3．研究開発領域

3.1　基礎理論

・情報科学技術と社会における基礎理論の役割

計算の基礎理論は情報科学技術の大きな方向性を左右する基盤であり、情報システムや情報社会を構築・評価し、進化あるいは革新させるものである。

古くはチューリングやゲーデルによる計算可能性・不可能性の理論と、それを示すためのモデルであるチューリング機械の提唱が、計算機が単なる学術的な数値計算のみでなく社会のありとあらゆる課題の解決に利用できるという万能性の担保になり、計算機開発の意義と価値を示した。実際にチューリングはその担保を具現化してドイツ軍のエニグマ暗号を破る専用計算機である Bombe 機械を作成し、電子的な計算機開発への道を拓いた。また、情報を計算機が扱う数値や論理式に変換する符号化理論は、シャノンにより情報通信の数学理論として発祥し、その後「情報理論」として発展し、画像や音声などを含む様々な情報を高速に通信・処理するのみならず、Web 検索に代表される高速データ検索などが行える「データ構造理論」へと派生発展し、計算機構やインフラの基盤となった。

その後、実社会の課題解決を数理的にモデル化する「最適化理論」が、フォンノイマンらが提唱したオペレーションズ・リサーチから発祥し、更に数理化された課題を解決する実行手順(アルゴリズム)の設計理論と品質保証理論である「アルゴリズム理論」と一体となって発達し、実社会の多くの課題が計算機によって解決される情報化社会を作りだしている。

現実の問題に対して数理的な定式化を見出し、優れたアルゴリズムや通信プロトコル（通信に関する手順）を開発することは、情報技術の核であり、新規産業を生み出す爆発的な力を持つ。例えば遺伝子検証における Smith-Waterman のアルゴリズムはゲノム産業の出現の大きな要因となり、データの関連性を見つけるアプリオリアルゴリズムはデータマイニングの概念を生んでデータ解析を産業化し、安定結婚マッチングアルゴリズムは物資や人員の配置メカニズムを変革し（ノーベル経済学賞受賞)、ページランクアルゴリズムと呼ばれる Web 検索を補助するアルゴリズムは Google 社を立ち上げて情報検索産業を革新した。また、公開鍵暗号プロトコルは「暗号理論」を変貌させて電子商取引を可能にし、誤り訂正符号やデータ圧縮手法は画像やビデオ等の高品質デジタル配信を可能にしている。いずれの例も産業的・社会的に計り知れない影響を及ぼしている。

また、一つのバグが大きな社会的な影響を与える現代において、プログラムやソフトウエアの正当性・安全性の保証は非常に重要であり、アルゴリズムやプロトコルを安全かつ効果的に実装し、計算機やネットワークで実行するために、プログラムの開発やソフトウエア自動検証に関する「プログラム基礎理論」が大きな役割を担っている。

・計算の本質の追究による IT の先導

情報や計算の構造は、順列やグラフ、論理関数等の「離散構造と組合せ論」を用いて数学的に抽象化され、数理的に定式化されて解析される。理論的に効率よく計算できる（専門的には多項式時間計算可能な）問題の体系が整理され、解法やアルゴリズムが与えられ

て実社会の課題解決の基盤になっている。一方で計算機や IT の応用の変遷に伴って解くべき課題は急速に拡大・進化し、理論的な解法の体系の進歩と整備は常に急務である。

チューリングらに発祥した計算の本質の追究は「計算複雑度理論」として発展し、計算理論全体の基本指針を与えている。現実に求解が切望される問題の中にも現実的な計算時間では計算できそうもない問題群があり、これらを類別し計算の困難性を解明するために、様々な計算モデルによる計算可能性が議論されている。特に非決定性計算というモデルで効率よく解ける問題群は、NP 問題と呼ばれる。これは答えを推理や経験で発見して、それを検証するという、人間が行う典型的な求解のプロセスでの問題解決をモデル化したものであり、計算機を用いて人間の智にどこまで迫れるかという疑問が、『NP 問題を通常の計算モデルでも効率よく解けるか？』という問題（P vs NP 問題）として定式化され、計算理論で最も有名な未解決問題となっている。NP 問題は NP 完全問題と呼ばれる問題群のどれかを解けばすべて解けるという理論（NP 完全性の理論）がクックやカープらによって 1970 年代に確立され、大きな影響を与えた。

