

**NII Interview**

ビッグデータ社会の課題に  
アルゴリズムと数理で挑む

**That's Collaboration**

「数理」と「アルゴリズム」と「熱血」が  
新しい扉を開く

**NII Special 1**

越境により進化する新アルゴリズムの威力

**NII Special 2**

脳から新しいアルゴリズムを抽出する

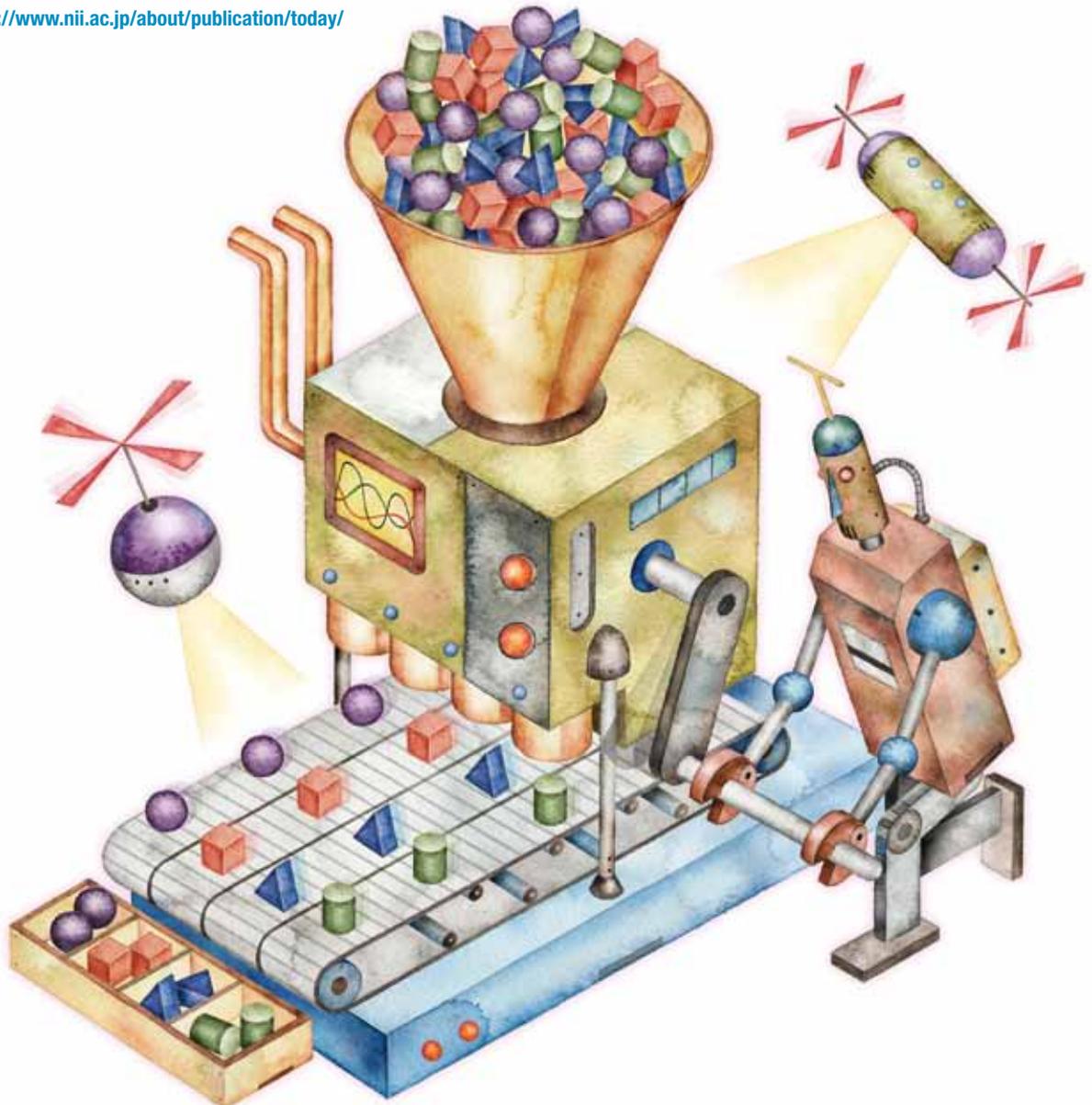
**Feature**

## アルゴリズムと 数理研究の融合



「NII Today」がデジタルブックになりました。

<http://www.nii.ac.jp/about/publication/today/>





NII Interview

# ビッグデータ社会の課題に アルゴリズムと数理で挑む

インターネットでの検索や購入履歴、スマートフォンの位置情報など、私たちの日々の暮らしは膨大なデータを生み出している。このビッグデータから社会を変革するような新しい価値を創出できるのか。アルゴリズムを研究するNIIの宇野毅明教授と、コンピュータ科学を理論面で支える離散数学が専門のNIIの河原林健一教授に、ビッグデータ時代の社会課題の解決にどのように挑むかを聞いた。

**滝田** 爆発的に増え続けるデータは、もはやコンピュータの処理能力を超えていると言われます。データをすばやく解析し、意味ある情報を引き出すためには、どのようなアプローチが必要でしょうか。

**宇野** まずはできるだけ計算回数を少なくする方法を考えます。アルゴリズムと

というのは、コンピュータに計算させる方法の設計図です。コンピュータのメモリを全部使ってしまうような膨大な計算をせずに、現実の問題に対して最適の解を導きたい。データの中で情報の乏しいノイズの部分捨てることも大事です。例えば、コールセンターの録音データから、お客様が怒っているのかどうかを知りたいのなら、「もしもし」の「も」の声だけでわかるかもしれない。後のデータは捨てていいわけです。

