ERATO/基盤(S) 離散構造処理系プロジェクト

# プロジェクトの近況と今後の展望

平成27年12月22日

研究総括　湊　真一
（北海道大学 情報科学研究科）

# 離散構造処理系プロジェクトの基本構想

- 様々な工学的応用を持つ基盤技術として
**「離散構造処理系」**に着目し、研究開発を行う

システム設計自動化　データマイニングと知識発見

大規模システム故障解析　機械学習と自動分類

制約充足問題　生命情報科学　web情報解析

工学的応用

→ 社会への影響大

本研究の対象領域

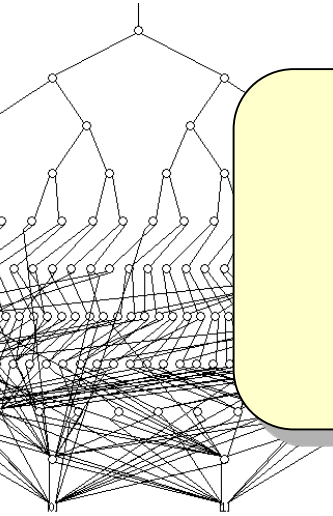
**離散構造処理系**

集合理論　記号論理　帰納的証明

組合せ論　グラフ理論　確率論

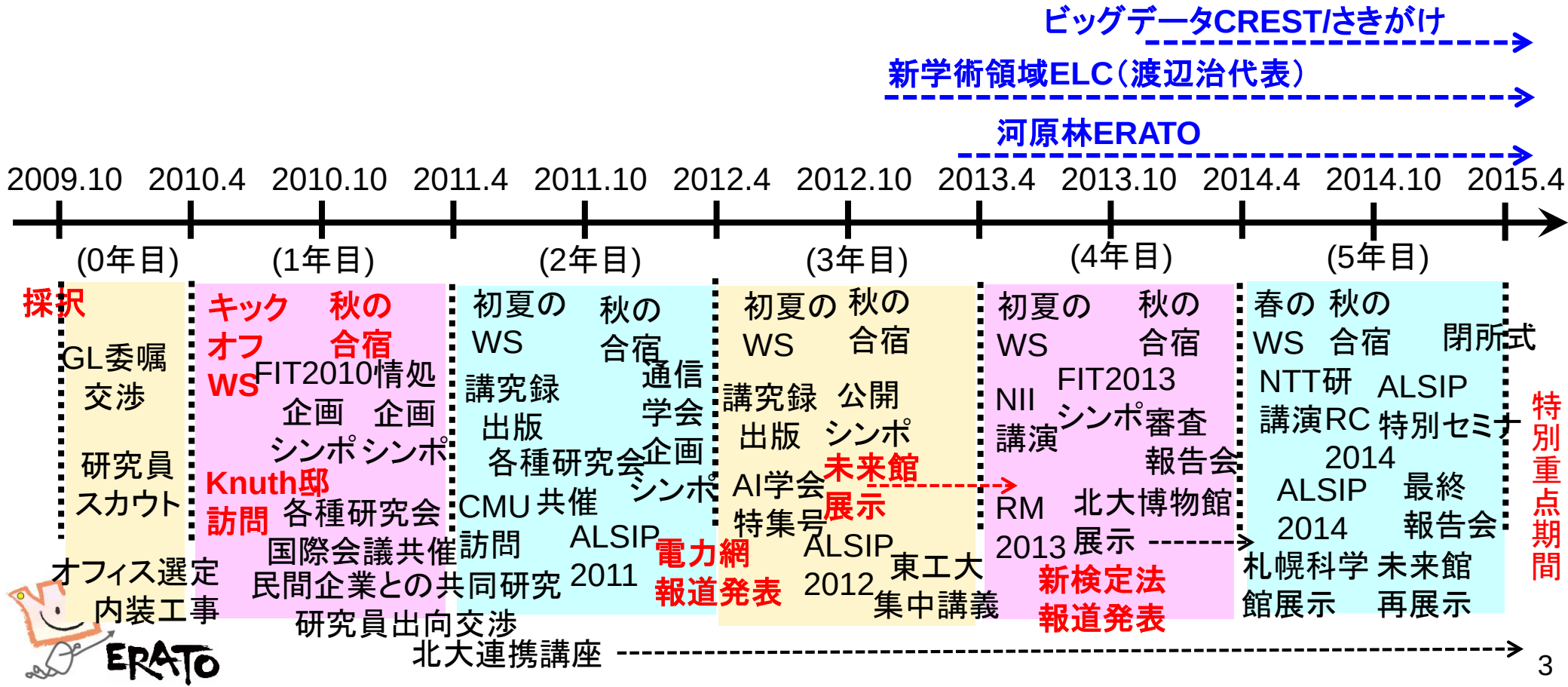
性能向上
(10倍～100倍以上)

数学的概念構造



# 湊ERATOプロジェクトの活動経過

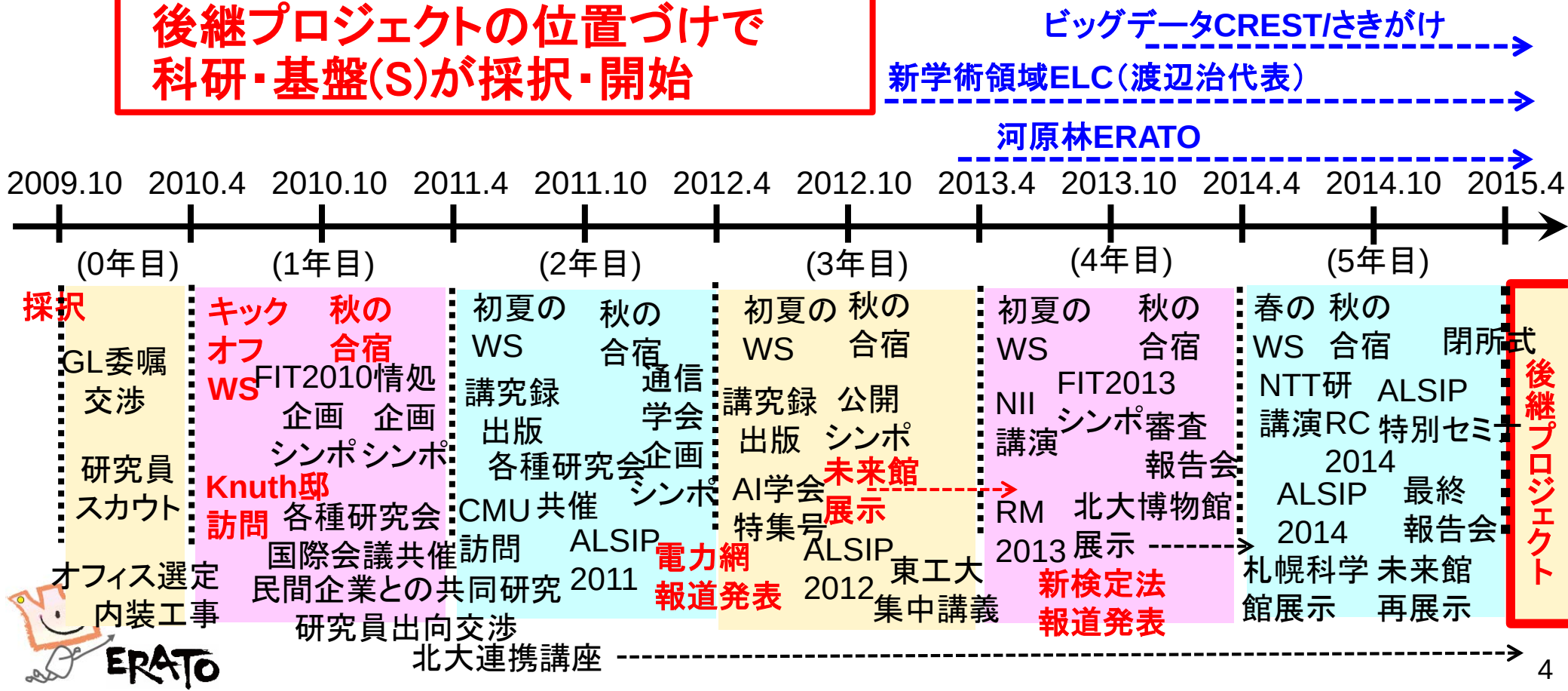
- 昨年度までで5年間のプロジェクト期間を成功裏に終了
  - 最終審査（予備審査）で暫定A評価を獲得
  - 現在、1年間のERATO「特別重点期間」が認められている。
    （ポスドク雇用ソフトランディングと後継プロジェクト準備の期間）



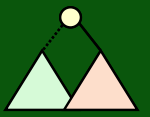
# 湊ERATOプロジェクトの活動経過

- 昨年度までで5年間のプロジェクト期間を成功裏に終了
  - 最終審査（予備審査）で暫定A評価を獲得
  - 現在、1年間のERATO「特別重点期間」が認められている。
    （ポスドク雇用ソフトランディングと後継プロジェクト準備の期間）

**後継プロジェクトの位置づけで**
**科研・基盤(S)が採択・開始**

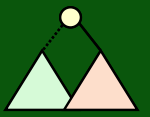


# 基盤(S)「離散構造処理系」の主要研究課題



- 離散構造処理系の基盤アルゴリズム技術の確立とソフトウェアの整備
  - 電力・道路・鉄道・通信，避難所割り当て，選挙区割りなどの社会インフラ問題への応用に必要な基盤アルゴリズムの研究開発
  - BDD/ZDDの新しい派生形とその処理系の研究開発・ツール化
- 離散構造処理系による効果的な探索法および列挙索引化法の開発
  - ERATOでの研究成果を受けつぎ，新しい探索・列挙索引化アルゴリズムの研究開発を継続して進める.
- Art層コミュニティの維持発展、および国家プロジェクトとの連携
  - セミナ、ワークショップ、学会共催イベント、国際ワークショップ開催，学会誌・単行本の執筆、アウトリーチ活動を積極的に行う.
  - ビッグデータCREST/さきがけ，電力CREST、河原林ERATO，新学術「計算限界」，海外研究機関との情報交流や研究者交流

# 基盤(S)「離散構造処理系」の主要研究課題



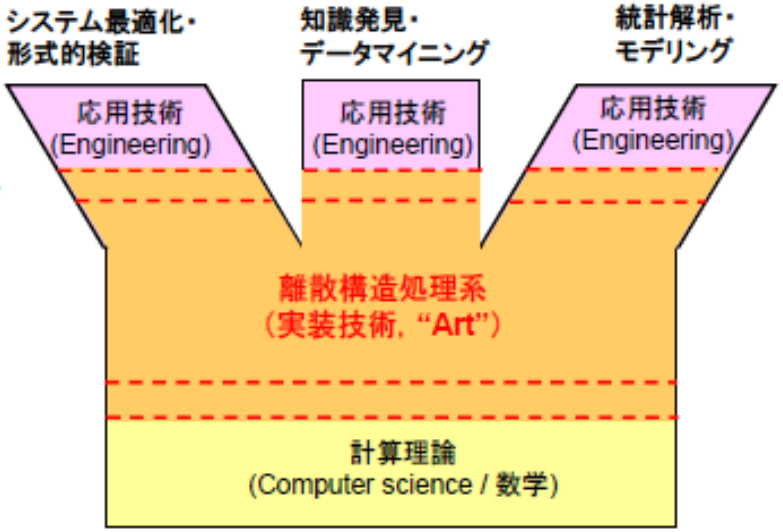
- 離散構造処理系の基盤アルゴリズム技術の確立とソフトウェアの整備
  - 電力・道路・鉄道・通信，避難所割り当て，選挙区割りなどの社会インフラ問題への応用に必要な基盤アルゴリズムの研究開発
  - BDD/ZDDの新しい派生形とその処理系の研究開発・ツール化
- 離散構造処理系による効果的な探索法および列挙索引化法の開発
  - ERATOでの研究成果を受けつぎ，新しい探索・列挙索引化アルゴリズムの研究開発を継続して進める.
- Art層コミュニティの維持発展、および国家プロジェクトとの連携
  - セミナ、ワークショップ、学会共催イベント、国際ワークショップ開催，学会誌・単行本の執筆、アウトリーチ活動を積極的に行う.
  - ビッグデータCREST/さきがけ，電力CREST、河原林ERATO，新学術「計算限界」，海外研究機関との情報交流や研究者交流