実際に、実社会で必要とされる問題の多くは、NP 問題を最適化問題として定式化した NP 最適化問題と呼ばれるものである。最適化問題においては、解の品質という概念がある。例としてカーナビゲーションで要求される、地図上での２点間の最短時間経路を求めよという問題は、典型的な最適化問題であり、経路の所要時間が解の品質になる。NP 完全性理論は、典型的な NP 最適化問題の解法を開発して汎用的なソルバーを作ることの可能性と重要性を示唆しており、その実用的・産業的価値は非常に高い。特に整数計画問題は、それを緩和(整数条件を除去して解く)してできる線形計画問題が効率よく解けるということが判明しているため、多くの場合に品質の良い解を得やすく、現在では産業界において多くの現実問題の解決にソルバーとして利用されている。また、NP の求解モデルをそのまま利用し、(多くは自然や人間の挙動を模倣して)解の候補を発見して、それを検証してよい解を探す発見的手法(ヒューリスティクス)が提案され、局所探索法、焼きなまし法、遺伝アルゴリズムなど数多くの種類が実用化されている。

さらに、人間の行うさらに進んだ求解手法である「多人数で手分けして問題を解く」「教師や専門家に相談する」「サンプル調査した結果を検証して判断する」「熟練を積む」「複数での合議を行う」「過去のデータから予測する」などのモデルに対して、並列・分散モデル、対話証明モデル、確率的検証証明モデル、学習モデル、強化学習モデル、オンラインモデル、統計的学習モデルなど様々な計算モデルが考案され、それぞれの計算力の強力さを理論的に解析・評価し、その評価に基づいて現実にこれらのモデルを活用した計算アルゴリズムや計算機構の構築が行われてきた。これらの計算理論によるモデル構築と理論解析は、IT や情報社会の水先案内人の役割を果たしており、現在の Web ビジネスやクラウド、ビッグデータ活用等の動機を与えている。産業界でも、米国などの情報関連企業では計算理論の研究者を集め、技術開拓の水先案内を担わせている。

・ビッグデータ時代に向けた計算理論の進化

科学において計算の大きな役割として、データの収集・整理・加工と可視化、理論的モデルに従った最適化、人間の作った仮説に対するシミュレーションによる検証が過去における代表的なものである。これらは課題を人間と機械の協業により解決するものであり、

観察・理論・実験の科学の古典的なパラダイムに従っている。一方、グレイにより提唱され近年注目されている新しい科学パラダイムであるデータ駆動型科学では、新しい計算の役割が必要になる。即ち自然や人間生活からセンサーや Web 情報発信などによって収集されるビッグデータと呼ばれる大量かつ非均質なデータを活用して、モデルを自動構築する、あるいはモデルを明示的に提示せずに知識や性質を取り出すビッグデータ技術の重要性が増している。

必然的に計算の基礎理論も対応した進化を要請され、データの統計的表現や隠れたモデルを用いたデータの構造化技術、データからの情報抽出技術であるデータマイニングや機械学習のためのモデル構築やアルゴリズムの設計・解析技術、予測やそれを用いた意思決定支援などを含む、いわゆる「データアナリシス」の計算理論基盤の開拓が重要なテーマになっている。ここでは、データの巨大化に従い、リアルタイム処理のための圧縮データ構造や次元削減手法、限られた記憶領域を用いてストリームデータを処理するストリームアルゴリズム、データの全体を読むよりも早く性質をとらえる劣線形時間アルゴリズム、高次元幾何学データにおける近傍・類似検索構造、クラウド上のデータを共有するためのセキュリティー・プライバシー保護のための理論など、アルゴリズム、データ構造、暗号理論、情報理論などすべての分野を巻き込んだ統合的革新が求められている。これは情報社会の変革につながる革新と考えられており、特に米国では重点的に推進されている。

・我が国の現状

歴史的には、第二次世界大戦および戦後の冷戦時に政策的に行われた、計算機の開発・活用及び最適化に関した数学的研究開発が計算基礎理論の探求・深化の契機であり、研究の最大拠点は米国である。日本においては過去に実施された特定領域研究等の効果で計算理論の研究レベルは全般的に国際的に高く評価され、特にアジアにおいてはトップリーダーの位置にある。一方で上記のデータアナリシスなど、産業界からの要請で爆発的に進化している分野では、我が国の人材育成の対応は遅れがちであると言われている。また、欧米においては応用数学の主要分野として計算理論や最適化理論、データアナリシスがあり、数学を志望する優秀な学生が本分野に参入し、博士を取得して産業界で優遇される。それに対して、我が国では情報科学と数学の協働関係が希薄であり、国際的に高レベルの数学能力を活かせていない。近年は ERATO や新学術領域等のプロジェクトで数学能力の高い学生の啓蒙や勧誘を行い、状況の改善を図っている。