**河原林** どこに注目すれば不要なデータと残すべきデータを判別できるのかを、数理の面から攻めて定式化できないかと考えています。私は鉄道網や道路網のように頂点と辺の集合である「グラフ構造」を研究していますが、データを眺めると、その構造が見えてきます。どこが

大事で、どこは省略しても良いかがわかるのです。長年の経験による部分が大いのですが、そういった力をアルゴリズムの研究者と組んで社会に役立てたいですね。

**滝田** 実際に企業との共同研究もされていますね。

**河原林** 日本のメーカーのビジネスは、半導体開発から情報通信技術（ICT）に移っていますが、ICTには半導体のような基礎理論がまだありません。完成した基礎理論の上に、もの作りの技術を発展させていった半導体開発の成功体験はICTでは通じない。土台部分の理論研究を科学者がもっと担わなければならないと思います。

**宇野** 企業の方々は、肌感覚でデータの持つ意味をわかっています。でも、データ量が大きくなると、その意味を数字なり数理で言語化しなければならない。それが難しいのです。現場と信頼関係を築き、私たち研究者は数理の面で良いモデルを提供したい。

**滝田** 米国が圧倒的に強いICT分野ですが、日本に勝機はありますか。

**河原林** 2020年の東京五輪は、日本にとってゲームチェンジになると思います。



宇野毅明

UNO Takeaki

国立情報学研究所

情報学プリンシプル研究系 教授

総合大学院大学

複合科学研究科 情報学専攻 教授

## 河原林健一

KAWARABAYASHI Ken-ichi

国立情報学研究所 情報学プリンシプル研究系 教授  
総合大学院大学 複合科学研究科 情報学専攻 教授

五輪では動画や写真など、大量のデータが蓄積されることになります。これをどう活用するかです。

**宇野** 五輪招致では「おもてなし」がキーワードになりましたが、日本のおもてなしは情報が関わっていますよね。店員さんはお客様の様子や会話から、何を求めているかを判断しています。そういった細やかな心配りや現場の知見を数理で言語化して、新たなサービスを創出できるのではないかと。例えば、競技会場周辺の表示板に、気温が上がったら飲み物がある場所を案内したり、カメラの情報から子どもが多いことがわかったら、子どもが遊べる場所を紹介したりするなどが考えられます。

**河原林** 五輪に備えたICTの基盤整備も大切ですが、集まったデータを活用できる人材をどれだけ育てられるかが今後のカギを握ると思います。数学的な物の見方というのは、研究者として何年も訓練して、たくさん問題にあたって、身に付く部分があります。企業から見れば短期で成果をもたらさないかもしれない。しかし、長期的にはこういった人材がいるかないかで、基礎体力が全然違ってきます。ICTで成功した米企業には、自由な雰囲気の中でゆとりを持って研究を楽しむ環境があるように思います。

**宇野** もの作りやサービス産業では、10人全員が同じように生産性を上げていくことが求められますよね。でも、ICTの

基盤となる科学の世界は、芸術やスポーツに近い。10人中で1人が良い結果を出せばいいのです。だからといってトップ1人を雇えば、成果ができるわけではない。残りの9人がいるから頂点ができるのです。トップスターが生み出したビジネスで、残りの人が食べていく。こういった発想が日本企業にはないと思います。

**滝田** 数理、数学のセンスを持った人材育成が求められているのですね。

**河原林** 「数学は実社会の役に立つのか」という議論が長くされてきました。最近では、役に立つということに異論はないと思います。でも、「数学者が役に立つ」とは思われていない。その認識を変えていかなければなりません。もちろん、これまでの数学者の側にもいろいろと問題はありました。企業の人と話している時に、「その定義は何ですか」とか聞くのは良くないですね。社会が抱えている問題を自分自身が発見し、消化し、定式化する。そうやって、数学者が社会の問題を解決できるのだということを示していきたい。

**宇野** 今までの数理は、ある現象の証明に取り組んで解決すれば終わりでした。それでは産業界で役立ちません。その証明をする時の考え方、物の見方というのを応用していくことが大切です。数理は経営の効率化という視点で使われること

が多いのですが、これからは新しい価値の創出に使われるべきです。数理は単なる計算ではなくて、人間の創造性に深く結びついているものです。その面白さが伝われば、世の中の見方も変わって思っています。

### インタビューの一言



お勤めの洋服や本、お出かけスポットをコンピュータが示すビッグデータ社会は、便利だが空虚さも漂う。身体感覚からかけ離れた解析のスピードに、居心地の悪さを覚えることもある。現場の主観や「おもてなしの心」を、アルゴリズムや数理に採り入れ、コンピュータではなく、人が中心のビッグデータ社会を築いてほしい。

### 滝田恭子

TAKITA Kyoko

読売新聞東京本社 論説委員 兼 編集委員

上智大学外国語学部卒業後、読売新聞社入社。八王子支局などを経て2000年、カリフォルニア大学バークレー校ジャーナリズム大学院修了。2002年より科学部で科学技術政策、IT、宇宙、環境、医学、災害を担当し、東日本大震災も取材した。2014年より現職。