# 基盤(S)申請書(抜粋)

予算規模はERATOの10分の1
→ ERATOと同じことはできない

…点及び予想される結果と意義

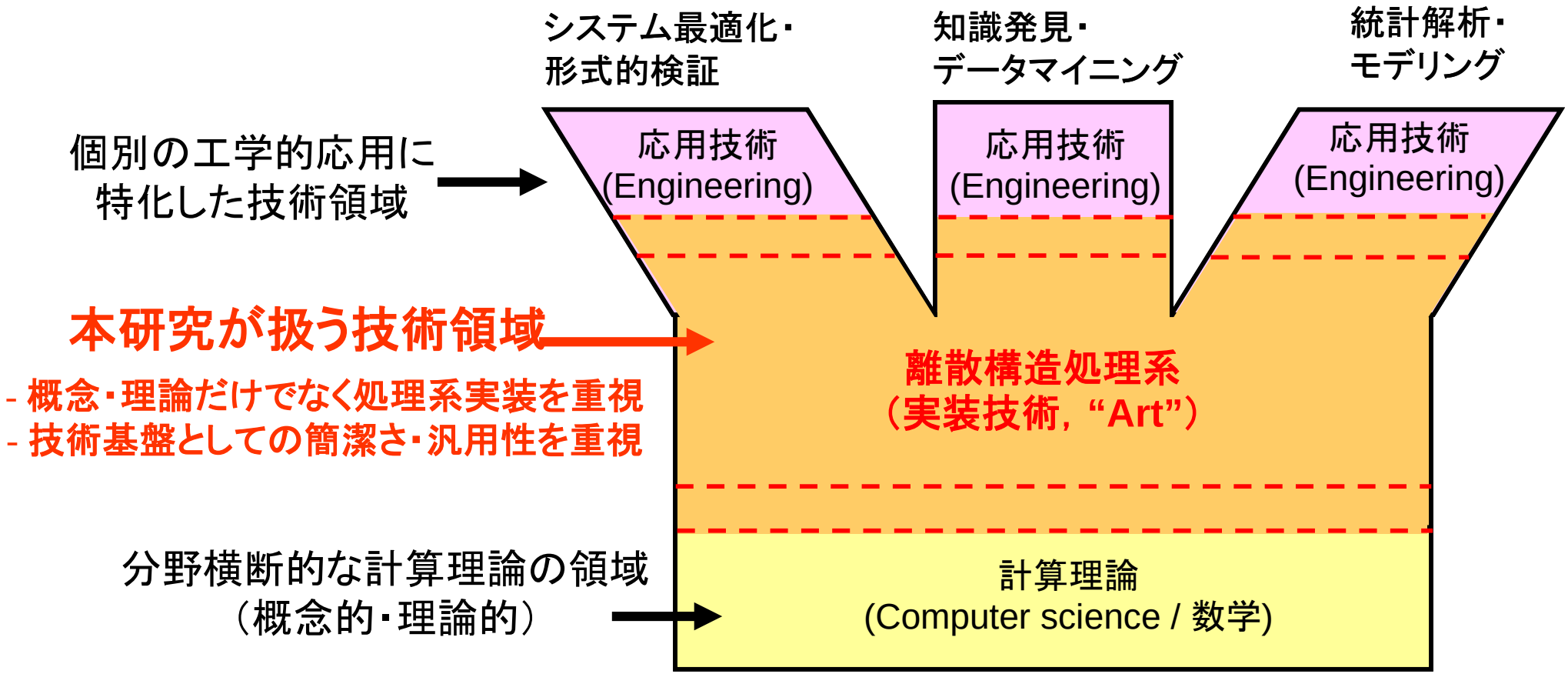
…示す．
…論の
領域がある．一方で様々な工学的応用に特化した
技術領域が並列に存在している．それらの中間層として，
離散構造処理系を扱う「Art 層」の研究者コミュニティを
育成してきた．ここでは，理論的成果だけでなく実用に
耐えるアルゴリズムを実装することを重視するが，各分野
に固有の問題にアドホックに対応するのではなく，技術
基盤としての簡潔さや汎用性を重視する．湊 ERATO
では**科学(Science)と工学(Engineering)をつなぐ「Art 層」を強化し，理論研究者と各応用分野の技術者が集まる「場」を創出すること**を提唱しており，さらにこれまで意思疎通がなかった異分野同士の技術交流により，新しい研究成果の源泉となるアイデアを醸成してきた．近年，河原林 ERATO や新学術「計算限界」など隣接分野の大型プロジェクトが開始され，最近ではビッグデータ CREST/さきがけに湊 ERATO 関係者が複数採択されるなど，**Art 層の周辺部では後継プロジェクトへの発展に成功しつつある．**それらの周辺領域については，それぞれのプロジェクトに研究開発を任せる一方で，Art 層のコアの部分では，これまでのように離散構造に関連する**研究者が集まる「場」を継続的に提供し，競争力の源泉となるアイデアを醸成し続けること**が，情報科学・計算機工学の分野全体の発展に資すると考える．




本研究の学術的な特色としては，種々の離散構造のモデルに関して，列挙と圧縮索引化をベースとする演算処理系を研究開発することが挙げられる．計算機の活用において，**「最適化と列挙は車の両輪」であり，最適化の技法に関しては世界的にかなり研究が進んでいるが，列挙系の技法はそれほどなされておらず，伝統的に日本が強い分野である．**これまでの ERATO で構築した「場」を活かして，列挙と圧縮索引化の分野で世界をリードし，さらに発展させることを本研究の目標とする．

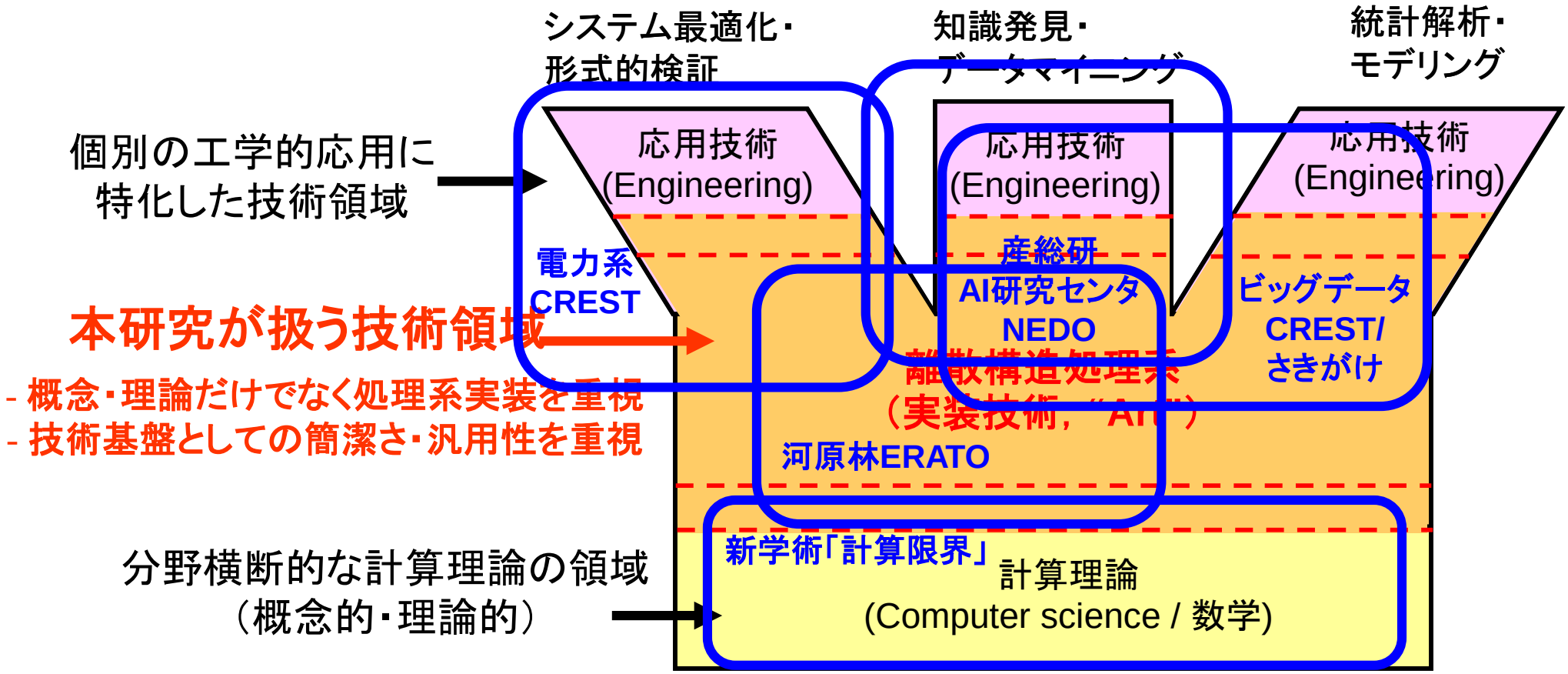
# 基盤(S)「離散構造処理系」の研究構想

**Art層の周辺部では後継プロジェクトへの発展に成功しつつある．**
**ボトムアップな科研費で<u>Art層のコア部分に研究者が集まる「場」を継続的に提供し</u>，競争力の源泉となるアイデアを醸成し続ける**



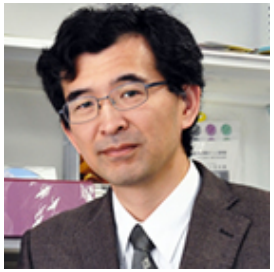
# 基盤(S)「離散構造処理系」の研究構想

**Art層の周辺部では後継プロジェクトへの発展に成功しつつある．**
**ボトムアップな科研費でArt層のコア部分に研究者が集まる「場」を継続的に提供し，競争力の源泉となるアイデアを醸成し続ける**



# 基盤(S)「離散構造処理系」の実施体制

代表者:



**湊　真一(北大)**
- 基盤アルゴリズムの研究開発
- SeqBDD、πDD処理系の開発
- ワークショップ・セミナ開催
- 研究計画全体の推進マネジメント

研究協力者:

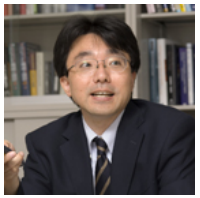
**渡辺治(東工大)** - 新学術「計算限界」
**河原林健一(NII)** - ERATO「巨大グラフ」
**林泰弘(早稲田大)** - 電力CREST最強チーム
**加藤直樹、羽室行信(関学大)**
- CREST「革新アルゴリズム」
**佐久間淳(筑波大)** - CREST「セキュリティ」
**竹田正幸(九大)** - サイバーフィジカル
**伊藤公人(北大)** - 人獣共通感染症
**R.Raman(レスター大)** - 簡潔データ構造
**A.Darwiche(UCLA)** - 確率グラフィカルモデル

分担者:

**津田 宏治(東大)**
- 統計多重検定(LAMP)への応用
- CREST「離散構造統計学」との連携

**鷲尾 隆(阪大)**
- 機械学習と統計モデリングへの応用
- 劣モジュラ解析・データ次元の解析

**宇野 毅明(NII)**
- 超高速列挙アルゴリズム技法
- CREST「データ粒子化」との連携



**有村 博紀(北大)**
- 超高速な半構造マイニングへの応用
- グラフ列挙と索引化の基礎技法

**堀山 貴史(埼玉大)**
- 種々の離散構造モデルへの拡張
- 新学術「計算限界解明」との連携

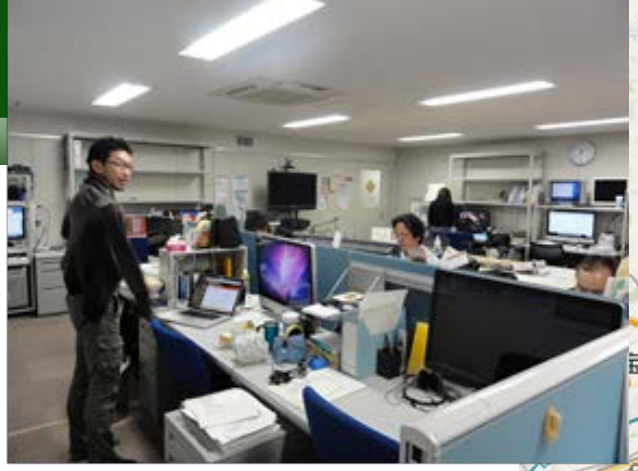
# ERATOでのサテライト拠点の設置

- 札幌にメインオフィスを設置
  - 北大キャンパス内に設置
  - 空港・駅からの交通至便
  - 北大の関連研究者とも交流が容易

  → **研究に集中できる優れた立地環境**
- 東京・関西地区にサテライト拠点を設置
  - **東工大と大阪梅田に設置**
  - 九大に連絡拠点
  - 高品質TV会議システムを導入し、十分な意思疎通を図る。