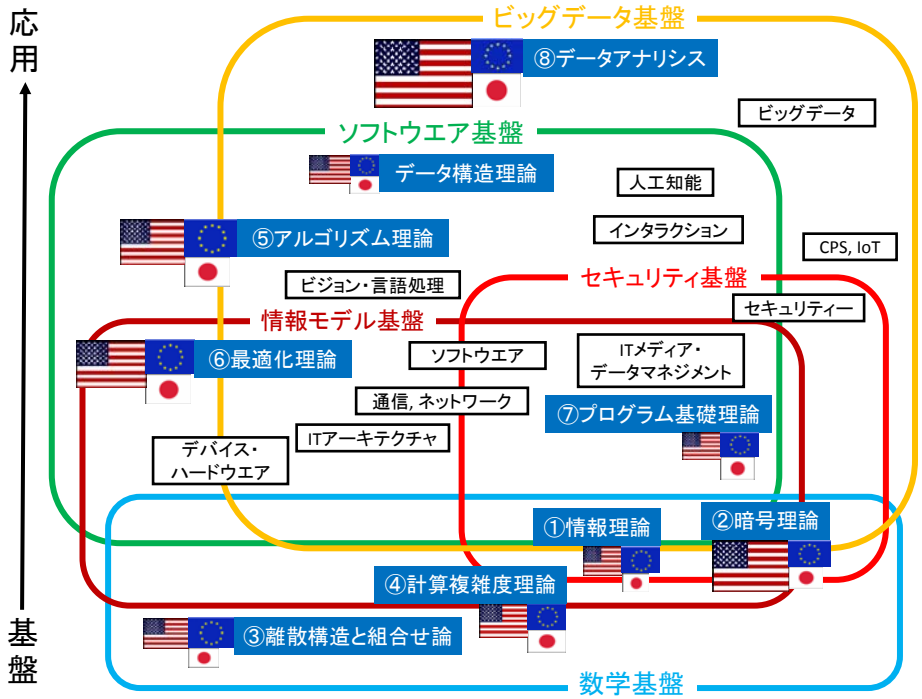

図 3.1.1 基礎理論の俯瞰図

情報科学技術の 5 つの基盤（数学基盤、情報モデル基盤、セキュリティー基盤、ソフトウエア基盤、ビッグデータ基盤）を支える基礎理論における研究開発領域を青色白抜きで示す。国旗は米国、EU、日本における戦略的重点化の現状を模式的に表現。

３．１．１　情報理論

（１）研究開発領域名

情報理論

（２）研究開発領域の簡潔な説明

情報理論は、近年のネットワーク情報社会の急激な技術進歩をささえる基本理論として欠く事はできない。情報理論は情報や情報の変換（情報処理や情報通信等）の本質を数理モデル化し、その数理モデル上で情報変換の性質や限界をエントロピー等の様々な情報量を用い明らかにすると共に、その限界を達成する最適な変換法（符号化，復号化等）を求めている。

物質の運動の本質が物理学によって解明されるように、情報の伝送や変換の本質は情報理論によって解明されており、情報理論は情報技術の基礎理論としての工学的側面のみでなく、科学的側面ももっている。文字、音声、画像はもちろん、知的情報や遺伝子等の生物情報まですべての情報を科学、工学両面から解明する情報理論は、情報通信，情報処理の問題のみならず様々な情報に関連する分野においてその有効性が確かめられている。

（３）研究開発領域の詳細な説明と国内外の動向

情報理論はシャノンの 1948 年の”A mathematical theory of communication”により誕生し、この論文では通信における①情報源符号②通信路符号③暗号の３つの問題を柱としているが、これらの問題は、情報の変換（処理）の基本問題をほとんど網羅しており、論文タイトルの communication の部分は現在では information と読み替えた方が適切であろう。計算の本質がチューリングマシンにより数理モデル化され解明されたのと同様に、情報理論では情報の変換の基本問題を数理モデル化し情報の本質を解明している。情報理論の研究領域は情報全てと言っても過言ではなく、非常に深く多岐に渡っているため、ここでは①と②の中の基礎的研究問題についてのみ説明する。