# 「数理」と「アルゴリズム」と「熱血」が 新しい扉を開く

That's Collaboration

膨大なデータの中から頻出するパターンとその相関関係を発見し、分類し、傾向を明らかにして未来予測にまでつなげる「発見的」なデータ分析技術、「データマイニング」。その研究の第一線で活躍するのが有村博紀教授と宇野毅明教授だ。2004年、両氏が開発した「世界最速」のデータ分析プログラムは国際的なデータマイニング実装コンテストで優勝。以来、二人三脚で数々のデータ分析アルゴリズムを開発し、ビッグデータ分析に多大な貢献を続けている。両氏に、その研究の歩みについて語っていただいた。「数理」「アルゴリズム」「応用」の密接な関係が見えてくる。

## 「知識創出学」の推進者と 数理科学研究者との出会い

——お2人が共同研究を進めるに至った経緯を教えてください。

**有村** 私はもともと科学好きの少年でしたが、実はあまり数学が得意ではありませんでした。ところがある時、連続した量の変化を扱う基本の数学とは違った数学があることを知りました。それが「離散数学」、いわば量よりも質、モノとモノとの関係を調べるための数学です。そこで、九州大学の計算機科学の権威で、後に九大総長も務められた有川節夫先生

の元を訪ねたのが、今の研究に携わるきっかけとなりました。先生のお話から、コンピュータでできることとできないことが数学的に証明できることや、微分・積分などの計算をしなくても、離散的に物事を捉えることで十分に社会の役に立つことを知り、「モノの数学」である離散数学の魅力にとりつかれてしまったのです。そのときすでに就職していましたが、有川先生の研究室に入り直し、研究者の道を選びました。以来、勉強をしているという意識もなく、離散数学とその応用を楽しんで過ごしてきた感じです。

その後、データマイニングと出会ったのは1996年頃のこと。アメリカでは大量に蓄積されたデータの「全体を見る」と、「個々の詳細を見る」ことの両方の研究が進んでいました。そこに魅力を感じ、人工知能や機械学習の研究とともに取り組んできました。北海道大学に移ってからは、異分野の科学研究者との交流を深め、膨大な量におよぶ生物の遺伝子の収集と解析や、深海調査艇から得られる深海生物の形態情報・遺伝子情報の統合、あるいはロボットや通信に関連するデータ解析など、領域をまたがる研究プロジェクトにおいて、データマイニング技術を応用して共同研究を進めています。異分野の知見を集めて新しい知識を創り出す研究を「知識創出学」と名付けて、取り組んでいるところです。

**宇野** 私のほうは小学生の頃からパズル好きで、コンピュータのプログラムづくりにも没頭していました。プログラムの



有村博紀

ARIMURA Hiroki

国立情報学研究所 客員教授  
北海道大学 大学院 情報科学研究科  
コンピュータサイエンス専攻 教授



## 宇野毅明

UNO Takeaki

国立情報学研究所  
情報学プリンシプル研究系 教授  
総合大学院大学  
複合科学研究科 情報学専攻 教授

背後にある数学、すなわち「数理最適化」にも面白さを感じていましたね。数理最適化とは、ある決まったルールとデータの枠の中で、最適な解を見つけること。究極の解ではないかもしれないが、問題解決に十分使える答えを見出すのです。

ちなみに同じ数理研究者でも、「数理解論」研究者と、「応用」研究者の2タイプがあるように思います。その違いは「モノ」に対する考え方。前者は「モノ」を前提にせず、究極の答を求める方法を考えます。後者はデータがあることを前提に、有限時間の中で処理する順番を考え、応用して役に立つ答えを求める。答は必ずしも1つとは限りませんが、限られた時間と少ない労力で最適な答えを導き出すための仕組みを考えるのです。

それが「アルゴリズム」のベースです。アルゴリズムは数理の一部ですが、時間と手続きの概念があるかどうか、数理解論の研究とは異なる。プログラムは、いわばアルゴリズムの塊であり、私はその研究にずっと携わってきました。

私と有村先生を引き合わせてくださったのは、NIIの佐藤健教授です。佐藤先生は、幅広い分野の問題意識を持つ有村先生と、問題を解決するための数理研究が得意な私、つまり「問題を見つける人」と「解答を出す人」を合わせれば凄いことができる、と。以来、有村先生が問題を提起し、私が解決法を提案するというスタイルで研究を行っています。また佐藤先生はもう1つの素敵な提案をしてくださいました。当時、イギリスで開催さ

れていた国際的なデータマイニング実装コンテストへの参加を促されたのです。

### 「数理」と「応用」のはざまで見出した最速アルゴリズム

——その国際コンテストで優勝されたのでしたね。

**有村** コンテストはあらかじめ用意された問題をデータマイニングプログラムでどれだけ高速に解けるかを競うものでした。データマイニングは、アルゴリズムの洗練度により1,000倍は速度に差が出る世界です。佐藤先生は、宇野先生が開発されていた「極大クリーク列挙」という高速アルゴリズムをデータマイニングに応用すれば、必ず勝ると言う。宇野先生は「できます！」と即答でしたね。