**フルタイムの若手研究者は極力、札幌に常駐させて、活性化・相乗効果を図る。**

# 大阪サテラボ



- ERATO研究成果のインキュベータ拠点として機能。
- 「関西ダブルCRESTラボ」として存続
  - 加藤CREST＆宇野CREST
- 金融ファンドマネジメントを行うMagne-Max 社を起業
  - 上記CRESTメンバが経営にコミット
  - 三菱UFJ信託への運用助言を受注
  - ヤフー（株）と資本提携（11月～）
  - ERATOメンバOBが執行役員として加入
    **（研究成果により雇用を創出！）**

# 新・東京サテライト拠点の設置

- 東京都心（JR神田駅・徒歩3分）に関連研究者が情報交換を行うための「連携研究拠点」を設置。
  - 民間オフィスビル・1フロア約80m$^2$
  - 新幹線から10分圏内。NIIから徒歩圏
    首都圏の多くの大学から好アクセス。
  - 割安な地区（大阪サテラボと同価格帯）
  - 基盤（S）予算で賃料を負担。分担者の宇野先生が現地責任者に（NIIが借主）
  - 10月に候補物件を絞り込み
    不動産業者と契約交渉を開始
    → **12月15日より正式に使用開始。**
    （現在、ライフライン・什器の整備中）
- **情報系研究者の交流拠点として提供**

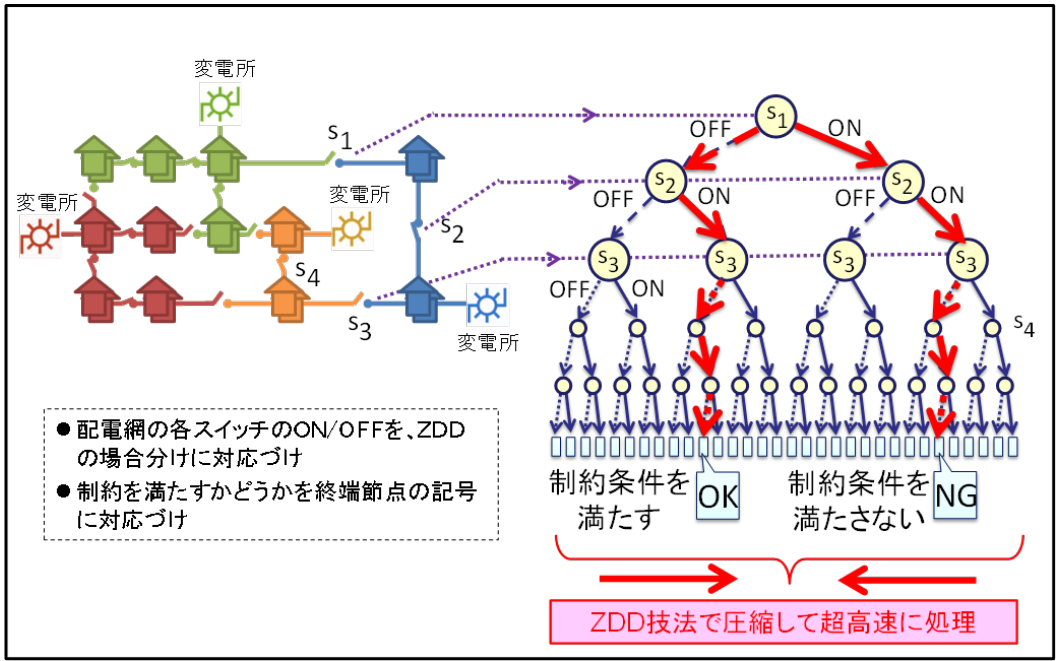
# 新・東京サテライト拠点の設置

# 電力網の設計・最適化への応用

- 林 泰弘 教授（早稲田大）との共同研究
  - 電力系スマートグリッド業界のリーダ的存在（国の委員など多数）
  - 電力網最適化の研究で1990年代より湊と協力関係
  - 大震災後、最重要の課題に（長期的に不足する電力を、再生可能エネルギーで補うために必須の電力網解析・制御技術）
- **可能な網構成の網羅的列挙に世界で初めて成功**
  - 損失最小化＆障害復旧可能性判定 [IEEE Trans.SmartGridに2編]

IF: 4.25

- 林教授が電力系CRESTの「最強チーム」メンバに選出
  - CREST+ERATO代表者が東京電力の研究開発部門を訪問し、協力関係を構築
  - ERATO研究成果を電力系CRESTを通じて現実の電力網に提供可能に



# 選挙区割りへの応用

川原先生（NAIST）、堀田先生（文教大）他との共同研究

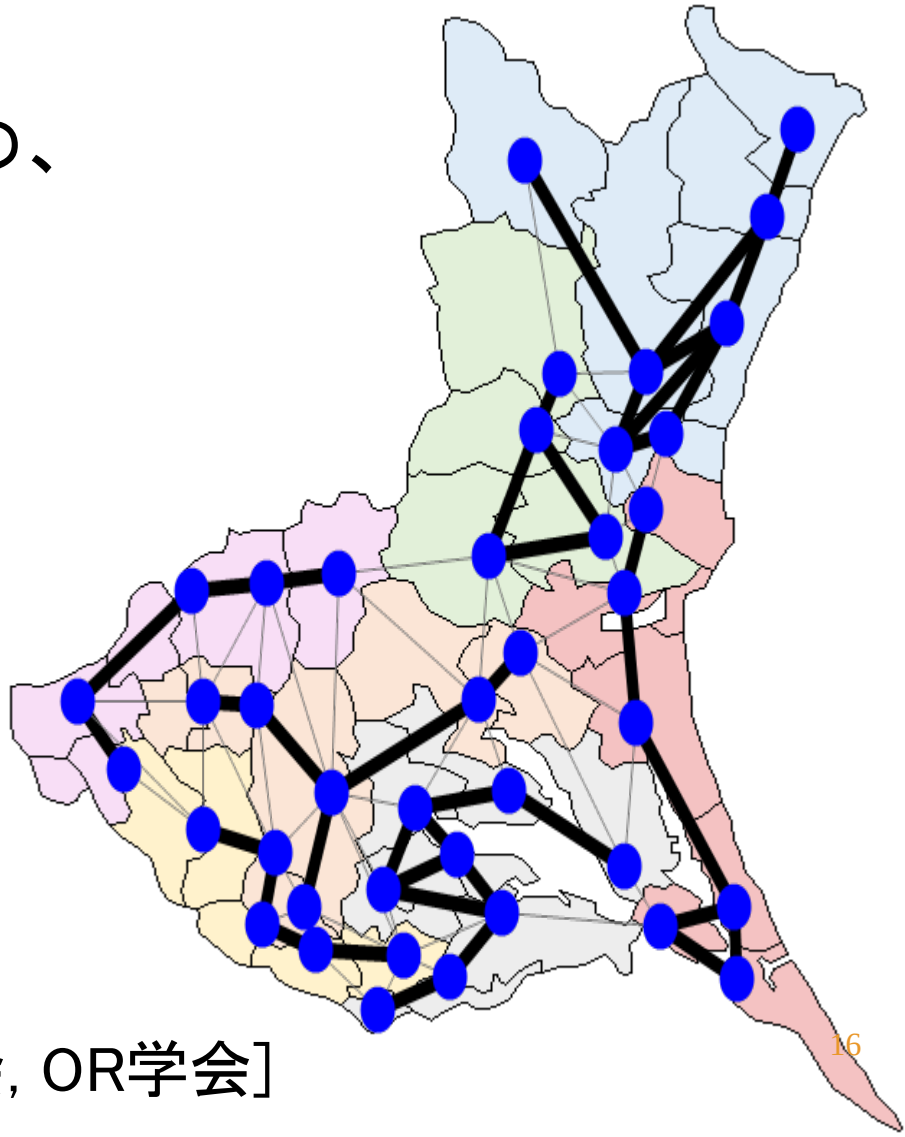
一票の格差最小だけではなく
地理的・社会的制約条件があるため、
列挙して解の分布を示すことに
社会的意味がある

茨城県の例：
頂点数41、辺数87、分割数7
（41市区町村,7選挙区）

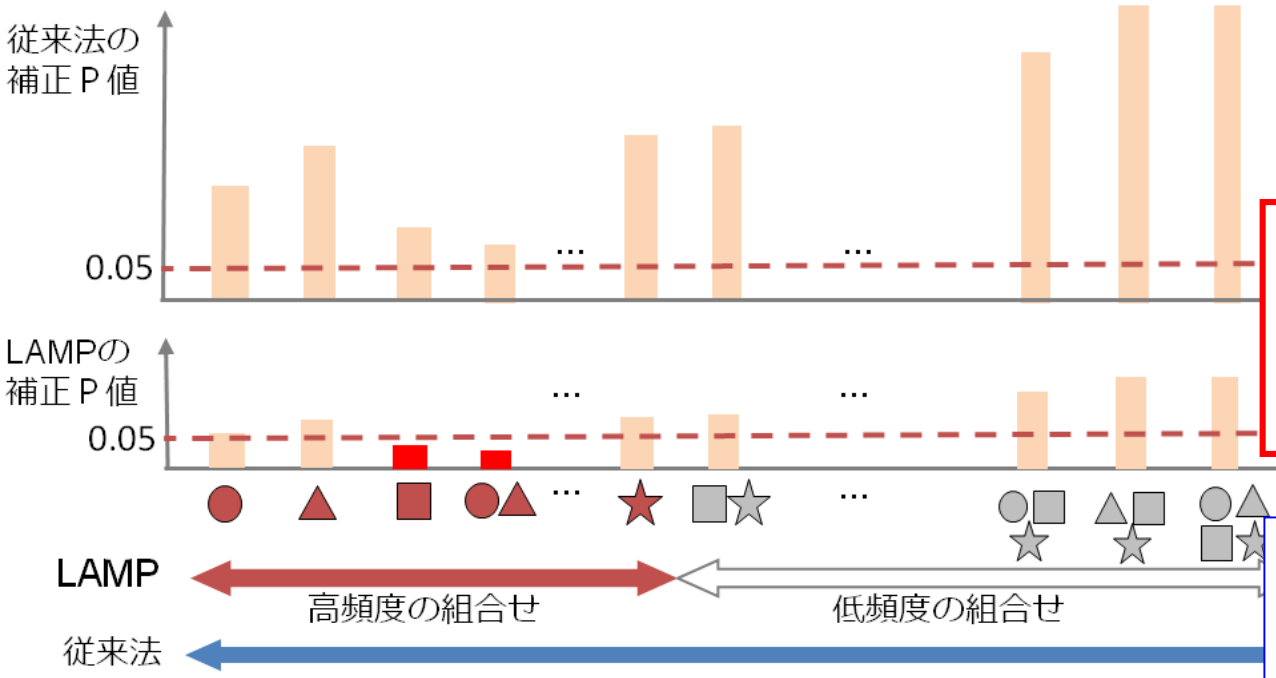
解の個数は**11,893,998,242,846個**

その中から格差が1.4倍以下のものを
抜き出すと **25,730,669個**
（1925.21秒で計算）

[2015選挙学会，OR学会]



# ビッグデータから新たな知識発見をもたらす統計手法




## 多重検定補正法「LAMP」の提案

**列挙アルゴリズム技術があらゆる実験系科学に貢献できる道が開かれた**

**米国科学アカデミー紀要（2013年8月）**

PNAS

# Statistical significance of combinatorial regulations

Aika Terada[a,b,c], Mariko Okada-Hatakeyama[d], Koji Tsuda[c,e,1], and Jun Sese[a,b,1]

[a]Department of Computer Science and [b]Education Academy of Computational Life Sciences, Tokyo Institute of Technology, Meguro-ku, Tokyo 152-8550, Japan; [c]Minato Discrete Structure Manipulation System Project, Exploratory Research for Advanced Technology, Japan Science and Technology Agency, Sapporo, Hokkaido 060-0814, Japan; [d]Laboratory for Integrated Cellular Systems, RIKEN Center for Integrated Medical Sciences (IMS-RCAI), Yokohama, Kanagawa 230-0045, Japan; and [e]Computational Biology Research Center, National Institute of Advanced Industrial Science and Technology, Koto-ku, Tokyo 135-0064, Japan