［情報源符号化問題］

情報源符号化問題はデータ圧縮技術等の基礎理論として有用で，情報源から発生する情報（文字、音声、画像等）をある記号列（符号系列）に変換（符号化）し、その系列を伝送や蓄積した後、受信者（利用者）が受信した系列を元の情報に変換（復号）する過程を数理モデル化することにより、情報とその変換の性質と限界を解明している。復号した情報が伝送したかった情報と一致する（歪みをゼロ）としたもとで、伝送できるレートの限界が情報源のエントロピーであることは良く知られている。

工学的には、エントロピーまで圧縮できる符号と復号法を構成することが目標となり、これは既に情報源の確率構造を巧みに利用したハフマン符号や算術符号等で実現されている。これらの基礎理論をベースに音声や画像に即した圧縮技術が開発され、情報化社会の巨大な情報の伝送、蓄積の効率化に不可欠の技術となっている。

他方、情報の性質を解明すると言う意味では、情報源や情報系列に含まれる情報の量を測る測度としてエントロピーが本質的であることが明確となり、言語や遺伝子等に含まれる情報の量の考察や、乱数をはじめ系列のランダム性の考察等にも無くてはならない理論となっている。

また、情報源の確率構造が未知の場合でも、ベイズ符号等のユニバーサル符号を用いることで、やはりエントロピーまで圧縮は可能であり、符号長がエントロピーに漸近する速度まで精密に求められている。この問題は未知のモデルの学習問題や時系列データ解析の問題等とも密接な関係があり、それらの問題の最適なアルゴリズムの構築やその限界の解明にベイズ符号等の漸近解析の結果が重要な役割を演じている。

［通信路符号化問題と符号理論］

符号理論は、伝送した信号に誤りが生じて受信側に伝送された場合にその誤りを訂正する技術の基礎理論として、すべての情報伝送と蓄積になくてはならない理論となっている。例えば、最も初期に提案された１ビットの誤りまで訂正できるハミング符号は、コンピュータ内部の情報蓄積、通信において欠くことの出来ない技術であり、その後もっと強力な誤り訂正能力をもったリードソロモン符号、畳み込み符号、LDPC（低密度パリティ検査）符号等が開発されている。もしこの技術がなければ、文字、音声、画像情報の携帯電話、地デジ、インターネットによる伝送や、ハードディスク、DVD の記録再生のような日常的な行為が不可能となってしまうほど、現代社会に浸透した技術である。

通信路符号化問題では、情報源から発生する情報を符号化した符号系列を、伝送や蓄積を行なう過程（通信路）で確率的に誤りが生じ、符号系列の一部が変化した系列を受信者が受信したもとで、その受信系列を元の情報に復号する問題を数理モデル上で扱っている。この問題において、誤りをゼロに限りなく近づけるもとで、通信の効率であるレートの上限は，通信路の相互情報量を入力分布で最大化した通信路容量となることが示されている。

相互情報量は二つの情報系列において，一方の情報系列が含んでいるもう一方の系列の情報量を表していると解釈され、二つの情報系列の関連性や距離の尺度として、通信の問題以外でも様々な研究分野で応用されている。例えば、パターン認識における特徴量とクラスの関連性を表す量として、遺伝子間の近さの尺度として等、学習理論や生物情報の分野等でも重要な役割を担っている。

［情報理論独特の研究アプローチ］

情報理論は対象とする問題を抽象化して数理モデル化し、その問題の限界をまず導出する。例えば、情報源符号の圧縮の限界であるエントロピーや通信路符号の伝送の限界である通信路容量等がそれに対応する。そして、その次にその限界を達成する方法（符号や復号法）を構成することをめざす。このような研究アプローチのため、技術者にとっては達成しなければならない目標が明確になり、現時点の技術が最適なものに比べどの程度であるかも定量的に把握できる。また、限界の証明や問題のもつ性質の理論的解析は、限界をめざす技術の開発に深い示唆を与えるものとして非常に役立っている。

例えば、通信路符号化の問題は、誤りなく伝送できるレートの限界として通信路容量があり、それを達成する符号と復号の存在はシャノンによって示された。しかし、それを具体的に構成する方法は示されておらず、達成限界を目指した研究がそれ以後開始され、現在も精力的に続けられている。リードソロモン符号や畳み込み符号は現在も多くの通信で利用される優れた特性を持った符号であるが、そのレートは通信路容量に遠く及ばなかった。ところが近年、ターボ符号や LDPC 符号が通信路容量の限界に迫る符号として注目が集まり、世