**宇野** そのアルゴリズムを応用すれば、データベースから頻出するパターンを高速に導き出すことができますから。つまり、膨大なデータの中から意味のある情報が簡単に引き出せるのです。

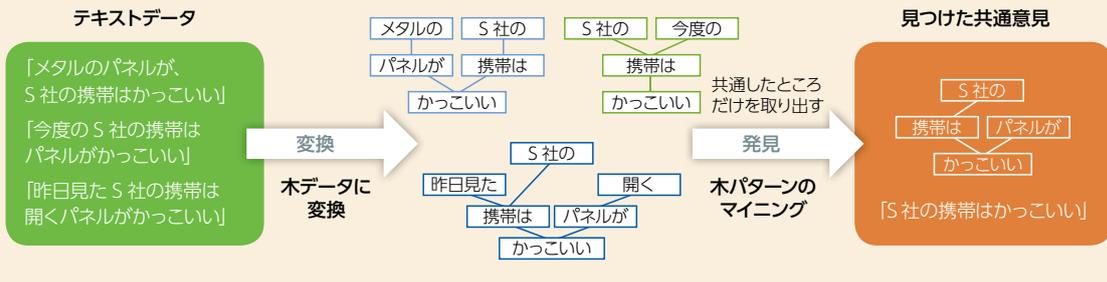
**有村** 2003年の最初の挑戦の時は、宇野先生がアルゴリズムを考え、私がモデルを考える形で研究を進めました。ところが、1位には届きませんでした。

**宇野** アルゴリズムの観点からは自信を持っていたのですが、正直にすべてのデータを対象に処理をしていたのが敗因です。他の参加者の多くはデータベース技術者であり、データの扱いに長けていて、大量のデータのうち使わないものはデータを読み込む段階で捨てていたのです。

**有村** その反省から2004年はデータをチューニングし、宇野先生がコーディングを、私ともう1人の共同研究者が実験を行う形で開発を進めました。最後には3交代制で昼夜を問わず実験とコーディングを繰り返しました。宇野先生は海外出張中でも、実験結果を伝えると8時間後には修正結果を戻してくれました。そんなハードワークを続けた結果、見事に優勝できました。

この「LCM (Linear time Closed itemset Miner)」というプログラムは、一般に公開されています。これに用いられたアルゴリズムは、アソシエーションデータマイニング領域で今でも世界最速とされています。この時の経験は、新しい扉を開くには「数理」と「アルゴリズム」にプラスして、「熱血」も必要だということを教えてくれましたね (笑)。

図1 木マイニングのテキストマイニングへの応用の例



ビッグデータから役に立つ情報を取り出す最新技術

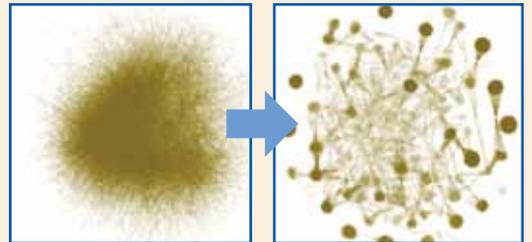
—ご研究の最近のトピックスをお聞かせください。

**有村** 現在、ビッグデータ分析が注目されています。SNSの書き込みやリンクデータ、移動端末の位置情報、建築物などに取り付けられた各種センサデータなど、多様で大量のデータが発生し、蓄積されていますが、そこから人間が理解できる「知識」を取り出すのは困難です。そこでデータマイニング技術では頻出パターンの発見や類似度を判定するクラスタリング、場合分けによってデータを分類する「決定木」などの手法を使って、人間にわかりやすい大きさ（粒度）の情報を取り出すことを目指しています。

例えばWebページのデータ、化合物

図2 企業間取引データをクラスター化して把握しやすくした例

	研磨前	研磨後
頂点数	3,282	3,282
枝数	35,168	73,132
クリーク数	32,953	343



のグラフ構造、自然言語のかかり受け木構造データなどから、特徴的な部分構造を、小さな木のパターンとして発見するアルゴリズムを開発しています。図1はオンラインデータベースの木構造から、くり返し出現するテンプレート構造を見つけた例です。

**宇野** ビッグデータから意味ある情報を得るには、データを研磨して、おおまかなクラスターに分類して把握しやすくする必要があります。その手法の1つにデー

タとデータのつながりに着目する方法があります。図2は企業間取引の大規模データを利用して、企業間の関係をグラフにしたものです。点は企業、線は取引を表します。左はデータをそのままプロットしたもの、右は最近発表したアルゴリズムによってデータを研磨した結果です。取引関係がわかりやすくなったと思います。これはSNSの「共通の友達」を探すのにも応用できます。仮に100万個のデータをすべて解析するなら約1兆回のデータ比較をしなければなりません、このアルゴリズムなら多くて1,000回程度で済む。スーパーコンピュータを使わなくても、普通のパソコンでできちゃうのがアルゴリズム研究の面白さです。

**有村** 現在の一般的なデータ分析は、仮説を立ててそれを検証する手法が主流です。しかしビッグデータを相手にするには限界がある。データの中にひそむ数理を、人間がわかる形でデータそのものに語らせた。アルゴリズム研究や機械学習の精度向上などを推し進めることで、その夢がかなうものと考えています。

(取材・文＝土肥正弘)



# アルゴリズムとは何か

アルゴリズムの語源をご存知でしょうか。9世紀前半、バグダードで活躍した数学者、アル・フワーリズムー、正しくはアブー＝アブドゥッラー・ムハンマド・イブン＝ムサー・アル＝フワーリズムーの名前から来ているといわれます\*。彼の書いた、インドにおける数の計算法の教科書のラテン語版の冒頭に、*algorithmi dicti* (アル・フワーリズムーに曰く)という一説があり、それがアルゴリズム (Algorithm) の語源となったという説が広く知られています。

