^{M-1}\frac{-\frac{i}{2^k}}{1-\frac{i}{2^k}}) \; (\because 1-x \geq \exp(\tfrac{-x}{1-x})) \\
= \; & \exp(\sum_{i=1}^{M-1}\frac{-i}{2^k-i})
\end{aligned}
$$

$$
\begin{aligned}
&> \exp\left(\frac{-1}{2^k-(M-1)}\sum_{i=1}^{M-1} i\right) \\
&> \exp\left(\frac{-1}{2^k}\times\frac{M(M-1)}{2}\right),
\end{aligned}
$$

ここで，$y^{\underline{x}} = y(y-1)\cdots(y-x+1)$ である．

よって，この割合は M が大きく，$k = (2+d)\log M$ (d は正の定数) であるとき，ほぼ 1 となる．

□

Theorem 16 は次の定理より，ほぼ厳密であるといえる．

Theorem 17 *(Randomized Lower bound)* $M \leq 2^{N-o(N)}$ とする．ほぼすべての関数について，論理関数の真理値表が M 個以上の *1* を持つとき，その関数に対する質問量は $\Omega(\log M/\log N+\sqrt{N})$ となる．

Theorem 17 の証明の概略

Theorem 15 において，$M > 2^{\sqrt{N}}$ のとき，ほぼすべての関数について，質問量の下界が $\Omega(\log M/\log N)$ となることを示した．よって，$M > 2^{\sqrt{N}}$ のとき，質問量の下界が $\Omega(\log M/\log N)$ になることを示せば十分である．

(Case: $M \leq 2^{cN}$ with $c < 1$) 真理値表に 1 を M 個持つ関数の個数は以下のように計算できる．

$$
\begin{aligned}
\binom{2^N}{M} &\geq (2^N/M)^M = 2^{NM-M\log M} \\
&\geq 2^{(1-c)NM} = 2^{\Omega(NM)}. \qquad (12.1)
\end{aligned}
$$

一方，Lemma 12 および Lemma 13 より，質問量が l の関数の個数は高々次の通りである．

$$
2^{O(D(N,l)lN^2)} = 2^{O(lN^l lN^2)} = 2^{O(l^2N^{l+2})} = 2^{O(N^{l+4})}
$$

ここで，

$$
D(N,l) = \sum_{k=0}^{l}\binom{N}{k} \leq lN^l.
$$

である．

$l = o(\log M/\log N) = o(\log(NM)/\log N)$ $(N < M)$ とすると，上記の値は，

$$
2^{O(N^{l+4})} = 2^{o(NM)}. \qquad (12.2)
$$

となる．

等式 (12.1) と (12.2) を比較すると，$l = o(\log M/\log N)$ であれば，質問量が l の関数の個数と，真理値表に 1 を M 個持つ関数の個数の比は N を無限大にすると 0 に近づく．よって，真理値表の 1 の個数が M であるようなほぼすべての関数について，質問量は $\Omega(\log M/\log N)$ となる．

(Case: $M = 2^{N-o(N)}$) 真理値表に 1 を M 個持つ関数の個数は，次のように計算できる．

$$
\begin{aligned}
&\binom{2^N}{M} && (12.3)\\
&= \frac{2^N!}{(2^N-M)!M!} && (12.4)\\
&\geq \frac{\sqrt{2\pi 2^N}\left(\frac{2^N}{e}\right)^{2^N}}{\sqrt{2\pi(2^N-M)}\left(\frac{2^N-M}{e}\right)^{2^N-M}\cdot\sqrt{2\pi M}\left(\frac{M}{e}\right)^M} \\
&= \frac{\sqrt{2^N}}{\sqrt{2\pi}\sqrt{(2^N-M)M}}\left(\frac{2^N}{2^N-M}\right)^{2^N}\left(\frac{2^N-M}{M}\right)^M, && (12.5)
\end{aligned}
$$

ここで，不等式はスターリンの公式より得られることに注意されたい．$M = 2^{N-o(N)}$ であることに注意すると，

$$
\begin{aligned}
&\left(\frac{2^N}{2^N-M}\right)^{2^N} \\
&= \left(\frac{1}{1-2^{-o(N)}}\right)^{2^N} \\
&= \left(1+\frac{2^{-o(N)}}{1-2^{-o(N)}}\right)^{2^N} \\
&= \left(1+\frac{1}{2^{o(N)}-1}\right)^{2^N} \\
&\geq e^{2^{N-o(N)}},
\end{aligned}
$$

また，

$$
\frac{\sqrt{2^N}}{\sqrt{2\pi}\sqrt{(2^N-M)M}}\left(\frac{2^N-M}{M}\right)^M = (2^{o(N)}-1)^M > 1.
$$

よって，式 (12.5) より，以下を得る．

$$\binom{2^N}{M} \geq e^{2^{N-o(N)}} \geq 2^{2^{N-o(N)}}. \qquad (12.6)$$

一方，Lemma 12 および Lemma 13 より，質問量が l の関数の個数は高々次のように計算できる．

$$2^{O(D(N,l)lN^2)} = 2^{O(N^{l+4})}.$$

$l = o(\log M/\log N)$ とすると，

$$2^{O(N^{l+4})} = 2^{o(M)} = 2^{o(2^{N-o(N)})}.$$

上記と，式 Eq.(12.6) を比較すると，真理値表の 1 の個数が M の関数の個数と質問量が $o(\log M/\log N)$ の関数の個数の比は N を無限大にすると 0 に近づく．よって，真理値表の 1 の個数が M 個のほぼすべての関数の質問量は $\Omega(\log M/\log N)$ である．

□

Comments.

Lemma 13 は d が大きいときは厳密であり，小さいときはそうでないように思える．極端な例として，$d = 2$ の場合を考えると，次数 2 のマルチリニア多項式の個数はおよそ $2^{\binom{N}{2}} \approx 2^{N^2}$ である．$M = N$ となるような関数の個数は $\binom{2^N}{N} \approx (2^N)^N = 2^{N^2}$ であり，前述の値とほぼ等しくなる．しかしながら，2 ステップの量子アルゴリズムではほとんど何も計算できない．よって，量子アルゴリズムのステップ数が少ないときは，対象となるマルチリニア多項式の個数を減らすことができるような新しい性質を見つける必要があると思われる．もし，そのような性質が見つかれば，Theorem 16 と Theorem 17 のギャップを埋めることができる可能性がある．

12.4 むすび

これまでに示した結果を以下にまとめる．(図 11 参照)

(1) 真理値表の 1 の個数が少ない場合 (例えば多項式個の場合)，質問量は $\Theta(\sqrt{N})$ となる．つまり，質問量は真理値表の 1 の個数によって完全に記述できる．

(2) そうでない場合，質問量の上界は $O(\sqrt{N\frac{\log M}{\log N}})$ となる．また，この上界は，閾値関数のような特定の関数については厳密であることを示した．この質問量はオラクル同定問題の質問量から導き出されたものであり，よって，その上界はオラクル同定問題の質問量と等しくなっている．$2^{N^d} < M < 2^{N/\log N}$ において，オラクル同定問題の質問量は $o(N)$ である．しかしながら，簡単のため，Theorem 13 において，$M > 2^{N^d}$ の場合を省略した．図 11 において，$M < 2^{N/\log N}$ のときの上界が $o(N)$ になることを図示している．

(3) ほぼすべての関数について，ほぼ厳密な上下界を示した．つまり，質問量の上下界はそれぞれ，$O(\log M)$，$\Omega(\log M/\log N)$ となることを示した．

以上より，真理値表の 1 の個数に基づいて関数計算の質問量をほぼ厳密に特徴付けることができたと言える．

上述の下界 $\Omega(\log M/\log N)$ は $O(\log M)$ に改良でき，その結果，上界に一致するのではないかと予想している．しかしながら，今のところ，文献 [2] の手法や，アドバーサリー法，多項式法を応用して所望の下界を得るには至っていないため，今後はこの下限の改良を行っていきたい．

参考文献

[1] A. Abdollahi and M. Pedram: “Analysis and synthesis of quantum circuits by using quantum decision diagrams”, Proc. of the Design, Automation and Test in Europe (DATE’06), pp. 317–322 (2006).

[2] A. Ambainis. A note on quantum black-box complexity of almost all boolean functions. *Ina. Process. Lett.*, 71(1):5–7, 1999.

[3] A. Ambainis, K. Iwama, A. Kawachi, R. Raymond, and S. Yamashita. Improved algorithms for quantum identification of boolean oracles. *Theor. Comput. Sci.*, 378(1):41–53, 2007.

[4] H. Barnum, C. Crepeau, D. Gottesman, A. Smith, and A. Tapp, “Authentication

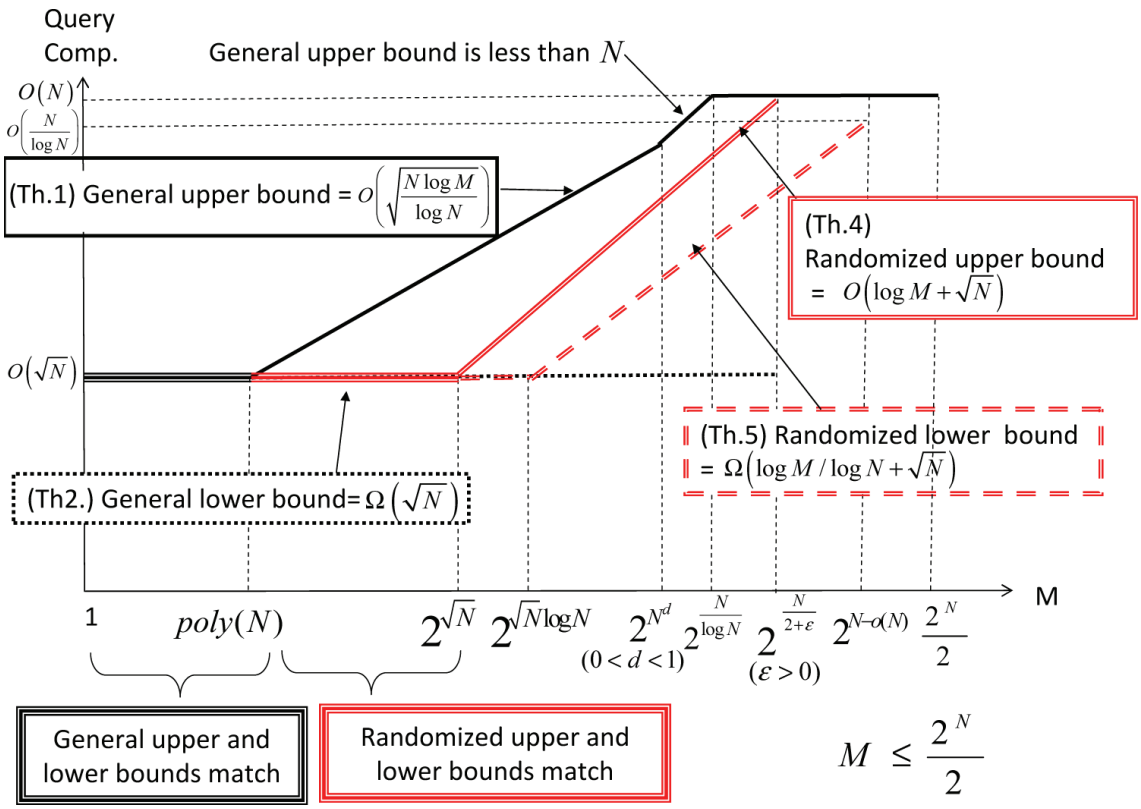


図 11: The number of 1's and quantum query complexity

of Quantum Messages," Proc. 43rd Annual IEEE Symposium on the Foundations of Computer Science, pp. 449-458, IEEE Press, 2002.

[5] R. Beals, H. Buhrman, R. Cleve, M. Mosca, and R. de Wolf. Quantum lower bounds by polynomials. In *Proc. 39th Annual IEEE Symposium on Foundations of Computer Science*, pages 352–361, 1998.

[6] R. E. Bryant: "Graph-based algorithm for Boolean function manipulation", IEEE Trans. Comput., **C-35**, 8, pp. 667–691 (1986).

[7] R. Cleve, D. Gottesman, and H. L. Lo, "How to Share a Quamtum Secret," Phys. Rev. Lett., 83, 648, 1999.

[8] C. Crepeau, D. Gottesman, and A. Smith, "Secure Multi-party Quantum Computation," Proc. 34th Annual ACM Symposium on Theory of computing, pp.643–652, 2002.

[9] C. Crepeau, D. Gottesman, and A. Smith, "Approximate Quantum Error-Correcting Codes and Secret Sharing Schemes," EUROCRYPT 2005, May, 2005.

[10] D. Gottesman, "Stabilizer Codes and Quantum Correction," quant-ph/9705052.

[11] L. K. Grover: "A fast quantum mechanical algorithm for database search", Proceedings of 28th ACM Symposium on Theory of Computing, pp. 212–219 (1996).

[12] G. P. Guo and G. C. Guo, "Quantum Secret Sharing without Entanglement," Phys. Lett. A, vol. 310/4, pp.247–251, 2003.

[13] G.-P. He, "Secure quantum string seal exists," quant-ph/0602159, 2006.

[14] M. Hillery, C. Buzek, and A. Berthiaume, "Quantum Secret Sharing," Phys. Rev. A., 59, 1829, 1999.

[15] A. Karlsson, M. Koashi, and N. Imoto, "Quantum Entanglement for Secret Sharing and Secret Splitting," Phys. Rev. A., 59, 162, 1999.

[16] K. Kurosawa, S. Obana, and W. Ogata, "t-Cheater Identifiable (k, n) Threshold Secret Sharing Schemes," Proc. of CRYPTO'95, LNCS, 963, pp.410–423, 1995.

[17] V.M. Khrapchenko. Complexity of the realization of a linear function in the case of Π-circuits. *Math. Notes Acad. Sciences*, 9:21–23, 1971.

[18] S. Laplante, T. Lee, and M. Szegedy. The quantum adversary method and classical formula size lower bounds. In *CCC '05: Proceedings of the 20th Annual IEEE Conference on Computational Complexity*, pages 76–90, Washington, DC, USA, 2005. IEEE Computer Society.

[19] H.-K. Lo and H. F. Chau, "Unconditional security of quantum key distribution over arbitrarily long distances," Science, vol. 283, 2050–2056, 1999.

[20] A. Shamir, "How to Share a Secret," Commun. ACM 22, 612, 1979.

[21] P. W. Shor, "Fault-tolerant quantum computation," Proc. 37th Symposium on Foundations of Computing, IEEE Computer Society Press, 1996, pp. 56–65, 1996.

[22] S. Tani and M. Nakanishi and S. Yamashita, A quantum protocol for the list-nonequality function. In *Proceedings of the Eleventh Quantum Information Technology Symposium (QIT11)*, QIT2004-51, pages 21–25, IEICE, 2004.

[23] P. Tiwari. Lower Bounds on Communication Complexity in Distributed Computer Networks. Journal of the ACM, 34(4):921–938, 1987.

[24] S. Aaronson and A. Ambainis, Quantum Search of Spatial Regions. In *Proceedings of the 44th Symp. on Foundations of Computer Science*, pp.200–209, 2003.

[25] C. H. Bennett, "Quantum cryptography using any two nonorthogonal states," Phys. Rev. Lett. 68, 3121 (1992).

[26] C. H. Bennett, and G. Brassard, "Quantum cryptography: public key distribution and coin tossing," In Proc. IEEE International Conf. Computers, Systems, and Signal Processing, pp. 175–179, (1984).

[27] K. Boström and T. Felbinger, "Deterministic Secure Direct Communication Using Entanglement," Phys. Rev. Lett. 89, 187902 (2002)

[28] A. K. Ekert, "Quantum cryptography based on Bell's theorem," Phys. Rev. Lett. 67, 661 (1991).

[29] L. Goldenberg and L. Vaidman, "Quantum cryptograph based on orthogonal states," Phys. Rev. Lett. 75, 1239 (1995).

[30] A. Kawachi, T. Koshiba, H. Nishimura, and T. Yamakami, "Computational indeistinguishability between quantum states and its cryptographic application," Proc. EuroCrypt 2005, LNCS 3494, pp. 268–284, (2005).

[31] T. Okamoto, K. Tanaka, and S. Uchiyama, "Quantum public-key cryptosystems," Proc. of CRYPTO 2000, LNCS 1880, pp.147–165, (2000).

[32] K. Iwama and R. Raymond and S. Yamashita, "General bounds for quantum biased oracles," IPSJ Journal, Vol. 46, No. 10, pp. 1234–1243, 2005.

[33] P. Høyer, M. Mosca, and R. de Wolf, "Quantum search on bounded-error inputs," ICALP, pp.291–299, 2003.

研究業績一覧

学術論文

1. M. Nakanishi:
"On the Power of One-Sided Error Quantum Pushdown Automata with Classical Stack Operations", Proc. of 10th Annual International Computing and Combinatorics Conference (COCOON 2004), LNCS 3106, pp. 179–187, 2004 年 8 月.
概要: One of important questions on quantum computing is whether there is a computational gap between the models that are allowed to use quantum effects and the models that are not. Several types of quantum computation models have been proposed, including quantum finite automata and quantum pushdown automata (with quantum pushdown stack). It has been shown that some quantum computation models are more powerful than classical counterparts and some are not since quantum computation models are required to obey some restrictions such as reversible state transitions. In this paper, we investigate the power of quantum pushdown automata whose stack is assumed to be implemented as a classical device, and show that they are strictly more powerful than classical counterparts in the one-sided error setting. That is, we show that there is a non-context-free language which quantum pushdown automata with classical stack operations can recognize with one-sided error.

2. S. Yamashita, M. Nakanishi, K. Watanabe:
"Toward a Practical Environment for Quantum Programming", Proc. of Asia-Pacific Conference on Quantum Information Science, pp. 133–141, 2004 年 12 月.
概要: This paper proposes a practical framework for quantum programming. In our framework, the parts of a program to be performed on a quantum computer are almost automatically determined, and the other parts are performed on a classical computer. We only consider Grover Search to be performed on a quantum computer in the framework because the other quantum algorithms known so far cannot be applied to general cases. By considering only Grover Search, we have several advantages that show our framework is really practical.

3. K. Iwama, A. Kawachi, R. Raymond H.P., S. Yamashita:
"Robust Quantum Algorithms for Oracle Identification", Workshop on Quantum Information Processing 2005, Poster Presentation.
概要: In this paper, we treat a general problem, called the *oracle identification problem (OIP)*, introduced by Ambainis et al., (STACS 04). An OIP is, given an oracle $f_i : \{0,1\}^n \to \{0,1\}$ and a candidate set $S = \{f_1, ..., f_M\}$, to determine which oracle in S is now in the black-box. Note that OIP is a generalization of several oracle-computation problems where we need to determine all the $N = 2^n$ values of the truth table for a given oracle f_i. For example, the Grover search has an S of size N and each f_i in S is defined as $f_i(j) = 1$ iff $i = j$. The STACS paper proves among others that the query complexity of OIP is $O(\sqrt{N})$ for *any* S if $|S| = N$, which includes the upper-bound for the Grover search as a special case. Our main results in the current paper include: (i) The above algorithm for OIP with any S of size N can be made robust with a sacrifice of only a constant factor, i.e., its query complexity is still $O(\sqrt{N})$. To do this, we need to bypass the classical (noise-free) search which plays an im-

portant role in the original algorithm. (ii) A similar extension is also possible for the case that $|S| > N$. (Note that $|S|$ can be as large as 2^N.) In this case, our new algorithm is not only robust but its query complexity is significantly better than the previous one.

4. S. Tani, M. Nakanishi, S. Yamashita: "Quantum Protocol for the List-nonequality Function", Workshop on Quantum Information Processing 2005, Poster Presentation.
概要: In this paper, we show that a quantum protocol can do much better than classical protocols for a special case of the direct product problem, where we focus on computing the OR of the values of l instances of $2k$ input bit function f, i.e.,$f^{\vee l} = \bigvee_{i=1}^{l} f(x^i, y^i)$. In particular, we show an efficient quantum protocol to compute the list-nonequality function $LNE(l, k)$: the negation of $\mathrm{EQ}_k^{\vee l}$ where EQ_k means the equality function of k bits. It should be noted that the recursive calls of amplitude amplification technique used in the protocol for DISJ_n by Høyer et al. gives us a quantum protocol with $O(c^{\log^* l}\sqrt{l}\log k)$ communication bits. However, this is not necessarily better than the classical bound, i.e., $\Theta(l + \log k)$. To get a better bound than the classical one, we develop a quantum version of the classical public-private coin conversion: if there is a bounded error quantum protocol that uses classical public coins, a bounded error quantum protocol that does not use classical coins exists. By carefully combining this conversion with a robust quantum search by Høyer et al., we can show a quantum protocol whose complexity is $O(\sqrt{l}\log l + \log k)$.

5. M. Nakanishi, K.Hamaguchi, T.Kashiwabara: "Expressive power of quantum pushdown automata with classical stack operations under the perfect-soundness conditions", IEICE Transactions on Information and Systems, Vol. E89-D, no. 3, pp. 1120–1127, 2005 年 3 月.
概要: One important question for quantum computing is whether a computational gap exists between models that are allowed to use quantum effects and models that are not. Several types of quantum computation models have been proposed, including quantum finite automata and quantum pushdown automata (with a quantum pushdown stack). It has been shown that some quantum computation models are more powerful than their classical counterparts and others are not since quantum computation models are required to obey such restrictions as reversible state transitions. In this paper, we investigate the power of quantum pushdown automata whose stacks are assumed to be implemented as classical devices, and show that they are strictly more powerful than their classical counterparts under the perfect-soundness condition, where perfect-soundness means that an automaton never accepts a word that is not in the language. That is, we show that our model can simulate any probabilistic pushdown automata and also show that there is a non-context-free language which quantum pushdown automata with classical stack operations can recognize with perfect soundness.

6. K. Iwama, A. Kawachi, S. Yamashita: "Quantum Sampling for Balanced Allocations", IEICE Transactions on Information and Systems, Vol. E89-D, no. 1, pp. 39–46, 2005 年 1 月.
概要: When we put N balls into N bins uniformally at random, we obtain the highest bin which contains $\Theta(\ln N/\ln\ln N)$ balls with high probability. This problem has a practical meaning such as load balanc-

ing. In this paper, we investigate to use a quantum sampling for this problem, based on Grover's quantum search algorithm. We show that our quantum algorithm creates a bin which has $\Omega(\sqrt{\ln N/\ln\ln N})$ balls with high probability. However, it is known that the highest bin contains only $O(\ln\ln N)$ if we choose two bins at random for each ball and put it into the one which is less full. This simple randomized algorithm significantly outperforms the quantum algorithm.

7. M. Adcock, R. Cleve, K. Iwama, R. Putra, S. Yamashita:
"Quantum Lower Bounds for the Goldreich-Levin Problem", Information Processing Letters, Vol.97(5), pp. 208–211, 2006 年 3 月.
概要: At the heart of the Goldreich-Levin Theorem is the problem of determining an n-bit string a by making queries to two oracles, referred to as IP (inner product) and EQ (equivalence). The IP oracle, on input x, returns a bit that is biased towards $a \cdot x$ (the modulo two inner product of a with x) in the following sense. For a random x, the probability that $IP(x) = a \cdot x$ is at least $\frac{1}{2}(1+\varepsilon)$. The EQ oracle, on input x, returns a bit specifying whether or not $x = a$. It has been shown that a quantum algorithm can solve this problem with $O(1/\varepsilon)$ IP and EQ queries, whereas any classical algorithm requires $\Omega(n/\varepsilon^2)$ such queries. Also, the quantum algorithm requires only $O(n/\varepsilon)$ auxiliary one- and two-qubit gates in addition to its queries. We show that the above quantum algorithm is optimal in terms of both EQ and IP queries. Specifically, $\Omega(1/\varepsilon)$ EQ queries are necessary, and $\Omega(1/\varepsilon)$ IP queries are necessary if the number of EQ queries is $o(\sqrt{2^n})$.

8. Y. Kawano, S. Yamashita, and M. Kitagawa:
"Explicit implementation of quantum circuits on a quantum-dot-cellular-automata-like architecture", Physical Review A, Vol. 72, No. 1, 012301, 2005 年 7 月.
概要: We present an efficient strategy to translate a normal quantum algorithm into a sequence of operations on the unidirectional periodic structure (UPS) originally proposed by S. Lloyd. The UPS assumes arrays of weakly coupled quantum systems where an interaction exists only between neighboring qubits, and can only perform the same quantum operation onto all the qubits. The sequence obtained by the strategy proposed by S. Lloyd needs at most $12n$ operations, where n is the number of qubits for the original circuit. The sequence obtained by our strategy needs at most $6n$ operations. We also clarified the relations between the upper bound of the number of translated operations and the period of the UPS, and between the upper bound of the number of qubits and the period of the UPS.

9. K. Iwama, R. Raymond, and S. Yamashita:
"General Bounds for Quantum Biased Oracles", Journal of IPSJ, Vol. 46, No. 10, pp. 2481–2491, 2005 年 10 月.
概要: An oracle with bias ϵ is an oracle that answers queries correctly with probability at least $1/2+\epsilon$. In this paper, we study the upper and lower bounds of quantum query complexity of oracles with bias ϵ. For general upper bounds, we show that for any quantum algorithm solving some problem with high probability using T queries of perfect quantum oracles, i.e., oracles with $\epsilon = 1/2$, there exists a quantum algorithm solving the same problem, also with high probability, using $O(T/\epsilon)$ queries of the corresponding biased quantum oracles. As corollaries we can show robust quantum algorithms and gaps between biased quantum and classical oracles, e.g.,

by showing a problem where the quantum query complexity is $O(N/\epsilon)$ but the classical query complexity is lower bounded by $\Omega(N \log N/\epsilon^2)$. For general lower bounds, we generalize Ambainis' quantum adversary argument to biased quantum oracles and obtain the first lower bounds with explicit bias factor. As one of its applications we can provide another proof of the optimality of quantum algorithms for the so-called quantum Goldreich-Levin problem which was proved before by Adcock et al. using different and more complicated methods.

10. K. Iwama, A. Kawachi and S. Yamashita: "Quantum Biased Oracles", Journal of IPSJ, Vol. 46, No. 10, pp. 2400–2408, 2005 年 10 月.
概要: This paper reviews researches on quantum oracle computations when oracles are not perfect, i.e., they may return wrong answers. We call such oracles *biased* oracles, and discuss the formal model of them. Then we provide an intuitive explanation how quantum search with biased oracles by Høyer et. al works. We also review the method, by Buhrman et. al, to obtain all the answers of a quantum biased oracle without any overhead compared to the perfect oracle case. Moreover, we discuss two special cases of quantum biased oracles and their interesting properties, which are not found in the classical corresponding cases. Our discussion implies that the model of quantum biased oracle adopted by the existing researches is natural.

11. Y. Murakami, M. Nakanishi, S. Yamashita and K. Watanabe: "Quantum versus Classical Pushdown Automata in Exact Computation", Journal of IPSJ, Vol. 46, No. 10, pp. 2471–2480, 2005 年 10 月.
概要: Even though quantum computation is useful for solving certain problems, classical computation is more powerful in some cases. Thus, it is significant to compare the abilities of quantum computation and its classical counterpart, based on such a simple computation model as automata. In this paper we focus on the quantum pushdown automata which were defined by Golovkins in 2000, who showed that the class of languages recognized by quantum pushdown automata properly contains the class of languages recognized by finite automata. However, no one knows the entire relationship between the recognitive abilities of quantum and classical pushdown automata. As a part, we show a proposition that quantum pushdown automata can deterministically solve a certain problem that cannot be solved by any deterministic pushdown automata.

12. M. Nakanishi: "Automata with quantum and classical resources", Journal of IPSJ, Vol. 46, No. 10, pp. 2384–2391, 2005 年 10 月.
概要: Quantum automata have been studied as simple quantum computation models. They can be considered models of small (or restricted) quantum computers. In this paper, we give descriptions of several kinds of quantum automata and show their power in comparison to their classical counterparts. We also give descriptions of quantum automata that have additional classical computational resources. Introducing classical computational resources can enhance the power of quantum automata, since this approach relaxes such restrictions as reversible state transitions.

13. S. Tani, M. Nakanishi, S. Yamashita: "Quantum Communication Complexity for the Distinctness Function on a Ring", Workshop on Theory of Quantum Computation, Communication, and Cryptography

2006, pp. 10–11, 2006年2月.

概要: This paper considers the problems of computing functions that depend on distributed inputs on a network whose underlying toplology is not necessarily a complete graph. We first establish a lower bound technique for the communication complexity of computing any Boolean function that depends on the inputs given to two parties on the network of any topology. Next we apply the technique to lower-bound the communication complexity of computing the Distinctness problem on an n-party ring, and also give two quantum algorithms for the problem to upper-bound the complexity. When the set of possible input is of size $O(n)$, our bound is tight.

14. A. Ambainis, K. Iwama, A. Kawachi, R. Raymond, S. Yamashita:
"Improved Algorithms for Quantum Identification of Boolean Oracles", Theoretical Computer Science, Vol. 378, Issue 1, pp. 41–53, 2007年6月.

概要: The oracle identification problem (OIP) was introduced by Ambainis *et al.* It is given as a set S of M oracles and a blackbox oracle f. Our task is to figure out which oracle in S is equal to the blackbox f by making queries to f. OIP includes several problems such as the Grover Search as special cases. In this paper, we improve the previous algorithms by providing a mostly optimal upper bound of query complexity for this problem: (i) For any oracle set S such that $|S| \leq 2^{N^d}$ $(d < 1)$, we design an algorithm whose query complexity is $O(\sqrt{N \log M / \log N})$, matching the known lower bound. (ii) Our algorithm also works for the range between 2^{N^d} and $2^{N/\log N}$ (where the bound becomes $O(N)$), but the gap between the upper and lower bounds worsens gradually. (iii) Our algorithm is robust, namely, it exhibits the same performance (up to a constant factor) against the noisy oracles as also shown in the literatures for special cases of OIP.

15. T. Suzuki, S. Yamashita, M. Nakanishi, K. Watanabe:
"Robust Quantum Algorithms Computing OR with ϵ-biased Oracles", IEICE transactions on Information and Systems, Vol. E90-D, No.2, pp. 395–402, 2007年2月.

概要: This paper considers the quantum query complexity of ε-*biased oracles* that return the correct value with probability only $1/2 + \varepsilon$. In particular, we show a quantum algorithm to compute N-bit OR functions with $O(\sqrt{N}/\varepsilon)$ queries to ε-biased oracles. This improves the known upper bound of $O(\sqrt{N}/\varepsilon^2)$ and matches the known lower bound. We also show a quantum algorithm to cope with the situation in which we have no knowledge about the value of ε. This contrasts with the corresponding classical situation, where it is almost hopeless to achieve more than a constant success probability without knowing the value of ε.

16. M. Hayashi, K. Iwama, H. Nishimura, R. Raymond, S. Yamashita:
"(4,1)-Quantum Random Access Coding Does Not Exist—One Qubit Is Not Enough To Recover One of Four Bits", New J. Phys., 8, 129, 2006年8月.

概要: An $(n, 1, p)-$quantum random access (QRA) coding, introduced by Ambainis et al., is the following communication system: the sender which has n-bit information encodes his/her information into one qubit, which is sent to the receiver. The receiver can recover any one bit of the original n bits correctly with probability at least p, through a certain decoding process based on positive operator-valued measures. Actually, Ambainis et al. shows the existence of a $(2, 1, 0.85)$-QRA coding and also proves the impossibility of its classical counterpart. Chuang immediately extends it to a

$(3, 1, 0.79)$-QRA coding and whether or not a $(4, 1, p)$-QRA coding such that $p > 1/2$ exists has been open since then. This paper gives a negative answer to this open question. Moreover, we generalize its negative answer for one-qubit encoding to the case of multiple-qubit encoding

17. M. Hayashi, K. Iwama, H. Nishimura, R. Raymond, S. Yamashita:
"(4,1)-Quantum Random Access Coding Does Not Exist", Proc. of the 2006 IEEE International Symposium on Information Theory (ISIT2006), pp. 446–450, 2006年7月.
概要: An $(n, 1, p)$-Quantum Random Access (QRA) coding, introduced by Ambainis, Nayak, Ta-shma and Vazirani in ACM Symp. on Theory of Computing 1999, is the following communication system: The sender which has n-bit information encodes his/her information into one qubit, which is sent to the receiver. The receiver can recover any one bit of the original n bits correctly with probability at least p, through a certain decoding process based on positive operator-valued measures. Actually, Ambainis et al. shows the existence of a $(2, 1, 0.85)$- QRA coding and also proves the impossibility of its classical counterpart. Chuang immediately extends it to a $(3, 1, 0.79)$- QRA coding and whether or not a $(4, 1, p)$-QRA coding such that $p > 1/2$ exists has been open since then. This paper gives a negative answer to this open question.

18. A. Ambainis, K. Iwama, A. Kawachi, R. Raymond and S. Yamashita:
"Improved Algorithms for Quantum Identification of Boolean Oracles", Prof. of the 10th Scandinavian Workshop on Algorithm Theory (SWAT 2006), pp. 280–291, 2006年7月.
概要: The oracle identification problem (OIP) was introduced by Ambainis et al. It is given as a set S of M oracles and a blackbox oracle f. Our task is to figure out which oracle in S is equal to the blackbox f by making queries to f. OIP includes several problems such as the Grover Search as special cases. In this paper, we improve the previous algorithms by providing a mostly optimal upper bound of query complexity for this problem: (i) For any oracle set S such that $|S| \leq 2^{N^d}$ $(d < 1)$, we design an algorithm whose query complexity is $O(\sqrt{N \log M / \log N})$, matching the known lower bound. (ii) Our algorithm also works for the range between 2^{N^d} and $2^{N/\log N}$ (where the bound becomes $O(N)$), but the gap between the upper and lower bounds worsens gradually. (iii) Our algorithm is robust, namely, it exhibits the same performance (up to a constant factor) against the noisy oracles as also shown in the literatures for special cases of OIP.

19. T. Suzuki, S. Yamashita, M. Nakanishi, K. Watanabe:
"Robust Quantum Algorithms with ϵ-biased Oracles", Proc. of 12th Annual International Computing and Combinatorics Conference (COCOON 2006), pp. 116–125, 2006年8月.
概要: This paper considers the quantum query complexity of *ε-biased oracles* that return the correct value with probability only $1/2 + \varepsilon$. In particular, we show a quantum algorithm to compute N-bit OR functions with $O(\sqrt{N}/\varepsilon)$ queries to ε-biased oracles. This improves the known upper bound of $O(\sqrt{N}/\varepsilon^2)$ and matches the known lower bound. We also show a quantum algorithm to cope with the situation in which we have no knowledge about the value of ε. This contrasts with the corresponding classical situation, where it is almost hopeless to achieve more than a constant success probability without knowing the value of ε.

20. H. Nishiyama, M. Nakanishi, S. Yamashita, Y. Nakashima:
"An Efficient Approximation of SU(d) Using Decomposition", Proc. of Asian Conference on Quantum Information Science 2006 (AQIS 2006), pp. 147–148, 2006 年 9 月.
概要: In this paper, we show the approximation method for n-dimensional unitary matrix using decomposition into the product of the two-level unitary matrices. The computation time for the preparation stage does not depend on d. We show that the length of circuit is asymptotically $O(d^2 \log^{3.97}(1/\varepsilon))$. We also show an exact comparison of the length of the circuit where we can see that our method can even provide smaller circuits than Dawson's method in some cases.

21. Y. Murakami, M. Nakanishi, M. Hagiwara, S. Yamashita, Y. Nakashima:
"Quantum Secure Direct Communication Protocols for Sending a Quantum State", 2006 International Symposium on Information Theory and its Applications (ISITA2006), 2006 年 10 月.
概要: We propose a new quantum secure direct communication protocol, which has an advantage of combining the following features over related works: The protocol allows us to send a quantum state in security; The proper sender and receiver can detect an eavesdropper efficiently; The protocol employs no entanglement resource, and thus it is relatively easy to be experimented in current technology.

22. K. Iwama, H. Nishimura, R. Raymond, S. Yamashita:
"Quantum Random Access Coding using Multiple Qubits", Workshop on Quantum Information Processing 2007, Poster Presentation.
概要: In 1999, Ambainis et al. first proposed the concept of *Quantum Random Access Coding (QRAC)* to show advantages of quantum information beyond the limitation of Holevo Bound. Concerning QRAC, Hayashi et al. show the following: There is no $(2^{2m}, m, p > 1/2)$–QRAC. In other words, for any $(n, m, p > 1/2)$–QRAC, it holds that $m > \frac{\log n}{2}$. However, it was not clear if $(2^{2m} - 1, m, > 1/2)$–QRACs do exist for any $m \geq 1$. This paper show that such QRACs actually exist hence showing the tightness of the above results by Hayashi et al.

23. K. Iwama, H. Nishimura, R. Raymond, S. Yamashita:
"Quantum Network Coding for General Graphs", Workshop on Quantum Information Processing 2007, Poster Presentation.
概要: Network coding is often explained by using a small network model called Butterfly. In this network, there are two flow paths, s_1 to t_1 and s_2 to t_2, which share a single bottleneck channel of capacity one. So, if we consider conventional flow (of liquid, for instance), then the total amount of flow must be at most one in total, say 1/2 for each path. However, if we consider information flow, then we can send two bits (one for each path) at the same time by exploiting two side links, which are of no use for the liquid-type flow, and encoding/decoding operations at each node. This is known as *network coding* and has been quite popular since its introduction by Ahlswede, Cai, Li and Yeung in 2000. In QIP 2006, Hayashi et al showed that *quantum* network coding is possible for Butterfly, namely we can send two *qubits* simultaneously with keeping their fidelity strictly greater than 1/2. In this paper, we show that the result can be extended to a large class of general graphs by using a completely different approach. The underlying technique is a new cloning method called *entanglement-free cloning* which does not

produce any entanglement at all. This seems interesting on its own and to show its possibility is an even more important purpose of this paper. Combining this new cloning with approximation of general quantum states by a small number of fixed ones, we can design a quantum network coding protocol which "simulates" its classical counterpart for the same graph.

24. Y. Murakami, M. Nakanishi, S. Yamashita, Y. Nakashima:
"A Quantum Secure Direct Communication Protocol for Sending a Quantum State and Its Security Analysis", Workshop on Quantum Information Processing 2007, Poster Presentation.
概要: We propose a new quantum secure direct communication (QSDC) protocol which allows us to (i) carry an arbitrary quantum state; (ii) detect eavesdroppers on a channel efficiently; and (iii) employ no entanglement resource. We also give its security analysis.

25. M. Nakanishi, S. Tani, S. Yamashita:
"Tight Bounds on Information Gains in Quantum Sealing Protocols", Workshop on Quantum Information Processing 2007, Poster Presentation.
概要: This paper investigates the ability of quantum string sealing protocols by giving an optimal quantum sealing protocol when the whole string is not necessarily recovered even for honest players. More specifically, in the followings, we show that the information leaks (i.e. information gains for a cheater) matches a trivial lower bound in our quantum sealing protocol. Thus we give a clear understanding of what we can do by quantum string sealing protocols.

26. M. Hayashi, K. Iwama, H. Nishimura, R. Raymond, S. Yamashita:
"Quantum Network Coding", Proc. of the 24th International Symposium on Theoretical Aspects of Computer Science (STACS 2007), pp. 610–621, 2007年2月.
概要: Since quantum information is continuous, its handling is sometimes surprisingly harder than the classical counterpart. A typical example is cloning; making a copy of digital information is straightforward but it is not possible exactly for quantum information. The question in this paper is whether or not *quantum* network coding is possible. Its classical counterpart is another good example to show that digital information flow can be done much more efficiently than conventional (say, liquid) flow. Our answer to the question is similar to the case of cloning, namely, it is shown that quantum network coding is possible if approximation is allowed, by using a simple network model called Butterfly. In this network, there are two flow paths, s_1 to t_1 and s_2 to t_2, which shares a single bottleneck channel of capacity one. In the classical case, we can send two bits simultaneously, one for each path, in spite of the bottleneck. Our results for quantum network coding include: (i) We can send any quantum state $|\psi_1\rangle$ from s_1 to t_1 and $|\psi_2\rangle$ from s_2 to t_2 simultaneously with a fidelity strictly greater than 1/2. (ii) If one of $|\psi_1\rangle$ and $|\psi_2\rangle$ is classical, then the fidelity can be improved to 2/3. (iii) Similar improvement is also possible if $|\psi_1\rangle$ and $|\psi_2\rangle$ are restricted to only a finite number of (previously known) states. (iv) Several impossibility results including the general upper bound of the fidelity are also given.

27. K. Iwama, H. Nishimura, R. Raymond, S. Yamashita:
"Unbounded-Error One-Way Classical and Quantum Communication Complexity", L. Arge et al. (Eds.): ICALP 2007, LNCS 4596, pp. 110–121, 2007年7月.
概要: This paper studies the gap between quantum one-way communication complex-

ity $Q(f)$ and its classical counterpart $C(f)$, under the *unbounded-error* setting, i.e., it is enough that the success probability is strictly greater than 1/2. It is proved that for *any* (total or partial) Boolean function f, $Q(f) = \lceil C(f)/2 \rceil$, i.e., the former is always exactly one half as large as the latter. The result has an application to obtaining (again an exact) bound for the existence of (m, n, p)-QRAC which is the n-qubit random access coding that can recover any one of m original bits with success probability $\geq p$. We can prove that $(m, n, > 1/2)$-QRAC exists if and only if $m \leq 2^{2n} - 1$. Previously, only the construction of QRAC using one qubit, the existence of $(O(n), n, > 1/2)$-RAC, and the non-existence of $(2^{2n}, n, > 1/2)$-QRAC were known.

28. S. Yamashita, M. Nakanishi:
"A Practical Framework to Utilize Quantum Search", Proc. of the 2007 IEEE Congress on Evolutionary Computation (CEC2007), (CD-ROM), 2007年9月.
概要: In this paper we propose a *practical* framework to utilize quantum computers in the future. To the best of our knowledge, this is the first paper to show a *concrete* usage of quantum computation in general programming. In our framework, we can utilize a quantum computer as a coprocessor to speed-up some parts of a program that runs on a classical computer. To do so, we propose several new ideas and techniques, such as a practical method to design a large quantum circuits for search problems and an efficient quantum comparator.

29. K. Iwama, H. Nishimura, R. Raymond, S. Yamashita:
"Unbounded-error classical and quantum communication complexity", Proc. of 18th International Symposium on Algorithms and Computation (ISAAC2007), Lecture Notes in Computer Science 4835, pp. 100–111, 2007年12月.
概要: Since the seminal work of Paturi and Simon, the unbounded-error classical communication complexity of a Boolean function has been studied based on the arrangement of points and hyperplanes. Recently, it was found that the unbounded-error *quantum* communication complexity in the *one-way communication* model can also be investigated using the arrangement, and showed that it is exactly (without a difference of even one qubit) half of the classical one-way communication complexity. In this paper, we extend the arrangement argument to the *two-way* and *simultaneous message passing* (SMP) models. As a result, we show similarly tight bounds of the unbounded-error two-way/one-way/SMP quantum/classical communication complexities, implying that all of them are equivalent up to a multiplicative constant of four. Moreover, the arrangement argument is used to show that the gap between *weakly* unbounded-error quantum and classical communication complexities is at most a factor of three.

30. M. Nakanishi, S. Tani, S. Yamashita:
"An information-theoretic security analysis of quantum string sealing", In Proc. of the 6th WSEAS International Conference on Information Security and Privacy (ISP'07), pp. 30–35, 2007年12月.
概要: A quantum sealing protocol is a quantum cryptographic protocol that can detect unsealing of a message. It is known that quantum *bit* sealing is insecure. However, this does not imply that quantum *string* sealing is also insecure. So the possibility of quantum string sealing has been studied. In this paper, we propose a quantum sealing protocol and give an information theoretic analysis. We also show that our protocol is almost optimal for uniformly distributed

inputs in the sense that the upper bound of information leaks almost matches the trivial lower bound.

31. Y. Murakami, M. Nakanishi, S. Yamashita, Y. Nakashima, M. Hagiwara:
"A quantum secure direct communication protocol for sending a quantum state and its security analysis", In Proc. of the 6th WSEAS International Conference on Information Security and Privacy (ISP'07), pp. 91–97, 2007 年 12 月.
概要: These days, quantum secret communication algorithms different from quantum key distribution protocol have been proposed, which are called quantum secure direct communication protocols. These do not aim to have a key agreement between two communicators, but, to send a secret message directly using quantum devices. In this paper, we propose a new quantum secret direct communication protoccol which has advantages over the current ones and also discuss its security analysis against the man-in-the-middle attack.

研究会等

1. 河内亮周, 山下茂, 岩間一雄:
"オラクル同定問題に対する頑健な量子アルゴリズム", 情報処理学会研究会報告 アルゴリズム研究会, Vol. 96-1, pp. 1–8, 2004 年 7 月.
概要: オラクル同定問題とは n ビット入力 1 ビット出力のオラクルの集合 $S = \{f_1, ..., f_M\}$ と S に属する未知のオラクル f_i が与えられたとき，未知のオラクル f_i を S の中から同定する問題であり，数多くの問題の一般化になっている．本稿では，与えられた未知のオラクルクエリーの結果が正しい値と誤った値が重ね合わせ状態として出力される場合のオラクル同定問題（但し，誤った計算結果が得られる確率は 1/2 より小さい定数確率，例えば 1/10 以下とする.）に対して以下を示す．(i) 与えられたオラクル集合のサイズが $N = 2^n$ の場合（すなわち $M = N$.），エラー付きオラクル同定問題をクエリー回数 $O(\sqrt{N})$ で少なくとも定数確率で解く量子アルゴリズムが構成できる．(ii) 与えられたオラクル集合のサイズが N の場合，各 $x \in \{0,1\}^n$ で $f_i(x) = 1$ を満たす i の数に関するパラメータ K に関してある条件を満たすとき，エラー付きオラクル同定問題をクエリー回数 $O(\sqrt{N/K})$ で少なくとも定数確率で解く量子アルゴリズムが構成できる．

2. 谷誠一郎, 中西正樹, 山下茂:
"リスト型非等価関数を計算する量子プロトコル", 電子情報通信学会 量子情報技術研究会, QIT2004-51, pp. 21–25, 2004 年 12 月.
概要: This paper considers the quantum communication complexity of the list-nonequality function $LNE(l, k)$, which is AND of the values of l instances of the $2k$-bit nonequality function, i.e., the negation of OR of the values of l instances of the $2k$-bit equality function. We give the quantum communication complexity of $LNE(l, k)$ is at most $O(\sqrt{l} \log l + \log k)$ in the bounded error setting, which is asymptotically better than the corresponding classical complexity, i.e., $\Theta(l + \log k)$.

3. 岩間一雄, 河内亮周, Rudy Raymond, 山下茂:
"オラクル同定問題に対するロバストな量子アルゴリズム", 電子情報通信学会 量子情報技術研究会, QIT2005-11, pp. 42–47, 2005 年 5 月.
概要: The oracle identification problem (OIP) was introduced by Ambainis et al., which is given as a set S of M oracles and a blackbox oracle f. Our task is to figure out which oracle in S is equal to the blackbox f by making queries to f. OIP includes several problems such as the Grover Search as special cases. In this paper, we design *robust* algorithms, i.e., those which are tolerant against noisy oracles, for OIP. Our results include: (i) For any oracle set S such

that $|S|$ is polynomial in N, $O(\sqrt{N})$ queries are enough to identify the blackbox oracle, which is obviously optimal since this OIP includes the Grover Search as a special case. (ii) For the case that $|S| \leq 2^{N^d}$ $(d < 1)$, we design an algorithm whose query complexity is $O(\sqrt{N \log M / \log N})$ and matches the known lower bound. (iii) We can furthermore design a robust algorithm whose complexity changes smoothly between the complexity of (ii) and the complexity of recovering all information about the blackbox oracle whose complexity is $O(N)$ as showed by Buhrman et al. Thus our new algorithms are not only robust but also better than the previous noiseless case in terms of query complexity.