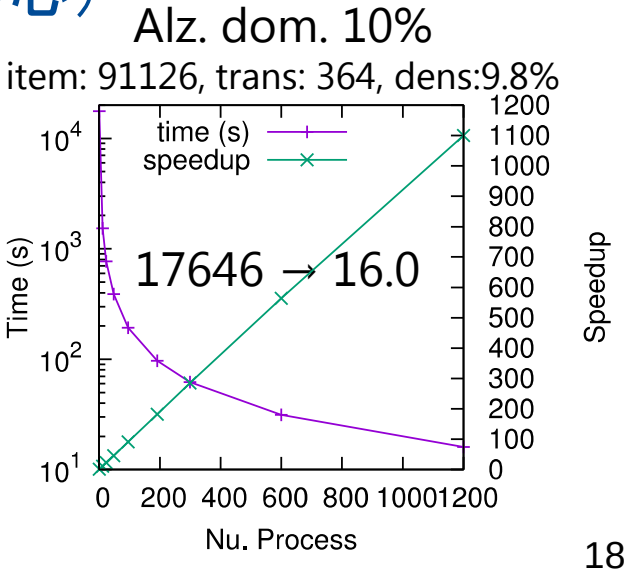
Edited by Wing Hung Wong, Stanford University, Stanford, CA, and approved July 3, 2013 (received for review February 4, 2013)

**More than three transcription factors often work together to enable cells to respond to various signals. The detection of combinatorial regulation by multiple transcription factors, however, is not only**

deliberately excluding such tests. Here, we propose an efficient branch-and-bound algorithm, called the "limitless arity multiple-testing procedure" (LAMP). LAMP counts the exact number of

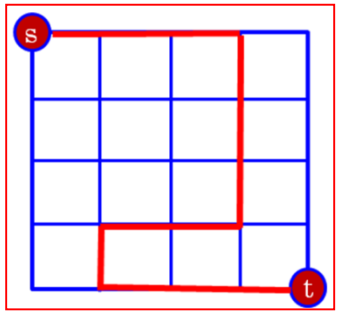
# LAMPの高速化

- Original LAMP: 基本原理の考案と初期実装 [PNAS2013]
  - 東京サブグループの成果（津田先生、瀬々先生、寺田さん他）
  - 「出現頻度が低い組合せは決して有意にならない」→ LCMの活用
- LAMP-ver.2: 単調性を利用した高速化 [ECML-PKDD2014]
  - 湊と宇野先生（LCM開発者）がアルゴリズム開発に参入
  - **今まで数日かかっていた例題が数分で解けることも**
- MP-LAMP: 分散並列LCMによる大規模並列化 [2015投稿中]
  - 東京サブグループの最近の成果（美添さん中心）
  - **1200コアで約1000倍の高速化**
    （安価なPCクラスタでも効果を発揮）
  - 分散並列LCMは、LAMP以外でも
    幅広い応用の可能性



# 「フカシギの数え方」の展示と数え上げ世界記録

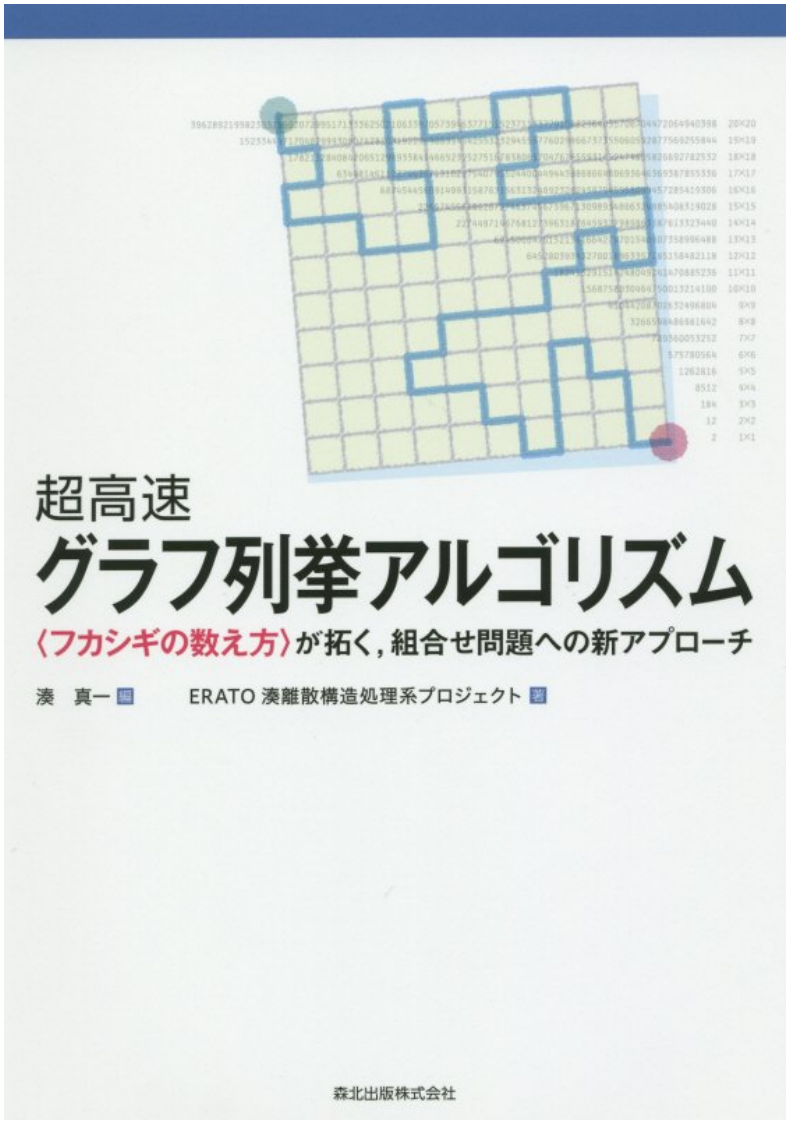
YouTube
未来館チャンネルで
**1,737,224ビュー**
（非公式サイト含むと
300万回超え）




**情報科学の重要性を
小中高生や一般市民に
強く印象付けることに成功**

# Graphillion専門書の出版

- ERATOプロジェクトの成果として超高速グラフ列挙索引化ソフト「Graphillion」を開発・公開
- Graphillionの活用法と内部のアルゴリズム技法をまとめた専門書を出版
  - ERATO関係者13名による共同執筆
- 2015年4月8日発売開始
  - 予約好調につき発売前に重版増刷、10月に第3刷発行
- **英語版の執筆も準備中**



# おねえさんの LINEスタンプ

**2015年10月より 公開中**

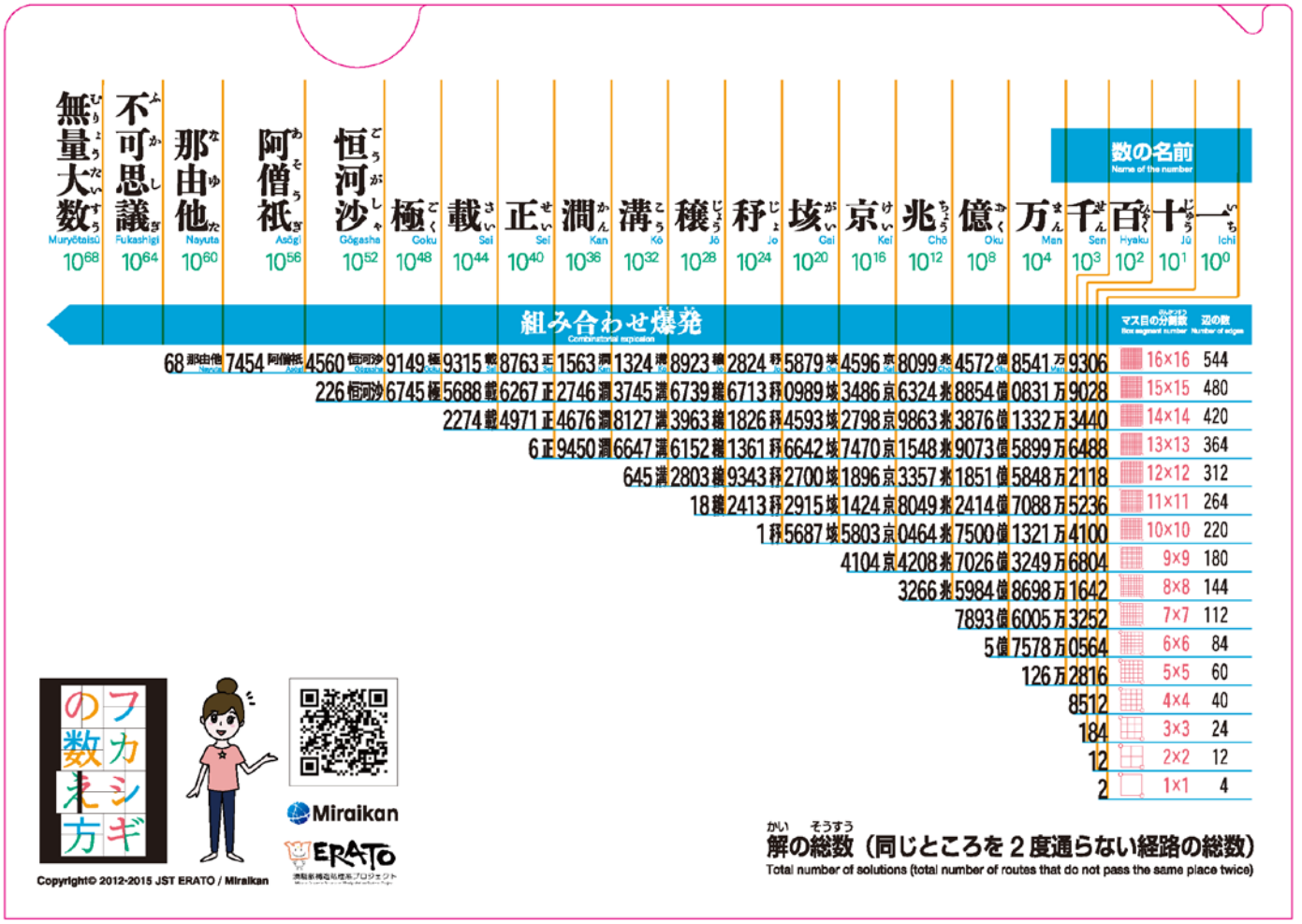
120円で40パタンが
無期限で使い放題

未来館で初の
公式LINEスタンプ
（JSTでも初）

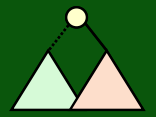
# フカシギ数表クリアファイルの商品化

- 未来館ショップで一般販売開始（12/15～）
  - 税込270円
  - 2面刷りで
    数字だけが
    見え隠れ

元素周期表や
太陽系の
図柄と並んで
巨大数が陳列棚に



# 本研究の技術的スコープ



この部分で
着実な成果

現在の
ZDD

基本的
成果

**BDD/ZDDが有効な
世界への適用拡大**

データマイニング 機械学習

高度な検索　　etc.