アルゴリズムの中には、コンピュータが誕生する以前、それも紀元前に発明されたアルゴリズムがあります。その代表例がユークリッドの「互除法」です。これは2つの整数の最大公約数を求める手順のことで、皆さんも中学校で習ったはずです。このユークリッドの互除法は、現代社会でも使われています。通信などの秘密を守る公開鍵暗号の基礎となり、デジタル署名などでも使われているのです。その発明者のユークリッドは、ユークリッド幾何学の名前で知られている、紀元前3世紀頃の数学者ですが、彼の生きた時代から2,000年

以上先の未来に互除法が貢献するとは夢にも思わなかったでしょう。

いまのアルゴリズムの定義、すなわちコンピュータや数学などの分野で、問題を解くための手順を定めた表現となるのは、1930年代になってからです。とはいえ、プログラムとして表現されたものだけがアルゴリズムではありません。例えば小学生時代に習った手計算による多数桁の足し算や、かけ算や割り算のやり方も、その手順に従えば答えが導き出せるわけですから、アルゴリズムといえます。また、料理のレシピも、料理を作るという問題を解く手順とみれば、アルゴリズムといえるでしょう。郵便も、ハガキや封書の郵便番号や住所を読み、所定の手順に従って、郵便局間でハガキや封書を送り、最終的に宛先に届けるという意味では、その配送手順はアルゴリズムといえます。そしてインターネットの通信パケット (小包) の配送は、郵便の手順、つまり郵便のアルゴリズムを元にしていますから、現実世界のアルゴリズムが通信のアルゴリズムになったこととなります。

ところで、アルゴリズムの研究とは何なのでしょう。いままで解け

なかった問題を解く方法を見つける研究もありますし、解く方法が知られていても、手間のかかる手順の場合、その手間を減らすことも重要な研究です。実際、科学データやビッグデータのように大量データの処理アルゴリズムの場合、その処理に数時間どころか数日かかることがあります。それが少しでも速くなれば大きな意味を持つのです。

さまざまなアルゴリズムが身近に使われているように、アルゴリズムという言葉も広く使われ始めています。10年ぐらい前までは、アルゴリズムという言葉は、コンピュータサイエンスの専門用語として使われることが多かったですが、いまはテレビ番組に登場する「アルゴリズム行進」から、金融取引で使われるアルゴリズム取引まで、日々の生活や社会に登場する言葉となりました。次の10年では、どんなアルゴリズムが生まれ、そしてアルゴリズムという言葉はどんなところで使われるのでしょうか。

(文＝佐藤一郎)

\*代数学を英語で Algebra といいます。これもこの数学者の名前のラテン語訳に由来するといわれています。



テレビ番組でおなじみの「アルゴリズム行進」

# 越境により進化する 新アルゴリズムの威力

## 物理、数学の知見を理論計算機科学に活用

吉田悠一助教の研究は、理論計算機科学の成果を駆使して、高速なアルゴリズムの性質を探るというものだ。中でも、吉田助教の中心的な研究テーマである「定数時間アルゴリズム」は、現実世界の問題を解く上で衝撃的な威力を持つ。この手法を使えば、例えば通常のやり方なら数万年もかかる計算を、条件を付けることでわずか数ミリ秒で解けるようになる。成果の秘訣は、自身の専門分野の「隣」の分野の知見に注目することだった。

### 「隣」の分野の知見に 注目して成果を出す

吉田悠一助教の専門分野は、理論計算機科学である。例えば、ある問題を計算機が解く場合の「計算量」を調べ、それ

が解ける問題なのか、解けない問題なのかを調べる学問分野だ。

その吉田助教は現在、「実用的なアルゴリズム」、つまり計算機で実際に動作させた時に高速なアルゴリズムの研究にも手を広げる。というのも、吉田助教が理論計算機科学の立場から計算機科学分野のアルゴリズムの世界を見た際に「驚

くほど手がついていない状態」であることに気づいたからだ。つまり、未解決の問題の宝庫に見えたというわけだ。

例えば、データベース分野で高速なアルゴリズムの研究は進んでいるが、それが「なぜ高速なのか」を理論的に解明する研究は手つかずの部分が多かった。

「理論に基づく高速なアルゴリズムができれば、より有用性が増します。また、理論の発展は、現実のコンピュータを相手にする計算機科学の研究者にも、大きな刺激となるはず。アルゴリズムに理論を持ち込むことにより、自分自身が、従来分野の研究者に刺激を与えるような“黒船”になりたいのです」と、吉田助教は話す。

吉田助教がこのように語る背景として、今、異なる研究分野を「越境」する研究が盛んになっていることがある。

「ブレイクスルーのためには、いろいろな分野に触れることが不可欠です。違う分野でこそ、良い問題、良い道具が見つかりやすい。実際に、数学の分野である表現論や代数幾何学の研究者が、理論計算機科学の問題を解くといったこともある。逆に、黒船に乗り込まれてしまうこともあるのです（笑）」