4. Y. Murakami, M. Nakanishi, S. Yamashita, K.Watanabe:
"Cheater Identifiable Quantum Secret Sharing Schemes", 情報セキュリティ研究会, vol. 105, pp. 90–92, 2005 年 7 月.
概要: In this paper, we show that there exists a cheater identifiable (k, n) threshold secret sharing scheme for a quantum secret. Suppose that there are at most t cheaters $(k \geq 3t + 1)$, k or more participants can identify who are cheating. Our scheme utilizes authentication codes based on an orthogonal array and stabilizer codes, which are unconditionally secure. Moreover, we discuss the relation between n and t.

5. T. Katsumata, M. Nakanishi, S. Yamashita, K. Watanabe:
"Quantum sealing schemes against collective measurement attacks", 電子情報通信学会 量子情報技術研究会, QIT2005-90, pp. 229–232, 2005 年 11 月.
概要: A quantum seal is a way of encoding a classical message into quantum states, so that everyone can read the message with no or little error, while authorized verifiers can detect that someone has read the message and broken the seal. We present two schemes that are more impervious to collective measurement than the Bechmann-Pasquinucci scheme. Both schemes have an advantage that the longer a classical message is encoded, the more difficult a collective measurement perform in practice. Moreover one of them, *scheme 2*, may make it impossible to try collective measurement depending on an encoded classical message.

6. K. Iwama, H. Nishimura, R. Raymond, S. Yamashita:
"Transmitting classical information on the quantum network efficiently", 電子情報通信学会 コンピュテーション研究会, vol. 105, pp. 15–20, 2005 年 12 月.
概要: The question in this paper is whether the quantum random access (QRA) coding, given by Ambainis et al., on the *network* is possible. This question is motivated by the network coding introduced by Ahlswede et al., which enables us to send classical information much more efficiently on the network by coding at intermediate nodes. We demonstrate that quantum network coding is possible for the QRA coding, by using a simple network model called the Butterfly network. In this network, there are two flow paths, s_1 to t_1 and s_2 to t_2, which shares a single bottleneck channel of capacity one. In the classical case, we can send two bits simultaneously, one for each path, in spite of the bottleneck. However, the quantum case makes this bottleneck much more severe than in the classical case because of the non-orthogonality of quantum states. We resolve this bottleneck and design a protocol which can send two classical bits from s_1 to t_1 (similarly from s_2 to t_2) but only one of them should be recovered.

7. 林正人, 岩間一雄, 西村治道, Rudy Raymond,

山下茂:
“量子ネットワーク上での効率的な情報の伝送”, 冬のＬＡシンポジウム, 2006 年 1 月.
概要: Since quantum information is continuous, its handling is sometimes surprisingly harder than the classical counterpart. A typical example is cloning; making a copy of digital information is straightforward but it is not possible exactly for quantum information. The question in this paper is whether or not *quantum* network coding is possible. Its classical counterpart is another good example to show that digital information flow can be done much more efficiently than conventional (say, liquid) flow. Our answer to the question is similar to the case of cloning, namely, it is shown that quantum network coding is possible if approximation is allowed, by using a simple network model called Butterfly. In this network, there are two flow paths, s_1 to t_1 and s_2 to t_2, which shares a single bottleneck channel of capacity one. In the classical case, we can send two bits simultaneously, one for each path, in spite of the bottleneck. We show that We can send any quantum state $|\psi_1\rangle$ from s_1 to t_1 and $|\psi_2\rangle$ from s_2 to t_2 simultaneously with a fidelity strictly greater than 1/2.

8. 誤り率を考慮したバイアス付量子オラクル計算量:
“鈴木智哉, 山下茂, 中西正樹, 渡邉勝正”, 冬のＬＡシンポジウム, 2006 年 1 月.
概要: This paper considers the quantum query complexity of *ε-biased oracles* that return the correct value with probability only $1/2 + \varepsilon$. In particular, we show a quantum algorithm to compute N-bit OR functions with $O(\sqrt{N}/\varepsilon)$ queries to ε-biased oracles. This improves the known upper bound of $O(\sqrt{N}/\varepsilon^2)$ and matches the known lower bound; we answer the known conjecture affirmatively. We also show a quantum algorithm to cope with the situation in which we have no knowledge about the value of ε. This contrasts with the corresponding classical situation, where it is almost hopeless to achieve more than a constant success probability without knowing the value of ε.

9. Geometric Characterization of Quantum Oracle Identification:
“河内亮周, 山下茂”, 冬のＬＡシンポジウム, 2006 年 1 月.
概要: We geometrically characterize the query complexity of the oracle identification problem in this paper. By defining an inner product between two oracles, we construct a quantum algorithm, which generalizes Bernstein-Vazirani and Grover algorithms, for a certain class of the oracle identification problem characterized by the inner product. We also show the optimality of this quantum algorithm.

10. M. Hayashi, K. Iwama, H. Nishimura, R. Raymond, S. Yamashita:
“(4,1)-Quantum Random Access Coding Does Not Exist”, 電子情報通信学会コンピュテーション研究会, COMP2006-14, pp.33–38, 2006 年 5 月.
概要: An $(n, 1, p)$-Quantum Random Access (QRA) coding, introduced by Ambainis, Nayak, Ta-shma and Vazirani in ACM Symp. on Theory of Computing 1999, is the following communication system: The sender which has n-bit information encodes his/her information into one qubit, which is sent to the receiver. The receiver can recover any one bit of the original n bits correctly with probability at least p, through a certain decoding process based on positive operator-valued measures. Actually, Ambainis et al. shows the existence of a $(2, 1, 0.85)$- QRA coding and also proves the impossibility of its classical counterpart. Chuang immediately extends it to a $(3, 1, 0.79)$- QRA coding and whether or not a $(4, 1, p)$-QRA coding such that $p > 1/2$

exists has been open since then. This paper gives a negative answer to this open question.

11. 村上ユミコ, 中西正樹, 萩原学, 山下茂, 中島康彦:
“量子状態を送信可能な量子秘密通信プロトコル”, 電子情報通信学会 量子情報技術研究会, QIT2006-28, pp. 133–134, 2006 年 5 月.
概要: We propose a new quantum secure direct communication protocol, which has an advantage of combining the following features over related works: The protocol allows us to send a quantum state in security; The proper sender and receiver can detect an eavesdropper efficiently; The protocol employs no entanglement resource, and thus it is relatively easy to be experimented in current technology.

12. 西山寛之, 中西正樹, 山下茂, 中島康彦:
“3 準位量子ゲート設計のためのユニタリ行列近似”, 電子情報通信学会 量子情報技術研究会, QIT2006-41, pp. 180–183, 2006 年 5 月.
概要: 近年，量子計算において多準位系を用いる手法が提案されている．本研究では，3 準位系を用いた量子ゲートの設計のためのユニタリ行列の任意精度の近似手法を提案する．本稿ではまず従来の 2 準位系を用いた手法について概観し，その発展として 3 準位でのユニタリ行列の近似法を述べる．そのために，3 準位系の Bloch ベクトルから作られる回転変換を定義し，その回転変換でユニタリ行列を近似する．

13. 竹谷昌敏, 中西正樹, 山下茂, 中島康彦:
“量子封印を用いた量子認証”, 電子情報通信学会 量子情報技術研究会, QIT2006-92, pp. 209–212, 2006 年 11 月.
概要: 現在，インターネットを利用する際に暗号技術は欠かせないものである．しかし，量子計算機で因数分解を高速に解ける事がわかったため，現代の数体暗号に換わるあたらしい暗号が必要となった．その中，量子を通信に利用する事により，BB84 等のプロトコルが考案された．その中で量子通信で認証や情報の完全性の保証をその計算量で保証するのではなく，量子の性質を利用し無条件安全なプロトコルの考案が要求された．本稿では現状の技術で実現が容易な量子認証プロトコルを提案する．

14. 山下茂, 中西正樹, 内山智香子, 相原正樹:
“Pure Dephasing を考慮した量子オラクル計算モデル”, 電子情報通信学会 量子情報技術研究会, QIT2006-97, pp. 229–232, 2006 年 11 月.
概要: 本稿では，量子オラクルと呼ばれる量子のブラックボックスがデコヒーレンスのエラーを受けた状況を計算科学の視点から考察する．我々の最終的な目標は，計算量（ある計算をするのに必要な量子オラクルの呼び出し回数）とデコヒーレンスのエラーの度合いとの関係を解析することである．その最初のステップとして，オラクルの出力に pure dephasing のエラーが起こってもそのようなオラクルを複数用いて pure dephasing のエラーを軽減する手法を提案する．これは，古典計算においてエラーのあるオラクルに対処するための最も基本的な手法である多数決による成功確率の増幅に対応する．我々の知る限りでは，我々の問題設定はオラクル計算の枠組みでデコヒーレンスの影響を考察する最初のものである．

15. 中西正樹, 谷誠一郎, 山下茂:
“量子文字列封印における復号化率と検出率のトレードオフについて”, 電子情報通信学会 量子情報技術研究会, QIT2006-98, pp. 233–236, 2006 年 11 月.
概要: 量子封印とは，乱数と量子ビットを用いてメッセージを符号化することにより，第三者がそのメッセージを復号すると量子状態が壊れ，メッセージが読まれたことが確認できるというものである．1 ビットの値を封印する量子ビット封印に関しては，安全なプロトコルが存在しないことがすでに知られている．しかし，量子ビット封印が安全でないからと言って量子文字列封印も安全でないとは限らない．そこで，量子文字列封印の可能性について様々な議論がなされている．しかし，

今まで盗聴者が盗聴できる情報量と検出率の関係については知られていなかった．本稿では，盗聴される情報量の上界が自明な下界に一致するプロトコルを提案し，盗聴者が盗聴できる情報量と検出率のトレードオフの関係を明らかにする．

16. S. Yamashita, D. Michael Miller:
"Decision Diagram Data Structure to Represent Quantum Circuit", 電子情報通信学会技術研究報告, VLD2006-58, pp. 41–46, 2006 年 11 月.
概要: In this paper, we propose a systematic design method for quantum circuits that compute Boolean functions. Our synthesis strategy is based on the previous method, but it is much simpler, and also may be more practical since we only use three different types of gates whereas the previous method may use many types of gates including rotation gates with *exponential* (with respect to the number of inputs of the desired function) accuracy. To formulate our synthesis strategy, we introduce and formulate a new concept, *matrix function*, which can describe a quantum circuit's behaviors very well. Then, in order to represent matrix functions efficiently, we naturally introduce decision diagram structures, which we call *(DDMF) Decision Diagram for a Matrix Function*. These formulation may be interesting on their own.

17. 岩間一雄，西村治道，ルディーレイモンド，山下茂:
"非有界誤り一方向量子および古典通信計算量", 電子情報通信学会 量子情報技術研究会, pp. 106–111, 2007 年 5 月.
概要: This paper studies the gap between classical one-way communication complexity $C(f)$ and its quantum counterpart $Q(f)$, under the *unbounded-error* setting, i.e., it is enough that the success probability is strictly greater than 1/2. It is proved that for *any* (total or partial) Boolean function f, $Q(f) = \lceil C(f)/2 \rceil$, i.e., the former is exactly (without an error of even ± 1) one half as large as the latter. The result has an application to obtaining (again an exact) bound for the existence of (m, n, p)-QRAC which is the n-qubit random access coding that can recover any one of m original bits with success probability $\geq p$. We can prove that $(m, n, > 1/2)$-QRAC exists if and only if $m \leq 2^{2n} - 1$.

18. S. Tani, M. Nakanishi, S. Yamashita:
"An analysis of quantum communication complexity depending on network topologies", 電子情報通信学会 量子情報技術研究会, pp. 150–153, 2007 年 11 月.
概要: This paper describes a general quantum lower bound technique for the communication complexity of a function that depends on the inputs given to two parties on an n-party network of any topology. The technique can also be employed to obtain a lower-bound of the communication complexity of some functions depending on the inputs distributed over all parties on the network. As an example, we apply our technique to the Distinctness problem on an n-party ring; almost matching upper bounds are also given when the input size is not too large.

19. Y. Murakami, M. Nakanishi, M. Hagiwara, S. Yamashita, Y. Nakashima:
"A quantum secure direct communication protocol for sending a quantum state and its security analysis", The 2008 Symposium on Cryptography and Information Secrity, 2D1-2, 2008 年 1 月.
概要: These days, quantum secret communication algorithms different from quantum key distribution protocol have been proposed, which are called quantum secure direct communication protocols. These do not aim to have a key agreement between two communicators, but, to send a secret

message directly using quantum devices. In this paper, we propose a new quantum secret direct communication protoccol which has advantages over the current ones and also discuss its security analysis against the man-in-the-middle attack.

学会大会等

1. Y. Murakami, M.Nakanishi, S.Yamashita, K.Watanabe:
   "No preshared key quantum secret communication protocol", 電子情報通信学会２００６年総合大会講演論文集, DS-1-12, 2006 年 3 月.

2. 本間知教, 中西正樹, 山下茂, 中島康彦:
   "耐故障性量子計算におけるエラー訂正回数の削減手法", 電子情報通信学会２００８年総合大会講演論文集, DS-1-4, 2008 年 3 月.

文部科学省科学研究費補助金 特定領域研究 (課題番号 16092219)

A08: 動的な構造をもつネットワーク上の資源割当て問題の研究

計算機ネットワークや物流ネットワークなど、我々が日頃かかわっているネットワークの多くは構造が動的に変化するという共通した特徴をもっている。ネットワーク構造の動的な変化は、それをベースに設計されている各種最適化問題の性質に少なからぬ影響を与える。本研究課題の目的は、動的な構造をもつネットワーク上で定式化される資源割当て問題の数理的な構造を明らかにするとともに、その根本的な解決のための道筋を示すことである。この目的を達成するため本研究では、1) 動的ネットワーク上の資源割当て問題のモデル化と特徴づけ、2) 実時間性を考慮した新たな評価基準の考案とその妥当性の検証、3) 考案した評価基準のもとで最適なアルゴリズムの提案とその有効性の評価、という 3 つの具体的な目標を設定し、研究を進めてきた。以下では、その具体的成果について詳述する。

研究組織

研究代表者： 藤田 聡　広島大学 大学院 工学研究科
研究分担者： 中野 浩嗣　広島大学 大学院 工学研究科 (平成 17–19 年度)
森本 康彦　広島大学 大学院 工学研究科 (平成 17–19 年度)

交付決定額 (配分額)

平成 16 年度	1,700,000 円
平成 17 年度	3,500,000 円
平成 18 年度	3,500,000 円
平成 19 年度	2,000,000 円
合　計	10,700,000 円

研究成果の概要

- 学術誌
  Journal of Interconnection Networks (2007 年), 日本データベース学会 Letters (2005 年), IEICE Trans. Information and Systems (2005 年), IEICE Trans. Information and Systems (2006 年), IEICE Trans. Information and Systems (2007 年), 情報処理学会論文誌 (2005 年) など 12 件
- 国際会議
  ISAAC (香港, 2004 年), PDCAT (シンガポール, 2004 年), PDCAT (大連, 2005 年), ICPADS (台北, 2005 年), ISPA (南京, 2005 年), I-SPAN (ラスベガス, 2005 年), ISAAC (海南, 2005 年), PDCAT (台北, 2006 年) など 34 件
- 招待講演
  回路とシステム軽井沢ワークショップ (2006 年, 2007 年), ネットワーク生態学シンポジウム (2008 年), Workshop on Algorithms and Computation (2008 年)

1　はじめに

本研究課題の目的は、動的な構造をもつネットワーク上の資源割当て問題の数理的な構造を明らかにし、その根本的な解決のための道筋を示すことである。補助金申請時に設定した本研究の柱は以下の 3 つである：

1) 動的ネットワーク上の資源割当て問題のモデル化と特徴づけ

2) 実時間性を考慮した新たな評価基準の考案とその妥当性の検証

3) 考案した評価基準のもとで最適なアルゴリズムの提案とその有効性の評価

上記の目的を達成するため、本研究課題では以下のような実施体制で 4 年間の研究を行った：まず研究代表者の藤田が、研究の総括と動的なネットワーク上の資源割当て問題に関する研究を担当した。研究分担者の中野教授と森本准教授には 2 年目の平成 17 年度から本研究課題に加わっていただき、中野教授にはアドホックネットワーク上の資源割当て問題に関する研究を、森本准教授にはセンサーネットワーク上の資源割当て問題に関する研究をそれぞれご担当いただいた。それぞれ専門の異なる 2 名の研究者と研究をすることで、研究の計画時に想定していた以上の独創的な研究成果を得ることができた。

本報告書では、本研究課題で得られた研究成果について順次説明していくが、その概要は、以下のようにまとめられる：

a) 支配集合をベースにした動的ネットワーク上の資源割当て問題のモデル化

ネットワーク上の資源割当てをグラフ上の支配集合 (dominating set) によってモデル化し、動的なネットワーク上であらわれるいくつかの問題に対して論考を加えた。まずサーバの位置が動的に変化する新しい計算モデルを考案し、そのモデル上で、ユーザへのサービスを停止することなくサーバを安全に移動させることが可能となるためのサーバ数の上下限を明らかにした (2 節)。次に複数のサービスを平行して提供する状況を支配集合分割 (domatic partition) によってモデル化した上で「欠陥のある支配集合分割」という新しい概念を提案し、その基本的な性質の解明を行った (3 節)。動的ネットワークではサーバの故障やネットワークからの離脱などのため支配集合分割を厳密に維持することは困難であることから、目標とする状況からの近さを明示的に導入することには大きな意義がある。またいくつかの具体的な状況では、欠陥を許さない通常の問題設定に比べて計算複雑度が下がることも理論的に確かめられた。

b) 現実のネットワークにおける実時間性を考慮した新しい尺度の検討

実時間性を考慮した新たな評価基準を探求するための手がかりとしてネットワーク上の輻輳制御問題とアドミッションコントロール問題に注目し、既存手法の性能を実験的に評価するとともに、それらの問題点について詳細な検討を行った。加えて、ネットワーク資源の公平な参加ノードが動的に変化する無線通信ネットワーク上で通信品質 (QoS) を保証するための各端末への通信機会 (TXOP) 割当て問題を定式化し、この問題に対する解法の提案を行った。次に、無線ネットワークにおける重要な課題のひとつである消費電力の軽減を実現するため、以下の研究を行った：b1) 空間内にばらまかれたセンサーノードに対してその正確な位置情報を自律分散的に推定させる問題を考え、具体的なアルゴリズムを提案するとともに、その性能を実機を用いて実証した (4 節)。b2) 無線ネットワーク上の通信帯域や電力などの効率的な資源配分方法について検討を行った。具体的には、ランキング問題やリーダ選挙問題について詳細に検討し、最適なアルゴリズムの提案を行った。b3) センサーネットワークの省電力ルーティングを実現するための具体的な手法のひとつとしてグラフの連結支配集合分割問題について検討した。具体的には、キュービックグラフに関する上下限と、交差グラフの一種であるパスグラフに対するアルゴリズムの提案を行った。

c) P2P オーバーレイネットワークの数理的構造に関する研究

分散ネットワーク上のアプリケーションでは、ルータを接続して得られる物理的なネットワークの上に、ホストやプロセスの間を論理的なリンクで結んで得られるオーバーレイネットワークを構成することで処理の効率化が図られている。構造

が動的に変化するP2P(ピア・ツー・ピア)ネットワーク上での共有資源の効率的な資源管理を目的として、以下の研究を行った:1) スキップグラフに基づく新しいオーバーレイネットワーク構造の提案とその理論的解析(5節)。2) DHT型P2P上で効率よく検索を行うための新しい手法の提案。一般にオーバーレイネットワーク上の任意の2点間の通信は、ルーティングテーブルにしたがった貪欲ルーティングによって効率よく実現できることが知られている。しかし、多くのアプリケーションでは複数経路を用いた情報のやり取りが行われており、そのような状況では、貪欲ルーティングを単純に並列に実行したのでは、十分な性能を引き出すことができない。また動的ネットワークでは、耐故障性の高さもルーティングアルゴリズムの満たすべき重要な要件のひとつである。この問題に対して本研究では、冗長化により使用されるネットワーク資源量の増加を抑える新しい手法を開発し、その効果をトランスポジショングラフ上のルーティング問題で確かめた(6節)。

以下では、上記の各項目について、背景と研究の位置づけを含めて、本研究課題で得られた結果の概要を述べる。そのほかの結果の詳細については、業績リスト中の論文を参照されたい。

2 移動サーバ割り当て問題

2.1 背景

近年、セルラーフォンシステムや無線LANなど様々なネットワーク環境下で、高品質のサービスを安定して提供することが求められている。このようなサービスの多くでは、ノードの物理的な位置やオペレーティングシステムなどに関して透過(transparent)であることが求められており、同一のサービスが“いつでもどこでも”受けられることが強く要請されている。このことは、与えられたコンピュータネットワーク上に特定のサービスを提供するサーバをどのように配置すればよいかという問題を提起する。具体的には、ネットワーク上のどのノード上でサーバを立ち上げれば、コンテンツ配送のための遅延や、配送のためのパスのバンド幅、あるいは各サーバからサービスを受けるクライアント数の均等化などの意味で「良い」サーバの配置が実現されるかが重要な問題となる。またそのようなメトリックの多くは、物理的な距離の近さやバンド幅制約等を満足するノード間を論理的なリンクで結んで得られるオーバレイネットワークを考え、その上で支配集合(dominating set)となるようなノード集合上にサーバを配置することによって自然に実現することができる。

頂点集合Vと辺集合Eをもつネットワーク$G = (V, E)$が与えられたとき、Gの**支配集合**Uとは、任意の頂点$u \in V$に対して、$u \in U$であるか$\{u, v\} \in E$かつ$v \in U$であるような頂点vが存在するようなVの部分集合として定義される。グラフの支配集合はこれまでに多くの研究者によって精力的に研究されており、グラフ理論的な特徴づけや、最小サイズの支配集合を求めるための計算複雑度に関する研究、さらには区間グラフやパーフェクトグラフなど特別なクラスのグラフに対する多項式時間アルゴリズムの提案など、これまでに様々な結果が得られている。また支配集合の概念がネットワーク上のサーバ配置問題と関連していることも古くから指摘されており、最近では、アドホックネットワークのための効率的なルーティングに(連結)支配集合をバックボーンとして利用しようという提案もなされている。

2.2 提案モデル

本研究では、支配集合の考え方に基づいた新しいサービス提供方式のモデルとして、あるホスト上に置かれたサーバが、サービス提供中にほかのホスト上に移動できるようなモデルを考える。そのようなサーバ機能の物理的な移動は、移動エージェントやプロセスマイグレーションの技術を用いることで実現することができるが、その際に問題となるのは、サーバの移動によってユーザに対するサービスが途切れないようにしなくてはならないという点である。すなわち、移動中を含む任意の時刻において、サーバの置かれた頂点集合が与えられたネットワークの支配集合となっていることが求められる(技術的な理由から、ここではリンク上を移動しているサーバはその両端点を同時に支配しているものとし、支配集合は多重集合であってもよいものと仮定する)。

以下では、2つの支配集合U_1, U_2 $(\subset V)$が以下

の条件を満たしているとき、支配集合間の関係の意味で隣接していると呼ぶことにする：

> 支配集合 U_1 中のある頂点 u を取り除き、$\{u, v\} \in E$ であるような頂点 v を加えることで、支配集合 U_2 が得られるとき。

また、支配集合間の隣接関係を推移的に適用して得られる関係を支配集合間の到達可能性と呼び、支配集合の集合 $\mathcal{D}$ ($\subset 2^V$) に属している任意の 2 つの支配集合が支配集合のみを経由して互いに到達可能であるとき、集合 $\mathcal{D}$ は相互変換可能であると呼ぶ (ただし経由する支配集合は、$\mathcal{D}$ の要素である必要はない)。

本研究では、どのような条件を満たせば支配集合の集合が相互変換可能になるのかについて詳細に検討を行った。特に、ここでは対象となるグラフのクラスを木とハミルトングラフに限定し、支配集合の集合が相互変換可能になるためには支配集合のサイズがどのくらい大きくなくてはならないか、あるいはどのくらい小さくても大丈夫なのかについてあきらかにした。具体的には、以下の定理が成立する：

定理 1 : n 個の頂点をもつ任意の木 T に対して、$k \geq \lfloor n/2 \rfloor$ 個の頂点からなるような T の支配集合たちは、相互変換可能である。また、任意の n に対して、最小支配集合のサイズが $\lfloor n/2 \rfloor$ であるような木 T が存在する。

定理 2 : 任意の $r \geq 2$ と $n \geq 1$ に対して、サイズ $\lceil (n+1)/3 \rceil - 1$ の支配集合全体からなる集合 $\mathcal{D}$ が相互変換可能ではないような n 頂点のハミルトン r-正則グラフ G が存在する。

定理 3 : n 頂点をもつ任意のハミルトングラフ G について、$k \geq \lceil (n+1)/3 \rceil$ 個の頂点からなるような G の支配集合たちは、相互変換可能である。

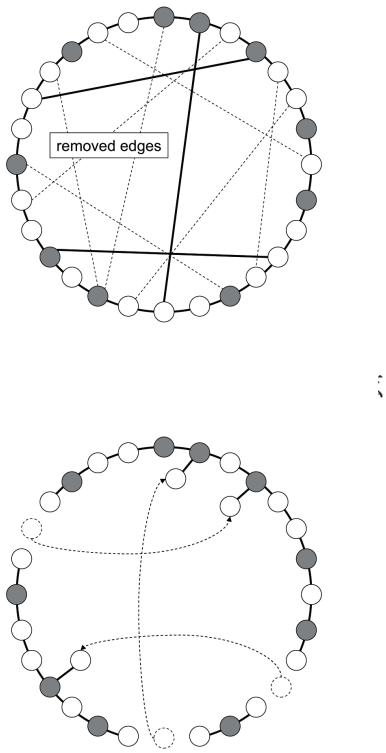
2.3 証明のスケッチ

以下では、それぞれの定理の証明のアイデアを簡単に説明する。

(定理 1 の証明のスケッチ) 葉 u を含むような支配集合が、u の代わりにその親頂点 v を含むような支配集合と (支配集合間の関係の意味で) 隣接している、という事実を利用する。頂点 v を支配集合に含めることで、少なくとも u, v の 2 頂点は (v によって) 支配されたことになるから、これらを木から取り除いて残りの部分に対して相互変換可能性を議論しても一般性を失わない。いま与えられた木に対する支配集合のサイズを $k \geq \lfloor n/2 \rfloor$ 個と仮定しているから、同様の議論を繰り返すことで、木に対する自明な支配集合に帰着させることができる。このことにより、定理の前半が証明できたことになる。後半部分については、中心となる根頂点がそれぞれの葉と長さ 2 のパスで結ばれているような木を考えればよい。葉の数は $\lfloor n/2 \rfloor$ 個あり、そのそれぞれが異なる頂点によって支配される必要があるため、このような木の最小支配集合のサイズは $\lfloor n/2 \rfloor$ となる。

(定理 2 の証明のスケッチ) $r = 1$ の場合は、サイズ $3m$ のリング (ただし m は整数) を考えればよい。サイズ $3m$ のリング上のサイズ $\lceil (n+1)/3 \rceil - 1 = m$ の支配集合は合計で 3 つ存在するが、それらの支配集合間を、支配集合のみを経由してお互いに行き来することはできない (サイズ m の支配集合に含まれる頂点のうちどれかひとつを隣接頂点に移動させると、支配集合ではなくなってしまう)。これは、サイズ m の支配集合に含まれる任意の頂点がそれぞれ 2 つの頂点を独占的に支配していることに起因する。定理 2 の証明では、この性質を利用したグラフの構成を行う。すなわち、サイズ m の支配集合に含まれる各頂点が、それぞれ 2 つの頂点を独占的に支配するような頂点数 $3m$ のグラフを具体的に構成することで、定理を証明することができる。

(定理 3 の証明のスケッチ) 証明は 2 つの部分から構成される。証明の前半部分では、支配集合のサイズが $\lceil (n+1)/3 \rceil$ 以上のとき、与えられたハミルトングラフ G とそのグラフに対する任意の支配集合 U から、G 上のハミルトン閉路 C の支配集合 U' に必ず到達させることができることを示す。後半部分では、サイズ n のリング上のサイズ $\lceil (n+1)/3 \rceil$ 以上の任意の 2 つの支配集合が、支配集合のみを経由してお互いに到達可能であることを示す。これらの 2 つの部分を組み合わせることにより、G 上のサイズ $\lceil (n+1)/3 \rceil$ 以上の任意の支配集合が、G 上のハミルトン閉路 C の支配集合を経由してお互いに到達可能であることを示すことができる。



(b) 前処理によって得られた森

図 1: 定理 3 のアイデア.

証明の前半部分で使うキーテクニックは、具体的な相互変換を行う前に、U が支配集合であるような G の部分グラフを、辺を順次取り除くことであらかじめ求めておくことである。図1に前処理のアイデアの一部を図示してある。そのようにして得られた G の部分グラフは森となることが示せるから、木における支配集合に対する適切な操作を施すことで、(多少の場合分けは必要ではあるが、)C に含まれない辺を少なくとも 1 本以上取り除くことができるような支配集合への変換を実現することができる。

一方、後半部分では、リングサイズ n の 3 に関する剰余による場合分けを行うことで見通しよく議論を行うことができる。具体的には、まず最初に 3 による剰余が 2 である場合を考えると、これはサイズ $3m+2$ のリングがサイズ $m+1$ の支配集合によって支配されている状況に対応し、そのような任意の支配集合は、2 つの頂点によって支配されている唯一の頂点 u によって明確に特徴づけられることがわかる。u を支配している 2 つの頂点のうちどちらか一方を u から遠ざかる方向に移動させることで (隣接する) 新たな支配集合を得ることができ、そのような操作を繰り返し適用することで、任意の支配集合間の到達可能性を示すことができる (いまリングのサイズが 3 で割り切れない状況について考えていることに注意)。リングサイズ n の 3 による剰余が 1 である場合は、2 つの頂点によって支配されている頂点 u とその隣接頂点の合計 3 つの頂点をひとつの仮想的な頂点とみなし、剰余が 2 である場合と同様の議論を適用することができる。また剰余が 0 である場合は、さらにもう一段仮想的な頂点を考えればよい。

3 欠陥のある支配集合分割

3.1 諸定義

無向グラフ $G=(V,E)$ を考える。各頂点 $u \in V$ について、u の閉じた近傍を $N[u] \stackrel{\text{def}}{=} \{u\} \cup \{v \in V : \{u,v\} \in E\}$ で定義する。V の部分集合 U は、任意の $u \in V$ について $N[u] \cap U \neq \emptyset$ であるとき G の支配集合であるといい、V の分割 $\mathcal{V} = \{V_1, V_2, \ldots, V_d\}$ は、各 V_i が G の支配集合であるとき G の**支配集合分割** (domatic partition) であるという (ここで $\mathcal{V}$ 中の各要素は、空集合ではないとする)。以下では、グラフ G のサイズ d の支配集合分割を d-DP と呼ぶ。ここで $\mathcal{V}$ が d-DP であるとき、任意の $u \in V$ について $|\{V_i \in \mathcal{V} : N[u] \cap V_i \neq \emptyset\}| = d$ となることに注意しよう。

本節では、**欠陥のある支配集合分割** (defective domatic partition: DDP) という概念を新たに定義し、その基本的な性質を調べる。グラフ G の頂点集合 V の分割 $\mathcal{V}$ は、1) $|\mathcal{V}| = d$ であり、かつ、2) 任意の $u \in V$ について $|\{V_i \in \mathcal{V} : N[u] \cap V_i \neq \emptyset\}| \geq d-k$ のとき、(d,k)-DDP であるという。ここで、欠けている高々k 個の支配集合の種類が頂点ごとに異なっていてもよいことにも注意しておこう。以下では (d,k)-DDP を、頂点あたり高々k 個の欠陥のある支配集合分割と呼ぶ。定義よりあきらかに、グラフ G が (d,k)-DDP をもつとき、任意の $i \geq 1$ について、G が $(d,k+i)$-DDP をもつことに注意されたい。

DDP は、センサーネット上の省電力ルーティングなどのように、支配集合分割によってモデル

化されるアプリケーションに対して自然に適用することができる。たとえば $(d,1)$-DDP は、各頂点が (d 個用意された支配集合のうち) 少なくとも $d-1$ 個の部分集合によって支配されていることをあらわしており、時分割でルータノードを切り替えるタイプのルーティング方式のもとでは、そのような分割を用いることで、$(d-1)/d$ の割合でルーティングサービスが提供されることを保証できることになる。

まず以下の補題が成立する：

補題 1： グラフ G が (d,k)-DDP をもつとき、任意の $1 \le i \le |V|-d$ について、G は $(d+i,k+i)$-DDP をもつ。

したがって、グラフ G の最大支配分割数が $d(G)$ のとき、任意の $1 \le i \le |V|-d(G)$ について、G は $(d(G)+i,i)$-DDP をもつことになる。

グラフ G の最大頂点次数を $\Delta(G)$、最小頂点次数を $\delta(G)$ であらわすこととしよう。頂点集合 V を $|V|$ 個の集合に分割すると、各頂点は少なくとも $\delta(G)+1$ 個の異なる頂点集合によって支配されるから、以下のことが言える：

性質 1： 任意のグラフ G は $(|V|,|V|-(\delta(G)+1))$-DDP をもつ。

また欠陥のない最大の支配集合分割数は $\delta(G)+1$ を超えることはないから、分割数 d が $\delta(G)+2$ 以上のとき、欠陥の数は少なくとも $d-(\delta(G)+1)$ 以上になる。したがって以下のことも直ちに言える：

性質 2： 任意のグラフ G は、任意の $i \ge 0$ に対して、$(\delta(G)+2+i,i)$-DDP をもたない。

考察の一つの方向は、上記の性質 1 を緩和していくことである。すなわち、すべての頂点に異なるラベルが与えられるという極端な状況から開始し、同じラベルを与えられた頂点がお互いに干渉しないように保ったままで、ラベルの総数をどこまで減少させていけるかを考える (3.2節参照)。もう一つの方向は、k の値を 0 から徐々に増やしていくことで、与えることのできるラベル数をどこまで増やしていけるかを考えていくことである (3.3節参照)。

3.2 ラベル数の上限

グラフ $G=(V,E)$ の平方 G^2 を以下のように定義する：$V(G^2)=V$ かつ、$E(G^2)=\{\{u,v\} \in V \times V : \{u,v\} \in E \text{ or } \exists w \text{ such that } \{u,w\},\{w,v\} \in E\}$. このとき、グラフ G^2 の独立点集合に含まれる任意の 2 頂点は、G において同一頂点の近傍に同時に含まれることはない。したがってこの性質を用いると、性質 1 におけるラベル総数を、d と k のギャップを保ったままで、$|V|$ から $\chi(G^2)$ に減少させることができる (ここで $\chi(G)$ は G の点彩色数)。Brooks の定理より、完全グラフではない任意のグラフ G の点彩色数は G の最大次数以下であることが知られているから、以下の性質が得られる (そのような点彩色は多項式時間で求めることができる)：

補題 2： G^2 が完全グラフではない頂点数 Δ^2+1 以上の任意のグラフ G は $(\Delta^2,\Delta^2-(\delta+1))$-DDP をもつ。ここで Δ は G の最大次数、δ は G の最小次数である。

したがってたとえば、頂点数 10 以上の任意のキュービックグラフが $(9,5)$-DDP をもつことを、上記の性質から直ちに導くことができる。

3.3 ラベル数の下限

本節では、正則グラフに対してラベル数の下限を調べる。任意のリングの最大支配集合分割数は 2 以上 3 以下であり、リングが 3-DP を持つための必要十分条件はリング中の頂点数が 3 の倍数になっていることである。以下の補題が成立する：

補題 3： リングが $(4,1)$-DDP を持つのは、リング中の頂点数が 5 でないときかつそのときに限られる。

この結果から、任意のリングが $(5,2)$-DDP をもつことが直ちにいえる (サイズ 5 の場合は性質 1 から、それ以外の場合は補題 1 と補題 3 から)。したがって性質 2 とあわせると、リングに関しては、すべての d,k の組合せに関して (d,k)-DDP の存在可能性と存在非可能性が示せたことになる。

次にキュービックグラフについて考える。任意のキュービックグラフ G の最大支配集合分割数 $d(G)$ は、2 以上 4 以下である。また「キュービッ

クグラフ G が 3-DP、すなわち $(3,0)$-DDP をもつかどうか」を問う問題の計算複雑さが NP 困難であることもすでにわかっている (証明には 3 次元マッチングからの帰着を用いる)。

以下の結果が成立する：

- 任意のキュービックグラフ G は、$(4,1)$-DDP をもつ。
- $(5,1)$-DDP をもたないような、ブリッジをもたないキュービックグラフが存在する。

性質 2 から、任意のキュービックグラフは $(5,0)$-DDP や $(6,1)$-DDP をもたない。また上記の結果と性質 1 から、ブリッジをもたない任意のキュービックグラフは、$(5,2)$-DDP や $(6,3)$-DDP をもつ。したがって、次に考えなくてはならないのは、以下の各項目である：

- キュービックグラフが $(5,1)$-DDP をもつかどうかを問う問題の計算複雑さについて
- キュービックグラフが $(6,2)$-DDP, $(7,3)$-DDP, $(8,4)$-DDP をもつかどうかを問う問題の計算複雑さについて

ちなみにキュービックグラフが (欠陥を許さない)4-DP をもつかどうかを問う問題の NP-困難性は、3-edge-coloring との等価性から示すことができる (つまり $(4,0)$-DDP をもつかどうかを調べる問題は NP 困難)。さきの議論から、任意のキュービックグラフが $(9,5)$-DDP をもつことはわかっているから、ギャップを 4 に保ったままでラベル数を 4 から 9 まであげていくどこかの段階で、計算複雑さが NP 困難から P に変化していることが予想される (ギャップが 3 の場合は、$k=0$ のときのみ NP 困難で、$k \geq 1$ のときはすべて P であることがわかっており、ギャップが 5 の場合は、そのような DDP は存在しない)。

上記の検討課題についてすべて解決したわけではないが、次の結果が得られている：

- $d \leq 7$ のとき、$(d, d-4)$-DDP が存在しないようなインスタンスが存在する (図 2を参照のこと)。
- 与えられたキュービックグラフが $(8,4)$-DDP をもつかどうかを問う問題は多項式時間で解くことができる。

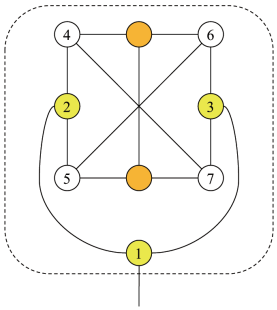


図 2: $(7,3)$ を持たないインスタンスの一部.

$(7,3)$-DDP が存在しなければ $(6,2)$-DDP や $(5,1)$-DDP ももちろん存在しない。また $d \geq 4$ のとき、$(d, d-4)$-DDP が存在するようなインスタンスが存在することから (自分と周囲のラベルがすべて異なるような、いわゆる “perfect domination” を考えればよい)、少なくとも $4 \leq d \leq 7$ の範囲においては、与えられたインスタンスが $(d, d-4)$-DDP をもつかどうかという問に対する答えは、YES であるようなインスタンスも NO であるようなインスタンスも存在するという意味で自明ではないことになる ($d=9$ のときは任意のキュービックグラフに対して $(9,5)$-DDP が存在するし、$d=4$ のときの計算複雑さが NP 困難であることもすでにわかっている)。

4 センサーノードへの位置情報配布問題

4.1 背景

ネットワーク上の資源配置のためのもうひとつの具体的な問題として、ユークリッド平面上に大量にばらまかれたセンサーノード (以下ノードと呼ぶ) たちに対して、各ノードの位置を無線通信機能を用いて正確に推定させる問題について研究を行った。この問題は一般に位置測位 (localization) 問題と呼ばれ、ユビキタスコンピューティングなどの分野で取り扱われる基本的な問題のひとつである。以下では、システム中には自身の正確な位置を知っている特別なデバイス (以下アンカーと呼ぶ) が複数存在し、ノードの位置測位は、アンカーたちとの通信を繰り返すことで実現されるも

のと仮定する。

本研究で対象とするのは、無線信号の送受信と、送受信装置間の距離変化による受信信号の強度変化の検出とができる程度の単純な構造をもったノードたちからなるセンサーネットワークであり、GPS (Global Positioning System) などで用いられている位相情報等に基づく距離推定機能は、チップサイズや消費電力などの制約から、ここでは仮定しない。提案する位置測位方式は、He らによって提案された PIT(Point-In-Triangle) テストに基づいている。PIT テストは、各アンカーとの距離を直接推定することなく各ノードの位置を測定することのできるいわゆるレンジフリー型の位置測位方式のための基本方式のひとつであり、信号を観測しているノードが 3 つのアンカーによって構成される三角形の内部に含まれているか否かを、以下の性質を利用して判定する：

> 観測点 M がアンカーによって構成される三角形 ABC の内部に存在するのは、M の近傍に、各アンカーから送られてくる信号の強度が同時に減少 (もしくは同時に増加) するような点が存在するときかつそのときに限られる。

PIT テストに基づく従来の位置測位方式 (APIT アルゴリズム) では、ノードの位置を微小変化させて信号強度の変化を見るのではなく、各ノードの周囲にすでに配置されているノードたちそれぞれの固定された位置における観測によって、ノード位置の変化を代替させるというアプローチがとられていた。しかしこのような方式では、観測位置の微小変化をほぼ同時刻に実現するためにはアンカーからの信号送信時刻を同期させるなどの付加機構が必要となり困難を伴うことから、送信時刻にばらつきが出てしまうことが予想され、そのことによって、観測結果が信号強度の時間変化の影響を直に受けてしまうという欠点がある。

4.2 提案モデル

提案手法では、観測位置の微小変化をノード側で実現するのではなく、アンカーの側で実現させる方法を新たに採用した。すなわち、複数のアンカーを少しずつ位置をずらして配置し、その位置関係と方角とを固定したまま空間内を移動できるようにして構成されたアンカーグループ (AG) という概念を導入し、PIT テストを行うための複数の三角形を、AG に空間内を時間をかけて移動させることによって実現することにした (もちろん同様の機能をもった AG を複数用意すれば、ノードの位置測位に要する時間は AG 数が増えた分だけ減少する可能性がある)。このことにより、測定位置の微小変化を AG に含まれるアンカーたちの同時送信によってほぼ同時刻にシミュレートすることが可能となり、PIT テストの精度を大幅に向上させることができる。この提案手法は、今年度、クラス 2 規格の Bluetooth デバイスを用いて実装され、いくつかの実証実験の結果、各観測点の位置測位が従来手法に比べて高精度に行えることが確かめられた。

またこの実装システムの成功をうけて、現在このモデル上で、以下のような理論的な問題についても取り組んだ：

- 位置測位に求められる分解能と移動時間や信号送信回数、AG の数などのコストが制約として与えられたとき、AG の移動スケジュールをどのようにすればよいか。
- 送信ポイントの数が与えられたとき、分解能をできるだけ向上させるような AG のスケジュールはどのようなものか。

スケジューラによって空間内に設定される送信ポイント間を線分で結ぶことによって、与えられた空間をいくつかの部分領域に分割することができるが、送信ポイント数を n としたとき、そのようにして得られる部分領域数の上限は $O(n^6)$ となることが知られている。この事実を利用することで、与えられたポイント数 n に対する分解能の下限を算出することができるが、最悪分解能という尺度で評価した場合、実際に得られる領域数は、それよりもはるかに少なくなる。具体的には、空間全体に関する最悪分解能は、送信ポイント集合から得られる凸包 (convex hull) 上の最長辺の長さよりも下回ることはない (その辺に内側から張り付いている部分領域の直径は辺の長さとほぼ等しくなる)。したがって今後は、AG の移動可能な領域と位置測位の対象となるノードの置かれてい

る領域とを別々に分けて議論するなどの、さらに踏み込んだ議論が必要になると考えられる。

5 検索コストの最小化をめざしたP2P オーバーレイ

5.1 背景

情報管理を集中的に行うサーバの存在を仮定しない「ピュア型 P2P システム」のためのオーバーレイネットワークは、その構成法によって、構造型と非構造型とに大別される。非構造型オーバーレイを用いた P2P システムの具体例として、Gnutella や Freenet などが知られている。非構造型オーバーレイのノード次数は一般に定数であり、生存時間 (TTL) で制御されるクエリのフラッディングによって、求めるべきコンテンツの検索が行われる。一方の構造型オーバーレイの具体例としては、分散ハッシュテーブル (Distributed Hash Table: DHT) が広く知られており、これまでに様々な座標空間に基づいた DHT 型 P2P システムが提案されている。DHT に基づく構造型オーバーレイ上でのインデックス検索には、非構造型オーバーレイの場合とは異なり、もし目的とするインデックスがシステム上に保存されていればそれを必ず取得することができるという大きなメリットがある。しかし一般に DHT 型オーバーレイでは、ノード次数と平均ルーティング時間との間にトレードオフの関係があり、したがってたとえば多項対数 (polylog) 時間でのクエリのルーティングを実現するためには、ノード次数を $\log\log N$ や $\log N$ などの “N に依存した値” に設定する必要がある。

本研究では、各ノードの平均次数が定数で、しかも多項対数時間でのクエリのルーティングを可能とするような、新しい構造型オーバーレイネットワークの提案を行った。提案するオーバーレイネットワークは、Aspnes と Shah によって提案されたスキップグラフの拡張となっている。スキップグラフは、順序関係のあるデータ集合に対する効率的なキー検索を実現するために提案された分散データ構造の一種であり、基本となる線形リストに対して平均次数が $O(\log N)$ となるようにショートカットリンクを付加して得られる構造をもっている。以下では、スキップグラフで用いられていたノードあたり $O(\log N)$ 本のショートカットリンクのうち、定数本のみをランダムに選択することで、$O(\log^2 N)$ ステップの平均ルーティング時間が実現できることを証明する。定理の証明には、Kleinberg が提案したグリッドグラフのためのショートカット手法と同様のアイデアが用いられる。実際、本稿で提案するオーバーレイネットワークは、スキップグラフのショートカット選択に対して Kleinberg の手法を近似的に応用したものとなっている。

5.2 スキップグラフ

スキップグラフ (以下、SG) は、順序関係のあるデータ集合に対する効率的な検索を実現するために提案された分散データ構造の一つである。N 個のノードからなるノード集合 V を考える。V 中の各ノード u には、固有の識別子のほかに、整数値 $key(u)$ がキーとして割り当てられているものとする。加えて各ノードには、ランダムかつ独立に選択された無限長のビット列が**メンバーシップベクトル**として割り当てられる。以下では、ノード u に割り当てられたメンバーシップベクトルを $m(u)$ と記す。

ある与えられたキー k を保持しているノードを V の中から探し出す問題を考える。この問題を効率よく解くために SG で用いられている基本アイデアの一つは、集合 V を、そこで保持されているキーの大きさに関して整列 (sorting) することである。整列の結果は、隣り合った要素をキーとしてもつ 2 ノード間を双方向リンク (無向辺) で結んで得られる線形リストの形で保存される。以下では、このリストのことを**基本リスト**と呼び、記号 C で表現する。キーの探索効率を高めるために SG で採用されているアイデアの二つ目は、基本リストに対して効果的なショートカットを付加することである。あるリスト L の**部分リスト** L' とは、L 中の要素のいくつか (もしくは全部) を、L 中で出てくる順番に、双方向リンクで結んで得られるリストのことである。また基本リスト C の**レベル i の部分リスト**とは、C の部分リストの中で、メンバーシップベクトルの「長さ i のプレフィックス」が同一であるようなノードたちのみを結ん

で得られるもののことをいう (レベル 0 の部分リストが C そのものであることに注意しよう)。ただしここでは部分リストとして、長さが 2 以上のリストのみを考えることにする[1]。無限の長さをもつビット列の長さ i の異なるプレフィックスの個数は 2^i であるから、レベル i の部分リストが、任意の C に対して高々2^i 個存在しうることに注意しよう。また各ノードは任意の i について高々一つのレベル i の部分リストにしか含まれず、しかもここでは各部分リストの長さが少なくとも 2 以上であると仮定しているので、レベル i の部分リスト数の上限は、値 $\lfloor N/2 \rfloor$ によっても与えることができる。