**ERATOプロジェクトから
波及した応用領域
プロジェクトが展開中**

さらに
成果を展開

(組合せ集合モデル)

(より高次のモデル)

**より高次の成果**

高度な
ZDD風の
離散構造

- ワード単位データ
- 文字列
- 順列
- 分割
- ツリー
- ネットワーク

**将来に
つながる
成果**

新しい演算
体系の構築

**より高次のデータモデルを
用いた実用的問題への応用**

文字列データ解析

数値データ処理

ツリーや半構造データの解析

# まとめと今後の展望

- 過去5年間のERATO湊プロジェクトの成果を引継ぎ、理論研究者と各応用分野の研究者が集まる「場」を、さらに発展強化させる。
- あらゆる計算機応用の基盤技術であり、理論と応用の交差点（→Art層）を形成している「離散構造処理系」にシャープに焦点を絞り、堅実に研究を進める。
- **「圧縮-列挙-索引化-演算」**の中堅若手研究者グループが活性化しているので、今後も燃料投下を続けたい。

北大（札幌）の地理的特徴を活かし、情報系の多才な人材が離合集散する「梁山泊」ネットワークを形成

# 全解列挙型SATソルバー

戸田貴久

toda.takahisa@is.uec.ac.jp

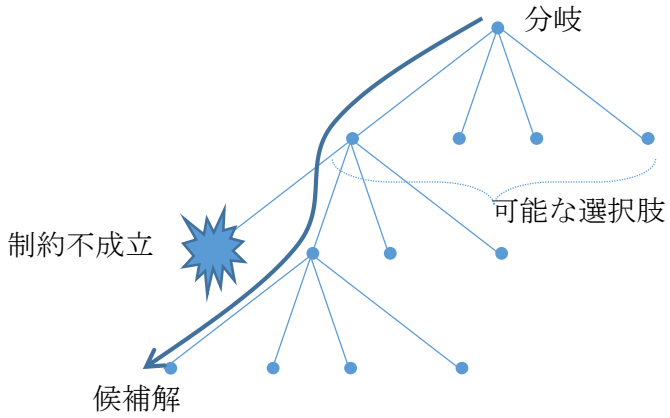
電気通信大学情報システム学研究科

共同研究：宋剛秀（神戸大学）

情報系WINTER FESTA　平成27年12月22－23日

## All Solutions SAT (AllSAT)

- 命題論理式の充足可能な値割当てをすべて計算する問題
- 単純な問題設定、幅広い問題を含む



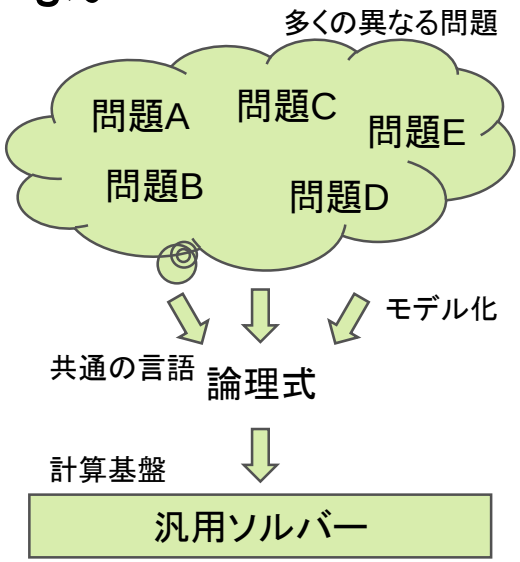

## 課題

- いくつかのアルゴリズムが提案されているが・・・
  - 網羅的な評価がない
  - Publicなソルバーがほぼ皆無
  - 応用の場で必ずしも最新技法使われていない

## これまで・これからの取り組み

- 主要な技法の調査、実装、評価結果の公開

  http://www.sd.is.uec.ac.jp/toda/code/allsat.html
- もっと多くの良い応用を探しています
- より大規模、困難な問題に挑戦



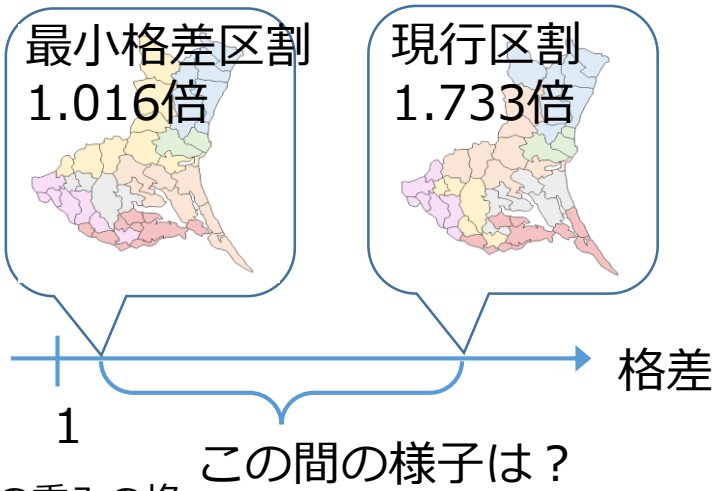
# 種々の評価尺度を用いた選挙区割の比較

川原 純　奈良先端科学技術大学院大学
堀田 敬介　文教大学
堀山 貴史　埼玉大学
湊 真一　北海道大学

- 研究目的: 一票の格差が小さな選挙区割を列挙
  - 最小の区割を1つ求めるのは可能 [根本, 堀田 2005]

画定作業支援、現行選挙区のより詳細な評価のため区割の候補列挙をしたい

例: 茨城県
41市区郡、7選挙区




根本俊男，堀田敬介，衆議院小選挙区制における一票の重みの格差の限界とその考察，選挙研究，vol. 20, 136--147, 2005.