例えば物理学もその1つ。現在、多く

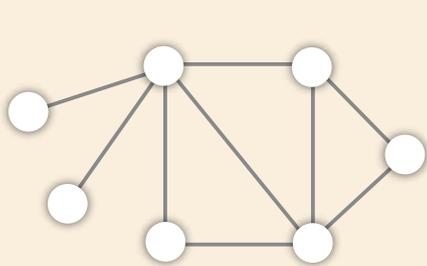


図1 7頂点と9辺からなるグラフ。グラフ理論では、このように頂点とそれらを結び辺で構成されるグラフの性質について調べている。

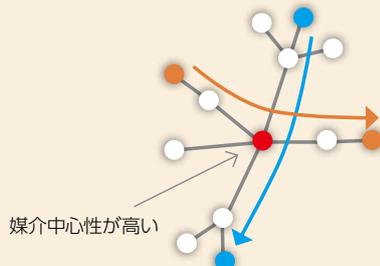


図2 媒介中心性の概念図。オレンジの頂点間の最短経路、青い頂点間の最短経路はどちらも赤い頂点を通っており、赤い頂点の媒介中心性が高い。

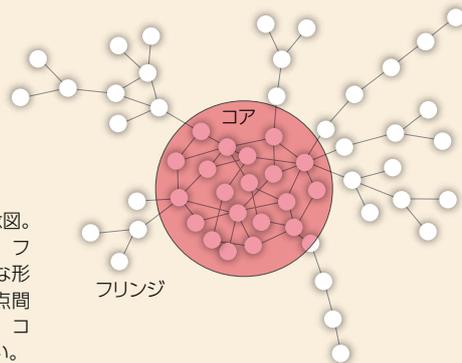


図3 コアとフリンジの概念図。コア部分は密になっており、フリンジ部分は「木」のような形をしている。フリンジの頂点間の最短経路はコアを通るため、コアの頂点の媒介中心性は高い。



## 吉田悠一

YOSHIDA Yuichi

国立情報学研究所  
情報学プリンシプル研究系 助教

の物理学者が複雑ネットワークに関心を持ち、その性質を研究している。この研究領域から出てきた概念に「グラフの頂点をどれだけ最短経路が通っているか」で頂点の重要性を評価する「媒介中心性」がある(図1)(図2)。感染症の流行のモデル化に使ったりと応用範囲が広いが、計算量が大きくなることから、高速アルゴリズムの研究対象となり得る。「他分野の知見をアルゴリズム研究に使わないのはもったいない」と吉田助教は力説する。

### 定数時間アルゴリズムの すごさ

吉田助教が研究者として取り組む中心的なテーマは、「定数時間アルゴリズム」の研究である。その基本的な考え方は、問題のサイズに関わらず、一定の時間で解くこと。つまり、問題のサイズが巨大になるほど、この定数時間アルゴリズムは恐るべき威力を発揮するようになる。

例えば、SNSのFacebookをモデル化してグラフとして表すと、ユーザー数と同じ約10億の「頂点」を持つことになる。このグラフの「直径」(グラフのすべての頂点の組み合わせの中で最大の距離)を正確に求めるには、コンピュータの処理速度を毎秒1億ステップとして約2万年もの計算時間が必要となる。現実的には解くことが不可能な問題なのである。

ところが、誤差  $\epsilon = 0.01$ 、つまり1%

の誤差を許容して、ある定数時間アルゴリズムを使って解くと、同じ問題の解が1.2ミリ秒と、ほぼ一瞬で求まる。2万年が1.2ミリ秒に短縮されるということは、現実には解けなかった問題が、解けるようになることを意味する。これが定数時間アルゴリズムの威力だ。

このアルゴリズムを現実世界に応用できれば、今までできなかった処理ができるようになる。有用性という観点で見ても、恐るべき可能性を秘めた理論だろう。

一方、吉田助教の関心は現実世界への応用よりも理論へ向いている。最近の研究では、ある問題が「どんな性質なら定数時間に解けるか、解けないか」を追究し、「その必要十分条件を与える」という成果を挙げた。定数時間アルゴリズムの本質に関わる理論を証明したのだ。その証明の過程では、「調和解析」という数学の道具を使った。ここでは逆に、理論計算機科学が数学の成果を取り込んだ格好だ。「これも隣の分野の成果を使った研究の例です」と吉田助教は話す。

### プログラミングコンテストが 取り持つ縁

現在、吉田助教は、「ERATO(戦略的創造研究推進事業・総括実施型研究)河原林巨大グラフプロジェクト」において、「複雑ネットワーク・地図グラフグループ」のグループリーダーを務めている。

このグループの構成メンバーは、「プログラミングコンテスト出身のアルゴリズムの達人揃い」なのだという。

この研究グループから生まれた新しいアルゴリズムとして、グラフ内の任意の2点を結び最短距離を既存手法の100倍も高速に解く「Pruned Landmark Labeling」がある。このアルゴリズムでは現実世界のWebを観察した知見を取り入れた。Webのリンクをモデル化した「Web Graph」は、多くの頂点が集まる「コア」と周辺領域にあたる「フリンジ」と呼ぶ構造を持つ(図3)。そしてグラフ上の2点の最短経路はこのコアを通ることが多い。この知見を元に理論を駆使し、高速化に成功したのだ。

このアルゴリズムの論文は、吉田助教と、東京大学大学院博士課程に在学中の秋葉拓哉氏、岩田陽一氏の3人で書かれた。皆、ACM-ICPC(ACM国際大学対抗プログラミングコンテスト)の出身者だ。アルゴリズムを理解し、プログラムを組む訓練を積んできた若者たちが今、理論計算機科学において、アルゴリズム研究の最前線を担っているというわけだ。

最新理論に基づくアルゴリズムの研究は「計算機ができること」の範囲を大きく広げ、現実の世界に衝撃的な変化をもたらす可能性がある。ITが社会の隅々まで浸透する中、理論計算機科学の成果への期待はますます高まっている。

(取材・文=星 暁雄)