いま、リスト C のレベル i の (長さ 2 以上の) 部分リスト全体で用いられている辺の集合を E_i としよう。ここで各辺にはその辺の所属している部分リストのレベルがラベルとして付与されており、同一のノード対を結んでいても、ラベルが異なるものは別の辺と見なすことにする。このときスキップグラフ G は、

$$G \stackrel{\text{def}}{=} \langle V, \bigcup_{i=0}^{\infty} E_i \rangle \tag{5.1}$$

のように定義される。文献では、N 個のノードをもつ SG の平均次数と直径の期待値がともに $O(\log N)$ であることが示されている。

5.3 提案手法

提案手法では、各ノードは集合 D $(= \{1, 2, \ldots, \log N\})$ から δ 個の要素を (重複なく) 等確率で選択し、選択された要素のそれぞれをショートカット生成のためのプレフィックス長として用いる (δ は生成されるオーバーレイの次数になる)。具体的には、選択された要素が ℓ のとき、そのノードは「長さ ℓ のプレフィックス」が同一となるような基本リスト上の次のノード v を特定し、そのノードに向けてショートカットリンクを張ることになる。

$k = \log N$ をメンバーシップベクトルの長さとし、$\delta = 1$ とする。以下では簡単のために、メッセージのルーティング方向は一方向とし、終点ノード t は基本リスト上の端点に位置するとする。したがって、各ノードは選択したプレフィックスによってその方向のみのショートカットを生成する。また、終点ノードとの距離によってノード集合を以下のように分ける: 1) 集合 B_0 を t からの距離が 1 となる 2^0 $(= 1)$ 個のノード集合とし、2) 各 $1 \le j \le \log N - 1$ について、集合 B_j を t からの距離が $2^{j-1} + 1$ から 2^j までの 2^{j-1} 個のノード集合とする (基本リスト上で存在しうる最長の距離が $2^{\lceil \log N \rceil} - 1$ であることに注意しよう)。

ノード u が ℓ $(\in D)$ をプレフィックス長として選択したとしよう。他のノード v が u と同じ長さ ℓ のプレフィックスを持っている確率は $1/2^\ell$ であるから、ノード u から集合 $\{t\} \cup \bigcup_{i=0}^{j} B_i$ 内のあるノードにショートカットが張られている確率は、

$$P_\ell = \left(1 - \frac{1}{2^\ell}\right)^{x-1} \left\{1 - \left(1 - \frac{1}{2^\ell}\right)^{2^j}\right\}$$

となる。ここで、x は u から B_j までの距離である。$y = 2^j$ のとき $1 \le x \le 2^j$ であるから、一般性を失うことなく、$x - 1 = 2^{j-1} = y/2$ と仮定することができる。このとき、

$$\begin{aligned} P_\ell &= \left(1 - \frac{1}{2^\ell}\right)^{y/2} \left\{1 - \left(1 - \frac{1}{2^\ell}\right)^{y}\right\} \\ &= \left(1 - \frac{1}{2^\ell}\right)^{y/2} - \left(1 - \frac{1}{2^\ell}\right)^{y+y/2} \\ &> \left(1 - \frac{1}{2^\ell}\right)^{y/2} - \left(1 - \frac{1}{2^{\ell-1}}\right)^{y/2} \end{aligned}$$

となるが、いま、各プレフィックスは確率 $1/k$ で選択されるため、u から $\{t\} \cup \bigcup_{i=0}^{j} B_i$ 内のあるノードにショートカットが張られている確率 P は、

$$\begin{aligned} P &> \left(\frac{1}{k}\right)(P_k + P_{k-1} + \cdots + P_1) \\ &> \left(\frac{1}{k}\right)\left\{\left(1 - \frac{1}{2^k}\right)^{2^{j-1}} - \left(\frac{1}{2}\right)^{2^{j-1}}\right\} \\ &\simeq \left(\frac{1}{k}\right)\left(\frac{1}{e}\right)^{1/2^{k-j+1}} \end{aligned}$$

となり、不等式の右辺は、ある定数 α を用いて α/k のように表現することができる。実際、

- $j = \log N$ のとき $P > 0.6065/\log N$、
- $j = \log N - 1$ のとき $P > 0.7788/\log N$、

[1] 長さ 1 のリストはショートカットとしての意味をなさないため。

- $j \leq \log N - 2$ のとき $P > 0.8825/\log N$

となる。したがって、部分集合 B_{j+1} 内でのルーティング時間の期待値は、

- $j = \log N$ のとき $1.65 \log N$、
- $j = \log N - 1$ のとき $1.29 \log N$、
- $j \leq \log N - 2$ のとき $1.14 \log N$

となる。このような部分集合は高々$\log N$ 個であるため、ルーティング時間の期待値は高々$1.65 \log^2 N$ となる。

δ の値を増やした場合は、上記の解析の中で各ノードが行っていたベルヌーイ試行の数が 1 から δ に増加するため、走査するノード数の期待値は、単純に $1/\delta$ に減ることになる。したがって、ルーティング時間の期待値は、高々$(1.65/\delta) \log^2 N$ となる。

5.4 拡張

前節では、各ノードの次数を定数に抑えたままで多項対数時間でのルーティングを実現するためのスキップグラフの拡張方法を述べた。以下ではスキップグラフのもうひとつの拡張として、より多くの次数を用いることで、二重対数時間でのルーティングを実現する方法を示す。ここで、各ノードの次数が $n-1$ のとき、定数時間のルーティングが自明に達成されることに注意しておこう。したがって以下では、次数をできるだけ抑えたままで二重対数時間のルーティングを行うことを具体的な目標とする。

基本アイデアは、$\lg \lg N$ 段のレベルをもつ再帰的な構造を用いることである (このような構造は、たとえば PRAM 上の中央値選択アルゴリズムなどで用いられている)。再帰の基本ステップは、与えられたサイズ M の集合をサイズ $\sqrt{M}$ の $\sqrt{M}$ 個の部分集合に分割し、再帰のより深いレベルに受け渡すことである。与えられた集合のサイズが N のとき、上記のステップを繰り返し適用することで、各レベルの部分集合のサイズは

$$N^{1/2}, N^{1/4}, N^{1/8}, N^{1/16}, \ldots$$

のように縮小していくことになる。したがって、部分集合のサイズが定数になるレベルの深さは $i = \lg \lg N$ となる。

メンバーシップの考え方を用いることで、そのような再帰的な分割を、以下のように行うことができる：1) 各ノードにそれぞれ長さ $\lg N$ のメンバーシップベクトルを与える。2) i 番目のレベルの部分集合を、メンバーシップベクトルの長さ $\sum_{j=1}^{i} (\lg N)/2^j$ のプレフィックスが同一のノードたちからなる集合として定義する。ここで $\sum_{j=1}^{i} (\lg N)/2^j = (\lg N)\left(\frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2^i}\right) = \lg N - \left(\frac{1}{2^i}\right) \lg N$ であることに注意しておこう。もし同一のレベルに属するすべてのノードを互いに (完全結合で) 結ぶことができれば、同一のレベルに属するノード間の通信をそれぞれ定数ステップで実現することができるから、任意の 2 ノード間のメッセージ転送を平均 $O(\log \log n)$ ステップで実現できるようなネットワークが構成されることになる (結合がランダムに与えられたメンバーシップベクトルに基づいて行われることに注意しよう)。したがって、平均通信時間をスキップグラフの $O(\log n)$ から $O(\log \log n)$ に大幅に減少することができる。ただしこのような直接的な方法では、根に最も近いレベルの部分集合のサイズが $\sqrt{N}$ であることから、ノードあたりの次数は、

$$\sum_{i=1}^{\lg \lg N} N^{1/2^i} = \sqrt{N} + O(N^{1/4})$$

にまでしか抑えることができない。

しかし上述のノード次数は、ステップ数をわずかに増加させることで、次のような方法を用いて減少させることが可能である。基本アイデアは、最も高いレベルのショートカットリンクを (同様の手法を最適的に用いることで) $\lg \lg N - 1$ ステップでシミュレートすることである。再帰の最も高いレベルのリンクが、同一部分集合に属する $N^{1/2}$ 個のノードを完全結合していたことに注意しよう。前述の手法が N 個のノード間のルーティングを平均 $O(\log \log n)$ ステップで実現していたことに注意すると、同様の手法を最も高いレベルに適用することで、ノードあたりの平均次数を

$$2 \sum_{i=2}^{\lg \lg N} N^{1/2^i} = 2N^{1/4} + O(N^{1/8}).$$

にまで減少させることができる。さらにこのような再帰的な手法を k 回適用することで、平均ルーティング時間を $2^k \lg\lg N$ に抑えたままで、ノードの次数を

$$2^k N^{1/2^k} + O(N^{1/2^{k+1}})$$

に減少させることができる。したがって、たとえば k の値を $k = \lg\lg\lg N$ のように設定すれば、ノード次数が $(\lg\lg N) \times N^{1/\lg\lg N}$ であるようなネットワーク上で平均 $(\lg\lg N)^2$ ステップのルーティングを実現することが可能となる。

6 トランスポジショングラフ上の点素なパスの構成法

本節では、Cayley グラフの一種であるトランスポジショングラフにおいて、与えられたソース頂点と目的頂点集合とを点素なパス集合によって結ぶためのアルゴリズムを示す。このアルゴリズムの実行時間はトランスポジショングラフの次元に関して多項式時間であり、得られるパスの長さは、$n \geq 7$ のとき、各目的頂点までの最短パスの長さに比べて高々14 しか長くならない。この結果はこれまでに知られていた最もよいアルゴリズムの大幅な改良になっている (従来手法では最短パスの長さからの増分は $3n-5$ であり、定数では抑えられない)。

6.1 準備

V_n を記号 $1, 2, \ldots, n$ 上の $n!$ 個の順列の集合とする。順列 $u \in V_n$ の第 i 桁の値を u_i と記す。ジェネレータ $g_{i,j}$ $(1 \leq i < j \leq n)$ とは、与えられた順列の i 桁と j 桁を交換する関数のことであり、与えられた順列が $u = u_1 u_2 \ldots u_n$ $(\in V_n)$ のとき、

$$g_{i,j}(u) \stackrel{\text{def}}{=} u_1 \ldots u_{i-1} \underline{u_j} u_{i+1} \ldots u_{j-1} \underline{u_i} u_{j+1} \ldots u_n$$

のように定義される。

定義 1 **(n-TG)** n 個のシンボル上のトランスポジショングラフ T_n とは、V_n を頂点集合とし、以下のように定義される E_n を辺集合とするような無向グラフである：$E_n \stackrel{\text{def}}{=} \{\{u, g_{i,j}(u)\} \mid u \in V_n,\ 1 \leq i < j \leq n\}$.

T_n 中の頂点 u の隣接頂点の集合を $N(u) \stackrel{\text{def}}{=} \{v \in V_n \mid \{u, v\} \in E_n\}$ とする。任意の $u \in V_n$ に対して $|N(u)| = n(n-1)/2$ であることに注意しよう。実際、T_4 においては、頂点 1234 は、次の 6 $(= 4 \times 3/2)$ 個の隣接頂点をもつ：$1243, 1324, 1432, 2134, 3241, 4231$。以下では表記を簡単にするため、$f_n \stackrel{\text{def}}{=} n(n-1)/2$ とする。また n-TG の対称性より、$1234...n$ をソース頂点とし、s と D との距離は 2 以上であると仮定する。さらに n-TG が次数 f_n $(= n(n-1)/2)$ の正則グラフであることから、$|D| = f_n$ を仮定する。

6.2 構成法

垂直クラスタリング

提案手法ではまず最初に、T_n を長さ 2 の suffix によって特徴づけられるような T_{n-2} のコピーたちに分割する。suffix ij によって特徴づけられるコピーを $TG(i,j)$ と記し、そのような $n(n-1)$ 個のコピー全体の集合を $\mathcal{T}$ とする (例えばソース $s = 123\ldots n$ は $TG(n-1, n)$ に属しており、s の f_n 個の隣接頂点のうち f_{n-2} 個は、s と同じコピーに属している)。以下では、$\mathcal{T}$ を V_n の**垂直クラスタリング**と呼び、その各要素を**垂直クラスタ**と呼ぶ。定義より、異なる垂直クラスタが交わりをもたないことに注意しよう。

提案手法の第 1 のアイデアは、D 中の各頂点を $\mathcal{T}$ 中の垂直クラスタに単射によって割り当てることである ($|\mathcal{T}| = 2|D|$ より、$\mathcal{T}$ がそのような単射を実現するのに十分な垂直クラスタを含んでいることに注意しよう)。以下では $i \neq n-1$ かつ $j \neq n$ であるような垂直クラスタ $TG(i,j) \in \mathcal{T}$ をソース頂点 s に関して**安全**なクラスタと呼ぶ。図 3は、T_4 における安全な垂直クラスタを示している。任意の安全な垂直クラスタ $TG(i,j)$、$TG(i',j')$ について、s と $TG(i',j')$ 中の任意の頂点を結ぶ最短パスは $TG(i,j)$ と頂点を共有しない。なぜならば、そのようなパスは $(n-1)$ 桁の値が $n-1$ であるか n 桁の値が n であるかによって特徴づけられるからである (つまりそのようなパスのすべての頂点は安全でない垂直クラスタに含まれている)。容易にわかるように、$\mathcal{T}$ は $2n-3$ $(= n-1+n-2)$ 個の安全でない垂直クラスタを含んでおり、その

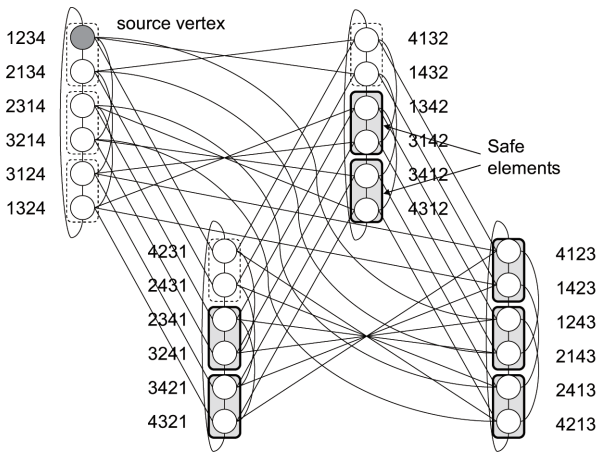


図 3: $\mathcal{T}$ 中の垂直クラスタ (ソース頂点は濃いグレーで、安全な垂直クラスタは明るいグレーで示されている).

数は $n \geq 2$ のとき $|\mathcal{T}|$ の半数を超えない。

パスの基本設計

提案手法で構成される s と D の各頂点を結ぶパスは、以下の 3 つのパートから成る：

Part 1: 最初のパートは、s と安全な垂直クラスタとを長さ 2 以下の最短パスで結ぶ (パスの長さが 2 のとき、その中間頂点は安全でない垂直クラスタに属している)。

Part 2: 2 番目のパートは、1 番目のパートの端点と 3 番目のパートの端点とを安全な垂直クラスタ内部で結ぶ。

Part 3: 3 番目のパートは D 中の頂点と安全な垂直クラスタの頂点とを結ぶ。

目的頂点の安全な垂直クラスタへの割り当てを厳密に述べるために、以下のような記法を導入する。M を以下のように定義される $n \times n$ 行列とする:

$$M[i,j] \stackrel{\text{def}}{=} |D \cap TG(i,j)| \text{ for each } i,j.$$

行列 M が長さ 2 の suffix に関する目的頂点の初期分布をあらわしていることに注意されたい。$|D| = f_n$ および $D \subset \bigcup_{i,j} TG(i,j)$ より、$\sum_{i,j} M[i,j] = f_n$ が成立する。与えられた行列 M に対し提案手法は、D から $\mathcal{T}$ への単射をあらわす 0,1 行列 M^* を構成する。$d(i,j) \in D$ を行列 M^* のもとで $TG(i,j)$ に対応付けられた頂点とする (ただし $i = j$ の場合や $TG(i,j)$ に何も割り当てられていない場合には、$d(i,j)$ を NULL とする。M^* が $f_n + n$ 個の 0 と f_n 個の 1 を含むことに注意しよう)。提案手法では、行列 M^* は初期分布 M から以下の 3 つの要件を満たすように構成される：

1) $TG(i,j)$ が安全で $M[i,j] = 1$ であるようなペア (i,j) が存在すれば、$D \cap TG(i,j)$ 中の (唯一の) 頂点が、M^* のもとで $TG(i,j)$ に対応する目的頂点として選ばれること。

2) 安全でない任意の垂直クラスタについて $M^*[i,j] = 0$ であること。

3) s と D を結ぶ以下のような点素なパスの集合が存在すること：任意の $(i',j') \neq (i,j)$ について、s と $d(i,j)$ を $TG(i,j)$ を経由して結ぶパスは、s と $d(i',j')$ を $TG(i',j')$ を経由して結ぶパスと点素であること。

水平クラスタリング

目的地までの点素なパスによる配分 (行列 M^* を実現するような点素なパス) は、以下で述べる水平クラスタリングという考え方を用いることで実現される。頂点 $v \in V_n$ は、その長さ $n-2$ の prefix が以下の操作によって c になるとき、クラス c であるという。

- もし v の $(n-1)$ 桁目が n でなければ、v の n 桁目を (必要であれば)n にし、その後で、$(n-1)$ 桁目を (必要であれば)$n-1$ にする。
- もし v の $(n-1)$ 桁目が n であれば、v の $n-1$ 桁目を (必要であれば)$n-1$ にし、その後で、n 桁目を (必要であれば)n にする。

すなわちクラス c は、$TG(n-1,n)$ 中の頂点 "$c\,n-1\,n$" から以下の手順で得られる頂点たちである：

1. 最初に $(n-1)$ 桁目を変更し、そのあとで n 桁目を変更する。
2. 最初に $(n-1)$ 桁目と n 桁目を交換し、その後で n 桁目を変更する。

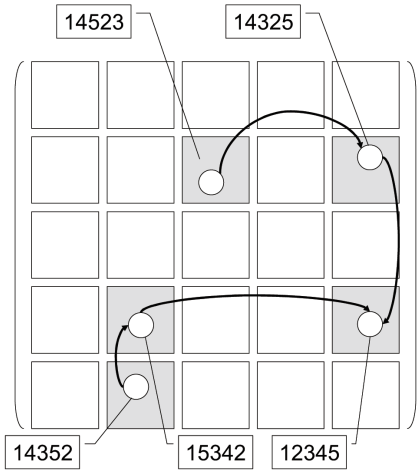


図 4: 例 1.

ここで T_n にはそのようなクラスが $(n-2)!$ 個あり、各クラスは $2f_n$ 個の頂点から成っていることに注意しよう。以下では頂点 $u \in V_n$ の属しているクラスを $c(u)$ であらわし、$(n-2)!$ 個のクラス全体の集合を $\mathcal{C}$ であらわすことにする。

例 1: T_5 中の頂点 14523 はクラス 123 である。なぜならば、この頂点に $g_{3,5}$ と $g_{2,4}$ を順に適用することで頂点 12345 になるからである。同様に、頂点 14352 もクラス 123 に含まれる。なぜならば、この頂点に $g_{2,3}$ と $g_{2,5}$ を順に適用することで頂点 12345 になるからである。図 4を参照されたい。ここで、 各クラスの頂点集合が、対応するジェネレータによって誘導される生成木によって特徴づけられることに注意しよう。

ここで注意しなくてはならないのは、異なるクラスに属する 2 つの頂点が与えられたとき、それらの最後の 2 桁 (長さ 2 の suffix) の変化は、安全でない垂直クラスターを経由した点素なパスによって実現されるということである。すなわち、図 4 に示すような二次元配列が各クラスごとに用意され、しかもそれらは互いに素になっている。したがって、もし D 中の頂点がそれぞれ異なるクラスに属していれば、D 中の各頂点の安全な垂直クラスターへの単射 (割り当て) は、点素なパスによって容易に実現できることになる。具体的には、パスの 3 番目のパートを図に示される生成木に沿ってつくればよい。したがって残る問題は、D 中の頂点にクラスの重複がある場合にどうするか、という点に絞られることになる。

目的頂点のクラス替え

以下では、以下の性質を満たす中間頂点の集合 I $(\subset V_n)$ を構成する：1) I 中の各頂点はそれぞれ別のクラスに属していること、2) D と I が互いに素なリンクの集合によって結ばれていること、3) 集合 D が (前節までに述べた)I と s を結ぶパスの集合と共通頂点をもたないこと。これにより、残されていた最後の 1 ホップが実現できる。

D 中の各頂点 u が各垂直クラスタの中で f_{n-2} 個の隣接頂点を持っていたことを思い出そう。そのような隣接頂点の集合を $N'(u)$ とする。頂点 u の隣接頂点から D を除いたものたちの属しているクラスの集合を $C(u)$ とする。すなわち $C(u)$ を

$$C(u) \stackrel{\text{def}}{=} \{c(w) : w \in N'(u) \setminus D\} \cup \{c(u)\}.$$

のように定義する。ここで $N(u) \cap D = \emptyset$ のとき $|C(u)| = f_{n-2}+1$ であることに注意しよう。次に以下のような二部グラフ H を考える：頂点集合が $D \cup \mathcal{C}$ で、頂点 $v \in D$ と頂点 $c \in \mathcal{C}$ は、$c \in C(v)$ のときかつそのときに限り辺で結ばれる。もし H が D 中の頂点をすべて被覆するようなマッチング E を持てば、中間頂点の集合 I を以下のように決定することができる：

- u が $c(u) \in \mathcal{C}$ と E 中の辺で結ばれていれば、u は自分自身を中間頂点とする。
- もし u が $c' \neq c(u)$ と E 中の辺で結ばれていれば、u は $c(v) = c'$ であるような隣接頂点 $v \in N'(u)$ を選び、v を中間頂点とする。

この構成法が上で述べたすべての条件を満たしていることに注意されたい。

以下の定理が得られる。

定理 4: $n \geq 7$ のとき、二部グラフ H が D のすべての頂点を被覆するマッチングをもつか、2 つ以上の目的頂点を含むような垂直クラスタが高々 $2n-5$ 個存在するかのいずれか一方が成立する。

証明の基本アイデアは、二部グラフ H が D を被覆するようなマッチングをもつのは、任意の X $(\subseteq D)$ に対して X に隣接している頂点数が $|X|$ 以上であるときかつそのときに限られるという、よく知られた性質を利用することである。

マッチングが存在していれば、マッチング中の辺を最後の 1 ホップとして用いることでパスを完

成させることができる。一方もしマッチングが存在しなければ、D 中の頂点に対応するクラス数が高々$2n-4$ $(=2n-5+1)$ であり、少なくとも $f_{n-2}+1$ 個の頂点が属しているようなクラス c が存在することから、考慮すべき目的頂点数を f_n から高々$4n-10$ $(=2(2n-5))$ に削減することができることになる (ちょうどひとつの頂点を含むような垂直クラスタに関しては、クラスの変換を行う必要はない)。

残された頂点のクラス変換は、高々$4n-10$ 個の頂点と $|\mathcal{C}|$ 個の頂点とを結ぶような別の二部グラフ H' をつくることで実現できる。ところが、$n \geq 7$ のとき $f_{n-2}+1 > 2n-4$ であるから、H' 上では、任意の部分集合 X について、X に隣接している頂点数は X 未満ではありえず、したがって、H' には必ずいっぽうの頂点集合を被覆するようなマッチングが存在することになる (つまり、この場合もやはり最後の 1 ホップを何とか実現できることになる)。得られるパスの長さが最短経路よりも高々14 ほど長いことは、構成法から容易に示すことができる。

研究業績一覧

学術論文

1. 森本康彦, 八木一光:
   "POS データからの売上変動パターンの発掘", 日本データベース学会 Letters (DBSJ Letters), Vol. 4, No. 2, 97–100, 2005 年 10 月.
   概要: ある商品の売上の増減は，他の商品の売上の増減に影響を与えるという仮説にもとづき，多数の商品の日々の売上が記録されている POS データから商品の売上増減事象に着目した時系列パターン発掘を行った. 本論文では，時系列パターンおよび推移的時系列パターンの確信度の最適化問題と，その空間効率のよいアルゴリズムについて簡単に述べ，実際の POS データから得られた解析結果と本手法の有用性について報告する.

2. D. Takemoto, S. Tagashira, S. Fujita:
   "Distributed Zone Partitioning Schemes for CAN and its Application to the Load Balancing in Pure P2P Systems", 情報処理学会論文誌, 46(2), 348–356, 2005 年 2 月.
   概要: In this paper, we propose several distributed zone partitioning schemes for Content-Addressable Networks (CAN), which is known as a pure peer-to-peer system based on the Distributed Hash Table (DHT). The main objective of the proposed schemes is to balance the load of nodes in the CAN system, in such a way that every node receives almost the same number of inquiries from the other nodes in the system. The result of simulations implies that, by using the proposed schemes instead of a randomized scheme originally implemented in the CAN system, we could reduce the response time for each inquiry to less than 75%.

3. H. Osada, S. Fujita:
   "CHQ: A Multiagent Reinforcement Learning Scheme for Partially Observable Markov Decision Processes", IEICE Trans. Information and Systems, E88-D(5), 1004–1011, 2005 年 5 月.
   概要: In this paper, we propose a new reinforcement learning scheme called CHQ that could efficiently acquire appropriate policies under partially observable Markov decision processes (POMDP) involving probabilistic state transitions, that frequently occurs in multi-agent systems in which each agent independently takes a probabilistic action based on a partial observation of the underlying environment. A key idea of CHQ is to extend the HQ-learning proposed by Wiering *et al.* in such a way that it could learn the activation order of the MDP subtasks as well as an appropriate policy under each MDP subtask. The goodness of the proposed scheme is experimentally evaluated. The result of experiments implies that it can acquire a deterministic policy with a sufficiently high success rate, even if

the given task is POMDP with probabilistic state transitions.

4. K. Nakano:
"An Energy Efficient Ranking Protocol for Radio Networks", IEICE Trans. on Fundamentals, E89-A(5), 1346–1354, 2006年5月.
概要: A radio network (RN for short) is a distributed system with no central arbiter, consisting of n radio transceivers, henceforth referred to as stations. We assume that the stations run on batteries and expends power while broadcasting/receiving a data packet. Thus, the most important measure to evaluate protocols on the radio network is the number of awake time slots, in which a station is broadcasting/receiving a data packet. We also assume that the stations are identical and have no unique ID number, and no station knows the number n of the stations. For given n keys one for each station, the ranking problem asks each station to determine the number of keys in the RN smaller than its own key. The main contribution of this paper is to present an optimal randomized ranking protocol on the k-channel RN. Our protocol solves the ranking problem, with high probability, in $O(\frac{n}{k} + \log n)$ time slots with every station being awake for at most $O(\log n)$ time slots. We also prove that any randomized ranking protocol is required to run in expected $\Omega(\frac{n}{k} + \log n)$ time slots with at least one station being awake for expected $\Omega(\log n)$ time slots. Therefore, our ranking protocol is optimal.

5. J. L. Bordim, Y. Ito, K. Nakano:
"An Energy Efficient Leader Election Protocol for Radio Network with a Single Tranceiver", IEICE Trans. on Fundamentals, E89-A(5), 1355–1361, 2006年5月.
概要: In this work we present an energy efficient leader election protocol for anonymous radio network populated with n mobile stations. Previously, Nakano and Olariu have presented a leader election protocol that terminates, with probability exceeding $1 - \frac{1}{f}$ $(f \geq 1)$, in $\log\log n + o(\log\log n) + O(\log f)$ time slots. As the above protocol works under the assumption that every station has the ability to transmit and monitor the channel at the same time, it requires every station to be equipped with two transceivers. This assumption, however, is unrealistic for most mobile stations due to constraints in cost, size, and energy dissipation. Our main contribution is to show that it is possible to elect a leader in an anonymous radio network where each station is equipped with a single transceiver. Quite surprisingly, although every station has only one transceiver, our leader election protocol still runs, with probability exceeding $1 - \frac{1}{f}$ $(f \geq 1)$, in $\log\log n + o(\log\log n) + O(\log f)$ time slots. Moreover, our leader election protocol needs only expected $O(n)$ total awake time slots, while Nakano and Olariu's protocol needs expected $O(n \log\log n)$ total awake time slots. Since every leader election protocol needs at least $\Omega(n)$ awake time slots, our leader election protocol is optimal in terms of the expected awake time slots.

6. S. Tagashira, S. Shirakawa, S. Fujita:
"Proxy-Based Index Caching for Content-Addressable Networks", IEICE Trans. Information and Systems, E89-D(2), 555–562, 2006年2月.
概要: Content-Addressable Network (CAN) provides a mechanism that could retrieve objects in a P2P network by maintaining indices to those objects in a fully decentralized manner. In the CAN system, index caching is a useful technique for reducing the response time of retrieving objects. The key points of effective caching techniques are to improve cache hit ratio by actively sharing caches distributed over

the P2P network with every node and to reduce a maintenance and/or routing overhead for locating the cache of a requested index. In this paper, we propose a new caching technique based on the notion of proxy-type caching techniques which have been widely used in WWW systems. It can achieve active cache sharing by incorporating the concept of proxy caching into the index access mechanism and locate a closer proxy cache of a requested index with a little routing overhead. By the result of simulations, we conclude that it can improve the response time of retrieving indices by 30conventional caching techniques.

7. H. Ochi, S. Tagashira, S. Fujita:
"A Localization Scheme for Sensor Networks Based on Wireless Communication with Anchor Groups", IEICE Trans. Information and Systems, E89-D(5), 1614-1621, 2006年5月.
概要: In this paper, we propose a new localization scheme for wireless sensor networks consisting of a huge number of sensor nodes equipped with simple wireless communication devices such as wireless LAN and Bluetooth. The proposed scheme is based on the Point-In-Triangle (PIT) test proposed by He *et al.* The scheme is actually implemented by using Bluetooth devices of Class 2 standard, and the performance of the scheme is evaluated in an actual environment. The result of experiments indicates that the proposed scheme could realize a localization with an error of less than 2 m.

8. D. Takemoto, S. Tagashira, S. Fujita:
"A Fault-Tolerant Content Addressable Network", IEICE Trans. Information and Systems, E89-D(6), 1923-1930, 2006年6月.
概要: In this paper, we propose a new method to enhance the fault-tolerance of the Content Addressable Network (CAN), which is known as a typical pure P2P system based on the notion of Distributed Hash Table (DHT). The basic idea of the proposed method is to introduce redundancy to the management of index information distributed over the nodes in the given P2P network, by allowing each index to be assigned to several nodes, which was restricted to be one in the original CAN system. To keep the consistency among several copies of indices, we propose an efficient synchronization scheme based on the notion of labels assigned to each copy in a distinct manner. The performance of the proposed scheme is evaluated by simulation. The result of simulations indicates that the proposed scheme significantly enhances the fault-tolerance of the CAN system.

9. S. Tagashira, M. Mito, S. Fujita:
"A Generic Solver based on Functional Parallelism for Solving Combinatorial Optimization Problems", IEICE Trans. Information and Systems, E89-D(6), 1940-1947, 2006年6月.
概要: This paper proposes a new class of parallel branch-and-bound (B&B) schemes. The main idea of the scheme is to focus on the functional parallelism instead of conventional data parallelism, and to support such a heterogeneous and irregular parallelism by using a collection of autonomous agents distributed over the network. After examining several implementation issues, we describe a detail of the prototype system implemented over eight PC's connected by a network. The result of experiments conducted over the prototype system indicates that the proposed parallel processing scheme significantly improves the performance of the underlying B&B scheme by adaptively switching exploring policies adopted by each agent participating to the problem solving.

10. 藤田聡:
“分散ネットワーク上の故障診断問題について”, 電子情報通信学会論文誌 A, J89-A(6), 430-437, 2006 年 6 月.
概要: Preparata らによって 1967 年に提案された分散システムのための自律的な故障診断モデルである PMC モデルは，近年，アルゴリズム理論や計算複雑さに関する研究者たちをまきこんで，新しい展開を見せている．本解説では，この問題の重要さと面白さについて，最新の成果を交えてわかりやすく解説する．

11. S. Fujita, A. Ohtsubo, M. Mito:
“Extended Skip Graphs for Efficient Key Search in Peer-to-Peer Environment”, Journal of Interconnection Networks (JOIN), 8(2), 119-132, 2007 年.
概要: In this paper, we propose three techniques to improve the cost/performance of the skip graph that was recently proposed by Aspnes and Shah. The skip graph, which is a distributed data structure that could efficiently support find, insert, and delete operations of a key drawn from a totally ordered set, consists of N nodes each of which is connected with exactly $\log_2 N$ nodes determined by a set of random binary vectors called membership vectors. In the following, we will extend the construction of the skip graph in the following two directions: 1) proposal of a subgraph of the skip graph which realizes a graceful degradation of the routing performance when the number of neighbors reduces from $\log_2 N$, and 2) proposal of a supergraph of the skip graph which realizes a significant performance improvement when the number of neighbors increases from $\log_2 N$. The performance of those extended graphs will be evaluated analytically.

12. K. Kobatake, S. Tagashira, S. Fujita:
“A New Caching Technique to Support Conjunctive Queries in P2P DHT”, IEICE Trans. Information and Systems, .
概要: P2P DHT (Peer-to-Peer Distributed Hash Table) is one of typical techniques for realizing an efficient management of shared resources distributed over a network, and a keyword search over such networks in a fully distributed manner. In this paper, we propose a new method for supporting conjunctive queries in P2P DHT. The basic idea of the proposed technique is to share a global information on past trials by conducting a local caching of search results for conjunctive queries, and by registering the fact to the global DHT. Such a result caching is expected to significantly reduce the amount of transmitted data compared with conventional schemes. The effect of the proposed method is experimentally evaluated by simulation. The result of experiments indicates that by using the proposed method, the amount of returned data is reduced by 60P2P DHT which does not support conjunctive queries.

13. S. Fujita, T. Araki:
“Three-Round Adaptive Diagnosis in Binary n-Cubes”, Algorithms and Computation, 15th International Symposium, ISAAC 2004, Hong Kong, China, December 2004, Proceedings (R. Fleischer, G. Trippen Eds.), LNCS 3341, 442–451, 2004 年 12 月.
概要: In this paper, we consider the problem of adaptive fault-diagnosis in binary n-cubes, and propose a scheme that completes a diagnosis in at most three test rounds, provided that $n \geq 3$ and the number of faulty vertices is at most n. The proposed scheme is optimal in the sense that: 1) three rounds are necessary for the adaptive diagnosis, 2) there exists a set of $n+1$ faulty vertices that can not be located by any diagnosis scheme, and 3) to identify n faulty vertices, the system must contain at least $2n+1$ vertices. Note that $n = 3$ is the smallest integer satisfying $2^n \geq 2n+1$.

14. S. Fujita, Y. Liang:
"How to Provide Continuous Services by Mobile Servers in Communication Networks", Parallel and Distributed Computing: Applications and Technologies, 5th International Conference, PDCAT 2004, Singapore, December 2004, Proceedings (K.-M. Liew, H. Shen, S. See, W. Cai, P. Fan, S. Horiguchi, Eds.), LNCS 3320, 326–329, 2004 年 12 月.
概要: In this paper, we propose a new framework to provide continuous services by "mobile servers" in interconnection networks. In particular, we prove that for the classes of ring networks with n vertices, $\lceil (n+1)/3 \rceil$ mobile servers are necessary and sufficient to realize mutual transfers among configurations dominating the network, where transfer between two dominating configurations is realized by moving the role of dominating vertex from a vertex to an adjacent vertex.

15. A. Ohtsubo, D. Takemoto, S. Tagashira, S. Fujita:
"The Power of Shortcuts in Greedy Routing in Content Addressable Networks", Embedded and Ubiquitous Computing, International Conference, EUC 2004, Proceedings (L. T. Yang, M. Gua, G. R. Gao, N. K. Jha, Eds.), LNCS 3207, 994–1003, 2004 年 8 月.
概要: In this paper, we propose a scheme to dynamically establish shortcut links in Content Addressable Networks (CAN), based on a probabilistic method suggested by Kleinberg in 2000. The proposed scheme consists of the following three steps: 1) estimation of the total number of nodes participating to the P2P system, 2) dynamic adjustment of the probability distribution for selecting the target point of shortcut based on the estimated number, and 3) the establishment of shortcut to the nodes corresponding to the selected points. The performance of the resulting scheme is evaluated by simulation. The result of simulations implies that the shortcut significantly reduces the average number of hop counts, and the rate of improvement increases as increasing the number of nodes; e.g., the proposed method improves the hop counts in conventional CAN by 49% when it contains 4096 nodes. In addition, dynamic adjustment of the probability distribution improves the performance of the overall scheme, although the amount of improvement is less than "one hop" for relatively small netwotks examined in the simulation.

16. H. Kutsuna, S. Tagashira, S. Fujita:
"Experimental Evaluation of Bimodal Congestion Control Scheme in High-Speed Networks with Broad Bandwidth", Proc. the 2004 IEEE International Midwest Symposium on Circuits and Systems (MWSCAS 2004), III-227–III-230, 2004 年 7 月.
概要: In this paper, we study congestion control schemes in high-speed networks. We focus on Bimodal congestion control scheme proposed by Attie *et al.*, as an alternative to AIMD that is used in the current TCP. By the result of extensive simulations, we found that although Bimodal could improve the throughput of AIMD by 10% in low-speed networks, the amount of improvements will be limited in high-speed networks because of the high frequency of retransmissions of packets.

17. M. Yamane, S. Tagashira, S. Fujita:
"Admission Control of Predictive Services under an Environment with Frequent Retransmissions of Packets", Proc. the 2004 IEEE International Midwest Symposium on Circuits and Systems (MWSCAS 2004), II-681–II-684, 2004 年 7 月.
概要: In this paper, we consider the admission control of communication flows under an environment with frequent retransmis-

sions of packets. The basic idea of the proposed scheme is to improve Jamin *et al.*'s predictive service so as to take into account the effect of retransmissions of packets in a parameterized manner. The goodness of the scheme is evaluated by simulation. The result of simulations indicates that the proposed scheme can guarantee a predetermined delay bound without degrading the link utilization rate by appropriately selecting the value of a configurable parameter.

18. J. L. Bordim, T. Hunziker, K. Nakano: "Adaptive Carrier Sensing and Packet Sending- An Alternative to Boost the Performance in Directional Communications", Proc. of 6th International Conference on PDCAT, 274–278, 2005 年 12 月.
概要: Despite the efforts to leverage the network performance through the use of Directional Communications, most protocols still rely on the basic mechanisms employed by the IEEE802.11 standards. While this might be desirable facilitate compatibility, nodes empowered with directional antennas should be able to explore the benefits of directional communications whenever possible. With that in mind we propose two mechanisms, termed Adaptive Carrier Sensing and Adaptive Packet Sending, aiming to allow nodes equipped with directional antennas to exploit the advantages of such systems to improve network performance.

19. Y. Morimoto: "Optimized Transitive Association Rule: Mining Significant Stopover between Events", Proc. Annual ACM Symposium on Applied Computing (SAC), 547–548, 2005 年 3 月.
概要: We consider a problem of finding optimized transitive association rules. A transitive association rule is a transitive sequence of different events. Each of two consecutive events in a transitive association rule is a conventional association rule. Transitive confidence is the product of confidence values of all the association rules in the event sequence. The optimized transitive association rule from a cause event to an effect event is the optimal sequence of events whose transitive confidence is maximum among all possible sequences from the cause to the effect. In this paper, we present space efficient algorithms for computing all of the optimized transitive association rules between events, whose transitive confidence is not less than a user specified minimum confidence value.

20. K. Yagi, Mai Soramoto, Miho Mokuda, Yasuhiko Morimoto: "Optimized Sequential Pattern Mining from Point Of Sales Data", Proc. IEEE International Special Workshop on Databases for Next Generation Researchers, 12–15, 2005 年 4 月.
概要: We consider an optimization problem of sequential pattern mining. We present space efficient algorithms for computing optimized sequential pattern that maximize transitive confidence from a cause event to an effect event. Those can compute optimized sequential patterns for all cause-effect pairs of sales fluctuation events in large POS data.

21. H. Ochi, S. Tagashira, S. Fujita: "A Localization Scheme for Sensor Networks Based on Wireless Communication with Anchor Groups", Proc. the 11th Int'l Conf. on Parallel and Distributed Systems (ICPADS05), 299–305, 2005 年 7 月.
概要: In this paper, we propose a new localization scheme for wireless sensor networks consisting of a huge number of sensor nodes equipped with simple wireless communication devices. The proposed scheme is based on the Point-In-Triangle (PIT) test proposed by He *et al.* The scheme is actually

implemented by using Bluetooth devices of Class 2 standard, and the performance of the scheme is evaluated in an actual environment. The result of experiments indicates that the proposed scheme could realize a localization with an error of less than 2 m.

22. S. Fujita, S. Tagashira, C. Qiao, and M. Mito:
"Distributed Branch-and-Bound Scheme for Solving the Winner Determination Problem in Combinatorial Auctions", Proc. Int'l Conf. on Advanced Information Networking and Applications (AINA05), 661–666, 2005 年 3 月.
概要: In this paper, we propose a new class of parallel branch-and-bound (B&B) schemes. The main idea of the scheme is to focus on the *functional parallelism* instead of conventional *data parallelism*, and to support such a heterogeneous and irregular parallelism by using a collection of autonomous agents distributed over the network. After examining several design issues toward the implementation of a prototype of the distributed B&B system, we illustrate the result of our preliminary experiments conducted to estimate the performance of the proposed scheme. The result shows that it could cause a significant performance improvement if each agent autonomously changes its function type according to the change of the underlying environment.

23. T. Tang, S. Tagashira, S, Fujita:
"LQ-Routing Protocol for Mobile Ad-Hoc Networks", Proc. Int'l Conf. on Computer and Information Science (ACIS-ICIS 2005), 441–446, 2005 年 7 月.
概要: In this paper, we propose a new routing protocol for mobile ad-hoc networks (MANET), which have attracted considerable attentions in recent years as a tool to realize flexible and robust communication environments in many critical application fields. The proposed protocol, called LQ-routing, is a proactive routing protocol based on the Q-routing protocol which was originally proposed for conventional wired networks. In order to reflect the mobility of hosts to the routing decision, which would cause frequent connection and disconnection of communication links between nearby hosts, we introduce the notion of lifetime to represent the stability of routes connecting to the destination. The performance of the proposed routing scheme is experimentally evaluated by using NS-2. The result of simulation indicates that by introducing the notion of lifetime to a simple combination of previous proactive routing scheme DSDV and the Q-routing, we could improve the performance of DSDV in terms of the packet loss rate and the average packet delay.

24. S. Fujita:
"On the relation between average and maximum cuts in linear arrangements of graphs", 2005 Korea-Japan Joint Workshop on Algorithms and Computation, 191–198, 2005 年 8 月.
概要: This paper studies the relationship between the average and the maximum cut of linear arrangement of graphs.

25. S. Tagashira, M. Mito, S. Fujita:
"Towards Generic A Branch-and-Bound Solver of Combinatorial Optimization Problems with Autonomous Agents in P2P Networks", Proc. Int'l Symp. on High Performance Computing (ISHPC'05), CD-ROM, 2005 年 9 月.
概要: In this paper, we propose a new class of parallel branch-and-bound (B&B) schemes. The main idea of the scheme is to focus on the *functional parallelism* instead of conventional *data parallelism*, and

to support such a heterogeneous and irregular parallelism by using a collection of autonomous agents distributed over the network. After examining several design issues toward the implementation of a prototype of the distributed B&B system, we illustrate the result of our preliminary experiments conducted to estimate the performance of the proposed scheme. The result shows that it could cause a significant performance improvement if each agent autonomously changes its function type according to the change of the underlying environment.

26. D. Takemoto, S. Tagashira, S. Fujita: "A Fault-Tolerant Content Addressable Network", Proc. Int'l Symp. on Parallel and Distributed Processing and Applications (ISPA'2005), Lecture Notes in Computer Science 3758, 792 – 803, 2005年11月.
概要: In this paper, we propose a new method to enhance the fault-tolerance of the Content Addressable Network (CAN), which is known as a typical pure P2P system based on the notion of Distributed Hash Table (DHT). The basic idea of the proposed method is to introduce a redundancy to the management of index information distributed over the nodes in the network, by allowing each index to be assigned to several nodes, which was restricted to be one in the original CAN system. To keep the consistency among several copies of indices, we propose an efficient synchronization scheme based on the notion of labels assigned to each copy in a distinct manner. The performance of the proposed scheme is evaluated by simulation. The result of simulations indicates that the proposed scheme really enhances the fault-tolerance of the CAN system.

27. S. Umezane, S. Fujita: "On Linkage Identification in EC for Solving Combinatorial Optimization Problems", Proc. The International Conference on Systems, Man and Cybernetics (SMC2005), CD-ROM, 2005年11月.
概要: This paper proposes a new linkage identification method based on the notion of evolution of linkages, that is designed for evolutionary computation schemes to solve combinatorial optimization problems. Linkage identification is a task to identify an appropriate fragment of genes contained in a chromosome that is expected to significantly increase the fitness of the overall individuals. As a concrete target problem of the proposed scheme, we focus on the traveling salesman problem (TSP) that is known as a typical combinatorial optimization problem. The performance of the scheme is experimentally evaluated by using TSPLIB, and is compared with the performance of previous schemes. The result of experiments shows that the proposed scheme really improves the performance of the previous scheme, and the amount of improvement is 42.7% in the best case.

28. S. Fujita, A. Ohtsubo, M. Mito: "Extended Skip Graphs for Efficient Key Search in P2P Environment", Proc. the 8th International Symposium on Parallel Architectures, Algorithms, and Networks (I-SPAN 2005), 256–261, 2005年12月.
概要: In this paper, we propose three techniques to improve the cost/performance of the skip graph that was recently proposed by Aspnes and Shah in 2003. More concretely, the skip graph, in which each node is connected with exactly $\log_2 N$ neighbors among N nodes contained in the system, is extended in the following two directions: 1) proposal of a subgraph of the skip graph which realizes a graceful degradation of the routing performance when the number of neighbors reduces from $\log_2 N$, and 2) proposal of a supergraph of the skip graph

which realizes a significant performance improvement when the number of neighbors increases from $\log_2 N$. The performance of those extended graphs will be evaluated analytically with concrete numerical results.

29. S. Fujita:
"A Tight Bound on the Number of Mobile Servers to Guarantee the Mutual Transferability among Dominating Configurations", Proc. Algorithms and Computation, 16th International Symposium, ISAAC 2005, Proceedings (Xiaotie Deng, Ding-Zhu Du Eds.). Lecture Notes in Computer Science, 563–572, 2005 年 12 月.
概要: In this paper, we propose a new framework to provide continuous services to users by a collection of mobile servers distributed over an interconnection network. We model those mobile servers as a subset of host computers, and assume that a user host can receive the service if at least one adjacent host computer (including itself) plays the role of a server; i.e., we assume that the service could not be routed via the interconnection network. The main results obtained in this paper are summarized as follows: For the class of trees with n hosts, $\lceil (n+1)/2 \rceil$ mobile servers are necessary and sufficient to realize continuous services by the mobile servers, and for the class of Hamiltonian graphs with n hosts, $\lceil (n+1)/3 \rceil$ mobile servers are necessary and sufficient.

30. J. L. Bordim, T. Hunziker, K. Nakano:
"Limiting the Effects of Deafness and Hidden Terminal Problems in Directional Communications", Proc. International Symposium on Parallel and Distributed Processing and Application, 195–206, 2006 年 12 月.
概要: It is known that wireless ad hoc networks employing omnidirectional communications suffer from poor network throughput due to inefficient spatial reuse. Although the use of directional communications is expected to provide significant improvements, the lack of efficient mechanisms to deal with deafness and hidden terminal problems makes it difficult to fully explore its benefits. The main contribution of this work is to propose a Medium Access Control (MAC) scheme which aims to lessen the effects of deafness and hidden terminal problems in directional communications without precluding spatial reuse. Unlike other proposals that focus on exploring the characteristics of the physical layer, the proposed MAC protocol relies on simple mechanisms that can be easily coupled with a directional antenna without requiring major modifications to the current MAC standard.