2010年国勢調査データ

# 比較　　例）　26. 京都府（6選挙区）

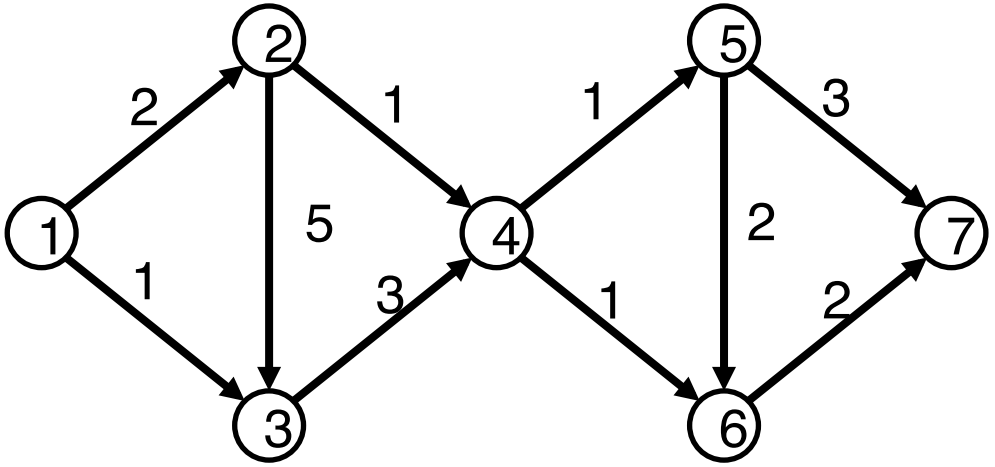
| | | 格差 | 人口頑健性 | | | | 親密度 | | | | | | | コンパクト | | | | | | | 標偏 | 乖離 | 未来 | | | 現在 | | | | 過去 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | 2010 | 2015 | 2020 | 2025 | maximin | | | | | | | | | | | | | | | | 頑健性 | | | 親密度 | コン | 標偏 | 乖離 | |
| | name | 解 | n | ratio | ratio | ratio | ratio | 1 | 2 | 3 | 4 | 5 | 6 | ave | 1 | 2 | 3 | 4 | 5 | 6 | ave | std | ave | 15 | 20 | 25 | mm | ave | ave | std | ave |
| 1 | 1 | 1 | 1 | 1.0815 | 1.0866 | 1.0936 | 1.1367 | 84.8% | 56.6% | 52.5% | 46.8% | 45.1% | 38.6% | 54.1% | 5 | 4 | 4 | 3 | 2 | 1 | 3.17 | 14,494 | 0.575 | 1 | 3 | 14 | 150 | 87 | 100 | 4 | 75 |
| 2 | 2 | 2 | 1 | 1.0881 | 1.0896 | 1.0918 | 1.1305 | 84.8% | 57.6% | 51.3% | 46.8% | 45.1% | 38.6% | 54.1% | 5 | 4 | 4 | 3 | 2 | 1 | 3.17 | 14,727 | 0.575 | 2 | 1 | 3 | 150 | 91 | 100 | 6 | 75 |
| 3 | 3_1 | 3 | 2 | 1.0900 | 1.0967 | 1.1304 | 1.1757 | 84.8% | 56.4% | 54.1% | 46.4% | 41.7% | 38.7% | 53.7% | 5 | 4 | 3 | 3 | 2 | 1 | 3.00 | 14,113 | 0.573 | 4 | 36 | 94 | 120 | 118 | 33 | 2 | 59 |
| 3 | 3_2 | | | 1.0900 | 1.0990 | 1.1304 | 1.1757 | 84.8% | 53.5% | 53.3% | 46.4% | 41.7% | 38.7% | 53.1% | 5 | 4 | 3 | 3 | 2 | 2 | 3.17 | 14,184 | 0.616 | 5 | 36 | 94 | 120 | 164 | 100 | 3 | 154 |
| 4 | 4_1 | 4 | 2 | 1.0992 | 1.1049 | 1.1097 | 1.1507 | 84.8% | 54.1% | 52.2% | 45.1% | 43.2% | 42.5% | 53.6% | 5 | 4 | 4 | 2 | 2 | 2 | 3.17 | 13,912 | 0.519 | 8 | 11 | 34 | 47 | 124 | 100 | 1 | 9 |
| 4 | 4_2 | | | 1.0992 | 1.1049 | 1.1097 | 1.1507 | 84.8% | 54.1% | 52.2% | 45.1% | 43.4% | 43.0% | 53.8% | 5 | 4 | 4 | 3 | 2 | 2 | 3.33 | 15,550 | 0.521 | 8 | 11 | 34 | 25 | 113 | 173 | 12 | 14 |
| 5 | 5 | 5 | 1 | 1.1013 | 1.0960 | 1.0919 | 1.1212 | 84.8% | 60.8% | 51.4% | 50.0% | 45.1% | 38.6% | 55.1% | 5 | 4 | 4 | 2 | 2 | 1 | 3.00 | 14,975 | 0.579 | 3 | 2 | 1 | 150 | 23 | 33 | 9 | 82 |
| 6 | 6 | 6 | 1 | 1.1031 | 1.1046 | 1.1304 | 1.1757 | 84.8% | 53.3% | 52.3% | 45.3% | 43.2% | 38.7% | 52.9% | 5 | 4 | 3 | 3 | 2 | 2 | 3.17 | 14,804 | 0.562 | 7 | 36 | 94 | 120 | 177 | 100 | 8 | 45 |
| 7 | 7_1 | 7 | 2 | 1.1035 | 1.1071 | 1.1105 | 1.1510 | 84.8% | 55.0% | 53.3% | 45.1% | 43.2% | 42.5% | 54.0% | 5 | 4 | 4 | 2 | 2 | 2 | 3.17 | 14,599 | 0.572 | 11 | 13 | 38 | 47 | 93 | 100 | 5 | 54 |
| 7 | 7_2 | | | 1.1035 | 1.1071 | 1.1105 | 1.1510 | 84.8% | 55.0% | 53.3% | 45.1% | 43.4% | 43.0% | 54.1% | 5 | 4 | 4 | 3 | 2 | 2 | 3.33 | 16,167 | 0.573 | 11 | 13 | 38 | 25 | 82 | 173 | 15 | 57 |
| 8 | 8_1 | 8 | 2 | 1.1036 | 1.1106 | 1.1378 | 1.1802 | 84.8% | 53.5% | 53.3% | 46.4% | 43.3% | 43.1% | 54.1% | 5 | 4 | 3 | 2 | 2 | 2 | 3.00 | 16,449 | 0.616 | 16 | 50 | 113 | 17 | 88 | 33 | 17 | 154 |
| 8 | 8_2 | | | 1.1036 | 1.1106 | 1.1378 | 1.1802 | 84.8% | 56.4% | 54.1% | 46.4% | 43.3% | 43.1% | 54.7% | 5 | 4 | 3 | 2 | 2 | 1 | 2.83 | 16,388 | 0.573 | 16 | 50 | 113 | 17 | 36 | 5 | 16 | 59 |
| 9 | 9 | 9 | 1 | 1.1043 | 1.1272 | 1.1477 | 1.1757 | 84.8% | 53.3% | 52.3% | 46.0% | 43.4% | 38.7% | 53.1% | 5 | 4 | 3 | 3 | 3 | 2 | 3.33 | 17,202 | 0.562 | 34 | 78 | 94 | 120 | 162 | 173 | 21 | 41 |
| 10 | 10 | 10 | 1 | 1.1047 | 1.1030 | 1.1028 | 1.1385 | 84.8% | 53.3% | 47.8% | 46.9% | 45.1% | 42.5% | 53.4% | 5 | 5 | 4 | 2 | 2 | 2 | 3.33 | 14,741 | 0.598 | 6 | 6 | 20 | 47 | 144 | 173 | 7 | 113 |
| 11 | 11_1 | 11 | 2 | 1.1051 | 1.1101 | 1.1145 | 1.1543 | 84.8% | 56.4% | 54.1% | 52.7% | 45.1% | 38.6% | 55.3% | 5 | 4 | 4 | 2 | 2 | 1 | 3.00 | 15,051 | 0.563 | 14 | 19 | 47 | 150 | 14 | 33 | 10 | 48 |
| 11 | 11_2 | | | 1.1051 | 1.1101 | 1.1145 | 1.1543 | 84.8% | 53.5% | 53.3% | 52.7% | 45.1% | 38.6% | 54.7% | 5 | 4 | 4 | 2 | 2 | 2 | 3.17 | 15,117 | 0.605 | 14 | 19 | 47 | 150 | 37 | 100 | 11 | 134 |
| 12 | 12 | 12 | 1 | 1.1099 | 1.1063 | 1.1036 | 1.1366 | 84.8% | 60.8% | 51.7% | 46.8% | 45.1% | 38.6% | 54.6% | 5 | 4 | 4 | 3 | 2 | 1 | 3.17 | 16,504 | 0.601 | 10 | 7 | 12 | 150 | 40 | 100 | 18 | 125 |
| 13 | 13_1 | 13 | 2 | 1.1125 | 1.1159 | 1.1209 | 1.1606 | 84.8% | 56.4% | 54.1% | 48.1% | 45.1% | 42.5% | 55.2% | 5 | 4 | 3 | 2 | 2 | 1 | 2.83 | 15,749 | 0.573 | 23 | 26 | 56 | 47 | 21 | 5 | 13 | 59 |
| 13 | 13_2 | | | 1.1125 | 1.1159 | 1.1209 | 1.1606 | 84.8% | 53.5% | 53.3% | 48.1% | 45.1% | 42.5% | 54.6% | 5 | 4 | 3 | 2 | 2 | 2 | 3.00 | 15,812 | 0.616 | 23 | 26 | 56 | 47 | 45 | 33 | 14 | 157 |
| 14 | 14 | 14 | 1 | 1.1159 | 1.1139 | 1.1127 | 1.1463 | 84.8% | 53.3% | 52.3% | 45.1% | 45.1% | 43.4% | 54.0% | 5 | 4 | 4 | 3 | 3 | 2 | 3.50 | 18,211 | 0.562 | 21 | 17 | 33 | 4 | 92 | 202 | 29 | 43 |
| 15 | 15 | 15 | 1 | 1.1169 | 1.1149 | 1.1378 | 1.1802 | 84.8% | 53.3% | 52.3% | 46.9% | 43.3% | 43.2% | 54.0% | 5 | 4 | 3 | 2 | 2 | 2 | 3.00 | 16,911 | 0.562 | 22 | 50 | 113 | 14 | 95 | 33 | 19 | 45 |
| 16 | 16_1 | 16 | 2 | 1.1176 | 1.1122 | 1.1085 | 1.1411 | 84.8% | 55.3% | 52.3% | 45.1% | 43.4% | 43.0% | 54.0% | 5 | 4 | 4 | 3 | 2 | 2 | 3.33 | 18,460 | 0.521 | 19 | 9 | 23 | 25 | 97 | 173 | 31 | 14 |
| 16 | 16_2 | | | 1.1176 | 1.1122 | 1.1085 | 1.1411 | 84.8% | 55.3% | 52.3% | 45.1% | 43.2% | 42.5% | 53.9% | 5 | 4 | 4 | 2 | 2 | 2 | 3.17 | 17,103 | 0.519 | 19 | 9 | 23 | 47 | 105 | 100 | 20 | 9 |
| 17 | 17_1 | 17 | 2 | 1.1243 | 1.1083 | 1.1044 | 1.1367 | 84.8% | 56.6% | 52.5% | 46.8% | 41.9% | 36.6% | 53.2% | 5 | 4 | 4 | 3 | 2 | 1 | 3.17 | 19,113 | 0.575 | 13 | 8 | 14 | 186 | 155 | 100 | 45 | 75 |
| 17 | 17_2 | | | 1.1243 | 1.1113 | 1.1027 | 1.1305 | 84.8% | 57.6% | 51.3% | 46.8% | 41.9% | 36.6% | 53.2% | 5 | 4 | 4 | 3 | 2 | 1 | 3.17 | 19,290 | 0.575 | 18 | 4 | 3 | 186 | 156 | 100 | 50 | 75 |
| 26 | 26 | 26 | 1 | 1.1322 | 1.1478 | 1.1615 | 1.1858 | 84.8% | 61.4% | 49.5% | 46.8% | 45.1% | 43.7% | 55.2% | 6 | 5 | 4 | 2 | 2 | 1 | 3.33 | 22,300 | 0.608 | 87 | 131 | 143 | 3 | 17 | 173 | 114 | 144 |
| 30 | 30_2 | | | 1.1343 | 1.1179 | 1.1027 | 1.1212 | 84.8% | 60.8% | 51.4% | 50.0% | 41.9% | 36.6% | 54.3% | 5 | 4 | 4 | 2 | 2 | 1 | 3.00 | 19,481 | 0.579 | 25 | 5 | 1 | 186 | 68 | 33 | 52 | 82 |
| 31 | 31_1 | 31 | 2 | 1.1362 | 1.1216 | 1.1107 | 1.1305 | 84.8% | 55.0% | 51.3% | 50.0% | 41.9% | 36.6% | 53.3% | 5 | 4 | 4 | 3 | 2 | 1 | 3.17 | 23,888 | 0.537 | 26 | 15 | 3 | 186 | 151 | 100 | 143 | 22 |
| 31 | 31_2 | | | 1.1362 | 1.1216 | 1.1107 | 1.1305 | 84.8% | 55.0% | 51.3% | 50.0% | 45.1% | 38.6% | 54.1% | 5 | 4 | 4 | 3 | 2 | 1 | 3.17 | 20,381 | 0.537 | 26 | 15 | 3 | 150 | 77 | 100 | 66 | 22 |
| 46 | 46_1 | 46 | 2 | 1.1442 | 1.1441 | 1.1444 | 1.1802 | 84.8% | 56.4% | 54.1% | 46.3% | 43.3% | 42.6% | 54.6% | 5 | 3 | 3 | 2 | 2 | 1 | 2.67 | 18,904 | 0.558 | 69 | 66 | 113 | 38 | 43 | 2 | 40 | 37 |
| 51 | 51_1 | 51 | 2 | 1.1455 | 1.1463 | 1.1499 | 1.1757 | 84.8% | 55.0% | 54.1% | 43.2% | 41.1% | 38.7% | 52.8% | 5 | 3 | 3 | 3 | 2 | 2 | 3.00 | 19,055 | 0.477 | 72 | 92 | 94 | 120 | 187 | 33 | 44 | 1 |
| 53 | 53_1 | 53 | 3 | 1.1464 | 1.1472 | 1.1484 | 1.1885 | 84.8% | 67.6% | 57.7% | 46.8% | 45.1% | 38.6% | 56.8% | 5 | 4 | 4 | 2 | 1 | 1 | 2.83 | 22,131 | 0.601 | 79 | 80 | 145 | 150 | 2 | 5 | 111 | 125 |
| 61 | 61_3 | | | 1.1519 | 1.1584 | 1.1658 | 1.1802 | 84.8% | 52.4% | 50.4% | 49.9% | 44.9% | 43.3% | 54.3% | 5 | 3 | 2 | 2 | 2 | 2 | 2.67 | 25,860 | 0.600 | 127 | 136 | 113 | 5 | 65 | 2 | 175 | 119 |
| 65 | 65_3 | | | 1.1565 | 1.1583 | 1.1591 | 1.1613 | 84.8% | 56.4% | 54.1% | 48.1% | 44.7% | 40.8% | 54.8% | 5 | 3 | 3 | 2 | 2 | 1 | 2.67 | 21,331 | 0.609 | 124 | 125 | 62 | 68 | 30 | 2 | 98 | 146 |
| 69 | 69_2 | | | 1.1598 | 1.1570 | 1.1645 | 1.1802 | 84.8% | 55.0% | 54.1% | 44.9% | 43.3% | 42.5% | 54.1% | 5 | 3 | 3 | 2 | 2 | 2 | 2.83 | 26,217 | 0.477 | 120 | 134 | 113 | 45 | 81 | 5 | 184 | 1 |
| 90 | 90 | 90 | 1 | 1.1716 | 1.1548 | 1.1569 | 1.1757 | 84.8% | 55.0% | 54.1% | 43.5% | 42.5% | 38.7% | 53.1% | 5 | 3 | 3 | 3 | 2 | 2 | 3.00 | 26,393 | 0.477 | 116 | 113 | 94 | 120 | 160 | 33 | 186 | 1 |
| 92 | 92 | 92 | 1 | 1.1728 | 1.1908 | 1.2044 | 1.2299 | 84.8% | 62.5% | 61.4% | 46.8% | 45.1% | 38.6% | 56.6% | 5 | 4 | 4 | 2 | 2 | 1 | 3.00 | 25,260 | 0.586 | 195 | 195 | 196 | 150 | 3 | 33 | 169 | 89 |
| 93 | 93_4 | | | 1.1741 | 1.1898 | 1.1991 | 1.2113 | 84.8% | 58.7% | 53.3% | 46.1% | 45.1% | 45.1% | 55.5% | 5 | 4 | 3 | 3 | 2 | 2 | 3.17 | 25,760 | 0.584 | 189 | 188 | 178 | 1 | 7 | 100 | 172 | 86 |
| 99 | 99_3 | | | 1.1769 | 1.1799 | 1.1845 | 1.1966 | 84.8% | 52.4% | 50.4% | 48.1% | 45.1% | 45.1% | 54.3% | 5 | 4 | 4 | 3 | 2 | 2 | 3.33 | 28,493 | 0.607 | 177 | 175 | 172 | 1 | 60 | 173 | 199 | 138 |
| 103 | 103_1 | 103 | 2 | 1.1776 | 1.1818 | 1.1930 | 1.2117 | 84.8% | 58.2% | 53.3% | 43.1% | 42.5% | 38.1% | 53.4% | 5 | 4 | 3 | 3 | 2 | 2 | 3.17 | 26,583 | 0.577 | 182 | 182 | 184 | 176 | 146 | 100 | 188 | 80 |
| 103 | 103_2 | | | 1.1776 | 1.1801 | 1.1848 | 1.2117 | 84.8% | 58.2% | 53.3% | 43.2% | 41.7% | 38.1% | 53.2% | 5 | 4 | 3 | 3 | 2 | 2 | 3.17 | 22,400 | 0.577 | 181 | 181 | 184 | 176 | 154 | 100 | 118 | 80 |
| 現 | cur13 | 現 | 1 | 1.8135 | 1.9109 | 1.9952 | 2.0731 | 94.3% | 59.5% | 53.1% | 52.8% | 46.6% | 43.1% | 58.2% | 3 | 3 | 3 | 3 | 2 | 1 | 2.50 | 88,077 | 0.000 | 204 | 204 | 204 | 24 | 1 | 1 | 204 | |

# 制約を伴うDAG最短路問題の高速解法

竹内文登[1] 西野正彬[2] 安田宜仁[1] 秋葉拓哉[3] 湊真一[1] 永田昌明[2]（北大[1] NTT[2] NII[3]）

制約を伴うDAG最短路問題とは？

- DAGの頂点 s から t までの最短路を求める。
- ただし、辺の取捨につく論理制約を満たさなければならない。




$$f = (e_{12} \wedge e_{45}) \vee e_{13}$$

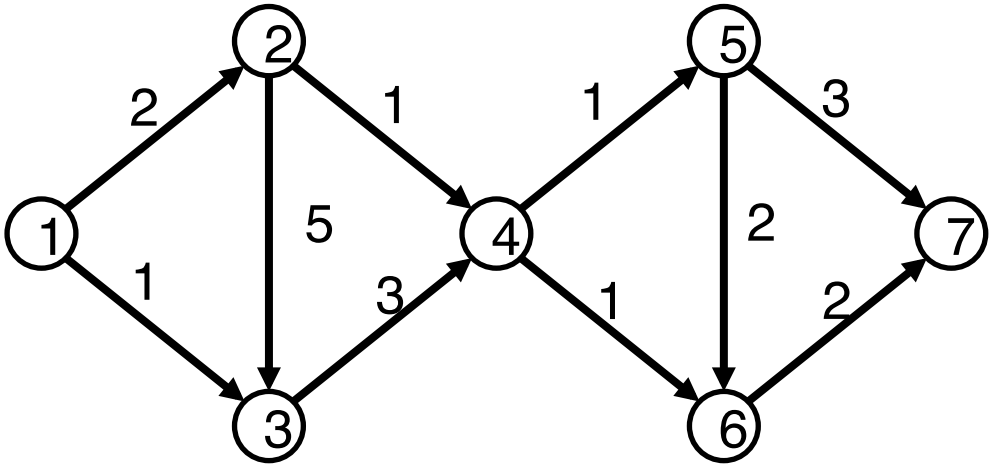
この問題が解けると？

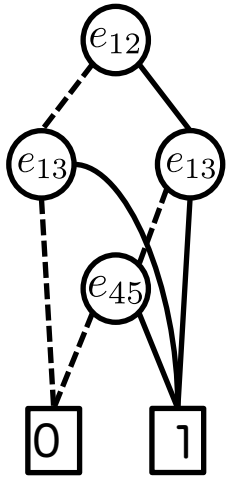
0/1ナップサック問題や文字列間の編集距離計算など、いくつかの動的計画法で解くことのできる問題に制約がついた問題を統一的に解くことができる。