# 脳から新しいアルゴリズムを 抽出する

## 脳科学と情報学の融合の成果を、 医療や新たな情報処理へ応用

脳科学と情報学との結びつきにより、新たなアルゴリズムの抽出に挑んでいるのが、NIIの小林亮太助教だ。多くの謎が潜んでいる脳機能の解明に情報学の知見を活用。脳の動きを知ることで、難病の克服といった医療分野への応用が期待されるほか、脳から抽出した新たなアルゴリズムによって、近未来には、コンピューティングパワーに頼らない高速な情報処理を実現することも可能になるという。

### 脳活動から情報が 取り出せる時代へ

脳科学と情報学との結びつきは、米国ではかなり前から進んでいるという。

「脳科学の課題は、脳のなかで実際に何が起きているのかを見極めるのが難しい点です。そうした中、ここ10年ほどの間に、脳科学者が持つ実験データを、コンピュータサイエンティストや数理工

学分野の研究者が解析することで、従来の理論を証明したり、新たな理論を導き出すことが頻繁に行われています」と、NIIの小林亮太助教は切り出す。

例えば、被験者に文字を見せた場合のf-MRI（磁気共鳴機能画像）データを大量に集めると、f-MRIデータを見るだけで、その人が何の文字を見ているのかがわかるという。脳科学におけるf-MRIの計測技術と、それを収集し、分析するデータマイニング技術などの情報学の知見が融合することで明らかになっ

た事例だ。

一方で昨今、「ディープラーニング」と呼ばれる取り組みも、脳科学と情報学の結びつきによって急速に進展している領域だ。これは、かつて人工知能研究の一領域としてニューラルネットワークと呼ばれていたもので、脳の動きを参考にしてアルゴリズムを改良することで性能が急速に発展しつつある。機械学習や音声認識などへの応用もその一例といえる。

だが、「脳科学と情報学の結びつきによる技術進化はみられますが、日本ではこの領域での研究者が少ない」と小林助教。欧米においては、政府主導の大規模な研究プロジェクトが開始されているが、残念ながら日本では現状、こうした取り組みを本格化させる動きはないという。

### シミュレーションの成果を、 脳疾患の治療に役立てる

小林助教が取り組んでいるのは、脳を構成する最小単位である神経細胞と、それらを結びつけるシナプスの動きを研究することで、脳活動をシミュレーションし、これを難病克服などの医療分野への



### 小林亮太

KOBAYASHI Ryota

国立情報学研究所  
情報学プリンシプル研究系 助教  
総合研究大学院大学  
複合科学研究科 情報学専攻 助教

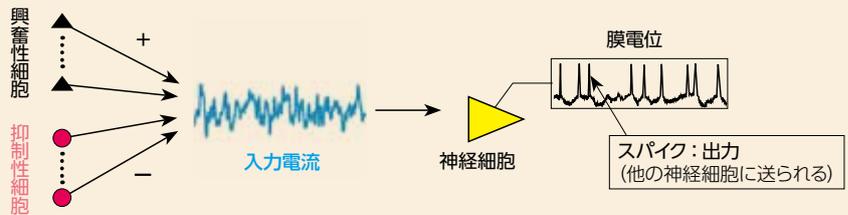


図1 神経細胞の情報処理の仕組み。神経細胞は数千もの興奮性細胞、抑制性細胞からの入力電流を受け取り、スパイクを生成し、他の神経細胞にスパイクの情報を送ることによって情報処理を行う。

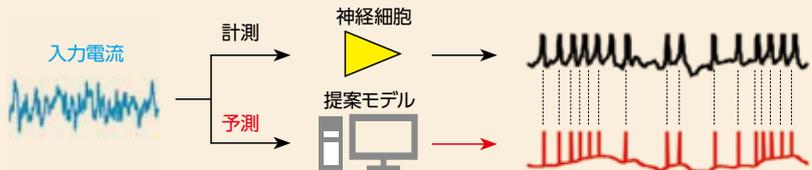


図2 神経細胞のスパイク予測。上段は、変動電流を神経細胞に入力した時の膜電位の計測データ(黒)、下段は、提案モデルによる予測結果(赤)である。誤差 0.002 秒以内に予測できたスパイクは点線で示した。このモデルは、約 90%のスパイクを予測可能であり、スイス連邦工科大学ローザンヌ校 (EPFL) で開催されたスパイク予測の国際コンテストで優勝を果たした。本研究は篠本 滋准教授(京都大学)、坪 泰宏准教授(立命館大学)との共同研究である。モデルの詳細については Kobayashi, Tsubo, Shinomoto, Front. Comput. Neurosci., 2009、コンテストの詳細については Gerstner & Naud Science 2009 を参照。

応用や、新たなコンピューティングモデルの創出などに生かすことである。

そうした中、小林助教は、スイス連邦工科大学ローザンヌ校 (EPFL) が、2007~2009年にかけて開催した神経細胞のスパイク予測コンテストにおいて優勝。同コンテストでは、神経細胞から出力される脳内通信信号であるスパイクの動きを予測。その精度を競うものだが、小林助教のシミュレーションは90%という精度でこれを的確に予測した。ノーベル賞を受賞したHodgkin-Huxleyモデルの60%を大幅に上回る実績により、全世界から、神経細胞モデルのシミュレーションモデル基盤を構築したと高い評価を得た。

さらに、神経細胞がどのように活動して情報処理を行っているのかを推定するデータ解析手法を開発。神経細胞同士の情報のやり取りに使われるシナプスの研究にも踏み出している。

「シナプス同士がどんな強さで結びつくのか、そのときに何が起きているのか。これを調べることで、脳活動がシミュレーションできるようになり、さまざまな領域への応用が可能になります」と小林助教は研究の狙いを語る。