31. J. L. Bordim, Y. Ito, K. Nakano:
"Randomized Leader Election Protocols in Noisy Radio Networks with a Single Transceiver", Proc. International Symposium on Parallel and Distributed Processing and Application, 246–256, 2006 年 12 月.
概要: In this work, we present leader election protocols for single-hop, single-channel noisy radio networks that do not have collision detection (CD) capabilities. In most leader election protocols presented so far, it is assumed that every station has the ability to transmit and monitor the channel at the same time, it requires every station to be equipped with two transceivers. This assumption, however, is unrealistic for most mobile stations due to constraints in cost, size, and energy dissipation. Our main contribution is to show that it is possible to elect a leader in an anonymous radio network where each station is equipped with a single transceiver. We first present a leader election protocol for the case the number n of stations is known beforehand. The protocol runs in $O(\log f)$ time slots with probability at least $1 - \frac{1}{f}$ for any $f > 1$. We then present a leader election protocol

for the case where n is not known beforehand but an upper bound u of n is known. This protocol runs in $O(\log f \log u)$ time slots with probability at least $1 - \frac{1}{f}$ for any $f > 1$. We also prove that these protocols are optimal. More precisely, we show that any leader election protocol elect a leader with probability at least $1 - \frac{1}{f}$ must run in $\Omega(\log f)$ time slots if n is known. Also, we proved that any leader election protocol elect a leader with probability at least $1 - \frac{1}{f}$ must run in $\Omega(\log f \log u)$ time slots if an upper bound u of n is known.

32. S. Fujita:
"Loose Covering by Graphs", 2006 Japan-Korea Joint Workshop on Algorithms and Computation, 9–15, 2006 年 7 月.
概要: This paper introduces a new notion of an embedding of local structure to a global network, defined as a loose covering by graphs. In this paper, we derive several basic properties of the notion including the computational complexity of the problem of finding a loose covering of a given graph by small cycles and paths.

33. H. Kutsuna, S. Tagashira, S. Fujita:
"A Fair and Efficient Congestion Control Scheme Based on Minority Game", Proc. The 14th IEEE International Conference on Networks (ICON '06), 98–103, 2006 年 9 月.
概要: In this paper, we propose a new congestion control scheme for high-speed networks. The basic idea of the scheme is to adopt a game theory called Minority Game (MG) to realize a *selective reduction* of the transmission speed of senders. More concretely, upon detecting a congestion, it plays a game among all senders participating in the communication, and reduces the transmission speed of each sender according to the result of the game. MG is a game that has recently attracted considerable attentions, and it is known to have a remarkable property such that the number of winners converges to a half of the number of players in spite of the selfish behavior of the players. By using such property of MG, we can realize a fair reduction of the transmission speed, which is more efficient than previous schemes in which all senders uniformly reduce their transmission speed. The effect of the proposed scheme is evaluated by simulation. The result of simulations indicates that the proposed scheme certainly realizes a selective reduction of the transmission speed, and it is sufficiently fair compared with simple randomized schemes and is sufficiently efficient compared with conventional schemes.

34. K. Kobatake, S. Tagashira, S. Fujita:
"A New Caching Technique to Support Conjunctive Queries in P2P DHT", Proc. IFIP Int'l Conf. on Network and Parallel Computing, 185–190, 2006 年 10 月.
概要: P2P DHT (Peer-to-Peer Distributed Hash Table) is a typical technique to realize an efficient keyword search in pure P2P systems. In this paper, we propose a new caching technique to support conjunctive queries in P2P DHT. The basic idea of the proposed technique is to share a global information on past trials by conducting a local caching of search results for conjunctive queries, and by registering the fact to the global DHT. Such a result caching is expected to significantly reduce the amount of transmitted data compared with conventional schemes. The effect of the proposed method is experimentally evaluated by simulation. The result of experiments indicates that it reduces the amount of transmitted data by 60% compared with conventional P2P DHT.

35. S. Fujita:
"Vertex-Disjoint Paths in Transposition Graphs", Proc. the 18th IASTED Int'l

Conf. on Parallel and Distributed Computing and Systems, 490–494, 2006 年 11 月.

概要: A transposition graph is a Cayley graph in which vertices correspond to permutations and edges are placed between permutations that differ by exactly one transposition. In this paper, we propose an algorithm to find a collection of vertex-disjoint paths connecting a given source vertex s and a given set of destination vertices D. The running time of the algorithm is polynomial to the number of destination vertices, and the resultant path connecting s and each destination is longer than the length of a shortest path connecting those vertices by at most fourteen.

36. K. Kobatake, S. Tagashira, S. Fujita: "Efficient Support of Conjunctive Queries in P2P DHT by using Bloom Filter", Proc. the 18th IASTED Int'l Conf. on Parallel and Distributed Computing and Systems, 202–207, 2006 年 11 月.

    概要: P2P DHT (Peer-to-Peer Distributed Hash Table) is a typical technique to realize an efficient keyword search in pure P2P systems. In this paper, we propose a new method to support conjunctive queries in P2P DHT. The basic idea of the proposed method is to cache the result of conjunctive queries to each peer, and to register the fact to the DHT. In addition, in order to reduce the number of meaningless queries to DHT, it registers cached conjunctions to Bloom filter, and circulates it among all peers. The effect of the proposed method is experimentally evaluated by simulation. The result of experiments indicates that by using the proposed method, the amount of returned data reduces to 40%, and the number of queries reduces to 66%, compared with conventional P2P DHT.

37. K. Yasuda, S. Tagashira, S. Fujita: "Adaptive MAC Scheme for Wireless LAN based on Dynamic Group Constructions", Proc. Int'l Conf. on Parallel and Distributed Computing, Applications and Technologies, 300–305, 2006 年 12 月.

    概要: In wireless LAN, several terminals could share a single communication channel by using an appropriate media access control (MAC) mechanism. However, existing MAC schemes for wireless LAN have serious drawbacks such that it significantly degrades the throughput if it frequently occurs a collision of frames or a useless assignment of communication bandwidth. In this paper, we propose a new MAC scheme for wireless LAN to realize an efficient utilization of the communication bandwidth while suppressing the probability of frame collisions. The basic idea of the scheme is to conduct a channel assignment for each group which is dynamically constructed based on the unique identifier that is assigned to each terminal beforehand. The scheme could be customized by appropriately tuning parameters according to the given communication environment. The effect of the scheme is estimated by simulation. By the result of simulations, it is confirmed that the proposed scheme could achieve a stable performance for any access frequency.

38. M. Yamane, S. Tagashira, S. Fujita: "An Efficient Assignment of Transmission Opportunity in QoS Guaranteed Wireless LAN", Proc. Int'l Conf. on Parallel and Distributed Computing, Applications and Technologies, 105–108, 2006 年 12 月.

    概要: IEEE802.11e standard adopts several important techniques to guarantee the quality of service (QoS) of multimedia communications in wireless LAN. A key technique introduced in the standard is an efficient media access control based on Hybrid Coordination Function (HCF), in which each flow is assigned an appropriate trans-

mission opportunity (TXOP) to satisfy several requirements on the flow such as the delay bound and the communication bandwidth. Although it could provide an efficient tool to guarantee the QoS in wireless LAN, HCF has a serious drawback such that a useless bandwidth assignment frequently occurs, which is primarily due to the limitation of the scheme such that all flows share the same service interval (SI). In this paper, we propose a new scheduling scheme that associates each flow with its own SI according to the hardness of the delay bound. The effect of the scheme is experimentally evaluated by simulation. The result of simulations indicates that the number of admitted flows and the data transmission rate certainly increase by the proposed scheme.

39. Y. Morimoto:
"Geometrical Information Fusion from WWW and Its Related Information", Proc. The 5nd International Workshop on Databases in Networked Information System, LNCS 4777, 16–32, 2007 年 10 月.
概要: We considered a spatial data mining method that extracts spatial knowledge by computing geometrical patterns from web pages and access log files of HTTP servers. There are many web pages that contain location information such as addresses, postal codes, and telephone numbers. We can collect such web pages by web-crawling programs. For each page determined to contain location information, we apply geocoding techniques to compute geographic coordinates, such as latitude-longitude pairs. Next, we augment the location information with keyword descriptors extracted from the web page contents. We then apply spatial data mining techniques on the augmented location information. In addition, we can use hyperlinks and access log files to find linkage between pages with location information to derive spatial knowledge.

40. A. Ohtsubo, S. Tagashira, S. Fujita:
"A Content Addressable Small-World Network", Proc. Parallel and Distributed Computing and Networks, 551-111, 2007 年 2 月.
概要: In this paper, we study the effect of shortcut links in overlay networks for distributed systems. In particular, we exam ine the effect of shortcuts in Content Addressable Network (CAN) which is known as a typical pure P2P network based on the notion of Distributed Hash Table (DHT). The pro posed scheme is an extension of Kleinberg's method which was originally proposed for the study of small-world net works. The result of simulations indicates that the proposed shortcutting scheme significantly reduces the average num ber of hopcounts in a greedy routing, and the rate of im provement increases as increasing the number of nodes.

41. K. Maruyama, S. Tagashira, S. Fujita :
"A P2P Information Monitoring System Supporting Conjunctive Continual Queries", Proc. Parallel and Distributed Computing and Networks, 551-109, 2007 年 2 月.
概要: In this paper, we propose a P2P information monitoring system to support the processing of conjunctive queries. The proposed system is an extension of PeerCQ, and our main contribution in this paper is the following three points: 1) we provide a detailed speci?cation of Conjunc tive Continual Queries (CCQs) that is an agent to check the satisfaction of a given conjunction, 2) we propose a map ping of CCQs to peers by considering the similarity of the referred data items, and 3) we improve the utility function used in PeerCQ in such a way that the trade-off between the load of each peer and the overall load of the system can be explicitly controlled. The effect

of the proposed method is evaluated by simulation.

42. W. Liu, S. Tagashira, S. Fujita:
"SwRED: A Robust Active Queue Management Scheme based on Load Level Prediction", Proc. Parallel and Distriuted Processing Techniques and Applications, Volume 1, 336–341, 2007年6月.
概要: Active queue management (AQM) is a novel technique to realize a smooth congestion control in communication networks, which is expected to work effectively by combining it with end-to-end congestion control schemes such as the window control adopted in TCP. Although there are a lot of AQM schemes proposed in the literature, an appropriate tuning of parameters used in those schemes has been remaining as a challenging problem to be overcome. In this paper, we propose a new AQM scheme based on a load level prediction of bottleneck links, and on a combination of queue-based and rate-based AQM schemes in multiple time scales. Different from most of conventional AQM schemes, the proposed scheme, called Switch-RED (SwRED), employs a new packet drop strategy based on a macroscopic analysis of the queue dynamics. The performance of the resultant scheme is evaluated by simulation. The result of simulations indicates that SwRED realizes a more stable behavior compared with previous schemes, i.e., it realizes a short queuing delay and a low packet loss rate while keeping a high link utilization.

43. X.-R. Wang, S. Tagashira, S. Fujita:
"FDAR: A Load-Balanced Routing Scheme for Mobile Ad-hoc Networks", Proc. 6th International Conference on Ad-Hoc, Mobile, and Wireless Networks, ADHOC-NOW 2007, LNCS 4686, 186–197, 2007年9月.
概要: In this paper, we propose an efficient and practical routing scheme for MANETs based on a new routing metric called the free-degree of nodes. In the proposed routing scheme named FDAR (Free-Degree Adaptive Routing), it is intended to deliver data packets circumventing congested routes, so as to realize a short end-to-end delay and a moderate load balancing of the overall network. In addition, it tries to avoid frequent and unnecessary invocation of route discovery process which would significantly degrade the routing performance of the MANET. The effectiveness of the proposed scheme is evaluated by simulation. The result of simulations indicates that it certainly outperforms previous load-balanced routing schemes including DLAR and LBAR, in terms of the packet loss rate and the average end-to-end delay.

44. K. Doi, S. Tagashira, S. Fujita:
"Proximity-Aware Content Addressable Network Based on Vivaldi Network Coordinate System", Proc. Fifth International Workshop on Databases, Information Systems and Peer-to-Peer Computing, .
概要: Distributed hash tables (DHTs) have attracted considerable attentions as a way to realize an efficient index search in P2P environments. However, most of conventional DHTs do not take into account the proximity of nodes in the physical network, which causes a performance degradation as increasing the size of the resultant overlay network. In this paper, we propose a new approach to construct a proximity-aware DHT-based P2P. More concretely, we introduce the notion of proximity to the Content Addressable Network (CAN) by using a mapping from a virtual coordinate space realized by the Vivaldi network coordinate system. The effect of the proposed method is evaluated by simulation. The result of simulations indicates that the proposed scheme reduces the

average physical distance between nodes in the overlay by 10% for random physical networks, and by 80% for grid-structured physical networks.

45. A. Sugimoto, S. Tagashira, S. Fujita:
"A Peer Selection Scheme for P2P Download Systems based on a Matchmaking Method", Proc. Fifth International Workshop on Databases, Information Systems and Peer-to-Peer Computing, .
概要: According to the recent advancement of network technologies, a demand for high-speed download of giga-scale contents is increasing. In this paper, we propose a new technique to improve the efficiency of parallel download systems of BitTorrent (BT) type. A key point of the proposed scheme is that: 1) each peer first conducts oblivious piece selections over a randomly organized overlay network, and 2) when the number of uncollected pieces becomes lower than a predetermined threshold, it reorganizes the overlay by conducting an explicit matchmaking. The performance of the proposed scheme is evaluated by simulation. The result of simulation shows that it certainly reduces the overhead without increasing the number of elapsed rounds in conventional schemes.

46. S. Fujita:
"Vertex Domination in Dynamic Networks", Proc. Workshop on Algorithms and Computation (WALCOM), LNCS 4921, 1–12, 2008 年 2 月.
概要: This paper studies a vertex domination problem in dynamic networks, which allows dynamic change of the set of vertices, the set of edges, and the set of weights. In particular, we will examine the following two theoretical issues arising in designing an adaptive vertex domination scheme for such networks: 1) How can we transfer a given configuration to a dominating configuration while keeping that any intermediate configuration is safe? 2) How can we reduce the computational complexity of the vertex domination problem by allowing defections in the domination?

研究会等

1. 中尾, 水戸, 田頭, 藤田:
"P2P オーバレイネットワークにおけるプローブを用いた共有情報の範囲検索法", グループウェアとネットワークサービスワークショップ 2004(GNWS2004), 69–74, 2004 年 11 月.
概要: P2P 環境では，様々な属性をもつ情報が多くの計算機 (ピア) により共有されている．これらの共有情報の中から，特定の属性に対して，ある範囲内に含まれる情報を発見する範囲検索技術が注目されている．このような検索を P2P 環境で実現するためには，頻繁に発生するピアの参加・離脱を想定した効率的な手法が必要である．本研究では，リング型のオーバレイネットワークをピア間で構築し，そのネットワーク上にショートカット (プローブ) を付加して P2P 環境に適した範囲検索手法を提案する．提案手法では，プローブを用いて効率的な検索を実現し，さらに，楽観的にプローブの付加と管理をおこなうことで，参加・離脱に伴うオーバレイネットワークの変化に柔軟に対応する．我々は提案手法を実験により評価し，その有効性を確認した．その結果から，提案手法は様々な属性値の分布に対して，平均 $O(\sqrt{n/d})$ ステップで目的とする情報を発見できることがわかった．

2. 梅實, 藤田:
"リンケージ同定を導入した TSP のための進化計算解法の提案", 電子情報通信学会 ニューロコンピューティング研究会, Vol.104, No.474, 49–54, 2004 年 11 月.
概要: 本稿では，遺伝的アルゴリズム (GA) のための新しいリンケージ同定法の提案と評価を行なう．GA は組合せ最適化問題のためのメタ解法のひとつとして注目されており，

これまでにその改良や実装に関する数多くの提案がなされている．提案手法では，組合せ最適化問題のための GA 解法においては “遺伝子間の関係” が特に強い意味をもっていることに着目し，それらの関係を考慮した高精度のリンケージ同定を DSMGA 法を改良することで実現している．提案手法の性能は，TSP のためのベンチマークセットである TSBLIB を用いて実験的に評価された．その結果，提案手法には比較的精度の高い解を少ないばらつきで生成する効果があることが明らかになった．

3. C. Qiao, M. Mito, S. Tagashira, S. Fujita:
“Parallel Branch-and-Bound Scheme for the Winner Determination Problem in Combinatorial Auctions”, 情報処理学会 計算機アーキテクチャ研究会, No.2004-ARC-160, 23–28, 2004 年 12 月.
概要: In this paper, we propose a new class of parallel branch-and-bound (B&B) schemes. The main idea of the scheme is to focus on the *functional parallelism* instead of conventional *data parallelism*, and to support such a heterogeneous and irregular parallelism by using a collection of autonomous agents distributed over the network. After examining several design issues toward the implementation of a prototype of the distributed B&B system, we illustrate the result of our preliminary experiments conducted to estimate the performance of the proposed scheme. The result shows that it could cause a significant performance improvement if each agent autonomously changes its function type according to the change of the underlying environment.

4. 越智，田頭, 藤田:
“無線センサーネットワークを利用した位置測位システムの構築”, 情報処理学会 計算機アーキテクチャ研究会, No.2004-ARC-160, 17–22, 2004 年 12 月.
概要: ユビキタスコンピューティング環境における重要な課題のひとつに，実空間におけるヒトやモノの位置情報を如何にして正確に把握するかという問題がある．本稿では，無線通信機能を搭載した微小なデバイスたちからなるセンサネットワークを対象とし，ある領域に分散された各センサノードが，外部から与えられる情報をもとに，自身の位置情報を自律的に獲得していくための新しい手法の提案と評価を行なう．

5. 小畠, 竹本, 田頭, 藤田:
“P2P システム上でのリザルトキャッシングを用いた条件付検索手法”, 電子情報通信学会 ネットワークシステム研究会, Vol.104, No.513, 7–12, 2004 年 12 月.
概要: P2P システム上で共有資源を効率的に管理するための手法のひとつとして，DHT(分散ハッシュテーブル) を用いた情報管理手法が注目されている．本稿では，DHT を用いた P2P システム (P2P DHT) 上で AND 検索や OR 検索のような条件付き検索を効率的に実現するための手法の提案を行う．提案手法では，一度検索した結果を各ピアがローカルにキャッシュし，その情報を将来ほかのピアが行う検索で再利用することで，検索コストの大幅な削減を図っている．また各ピアから出された検索内容をキャッシュ内に存在する情報とうまく関連付けることで，キャッシュの利用率を向上させる試みもあわせて行っている．提案手法の有効性はシミュレーションにより評価される．

6. S. Fujita, Y. Liang:
“On the mutual transferability among dominating sets in graphs”, 情報処理学会 アルゴリズム研究会, 2005-AL-099, 2005 年 1 月.
概要: In this paper, we propose a new framework to provide continuous services by “mobile servers” in interconnection networks. We model those mobile servers as a subset of host nodes, and assume that a node can receive the service if at least one adjacent node plays the role of a server (i.e., we assume that the service could not be routed via the interconnection network).

The main results are summarized as follows: For the classes of trees with n vertices, $\lceil (n+1)/2 \rceil$ mobile servers are necessary and sufficient to realize continuous services by the mobile servers, and for the classes of Hamiltonian graphs with n vertices, $\lceil (n+1)/3 \rceil$ mobile servers are necessary and sufficient.

7. T. Tao, S. Tagashira, S. Fujita:
"Performance Evaluation of Q-routing Protocol for Mobile Ad-hoc Networks", 情報処理学会 システム評価研究会, Vol.2005, No.34, 13–18, 2005 年 3 月.
概要: In this paper, we propose a new routing protocol for mobile ad-hoc networks (MANET), that have attracted considerable attentions in recent years as a tool to realize flexible and robust communication environments in many critical application fields. The proposed protocol, called LQ-routing, is a proactive routing protocol based on the Q-routing proposed by Littman and Boyan in 1993. In order to reflect the mobility of hosts to the routing decision, that would cause frequent connection and disconnection of communication links between nearby hosts, we introduce the notion of lifetime to represent the stability of routes connecting to the destination. The performance of the proposed routing scheme is experimentally evaluated by using NS-2. The result of simulation implies that by introducing the notion of lifetime to a simple combination of previous proactive routing scheme DSDV and the Q-routing, we could improve the performance of DSDV in terms of the packet loss rate and the average packet delay.

8. 藤田 聡:
"P2P ファイル検索のための拡張スキップグラフ", 第 19 回　回路とシステム軽井沢ワークショップ, 295–300, 2006 年 4 月.
概要: 本稿では，各ノードの平均次数が定数で，しかも多項対数時間でのクエリのルーティングを可能とするような，新しい構造型オーバーレイネットワークの提案を行う．提案するオーバーレイネットワークは，Aspnes と Shah によって提案されたスキップグラフの拡張となっている．スキップグラフで用いられていたノードあたり $O(\log N)$ 本のショートカットリンクのうち，定数本のみをランダムに選択することで，$O(\log^2 N)$ ステップの平均ルーティング時間が実現できることを証明する．定理の証明には，Kleinberg が提案したグリッドグラフのためのショートカット手法と同様のアイデアが用いられる．実際，本稿で提案するオーバーレイネットワークは，スキップグラフのショートカット選択に対して Kleinberg の手法を近似的に応用したものとなっている．

9. 田頭 茂明, 忽那 寛, 藤田 聡:
"Minority Game を利用した輻輳制御手法の評価", 情報処理学会システム評価研究会, Vol.2006 No.66, 1–6, 2006 年 6 月.
概要: 本論文では，Minority Game というゲーム理論を転送速度の調節に利用する新しい輻輳制御手法の提案をおこなう．Minority Game は，複雑系などの分野で注目されているゲームのひとつであり，利己的に振舞う参加者間で協調的行動が発現するという特性をもつ．したがって輻輳制御に Minority Game を利用することで，利己的に振舞う送信者間で協調的行動が発現し，効率的なパケット通信が実現されるのではないかと期待される．シミュレーションの結果，提案手法は，従来手法に比べて，公平で高効率な帯域の利用を可能にすることが明らかとなった．

10. 田頭 茂明, 齋藤 美幸, 藤田 聡:
"パッシブ型 RFID を用いた方向検出システム", 電子情報通信学会モバイルマルチメディア通信研究会, Vol.106, No.161, 25–30, 2006 年 7 月.
概要: 本稿では，リーダの読み取り方向を検出できるパッシブ型 RFID を用いた方向検出システムを提案する．方向の検出には，リーダとタグとの角度が通信範囲に影響を与える

という電磁誘導方式の特性を利用している．具体的には，この電磁誘導方式の特性を利用して，リーダの読めるタグが識別可能な各方向で異なるように複数枚のタグを配置する．方向の検出時には，リーダが読むことができたタグから，リーダの読み取り方向を推定する．提案する方向検出システムのプロトタイプシステムを周波数 13.56MHz 帯のパッシブ型 RFID を用いて実際に作成し，提案方式の性能評価を行った．実験の結果から，提案システムを用いて 8 方向の検出を 9cm 以内で安定してできることがわかった．

11. 徳永 達也, 田頭 茂明, 藤田 聡:
“DART を用いたアドホックネットワークのアドレス空間均等化手法”, 情報処理学会システムソフトウェアとオペレーティング・システム研究会, Vol.2006, No.86, 17–24, 2006 年 7 月.
概要: 本稿では, 大規模アドホックネットワーク向けに設計されたルーティングプロトコル DART（Dynamic Address RouTing）において，アドレス空間の均等化手法を提案する．DART は，動的な階層型アドレッシング手法とシンプルなルーティング手法を提供することにより，大規模ネットワークに対しても効率の良いルーティングを実現している．しかしながら，各ノードがアドレス空間を自律的に割り当てるために，ノード間で割り当てられるアドレス空間に偏りが発生する問題がある．そこで，アドレス空間の偏りを判断し，アドレスの再割り当てをおこなうことでアドレス空間を均等化する手法を提案する．提案手法をシミュレーションにより評価した．その結果，均等化の効果により，提案手法では従来手法よりも経路長を約 10%短くでき，また，経路表のサイズも減少することができた．

12. 圓山 康太, 田頭 茂明, 藤田 聡:
“PeerCQ における効率的な条件付き情報監視手法”, 情報処理学会システムソフトウェアとオペレーティング・システム研究会, Vol.2006, No.86, 1–8, 2006 年 7 月.
概要: 本稿では，PeerCQ に基づいた新しい P2P 型の情報監視システムを提案する．提案システムでは，AND 条件をサポートする情報監視に着目し，そのような情報監視を効率的に実現する枠組みを提供する．具体的には，AND 条件をサポートするために，PeerCQ に以下の拡張を行った．1) 与えられた AND 条件の成立を確認するエージェントとして，Conjunctive Continual Query(CCQ) を導入し，その詳細な仕様を提供した．2) 監視するデータ項目の類似性を考慮した CCQ のビアへの割り当て手法を提案した，3) 各ビアの負荷とシステム全体の負荷との間のトレードオフを制御できるように，PeerCQ において使われるユーティリティ関数を改良したい提案手法の効果をシミュレーションにより評価した．シミュレーションの結果，提案するユーティリティ関数によりシステムの性能は向上し，性能のトレードオフはパラメータを調整することにより制御できることを示した．

13. XiaoRan Wang, 田頭 茂明, 藤田 聡:
“ノードの Free Degree に基づいた MANET のための適応的ルーティングプロトコルの評価”, 情報処理学会システム評価研究会, Vol.2006 No.89, 9–14, 2006 年 8 月.
概要: 本稿では，リアクティブ型のプロトコルである Dynamic Source Routing (DSR) に着目し，ノードの”Free Degree”(FD) の概念に基づいた Free Degree Adaptive Routing (FDAR) プロトコルを提案する．FDAR は，ネットワークの負荷を分散させて，パケットを低遅延で転送するために，ルーティングメトリックとして FD を採用し，混雑している経路を迂回するようにパケットをルーティングする．また，ネットワークトポロジの変更時に必要な経路変更フェーズの過剰な実行を避けるために，寛容な経路変更ポリシーを確立した．FDAR の有効性をシミュレーションにより評価した．評価の結果，FDAR はパケットの転送遅延およびスループットにおいて，DSR より良い性能を示すことを確認した．

14. 田頭 茂明, 小畠 功士, 藤田 聡:
“分散ハッシュテーブル型 P2P システムにお

けるブルームフィルタを用いた高速連言検索手法の評価”, 情報処理学会システム評価研究会, Vol.2006 No.89, 15–22, 2006 年 8 月.
概要: 分散ハッシュテーブル型 P2P システム (P2P DHT) は，分散ネットワーク上に存在する共有資源の管理とキーワード検索とをピア・ツー・ピア (P2P) で効率よく実現するための代表的な手法の一つである．本稿では，P2P DHT 上で AND 検索を効率よくサポートするための新しい手法の提案と評価を行う．提案手法では，ブルームフィルタ (Bloom filter) を用いて，各ピアがローカルにキャッシュする過去の検索結果の情報を伝搬し，ピア間でそれらのキャッシュを効果的に共有する．提案手法の有効性を，シミュレーションにより実験的に評価した．実験の結果，システムに対する問合せ回数で約 66%改良できることを確認した．

15. 山内 亮, 田頭 茂明, 藤田 聡:
“ベンディング法を用いたセンサネットワークのためのナビゲーション手法”, 電子情報通信学会情報ネットワーク研究会, Vol.106, No.420, 79–84, 2006 年 12 月.
概要: センサネットワークにおいて，センサが危険を検知した場合，ユーザをそれらの危険から回避させつつ安全な目的地へできるだけ早く導くことができるようなパスを発見するナビゲーションシステムを提案する．これまでに提案されているおもな経路発見アルゴリズムは，折り返し探索法と網羅的な探索法とに大別される．しかしいずれの手法も，危険領域の大きさや形状によっては，得られるパスの長さや通信コストの点で，その性能が大きく劣化する可能性がある．本稿では折り返し法と網羅探索法の長所を併せ持った新しい経路探索アルゴリズムを提案する．提案手法の基本アイデアは，センサで覆われた領域内に複数のチェックポイントを系統的に配置し，それぞれのチェックポイントを通過する経路を並行して探索することである．提案手法の効果はシミュレーションにより実験的に評価される．実験の結果，提案手法は，非常に少ない通信コストで，網羅探索に匹敵するよさのパスを発見することがわかった．

16. 土井 賢治, 田頭 茂明, 藤田 聡:
“ノードの近接性を考慮した P2P DHT の構築手法”, 電子情報通信学会情報ネットワーク研究会, Vol.106, No.420, 139–144, 2006 年 12 月.
概要: 分散ハッシュテーブル (Distributed Hash Table: DHT) は，P2P システムにおいてネットワーク上に遍在する資源に対して効率的なアクセスを実現する手法のひとつとして高い注目を集めている．しかし，DHT には P2P オーバーレイネットワーク上のピア間の隣接関係が物理ネットワークにおけるノードの近接性には無関係に決められているという問題がある．本稿では，ピアが DHT に参加する際の手続きにおいて物理ネットワークにおけるノード間の近接性の全体的な関係を考慮できる手法を用いることで，ノードの近接性を考慮した P2P DHT の構築手法を提案する．提案手法の有効性はシミュレーションにより評価される．

17. 田頭 茂明, 安田 健一, 藤田 聡:
“グループの動的形成に基づく無線 LAN のための適応的メディアアクセス制御手法”, 電子情報通信学会モバイルマルチメディア通信研究会, Vol.106, No.498, 25–30, 2007 年 1 月.
概要: 無線 LAN においてメディアアクセス制御機構 (MAC) を用いることで，単一の無線通信チャネルを複数の端末間で共有することができる．しかしながら，既存の MAC 手法では，フレーム衝突の発生や不必要な通信帯域の割り当てが問題となり，通信帯域を有効に利用できない状況がある．本稿では，フレームの衝突確率を抑えながら，通信帯域を効率的に利用する無線 LAN のための新しい MAC 手法を提案する．基本的なアイデアは，事前に端末に割り当てた識別子からグループを動的に形成し，各グループに対してチャネルを割り当てていくことである．提案手法は，パラメータを適切に調節することにより，通信環境に応じたチャネル割り当てを行うことができる．シミュレーションの結果から提案手法は安定した性能を達成できることを確認できた．

18. A. Boulenouar, A. Regidor, S. Fujita:
"On Some Computational Aspect of Point Configurations in the Enclidean Space", 情報処理学会アルゴリズム研究会, Vol.2007 No.5, 63–70, 2007 年 1 月.
概要: Given a finite set $S \subset R^m$ in general position and a point $p \in R^m$, we devise an algorithm that checks if $S \cup \{p\}$ is also in general position, and if it is not, given a real number $\epsilon > 0$, finds another point p' within ϵ-distance from p so that $S \cup \{p_1\}$ is in general position. The worst-case running time of our algorithm is $O(n^{m-1} \log n)$, where n is the number of points in S, hence extending [l]'s result in dimension 2 and 3 to higher dimensions

19. W. Liu, S. Tagashira, S. Fujita:
"SwRED: A Robust Active Queue Management Scheme based on Load Level Prediction", 電子情報通信学会情報ネットワーク研究会, Vol.106, No.578, 267–272, 2007 年 3 月.
概要: Active queue management (AQM) is a novel technique to realize a smooth congestion control in communication networks, which is expected to work effectively by combining it with end-to-end congestion control schemes such as the window control adopted in TCP. Although it has been proposed a lot of AQM schemes in the literature, an appropriate tuning of parameters used in those schemes is recoginzed as a challenging problem to be overcome. In this paper, we proposed a new AQM scheme based on a load level prediction of bottleneck links, and on a combination of queue-based and rate-based AQM schemes in multiple time scales. Different from most of conventional AQM schemes, the propoesed scheme, called Switch-RED (SwRED), employs a new packet drop strategy based on a macroscopic analysis of the queue dynamics. The performance of the resultant scheme is evaluated by simulation. The result of simulations indicates that SwRED realizes a more stable queue behavior compared with previous schemes, i.e., it realizes a low queuing delay and a low packet loss rate while keeping a high link utilization.

20. 杉本 あい, 田頭 茂明, 藤田 聡:
"BitTorrent 型並列ダウンロードシステムにおける効率的なマッチメーキング手法の提案", 電子情報通信学会情報ネットワーク研究会, Vol.106, No.578, 167–172, 2007 年 3 月.
概要: 本稿では, P2P ネットワークにおいて BitTorrent に基づいた並列ダウンロード手法を提案する. 提案手法では, 従来手法と同程度のダウンロード性能を維持しながらピース情報の交換コストを削減することを試みる. 提案手法の基本的なアイデアは, 従来手法におけるピース選択アルゴリズムをローカルなピース情報だけを用いて実現するように改良することである. また, 改良によって起こりうるダウンロード時間の増加は, トラッカーで実行されるマッチメーキング手法を新たに導入することで抑制する.

21. 藤田 聡:
"定数ラウンドの適応的分散故障診断手法について", 第 20 回　回路とシステム軽井沢ワークショップ, 715–720, 2007 年 4 月.
概要: システムを構成する各基本ユニットの故障診断をシステム内で自律的に行うための基本モデルとしては、1967 年に Preparata, Metze, Chien の 3 名によって提案された PMC モデルが最も広く知られている。このモデルの基本アイデアは、システムを構成している各ユニットの状態をシステム外部から (直接的に) 検査するのではなく、各ユニットがシステム内部でお互いを検査しあい、それらの検査結果のみを外部から読み出すことで、どのユニットが故障しているのかを (間接的に) 診断するというものである。本稿では、PMC モデルに関する研究の中でも特に、Nakajima によって提案された適応的な故障診断手法の概説を行った。具体的には、適応的なモデルのもとで故障ユニットを正しく特

定するために必要なラウンド数の上下限について、これまでにわかっているおもな結果とそこで用いられている基本的なアイデアとをできるだけわかりやすく俯瞰している。

学会大会等

1. 伊達和宏，中野浩嗣:
“トラフィック分散を考慮したアドホックルーティングプロトコル”, 情報科学ワークショップ論文集, 10–11, 2005 年 9 月.

2. 山地敦大，中野浩嗣:
“モバイルアドホックネットワークにおける Localized Power Aware Routing”, 情報科学ワークショップ論文集, 12–13, 2005 年 9 月.

3. 山本亮平，中野浩嗣:
“指向性アンテナを用いたルーティングプロトコル”, 情報科学ワークショップ論文集, 14–15, 2005 年 9 月.

4. 松野博之，中野浩嗣:
“動的なアドホックネットワークにおけるクラスタリングアルゴリズム”, 情報科学ワークショップ論文集, 16–17, 2005 年 9 月.

文部科学省科学研究費補助金 特定領域研究 (課題番号 16092220)

A09: 非明示的表現に対するアルゴリズムの開発

本研究は，非明示的に表現されたオブジェクトに対するアルゴリズムの開発とその計算量の解析を行うことを目的し，また同時に，効率のよいアルゴリズムの開発に適した非明示的な表現法を文字列やグラフなどの離散構造を対象として研究を展開した．圧縮されたデータを陽に展開することなくそのまま照合を行う圧縮パターン照合問題を文脈自由文法の部分クラスや順序付き二分決定ダイアグラムとして表現された文字列を対象として，効率のよいアルゴリズムを開発した．また１次元，２次元に散らばった点の集合に対する近似的な照合を行うアルゴリズムや，最長共通部分列を漸増的に計算する手法，文字列の効率のよい索引構造，木パターンの照合アルゴリズムなどの開発，並びに文字列の繰り返し構造の解明を行った．さらに，XML テキストをパストライとバイナリ XML データの対に変換することによって，検索の領域効率と実行速度を向上させることに成功した．

研究組織

研究代表者：	篠原 歩	東北大学 大学院 情報科学研究科
研究分担者：	竹田 正幸	九州大学 大学院 システム情報科学研究院
	下薗 真一	九州工業大学 情報工学部
	坂本 比呂志	九州工業大学 情報工学部
	石野 明	東北大学 大学院 情報科学研究科 (平成 18–19 年度)

交付決定額 (配分額)

平成 16 年度	3,500,000 円
平成 17 年度	3,500,000 円
平成 18 年度	3,500,000 円
平成 19 年度	3,600,000 円
合　計	14,100,000 円

研究成果の概要

- 学術誌
  TCS(2005), JDA (2005), IJFCS (2005), Trans. on Computation on Biology and Bioinformatics (2004), DBSJ Letters (2005), 人工知能学会論文誌 (2006) など
- 国際会議
  CPM (Istanbul, 2004 年, Barcelona 2006 年), PSC (Prague, 2004 年, Prague, 2006 年), WABI (Begen, 2004 年) DS (Padova, 2004 年, Singapore, 2005 年, Barcelona, 2006, Sendai, 2007), ALT (Padova, 2004 年), SPIRE (Padova, 2004 年, Buenos Aires, 2005 年, Glasgow, 2005 年), AIRS (Beijing, 2004 年), DLT (Auckland, 2004 年), SOFSEM (Liptovsky Jan, 2005 年 Nový Smokovec, 2008 年), FCT (Luebeck, 2005 年), WEA (Santorini Island, 2005), ISAAC (Kalkata, 2006), DCC(Salt Lake City, 2007), RoboCup (Bremen, 2006), など

1 はじめに

アルゴリズムの効率のよさや問題の計算量は，入力長に対する関数として評価される．したがって，対象とする問題をどのように表現するかによって，その計算量は大きく変わることがある．例えば，単純な文字列照合問題を例にとってみよう．よく知られているように，Knuth-Morris-Pratt アルゴリズムを用いることによって，入力として与えられたテキスト長 N とパターン長 M に対して，$O(N+M)$ 時間でこのパターンがテキスト中に含まれている位置の検出を行うことができる．しかしながら，もしもこのテキストが圧縮された形式で与えられる場合には，状況が一変する．圧縮されたテキストの表現長 n がもとのテキストの大きさ N と比較して指数的に小さい場合，いったんこの圧縮テキストを展開するだけでも入力長 n に対する指数時間・領域を要することになり，展開後にいかに効率のよい方法を用いたとしても多項式時間アルゴリズムを構築することはできない．さらに，パターンも圧縮された形式で与えられる場合には，より洗練された手法を開発しなければ，入力長に対して多項式時間・領域で問題を解決することができない．すなわち，新たな評価基準のもとでの効率のよいアルゴリズムが求められることになる．

本研究では，このような着想を一般化し，計算の対象が明示的に与えられるのではなく，それを生成する文法や論理式，数式などを用いて非明示的に表現されて与えられたときに，それを効率よく処理するアルゴリズムの開発と計算量の解析を行うことを目的とするものである．同時に，効率のよいアルゴリズムの開発に適した非明示的な表現法そのものについて，文字列やグラフなどの離散構造に力点をおいて研究を展開した．

次章で，研究の概要を各項目について説明した後，いくつかの話題についてより詳細な内容を記述する．

2 研究の概要

本章では，得られた研究成果を内容に応じて大別し，簡潔に述べる．

2.1 データ圧縮とパターン照合

グラフの非明示的な表現としていくつかの応用例が知られている順序付き二分決定ダイアグラム (OBDD) に着目し，文字列を OBDD で表現する方法を提案した．これを文法の形で定式化しなおし，こうして非明示的に表されたパターン文字列とテキスト文字列が与えられたときに，それを陽に展開せずにパターン照合を行うアルゴリズムを開発した．また，コラージュシステムと呼ばれる文法型式で与えられた文字列に関しても，その部分族に関して既存のものよりも高速に照合を行えるアルゴリズムを与えた．さらに，これを対応する文脈自由文法の部分クラスにおいても同様の解析を行った．逆に，与えられた文字列からそれを表す小さなサイズの文法を見つける問題はデータ圧縮の一種とみなせるが，我々は領域効率のよい線形時間の圧縮アルゴリズムを開発することに成功した．

また，圧縮されデータを陽に展開することなく，そのままパターン照合を行う，圧縮文字列照合の研究に関して，形式文法に基づく圧縮法に対する効率のよいアルゴリズムを与えた．さらに，文字列中に共通に現れる共通部分文字列を新たな記号に置き換える作業を繰り返す事による圧縮法に関して，線形時間の圧縮アルゴリズムを与えた．

2.2 平滑化解析

最も基本的な文字列照合アルゴリズムである，素朴なアルゴリズムについて，その平滑化解析を行った．文字列照合は文字列処理のなかでも最も基本的な問題の一つである．照合を行うテキスト T の長さが n，パターン P の長さが m であるとき，文字列照合アルゴリズムの中でも最も基本的なアルゴリズムである素朴なアルゴリズムに於いて，最悪計算時間は $O(mn)$ であることがよく知られている．しかしながら，テキストやパターンが一様分布にしたがって完全にランダム選ばれたと仮定すると，このアルゴリズムの平均計算時間は，パターン長と無関係に $O(n)$ となる事も同時に知られている．すなわち，最悪時の計算時間と，平均時の計算量の間に大きなギャップが存在する．そこで，本研究において我々は，Spielman と Teng [6] によって提唱された，平滑化解析 (smoothed analysis) を

文字列処理アルゴリズムの評価に適応させ，そのギャップを埋める解析を行った．

2.3 近似照合と類似性の計算

１次元または２次元に散らばった点の集合に対する近似的な照合を行うアルゴリズムを開発した．

一方，文字列照合を行う非決定性の有限オートマトンの状態遷移をビットベクトルで表し，ビット演算によって高速な照合を行うビット並列の手法を用いて，日本語のような複数バイト文字列に対してもそのまま効率よく働くアルゴリズムの与えた．また，挿入と削除のみを許し，置換を許さない状況下での近似文字列照合に関して，既存のものよりも高速なビット並列アルゴリズムも示し，これらのアルゴリズムの有効性を計算機実験によって実証した．

さらに，文字列の類似性を測るためによく用いられている，最長共通部分列の漸増的計算方法について議論した．有限アルファベット Σ 上の文字列 $A = a_1a_2 \ldots a_n$ に対し，$|A|$ はその長さ n を表し，$A[i] = a_i$ とする．列 $A[i_1]A[i_2] \ldots A[i_k]$ は，$1 \leq i_1 < i_2 < \ldots < i_k \leq n$ のとき，A の長さ k の部分列という．文字列 A と B に共通して現れる部分列の中で，長さが最長のものを最長共通部分列 *(Longest Common Subsequence)* とよび，その長さを $LCS(A, B)$ と表す．$LCS(A, B)$ は，動的計画法を用いて $O(nm)$ 時間で求められることが知られている [1]．ここに n と m はそれぞれ A と B の長さを表す．我々は，$LCS(A, B)$ が既に求まっているときに，A や B の前または後に文字 c が追加された文字列に対して，$LCS(cA, B)$, $LCS(A, cB)$, $LCS(Ac, B)$, $LCS(A, Bc)$ を効率よく求める問題を考える．$LCS(A, B)$ が計算済みであるとき，$LCS(cA, B)$ と $LCS(A, cB)$ はそれぞれ $O(L)$ 時間と $O(n)$ 時間で計算できる．$LCS(Ac, B)$ と $LCS(A, Bc)$ は，それぞれ $O(L)$ 時間と $O(n)$ 時間で容易に計算できる．ここに $L = LCS(Ac, B)$ である．

2.4 XML 検索

XML テキストの高速処理に関する技法に関する技法についての研究を行った．XML は，あらゆるデータをタグ付きのテキスト文字列として行うことによって　可搬性と機械可読性を備えた汎用フォーマットとして，広範囲に活用されている．本研究では，与えられた XML テキストを，パストライとバイナリ XML データの対に変換することによって，領域効率と実行速度を向上させることに成功した．

また XQuery に関して，文字列検索技術を応用した単純で高速な処理手法を提案した．XML データから Aho-Corasick 法を拡張した XML データ走査器によって高速にタグやキーワードの検出を行い，パスに応じた処理を行うことで XQuery の質問式を処理した．特にパスの照合を高速に処理するために，XML データから得られるパススキーマを用いることで決定性有限オートマトンを構築するパスプルーニングを提案し，その効果を計算機実験によって検証した．さらに，半構造データに対する高速な XPath 処理法を提案した．本研究では，算術符号化を半構造データの検索に応用した逆算術符号化にノード間の先祖子孫関係を判定可能な範囲ラベルを導入することにより，より複雑な問い合わせ処理を高速に実現するための手法を提案した．評価実験の結果，300MB 程度の XML データに対してテキストを直接処理する既存の手法と比較し，数十から百倍の高速化を達成した．

2.5 索引構造

文字列に対する索引構造に関して，単語単位の接尾辞木 (suffix tree) や有向文字列グラフ (DAWG), 圧縮有向文字列グラフ (CDAWG) をそれぞれ線形時間で構築する，簡潔なアルゴリズムを開発した．

2.6 文字列の組み合わせ構造

また，種々の文字列処理アルゴリズムに対して，最悪実行時間の解析に活用されることの多いフィボナッチ文字列について，その繰り返し構造に関する組み合わせ的性質を解析した．

2.7 最適パターンの発見

文字列とそれに関連付けられた数値属性値との対の集合から最適パターンを発見する問題を扱った．パターンの「良さ」として，個々の文字列におけるパターンの生起回数とその文字列に関連付けられた数値属性値との間の相関を用い，接尾辞木に基づく二つのアルゴリズムを提示した．これらのアルゴリズムは，最適部分文字列パターンをそれぞれ $O(Nn)$ 時間と $O(N^2)$ 時間で求めることができる．ここに，n は文字列の個数であり，N は文字列長の総和である．また，より複雑なパターン族を扱う際に用いることのできる一般的な分枝限定手法をを示した．さらに，上述の $O(N^2)$ 時間アルゴリズムと分枝限定法を組み合わせることにより，著しく効率が良くなることを示した．

2.8 木パターンの照合

木パターン照合問題とは，二つのラベル付木 P と T（それぞれパターン木、ターゲット木とよぶ）が与えられて，T における P のすべての生起を見つける問題である．この問題に対して，順序木と無順序木の両方の場合について多くの研究がなされている．柔軟な照合を行うために，一種の VLDC が導入されている．XML 文書の部分を指し示すための言語である XPath においては，経路 VLDC が用いられている．本研究では，水平方向の VLDC を導入した．この VLDC は，順序木において連続した兄弟を根にもつ部分木の系列にマッチする．我々は，パターンが水平方向の VLDC をもつ場合の木パターン照合問題を扱った．この設定では，ターゲット木は，XML データストリームのように，タグを含む文字列として与えられる．この問題を $O(mn)$ 時間，$O(mh)$ 領域で解くアルゴリズムを提案した．ここに，m と n は，それぞれ，P と T の長さであり，h は T の高さである．さらに，ビット並列化技法を用いることにより、実用的に高速なアルゴリズムを得ることができた．

また，畳み込みカーネルのアイディアに基づいた，ラベル付き順序木に対する新しいカーネル関数を提案した．畳み込みカーネルの枠組みにおいてラベル付き順序木に対して任意の部分グラフを部分構造として用いた場合の，効率の良いカーネル計算のアルゴリズムを提案し，曖昧なラベルや構造を取り込むような拡張を行い，より一般的な木構造として，順序のないラベル付き根付き木に対するカーネルを考えた場合には，カーネルの計算が#P-完全問題であることを示した．

2.9 機械学習とロボティクスへの応用

自律ロボットによるサッカー競技である RoboCup サッカーを題材に，自律ロボットの学習に焦点をあてた研究を行った．RoboCup サッカーは人工知能とロボット工学の融合を目指して，提唱された国際プロジェクトであり，我々はそのなかの四足ロボットリーグに 2003 年から参加してきた．四足ロボットリーグは，すべてチームが同じロボットを使用する，他のリーグにはない大きな特徴を持っている． そのため，ハードウェアの優劣によらず純粋にソフトウェアの優劣を競うことができ，開発したアルゴリズムが実時間で期待通りのパフォーマンスを示すことができるかどうかを確かめるよいテストの場となる．我々は，ロボットによるサッカーの重要な要素技術の一つとして，動いてくるボールをよいタイミングでとめる，ボール補足の技術を強化学習を用いた自律学習によって，実際のロボットで行えることを示し，それを実戦にも用いて効果を確認した．

3 データ圧縮とパターン照合

テキスト文字列の中から目的とするパターン文字列の出現を検出する文字列パターン照合問題は，情報検索における最も基本的な問題として古くから盛んに研究され，Knuth-Moriss-Pratt 法や Boyer-Moore 法をはじめとして多くの効率的なパターン照合アルゴリズムが提案されてきた．しかし，一般にテキストはディスクなどの二次記憶装置に格納されるため，処理時間の大半がテキストのデータ転送に費やされることになる．したがって，パターン照合アルゴリズム自体の改良によって処理時間を短縮することには限界ある．

そこで，テキストをあらかじめ圧縮しておくことによってデータの転送量を減らして高速化を測る方式も提案されてきた．その際，圧縮テキスト

を復号した上でパターン照合を行ったのでは復号する手間がかかり高速化の妨げとなるため，圧縮テキストを陽に復号せずにそのままパターン照合処理を行おうという，**圧縮パターン照合**の技法に注目が集まっている．すなわち，入力となるテキストが，文字列そのものではなく，他の構造であり，したがって，その入力長そのものの測り方が変わってくることになり，その様相は一変する．また場合によってはテキストのみならず，パターンも他の構造で圧縮表現されたものを考える場合もある．本研究では，以下に示すとおり，文字列や点列を対象とした圧縮パターン照合アルゴリズムや，圧縮アルゴリズムの高速化を行った．

3.1 文法を用いた文字列圧縮

データ圧縮法の性能は，圧縮率，圧縮/展開時間，および主記憶量で決まるが，一般にこれらはトレードオフの関係にあり，すべてを高水準に保つことは難しい．例えばBW変換を利用した圧縮法は圧縮率こそ良いものの，原理的に線形時間での圧縮は不可能である．また，算術符号化も圧縮率はよいものの，時間と主記憶の両方を大量に消費する．本研究では，文字列を一意に表す文法を用いて，これらのリソースの消費を最小限に抑えつつ，理論的な限界に近い圧縮率を達成する高速な圧縮アルゴリズムを考案し，実装した．

さらに，以下のように定義される文法表現に対する圧縮文字列照合アルゴリズムを考案した．次のような変数代入の列によって表される文法表現 $\mathcal{T}$ を *MPM* とよぶ．

$$\begin{aligned} X_1 &= expr_1, \\ X_2 &= expr_2, \\ &\vdots \\ X_n &= expr_n \end{aligned}$$

ここにおのおのの X_i は変数であり，$expr_i$ は次のような形の式である：

- $expr_i = a$ $(a \in \Sigma)$, または
- $expr_i = X_\ell X_r$ $(\ell, r < i)$ ただし $|X_\ell| \geq |X_r|$ と $|X_\ell|$ は2のベキ数する．

MPM $\mathcal{T}$ は，$L(\mathcal{T}) = \{T\}$ となる，チョムスキー標準型の文脈自由文法の一種であるとみなすことができる．$\mathcal{T}$ のサイズは n であり，$\|\mathcal{T}\|$ と表す．例えば，文字列 $T =$ abacabbcabacc に対する MPM $\mathcal{T}$ は次のようになる．

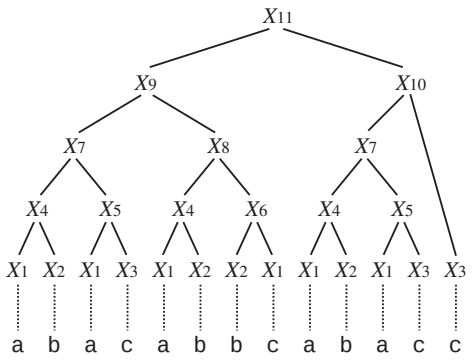


図 1: 文字列 abacabbcabacc に対する MPM の導出木.

$$\begin{aligned} X_1 &= \mathtt{a}, \\ X_2 &= \mathtt{b}, \\ X_3 &= \mathtt{c}, \\ X_4 &= X_1X_2, \\ X_5 &= X_1X_3, \\ X_6 &= X_2X_3, \\ X_7 &= X_4X_5, \\ X_8 &= X_4X_6, \\ X_9 &= X_7X_8, \\ X_{10} &= X_7X_3, \\ X_{11} &= X_9X_{10} \end{aligned}$$

図. 1にこの $\mathcal{T}$ に対する導出木を示す．

このとき，$\|\mathcal{T}\| = 11$ であり，その表す文字列の長さは13であることに注意されたい．一般に，$\mathcal{T}$ の表現長 $\|\mathcal{T}\|$ に対して，$\mathcal{T}$ が表す文字列 T の長さは指数的に長くなりうる．一方で，ここで考えている文法では,「$|X_\ell| \geq |X_r|$ と $|X_\ell|$ は2のベキ数する」という制約条件が付けられており，この制限によって構文木の形が大きく制限されることになる．

MPMに対するパターン照合問題とは，与えられた2つのMPM $\mathcal{T}$ と $\mathcal{P}$ に対し，それぞれを展開して得られる文字列 T と P をそれぞれテキストとパターンとみなして，T の中に P が現れる位置をすべて見つけ出す問題である．以下では，$\|\mathcal{T}\| = n$ および $\|\mathcal{P}\| = m$ とする．

上記の制約条件がない場合には，宮崎ら [5] による $O(m^2n^2)$ 時間 $O(mn)$ 領域のアルゴリズムが現在知られている中で最も高速な圧縮文字列照合アルゴリズムである．それに対し，我々は，こ

の制約を巧みに利用することによって計算時間を $O(mn^2)$ 時間に短縮することができた.