# 制約を伴うDAG最短路問題の高速解法

竹内文登[1] 西野正彬[2] 安田宜仁[1] 秋葉拓哉[3] 湊真一[1] 永田昌明[2]（北大[1] NTT[2] NII[3]）

従来法にBDD-Constrained Search (BCS) [Nishino 2015]がある。




BDD-ConstrainedにA*探索を適用することで、さらに高速化。

- 制約を伴うDAG最短路問題を緩和した問題の解をヒューリスティック関数とする。（3つのヒューリスティック関数を提案）

実験結果より、提案法は多くの場合BCSより高速（最大50倍高速）

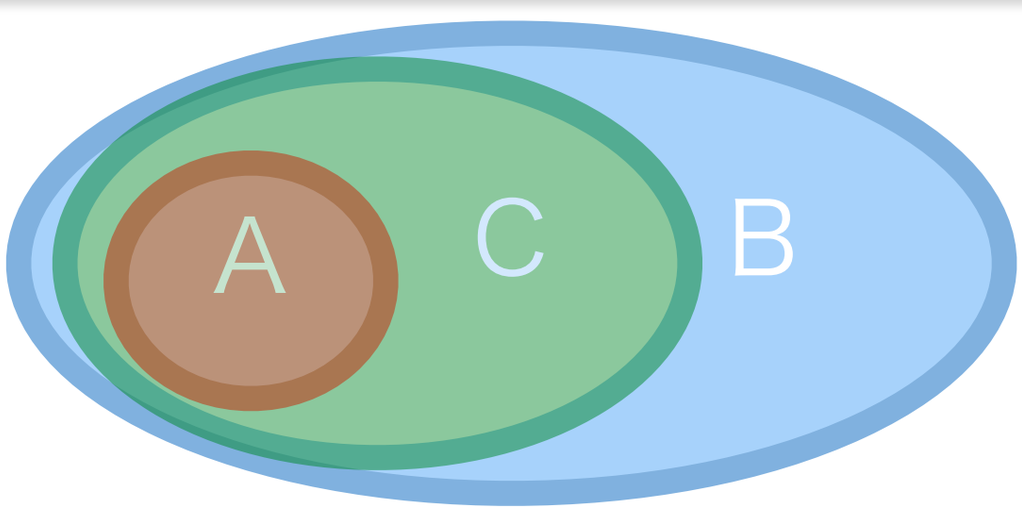
# 上下界集合制約を満たすコンパクトな決定グラフの構築

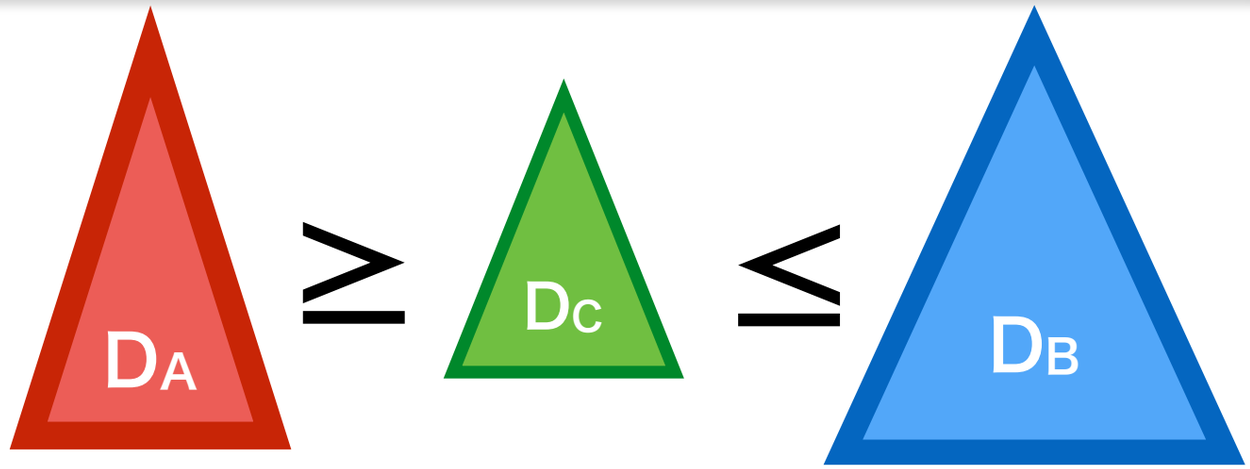
北海道大学大学院　情報科学研究科　　井上祐馬

## 問題

・$A \subseteq B$であるような集合A, Bを表す決定グラフ$D_A$, $D_B$が与えられる

・$A \subseteq C \subseteq B$を満たすような集合Cに対応する決定グラフ$D_C$のうち、節点数最小のものを求めよ




## 応用

・条件を満たす集合なら何を持っていてもよい場合、よりコンパクトになり演算も高速になる

- 例えば、群のケイリーグラフを探索する問題への応用で使える

# 上下界集合制約を満たすコンパクトな決定グラフの構築

北海道大学大学院　情報科学研究科　　井上祐馬

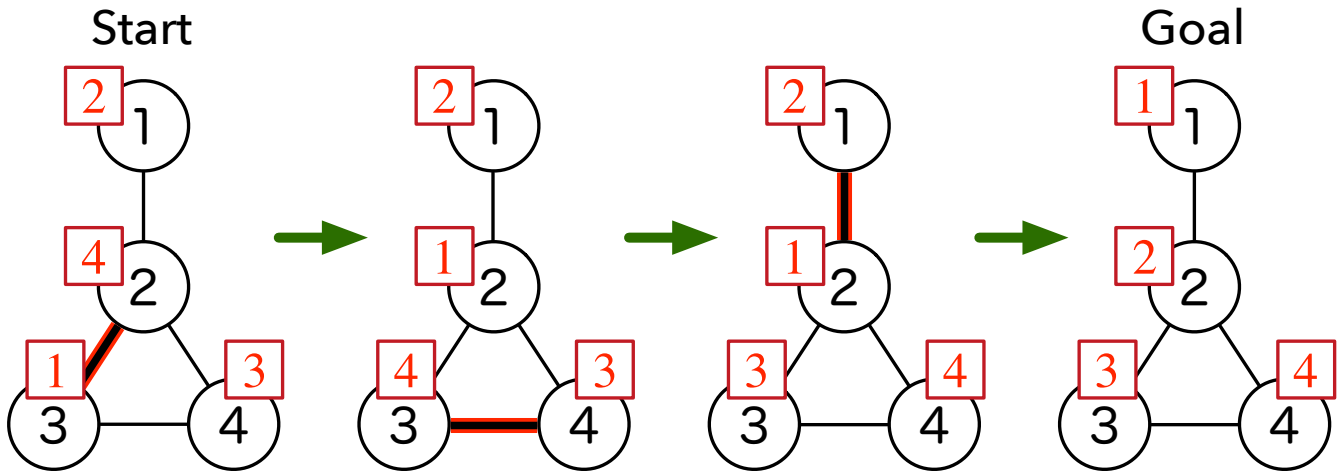
## 今日の内容

- ○ 現状、特によい結果が得られているわけではない
- ○ 予備実験の結果と考察
- ○ 今後についての議論
  - ○ 問題はどのくらい難しいか？ (計算複雑さ)
  - ○ 計算の軽いよいヒューリスティクスは考えられるか？
  - ○ 実行時間の更なる解析
    - ○ サイズを小さくできても演算が遅くなることがあった。なぜか？
    - ○ 大幅にサイズ削減できた場合でも遅くなることはある？

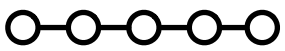
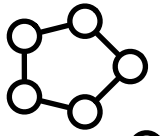
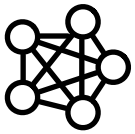
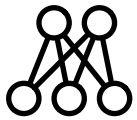
# On the Time Complexity of the Token Swapping Problem

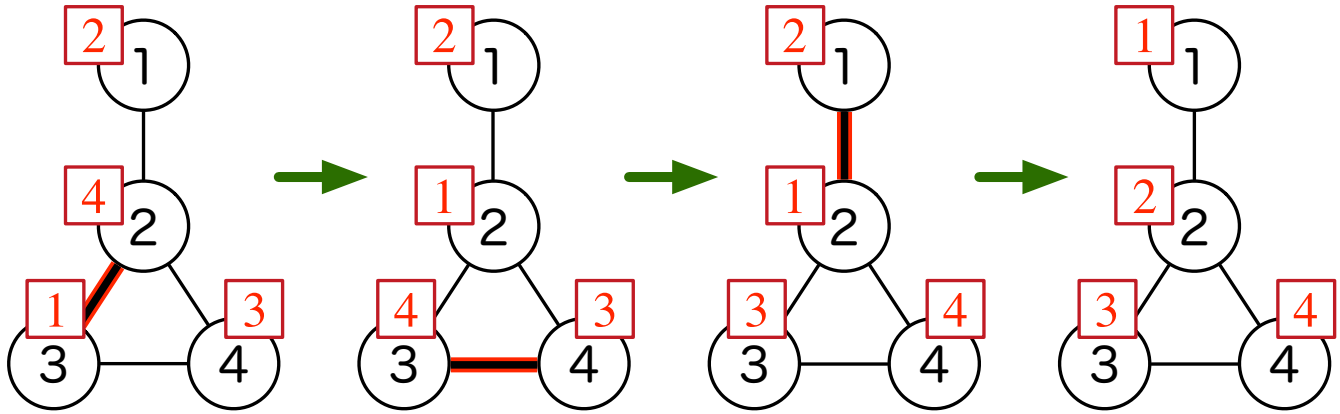
*J. Kawahara (NAIST), T. Saitoh (Kobe U), R. Yoshinaka (Kyoto U)*

- Token Swapping Problem (Yamanaka et al. 2014)
  - A token is on each node.
  - Tokens on adjacent nodes can be swapped.
  - How many swaps are needed to achieve the goal configuration?



# Known results and our results

| Classes | Complexity | |
|---|---|---|
| General | NP | Yamanaka et al. 2014 |
| Path  | P | Jerrum 1985 |
| Cycle  | P | Jerrum 1985 |
| Complete  | P | Cayley 1849 |
| Complete Bipartite  | P | Yamanaka et al. 2014 |
| Bipartite, degree at most 3 | NP-hard | Our results |
| a small class | P | |



# 頻出パターンマイニングの大規模並列化によるLAMPの高速化

美添 一樹 (東京大学)

購買データ
item: 商品
trans.: 顧客
x: 買ったもの

| transaction \ item | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| A | x | x | x | x | x | x |
| B | | x | x | | x | |
| C | | x | | | x | |
| D | x | x | | x | x | x |
| E | | x | | x | | |
| F | x | | | x | | x |
| G | | | x | x | | x |

Genome Wide Association Studies

どういう変異(SNP)を持つ人が positive か　pos. / neg.