一例としてあげられるのが、パーキンソン病の治療への応用だ。

パーキンソン病の患者の脳の異常な動きをしている部分に電極を埋め込んで刺激することにより、症状を改善できるという。これは脳深部刺激療法 (DBS) という、注目の治療法の1つだ。ここでは、DBSを行っている際の脳活動をシミュ

レーションすることで治療法の効率化を図るといった研究も開始されている。

さらに脳波を計測したデータを解析することで、従来は難しかった認知症やてんかんの診断にも活用できるという。とくに認知症診断では正答率は83%と高く、早期に診断・投薬ができれば、通常的生活を送ることができる実質寿命の長期化にも大いに貢献できるだろう。

「脳疾患のほか、躁鬱病や自閉症などの精神疾患においても、脳科学と情報学の融合手法を用いた研究が行われ始めています」(小林助教)。

### 新たなコンピューティングモデルを導く

脳活動の研究は、新たなコンピューティングモデルの創出にもつながる。

小林助教は、「人間の脳の中に神経細胞は2000億個、シナプスは1000兆個もあり、脳内の1秒間の処理をシミュレーションするにはスーパーコンピュータ「京」を利用しても40分かかる。人間の脳はスーパーコンピュータよりも優れて

いるのです」と前置きをした上で、「学習の際、神経細胞同士をつなぐシナプスが変化することがわかっていますが、こうした複雑な仕組みがわかれば、そのアルゴリズムを応用した新たなコンピューティングモデルを創出できる」という。

ビッグデータ時代を迎え、そこから数々の知見を導き出すには、膨大なコンピューティングパワーが必要である。だが、脳活動をもとにしたアルゴリズムが開発されれば、これまでのコンピュータにはできない特徴を抽出できるほか、新たな学習方法により、さらに高度な情報処理技術が開発できるだろう。また、脳の効率的な活動に倣えば、究極の省エネコンピュータを開発することも夢ではない。いわば、コンピューティングパワーに頼らない高性能コンピュータの創出が実現されるというわけだ。

今後いっそう脳科学と情報学が結びつくことにより、医療をはじめとするさまざまな領域へのアルゴリズムの活用が進み、また、新たなコンピューティングモデルが誕生することで、我々の生活がより豊かなものになるのは間違いなさそうだ。

(取材・文=大河原克行)

# リチャードソンの夢

速水 謙

HAYAMI Ken

国立情報学研究所 情報学プリンシプル系 教授

季節はずれの大きな台風が二つ、日本列島を駆け巡って行った。その威力を感じるとともに、昔に比べて台風の進路の予報が正確になったように感じた。

## 数値気象予測事始め

その昔、イギリスの数学者で気象学者のリチャードソンはコンピュータのない時代に、流体の方程式をもとに、天気を予測する計算方法を着想し、それを実現しようとした。空間を升目のように区切り、気圧の高い升目から低い升目に空気が流れるように計算するのである。しかし、計算尺などに頼った手計算では到底予報に間に合うようには計算ができなかった。6時間後の予報を行うのに6週間かかったのである。しかもその予測は現実とは程遠かった。

そこで、彼は劇場のような大きな建物の座席に60,000人の“コンピュータ”（計算をする人）を配置し、各升目の計算を担当させ、“指揮者”の統率のもとに全体の計算を並行して行わせることにより予報に間に合うように計算を高速化することを夢想した。これを“リチャードソンの夢”という。

## 数理モデルとアルゴリズムの役割

この夢は、電子計算機の出現とともに具体化し、

今日、スーパーコンピュータや並列計算機により現実のものとなり、かなり正確な天気予報や、中期的な気象予測が可能になっている。ただ、リチャードソンの夢が彼の時代に実現しなかった理由はもう一つある。数理モデルである方程式が現実の気象現象を的確に表現していなかったのである。その後、計算機の発達とともに、モデルの改良が積み重ねられた。また、予測のもとになる観測データが不十分な場合にいかにしてそれを補うかといった工夫や、並列計算機の実力を十分に発揮させるアルゴリズム（計算法）の開発など、数理モデル、アルゴリズム、計算機の発達の相乗効果によって今日のような天気予報や気象予測が可能になってきているのである。

この例が示すように、計算機を使った予測は、数理モデル、アルゴリズム、計算機を総動員して、初めて可能になるのである。

因みに、リチャードソンの興味は気象にとどまらず、クエーカー教徒で平和主義者の立場から数理モデルによる戦争の解析を行ったり、国境線の長さの解析を行う中からフラクタルという数学の概念の端緒をつかむというように、数学を様々な実社会の問題に応用し、またその中から新しい数学やアルゴリズムを生み出したことは興味深い。

情報から知を紡ぎだす。

NII

国立情報学研究所ニュース [NII Today] 第66号 平成26年12月

発行：大学共同利用機関法人 情報・システム研究機構 国立情報学研究所  
〒101-8430 東京都千代田区一ツ橋2丁目1番2号 学術総合センター

発行人：喜連川 優 監修：佐藤一郎

表紙画：城谷俊也 写真撮影：佐藤祐介

デスク：田井中麻都佳 制作：ノーバジェット株式会社

本誌についてのお問い合わせ：総務部企画課 広報チーム

TEL：03-4212-2164 FAX：03-4212-2150 e-mail：kouhou@nii.ac.jp

[NII Today] の  
デジタルブックもよろしく!



情報犬ビットくん  
(NIIキャラクター)

<http://www.nii.ac.jp/about/publication/today/>