3.2 点列の圧縮パターン照合

通常，我々が文字列について考察する場合，有限アルファベット上のものを対象とする．一方，点列（数値列）は，文字列のアルファベットを無限に拡張したものとみなすことができる．しかしながら，その点列の表現長は，実際に点列が空間を占める長さに対して log オーダーとなりうるため，ランレングス・コーディングと同様な圧縮された表現ととらえることができる．そのため，パターン照合を行う際には通常の文字列検索とは異なる様相を呈する．我々は，その違いを克服して文字列の近似照合と同じオーダーで動作する点列のための近似照合アルゴリズムについて，現実的な解法を与えるという観点から議論した．また逆に，文字列に対する照合アルゴリズムの高速化として用いられるテクニックを適用するための条件と，応用方法についても議論した.

線形変換による**点列パタン照合問題**は, ユークリッド空間上の 2 つの点集合パタン P とテキスト T が与えられた時, P を T の部分集合に写像する線形変換を見つける問題である. 本研究ではまず, 主軸について昇順の 1 次元の点列に対して, 平行移動と点の間の空間の挿入 · 削除による編集距離とそれに基づく近似点列パタン照合問題を定義した．次に P のサイズ n, T のサイズ m について時間計算量 $O(nm^2)$ で動作する近似点列パタン照合問題のアルゴリズムを紹介し, 高速化のための改良を加えたアルゴリズムが有限解像度を持つ P, T に対して $O(nm)$ で動作することを証明した. 更に「4 人のロシア人の方法」として知られている手法を応用して $O(\frac{nm}{\log m} + m \log^2 m)$ 時間に高速化する技法を示した．そして, 多項式時間で解くことができる 2 次元近似点列パタン照合を提案した．これらの方法は，２次元の点列の１次元目の値を音の発音のタイミングとし，また２次元目の値を音高とすれば，音楽データに対する近似照合問題の定式化のひとつとみなすことができ，この方面への応用も見込まれる.

3.3 XML データの圧縮パターン照合

XML データを圧縮する試みは 2000 年頃から行われているが，初期の頃に提案された XMill や xmlppm などは，圧縮のみに焦点を置いた研究であり，データ検索を行うことはできなかった．その後，圧縮データを展開しながら検索を行うことができる XGrind などが提案されたが，データを展開することがボトルネックとなっており，データを大規模化することができなかった．こうした背景の中で，データを圧縮したまま検索するための手法として，XPress が提案された．この手法は高速ではあるが，いくつかの欠点を持っている．まず，この手法は算術符号化のある性質に依存しているため，ごく限定された形式の検索しかできない．また，元データの構造を一部破棄してしまうため，データの追加/削除が難しく，特に，データの部分的な更新は不可能である．このような理由から，これらの手法を実用的な半構造データベースに応用することは困難であった．そのため我々は，XPress の手法を改良して，データ検索の能力を向上させることに成功した．さらに，そのことによる検索速度の落ち込みも回避しているため，検索能力は XPress を上回っているといえる.

4 ビット並列化

ビット並列化と呼ばれる手法は，非決定性の有限オートマトンの状態をビットベクトルでコンパクトに表現し，AND や OR, NOT といったビット演算を駆使することによって状態遷移を高速に実現する手法である（図 2）．これまでにも通常の文字列照合のみならず，近似文字列照合や正規表現へのマッチングなどへの応用が知られている．特に，32 ビットや 64 ビットの CPU が普及し，また MMX や SSE といったベクトル命令によって一度に多くのビット演算を行うことが容易になるにつれ，より大きなサイズの有限オートマトンを表現でき，また複雑な操作が可能になることから，より注目が高まっている．しかしながら，これまでは，いずれも英文テキストや遺伝子配列などのようにアルファベットサイズが小さい場合にその応用が限られており，日本語テキストのように複数バイトの文字列に対しては適用が困難であった.

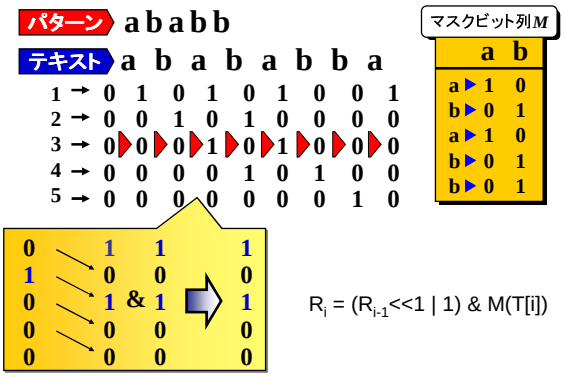


図 2: ビット並列化の手法．テキスト文字列 abababba に対してパターン文字列 ababb を検索する様子．

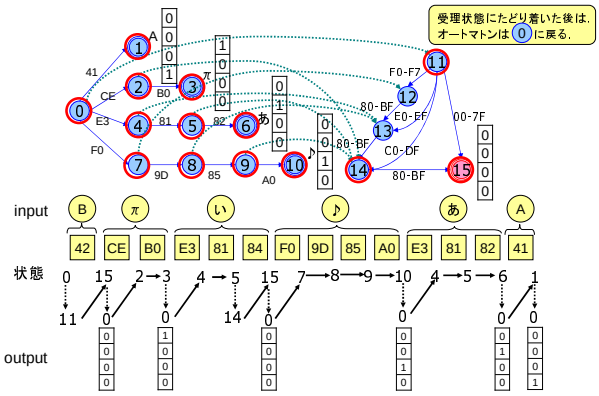


図 3: パターン集合 $P = \{$ π, あ, ♪,A$\}$ に対するコードオートマトン（UTF-8）．

これは，それぞれの文字に対応させたビットマスクを用意する必要があるためであり，そのまま日本語テキストに対して適用するには領域効率が悪すぎるのが問題になるからである．そこで本研究では，複数バイトのテキストをそのまま走査する有限オートマトンを考え，(図 3）直接その上で照合を行う方式を採用することによって，複数バイト文字列でもビット演算手法が有用であることを示した．

また，挿入と削除のみを許した枠組みでの近似文字列照合アルゴリズムについて，ビット演算を活用した新たなアルゴリズムの開発に成功した．

5 最長共通部分列の漸増計算

文字列 A と B に対し $(m+1)\times(n+1)$ の二次元配列 DP を次のように帰納的に定義する．

$$DP[i,j] = \begin{cases} 0 & i=0 \text{ or } j=0 \\ \max\{DP[i-1,j], DP[i,j-1]\} & A[j] \neq B[i] \\ DP[i-1,j-1]\} & A[j] = B[i] \end{cases}$$

このとき，$DP[m,n] = LCS(A,B)$ であり，これは $O(mn)$ 時間と領域で計算できる（図 4上）．

5.1 Landau らのアルゴリズム

Landau ら [3] は，$L = LCS(A,B)$ が計算済みであるとき，任意の文字 c に対して $LCS(cA,B)$ を効率よく計算するアルゴリズムを提案した．彼らのアルゴリズムは，DP 全体を再計算するのではなく，次のように定義される分割点の集合 P のみを計算するため，従来の手法より高速かつ省領域で動作する．$A[j] = B[i]$ のとき，対 (i,j) を A と B の間の一致点という．また，$DP[i,j] = DP[i-1,j]+1$ が成り立つとき，対 (i,j) を DP の分割点とよぶ．DP のすべての分割点からなる集合を P で表す．$DP[i,j] = v$ となる分割点 $(i,j) \in P$ に対し，$P[v,j] = i$ と定義する．すなわち，$P[v,j]$ は DP の第 j 列において，初めて v になる行を表す（図 4）．

Theorem 1 (Landau ら [3]) $L = LCS(A,B)$ が計算済みであるとき，任意の文字 c に対して，$LCS(cA,B)$ を $O(L)$ 時間で計算できる．

5.2 $LCS(A,cB)$ の計算

$LCS(A,B)$ が計算済みであるとし，そのときの DP 表を DP とし，分割点の集合を P とおく．B の先頭に文字 c が追加されたときの $LCS(A,cB)$ に対する DP 表を DP^{Bh} とし，分割点の集合を P^{Bh} とする．この変化に着目するために，$k = \min\{j \mid A[j] = c\}$ とする．k は DP^{Bh} の最初の行のなかの一致点を表す列である．

		a	d	b	d	c	d
b		0	0	**1**	1	1	1
c		0	0	1	1	**2**	2
b		0	0	**1**	1	2	2
b		0	**1**	1	**2**	2	**3**

	b	a	d	b	d	c	d
b	**1**	1	1	**1**	1	1	1
c	1	1	1	1	1	**2**	2
b	**1**	1	1	**2**	2	2	2
b	1	1	**2**	2	**3**	3	**3**

図 4: $A = \texttt{adbdcd}$, $B = \texttt{bcbd}$ のときの DP（上図）と，A の先頭に b を追加したときの DP^{Ah}（下図）．簡単のため，$i > 0$, $j > 0$ の部分のみ記載する．一致点を太字で表し，分割点を□で囲み，DP^{Ah} で新たに分割点になったところを灰色で塗っている．

Lemma 1 任意の列 $j \geq k$ に対し，次式を満たす行 E_j が存在する．

$$DP^{Bh}[i,j] = \begin{cases} DP[i,j]+1 & (i < E_j), \\ DP[i,j] & (i \geq E_j). \end{cases}$$

補題1より，P を P^{Bh} に更新する際，各列 $j \geq k$ で行 E_j のみに着目すればよいことがわかる．ここで (E_j, j) は P を P^{Bh} に更新する際に，列 j において削除される分割点である．次に示す補題2と補題3より，削除すべき分割点 (E_j, j) を効率よく求めることができる．

Lemma 2 分割点集合 P を P^{Bh} に更新する際に *j-1* 列および j 列で削除される分割点をそれぞれ $(E_{j-1}, j-1)$ および (E_j, j) とし，$DP^{Bh}[E_{j-1}, j-1] = v$ とする．このとき，$E_{j-1} \leq E_j \leq P^{Bh}[v+1, j-1]$ が成り立つ．(図 *5*参照)．

Lemma 3 分割点集合 P を P^{Bh} に更新する際に $j-1$ および j 列で削除される分割点をそれぞれ $(E_{j-1}, j-1)$ および (E_j, j) とし，$DP^{Bh}[E_{j-1}, j-1] = v$ とする．このとき，$P^{Bh}[v, j-1] < x \leq$

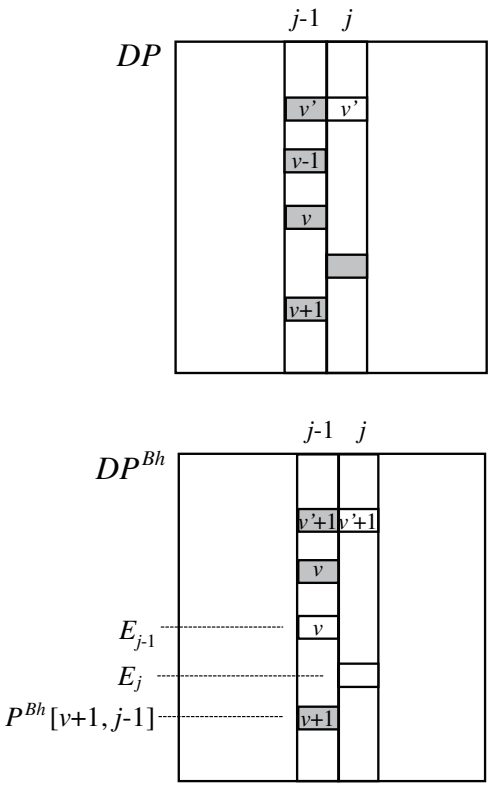


図 5: P を P^{Bh} に更新する際に第 j 列の分割点 E_j が存在する範囲．灰色で表した点が分割点を表す．

$Ej-1$ を満たす一致点 (x, j) が存在するときは，$E_j = E_{j-1}$, そうでなければ $E_j = P[v+1, j]$ が成り立つ．

Lemma 4 $k = \min\{j \mid A[j] = c\}$ とする．DP の j 列に分割点が存在するときは $E_k = null$, そうでなければ $E_k = P[1, k]$ である．

以上をまとめて，次の定理が得られる．

Theorem 2 $L = LCS(A, B)$ が計算済みであるとき，任意の文字 c に対して，$LCS(A, cB)$ を $O(n)$ 時間で計算できる．

表 1に本論文の提案手法と従来手法の計算時間の比較結果を示す．$LCS(A, B)$ が計算済みであるとき，$LCS(cA, B)$ と $LCS(A, cB)$ はそれぞれ $O(L)$ 時間と $O(n)$ 時間で計算できる．$LCS(Ac, B)$ と $LCS(A, Bc)$ は，それぞれ $O(L)$ 時間と $O(n)$ 時間で容易に計算できる．

		動的計画法	KP 法	提案手法
計算時間	$LCS(cA, B)$	$O(mn)$	$O(m+n)$	$O(L)$
	$LCS(Ac, B)$	$O(m)$	$O(m)$	$O(L)$
	$LCS(A, cB)$	$O(mn)$	$O(m+n)$	$O(n)$
	$LCS(A, Bc)$	$O(n)$	$O(n)$	$O(n)$
作業領域		$O(mn)$	$O(mn)$	$O(nL+m)$

表 1: 計算量の比較．KP 法は文献 [2] を応用したものである．$|A| = n$, $|B| = m$ であり，$L = LCS(A, B) \leq \min(n, m)$ が常に成り立つことに注意.

6 文字列照合アルゴリズムの平滑化解析

文字列照合は文字列処理のなかでも最も基本的な問題の一つである．照合を行うテキスト T の長さが n，パターン P の長さが m であるとき，文字列照合アルゴリズムの中でも最も基本的なアルゴリズムである素朴なアルゴリズム (naïve matching algorithm) に於いて，最悪計算時間は $O(mn)$ であることがよく知られている．しかしながら，テキストやパターンが一様分布にしたがって完全にランダム選ばれたと仮定すると，このアルゴリズムの平均計算時間は，パターン長と無関係に $O(n)$ となる事も同時に知られている．すなわち，最悪時の計算時間と，平均時の計算量の間に大きなギャップが存在する.（もちろん，最良時の計算時間も $O(n)$ であることは明かであるが，このことはあまり重要ではない.）そこで，本研究において我々は，Spielman と Teng [6] によって提唱された，平滑化解析 (smoothed analysis) を文字列処理アルゴリズムの評価に適応させ，そのギャップを埋める解析を行うことにした．平滑化解析は，アルゴリズムの最悪計算時間と平均時間の不一致を説明するための比較的新しい手法である．解析の方法としては，最悪時間となる入力を与えたり，完全にランダムな入力を与える代わりに，最悪時間を与える入力にランダムな部分摂動を行う．この解析法は，最悪計算時間と平均計算時間の間を補完する事ができるだけでなく，実際のアプリケーションなどの高速化のヒントにもなり得るという点でとても有用である．実際，Spielman と Teng らは，線形計画法を実用的によく用いられている，シンプレックス法（単体法）の平滑化解析を行い，その実用性を裏付ける理論的結果を得ている．シンプレックス法は，最適解が多面体の頂点に現れることを利用し，最適解に達するまで多面体の辺をたどってより高い目的関数の値を次々にたどることで線形計画問題を解く．このアルゴリズムは実際にかなり効率がよいことが知られており，巡回していないかに注意を払えば大域的最適解を見つけることが保証される．しかしながら，シンプレックス法は，最悪時には指数時間かかってしまう．一方，内点法や射影変換法など，最悪時でも多項式時間で動作することが保証されているアルゴリズムもあるが，実装のしやすさや，実際の問題に適用した際の実行時間の計測により，依然としてシンプレックス法はよく用いられている．Spielman と Teng は，新たに導入した摂動モデルを用いて，最悪時の入力に対しても，ほんの少しだけ摂動をかけることによって，実行時間が劇的に改善されることを示した．このことから，素朴な文字列照合アルゴリズムは，最悪時の実行時間は $O(nm)$ 時間であるものの，実質的にそのような「悪い」入力となるインスタンスは，問題空間の中で「孤立」していることが明かとなった．すなわち，実質的にはそのような例はほとんど発生しないため，素朴な文字列照合アルゴリズムは，実用的にも十分有用であることの理論的な裏付けが得られたことになる．

$\sigma = |\Sigma|$ とし，文字列 $w \in \Sigma^*$ と二つの整数 i, j 対して，$w[i, j]$ を w の $w[i]$ から $w[j]$ までの**部分文字列**とする．**文字列照合問題**は，長さ n のテキスト文字列 T と長さ m のパターン文字列 P が入力として与えられたとき，T の中に P が出現する全ての位置を求める問題である．

素朴な文字列照合アルゴリズムは 2 重構造のループからなるとても単純なアルゴリズムで，外側のループでテキストの全ての位置を訪れ，内側のループでそこから始まる部分文字列とパターンが等しいかを一文字ずつ検証する．不一致が見つかった時点で内側のループは終了する．よく知られているように，このアルゴリズムの計算時間が最悪となるのは $T = aaaa \cdots aaaa$, $P = aa \cdots aab$ のような入力に対してである．このとき，$O(mn)$ であるが，一様ランダムな文字列に対しては，平均計算時間は $O(n)$ である．

文字列 $w \in \Sigma^*$ の各位置において，確率 p で Σ から一様ランダムに選ばれた文字に置換する，**部**

分摂動を行う．$p=0$ の時，変異は起こらず w そのままになり，$p=1$ の時は w とは関係のない完全なランダム文字列になる．$|w|$ と $|w'|$ のハミング距離の期待値は $|w|p(1-\frac{1}{\sigma})$ となる．

$\mu(i,j)$ を，テキスト T の i 番目の文字 $T[i]$ と，パターン P の j 番目の文字 $P[j]$ が一致する確率と定義する．また，l を $1 \le l \le |P|$ である整数，i の範囲をそれぞれ $1 \le i \le |T|$ とする．この時，長さ $l-1$ のテキスト T の部分文字列 $T[1,l-2]$ にパターン P のプレフィックス $P[1,l-1]$ が一致し，かつテキスト中の文字 $T[i+l-1]$ とパターン中の文字 $P[l]$ が一致しない確率 $P_F(i,l)$ は次式で表される．

$$\begin{aligned} &P_F(i,l) = \\ &\quad (\mu(i,1)\cdot\mu(i+1,2)\cdots\mu(i+l-2,l-1)) \\ &\quad \cdot(1-\mu(i+l-1,l)) \end{aligned} \tag{6.1}$$

また，パターン P がテキスト T の i 番目からの部分文字列と完璧に一致するときの確率 $P_S(i,l)$ は，次式で表される．

$$\begin{aligned} &P_S(i,l) = \\ &\quad \alpha(i,1)\cdot\alpha(i+1,2)\cdots\alpha(i+l-2,l-1) \\ &\quad \cdot\alpha(i+l-1,l) \end{aligned} \tag{6.2}$$

テキストとパターンが一様ランダムに選ばれたとき，素朴な文字列照合アルゴリズムの総計算時間の期待値 $Ex(T,P)$ は次式で表される．

$$\begin{aligned} &\mathrm{Ex}(T,P) = \\ &\quad \sum_{i=1}^{n}\left(\sum_{l=1}^{m}(P_F(i,l)\cdot l) + P_S(i,m)\cdot m\right) \end{aligned} \tag{6.3}$$

特に，$\Sigma=\{\mathtt{a},\mathtt{b}\}$ の場合，$P_F(i,l)=P_S(i,l)=\frac{1}{2^l}$ より式 (6.3) は以下のようになる．

$$\mathrm{Ex}(T,P) = \sum_{i=1}^{n}\left(\sum_{l=1}^{m}\frac{l}{2^l} + \frac{m}{2^m}\right) \tag{6.4}$$

ここで，$\sum_{l=1}^{m}\frac{l}{2^l}$ について，

$$a_m = \sum_{l=1}^{m}\left(\frac{l}{2^l}\right) = \frac{1}{2}+\frac{2}{2^2}+\frac{3}{2^3}+\cdots+\frac{m}{2^m} \tag{6.5}$$

とする．(6.5) の両辺に $\frac{1}{2}$ をかけると，

$$\frac{1}{2}a_m = \frac{1}{2^2}+\frac{2}{2^3}+\cdots+\frac{m-1}{2^m}+\frac{m}{2^{m+1}} \tag{6.6}$$

となる．ここで (6.5) の両辺から (6.6) の両辺の差をとると，

$$\frac{1}{2}a_m = \frac{1}{2}+\frac{1}{2^2}+\frac{1}{2^3}+\cdots+\frac{1}{2^m}-\frac{m}{2^{m+1}}$$

となるので，a_m は

$$\begin{aligned} a_m &= 1+\frac{1}{2}+\frac{1}{2^2}+\cdots+\frac{1}{2^{m-1}}-\frac{m}{2^m} \\ &= \frac{1-\frac{1}{2}^m}{1-\frac{1}{2}}-\frac{m}{2^m} \\ &= 2\cdot\left(1-\frac{1}{2^m}\right)-\frac{m}{2^m} \end{aligned}$$

と表せる．a_m を (6.4) に代入すると，

$$\mathrm{Ex}(T,P) = 2n\cdot\left(1-\frac{1}{2^m}\right) \tag{6.7}$$

となる．

6.1 最悪時の入力に部分摂動を行った場合の期待値

p を T と P の各文字に対して置換が起こる確率とする．$0 \le p \le 1$ である．特に，$p=0$ の時は入力文字列に変化はなく，$p=1$ の時は入力に関係ない完全にランダムな文字列になる．ここで，$\Sigma=\{a_1,\cdots,a_\sigma\}$, $|\Sigma|=\sigma$ であるとき，Σ に含まれる a_1 のみからなる文字列 w があるとする．この文字列に含まれる各文字が，p の確率で Σ 内の文字のどれかに置き換わった文字列を w' とする．この時，a_1 のみの元の文字列と w' とのハミング距離の期待値，すなわち w' 内の文字で a_1 ではない文字の数は $p\cdot\frac{\sigma-1}{\sigma}\cdot|S(p)|$ で表される．

これと同じ事を素朴なアルゴリズムの入力 T と P についても考える．a_1 のみで構成される長さ n のテキスト T と，長さ m のパターン P それぞれについて前述のように置換を行うと，T の i 番目の文字 $T[i]$ と，P の j 番目の文字 $P[j]$ が一致するのは次の 3 つの場合に分けられる．

(1) $T[i]$ と $P[j]$ が両方とも置換されない

(2) $T[i]$ と $P[j]$ がが両方とも置換され，かつ両方とも同じ文字になる

(3) $T[i]$ と $P[j]$ の片方のみが置換され，かつもう一方の文字と同じ文字になる

まず (1) について，テキストとパターン双方において，文字が置換されない確率は $(1-p)$ であるため，(1) の確率は $(1-p)\cdot(1-p)$ である．

次に (2) について，文字が置換される確率は p であるため，文字が等しくなるためには置換され，かつ同じ文字にならなければならない．$|\Sigma| = \sigma$ であるとき，条件を満たす確率は $p\frac{1}{\sigma}\cdot p\frac{1}{\sigma} + p\frac{1}{\sigma}\cdot p\frac{1}{\sigma} + \cdots = p\cdot p\frac{1}{\sigma}$ となる．

最後に (3) について，片方が置換されもう片方が置換されず文字が等しくなる確率は $p\cdot(1-p)\frac{1}{\sigma}$ である．置換されるのがテキストの場合とパターンの場合 2 通りあるため求める確率は $2\cdot p\cdot(1-p)\frac{1}{\sigma}$ となる．

この 3 つをまとめると，$T[i]$ と $P[j]$ が一致する確率 $\alpha(i,j)$ は以下のように表される．

$$
\begin{aligned}
&\alpha(i,j)\\
&= (1-p)(1-p) + p\cdot p\frac{1}{\sigma} + 2\cdot(1-p)\cdot p\frac{1}{\sigma}\\
&= 1 - p(2-p)\left(1-\frac{1}{\sigma}\right)\\
&= 1 - \left(1-(1-p)^2\right)\left(1-\frac{1}{\sigma}\right) \qquad (6.8)
\end{aligned}
$$

ここで $\alpha(i,j)$ について，σ は定数であり，$\alpha(i,j)$ は i, j によらない p の関数であるので $\alpha(i,j)$ を α_p と表すことにする．これを用いると $P_F(i,l)$ と $P_S(i,l)$ は以下のようになる．

$$
\begin{aligned}
P_F(l) &= \alpha_p^{l-1}\cdot(1-\alpha_p) \qquad (6.9)\\
P_S(l) &= \alpha_p^{l} \qquad (6.10)
\end{aligned}
$$

これらと式 (6.8) を式 (6.3) に代入し式 (6.7) と同じように式変形を行うと，期待値 $\mathrm{Ex}(T,P)$ は

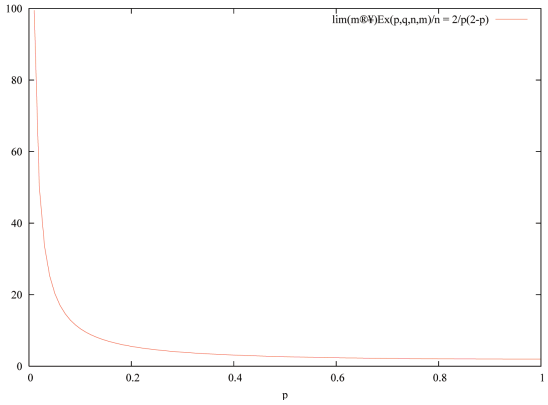


図 6: 最悪時の入力に部分摂動を行った場合の平滑化解析 ($\sigma = 2$)

$$
\begin{aligned}
&\mathrm{Ex}(T,P)\\
&= \sum_{i=1}^{n}\left(\sum_{l=1}^{m}\left(\alpha_p^{l-1}\cdot(1-\alpha_p)\cdot l\right) + \alpha_p^m\cdot m\right)\\
&= \frac{1-\alpha_p^m}{1-\alpha_p}\cdot n\\
&= \frac{n\cdot\left(1-\left(1-p\,(2-p)\left(1-\frac{1}{\sigma}\right)\right)^m\right)}{\left(1-(1-p)^2\right)\left(1-\frac{1}{\sigma}\right)}
\end{aligned}
$$

と表される．

例えば $\sigma = 2$ の時は $\alpha_p = 1 - p + \frac{1}{2}p^2$ となるため，$\mathrm{Ex}(T,P)$ は

$$
\mathrm{Ex}(T,P) = \frac{n}{p-\frac{p^2}{2}}\cdot\left(1-\left(1-\left(p-\frac{p^2}{2}\right)\right)^m\right)
$$

となる．

またこれより，

$$
\lim_{m\to\infty}\frac{\mathrm{Ex}(T,P)}{n} = \begin{cases}\frac{2}{p(2-p)} & (0 < p \le 1)\\ \infty & (p = 0)\end{cases}
$$

となる．この式を図 6に表す．最悪計算時間時の入力ににほんの少しノイズを混ぜただけで，計算時間は劇的に速くなる事がわかる．

6.2 $p \neq q$ の時の平滑化関数の導出

前節まではテキスト T とパターン P に対し摂動の起こる確率が等しいとして考えてきたが，次

は T と P それぞれに対し，摂動が起こる確率が違う場合について考える．T の各文字に摂動が起こる確率を p，P の各文字に摂動が起こる確率を q とする．p, q 共に $0 \le p \le 1, 0 \le q \le 1$ である．

この時，前節の (6.8) と比較して $T[i]$ と $P[j]$ が一致する確率 $\alpha(i,j)$ は以下のようになる．

$$
\begin{aligned}
&\alpha(i,j) \\
&= (1-p)(1-q) + p \cdot q \cdot \frac{1}{\sigma} \\
&\quad +(1-p) \cdot q \cdot \frac{1}{\sigma} + p \cdot (1-q) \cdot \frac{1}{\sigma} \\
&= (1-p)(1-q)(1-\frac{1}{\sigma}) + \frac{1}{\sigma} \qquad (6.11)
\end{aligned}
$$

よって，式 (6.3) に式 (6.9) と (6.10), (6.11) を代入すると求める期待値 $\mathrm{Ex}(T,P)$ は

$$
\mathrm{Ex}(T,P) = \frac{n \cdot \left(1 - (1-\frac{1}{\sigma})^m \left((1-p)(1-q) + \frac{1}{\sigma-1}\right)^m\right)}{(1-(1-p)(1-q))(1-\frac{1}{\sigma})}
$$

となり，また $\lim_{m\to\infty} \frac{\mathrm{Ex}(T,P)}{n}$ は，

$$
\lim_{m\to\infty} \frac{\mathrm{Ex}(T,P)}{n} = \begin{cases} \frac{1}{(1-(1-p)(1-q))(1-\frac{1}{\sigma})} & (0 < p \le 1) \\ \infty & (p = 0) \end{cases}
$$

となる．この式を図 7に表す．

図より，T と P に摂動の起こる確率が異なる場合でも少しノイズを混ぜただけで計算時間は劇的に速くなっている．p または q が 0.15 付近で計算時間はほぼ平均時間となっている．

本年度は，まず，XML テキストの高速処理に関する技法に関する技法についての研究を行った．XML は，あらゆるデータをタグ付きのテキスト文字列として行うことによって可搬性と機械可読性を備えた汎用フォーマットとして，広範囲に活用されている．本研究では，与えられた XML テキストを，パストライとバイナリ XML データの対に変換することによって，領域効率と実行速度を向上させることに成功した．また，圧縮されデータを陽に展開することなく，そのままパターン照合

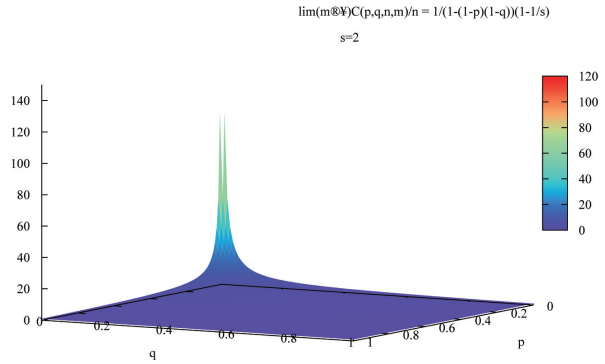


図 7: T と P に置換が起こる確率が異なる時の平滑化解析 ($\sigma = 2$). T と P に摂動の起こる確率が異なる場合でも少しノイズを混ぜただけで計算時間は劇的に速くなることがわかる．

を行う，圧縮文字列照合の研究に関して，形式文法に基づく圧縮法に対する効率のよいアルゴリズムを与えた．さらに，文字列中に共通に現れる共通部分文字列を新たな記号に置き換える作業を繰り返す事による圧縮法に関して，線形時間の圧縮アルゴリズムを与えた．一方，文字列に対する索引構造に関して，単語単位の接尾辞木 (suffix tree) や有向文字列グラフ (DAWG), 圧縮有向文字列グラフ (CDAWG) をそれぞれ線形時間で構築する，簡潔なアルゴリズムを開発した．また，種々の文字列処理アルゴリズムに対して，最悪実行時間の解析に活用されることの多いフィボナッチ文字列について，その繰り返し構造に関する組み合わせ的性質を解析した．

7 XML 文書フィルタリングのための軽量な高速化技法

インターネット上のニュース記事や広告等のメディア，日々観測されるセンサー情報や株価表示器など，膨大な量の情報が絶え間なく生成されている．これらの情報を有効に活用するための技術が強く求められている．SDI (Selective Dissemination of Information) は，そのようなサービスの一つであり，日々生成される大量の機械可読文書

を購読者が選択的に，かつ，リアルタイムに取得することができる．このような publish/subscribe 型のデータ配送システムは，従来，キーワードや “bag-of-words” に基づくフィルタリングによる比較的単純なものであった．

しかし近年，XML (eXtensible Markup Language) の台頭により，XPath を用いたより高度なフィルタリングが可能となった．XPath は，W3C の勧告による，XML 文書の特定の部分 (要素，属性，テキストなど) を指定するための言語で，事実上の標準となっている．

XML データストリームのフィルタリングに関する研究の初期段階では，DOM 木を用いた手法が主流であった．しかし，DOM 木の構築に多大な時間と主記憶領域を要するため，このような手法では，50–200MB 程度の XML データになると事実上適用不可能である．そこで，近年はストリーム処理に基づくフィルタリング手法が主流になっているストリーム処理とは，XML データを一方向に走査することによって出力を得る方式を指し，DOM 木を用いた方式に比べある一定サイズのバッファしか必要とせず，次々と生成されるデータに対して即座に対応可能といった特徴をもつ．

代表的な XML データストリーム処理器である YFilter や XMLTK は，非常に高速に XPath を処理できるが，扱う軸は順方向軸 (forward axes) のみと限定している．これに対して，最近では先祖軸 (ancestor axes)，兄弟軸 (sibling axes) にも対応したストリーム処理手法が提案されているが，順方向軸のみに限定した手法に比べ規模耐性，処理速度等を犠牲にしている．本研究で提案する手法は，扱う軸は順方向軸のみで，オートマトンを用いて XML データを走査する点では YFilter や XMLTK と同様であるが，XML データに前処理を施すことによって高速化・軽量化を可能とした点で異なる．

YFilter は，XFilter のパス照合処理を改良したもので，質問式毎に非決定性有限オートマトン (NFA) を構築するのではなく，質問式全体の共通接頭辞パスを共通の状態として束ねたもので，XFilter よりも質問数に関して規模耐性があり，高速に動作する．しかしながら，NFA は同時に複数の状態遷移を処理しなければならないため，決定性オートマトン (DFA) と比較して照合速度は遅い．NFA を等価な DFA に変換すると，一般に状態数が指数的に増大する．そこで，XMLTK では，XML データ走査中に必要に応じて NFA を部分的に DFA に変換する *Lazy-DFA* と呼ばれる手法を採用している．XMLTK は，質問数に関して規模耐性があり高速であるが，一方，DFA の状態数が増大し大量のメモリを必要とするという欠点がある．

本研究では，配信する XML データにあらかじめ簡単な前処理を施しておくことにより，さらなる規模耐性と高速化・省メモリ化を達成する．ここで提案する前処理手法は，XML データをパストライとバイナリ XML データの対に変換する．パストライは，XML データ中の根から葉へ向かうパスに沿って現れるタグ名の列全体を表すトライ構造である．また，バイナリ XML データは，XML データ中の開始タグと終了タグを各々特殊コードで置き換えたもので，開始タグを表すコードの直後に，タグ名を表す番号ではなく，対応するパストライ節点を識別するための番号（以下，パストライ節点番号）を埋め込んでいる．データ送信サーバは，このパストライとバイナリ XML データを対として送信する．データ受信側では，まず，パストライを対象にパス照合を行い，その結果をパストライの節点に付加する．これにより，バイナリ XML データの走査時にはもはやパス照合の必要はなく，バイナリ XML データ中に埋め込まれたパストライ節点番号を頼りにパス照合の結果を参照することが可能となる．通常，パストライは XML データ全体の木に比べて十分小さいため，処理速度は劇的に向上する．この前処理で行うデータの変換は，以下の望ましい性質を有する．

1. 変換に要する時間が極めて小さい．
2. 変換して得られたデータの大きさは，もとの XML データ以下に抑えられる．

さらに，この手法を SIX と併用することで，さらなる高速化が達成できる．

7.1 パストライを用いた高速なフィルタ処理

本研究では，XQuery の部分族をパスプルーニングと DFA を用いて効率的に処理する手法において，入力となる XML データを事前に調べてパストライを得るが，パスパターン照合をデータ走査時に実行するのではなく，質問式を得た段階で完了させておくことによってより高速なフィルタ処理を実現することに成功した．

以下では用語を定義したあと，フィルタ処理に必要な問題を定式化し，本手法のアルゴリズムを紹介する．

Σ を文字の有限集合とし，$\mathcal{N}$ をタグ名から成る集合とする．*XML* **木**とは，$\mathcal{N}$ のタグ名を内部節点のラベルとし，Σ 上の文字列を葉 (テキスト節点) のラベルとするようなラベル付き順序木をいう．

単純パスパターンとは $\mathcal{N}$ のタグ名と特殊な記号 “*” から成る記号列をいい，各々の記号は “/” または “//” で区切られているものとする．ここで，“/” と “//” は，それぞれ，親子関係，先祖-子孫関係に対応する．単純パスパターン π が XML 木のパスに**マッチする**とは，“*” を任意のタグ名にマッチするワイルドカード，“//” を任意のタグ名列にマッチする可変長ワイルドカードとみなした際に，π がそのパスに沿ったタグ名の列に合致するときをいう．

パスパターンとは単純パスパターン π_1，π_2 の順序対であり，$\pi_1[\pi_2]$ と表す．XML 木の任意の節点 x とその子孫 y（$x = y$ の場合も含む）に対し，パスパターン $\pi_1[\pi_2]$ が**位置** (x, y) **に出現する**とは，π_1，π_2 が，それぞれ，根から x へのパスと x' から y へのパスにマッチするときをいう．ここで，x' は，根から y へ向かうパスに沿った x の子節点である．

$e = e(w_1, \ldots, w_m)$ を，Σ 上の空でない文字列 $w_1, \ldots, w_m$ の生起の有無に関する論理式であるとする．Σ 上のテキスト d が論理式 e を**充足する**とは，d におけるキーワード $w_1, \ldots, w_m$ の生起に対応した真偽値割り当て (truth assignment) のもとで e が真となるときをいう．

XPattern とは，パスパターン $\pi_1[\pi_2]$ と論理式 e の組をいい，$\pi_1[\pi_2 : e]$ と表す．XPattern $\pi_1[\pi_2 : e]$ が XML 木の**節点** x **に出現する**とは，x の子孫 y が存在して，パスパターン $\pi_1[\pi_2]$ が位置 (x, y) に出現し，かつ，y の子であるテキスト節点の少なくとも一つが e を充足するときをいう．この定義より，パスパターン $\pi_1[\pi_2]$ は，XPattern $\pi_1[\pi_2 : \mathbf{true}]$ と見なすことができる．図 8 に，XPattern $\pi_1[\pi_2 : e]$ とその生起位置 (x, y) との関係を図示した．

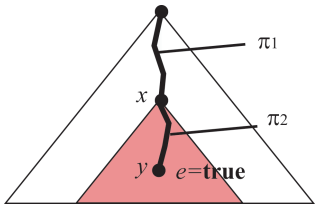


図 8: XML 木の位置 (x, y) における XPattern $\pi_1[\pi_2 : e]$ の生起．

本研究では，以下の問題に取り組む．

Definition 1
Given: *XML* データ T.
Query: *XPattern* $P_1, \ldots, P_\ell$.
Answer: 各 $i = 1, \ldots, \ell$ に対し，T に対応する *XML* 木において P_i が生起する節点 x を返せ．

ここで，入力として与えられる XML データ T は，複数の XML 文書の系列（XML 木の系列）となっていることを仮定している．なお，フィルタリング問題では，該当の XML 文書が質問に合致するかどうかだけを考えればよいため，上の定義で節点 x を求める必要はないように見える．しかし，複数の XPattern を組み合わせた複雑な質問を処理するためには，このような節点 x を求めることが必要である．

パストライとバイナリデータ

本手法では XML データに対して一度だけ前処理を行い，パストライと，バイナリ XML ファイ

表 2: DBLP と xmlgen によってランダムに生成された入力 XML データの大きさ．

	XML データ		
	サイズ (MB)	節点数	タグ数
DBLP	352	8,632,812	35
random	111	1,666,310	74

表 3: 前処理で構築されるバイナリ XML データとパストライのサイズ及びその構築時間.