…GTCTAAAACATGATT…　0

…GTCTGAATCATGATT…　1

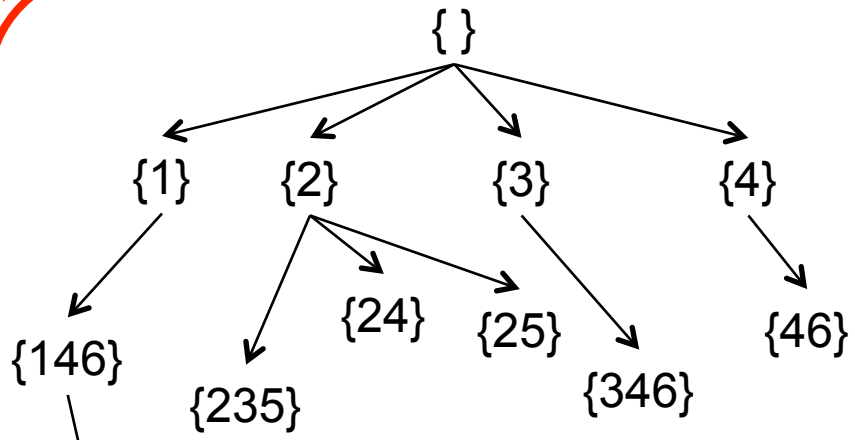
…GTCTGAAACATGATT…　0

…GTCTGAATCATCATT…　1

例えば、どういう変異の組合せがアルツハイマーの原因になっていそうか？

[Terada, Okada-Hatakeyama, Tsuda, Sese, 2014]

# 頻出パターンマイニングの大規模並列化によるLAMPの高速化




## LCM

Linear time Closed pattern Miner

[Uno, Asai, Uchida, Arimura 2004]

[Uno, Kiyomi, Arimura 2004]

深さ優先探索で非常に高速
でも遺伝データは手強い

[Yoshizoe, Terada, Tsuda 2015]

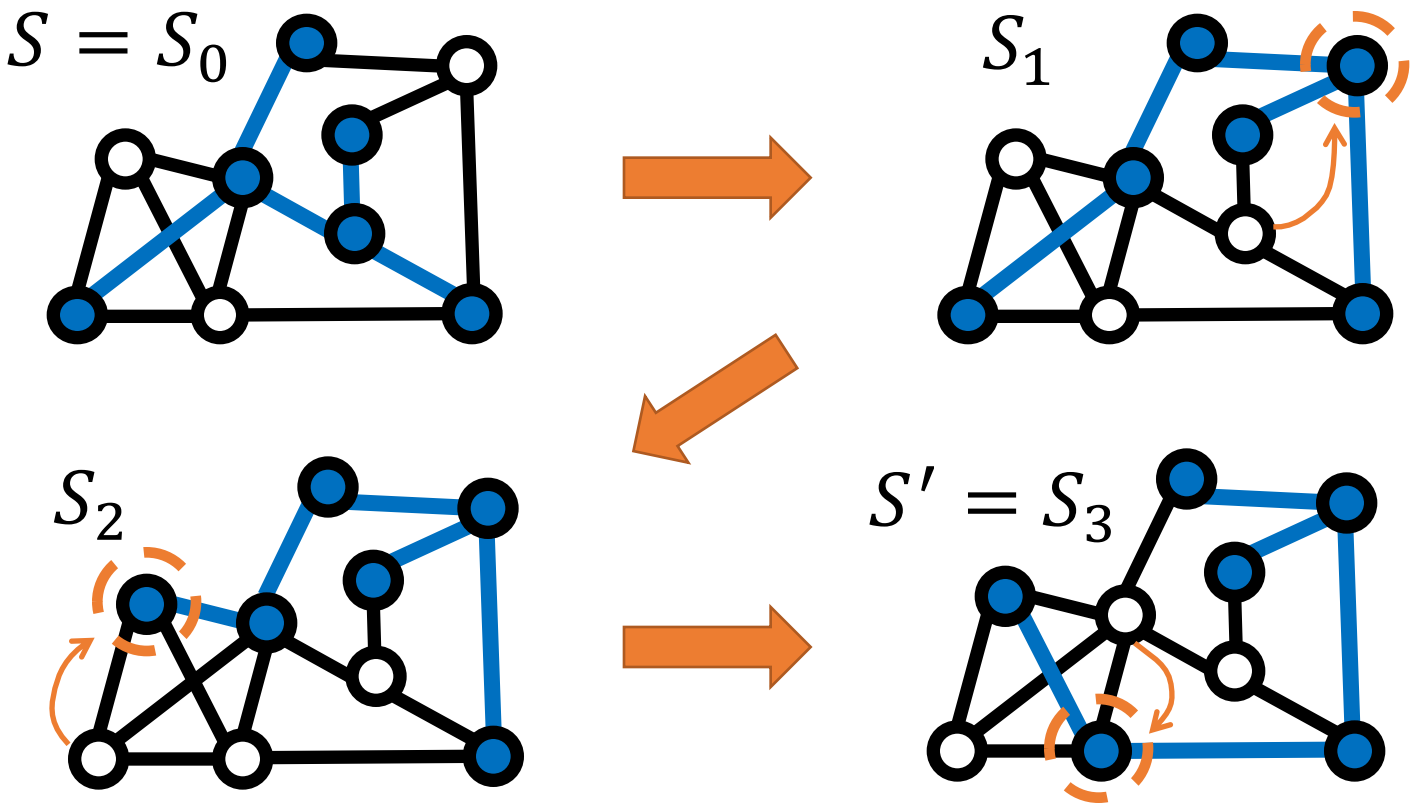
並列版LCMを自作。
東工大TSUBAMEの
最大**1,200コア** を用い、
最高で、LCMを用いた場合の
**1,030倍**の速度を達成

TSUBAME@東工大

# 誘導木遷移問題

○和佐 州洋[1], 山中 克久[2], 有村 博紀[1]
1: 北海道大学 2: 岩手大学

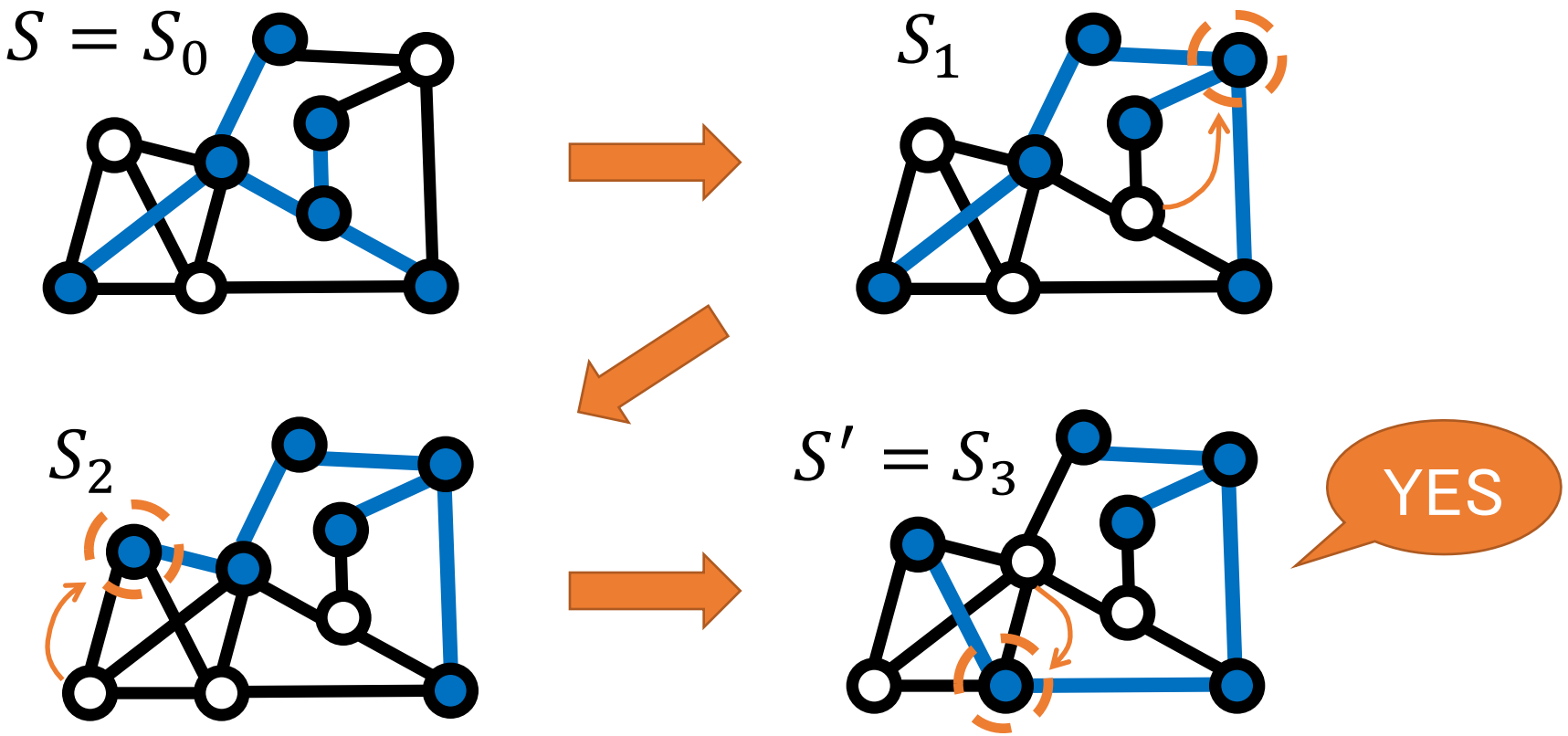
入力: グラフ$G$と, 2つの誘導木$S, S'$.
問い: $S$から$S'$への遷移列 $(S = S_0\ S_1, \ldots, S_\ell = S')$は存在するか.



# 誘導木遷移問題

○和佐 州洋[1], 山中 克久[2], 有村 博紀[1]
1: 北海道大学 2: 岩手大学

入力: グラフ$G$と, 2つの誘導木$S, S'$.
問い: $S$から$S'$への遷移列 $(S = S_0\ S_1, \ldots, S_\ell = S')$は存在するか.



# 主結果

定理 1. 誘導木遷移問題はPSPACE完全

定理 2. 極大誘導木遷移問題はPSPACE完全

定理 3. 誘導木遷移問題は$k + \ell$をパラメタ化し，TJのもとで，W[1]-hard.

定理 4. 誘導木遷移問題は$k + \Delta$をパラメタ化したもとで，FPT.

# 仮想ネットワーク埋込高速化のための ネットワーク変換

NTT未来ねっと研究所

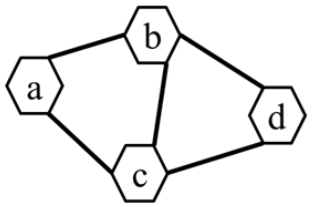
**間野暢**、**井上武**、水谷后宏、明石修

入力

**仮想ネットワーク**

$$N = (G_V, r)$$

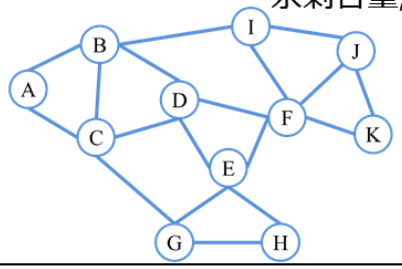
要求容量




**物理ネットワーク**

$$P = (G_P, a, c)$$

余剰容量, コスト




出力

**頂点埋込** $M_N : V_V \to V_P$

**枝埋込** $M_L : V_E \to \mathcal{P}_P$

**min. 埋込コスト**

$$c(M_N, M_L) = \sum_{v \in V_V} r(v)c(M_N(v)) + \sum_{e \in E_V} \sum_{f \in M_L(e)} r(e)c(f)$$

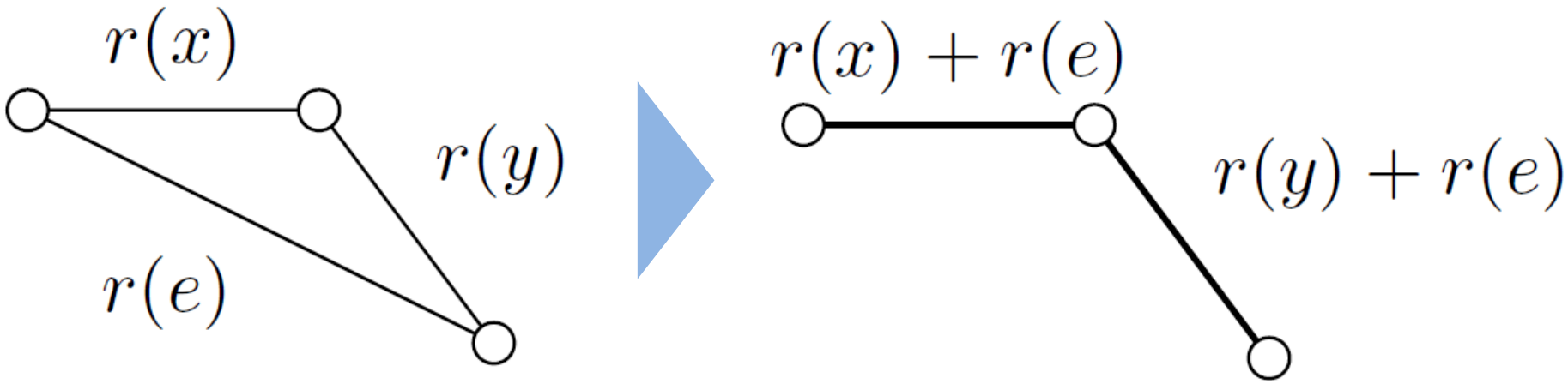
本研究

高速化のために

ネットワークを変換する

Copyright©2015 NTT corp. All Rights Reserved.

# 変換例1. (Single) Link Reduction




1. 枝 $e$ の重みを“隣” の枝 $x$, $y$ に移す
2. 枝 $e$ を削除する

操作を繰り返し、全体重みの変化を考える

Copyright©2015 NTT corp. All Rights Reserved.