	バイナリ XML ファイル	パストライ	CPU 時間
	サイズ (MB)	節点数	(sec)
DBLP	208	138	100.41
random	84	515	73.26

ルを作成する．XML データと前処理によって作成されるパストライ及びバイナリ XML データとの関係を図 9に示す．

この前処理に要する時間は $O(n \log |\mathcal{N}| + |T|)$ であり，表 3に示すとおり高速な処理が可能である．ここで，n は XML データ T 中の開始タグの個数を表し，$|\mathcal{N}|$ はタグ集合の要素数を，$|T|$ は XML データのサイズを表す．また，XML データと生成されるデータサイズの関係を表 2, 3 に示す．このことから，パストライは，XML 木に対して十分に小さく，バイナリ XML データに関しても入力ファイルサイズの 76%と小さくなっていることが分かる．

バイナリ XML データについては，開始タグを表す特殊記号の直後に，タグを表す番号ではなく，根からのパスに沿ったラベルの列を表すパストライ節点番号を埋め込むことに注意されたい．このことによって，パスパターン照合はパストライを対象としたパターン照合を行うだけで良く，パストライが XML データに比べ非常に小さいため，高速に処理を完了することが可能となる．

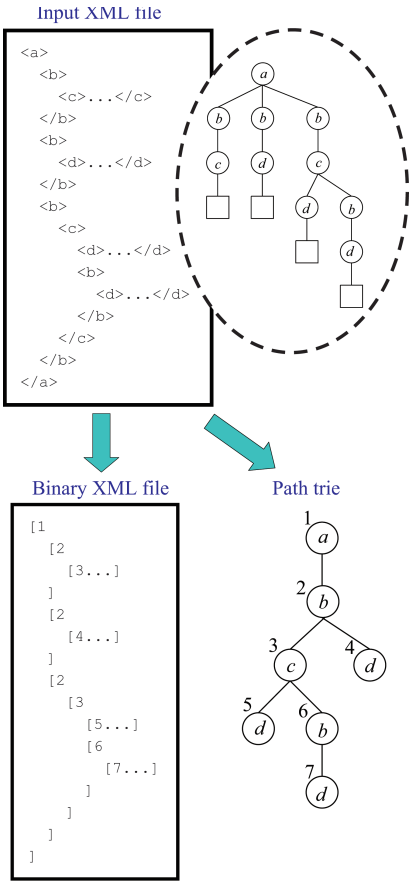


図 9: XML データに対して行う前処理の様子を図示した．XML データと対応する XML 木を上側に示した．正方形はテキスト節点を表す．この XML データから作成されるバイナリ XML データとパストライを，それぞれ，左下と右下に示した．パストライの節点の傍にある数は，その節点の ID である．バイナリ XML データにおいて，開始タグと終了タグは，それぞれ，特殊記号 “[” と “]” に置き換えられ，“[” の直後にはパストライの節点 ID が埋め込まれていることに注意しよう．

パスパターンの照合

本研究では，XML 木に対するパスパターン照合の問題を，パストライに対するパスパターン照合に置き換える．

図 9で示した XML データを例にとり，XML 木中，および，パストライ中におけるパスパターン $P = \mathrm{a//b[//d]}$ の生起の様子を図 10に示した．この例では，パスパターン $P_1 = \mathrm{a//b[//d]}$ は XML 木中の位置 $(x, y) = (4, 5), (6, 8), (6, 10), (9, 10)$ で生起している．一方，パストライ中においては，生起位置は $(x, y) = (2, 5), (2, 7), (6, 7), (2, 4)$ である．そこで，パストライ中の生起位置 (x, y) の節点 y にパスパターンの番号と節点 x, y の深さの差の情報を付加する．すると，XML 木中の生起位置は，パストライに付加された情報から得ることができる．

次に，パストライに対するパスパターン照合の手法について述べる．質問式に含まれる各パスパターンに対し，パストライ中の出現位置 (x, y) をすべて求めなければならない．このために，各パスパターンから非決定有限オートマトン (NFA) をそれぞれ作成し，それらをパストライ中を深さ優先で巡回させることにより照合を行う．NFA の実装にはビット並列化手法を用いる．多くの場合，各 NFA の状態数がワード長である 32（あるいは

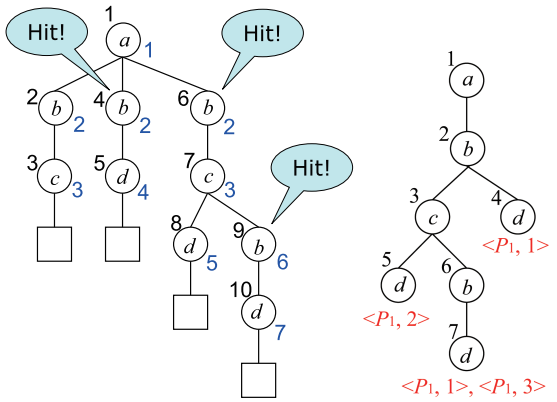


図 10: パスパターン P_1 = a//b[//d] は，図左の XML 木の位置 $(x, y) = (4, 5), (6, 8), (6, 10), (9, 10)$ に生起している．これらの位置は，図右に示したパストライの節点に付加された情報から得ることができる．たとえば，XML 木中の位置 $(6, 8)$ における P_1 の生起は，パストライの節点 5 のもつ出力 $\langle P_1, 2\rangle$ から得ることができる．すなわち，バイナリ XML データ中において，XML 木の節点 $y = 8$ に対応する箇所にはパストライ節点番号 5 が埋め込まれており，パストライの節点 5 に付加された情報 $\langle P_1, 2\rangle$ によって，y から 2 つ上の先祖である節点 $x = 6$ が得られる．

64）以内に収まるため，非決定的な状態遷移を並列に高速に処理することが可能となる．また NFA の構築も極めて高速である．これら二つの NFA を用いて，パストライ中を深さ優先で巡回し，パスパターン $P_i = \pi_1[\pi_2]$ の出現位置 (x, y) をすべて求める．出現位置 (x, y) がみつかる度に，節点 y に出力 $\langle P_i, d\rangle$ を付加する．ここに，d は節点 x, y の深さの差である．

本手法を用いることで，バイナリ XML データ走査時のパスパターン照合処理は不要となり，XML データ走査時にパスパターンの照合を行っている YFilter や XMLTK と比べ，処理時間を大幅に短縮することができる．

キーワード照合

パスパターン照合は，フィルタリング問題の一部に過ぎない．XPath の仕様から明らかなように，述語はパス以外に，テキストデータや属性，ポジションを含むことができる．このうち，テキストデータや属性は，/a/b[@month=December]，/a/b[containts(name, “mickey”)] などのように記述され，前者はパス/a/b の属性値が December であること，後者は/a/b/name のテキスト節点に，キーワード “mickey” があることを調べるときに用いられる．

フィルタリングにおいてこのような問い合わせの処理を考えるとき，XML データ中のすべてのテキスト節点における複数キーワードの出現を求める問題として捉えることができる．著者らは，この問題に対して古典的手法である Aho-Corasick(AC) 法を用いた．AC 法は同時に複数キーワードの照合を行うことができ，キーワード数の増加に伴うテキスト走査速度の低下は特にない．また，索引を用いる方法では，キーワード毎に出現するテキストの番号のリストを求め，ブール式に応じた集合演算を行うが，AC 法で複数のキーワードを同時に照合する場合には，テキスト節点毎にブール演算を行えばよいため，集合演算のために領域・時間を消費しないという利点もある．フィルタリングにおいては，大量の質問式から集めたキーワードの総数は膨大なものになり得るため，これらの利点のもつ意義は大きい．

バイナリ XML データ走査アルゴリズム

バイナリ XML データを走査してキーワード照合を行い，その結果とパスパターン照合の結果を結合して質問処理を行うアルゴリズムの概略を以下に示す．

与えられた XPattern $P_1, \ldots, P_\ell$ の論理式 e 中に含まれるキーワードからなる集合を $W = w_1, \ldots, w_m$ とする．W から AC 機械 M を構築する．バイナリ XML ファイル先頭からの位置を表す変数 *depth* と，入力を XML 木として見たときの節点の深さを表す変数 *depth* を用意する．ブール値を格納する 2 次元配列 Occ と Q を用意する．配列 Occ は，現在巡回している深さ d の節点の子であるテキスト節点にキーワード w_i が生起するときに $Occ[d][i]$ の値が真となるというものであり，配列 Q は，現在巡回している深さ d の節点に XPattern P_q が出現するときに $Q[d][q]$ の値が真となる．本アルゴリズムは，バイナリ XML ファイルを 1 バイトずつ走査し，読み込んだバイト c

に対して以下を実行する．

- c が開始タグを表す特殊記号 “[” であるとき．パストライ節点番号 v を読み込み，$(v, depth)$ をスタック S にプッシュする．$depth$ を 1 インクリメントする．$Occ[depth][1 \ldots m]$ と $Q[depth][1 \ldots l]$ を偽に初期化する．M の状態を初期状態に設定する．
- c が終了タグを表す特殊記号 “]” であるとき．スタック S より $(v, depth)$ をポップする．もしパストライのノード v が出力をもっていたら，それぞれの $\langle q, d \rangle$ に対して，XPattern P_q の論理式 e を，$Occ[depth][1 \ldots m]$ の値に基づいて計算し，もし e が真ならば，$Q[depth-d][q]$ を真と設定する．$depth$ を 1 デクリメントする．M の状態を初期状態に戻す．
- c がそれ以外のバイトコードであるとき．M を文字 c で状態遷移させる．もし遷移先の状態が出力をもつならば生起したキーワードに対応して $Occ[depth][1 \ldots m]$ を更新する．

実験の詳細については，本報告からは割愛するが，パストライがフィルタ処理の高速化に寄与することを示し，本手法が YFilter や XMLTK と比較して高い性能を有することわかった．

8 文字列の繰り返し構造の解析

隣接 2 項の漸化式で定義される**フィボナッチ数列**は，

$$\Phi_0 = 1, \quad \Phi_1 = 2,$$
$$\Phi_m = \Phi_{m-1} + \Phi_{m-2} \quad (m \geq 2)$$

である．漸化式 $\Phi_m = \Phi_{m-1} + \Phi_{m-2}$ から導かれる特性方程式 $x^2 = x + 1$ の正の実数解 $(1+\sqrt{5})/2$ は**黄金比**として知られており，ϕ と表記される．また，これは**フィボナッチ定数**とも呼ばれ，フィボナッチ数の隣接 2 項の比の極限がこの値と等しくなる．

フィボナッチ数列は隣接 2 項の数の和で生成されるが，これを文字に置き換え，次のように隣接 2 項の文字列の連結として生成される文字列がフィボナッチ文字列 F_m である．

$$F_0 = \mathtt{a}, \quad F_1 = \mathtt{ab}, \quad F_m = F_{m-1}F_{m-2} \quad (m \geq 2).$$

よって，任意の $m \geq 0$ について $|F_m| = \Phi_m$ となる．フィボナッチ文字列の初めの数項は次のようになる．

$$F_2 = \mathtt{aba},$$
$$F_3 = \mathtt{abaab},$$
$$F_4 = \mathtt{abaababa}.$$

また，文字列 w における部分文字列 u の繰り返し（連続出現），および繰り返し数を次のように定義する．文字列 u，v，w が

$$w = uu...uv = u^k v \quad (v \text{ は } u \text{ の接頭辞})$$

であり，$k \geq 2$ であるとき，文字列 w は u の繰り返しであるといい，

$$e = \frac{|w|}{|u|} = k + \frac{|v|}{|u|}$$

を文字列 u の繰り返し数という．

フィボナッチ数は自然現象のあらゆるところに出現し，面白い性質を示すことが知られている．フィボナッチ文字列も部分文字列のグラフ化問題や組み合わせ問題等においてさまざまな性質を持つことが知られている．そのひとつとして，Knuth-Morris-Pratt アルゴリズムや Boyer-Moore アルゴリズム，Aho-Corasick アルゴリズムといった文字列のパターンマッチングアルゴリズムにおいて最悪の入力インスタンスとなることが知られている．特に，Knuth-Morris-Pratt アルゴリズムでは，マッチングに失敗したときにキーワードの何文字目と照合すればよいかという情報を用いて再照合を行うが，キーワードの長さを z としたとき，この再照合の回数が高々 $\log_\phi z + 1$ 回であることが証明されている．

Mignosi と Pirillo は無限フィボナッチ文字列中における m 番目の有限フィボナッチ文字列の最大繰り返し数について，無限フィボナッチ文字列中に $2 + \phi \approx 3.618$ 回を超える部分文字列の繰り返し出現はないことを証明した [4]．

本研究では，隣接 n 項の文字列を連結したものとして定義される n-ボナッチ文字列と n-ボナッチ定数 $\phi^{(n)}$ との関連を明らかにした．フィボナッチ文字列の場合 $n = 2$ となり，その最大繰り返し数の極限は $2 + \phi^{(2)}$ である．では，$n > 2$ における一般的な最大繰り返し数の極限はどのように

表現されるだろうかが興味の対象となる．フィボナッチ文字列の場合ではフィボナッチ定数 $\phi^{(2)}$ がそのままの形で現れたので，直感的には $2+\phi^{(n)}$ となりそうであるが，実際には $2+1/(\phi^{(n)}-1)$ となることを本研究で示した．特に $n=2$ のとき，$\phi^{(2)}=1/(\phi^{(2)}-1)$ であり，これは Mignosi と Pirillo の結果と一致する．

ϵ を空文字とする．文字列 u_0，u_1，u_2，... の連結を，整数 i，j を用いて次のように表現する．

$$u_{[i\nearrow j]}=\begin{cases} u_iu_{i+1}u_{i+2}...u_{j-1}u_j & (i\le j)\\ \epsilon & (i>j),\end{cases}$$

$$u_{[i\searrow j]}=\begin{cases} u_iu_{i-1}u_{i-2}...u_{j+1}u_j & (i\ge j)\\ \epsilon & (i<j).\end{cases}$$

さらに，$u_iu_{i+1}u_{i+2}...$ のように文字列を無限に連結してできる文字列を $u_{[i\nearrow\infty]}$ と表す．

n-ボナッチ数は次のように定義される．

$$\Phi_0^{(n)}=1,$$

$$\Phi_m^{(n)}=\begin{cases} \displaystyle\sum_{i=0}^{m-1}\Phi_i^{(n)}+1=2^m & (1\le m<n),\\ \displaystyle\sum_{i=m-i}^{m-1}\Phi_i^{(n)} & (m\ge n).\end{cases}$$

これを元に，n-ボナッチ文字列を次のように定義する．

Definition 2 *n を 2 以上の整数とし，アルファベット $\Sigma=\{\sigma_1,\ \sigma_2,\ \sigma_3,\ ...,\ \sigma_n\}$ とする．次の漸化式で定義される文字列を n-**ボナッチ文字列**という．*

$$F_0^{(n)}=\sigma_1,$$

$$F_m^{(n)}=\begin{cases} \epsilon & (n<0),\\ F^{(n)}_{[m-1\searrow 0]}\sigma_{m+1} & (0<m<n),\\ F^{(n)}_{[m-1\searrow m-n]} & (m\ge n).\end{cases} \tag{8.1}$$

Example 1 *$n=3$ のとき，アルファベット $\Sigma=\{\mathtt{a},\ \mathtt{b},\ \mathtt{c}\}$ とすると，*

$$F_0^{(3)}=\mathtt{a},$$
$$F_1^{(3)}=F_0^{(3)}\mathtt{b}=\mathtt{ab},$$
$$F_2^{(3)}=F_1^{(3)}F_0^{(3)}\mathtt{c}=\mathtt{abac},$$
$$F_m^{(3)}=F_{m-1}^{(3)}F_{m-2}^{(3)}F_{m-3}^{(3)}\quad(m\ge 3).$$

となる．また，初めの数項は以下のようになる．

$$F_3^{(3)}=\mathtt{abacaba},$$
$$F_4^{(3)}=\mathtt{abacabaabacab},$$
$$F_5^{(3)}=\mathtt{abacabaabacababacabaabac}.$$

これはトリボナッチ文字列とも呼ばれる． ■

n-ボナッチ数と n-ボナッチ文字列には，$\left|F_m^{(n)}\right|=\Phi_m^{(n)}$ という関係がある．また，無限 n-ボナッチ文字列 $F_\infty^{(n)}$ も定義でき，全ての有限 n-ボナッチ文字列を接頭辞として持っている．

Example 2 *$n=3$ のとき，無限トリボナッチ文字列は次のようになる．*

$$F_\infty^{(3)}=\mathtt{abacabaabacababacabaabacabacabaab}...$$

■

Theorem 3 *$m-k\ge n+1$ および $n\ge 2$ を満たす任意の m，n，k において，$F_m^{(n)}$ 中に出現する $F_k^{(n)}$ の最長の連続出現は $F_k^{(n)}F_k^{(n)}F_k^{(n)}G_{k-n+1}^{(n)}$ である．*

Theorem 4 *$m-k\ge n+1$ および $n\ge 2$ を満たす任意の m，n，k において，$F_m^{(n)}$ 中の $F_k^{(n)}$ の最大繰り返し数の極限は $2+1/(\phi^{(n)}-1)$ である．*

Theorem 5 *整数 n，$k\to\infty$ のとき，n-ボナッチ文字列 $F_k^{(n)}$ の最大繰り返し数はちょうど 3 に近づく．*

参考文献

[1] A. Apostolico. String editing and longest common subsequences. *Handbook of Formal Languages*, 2:361–398, 1997.

[2] S. Kim and K. Park. A dynamic edit distance table. In *Proc. 11th Ann. Symp. on Combinatorial Pattern Matching*, volume 1848 of *Lecture Notes in Computer Science*, pages 60–68. Springer-Verlag, 2000.

[3] G. Landau, E. Myers, and M. Ziv-Ukelson. Two algorithms for LCS consecutive suffix alignment. In *Proc. 15th Ann. Simp. on Combinatorial Pattern Matching*, volume 3109 of *Lecture Notes in Computer Science*, pages 173–193. Springer-Verlag, 2004.

[4] F. Mignosi and G. Pirillo. Repetitions in the Fibonacci infinite word. *RAIRO Theoret. Inform. Appl.*, 26:199–204, 1992.

[5] M. Miyazaki, A. Shinohara, and M. Takeda. An improved pattern matching algorithm for strings in terms of straight line programs. *J. Disc. Algo.*, 1(1):187–204, 2000.

[6] D. A. Spielman and S.-H. Teng. Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time. In *Proc. of the 33rd Ann. ACM Symp. on Theory of Computing (STOC)*, pages 265–305, 2001.

研究業績一覧

学術論文

1. T. Suga and S. Shimozono:
"Approximate Point Set Pattern Matching on Sequences and Planes", Proc. 15th Annual Symposium on Combinatorial Pattern Matching (CPM2004), Lecture Notes in Computer Science 3109, pp. 89–101, July 2004.
概要: The point set pattern matching problem is, given two sets "pattern" and "text" of points in Euclidean space, to find a linear transformation that maps the pattern to a subset of the text. We introduce an approximate point set pattern matching for axis-sorted point sequences that allows a translation, space insertions and deletions between points. We present an approximate pattern matching algorithm that runs with pattern size n and text size m in $O(nm^2)$ time in general, and in $O(nm)$ time if distances between two contiguous points in texts and patterns are finite. A variant of the four-Russian technique achieving $O(nm/\log n + n\log n)$ time is also provided. Furthermore, as a natural extension we present an approximate point set pattern matching on the plane, and give a polynomial-time algorithm that solves this problem.

2. S. Inenaga, A. Shinohara, and M. Takeda:
"A Fully Compressed Pattern Matching Algorithm for Simple Collage Systems", Proc. The Prague Stringology Conference 2004 (PSC2004), Czech Technical University Press, pp. 98–113, September 2004.
概要: We study the fully compressed pattern matching problem (FCPM problem): Given T and P which are descriptions of text T and pattern P respectively, find the occurrences of P in T without decompressing T or P. This problem is rather challenging since patterns are also given in a compressed form. In this paper we present an FCPM algorithm for simple collage systems. Collage systems are a general framework that can represent various kinds of dictionary-based compressions, and simple collage systems are a subclass that includes LZW and LZ78 compressions. Collage systems are of the form $\langle D, S\rangle$, where D is a dictionary and S is a sequence of variables from D. Our FCPM algorithm performs in $O(||D||^2 + mn\log|S|)$ time, where $n = |T| = ||D|| + |S|$ and $m = |P|$. This is faster than the previous best result of $O(m^2n^2)$ time.

3. H. Bannai, H. Hyyrö, A. Shinohara, M. Takeda, K. Nakai, and S. Miyano:
"Finding Optimal Pairs of Patterns", Proc. 4th Workshop on Algorithms in Bioinformatics (WABI2004), Lecture Notes in Bioinformatics 3240, pp. 450–452, September 2004.

概要: We consider the problem of Finding the optimal pair of string patterns for discriminating between two sets of strings, i.e. finding the pair of patterns that is best with respect to some appropriate scoring function that gives higher scores to pattern pairs which occur more in the strings of one set, but less in the other. We present an $O(N^2)$ time algorithm for finding the optimal pair of substring patterns, where N is the total length of the strings. The algorithm looks for all possible Boolean combination of the patterns, e.g. patterns of the form $p \wedge \neg q$, which indicates that the pattern pair is considered to match a given string s, if p occurs in s, AND q does NOT occur in s. The same algorithm can be applied to a variant of the problem where we are given a single set of sequences along with a numeric attribute assigned to each sequence, and the problem is to find the optimal pattern pair whose occurrence in the sequences is correlated with this numeric attribute. An efficient implementation based on suffix arrays is presented, and the algorithm is applied to several nucleotide sequence datasets of moderate size, combined with microarray gene expression data, aiming to find regulatory elements that cooperate, complement, or compete with each other in enhancing and/or silencing certain genomic functions.

4. S. Inenaga, H. Bannai, H. Hyyrö, A. Shinohara, M. Takeda, K. Nakai, and S. Miyano: "Finding Optimal Pairs of Cooperative and Competing Patterns with Bounded Distance", Proc. 7th International Conference on Discovery Science (DS2004), Lecture Notes in Artificial Intelligence 3245, pp. 32–46, October 2004.
概要: We consider the problem of discovering the optimal pair of substring patterns with bounded distance α, from a given set S of strings. We study two kinds of pattern classes, one is in form $p \wedge_\alpha q$ that are interpreted as cooperative patterns within α distance, and the other is in form $p \wedge_\alpha \neg q$ representing competing patterns, with respect to S. We show an efficient algorithm to find the optimal pair of patterns in $O(N^2)$ time using $O(N)$ space. We also present an $O(m^2N^2)$ time and $O(m^2N)$ space solution to a more difficult version of the optimal pattern pair discovery problem, where m denotes the number of strings in S.

5. A. Shinohara: "String Pattern Discovery", Proc. 15th International Conference Algorithmic Learning Theory (ALT2004), Lecture Notes in Artificial Intelligence 3244, pp. 1–13, October 2004.
概要: Finding a good pattern which discriminates one set of strings from the other set is a critical task in knowledge discovery. In this paper, we review a series of our works concerning with the string pattern discovery. It includes theoretical analyses of learnabilities of some pattern classes, as well as development of practical data structures which support efficient string processing.

6. H. Sakamoto, T. Kida, and S. Shimozono: "A Space-Saving Linear-Time Algorithm for Grammar-Based Compression", Proc. 11th International Symposium on String Processing and Information Retrieval (SPIRE2004), Lecture Notes in Computer Science 3246, pp. 218–229, October 2004.
概要: A space-efficient linear-time approximation algoirithm for the grammar-based compression problem, which requests for a given string to find a smallest context-free grammar deriving the string, is presented. The algorithm consumes only $O(g_* \log g_*)$ space and achieves the worst case approximation ratio $O(\log g_* \log n)$, with the size n of an input and the optimum grammar size g_*. Experimental results for

typical benchmarks demonstrate that our algorithm is practical and efficient.

7. H. Bannai, H. Hyyrö, A. Shinohara, M. Takeda, K. Nakai, and S. Miyano: "An $O(N^2)$ Algorithm for Discovering Optimal Boolean Pattern Pairs", IEEE/ACM Transactions on Computational Biology and Bioinformatics, Vol. 1, No. 4, pp. 159–170, December, 2004.
概要: We consider the problem of finding the optimal combination of string patterns, which characterizes a given set of strings that have a numeric attribute value assigned to each string. Pattern combinations are scored based on the correlation between their occurrences in the strings and the numeric attribute values. The aim is to find the combination of patterns which is best with respect to an appropriate scoring function. We present an $O(N^2)$ time algorithm for finding the optimal pair of substring patterns combined with Boolean functions, where N is the total length of the sequences. The algorithm looks for all possible Boolean combinations of the patterns, e.g., patterns of the form $p \wedge \neg q$, which indicates that the pattern pair is considered to occur in a given string s, if p occurs in s, AND q does NOT occur in s. An efficient implementation using suffix arrays is presented, and we further show that the algorithm can be adapted to find the best *k-pattern* Boolean combination in $O(N^k)$ time. The algorithm is applied to mRNA sequence data sets of moderate size combined with their turnover rates for the purpose of finding regulatory elements that cooperate, complement, or compete with each other in enhancing and/or silencing mRNA decay.

8. H. Hyyrö, J. Takaba, A. Shinohara, and M. Takeda: "On Bit-Parallel Processing of Multibyte Text", Proc. The First Asia Information Retrieval Symposium (AIRS2004), Lecture Notes in Computer Science 3411, pp. 289–300, February 2005.
概要: There exist practical bit-parallel algorithms for several types of pair-wise string processing, such as longest common subsequence computation or approximate string matching. The bit-parallel algorithms typically use a size-σ table of match bit-vectors, where the bits in the vector for a character λ identify the positions where the character λ occurs in one of the processed strings, and σ is the alphabet size. The time or space cost of computing the match table is not prohibitive with reasonably small alphabets such as ASCII text. However, for example in the case of general Unicode text the possible numerical code range of the characters is roughly one million. This makes using a simple table impractical. In this paper we evaluate three different schemes for overcoming this problem. First we propose to replace the character code table by a character code automaton. Then we compare this method with two other schemes: using a hash table, and the binary-search based solution proposed by Wu, Manber and Myers. We find that the best choice is to use either the automaton-based method or a hash table.

9. S. Inenaga, A. Shinohara, M. Takeda: "An Efficient Pattern Matching Algorithm on a Subclass of Context Free Grammars", Proc. 8th International Conference on Developments in Language Theory (DLT2004), Lecture Notes in Computer Science 3340, pp. 225–236, December 2004.
概要: There is a close relationship between formal language theory and data compression. Since 1990's various types of grammar-based text compression algorithms have been introduced. Given an input string, a grammar-based text com-

pression algorithm constructs a context-free grammar that only generates the string. An interesting and challenging problem is pattern matching on context-free grammars $\mathcal{P}$ of size m and $\mathcal{T}$ of size n, which are the descriptions of pattern string P of length M and text string T of length N, respectively. The goal is to solve the problem in time proportional only to m and n, not to M nor N. Kieffer et al. introduced a very practical grammar-based compression method called multilevel pattern matching code (MPM code). In this paper, we propose an efficient pattern matching algorithm which, given two MPM grammars $\mathcal{P}$ and $\mathcal{T}$, performs in $O(mn^2)$ time with $O(mn)$ space. Our algorithm outperforms the previous best one by Miyazaki et al. which requires $O(m^2n^2)$ time and $O(mn)$ space.

10. H. Hyyrö, Y. Pinzon and A. Shinohara:
"Fast Bit-Vector Algorithms for Approximate String Matching under Indel-Distance", Proc. The 31st Annual Conference on Current Trends in Theory and Practice of Informatics (SOFSEM2005), Lecture Notes in Computer Science 3381, pp. 380–384, January 2005.
概要: The approximate string matching problem is to find all locations at which a query p of length m matches a substring of a text t of length n with at most k differences (insertions, deletions, substitutions). The fastest solutions in practice for this problem are the bit-parallel NFA simulation algorithms of Wu & Manber and Baeza-Yates & Navarro, and the bit-parallel dynamic programming algorithm of Myers. In this paper we present modified versions of these algorithms to deal with the restricted case where only insertions and deletions (called indel for short) are permitted. We also show test results with the algorithms.

11. H. Sakamoto:
"A Fully Linear-Time Approximation Algorithm for Grammar-Based Compression", Journal of Discrete Algorithm, Vol. 3, Issues 2–4, pp. 416–430, June 2005.
概要: A linear-time approximation algorithm for the *grammer-based compression* is presented. This is an optimization problem to minimize the size of a context-free grammar deriving a given string. For each string of length n, the algorithm guarantees $O(\log \frac{n}{g_*})$ approximation ratio without suffix tree construction.

12. T. Maita and H. Sakamoto:
"A Simple Extension of Queriable Compression for XML Data", Proc. The 2005 International Conference on Active Media and Technology (AMT2005), pp. 91–95, May 2005.
概要: Recently, the method of reverse arithmetic encoding for effective compression and fast query process for semi-structured data was proposed. We apply this method and develop a new method for more complicated query process. By the reverse arithmetic encoding, we can check the dependency between any two paths without decoding and using this property, a fast query process over tree structures is obtained. However, it is difficult to handle general queries since the type of query is very restricted. So, we expand the algorithm to realize the Boolean operations "NOT" and "OR" in the query expressions.

13. S. Inenaga, A. Shinohara, and M. Takeda:
"A Fully Compressed Pattern Matching Algorithm for Simple Collage Systems", International Journal of Fundations of Computer Science, Vol. 16, No. 6, pp. 1155–1166, December 2005.
概要: We study the fully compressed pattern matching problem (FCPM problem): Given $\mathcal{T}$ and $\mathcal{P}$ which are descriptions of

text T and pattern P respectively, find the occurrences of P in T without decompressing $\mathcal{T}$ or $\mathcal{P}$. This problem is rather challenging since patterns are also given in a compressed form. In this paper we present an FCPM algorithm for simple collage systems. Collage systems are a general framework representing various kinds of dictionary-based compressions in a uniform way, and simple collage systems are a subclass that includes LZW and LZ78 compressions. Collage systems are of the form $\langle \mathcal{D}, \mathcal{S} \rangle$, where $\mathcal{D}$ is a dictionary and $\mathcal{S}$ is a sequence of variables from $\mathcal{D}$. Our FCPM algorithm performs in $O(||\mathcal{D}||^2 + mn \log |\mathcal{S}|)$ time, where $n = |\mathcal{T}| = ||\mathcal{D}||$ and $m = |\mathcal{P}|$. This is faster than the previous best result of $O(m^2 n^2)$ time.

14. Z. Troníček and A. Shinohara:
"The Size of Subsequence Automaton", Theoretical Computer Science, Vol. 341, Issue 1–3, pp. 379-384, September 2005.
概要: Given a set of strings, the subsequence automaton accepts all subsequences of these strings. We derive a lower bound for the maximum number of states of this automaton. We prove that the size of the subsequence automaton for a set of k strings of length n is $\Omega(n^k/(k+1)^k k!)$ for any $k \geq 1$. It solves an open problem because only the case $k \leq 2$ was shown before.

15. Y. Ishida, S. Inenaga, A. Shinohara, and M. Takeda:
"Fully Incremental LCS Computation", Proc. The 15th International Symposium of Fundamentals of Computation Theory (FCT2005), Lecture Notes in Computer Science 3623, pp. 563–574, August 2005.
概要: Sequence comparison is a fundamental task in pattern matching. Its applications include file comparison, spelling correction, information retrieval, and computing (dis)similarities between biological sequences. A common scheme for sequence comparison is the longest common subsequence (LCS) metric. This paper considers the fully incremental LCS computation problem as follows: For any strings A, B and characters a, b, compute $LCS(aA, B)$, $LCS(A, bB)$, $LCS(Aa, B)$, and $LCS(A, Ba)$, provided that $L = LCS(A, B)$ is already computed. We present an efficient algorithm that computes the four LCS values above, in $O(L)$ or $O(n)$ time depending on where a new character is added, where n is the length of A. Our algorithm is superior in both time and space complexities to the previous known methods.

16. H. Hyyrö, Y. Pinzon, and A. Shinohara:
"New Bit-Parallel Indel-Distance Algorithm", Proc. 4th International Workshop on Efficient and Experimental Algorithms (WEA2005), Lecture Notes in Computer Science 3503, pp. 380–390, May 2005.
概要: The task of approximate string matching is to find all locations at which a pattern string p of length m matches a substring of a text string t of length n with at most k differences. It is common to use Levenshtein distance, which allows the differences to be single-character insertions, deletions, substitutions. Recently, the IndelMYE, IndelWM and IndelBYN algorithms where introduced as modified version of the bit-parallel algorithms of Myers, Wu&Manber and Baeza-Yates&Navarro, respectively. These modified versions where made to support the indel distance (only single-character insertions and/or deletions are allowed). In this paper we present an improved version of IndelMYE that makes a better use of the bit-operations and runs 24.5 percent faster in practice. In the end we present a complete set of experimental results to support our findings.

17. H. Bannai, K. Hatano, S. Inenaga, and M. Takeda:
"Practical Algorithms for Pattern Based Linear Regression", Proc. 8th International Conference on Discovery Science (DS2005), Lecture Notes in Computer Science 3735, pp. 44–56, October 2005.
概要: We consider the problem of discovering the optimal pattern from a set of strings and associated numeric attribute values. The goodness of a pattern is measured by the correlation between the number of occurrences of the pattern in each string, and the numeric attribute value assigned to the string. We present two algorithms based on suffix trees, that can find the optimal substring pattern in $O(Nn)$ and $O(N^2)$ time, respectively, where n is the number of strings and N is their total length. We further present a general branch and bound strategy that can be used when considering more complex pattern classes. We also show that combining the $O(N^2)$ algorithm and the branch and bound heuristic increases the efficiency of the algorithm considerably.

18. H. Tsuji, A. Ishino, and M. Takeda:
"A Bit-Parallel Tree Matching Algorithm for Patterns with Horizontal VLDC's", Proc. 12th International Conference on String Processing and Information Retrieval (SPIRE2005), Lecture Notes in Computer Science 3772, pp. 388–398, November 2005.
概要: The tree pattern matching problem is, given two labeled trees P and T, respectively called pattern tree and target tree, to find all occurrences of P within T. Many studies have been undertaken on this problem for both the cases of ordered and unordered trees. To realize flexible matching, a kind of variable-length-don't-care's (VLDC's) have been introduced. In particular, the path-VLDC's appear in XPath, a language for addressing parts of an XML document. In this paper, we introduce horizontal VLDC's, each matches a sequence of trees whose root nodes are consecutive siblings in ordered trees. We address the tree pattern matching problem for patterns with horizontal VLDC's. In our setting, the target tree is given as a tagged sequence such as XML data stream. We present an algorithm that solves the problem in $O(mn)$ time using $O(mh)$ space, where m and n are the sizes of P and T, respectively, and h is the height of T. We adopt the bit-parallel technique to obtain a practically fast algorithm.

19. 石野 明，竹田 正幸:
"パスプルーニングによる決定性有限オートマトンを用いた XQuery 処理の提案", DBSJ Letters, Vol. 4, No. 4, pp. 17–20, 2006.
概要: 本稿では，文字列検索技術を応用した単純で高速な XQuery 処理手法を提案する．XML データから Aho-Corasick 法を拡張した XML データ走査器によって高速にタグやキーワードの検出を行い，パスに応じた処理を行うことで XQuery の質問式を処理する．特にパスの照合を高速に処理するために，XML データから得られるパススキーマを用いることで決定性有限オートマトンを構築するパスプルーニングを提案する．また，実際に本手法の実装を行い，その実験結果について述べる．

20. 鹿島 久嗣, 坂本 比呂志, 小柳 光生:
"木構造データに対するカーネル関数の設計と解析", 人工知能学会論文誌, Vol. 21, No. 1, pp. 113–121, 2006.
概要: 本研究では，畳み込みカーネルのアイディアに基づいた，ラベル付き順序木に対する新しいカーネル関数を提案する．まず，畳み込みカーネルの枠組みにおいてラベル付き順序木に対して任意の部分グラフを部分構造として用いた場合の，効率の良いカーネル計算のアルゴリズムを提案し，曖昧なラベルや構造を取り込むような拡張を行った．さらに，

より一般的な木構造として，順序のないラベル付き根付き木に対するカーネルを考えた場合には，カーネルの計算が#P-完全問題であることを示した．

21. S. Inenaga and M. Takeda:
"On-line Linear-time Construction of Word Suffix Trees", Proc. 17th Annual Symposium on Combinatorial Pattern Matching (CPM2006), Lecture Notes in Computer Science 4009, pp. 60–71, 2006 年 7 月.
概要: Suffix trees are the key data structure for text string matching, and are used in wide application areas such as bioinformatics and data compression. Sparse suffix trees are kind of suffix trees that represent only a subset of suffixes of the input string. In this paper we study word suffix trees, which are one variation of sparse suffix trees. Let D be a dictionary of words and w be a string in D^+, namely, w is a sequence $w_1 \ldots w_k$ of k words in D. The word suffix tree of w w.r.t. D is a path-compressed trie that represents only the k suffixes in the form of $w_i \ldots w_k$. A typical example of its application is word- and phrase-level search on natural language documents. Andersson et al. proposed an algorithm to build word suffix trees in $O(n)$ expected time with $O(k)$ space. In this paper we present a new word suffix tree construction algorithm with $O(n)$ running time and $O(k)$ space in the worst cases. Our algorithm is on-line, which means that it can sequentially process the characters in the input, each by each, from left to right.

22. S. Inenaga and M. Takeda:
"Sparse Compact Directed Acyclic Word Graphs", Proc. Prague Stringology Conference 2006 (PSC2006), pp. 195–211, 2006 年 8 月.
概要: The suffix tree of string w represents all suffixes of w, and thus it supports full indexing of w for exact pattern matching. On the other hand, a sparse suffix tree of w represents only a subset of the suffixes of w, and therefore it supports sparse indexing of w. There has been a wide range of applications of sparse suffix trees, e.g., natural language processing and biological sequence analysis. Word suffix trees are a variant of sparse suffix trees that are defined for strings that contain a special word delimiter #. Namely, the word suffix tree of string $w = w_1 w_2 \ldots w_k$, consisting of k words each ending with #, represents only the k suffixes of w of the form $w_i \ldots w_k$. Recently, we presented an algorithm which builds word suffix trees in $O(n)$ time with $O(k)$ space, where n is the length of w. In addition, we proposed sparse directed acyclic word graphs (SDAWGs) and an on-line algorithm for constructing them, working in $O(n)$ time and space. As a further achievement of this research direction, this paper introduces yet a new text indexing structure named sparse compact directed acyclic word graphs (SCDAWGs). We show that the size of SCDAWGs is smaller than that of word suffix trees and SDAWGs, and present an SCDAWG construction algorithm that works in $O(n)$ time with $O(k)$ space and in an on-line manner.

23. Y. Higa, H. Bannai, S. Inenaga, and M. Takeda:
"Reachability on Suffix Tree Graphs.", Proc. Prague Stringology Conference 2006 (PSC2006), pp. 212–225, 2006 年 8 月.
概要: We analyze the complexity of graph reachability queries on ST-graphs, defined as directed acyclic graphs (DAGs) obtained by merging the suffix tree of a given string and its suffix links. Using a simplified reachability labeling algorithm presented by Agrawal et al. (1989), we show that for a random string of length n, its ST-graph can be preprocessed in $O(n \log n)$ expected time and space to answer reachability queries in

$O(logn)$ time. Furthermore, we present a series of strings that require $\Theta(n\sqrt{n})$ time and space to answer reachability queries in $O(\log n)$ time for the same algorithm. Exhaustive computational calculations for strings of length $n \leq 33$ have revealed that the same strings are also the worst case instances of the algorithm. We therefore conjecture that reachability queries can be answered in $O(\log n)$ time with a worst case time and space preprocessing complexity of $\Theta(n\sqrt{n})$.

24. S. Inenaga and M. Takeda:
"Sparse Directed Acyclic Word Graphs", Proc. 13th International Symposium on String Processing and Information Retrieval (SPIRE2006), Lecture Notes in Computer Science 4209, pp. 61–73, 2006 年 10 月.
概要: The suffix tree of string w is a text indexing structure that represents all suffixes of w. A sparse suffix tree of w represents only a subset of suffixes of w. An application to sparse suffix trees is composite pattern discovery from biological sequences. In this paper, we introduce a new data structure named sparse directed acyclic word graphs (SDAWGs), which are a sparse text indexing version of directed acyclic word graphs (DAWGs) of Blumer et al.We show that the size of SDAWGs is linear in the length of w, and present an online linear-time construction algorithm for SDAWGs.

25. Y. Higa, S. Inenaga, H. Bannai, and M. Takeda:
"A New Family of String Classifiers Based on Local Relatedness", Proc. 9th International Conference on Discovery Science (DS2006), Lecture Notes in Artificial Intelligence 4265, pp. 114–124, 2006 年 10 月.
概要: This paper introduces a new family of string classifiers based on local relatedness. We use three types of local relatedness measurements, namely, longest common substrings (LCStr's), longest common subsequences (LCSeq's), and window-accumulated longest common subsequences (wLCSeq's). We show that finding the optimal classier for given two sets of strings (the positive set and the negative set), is NP-hard for all of the above measurements. In order to achieve practically efficient algorithms for finding the best classifier, we investigate pruning heuristics and fast string matching techniques based on the properties of the local relatedness measurements.

26. S. Maruyama, H. Miyagawa and H. Sakamoto:
"Improving Time and Space Complexity for Compressed Pattern Matching", Proc. of The 17th International Symposium on Algorithms and Computation (ISAAC2006), Lecture Notes in Computer Science 4288, pp. 484–493, 2006 年 12 月.
概要: The compressed pattern matching problem is to find all occurrences of a given pattern in a compressed text. In this paper an efficient grammar-based compression algorithm is presented for the compressed pattern matching. The algorithm achieves the worst-case approximation ratio $O(g_* \log g_* \log n)$ for the optimum grammar size g_* with an input text of length n. This upper bound improves the complexity of the compressed pattern matching problem to $O(g_* \log g_* \log m + \frac{n}{m} + m^2 + r)$ time and $O(g_* \log g_* \log m + m^2)$ space for any pattern shorter than m and the number r of pattern occurrences.

27. R. Nakamura, H. Bannai, S. Inenaga, and M. Takeda:
"Simple Linear-Time Off-Line Text Compression by Longest-First Substitution", Data Compression Conference 2007 (DCC2007), pp. 123–132, 2007 年 3 月.

概要: We consider grammar based text compression with longest first substitution, where non-overlapping occurrences of a longest repeating substring of the input text are replaced by a new non-terminal symbol. We present a new text compression algorithm by simplifying the algorithm presented in {S. Inenaga, T. Funamoto, M. Takeda, and A. Shinohara, Linear-time off-line text compression by longestfirst substitution, Proc. SPIRE2003}. We give a new formulation of the correctness proof introducing the sparse lazy suffix tree data structure. We also present another type of longest first substitution strategy that allows better compression. We show results of preliminary experiments comparing grammar sizes of the two versions of the longest first strategy and the most frequent strategy.

28. H. Kobayashi, T. Osaki, E. Williams, A. Ishino, and A. Shinohara:
"Autonomous Learning of Ball Trapping in the Four-legged Robot League", RoboCup 2006: Robot Soccer World Cup X, Lecture Notes in Computer Science 4434, pp. 86–97, September, 2007.
概要: This paper describes an autonomous learning method used with real robots in order to acquire ball trapping skills in the four-legged robot league. These skills involve stopping and controlling an oncoming ball and are essential to passing a ball to each other. We first prepare some training equipment and then experiment with only one robot. The robot can use our method to acquire these necessary skills on its own, much in the same way that a human practicing against a wall can learn the proper movements and actions of soccer on his/her own. We also experiment with two robots, and our findings suggest that robots communicating between each other can learn more rapidly than those without any communication.

29. H. Kobayashi, K. Hatano, A. Ishino, and A. Shinohara:
"Reducing Trials by Thinning-Out in Skill Discovery", ,, Proc. The 10th International Conference on Discovery Science (DS2007), Lecture Notes in Computer Science 4755, pp. 127–138.
概要: , In this paper, we propose a new concept, thinning-out, for reducing the number of trials in skill discovery. Thinning-out means to skip over such trials that are unlikely to improve discovering results, in the same way as "pruning" in a search tree. We show that our thinning-out technique significantly reduces the number of trials. In addition, we apply thinning-out to the discovery of good physical motions by legged robots in a simulation environment. By using thinning-out, our virtual robots can discover sophisticated motions that is much different from the initial motion in a reasonable amount of trials.

30. W. Matsubara, S. Inenaga, A. Ishino, A. Shinohara, T. Nakamura, and K. Hashimoto:
"Computing longest common substring and all palindromes from compressed strings", Proc. The 34th Annual Conference on Current Trends in Theory and Practice of Informatics (SOFSEM2008), Lecture Notes in Computer Science 4910, pp. 364–375, January 2008.
概要: This paper studies two problems on compressed strings described in terms of *straight line programs* (*SLPs*). One is to compute the length of the longest common substring of two given SLP-compressed strings, and the other is to compute all palindromes of a given SLP-compressed string. In order to solve these problems efficiently (in polynomial time w.r.t. the compressed size) decompression is never

feasible, since the decompressed size can be exponentially large. We develop combinatorial algorithms that solve these problems in $O(n^4 \log n)$ time with $O(n^3)$ space, and in $O(n^4)$ time with $O(n^2)$ space, respectively, where n is the size of the input SLP-compressed strings.

研究会等

1. 小林　隼人，井上　淳，石野　明，篠原 歩:
“スクリプト言語 Lua を用いたロボカップ四足ロボットリーグシミュレータ”, 人工知能学会第 21 回 SIG-Challenge 研究会, 2005 年 5 月.
概要: In RoboCup four-legged robot league, no modifications or additions to the robot hardware are allowed; hence, there is no difference between teams without software. For optimum results, to improve the software development process is critical. In this paper, we introduce a method of using the scripting language Lua in soccer robots, and we show the soccer simulator, in which Lua scripts are executed as in robots. Using the scripts in both the environments has made it easy to develop soccer robots.

2. 柳町　修平，大崎　嗣豊，小林　隼人，石野　明，篠原　歩:
“実ロボットによる揺動運動のスキル獲得”, 人工知能学会第 62 回人工知能基本問題研究会, 2006 年 3 月.
概要: This paper presents acquisition of swing movement skills by using only acceleration sensors on real robots. Acquisition of swing movement skills is an interesting theme of controlling real robots. First, we create a hand-coding policy for swing movement and confirm that our robots can control by using only some acceleration sensors. Second, we make our robots learn in such a restricted environment. Finally, we compare results of the hand-coding policy with those of reinforcement learning.

3. 御手洗秀一，石野　明，竹田　正幸:
“XML 文書フィルタリングのための軽量な高速化技法”, 電子情報通信学会データ工学ワークショップ論文集, L5-4, 2007 年 3 月.
概要: XML データに極めて単純な前処理を施すことで，質問数の増大に頑健な XML フィルタリング方式を開発した．その手法は，データ配信側で XML データを事前に走査し，根から葉へ向かうパスに沿ったタグ名の列全体を表すパストライと，XML 文書中の開始タグと終了タグを特殊な記号で置き換えたバイナリ XML データを作成する．開始タグを表すコードの直後には，タグを表す番号ではなく，対応するパストライ節点の番号を埋め込む．データ受信側では，このパストライとバイナリ XML データを受け取るものとする．質問式に現れるパスパターンの照合は，パストライを対象として行う．パストライは XML データに比べ非常に小さいため，高速に処理を完了することが可能である．パスパターンの照合結果はバイナリ XML データ走査時に参照することができ，これをテキスト節点を対象としたキーワード照合の結果と結合させることで質問処理は完了する．代表的なストリーム型 XML プロセッサである XMLTK との比較において，実行速度で 4〜6 倍高速であり，メモリ使用量は 1/6 程度であることが判明した．

4. 奥山　哲郎，Masayoshi Sugimoto，大崎　嗣豊，石野　明，篠原　歩:
“ロボカップ４足サッカーロボットのためのファジィ論理を用いた意思決定エンジン”, 人工知能基本問題研究会, SIG-FPAI-65, pp. 65–70, 2007 年 3 月.
概要: In this paper, we introduce a decision engine for autonomous robots in RoboCup fourlegged robot league. The whole system is a platform which can be used for all programs which need a decision process between numerical entries and a set of actions. Robots are particularly appropriate to use a decision process because they have to choose a motion accordingly to their environment. We have implemented the de-

cision engine in agent players. The rules for the decision engine are quite simple and rapid to construct compared to our previous if-then-rule-based system. In this way, each developer can create and improve behaviors solely.

5. 佐々木 崇人，大崎 嗣豊，石野 明，篠原 歩: “無限 n-ボナッチ文字列の繰り返し構造について”, コンピュテーション研究会, 電子情報通信学会技術報告 Vol. 107, No. 24, pp.51–61, 2007 年 4 月.
概要: 隣接 2 項の漸化式で表されるフィボナッチ文字列を一般化し，隣接 n 項の漸化式で n-ボナッチ文字列を定義する．本研究では，任意の $n \geq 2$ に対して，無限 n-ボナッチ文字列における第 k 項目の有限 n-ボナッチ文字列の最大連続出現数は $k \to \infty$ のとき $2 + 1/(\phi^{(n)} - 1)$ に収束することを証明する．ここに，$\phi^{(n)}$ は n-ボナッチ定数であり，特に $\phi^{(2)} = (1 + \sqrt{5})/2$ は黄金比として知られている．さらに，n を増加させると，最大連続数はちょうど 3 に近づくことを示す.

学会大会等

1. 石田　祐介，稲永　俊介，篠原　歩，竹田　正幸:
“漸増的最長共通部分列問題”, 日本応用数理学会 2005 年度年会　講演予稿集, pp. 328–329, 2005 年 9 月.

2. 舞田哲哉, 坂本比呂志:
“算術符号化を利用した XML データへの高速な問い合わせ処理の改良”, 電子情報通信学会第 17 回データ工学ワークショップ (DEWS2006), 7B-o4, 2006 年 3 月.

3. 畑瀬　卓司，大橋　健士郎，下薗　真一:
“幾何点列近似照合を用いたタンパク質二次構造近似検索の実現”, 火の国情報シンポジウム 2007, B-2-4, 2007 年 3 月.

4. 里中　政儀，下薗　真一:
“SQL による評価関数の外部定義が可能な最適時間割作成システム”, 火の国情報シンポジウム 2007, C-2-5, 2007 年 3 月.

5. ミュラ アーノルド，下薗 真一，篠原 武:
“空間索引を用いたバイナリプログラムの高速照合”, 火の国情報シンポジウム 2007, C-5-3, 2007 年 3 月.

6. 今村 安，伸市来亮，二宮大輔，下薗　真一，篠原　武:
“文字列編集距離の成分分解を用いた類似文字列検索辞書の実現”, 火の国情報シンポジウム 2007, C-5-4, 2007 年 3 月.

7. 篠原　歩:
“文字列データからのパターン発見”, 日本機械学会ロボティクス・メカトロニクス部門 平成１９年度第２回地区特別講演会「知能発現と知識発見のロボティクス」, 2007年11月.

文部科学省科学研究費補助金 特定領域研究 (課題番号 16092223)

A10: 変移する要素間の関係を条件とする組合せ最適化モデル

本研究では，入力要素の値が時間とともに連続的に変移する環境において，要素間の関係が時々刻々と変化するような様々な問題を扱い，数理最適化モデルとして定式化する統一的な手法の検討と，精度の高い近似解及び高速性能を保証するアルゴリズムの開発を目的とする．検討を行った具体的な問題は，移動系における経路最適化問題，オンライン環境における経路最適化問題，ディスク・メモリ間のデータ・ブロック転送に対するオフライン計算モデル，データ分類におけるバンプハンティング問題とその確率的解法と統計的評価法，グラフ有向化の頂点出次数最適化問題のオフラインモデル，ブックマーク最適化問題のオフラインモデル，生産順序最適化問題のオフラインモデルとオンラインモデル，終了時刻最適化問題の確率的モデル，2 クリーク最適化問題のオフラインモデル，時間割作成最適化問題である．

研究組織

研究代表者： 宮野 英次　九州工業大学 情報工学部
研究分担者： 朝廣 雄一　九州産業大学 情報科学部
廣瀬 英雄　九州工業大学 情報工学部 (平成 17–19 年度)

交付決定額 (配分額)

平成 16 年度	2,200,000 円
平成 17 年度	4,100,000 円
平成 18 年度	3,800,000 円
平成 19 年度	3,000,000 円
合　計	13,100,000 円

研究成果の概要

- 学術誌
  IEICE (2006)，DAM (2006 年)，IJFCS (2007 年)，TOCS (2008 年予定)
- 国際会議
  WMSCI (アメリカ，2005 年)，IFORS (アメリカ，2005 年)，ICTCS (イタリア，2005 年)，CATS (オーストラリア，2006 年)，CIAC (イタリア，2006 年)，CITSA (アメリカ，2006 年)，SOFSEM (チェコ，2007 年)，FUN (イタリア，2007 年)，AAIM (アメリカ，2007 年)，CIT (日本，2007 年)，MFCS (チェコ，2007 年)，CATS (オーストラリア，2008 年) など
- 国際会議等のプログラム委員
  APDCM 2005, APDCM 2006, APDCM 2007, APDCM 2008, ISPA 2006, ISPA 2007, HICSS 2007, I-SPAN 2008
- 受賞
  2005 年度 IEEE Reliability Society Japan Chapter Award (Best Paper)

1　研究目的

本研究では，入力要素の値が時間とともに連続的に変移する環境において，要素間の関係が時々刻々と変化するような様々な問題を扱い，数理最適化モデルとして定式化する統一的な手法の検討と，精度の高い近似解及び高速性能を保証するアルゴリズムの開発を目的とする．

現実世界に現れる多くの情報工学的問題は，離散的な構造を有しており，多くの場合に組合せ計画モデルとして定式化できる．例えば，車両配送計画問題を従来の数理計画モデルに従って定式化する場合には，車両台数，運搬量，配送コスト，締め切り等の制約となる属性パラメータ値を入力要素として捉え，完全かつ固定的に，与えられた要素間の関係ならびに制約条件を数式化する．配送計画作成アルゴリズムは，事前に作成した条件に基づき作業計画を決定することになる．しかし，近年の社会環境の変化は激しく，時間をかけて作成した完全な作業計画が時をおかずに陳腐化してしまうという状況があり，従来の固定的に制約が与えられる計算モデルでは対応できないという問題が生じている．この事態を打開する期待が持てるものとして，オンライン計算モデルが活発に研究されている．オンライン計算モデルでは計算過程の中で必要な情報が徐々に得られ，かつ予測はできないので，随時対応していく形式で処理が進められる．しかし，現実の社会活動に現れる問題を見てみると，問題を解く前に完全な値はわからないまでも，統計的または経験的にある程度の予想を立てることが可能な場合も多い．例えば車両配送計画における車両台数は，故障や追加配備により動的に変化する可能性があるが，制約の変化に随時対応するのではなく，予め故障頻度などの予想を制約に考慮したモデル化を行うことで，将来の変化に対して柔軟な配送計画を作成することが可能になる．

本研究が目標とするのは，以上のような状況に対するモデル化手法とアルゴリズム開発であり，従来の数理計画モデルとオンライン計算モデルの溝を埋めるものと位置づけられる．本研究成果報告では，移動系における経路最適化問題，オンライン環境における経路最適化問題，ディスク・メモリ間のデータ・ブロック転送に対するオフライン計算モデル，データ分類におけるバンプハンティング問題とその確率的解法と統計的評価法，グラフ辺有向化の頂点出次数最適化問題のオフラインモデル，ブックマーク最適化問題のオフラインモデル，生産順序最適化問題のオフラインモデルとオンラインモデル，終了時刻最適化問題の確率的モデル，2 クリーク最適化問題のオフラインモデル，時間割作成最適化問題に関して，それぞれの最適化モデルとしての定式化，研究成果概要について述べる．

2　移動系における経路最適化問題

概要． 巡回セールスパーソン問題 (TSP) は，計算機科学の分野で古くから知られる問題の一つである．行商人は入力として与えられる都市をすべて巡回して出発位置に戻ってこなければならないが，その時通った道のりの総和をできるだけ小さくしたい．TSP については多くの研究成果が知られておりいまだに活発な研究対象である．本研究課題では，TSP を動的な環境へ拡張した，すなわち巡回対象が都市のように静止しておらず移動している問題である移動物体回収問題を扱っている．これまでに得られた研究成果については，電子情報通信学会・2007 年 3 月号において解説記事として掲載された (学術論文 [13]).

2.1　はじめに

以下のジョブ・スケジューリング問題を考える．ジョブの集合を $J_1, \ldots, J_n$ とする．ここで, ジョブ J_i の実行にかかるコスト $c_i(t, j)$ は実行開始時刻 t と J_i の直前にスケジュールされるジョブ J_j から決まるものとする．任意の i, j, t に対して $c_i(t, j) = c_j(t, i) = d_{ij}$ (d_{ij} は正の定数) とすると, 実行にかかる総コストを最小化する問題は有名な $\mathcal{NP}$-困難問題の一つである巡回セールスパーソン問題 (TSP) である．コスト $c_i(t, j)$ が i と j だけでなく実行開始時刻 t にも依存する場合は, 都市の位置が時刻に依存して変わる TSP 問題に対応し, Moving TSP と呼ばれている．

Moving TSP でモデル化される事例をいくつか見てみよう．自然な例としては配送経路を最適化する問題がある．ただし, 経路の混み具合は時刻

に依存して変化するので (配送時間のような) コストは時刻に依存する．例えば, 工場内に複数のベルトコンベアがあり, その上を製品が流れているとしよう．欠陥のある製品を発見したときに, ベルトコンベアを止めずにロボットを用いて欠陥製品を回収することにした．このとき, 与えられた欠陥製品の集合を最小時間で回収する問題は Moving TSP でモデル化できる．壮大な事例が好きならば, 宇宙ステーション間を宇宙船で行商する行商人を考えることもできる．

Moving TSP ではコストは有限であると仮定されている．しかし, 時刻に依存してコストが無限大になるような事例も考えられるだろう．たとえば, デッドラインが決まっている場合である．このような場合には, すべてのジョブを実行するための総コストを最適化する問題と共に (以前に), 有限コストで実行可能なジョブ数を最大化する問題が重要になってくる．このような問題の一つに本研究課題で扱っている移動物体回収問題がある．

移動物体回収問題を詳しく説明する．n 個の移動物体が座標平面上に存在し，それぞれの経路と速さは事前に与えられる．移動ロボット R によりこれらの物体を回収する．ここで，R と物体 b が同じ時刻に同じ点に存在したら R が b を回収したと考える．物体がすべて停止しており，かつ R は全方向に移動可能という条件下では TSP と同じ問題になる．つまり，移動物体回収問題は TSP を拡張した問題であり一見して難しそうである．

先に述べたようにデッドラインが決まっている場合や，R の速さが物体よりも遅い場合には，すべての物体を回収できないかもしれない．そのため，R の移動能力や物体の経路の条件により 2 種類の問題設定がある：(i) できるだけ早くすべての物体を回収 (総コストを最小化) することが目的の時間最小化と，(ii) できるだけ多くの物体を回収することが目的の回収数最大化．(i) は TSP に対する単純かつ自然な拡張であるが，(ii) は R の移動能力に対する制約を設けることで，問題の本質が大きく異った状況になる．以降では，本研究課題で得られた (ii) に関する研究成果 (学術論文 [10]，[21]) を簡単にまとめる．(i) については例えば文献 [1, 2, 3] を参照されたい．

2.2 アルゴリズムの性能評価

移動物体回収問題を解くアルゴリズムは，移動物体の初期配置，経路，速さの情報を用いて R の行程 (まずどの点に移動して，次はどの点に移動して…というスケジュール) を決定する．R の行程を求めることを，単に「解く」と言うことにする．計算機で実行することを想定しているので，アルゴリズムの性能評価のための指標の一つは，物体の個数 n に対してどの程度の時間で R の行程を決定できるか (実行時間と呼ぶ) である．いい加減に説明すると「ほぼ n^2 に比例する時間で決定できる」ということを 「$O(n^2)$ 時間で実行できる」などと言う．注意して欲しいのは，アルゴリズムの実行時間とは R の行程を計算機で求めるために必要な時間であり，R の回収作業自体にかかる時間ではない点である．

上で「指標の一つ」と断った理由を述べる．例えば，TSP に対して理論的に最良な行程を得るのは非常に困難である．専門的には「$\mathcal{NP}$ 困難である」と言うが，NP 困難な問題に対しては n の多項式程度の時間 (多項式時間) で解く方法が知られていない．例えば 2^n 程度の指数時間かければ解けたりするのだが，そのようなアルゴリズムを計算機上に実装すると，n が数百程度になると非現実的な時間がかかってしまうことになる．

アルゴリズムの性能評価のためのもう一つの指標について説明する．$\mathcal{NP}$ 困難な問題に対する一つのアプローチとして近似アルゴリズムがある．これは，最良の行程は求められないかもしれないが，多項式時間で実行できる．もちろん，余りにも悪い行程を求めてもしょうがないので，近似度による性能保証がついていることが多い：理論的に最良な行程の値を OPT，アルゴリズムが出力した行程の値を A としたときに，OPT/A が常に r 以下であるとき，このアルゴリズムは近似度 r であるという．ここで行程の値とは回収した物体の個数になる．例えば近似度 2 のアルゴリズムは，回収できる物体数の理論的最大値の半分以上を回収できる行程を出力する．

2.3 回収数最大化

回収数最大化を目的とする，4 つの問題設定とそれぞれに関する研究成果について紹介する．問

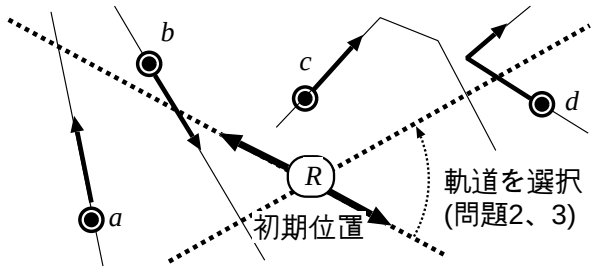


図 1: (問題 1) ロボット R は直線軌道上を移動し，物体 (a, b) も直線上を移動．(問題 2) R の軌道を選択．(問題 3) 1 個回収するたびに R は初期位置に戻り軌道を再選択．(問題 4) 物体 (c, d) が折れ線上を移動．

題 1 と 2 については学術論文 [10]，問題 3 と 4 については学術論文 [21] による．それぞれの問題設定を図 1 に簡単に図示する．

問題 1: R は 1 本の固定された直線 (軌道) 上を往復し，物体はそれぞれ別の直線上を一定の速さで移動する．

結果：すべての物体を回収するには，軌道を通過する時刻の小さい順に物体を整列し，その通過位置に R が随時移動すればよい．軌道が直線に制限されているため全部を回収できないかもしれないが，全部回収できるかを判断し，可能な場合にはその行程を求めるアルゴリズムは $O(n \log n)$ 時間で実行できる．すべての物体を回収できない時は，物体間の関係をうまくモデル化することで最大個数回収する行程を求めるアルゴリズムを構築でき，その実行時間は $O(n^2)$ である．また工夫により，$O(n \log n)$ 時間で実行できるアルゴリズムも構築できる．

問題 2: 問題 1 では固定されていた軌道を R は自由に選ぶことができる．軌道を決めれば，その軌道を用いて最大個数回収する行程は問題 1 の方法で得られるので，できるだけ多くの物体を回収できる軌道を求めたい．

結果：R の初期位置を通る直線すべてが軌道の候補になり無限に存在しそうである．しかし実は，n^2 個程度の直線を調べれば最大個数を回収できる軌道を発見でき，結果として R の行程を $O(n^3 \log n)$ 時間 ($n^2 \times n \log n$ 程度) で求めることができる．さらに工夫することで，$O(n^2 \log n)$ 時間で実行できるアルゴリズムも構築できる．

問題 3：R は物体を 1 個回収するたびに初期位置に戻る必要がある．ただし，先の 2 つの問題と異り，初期位置に戻るたびに R は軌道を回転できる．R の速さは物体よりも遅く，物体の経路は問題 1，2 と同様に直線とする．

結果：初期位置に戻ってくる時刻に関する貪欲法，すなわち未回収物体中で最も早く原点に戻ってこれるものから順に回収するアルゴリズムで最大個数の回収ができる．その実行時間は単純な実現方法だと $O(n^2)$ であるが，工夫により $O(n \log n)$ に減らすことができる．

問題 4：物体の経路として直線ではなく折れ線を考える．それ以外は問題 3 と同様．

結果：直感的には直線の場合よりも，折れ線の方が難しそうに思える．本当に難しいのかが興味の 1 つであるが，各物体が 1 回以上曲がると $\mathcal{NP}$ 困難になる．問題 3 で説明した貪欲法と同じ方針のアルゴリズムが近似度 2 であり，その実行時間は $O(mn^2)$ である．ここで m は最もたくさん曲がる物体の，その曲がる回数である．

文献

[1] P. Chalasani, R. Motowani, and A. Rao, “Approximation algorithms for robot grasp and delivery,” Proc. International Workshop on Algorithmic Foundations of Robotics, pp.347-362, 1996.

[2] M. Hammar and B.J. Nilsson, “Approximation results for kinetic variants of TSP,” Lecture Notes in Computer Science, vol.1644, pp. 392-401, 1999.

[3] C.S. Helvig, G.Robins, and A.Zelikovsky, “The moving-target traveling salesperson problem,” J. Algorithms, vol.49, no.1, pp.153-174, 2003.

3 オンライン環境における経路最適化問題

概要． オンライン環境における経路最適化問題とは，探索者が未知のグラフの全ての頂点を訪問することによりグラフ構造を調査する問題である．探索者は辺の存在とその長さをその端点を訪れるまで判らないとする．探索者は，できるだけ短い経路を通ることにより全ての頂点と辺を調査して，出発点まで戻って来なければならない．本研究課題では，このオンライン環境における経路最適化問題に対するオンラインアルゴリズムの設計とその競合比解析を行い，得られた研究成果については，第 33 回 Conference on Current Trends in Theory and Practice of Informatics (SOFSEM 2007) において公表を行った (学術論文 [11]).

3.1 はじめに

TSP は，オペレーションズ・リサーチや組合せ最適化の分野において最も有名な問題の一つである．TSP では，頂点数，辺の長さ，頂点の隣接関係・ネットワーク構造などの入力例題の完全な環境・情報が予め与えられており，それらを TSP の巡回者が巡回経路を決定するときに利用することができる．巡回者の目標は，それらの完全情報を利用して，移動距離を最短にするような経路を発見することである．しかし，このような**オフラインモデル** では実世界の問題を表現できない場合がある．すなわち，多くの経路選択問題やスケジューリング問題では，経路決定を行うための環境に関する全体情報が前もっては与えられずに，不完全な一部の情報のみが**オンライン**に与えられる場合がある．

上のような未知の環境で最適解を求めるオンライン経路選択問題は，**探索問題** (exploration problem) や**地図作成問題** (map construction) として知られている [1]．本稿では，次のようなグラフ上での地図作成問題を考える．本問題は，文献 [3] において，Kalyanasundaram と Pruhs によって定式化された問題である．探索者 (他の応用においては巡回者，ロボット，またはエージェントと呼ばれることがある) ができるだけ短い経路を移動することにより，未知の環境の完全な地図を構成することを考える．多くの場合に，この未知の環境は重み付きグラフ G で形式化することができる．探索者はある特定の出発点から地図作成を始める．探索者が頂点 v を訪問するとき，v を端点とする辺とその辺によって隣接する頂点までの距離情報を得る．頂点の個数や頂点間の隣接関係を知らないため，環境全体を調べるためには，探索者はすべての頂点を少なくとも一度は訪れなければならない．探索者は十分に多くの情報を記憶することができると仮定する．これにより，既に訪問した頂点やその頂点を端点とする辺を再び調べたときには，既に訪問済みであることが判るとする．従って，すべての頂点を訪問したことを知ることができ，探索者は最後に出発点に戻って来る．オンラインアルゴリズムの評価には**競合比**を用いる [2]．競合比は，すべての頂点を訪問し最後に出発点に戻って来る巡回経路の中で最短の経路の移動距離と，オンラインアルゴリズムが出力する出発点まで戻って来る巡回経路の移動距離の比の最大値である．

オンラインアルゴリズムの最初には，出発点のみが訪問済みであり，他の頂点は未訪問である．次のステップでは，出発点に隣接する頂点の中から 1 つの頂点が選ばれ，探索者は選択した頂点に移動する．アルゴリズムが進むに従って，既知となるグラフ構造が大きくなっていく．すなわち，アルゴリズムのあるステップで実行しなければならないことは，それまでに得られた知識を利用することによって，訪問済みの領域に隣接する未訪問頂点の中から，どの新しい頂点に移動すべきであるかを決定することである．最も単純な戦略の 1 つは，最近傍アルゴリズム (NN) である．NN は常に探索者の現在の位置から最も近い頂点を選択する．ただし，頂点の数を n とすると平面グラフにおいて競合比は $\Omega(\log n)$ まで大きくなることが知られている [4]．他の戦略としては，深さ優先アルゴリズムが有名であるが，サイクルグラフに対してさえ競合比が非常に大きくなってしまう．本問題に対する最初の競合アルゴリズムは，Kalyanasundaram と Pruhs により提案された [3]．彼らは，*ShortCut* と呼ばれる重み付き最近傍アルゴリズム (WNN) を提案し，平面グラフに対して ShortCut アルゴリズムの競合比が 16 になることを示した．

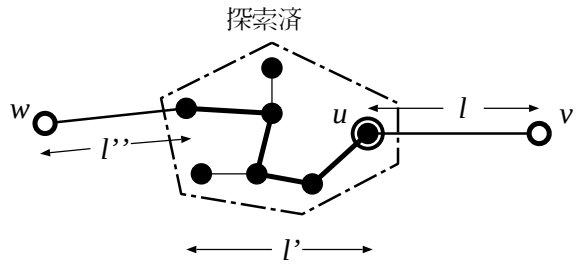


図 2: 重み付き最近傍アルゴリズム

WNN は固有の正定数 Δ の値に従って動作する．すなわち，WNN は Δ を用いた重み付き距離を用いて，次に訪問すべき頂点を決定する．図 2を用いてアイデアを説明する．探索者が，ある時点で，頂点 u にいるとする．また，頂点 v は，長さが ℓ であるような辺 (u, v) により頂点 u に隣接していると仮定する．さらに，他にまだ訪問していない新しい頂点 w を訪れるために，既に通ったことのある長さ ℓ' の経路と頂点 w を端点とする長さが ℓ'' の辺を通らなければならないと仮定する．重み付き最近傍戦略に従って動作するとき，探索者は $\Delta\ell \leq \ell' + \ell''$ ならば頂点 v を訪れ，そうでなければ，頂点 w を訪れる．上述の平面グラフに対する 16 競合比は，$\Delta = 3$ とすることによって得られる．すなわち，平面グラフの場合には，目前の新しい未訪問の頂点に貪欲的に進むよりも，以前通った経路を逆戻りして新しい未訪問頂点に行くことを優先するという戦略により，良い競合比を得られるということを意味している．

3.2 結果

入力グラフ G がもし木であるならば深さ優先アルゴリズムによって最適な経路を得ることができることは容易に分かる．しかし，もし入力がサイクルを持つグラフであるならば，この問題は単純ではなくなる．本稿では，WNN をサイクルグラフに適用することにより，WNN の能力を調べることを目的とする．本稿の結果は以下である．

- NN (すなわち，$\Delta = 1$ とした WNN) は 1.5 競合アルゴリズムである．
- NN の競合比が 1.5 となるようなサイクルグラフが存在し，我々の 1.5 競合比の解析は厳密である．
- WNN の中で，重みを $\Delta = 1$ とおくことが最適である．すなわち，Δ の値を $\Delta > 1$ または $0 < \Delta < 1$ とすると，WNN の競合比は少なくとも 1.5 となる．
- 任意の決定性オンラインアルゴリズムは，1.25 よりも小さい競合比を持たない．

文献

[1] P. Berman. On-line searching and navigation. In *Online Algorithms: The State of the Art* (Fiat and Woeginger Ed., Springer), pp.232–241 (1998)

[2] A. Fiat and G.J. Woeginger. Competitive analysis of algorithms. In *Online Algorithms: The State of the Art* (Fiat and Woeginger Ed., Springer), pp.1–12 (1998)

[3] B. Kalyanasundaram and K.R. Pruhs. Constructing competitive tours from local information. *Theoretical Computer Science* **130**, pp.125–138 (1994)

[4] D.J. Rosenkrantz, R.E. Stearns, P.M. Lewis. An analysis of several heuristics for the traveling salesman problem. *SIAM J. Computing* **6** (3), pp.563–581 (1997)

4 データ分類におけるバンプハンティング

概要． 顧客データベースから一定のルールに従う優良顧客 (あるいは非優良顧客) を抜き出すことを考え，本研究課題では，雑然として分類が一見困難と思われる顧客データベースからできるだけ有用な情報を引き出すようなバンプハンティング・モデルとしての定式化を行った．本研究課題で得られた研究成果については，第 3 回 International Conference on Cybernetics and Information Technologies (学術論文 [7])，IEICE Transactions on Information and Systems (学術論文 [8])，第 6 回および第 7 回 Hawaii International

Conference on Statistics, Mathematics and Related Fields (学術論文 [12, 19])，第 7 回 IEEE International Conference on Computer and Information Technology (学術論文 [18]) で公表を行った．

4.1 モデル

対象データは 0 および 1 の 2 値のレスポンスを持ち，z 次元の説明変数は連続あるいは属性変数である．これらのデータは混在しており，本研究課題での目標は，2 値データを分離する境界を求めるのではなく，レスポンス 1 となるデータの密度の濃い部分を切り出すことである．すなわち，入力としてデータ密度 d を指定し，切り出した領域のレスポンス 1 が含まれる密度が d 以上のできるだけ広い領域を見つけることが目標となる．このような領域をバンプ領域と呼ぶことにする．d が大きくなるとバンプ領域は減少する特性があり，これを獲得密度-獲得数のトレードオフ曲線と呼び，本曲線を求めることも目標の一つである．

4.2 結果

本研究課題では，ある顧客データベースの特徴に近い人工的なデータケース，およびリアルデータに対して，バンプ領域の切り出しを行う際に決定木を基本手法とし，これに確率的探索法である遺伝的アルゴリズムを組み合わせた方法を用いた．また，確率的手法の収束値の性格から，いくつかの収束値を用いて極値統計を応用して，最適に近い解の精度を保証した上で，獲得密度-獲得数のトレードオフ曲線の近似解を求める方法を提案した．

さらに，上で述べた方法により得られたトレードオフ曲線の精度について詳しく調べた．トレードオフ曲線を求める際，遺伝的アルゴリズムにおける世代交代では各世代でテストデータを評価する方法を用い，極値統計手法がトレードオフ曲線の上限を精度よく保証するというアルゴリズムを開発した．計算コストを押さえるため，テストサンプル法とブートストラップ法のメリットを併せ持つ方法を提案した．

5 ブロック転送問題に対するオフライン計算モデル

概要． 本研究課題では次のようなグラフのレイアウト問題の一種を考える．入力として有向非巡回グラフ (DAG) が与えられたとき，DAG 上の根から葉までのすべての道を見て，その外部有向辺数の最大値を最小にするように，それぞれのノードをサイズが高々B であるブロックに分割することを目標とする．ここで，外部有向辺とは異なるブロックに属すノード間の有向辺とする．本問題は一般的には $\mathcal{NP}$ 困難となり，B の値が 2 で DAG の高さが 3 の場合にも $\mathcal{NP}$ 困難のままであることを示すことができる．本研究課題では，$B = 2$ で DAG の高さが 2 の場合には多項式時間アルゴリズムが存在することを示した．この多項式時間アルゴリズムを利用することにより，DAG の高さが一般のときに，$B = 2$ の場合には，$\frac{3}{2}$ 近似アルゴリズムが存在し，近似精度保証の観点からこの $\frac{3}{2}$ の近似率は最適であることを示した．これまでに知られた近似率は 2 であった．また，$B \geq 3$ の場合にも，$\mathcal{P} \neq \mathcal{NP}$ と仮定すると，任意の正数 ε に対して，$\frac{3}{2} - \varepsilon$ 近似アルゴリズムが存在しないことを示した．さらに，$B = 3$ の場合，入力グラフを階層グラフに限定すると，2 近似アルゴリズムが存在することを示した．これらの研究成果については，第 6 回 International Conference on Algorithms and Complexity (CIAC 2006) において公表を行った (学術論文 [6])．

6 グラフ有向化の頂点出次数最適化問題

概要． 与えられた無向グラフのすべての辺について方向付けを行うことをグラフ有向化 (orientation) と呼び，グラフ理論や組合せ最適化の分野において古くから研究が行われている．例えば，Robbins は，グラフ有向化とグラフ連結度の関係について調べている [1]．Chvátal と Thomassen は，グラフの直径との関係，および計算複雑さについて研究を行っている [2]．本研究課題では，グラフ有向化の際に，出次数 (または入次数) の大きさを目的関数とするような問題のモデル化を行

い，計算複雑さ及び近似アルゴリズムに関する検討を行った．

頂点集合を V，辺集合を E，辺の重み集合を $w : E \to \mathbb{Z}^+$ とするとき，重みつき無向グラフを $G = (V, E, w)$ とする．ここで，頂点 $u, v \in V$ について，$\{u, v\}$ により辺を表す．グラフ G の有向化 Λ とは，辺 $\{u, v\} \in E$ の向きを決めることである．今年度は，以下のような最大出次数を最小化するグラフ有向化問題 (MMO 問題, Minimum Maximum Outdegree Orientation Problem) としてモデル化を行った．

MMO

入力： 重みつき無向グラフ
$G = (V, E, w)$.

問題： 頂点の最大出次数が最小となるグラフ G の有向化 Λ を求める．

第10回 Computing: The Australasian Theory Symposiumnt (学術論文 [5]) および International Journal of Foundation of Computer Science (学術論文 [14]) の成果は以下である．

- MMO 問題は $\mathcal{NP}$ 困難である．
- 次のような特別な場合には多項式時間で動作する最適なグラフ有向化アルゴリズムを設計することができる．(i) 入力グラフが重み無しグラフの場合．(ii) 入力グラフが木の場合．
- 上記の (i) および (ii) に対する多項式時間アルゴリズムを用いることにより，$\min\{\frac{w_{max}}{w_{min}}, (2-\varepsilon)\}$ の近似率を持つ組合せアルゴリズムを設計することができる．ここで，w_{max} および w_{min} はそれぞれ辺重みの最大値および最小値であり，ε は入力に依存するある正数である．

第33回 International Conference on Algorithmic Aspects in Information and Management (学術論文 [16]) の成果は以下である．

- 辺重みが 1 または $k \geq 2$ ときには，MMO は強 $\mathcal{NP}$ 困難である．
- 辺重みが 1 または $k \geq 2$ ときには，$\mathcal{P} \neq \mathcal{NP}$ の仮定のもとで，$1 + 1/k$ より小さい近似率を持つ多項式時間アルゴリズムを設計することはできない．
- 辺重みの最大値が k のとき $(2 - 1/k)$ の近似率を持つ多項式時間アルゴリズムを設計することができる．
- 辺重みが 1 または k のとき，$k = 2$ のときには 3/2 の近似率を持つ多項式時間アルゴリズムを設計できる．すなわち，近似率の意味で最適なアルゴリズムが設計できる．また，$k \geq 3$ のときには $(2 - 2/(k+1))$ の近似率を持つ多項式時間アルゴリズムを設計できる．

入力グラフに制限をつけた場合に関する検討も行い，第 12 回 Computing: The Australasian Theory Symposiumnt において以下の結果の公表を行った (学術論文 [20])．

- 入力グラフがカクタスの場合，MMO 問題は $\mathcal{P}$ である．
- 入力グラフがカクタスの最小スーパークラスである外部平面グラフの場合弱 $\mathcal{NP}$ 困難である．
- 入力グラフがカクタスの最小スーパークラスである $P_4$2 部グラフの場合弱 $\mathcal{NP}$ 困難である．
- 入力グラフが diamond-free グラフ，house-free グラフ，直並列グラフ，2 部グラフ，平面グラフの場合，いずれも $\mathcal{NP}$ 困難である．

文献

[1] H.E. Robbins, “A theorem on graphs with an application to a problem of traffic control,”

[2] V. Chvátal and C. Thomassen, “Distances in orientation of graphs,” *J. Combinatorial Theory, series B* **24** (1978) 61–75.

7 ブックマーク最適化問題

概要． ハイパーリンクの集合 E で連結されたウェブページの集合 V を表した根付き有向非巡回グラフ $G = (V, E)$ を考える．それぞれのノード v は，ユーザがノード v を閲覧するアクセス頻度を持っている．アクセスコストは，根ノード r からあるノード v に到達するためのステップ数の期待値で定義される．ブックマークは，グラフ G の根ノード r に付加するショートカットリンクであり，アクセスコストを小さくする効果がある．ブックマーク問題は，アクセスコストを最も改善するようなブックマーク集合を求める問題である．本研究課題では，ブックマーク問題に対する近似上限と近似下限に関する検討を行い，得られた研究成果を第32回 International Symposium on Mathematics Foundations of Computer Science において公表した (学術論文 [17])．

7.1 モデル

World Wide Web は，最も有名で巨大な情報源であり，ウェブページとハイパーリンクから構成されている．ハイパーリンクはあるページから別のページへの参照リンクであり，これにより非線形な情報探索を実現している．ユーザはホームページから始めて，ハイパーリンクをクリックして次々にウェブページを表示しながらインターネット上を巡回して，目的のページに辿り着くことができる．できるだけ少ない数のハイパーリンクをクリックすることにより希望するページを表示するために最もよく行われている方法の一つは，ブックマークと呼ばれるショートカットリンクを付加することである．

希望するウェブページへのアクセスを改善するための上で述べたような自然な解決方法は次のような最適化問題として形式化できる．本問題は k ブックマーク問題 (k-Bookmark Assignment Problem，以下では k-BAP と表す) と呼ばれ，文献 [3] で提案されている． 根ノード r を持つ根付き連結有向非巡回グラフ (DAG) を $G = (V, E)$ とし，ウェブの階層構造を表すとする．ノード $v \in V$ が有するアクセス頻度を p_v で表す．ここで，$\sum_{v \in V} p_v = 1$ および $p_v \geq 0$ が成り立つとする．ノード v について，根ノード r からノード v までの距離，すなわち，最短の有向パスの長さを $d(v)$ で表す．根ノード r からあるノードへ到達するまでの期待ステップ数を $\sum_{v \in V} p_v \cdot d(v)$ で定義する．ノード v へのブックマークは付加する有向辺 $(r, v) \notin E$ である．簡単のため，本稿では，ブックマーク (r, v) をノード v 自身で表すことにする．$b_i \in V$ について $B = \{b_1, b_2, \cdots, b_k\}$ をブックマークの集合とし，グラフ G にブックマーク集合 B を追加することで得られたグラフを $G \oplus B$ で表す．グラフ $G \oplus B$ における根ノード r からノード v への距離を $d_B(v)$ で表す．また，$G \oplus B$ において根ノードからあるノードへ到達するための期待ステップ数を $\sum_{v \in V} p_v \cdot d_B(v)$ と表す．ブックマーク集合 B のゲインを次のように表す．

$$g(B) = \sum_{v \in V} p_v \left(d(v) - d_B(v) \right).$$

ブックマーク問題は次のように形式化される．

k ブックマーク問題 (k-BAP):

入力: 有向非巡回グラフ $G = (V, E)$ および正整数 $k \leq |V|$.

目的: ゲイン $g(B)$ を最小化するようなサイズ k 以下のブックマーク集合 $B \subseteq V$ を求める．

ブックマーク問題は，インターネットにおけるウェブプロキシサーバ配置問題として形式化される場合がある．特別なネットワーク構造の場合，例えば単純パスや木の場合には多項式時間で最適解を求めるアルゴリズムが知られている．単純パスについては，Li，Deng，Golin，Sohraby によって，$O(k|V|^2)$ 時間で最適解を求めるアルゴリズムが提案されている [2]．木の場合には，Li，Golin，Italiano，Deng，Sohraby によって，文献 [4] において $O(k^2|V|^3)$ 時間で動作するアルゴリズムが示されている．特に，もし入力グラフが完全二分木であり，$k \leq \sqrt{|V| + 1}$ である場合には，より高速な $O(|V|)$ 時間で動作するアルゴリズムが Czyzowicz，Kranakis，Krizanc，Pelc，Martin により設計されている [1]．一方，Martin は，入力が有向非巡回グラフである場合には k-BAP が $\mathcal{NP}$ 困難となることを示している [3]．これまでに k-BAP に対して近似可能性および近似不可能性に関する結果は示されていない．

7.2 結果

本研究課題においては，k-BAP の近似可能性および近似不可能性について議論を行い，以下の結果を示した．

- 近似比 $(1-1/e)$ を持つ多項式時間近似アルゴリズムが存在する．
- 入力のサイズを N とするとき，$\mathcal{NP} \subseteq \mathcal{DTIME}(N^{O(\log\log N)})$ でないという仮定の下で，$(1-1/e)$ よりも良い近似比を持つ多項式時間近似アルゴリズムは存在しない
- δ をある小さな正定数とするとき，より弱い $\mathcal{P} = \mathcal{NP}$ でないという仮定の下で，$(1-\delta)$ よりも良い近似比を持つ多項式時間近似アルゴリズムは存在しない．

文献

[1] J. Czyzowicz, E. Kranakis, D. Krizanc, A. Pelc, and M.V. Martin. Assigning bookmarks in perfect binary trees. *Ars Combinatoria*, **LXXXII** (2007)

[2] B. Li, X. Deng, M.J. Golin, and K. Sohraby. On the optimal placement of web proxies in the Internet: Linear Topology, In *Proc. HPN'98*, pp.485–495 (1998)

[3] M. V. Martin. Enhancing Hyperlink Structure for Improving Web Performance. *PhD thesis, School of Computer Science, Carleton University* (2002)

[4] B. Li, M.J. Golin, G.F. Italiano, X. Deng, and K. Sohraby. On the optimal placement of web proxies in the Internet. In *Proc. INFOCOMM'99*, pp.1282–1290 (1999)

8 生産順序最適化問題

8.1 ペイントショップ最適化問題

概要． ペイントショップに運ばれてきた自動車の車種の順番を変えることなく塗る色のみを入れ替えることで，色の交換回数を最小化しようというオフライン最適化問題についての検討を行った．本問題は一般的な場合に $\mathcal{NP}$ 完全である．また，入力で各車種がある 2 色を 1 台ずつ注文を受けたという制限を加えた問題 1R2C (1Regular2Color) ペイントショップ最適化問題についても $\mathcal{NP}$ 完全であることが知られている．本研究課題では，1R2C ペイントショップ最適化問題を対象に貪欲法と近似解法の性能評価を行い，得られた研究成果を INFORMS Annural Meeting において公表した (学会大会 [17])．

モデル． 入力としてアルファベット集合 S と $n = 2|S|$ の長さを持つアルファベットの列 w が与えられる．ここで，それぞれのアルファベットは正確に 2 回ずつ w の中に出現しているものとする．このとき，各アルファベットについて，1 つは赤に，もう 1 つは青に塗り，赤から青または青から赤への色交換数を最小化するような塗り方を求める問題を考える．例えば，アルファベット集合 $S = \{a, b, c, d\}$，列 $w = abccbdad$ が入力として与えられたとする．また，例えば，a を赤で塗るとき a，青で塗るとき $\bar{a}$ で表すものとする．出力列として $a\bar{b}c\bar{c}bd\bar{a}\bar{d}$ とすると色交換数は 5 回になるが，$abc\bar{c}\bar{b}d\bar{a}d$ とすると色交換数は 2 回になる．

結果． 1R2C ペイントショップ最適化問題を対象に貪欲法と近似精度に関して以下の結果を得た．

- 任意のアルファベット a の第 1 回目の出現順序を a_1，2 回目の出現位置を a_2 とするとき，任意の 2 つアルファベットの a および b について，$a_1 < b_1 < b_2 < a_2$ または $a_2 < b_1$ となるような入力列 w に対して多項式時間アルゴリズムを設計することができる．
- 一般の入力に対して $O(\sqrt{\log n})$ の近似率を持つ多項式時間アルゴリズムを設計することができる．

8.2 ソーティングバッファ最適化問題

ペイントショップ問題と同様に色の交換回数を目的関数とする最適化問題に関して異なるモデル化を行ったソーティングバッファ最適化問題に関してオンラインモデルとオフラインモデルに関する検討を行った．

オンラインモデル． 以下のオンラインソーティングバッファ問題を扱う．サービスステーションとソーティングバッファがあり，多種類の品物の列が入力として与えられる．ソーティングバッファはランダムアクセス可能なバッファであり，k 個までの品物を (一時的に) 格納できる．サービスステーションにおいてできるだけ同種の品物を連続して処理をするために，ソーティングバッファを利用してで品物の列の置き換えを行う．目的は，ある品物の列から異なる品物の列に変わる回数をできるだけ少なくすることである．形式的には次のようにモデル化できる: 入力として数種類の品物の列 $I = \sigma_1\sigma_2\cdots\sigma_n$ がオンライン的に与えられ，それらの列を k 個の容量を持つソーティングバッファで置換して，品物の列 $S = \rho_1\rho_2\cdots\rho_n$ を出力する．ここでは，ペイントショップ問題と同様に，品物の種類を色と呼ぶことにする．目的は，出力列 S ができるだけ同色の品物が連続して現れるように列の置換を行うことである．ここで，連続する同色の品物の列を**ブロック**と呼ぶことにする．入力 I に対して出力列に含まれるブロック数を $C(I)$ と表すと，オンラインソーティング問題は $C(I)$ の最小化問題となる．

オンラインソーティングバッファ問題に対して，最も標準的で単純な手法として FIFO (First In First Out) がある．FIFO の競合比については $\Omega(\sqrt{k})$ の下限がある [1]．本研究課題では，FIFO の競合比が $O(k)$ であることを示した．また，LIFO (Last In First Out) の競合比の下限が $2k-1$ となることを示した．

文献

[1] H. Räcke, C.Sohler, M.Wetermann. "Online scheduling for sorting buffers," *ESA '02*, 820–832 (2002)

9 終了時刻最適化問題

概要． 終了時刻最適化問題題は，複数のジョブをできるだけ短時間で全ての処理を完了できるように複数のマシンに割り当てるコスト最小化問題である．ここでどのジョブを処理しなければならないのかという完全情報を仮定しているが，実際の場面では，入力ジョブ集合が確率的または統計的に与えられることがある．そこで近年，確率モデルに対する多くの研究成果が報告されている．本研究の目的は未来の変化に対して柔軟なモデルを構築することとアルゴリズムの設計であり，終了時刻最適化問題の確率モデルに関する検討を行った．

終了時刻最適化問題の確率モデルとして，次の 2 段階確率モデルを考える．初めに統計的データを元に処理しなければならないジョブ集合を予測し，それらを事前に処理しておき (第 1 段階)，次に実際与えられたジョブ集合の中で未処理のジョブに対して追加の処理を行う (第 2 段階) というモデルである．なお第 2 段階でジョブを処理する際には，緊急にジョブを処理をする必要があるという意味で，第 1 段階でジョブを処理をするのに比べて $\lambda \geq 1$ 倍の処理時間が必要となってしまうとしてモデル化する．つまりこの総コスト (総処理時間) は (第 1 段階のコスト) + λ(第 2 段階のコスト) と表すことができる．実際にしなければならないジョブ集合は統計的に与えらるので，目的関数を (第 2 段階のコスト) + λ·(第 2 段階のコストの期待値) と定式化し，本研究課題ではこの値を最小化するアルゴリズムを求めることを目的とする．なお評価指標は (アルゴリズムの出力)/(最適値) とする．この最適値は全数探索アルゴリズムによって導出される．すなわち，(第 1 段階のコストの最適値) + λ·(第 2 段階のコストの期待値の最適値) と定義される．本研究課題では，終了時刻最適化問題の確率的モデルに対して，近似率 3 の多項式時間アルゴリズムを設計した．

10 2 クリーク最適化問題

概要． 最大 2 クラン問題に対する近似可能性と近似困難性について考える．2 クランとは，無向グラフ中の直径が 2 であるような部分グラフであり，最大 2 クラン問題とは，無向グラフ $G = (V, E)$ ($|V| = n$) が与えられたとき，含まれる頂点数が最大となるような 2 クランを見つける最適化問題である．これまでに本問題の $\mathcal{NP}$ 困難性が示されているが，近似アルゴリズムや近似困難性に関する議論はない．本研究課題では，グラフ G の最大次数を k とした場合の $(k-1)$ 近似アルゴリズム

を設計，$\mathcal{P} \neq \mathcal{NP}$ の仮定のもとでは，多項式時間で動作する n^{σ} $(\sigma > 0)$ 近似アルゴリズムが存在しないことを示した．

11　時間割作成最適化問題

概要． 本稿で取り扱う問題は，大学における時間割作成問題 (timetabling problem) の一つで，試問予定表作成問題と呼ぶ．本問題では，発表する学生 1 人に付き 3 人の先生が発表を審査している状況を考える．それぞれの先生は複数の学生の審査をするように予め担当学生が決められている．例年は，1 つの会場で論文発表を行っていたが，昨年度は，論文発表を行う学生数が増えたため，2 つの会場で行う必要があった．このときに，審査を行う先生が 2 つの会場を移動することが無いようにスケジュールする必要が出て来た．そのため，先生がなるべく会場を移動せずに発表の採点がスムーズに出来るようにしたい．このような背景から試問予定表作成問題が生まれた．現実に扱われる問題をモデルにしているため，実用性の高い問題である．

試問予定表作成問題に対して試問スケジュールを作成する際には，いずれの先生も同時刻に審査を担当する学生が割り当てられることが無いようにする必要があり，この条件を満たす試問スケジュールを正当なスケジュールと呼ぶことにする．試問予定表作成問題は，学生の試問会場の割り当てに対して，試問する先生の会場移動回数を最小化するような正当なスケジュールを求めることを目的としたコスト最小化問題である．

ここで，先生の移動回数を考慮しなければ，正当なスケジュールを求めること自体は簡単であることに注意しておく．すなわち，正当なスケジュールが存在するならば，それを求めるのは多項式時間で計算可能である．これは，頂点を学生として，その学生を審査する先生が異なるような学生に対応した頂点の間に辺を張ったグラフを考え，このグラフの最大マッチングを求めれば良い．最大マッチングは多項式時間で求めることが可能であり，得られたマッチングが完全マッチング (ただし，頂点数は偶数の場合) であれば，それが一つの正当なスケジュールになっている．

本稿で考える試問予定表作成問題では，先生の移動回数を尺度とした最適化問題を考え，本問題の計算複雑さの解明を目標とする．試問予定表作成問題を考える第一ステップとして，2 つの制約条件を与える．一つ目は，学生の発表時刻は入力のスケジュールから変わらず，同時刻に割り当てられた 2 人の学生の発表会場の交換のみを行うという制約である．この制約により定義された問題を，Room 問題と呼ぶ．2 つ目は，会場の交換は行わず，発表時刻のみを入れ替える制約である．このとき，入力スケジュールでの同時刻の学生ペアは，出力スケジュールでも同時刻でペアにならなければならない．この制約により定義された問題を，Order 問題と呼ぶ．いずれの問題においても，ある先生が，同時刻に異なる部屋において異なる学生を審査するようなことが無いような正当な初期入力割り当てが与えられた場合には，スケジュールの正当性を考慮する必要はない．

本稿では，Room 問題と Order 問題の計算複雑さについて考える．まず，それぞれの先生が担当する学生の数は s 以下であり，それぞれの学生を担当する先生の数は t 以下であると仮定する．このように制限した Room 問題を $\textsc{Room}(s,t)$ 問題とする．$\textsc{Order}(s,t)$ 問題についても同様の制限を加えた問題であるとする．

$\textsc{Room}(1,t)$ と $\textsc{Order}(1,t)$ は各先生の担当する学生が 1 人のみであるので，多項式時間で解くことが出来るのは明らかである．本研究課題の主要な研究成果は以下であり，第 6 回 International Conference on the Practice and Theory of Automated Timetabling (PATAT 2006) において公表を行った (学術論文 [9])．

- $\textsc{Room}(2,1)$ は，クラス $\mathcal{P}$ に属する．
- $\textsc{Room}(2,2)$ は $\mathcal{NP}$ 困難である．よって，$s \geq 2$ かつ $t \geq 2$ について $\textsc{Room}(s,t)$ も $\mathcal{NP}$ 困難となる．
- $s \leq 2$ かつ $t \geq 1$ について $\textsc{Order}(s,t)$ はクラス $\mathcal{P}$ に属する．
- $\textsc{Order}(3,1)$ もクラス $\mathcal{P}$ に属する．

最適化問題の一つにスケジューリング問題がある．スケジューリングとは，複数の仕事があるときに，どういう順番で仕事を進めればもっとも早

く仕事が終了するか，あるいは都合が良いかなど，その最適な仕事を行う順番を求める問題のことを言う．スケジューリング問題は，最適解を見つけるようなプログラムを作っても膨大な計算時間を必要とするため，一般的に難しい問題に分類される．また，スケジューリング問題の中で特に実際の場面で頻繁に見られる問題として，授業の時間割を決める問題や，従業員の勤務表作成問題などの時間割作成問題がある．時間割作成問題は，教育機関や，医療施設，交通機関，スポーツの試合など，応用範囲は幅広い．

教育機関における時間割作成は，学校の授業時間割作成，大学のコースや試験の日程や時間割決定などにおいて用いられている [4]．例えば，De Werra は学校の時間割作成問題の枠組みを与え [1]， Junginger は，ドイツにおける時間割作成問題の研究について報告している [2]．また，Petrovic や Burke らによって，大学の講義や試験の予定表作成問題が調査されている [3]．このような多くの時間割作成問題では，担当教員や教室配置といった，多くのスケジュールや制約条件を考慮しなければならないため，非常に困難な問題であり，大学における時間割作成問題は，40 年以上も研究され続けている重要な研究課題である．

文献

[1] D. deWerra, “ An introduction to timetabling, ” European J. Operational Research, Vol.19, No.2, pp.151.162, 1985.

[2] W. Junginger, “ Timetabling in Germany - a survey, ”Interfaces, Vol.16, pp.66.74, 1986.

[3] S. Petrovic and E. Burke. “ University Timetabling, ”Handbook of Scheduling: Algorithms, Models, and Performance Analysis. Josepy Y-T.Leung(Ed.), Chapman and Hall/CRC, 2004.

[4] A. Schaerf, “ A survey of automated timetabling, ”Artificial Intelligence Review, Vol.13, No.2, pp.87.127, 1999.

研究業績一覧

著書

1. 牛島和夫編著，相利民，朝廣雄一著:
“離散数学 (コンピュータサイエンス教科書シリーズ)”, コロナ社, 2006 年 8 月.
概要: 本書は，計算機科学・情報工学の基礎である離散数学を扱った教科書であり，学部生対象の講義での利用を目的としている．教科書本体は概念・定義や定理そのものの説明に重点をおいており，定理の証明や演習問題の解答は付録 CD- ROM にて提供されている．本書は全 4 章構成であり，各章の内容は次の通りである: 第 1 章：集合論，第 2 章：代数系，第 3 章：数理論理，第 4 章：グラフ理論.

2. Hideo Hirose:
““The Bump Hunting by the Decision Tree with the Genetic Algorithm” in Advances in Computational Algorithms and Data Analysis”, Springer, (2008 予定).
概要: In difficult classification problems of the z-dimensional points into two groups giving 0-1 responses due to the messy data structure, it is more favorable to search for the denser regions for the response 1 points than to find the boundaries to separate the two groups. To such problems often seen in customer databases, we have developed a bump hunting method using probabilistic and statistical methods as shown in the previous study. By specifying a pureness rate in advance, a maximum capture rate will be obtained. In finding the maximum capture rate, we have used the decision tree method combined with the genetic algorithm. Then, a trade-off curve between the pureness rate and the capture rate can be constructed. However, such a trade-off curve could be optimistic if the training data set alone is used. Therefore, we should be careful in assessing the accuracy of the trade-off curve. Using the accuracy evalu-

ation procedures such as the cross validation or the bootstrapped hold-out method combined with the training and test data sets, we have shown that the actually applicable trade-off curve can be obtained. We have also shown that an attainable upper bound trade-off curve can be estimated by using the extreme-value statistics because the genetic algorithm provides many local maxima of the capture rates with different initial values. We have constructed the three kinds of trade-off curves; the first is the curve obtained by using the training data; the second is the return capture rate curve obtained by using the extreme-value statistics; the last is the curve obtained by using the test data. These three are indispensable like the Trinity to comprehend the whole figure of the trade-off curve between the pureness rate and the capture rate. This paper deals with the behavior of the trade-off curve from a statistical viewpoint.

学術論文

1. Yuichi Asahiro, Eiji Miyano, Shinichi Shimoirisa:
"k-Collect Tours for Moving Objects with Release Times and Deadlines", Proceedings of the 9th World Multi-Conference on Systemics, Cybernetics and Informatics, Vol.III, pp.192–197, (2005 年 7 月).
概要: In this paper, we consider a formalization of the following scheduling problem motivated by a variety of practical applications: The input to the problem consists of a robot arm of capacity k (i.e., it can grasp at most k products at a time) and n products, each of which is associated with a release time and a deadline and moves on a predefined path with a fixed constant speed. In each round (tour), the robot arm grasps at most k products by their deadlines and delivers to a specified delivery point. The goal of the problem is to find a collection of tours such that the robot arm of capacity k grasps by their deadlines and delivers as many products as possible. In this paper we show that there is an $O(n\alpha(n)\log n)$ algorithm for the case $k = 1$ if all the products move on given straight lines with faster speeds than the robot arm, where $\alpha(n)$ is the inverse Ackermann function. However, for the similar setting for the paths of the products, we prove that if some of the products' speeds are slower than the robot arm's one, then the problem remains $\mathcal{NP}$-hard.

2. Yuichi Asahiro:
"Simple Greedy Methods for DNA Word Design", Proceedings of the 9th World Multi-Conference on Systemics, Cybernetics and Informatics, Vol.III, pp.186–191, (2005 年 7 月).
概要: For designing a large set of DNA strands, we tested three algorithms based on traditional greedy heuristics. The constraints to be considered for the set are Hamming Distance (HD) constraint, Reverse Complement (RC) constraint, and GC Content (GC) constraint. The variety of the combinations of these three constraints is the most fundamental target of designing DNA strands. By the algorithms, we found larger sets than previously known ones for some settings of the length n and Hamming Distance d, e.g., $n = 14$ and $d = 3, 10$ with HD, GC, and RC constraints, and $n = 12$ and $d = 5, 6, 7, 8$ with HD and RC constraints. In addition to that, the theoretical complexity of the problem is shown, which does not seem to be explained in the previous studies.

3. Hideo Hirose:
"Optimal boundary finding method for the bumpy regions", Proceedings of the International Federation of Operational Research Societies Triennial 2005 Confer-

ence, FD-19 (Data Mining and Database Modeling)-3, p.l39, (2005 年 7 月).
概要: Assuming that the points having response 0 and the points response 1 are very similarly distributed except some regions, we want to find the boundaries for the bumpy regions in which the points with response 1 are much more frequently observed than in the whole support region.

4. Yuichi Asahiro, Eiji Miyano, Shinichi Shimoirisa:
"Pickup and Delivery for Moving Objects on Broken Lines", Proceedings of the 9th Italian Conference on Theoretical Computer Science, Lecture Notes in Computer Science, Vol.3701, pp.36–50, 2005 年 10 月.
概要: We consider the following variant of the Vehicle Routing Problem that we call the *Pickup and Delivery for Moving Objects (PDMO)* problem, motivated by robot navigation: The input to the problem consists of n products, each of which moves on a predefined path with a fixed constant speed, and a robot arm of capacity one. In each round, the robot arm grasps and delivers one product to its original position. The goal of the problem is to find a collection of tours such that the robot arm grasps and delivers as many products as possible. In this paper we prove the following results: (i) If the products move on broken lines with at least one bend, then the PDMO is MAXSNP-hard, and (ii) it can be approximated with ratio two. However, (iii) if we impose the "straight line without bend" restriction on the motion of every product, then the PDMO becomes tractable.

5. Yuichi Asahiro, Eiji Miyano, Hirotaka Ono, Kouhei Zenmyo:
"Graph Orientation Algorithms to Minimize the Maximum Outdegree", Proceedings of the Computing: The Australasian Theory Symposium(CATS), Conferences in Research and Practice in Information Technology, Vol.51, pp.11–20, 2006 年 1 月.
概要: We study the problem of orienting the edges of a weighted graph such that the maximum weighted outdegree of vertices is minimized. This problem, which has applications in the guard arrangement for example, can be shown to be $\mathcal{NP}$-hard generally. In this paper we first give optimal orientation algorithms which run in polynomial time for the following special cases: (i) the input is an unweighted graph, or more generally, a graph with identically weighted edges, and (ii) the input graph is a tree. Then, by using those algorithms as sub-procedures, we provide a simple, combinatorial, $\min\{\frac{w_{max}}{w_{min}}, (2-\varepsilon)\}$-approximation algorithm for the general case, where w_{max} and w_{min} are the maximum and the minimum weights of edges, respectively, and ε is some small positive real number that depends on the input.

6. Yuichi Asahiro, Tetsuya Furukawa, Keiichi Ikegami, Eiji Miyano:
"How to Pack Directed Acyclic Graphs into Small Blocks", Proceedings of the 6th International Conference on Algorithms and Complexity, Lecture Notes in Computer Science, Vol.3998, pp.272–283, 2006 年 5 月.
概要: The paper studies the following variant of clustering or laying out problems of graphs: Given a directed acyclic graph (DAG for short), the objective is to find a mapping of its nodes into blocks of size at most B that minimizes the maximum number of external arcs during traversals of the acyclic structure by following paths from the roots to the leaves. An external arc is defined as an arc connecting two distinct blocks. The problem can be shown to be NP-hard generally, and to remain intractable even if $B = 2$ and the height of DAGs is three. In this paper we provide a $\frac{3}{2}$ factor linear time approximation algo-

rithm for $B = 2$, and prove that the $\frac{3}{2}$ ratio is optimal in terms of approximation guarantee. In the case of $B \geq 3$, we also show that there is no $\frac{3}{2} - \varepsilon$ factor approximation algorithm assuming P $\neq$ NP, where ε is arbitrarily small positive. Furthermore, we give a 2 factor approximation algorithm for $B = 3$ if the input is restricted to a set of layered graphs.

7. Hideo Hirose, Takahiro Yukizane, Eiji Miyano:
"Boundary Detection for Bumps Using the Gini's Index in Messy Classificatin Problems", Proceedings of the 3rd International Conference on Cybernetics and Information Technologies, Systems and Applications, pp.293–298, 2006 年 7 月.
概要: Suppose that we are interested in classifying the points into two groups giving 0-1 responses in z-dimensional space which corresponds to z explanation variables. So far a lot of classification tools which adopt the misclassification rate to evaluate how well the discrimination be achieved have been proposed and played a key role in the field of statistical learning or data mining. In the real world, however, we often encounter the cases that the misclassification rate remains considerably high no matter how carefully we search for the discriminating boundaries if we want to capture the points to some extent with rather simpler rules. In such a messy case, the optimization criterion in classification goes to obtaining the larger number of points of response 1, we are interested in, with a specified pureness rate. Such a region is called to be a bump or a hotspot. Our primary objective is to find the bumps. To use the information of the boundaries of the bumps with ease, it would be beneficial that the bumps have much simpler shapes of their boundaries such as the z-dimensional box located parallel to some explanation variable axes. The problem raised is, however, how to find such a boundary for some bump. The decision tree algorithms find the splitting point for an explanation variable to construct much purer sub-regions using the entropy or the Gini's index criterion, from the top node to the bottom leafs step by step. Regarding that to produce the purer regions is just the same as to exclude those regions, this splitting criterion may also be applicable to produce the denser region for response 1 including response 0 points to some extent. This paper deals with this point. Mimicking the real data case in a customer database of a correspondence course in Japan, we investigate the fundamental data cases such that the marginal distributions for response 1 are rather simple shaped, such as monotonic or unimodal, and those for response 0 is almost uniformly distributed. The unimodal and monotonic cases are studied by using the normal distributions located at the center of and on the corner of the uniform distributions, respectively. Under such restricted conditions, we have found that the Gini's index can detect the boundaries for the bumps as if eyes of human beings do so.

8. Takahiro Yukizane, Shin-ya Ohi, Eiji Miyano, Hideo Hirose:
"The Bump Hunting Method Using the Genetic Algorithm with the Extreme-Value Statistics", IEICE Transactions on Information and Systems, Invited Papers from New Horizons in Computing, Vol.E89-D, No.8, pp.2332–2339, 2006 年 8 月.
概要: In difficult classification problems of the z-dimensional points into two groups giving 0-1 responses due to the messy data structure, we try to find the denser regions for the favorable customers of response 1, instead of finding the boundaries to separate the two groups. Such regions are called the bumps, and finding the boundaries of

the bumps is called the bump hunting. The main objective of this paper is to find the largest region of the bumps under a specified ratio of the number of the points of response 1 to the total. Then, we may obtain a trade-off curve between the number of points of response 1 and the specified ratio. The decision tree method with the Gini's index will provide the simple-shaped boundaries for the bumps if the marginal density for response 1 shows a rather simple or monotonic shape. Since the computing time searching for the optimal trees will cost much because of the NP-hardness of the problem, some random search methods, e.g., the genetic algorithm adapted to the tree, are useful. Due to the existence of many local maxima unlike the ordinary genetic algorithm search results, the extreme-value statistics will be useful to estimate the global optimum number of captured points; this also guarantees the accuracy of the semi-optimal solution with the simple descriptive rules. This combined method of genetic algorithm search and extreme-value statistics use is new. We apply this method to some artificial messy data case which mimics the real customer database, showing a successful result. The reliability of the solution is discussed.

9. Yuuki Kiyonari, Eiji Miyano, Shuichi Miyazaki:
"Computational Complexity Issues in University Interview Timetabling", Proceedings of the 6th International Conference on the Practice and Theory of Automated Timetabling, pp.448–453, 2006 年 8 月.
概要: This paper introduces a new timetabling problem on universities, called interview timetabling. In this problem, some constant number, say, three, of referees are assigned to each of $2n$ graduate students. It requires to assign these students to an $n \times 2$ array, consisting of n timeslots and two rooms, so that two students sharing the same referee must be assigned to different timeslots. Furthermore, we want to minimize the total number of movements of all referees between two rooms. This is a problem actually having arisen in the interview timetabling in Kyoto University. We propose two restricted models of this problem, and investigate their time complexities.

10. Yuichi Asahiro, Takashi Horiyama, Kazuhisa Makino, Hirotaka Ono, Toshinori Sakuma, Masafumi Yamashita:
"How to Collect Balls Moving in the Euclidean Plane", Discrete Applied Mathematics, Vol.154, No.16, pp.2247–2262, 2006 年 11 月.
概要: In this paper, we study how to collect n balls moving with a fixed constant velocity in the Euclidean plane by k robots moving on straight track-lines through the origin. Since all the balls might not be caught by robots, differently from Moving-target TSP, we consider the following 3 problems in various situations: (i) deciding if k robots can collect all n balls; (ii) maximizing the number of the balls collected by k robots; (iii) minimizing the number of the robots to collect all n balls. The situations considered in this paper contain the cases in which track-lines are given (or not), and track-lines are identical (or not). For all problems and situations, we provide polynomial time algorithms or proofs of intractability, which clarify the tractability-intractability frontier in the ball collecting problems in the Euclidean plane.

11. Yuichi Asahiro, Eiji Miyano, Shuichi Miyazaki, Takuro Yoshimuta:
"Weighted Nearest Neighbor Algorithms for the Graph Exploration Problem on Cycles", Proceedings of the 33rd Conference on Current Trends in Theory and Practice of Infor-

matics, Lecture Notes in Computer Science, Vol.4362, pp.164–175, 2007年1月.

概要: In the graph exploration problem, a searcher explores the whole set of nodes of an unknown graph. The searcher is not aware of the existence of an edge until he/she visits one of its endpoints. The searcher's task is to visit all the nodes and go back to the starting node by traveling as a short tour as possible. One of the simplest strategies is the nearest neighbor algorithm (NN), which always chooses the unvisited node nearest to the searcher's current position. The weighted NN (WNN) is an extension of NN, which chooses the next node to visit by using the weighted distance. It is known that WNN with weight 3 is 16-competitive for planar graphs. In this paper we prove that NN achieves the competitive ratio of 1.5 for cycles. In addition, we show that the analysis for the competitive ratio of NN is tight by providing an instance for which the bound of 1.5 is attained, and NN is the best for cycles among WNN with all possible weights. Furthermore, we prove that no online algorithm is better than 1.25-competitive.

12. Hideo Hirose, Shin-ya Ohi, Takahiro Yukizane:
"Assessment of the Prediction Accuracy in the Bump Hunting Procedure", Proceedings of the 6th Hawaii International Conference on Statistics, Mathematics and Related Fields, pp.1–23, 2007年1月.

概要: In a series of the previous study (Hirose 2005a, Hirose 2005b, Hirose et. al. 2006a, Yikizane et. al. 2006), we have developed a bump hunting method using probabilistic and statistical methods, where we are interested in finding denser regions for response 1 among regions for responses 0 and 1 mixed under a condition that the pureness rate is pre-specified; the pureness rate is the ratio of the number of response 1 to the number of responses 0 and 1 in the target region. A global maximum capture rate is estimated by using locally obtained maximum captured response 1 points, where the capture rate is the ratio of the number of response 1 to the number of responses 0 and 1 in the region defined. The accuracy of the estimated maximum capture rate was assessed by using the simple bootstrap method without correction formula. We have not thought seriously of the bias and the variance to the predicted estimate. However, we are now aware of that we should treat the value of the predicted estimate very carefully. In the previous study, we have dealt with all the data as the training data or the learning data. As is well-known in machine learning world, the best fitted model using the training data would not necessarily be accurate when the test data are used. A typical research in assessing the misclassification error in classification problems is reported by Kohavi (1995), and a more general explanation is found in the literature (Hastie, et. al. 2001), where the test sample method, the cross-validation, and the bootstrap method are often used to assess the accuracy of the predicted estimate and the model. It is recommended to take into account this kind of assessment to our problem too. However, difficulties lie on the bump hunting method that we have proposed because we are interested in not only estimating the maximum capture rate but also in finding the corresponding (binary decision) rule to capture the points of response 1. Since we consider that the rule should also capture a large amount of response 1 points to the future data, a combination of the maximum capture rate and the corresponding rule becomes important. In the example in the previous study, however, we have found that the optimized rule to capture the maximum

points for response 1 to the training data did not necessarily capture a large amount of response 1 points to the test data. In this paper, we show first how badly the optimality collapses to the test data. We then discuss how to circumvent this kind of difficulty.

13. 朝廣雄一，小野廣隆，宮野英次，山下雅史: “移動物体回収問題”, 電子情報通信学会誌, Vol.90, No.3, pp.245–247, 2007 年 3 月.
概要: 巡回行商人問題 (TSP) は，計算機科学の分野で古くから知られる問題の一つである．行商人は入力として与えられる都市をすべて巡回して出発位置に戻ってこなければならないが，その時通った道のりの総和をできるだけ小さくしたい．TSP については多くの研究成果が知られておりいまだに活発な研究対象である．本記事では，TSP を動的な環境へ拡張した，すなわち巡回対象が都市のように静止しておらず移動している場合を扱う，移動物体回収問題についての理論的な研究成果を紹介する．

14. Yuichi Asahiro, Eiji Miyano, Hirotaka Ono, Kouhei Zenmyo:
“Graph Orientation Algorithms to Minimize the Maximum Outdegree”, International Journal of Foundation of Computer Science, Vol.18, No.2, pp.197–215, 2007 年 4 月.
概要: This paper studies the problem of orienting all edges of a weighted graph such that the maximum weighted outdegree of vertices is minimized. This problem, which has applications in the guard arrangement for example, can be shown to be $\mathcal{NP}$-hard generally. In this paper we first give optimal orientation algorithms which run in polynomial time for the following special cases: (i) the input is an unweighted graph, and (ii) the input graph is a tree. Then, by using those algorithms as subprocedures, we provide a simple, combinatorial, $\min\{\frac{w_{max}}{w_{min}}, (2-\varepsilon)\}$-approximation algorithm for the general case, where w_{max} and w_{min} are the maximum and the minimum weights of edges, respectively, and ε is some small positive real number that depends on the input.

15. Kazuo Iwama, Eiji Miyano, Hirotaka Ono: “Drawing Borders Efficiently”, Proceedings of the 4th International Conference on Fun with Algorithms, Lecture Notes in Computer Science, Vol.4475, pp.213–226, 2007 年 6 月.
概要: A spreadsheet, especially MS Excel, is probably one of the most popular software applications for personal-computer users and gives us convenient and user-friendly tools for drawing tables. In practical situations, we often wish to draw several vertical and horizontal black lines on selective gridlines to enhance the readability of our spreadsheet. Such situations we frequently encounter are formulated as the Border Drawing Problem (BDP). Given a layout of black line segments, we study how to draw it efficiently from an algorithmic view point, by using a set of border styles and investigate its complexity. (i) We first give a formal model based on MS Excel, and then (ii) show that the problem is $\mathcal{NP}$-hard for the set of the Excel border styles and for any reasonable subset of the styles. Moreover, in order to provide potentially more efficient drawing, (iii) we propose a new compact set of border styles and show a necessary and sufficient condition of its drawability.

16. Yuichi Asahiro, Jesper Jansson, Eiji Miyano, Hirotaka Ono, Kouhei Zenmyo:
“Approximation Algorithms for the Graph Orientation Minimizing the Maximum Weighted Outdegree”, Proceedings of the 3rd International Conference on Algorithmic Aspects in Information and Management, Lecture Notes in Computer Science,

Vol.4508, pp.167–177, 2007 年 6 月.

概要: Given an undirected graph $G = (V, E)$ and a weight function $w : E \to \mathbb{Z}^+$, we consider the problem of orienting all edges in E so that the maximum weighted outdegree among all vertices is minimized. In this paper (1) we prove that the problem is strongly NP-hard if all edge weights belong to the set $\{1, k\}$, where k is any integer greater than or equal to 2, and that there exists no pseudo-polynomial time approximation algorithm for this problem whose approximation ratio is smaller than $(1 + 1/k)$ unless P=NP; (2) we present a polynomial time algorithm that approximates the general version of the problem within a factor of $(2 - 1/k)$, where k is the maximum weight of an edge in G; (3) we show how to approximate the special case in which all edge weights belong to $\{1, k\}$ within a factor of 3/2 for $k = 2$ (note that this matches the inapproximability bound above), and $(2 - 2/(k + 1))$ for any $k \geq 3$, respectively, in polynomial time.

17. Yuichi Asahiro, Eiji Miyano, Toshihide Murata, Hirotaka Ono:
"On Approximation of Bookmark Assingments", Proceedings of the 32nd International Symposium on Mathematics Foundations of Computer Science, Lecture Notes in Computer Science, Vol.4708, pp.115–124, 2007 年 8 月.

概要: Consider a rooted directed acyclic graph $G = (V, E)$ with root r, representing a collection V of web pages connected via a set E of hyperlinks. Each node v is associated with the probability that a user wants to access the node v. The access cost is defined as the expected number of steps required to reach a node from the root r. A bookmark is an additional shortcut from r to a node of G, which may reduce the access cost. The bookmark assignment problem is to find a set of bookmarks that achieves the greatest improvement in the access cost. For the problem, the paper presents a polynomial time approximation algorithm with factor $(1-1/e)$, and shows that there exists no polynomial time approximation algorithm with a better constant factor than $(1 - 1/e)$ unless $\mathcal{NP} \subseteq \mathcal{DTIME}(N^{O(\log \log N)})$, where N is the size of the inputs.

18. Hideo Hirose, Takahiro Yukizane, Takuya Deguchi:
"The Bump Hunting Method and Its Accuracy Using the Genetic Algorithm with Application to Real Customer Data", Proceedings of the IEEE 7th International Conference on Computer and Information Technology, pp.128–132, 2007 年 10 月.

概要: Suppose that we are interested in searching for denser regions showing response 1 with many feature variables (explanation variables) in a z-dimensional space, where each point is assigned response 1 or response 0 as its target value; such a region is called the bump. In a series of previous studies, we have shown that the bump hunting method using the decision tree combined with the genetic algorithm is useful for certain smaller simulated data case mimicked to a real customer case. We have developed a trade-off curve with its accuracy evaluation between the pureness rate and the capture rate to the simulated data. This paper deals with a real customer data case, and we have found that it is crucial to know the relation between the number of feature variables and the number of samples.

19. Hideo Hirose, Takahiro Yukizane:
"The Accuracy of the Trade-Off Curve in the Bump Hunting", Proceedings of the 7th Hawaii International Conference on Statistics, Mathematics and Related Fields, pp.17–19, 2008 年 1 月.

概要: In difficult classification problems of the z-dimensional points into two groups having 0-1 responses due to the messy data structure, it is more favorable to search for the denser regions for the response 1 assigned points than to find the boundaries to separate the two groups. To such problems often seen in customer databases, we have developed a bump hunting method using probabilistic and statistical methods as shown in the previous study. By specifying a pureness rate in advance, a maximum capture rate will be obtained. In finding the maximum capture rate, we have used the decision tree method combined with the genetic algorithm. Then, a trade-off curve between the pureness rate and the capture rate can be constructed. However, such a trade-off curve could be optimistic if the training data set alone is used; we might obtain an illusional curve when the number of samples is small and the number of explanation variables is large to some extent. Therefore, we should be careful in assessing the accuracy of the trade-off curve. Using the accuracy evaluation procedures such as the cross validation or the bootstrapped hold-out method combined with the training and test data sets, we have shown that the actually applicable trade-off curve can be obtained. We have also shown that an attainable upper bound trade-off curve can be estimated by using the extreme-value statistics because the genetic algorithm provides many local maxima of the capture rates with different initial values. However, this curve may also be optimistic because we use the training data set alone. In this paper, we propose a new genetic algorithm to tackle this problem. In the procedure of the new genetic algorithm, the evaluation to each evolution procedure is applied to the test data results rather than to the training data results. Due to this expansion for the genetic algorithm, we can expect the accuracy for the actually applicable upper bound for the trade-off curve by applying the extreme-value statistics by the new accuracy evaluation procedures such as the cross validation or the bootstrapped hold-out.

20. Yuichi Asahiro, Eiji Miyano, Hirotaka Ono: "Graph Classes and the Complexity of the Graph Orientation Minimizing the Maximum Weighted Outdegree", Proceedings of the Computing: The Australasian Theory Symposium(CATS), Conferences in Research and Practice in Information Technology, Vol.77, pp.97–106, 2008 年 1 月.
概要: Given an undirected graph with edge weights, we are asked to find an orientation, i.e., an assignment of a direction to each edge, so as to minimize the weighted maximum outdegree in the resulted directed graph. The problem is called MMO, and is a restricted variant of the well-known minimum makespan problem. As previous studies, it is shown that MMO is in $\mathcal{P}$ for trees, weak $\mathcal{NP}$-hard for planar bipartite graphs, and strong $\mathcal{NP}$-hard for general graphs. There are still gaps between those graph classes. The objective of this paper is to show tight thresholds of complexity: We show that MMO is (i) in $\mathcal{P}$ for cactuses, (ii) weakly $\mathcal{NP}$-hard for outerplanar graphs, and also (iii) strongly $\mathcal{NP}$-hard for P_4-bipartite graphs. The latter two are minimal superclasses of the former. Also, we show the $\mathcal{NP}$-hardness for the other related graph classes, diamond-free, house-free, series-parallel, bipartite and planar.

21. Yuichi Asahiro, Eiji Miyano, Shinichi Shimoirisa:
"Grasp and Delivery for Moving Objects on Broken Lines", Theory of Computing Systems, (掲載予定).
概要: This paper studies the following vari-

ant of the Vehicle Routing Problem that we call the *Grasp and Delivery for Moving Objects (GDMO)* problem, motivated by robot navigation: The input to the problem consists of n products, each of which moves on a predefined path with a fixed constant speed, and a robot arm of capacity one. In each round, the robot arm grasps one product and then delivers it to the depot. The goal of the problem is to find a collection of tours such that the robot arm grasps and delivers as many products as possible. In this paper we prove the following results: (i) If the products move on broken lines with at least one bend, then the GDMO is MAXSNP-hard, and (ii) it can be approximated with ratio 2. However, (iii) if we impose the "straight line without bend" restriction on the motion of every product, then the GDMO becomes tractable.

研究会等

1. 下入佐真一，朝廣雄一，宮野英次:
"移動系における最大個数巡回アルゴリズム", 数理解析研究所講究録 (計算機科学基礎理論の新展開), 1375, pp.260–266, 2004 年 5 月.
概要: 平面上を移動する n 個の物体を巡回する経路選択問題に対して，次のような条件を付けた問題を考える．(i) n 個の物体の平面上での初期座標はあらかじめ与えられいる．(ii) 移動物体は一定の方向と一定の速さで等速直線移動しており，これらについてもあらかじめ与えられている．(iii) 巡回者の数は 1 とする．(iv) 巡回者は一定の速さで直線的に移動可能で，時間 0 で移動方向を換えることができる．例えば，一定の速度で生産ライン上を流れて来る部品を，ロボットアームにより加工する際の最適スケジュールなど，制御ロボットへの応用が考えられる．本稿では，巡回者の可動方向を 3 方向 (上下左) および 4 方向 (上下左右) とした場合に対する移動物体巡回問題を考える．また，巡回者と n 個の移動物体の速さはすべて等しく，移動物体の可動範囲は平行な 2 直線上に限られているとする．本稿では，巡回者が 4 方向に移動可能であるときに，巡回個数を最大とするような経路を求める $O(n + T_s(n))$ 時間アルゴリズムを示す．ここで $T_s(n)$ は n 要素のソーティングに必要な時間とする．また，巡回者の可動方向が 3 に制限した場合に n 個の移動物体を全て巡回可能か否かを判定する $O(n^3)$ 時間アルゴリズムが存在することを示す.

2. Kazuo Iwama, Kousuke Izumi, Eiji Miyano, Hirotaka Ono:
"Drawing Borders Efficiently", 電子情報通信学会技術研究報告, 104 巻・147 号, pp.47–54, 2004 年 6 月.
概要: A spreadsheet, especially MS Excel, is probably one of the most popular software for personal-computer users and gives us convenient and user-friendly tools for drawing tables. In practical situations, we often wish to draw several vertical and horizontal black lines on selective gridlines to enhance the readability of our spreadsheet. Such situations we frequently encounter are formulated as the Border Drawing Problem (BDP). Given a layout of black line segments, from an algorithmic view point we study how to draw it efficiently by using a set of border styles and investigate its complexity. (i) We first give a formal model based on MS Excel, and then (ii) show that the problem is $\mathcal{NP}$-hard for the set of the Excel border styles and for any reasonable subset of the styles. Moreover, in order to provide potentially more efficient drawing, (iii) we propose a new compact set of border styles and show a necessary and sufficient condition of its drawability.

3. Kazuo Iwama, Kousuke Izumi, Eiji Miyano, Hirotaka Ono:
"Tiling Problems with the Edge-Overwriting Rule", 情報処理学会研究報告, 2004 巻・101 号, pp.61–66, 2004 年 10 月.

概要: The tiling problems have received considerable attention for the last decades, and a lot of results are reported. One of the most popular variants of the problems could be tilings of regions with Wang tiles, which are 1×1 squares with colored edges, by using the following rule: If two tiles touch, then the color of their shared edges must be the same. In this paper, we consider a new tiling problem under the edge-overwriting rule in which any pair of tiles can touch but the color of the shared edges after tiling is determined by the edge color of the tile placed later. The problem input is a pair of a target figure and a tile set to be used, and the output is a sequence of tiles which realizes the target figure under the edge-overwriting rule. We show the sufficient and necessary condition of a tile set to realize any target figure, and then show $\mathcal{NP}$-hardness proof and an approximation algorithm for the problem of minimizing the number of tiles realizing target figures.

4. Yuichi Asahiro, Eiji Miyano, Shinichi Shimoirisa:
"Collect Tours for Moving Objects with Release Times and Deadlines", 電子情報通信学会技術研究報告, 104 巻・743 号, pp.79–85, 2005 年 3 月.
概要: In this paper, we consider a formalization of the following scheduling problem motivated by a variety of practical applications: The input to the problem consists of a robot arm of capacity k (i.e., it can grasp at most k products at a time) and n products, each of which is associated with a release time and a deadline and moves on a predefined path with a fixed constant speed. In each round (tour), the robot arm grasps at most k products by their deadlines and delivers to a specified delivery point. The goal of the problem is to find a collection of tours such that the robot arm of capacity k grasps by their deadlines and delivers as many products as possible. In this paper we show that there is an $O(n\alpha(n)\log n)$ algorithm for the case $k = 1$ if all the products move on given straight lines with faster speeds than the robot arm, where $\alpha(n)$ is the inverse Ackermann function. However, for the similar setting for the paths of the products, we prove that if some of the products' speeds are slower than the robot's one, then the problem is $\mathcal{NP}$-hard.

5. Yuichi Asahiro, Eiji Miyano, Shinichi Shimoirisa:
"Hardness of Pickup and Delivery for Moving Objects on Broken Lines", 電子情報通信学会技術研究報告, 105 巻・72 号, pp.9–16, 2005 年 5 月.
概要: We consider the following variant of the Vehicle Routing Problem that we call the *Pickup and Delivery for Moving Objects (PDMO)* problem, motivated by robot navigation: The input to the problem consists of n products, each of which moves on a predefined path with a fixed constant speed, and a robot arm of capacity one. In each round, the robot arm grasps and delivers one product to its original position. The goal of the problem is to find a collection of tours such that the robot arm grasps and delivers as many products as possible. In this paper we prove the following results: (i) If the products move on broken lines with at least one bend, then the PDMO is MAXSNP-hard, and (ii) it can be approximated with ratio two. However, (iii) if we impose the "straight line without bend" restriction on every motion of the product, then the PDMO becomes tractable.

6. 廣瀬英雄:
"分類問題への極値分布の応用", 統計数理研究所研究会，極値理論の工学への応用 (3), p.54 – 65, 2005 年 9 月.
概要: In difficult classification problems

of the z-dimensional points into two groups giving 0-1 responses due to the messy data structure, we try to find the denser regions for the favorable customers of response 1, instead of finding the boundaries to separate the two groups. Such regions are called the bumps, and finding the boundaries of the bumps is called the bump hunting. Specifying the ratio of the number of points of response 1 to the total, we try to find the largest regions of the bumps. Then, we may obtain a trade-off curve between the number of points of response 1 and the specified ratio. The decision tree method with the Gini's index will provide the simple-shaped boundaries for the bumps if the marginal density for response 1 shows rather simple or monotonic shape. Since the computing time searching for the optimal trees will cost much because of the NP-hard property of the problem, some random search methods, e.g. the genetic algorithm adapted to the tree, are useful. Due to the existence of many local maxima unlike the ordinary genetic algorithm search results, the extreme-value statistics will be useful to estimate the global optimum number of captured points; this guarantees the accuracy of the semi-optimal solution with the simple descriptive rules. We show a successful result in applying this method to some artificial messy data case which mimics the real customer database.

7. 廣瀬英雄，行實隆広，大井伸哉，宮野英次: "Bump hunting 問題における極値統計の応用", 日本計算機統計学会 19 回シンポジウム論文集, pp.55–58, 2005 年 10 月.
概要: 分類困難なデータベースからの分類問題を扱う際に，分類ルールの明示化，分類の最適化などの要求に応えるのに確率的な決定木を作成する必要が起こった．確率的な木の作成に遺伝的アルゴリズムを使用し，更に極値統計を援用することで，所望する分類アルゴリズムへの見通しがついた．

8. Yuichi Asahiro, Eiji Miyano, Hirotaka Ono, Kouhei Zenmyo:
"On the Graph Orientation of Minimizing the Maximum Outdegree", 電子情報通信学会技術研究報告，105 巻・34 号，pp.53–60, 2005 年 10 月.
概要: We study the problem of orienting the edges of a weighted graph such that the maximum weighted outdegree of vertices is minimized. This problem, which has applications in the guard arrangement for example, can be shown to be $\mathcal{NP}$-hard generally. In this paper we first give optimal orientation algorithms which run in polynomial time for the following special cases: (i) the input is an unweighted graph, or more generally, a graph with identically weighted edges, and (ii) the input graph is a tree. Then, by using those algorithms as sub-procedures, we provide a simple, combinatorial, $\min\{\frac{w_{max}}{w_{min}}, (2-\varepsilon)\}$-approximation algorithm for the general case, where w_{max} and w_{min} are the maximum and the minimum weights of edges, respectively, and ε is some small positive real number that depends on the input.

9. Yuichi Asahiro, Tetsuya Furukawa, Keiichi Ikegami, Eiji Miyano:
"Approximability and Non-approximability of the Minimum Block Transfer Problem", 電子情報通信学会技術研究報告, 105 巻・679 号, pp.17–24, 2006 年 3 月.
概要: The paper deals with the following problem: Given a directed acyclic graph, the objective is to find a mapping of its nodes into blocks of size at most B that minimizes the maximum number of external arcs during traversals of the acyclic structure by following paths from the roots to the leaves. An external arc is defined as an arc connecting two distinct blocks. The problem can be shown to be NP-hard. In this paper we provide a $\frac{3}{2}$ factor approximation algorithm for $B = 2$, and prove

that the $\frac{3}{2}$ ratio is optimal in terms of approximation guarantee, while the previous approximation ratio was 2. In the case of $B \geq 3$, we also show that there is no $\frac{3}{2} - \varepsilon$ factor approximation algorithm assuming P $\neq$ NP, where ε is arbitrarily small positive. Furthermore, we give a 2 factor approximation algorithm for $B = 3$ if the input is a layered graph, which improves the known ratio of 3.

10. 清成悠貴，宮野英次，宮崎修一:
"試問予定表作成問題の計算複雑さ"，電子情報通信学会技術研究報告, 106 巻・128 号, pp.7–14, 2006 年 6 月.
概要: 本稿では，試問予定表作成問題と呼ばれる大学における時間割表作成問題を考える．この問題では，m 人の審査員の集合，$2n$ 人の学生の集合，各学生について審査する複数の審査員の担当割当 (例えば各学生を 3 人ずつの審査員が担当) が与えられる．審査は n 人ずつ 2 つの部屋で同時に行われるため，学生と担当審査員の組は，n 個の時刻と 2 つの部屋から成る $n \times 2$ 行列に割り当てられる必要がある．ここで，同じ審査員が担当する異なる 2 人の学生は異なる時刻に割り当てられなければならない．さらに，すべての審査員が 2 つの部屋を移動する回数の合計をできるだけ少なくしたい．本問題は，実際に京都大学の試問予定表作成時に起きた問題を形式化したものである．この問題について 2 つの制約付きモデルを提案し，それらの計算複雑さについて調査する．

11. Yuichi Asahiro, Jesper Jansson, Eiji Miyano, Hirotaka Ono, Kouhei Zenmyo:
"(In)approximability of Graph Orientation to Minimize the Maximum Weighted Outdegree", 情報処理学会アルゴリズム研究会情処研報, 2006 巻・71 号, pp.21–28, 2006 年 7 月.
概要: Given an undirected graph $G = (V, E)$ and a weight function $w : E \to \mathbb{Z}^+$, we consider the problem of orienting all edges in E so that the maximum weighted outdegree among all vertices is minimized. It has previously been shown that the unweighted version of the problem is polynomially solvable while the weighted version is (weakly) NP-hard. In this paper, we strengthen these results as follows: (1) We prove that the weighted version is strongly NP-hard even if all edge weights belong to the set $\{1, k\}$, where k is any integer greater than or equal to 2, and that there exists no pseudo-polynomial time approximation algorithm for this problem whose approximation ratio is smaller than $(1 + 1/k)$ unless P=NP; (2) we present a new polynomial-time algorithm that approximates the general version of the problem within a factor of $(2 - 1/k)$, where k is the maximum weight of an edge in G; (3) we show how to approximate the special case in which all edge weights belong to $\{1, k\}$ within a factor of 3/2 for $k = 2$ (note that this matches the inapproximability bound above), and $(2 - 2/(k + 1))$ for any $k \geq 3$, respectively, in polynomial time.

12. 朝廣 雄一，宮野 英次，宮崎 修一，吉牟田 拓朗:
"サイクルグラフ上での地図作成問題に対する重み付き最近傍アルゴリズム", 電子情報通信学会技術研究報告, 106 巻・405 号, pp.15–22, 2006 年 12 月.
概要: グラフ上での地図作成問題とは，探索者が未知のグラフの全ての頂点を訪問することによりグラフ構造を調査する問題である．探索者は辺の存在とその長さをその端点を訪れるまで判らないとする．探索者は，できるだけ短い経路を通ることにより全ての頂点と辺を調査して，出発点まで戻って来なければならない．本問題に対する最も単純な方法の一つは，最近傍アルゴリズム (NN) であり，まだ訪れていない頂点の中で探索者の現在の場所から最も近い場所に移動する戦略である．重み付き最近傍アルゴリズム (WNN) は，NN の拡張であり，ある重み付きの距離により次の移動場所を決める．平面グラフにおいては，

重み 3 である WNN が 16 競合であることが知られている．本稿ではサイクルグラフについては，NN の競合比が 1.5 となること，その解析が厳密であることを示す．また，サイクルグラフに対しては WNN の中で NN が最適であることを示す．さらに，本問題に対しては，1.25 競合よりも良いアルゴリズムが存在しないことを示す．

13. 行實隆広，廣瀬英雄，大井伸哉，宮野英次: “バンプ探索における解の精度”, 情報処理学会数理モデル化と問題解決研究会, 2006 巻・135 号, pp.13–16, 2006 年 12 月.
概要: 多次元空間内に特徴量を持ち 2 値 (0/1) 反応をとる N 個の中から反応 1 を示す点が他に比べて密な領域 (バンプ，ホットスポット) を探索する問題を考える．これまに，探索結果を予測に使いやすくするためには決定木を用いたバンプ探索法が有効であり，また最適値を求めるためには確率的探索法 (GA) に加え，極致統計法を用いる方法を新しく提案した．しかし，得られた結果がどのような精度を持っているかはまだ分かっていなかった．ここでは，学習データとテストデータを使うことで，この問題がいかに深刻であるかを指摘し，次に最適に探索された結果の精度について述べる．テストサンプル法とブートストラップのメリットを併せ持つ方法も提案する．

14. 朝廣雄一，宮野英次，小野廣隆，村田俊英: “ブックマーク問題の近似について”, 電子情報通信学会技術研究報告, 107 巻・73 号, pp.1–16, 2007 年 5 月.
概要: ハイパーリンクの集合 E で連結されたウェブページの集合 V を表した根付き有向非巡回グラフ $G = (V, E)$ を考える．それぞれのノード v は，ユーザがノード v を閲覧するアクセス頻度を持っている．アクセスコストは，根ノード r からあるノード v に到達するためのステップ数の期待値で定義される．ブックマークは，グラフ G の根ノード r に付加するショートカットリンクであり，アクセスコストを小さくする効果がある．ブックマーク問題は，アクセスコストを最も改善するようなブックマーク集合を求める問題である．本稿では，ブックマーク問題に対する $(1-1/e)$ 近似アルゴリズムを示す．また，$\mathcal{NP} \subseteq \mathcal{DTIME}(N^{O(\log\log N)})$ を満たさない場合には，$(1-1/e)$ よりも良い近似比を持つ多項式時間のアルゴリズムは存在しないことを示す．ここで，N は入力サイズを表している．

15. 出口拓也，行實隆広，宮野英次，廣瀬英雄: “顧客データベースにおける bump hunting とその精度”, 日本計算機統計学会 21 回シンポジウム論文集, pp.115–118, 2007 年 5 月.
概要: 分類の困難なデータベースからの分類問題を扱う際に，分類ルールの明示化と分類の最適化を同時に満たす方法として，確率的な決定木の作成に遺伝的アルゴリズムを使った bump hunting 法の提案を行い，シミュレーションデータで確認を行ってきた．ここでは，この方法を実データに適用し，その有効性を検証する．また，推定値の精度についても述べる．

16. 廣瀬英雄: “極値分布のバンプ探索への適用”, 統計数理研究所研究会 極値理論の工学への応用, pp.1–11, 2007 年 12 月.
概要: In difficult classification problems of the z-dimensional points into two groups having 0-1 responses due to the messy data structure, it is more favorable to search for the denser regions for the response 1 assigned points than to find the boundaries to separate the two groups. To such problems often seen in customer databases, we have developed a bump hunting method using probabilistic and statistical methods as shown in the previous study. By specifying a pureness rate in advance, a maximum capture rate will be obtained. In finding the maximum capture rate, we have used the decision tree method combined with the genetic algorithm. Then, a trade-off curve between the pureness rate and the capture rate can be constructed. However, such a

trade-off curve could be optimistic if the training data set alone is used; we might obtain an illusional curve when the number of samples is small and the number of explanation variables is large to some extent. Therefore, we should be careful in assessing the accuracy of the trade-off curve. Using the accuracy evaluation procedures such as the cross validation or the bootstrapped hold-out method combined with the training and test data sets, we have shown that the actually applicable trade-off curve can be obtained. We have also shown that an attainable upper bound trade-off curve can be estimated by using the extreme-value statistics because the genetic algorithm provides many local maxima of the capture rates with different initial values. However, this curve may also be optimistic because we use the training data set alone. In this paper, we propose a new genetic algorithm to tackle this problem. In the procedure of the new genetic algorithm, the evaluation to each evolution procedure is applied to the test data results rather than to the training data results. Due to this expansion for the genetic algorithm, we can expect the accuracy for the actually applicable upper bound for the trade-off curve by applying the extreme-value statistics by the new accuracy evaluation procedures such as the cross validation or the bootstrapped hold-out.

17. 朝廣雄一，小野廣隆，宮野英次:
“最小出次数最小化問題の各種グラフクラスに対する計算複雑さ”, 情報処理学会アルゴリズム研究会情処研報, 2008 年 3 月.
概要: Given an undirected graph with edge weights, we are asked to find an orientation, i.e., an assignment of a direction to each edge, so as to minimize the weighted maximum outdegree in the resulted directed graph. The problem is called MMO, and is a restricted variant of the well-known minimum makespan problem. As previous studies, it is shown that MMO is in $\mathcal{P}$ for trees, weak $\mathcal{NP}$-hard for planar bipartite graphs, and strong $\mathcal{NP}$-hard for general graphs. There are still gaps between those graph classes. The objective of this paper is to show tight thresholds of complexity: We show that MMO is (i) in $\mathcal{P}$ for cactuses, (ii) weakly $\mathcal{NP}$-hard for outerplanar graphs, and also (iii) strongly $\mathcal{NP}$-hard for P_4-bipartite graphs. The latter two are minimal superclasses of the former. Also, we show the $\mathcal{NP}$-hardness for the other related graph classes, diamond-free, house-free, series-parallel, bipartite and planar.

学会大会等

1. 下入佐真一，朝廣雄一，宮野英次:
“作業時間制約付き移動物体回収問題の $\mathcal{NP}$ 困難性”, 電気関係学会九州支部連合大会講演論文集, 12-1A-01, 2004 年 9 月.

2. 池上佳一，宮野英次:
“三者が互いに関連付けられた間接認証方式の提案”, 火の国情報シンポジウム 2005, 2005 年 3 月.

3. 朝廣雄一，古川哲也，宮野英次:
“サイズ 2 の最小ブロック転送問題に対する近似アルゴリズム”, 夏の LA シンポジウム, pp.23-1 – 10, 2005 年 7 月.

4. 大井伸哉，行實隆広，廣瀬英雄:
“GA を用いた Bump hunting 法に関する一考察”, 電気関係学会九州支部第 58 回連合大会, p.13-1P-06, 2005 年 9 月.

5. 行實隆広，大井伸哉，廣瀬英雄:
“GA を用いた Bump hunting の特性と最適値の統計的近似計算法”, 電気関係学会九州支部第 58 回連合大会, p.13-1P-07, 2005 年 9 月.

6. 廣瀬英雄:
“2005 年度 IEEE Reliability Society Japan

Chapter Awards (論文賞)", 情報システム学研究科シンポジウム，第 10 回信頼性とシステム安全学, 2006 年 2 月.

7. 廣瀬英雄，北村葵:
"性格から血液型を分類する危うさが示唆するもの: 決定木への注意", 電気関係学会九州支部第 59 回連合大会, p.12-2A-01, 2006 年 9 月.

8. 大井伸哉，行實隆広，宮野英次，廣瀬英雄:
"最適 tree 探索の確率的一方法", 電気関係学会九州支部第 59 回連合大会, p.12-2A-05, 2006 年 9 月.

9. 朝廣雄一，古川哲也，池上佳一，宮野英次:
"サイズ 3 の最小ブロック転送問題の近似困難性", 電気関係学会九州支部第 59 回連合大会, p.10-2A-08, 2006 年 9 月.

10. 吉牟田拓朗，宮野英次，宮崎修一:
"オンライン TSP アルゴリズムに対する下限について", 電気関係学会九州支部第 59 回連合大会, 確率的な終了時刻最小化スケジューリング問題の近似可能性, pp.10-2A-07, 2007 年 9 月.

11. 小川智雄，宮野英次:
"確率的な終了時刻最小化スケジューリング問題の近似可能性", 電気関係学会九州支部第 60 回連合大会, p.09-1A-03, 2007 年 9 月.

12. 清成悠貴，宮野英次，宮崎修一:
"試問予定表作成問題の制約付きモデルに対する NP 困難性", 電気関係学会九州支部第 60 回連合大会, p.09-1A-04, 2006 年 9 月.

13. 三溝和明，宮野英次:
"最大 2 クラン問題の近似可能性と近似不可能性", 電気関係学会九州支部第 60 回連合大会, p.09-1A-05, 2007 年 9 月.

14. 河原憲一，朝廣雄一，宮野英次:
"オンラインソーティングバッファに対する FIFO アルゴリズム", 電気関係学会九州支部第 60 回連合大会, p.09-1A-06, 2007 年 9 月.

15. 中村拓，朝廣雄一，宮野英次:
"2 色限定ペイントショップ問題に対する貪欲法と近似解法", 電気関係学会九州支部第 60 回連合大会, p.09-1A-07, 2007 年 9 月.

16. 廣瀬英雄，行實隆広，宮野英次:
"bump hunting における trade-off 曲線と recall precision 曲線との関係", 電気関係学会九州支部第 60 回連合大会, p.11-1A-12, 2007 年 9 月.

17. Yuichi Asahiro, Eiji Miyano, Taku Nakamura:
"Note on Approximation of 1-Regular 2-Color Paintshop Problem", INFORMS Annual Meeting, 2007 年 11 月.

18. 廣瀬英雄，行實隆広:
"テストデータ評価を用いた新しい遺伝的アルゴリズムによるバンプ探索法", 電子情報通信学会 2008 年総合大会, D-4-5, 2008 年 3 月.

19. 朝長英彦，行實隆広，廣瀬英雄:
"新しい遺伝的アルゴリズムによるバンプ探索トレードオフ曲線の精度: シミュレーションデータ", 電子情報通信学会 2008 年総合大会, D-4-6, 2008 年 3 月.

20. 行實隆広，朝長英彦，廣瀬英雄:
"新しい遺伝的アルゴリズムによるバンプ探索トレードオフ曲線の上限について：シミュレーションデータ", 電子情報通信学会 2008 年総合大会, D-4-7, 2008 年 3 月.

21. 出口拓也，行實隆広，廣瀬英雄:
"新しい遺伝的アルゴリズムによるバンプ探索のトレードオフ曲線：実データの場合", 電子情報通信学会 2008 年総合大会, D-4-8, 2008 年 3 月.