界中の研究者によって研究が進められ、急速に実用化されてきている。最近では、通信路容量を達成する符号として空間結合符号やポーラ符号が提案され、限界を達成する目標は、ある条件を満たした通信路に対して到達したが、一般の通信路や有限の符号長での性能向上等、まだまだ多くの問題が残されている。

（４）科学技術的・政策的課題

ここまでは、多岐にわたる情報理論の取り扱う問題の中で，情報源符号化と通信路符号化の問題において、送信者と受信者が１対１の通信を行なう最も基本的な数理モデルの問題について説明してきたが、従来から多対多の通信を扱う多端子情報理論（ネットワーク情報理論）も盛んに研究が行われている。特にインターネットをはじめ、コンピューター、センサー、データ等が大規模に分散するネットワーク情報化社会において、ネットワーク情報理論はこれらの問題の基礎理論として重要視されている。しかし、１対１の問題とは異なり、非常に複雑な数理モデルとなるために多くの問題は未解決であり、圧縮や伝送の限界すら特別な条件のもとでしか求まっていない。ネットワーク上での情報の変換（処理）の技術課題は急速に増加しているにも関わらず。それ支える基礎理論であるネットワーク情報理論の進歩が追いついていないというのが実情である。

また、情報理論において、圧縮や伝送の限界や最適な符号、復号法を求める場合に、情報系列や符号系列を無限長まで延ばした場合の漸近的な解析が多く用いられてきた。この漸近理論による解析結果でも、実用的な符号、復号法や情報処理アルゴリズムを構成するために十分有用であったが、我々は実際には有限長の情報や符号系列を扱っており、さらに厳密な解析が求められていた。近年、今まで漸近的解析を行なっていた問題を有限長でも解析できる理論が提案され、盛んに研究が行われている。しかし、情報理論の多岐にわたる問題のうち、有限長で解析された問題は限られており、実用技術と理論との橋渡しを担うこの有限長解析の理論の更なる発展が望まれている。

このような情報理論の更なる進歩のためには、周辺研究分野との連携は必須であり、各種数学はもちろん、統計理論、最適化理論、計算量理論等、従来から関連していた分野とより一層密接に研究を進めていくことが必要であろう。また、情報の変換、処理の本質を解明する情報理論は、様々な情報を扱う分野の基礎理論として非常に有用であるにもかかわらず、まだまだ通信、情報セキュリティ、学習等の周辺分野の研究者しかその有用性に着目しておらず、情報を扱う研究分野の研究者ですら、情報理論の扱える問題の広さや最近の進歩について十分な理解がされていないのが実情である。これらの分野の研究者や技術者と情報理論研究者のコラボレーションにより、その分野の進歩の加速化やブレークスルーが期待されるので、そのような仕組みづくりが望まれる。

日本の情報理論研究者は、世界的に見ても独自性の高いトップレベルの成果を出し続け、一つのステータスを近年までに築いてきていた。しかし、情報理論のような応用数学の研究者層がもともと薄い日本において、近年、実用研究重視の傾向や、企業の基礎研究に対する予算削減等もあり、世界と比較して相対的な研究レベルの低下が危惧されている。

欧米は情報理論研究の歴史も古く、研究者層がもともと厚い上、近年は中国、インド、韓国等の優秀な留学生や研究者が、欧米の情報理論研究の盛んな主要大学に膨大な数流入し、さらに層が厚くなると同時に研究論文の量、質ともさらに伸びてきている。これらの留学生

や研究者は、各国から、大学や様々な企業から、豊富な研究のサポートを受けられており、最新の研究成果の多くは彼らが担っていると言える。

残念ながらこの分野での日本から欧米への留学生や研究者は最近では減少する傾向にあり、このような世界的潮流にも乗り遅れてしまいつつある。研究者のポテンシャルでは、日本はまだ米国、ヨーロッパに次ぐ３番手と目されおり、日本国内で研究者個々としては精力的に研究活動を行なっているとはいえ、このままでは、あっという間に取り残されてしまう可能性を含んだ危機的状況にある。日本における情報理論のような基礎研究の衰退が、日本の技術、ひいては日本社会に与える影響が非常に危惧される。

（５）注目動向（新たな知見や新技術の創出、大規模プロジェクトの動向など）

［通信路容量を達成する符号，復号法］

伝送の限界である通信路容量を達成する、あるいはそれに迫る性能を持った符号、復号法が、(3)でも述べたように近年相次いで発表され、盛んに研究が行われている。また、この研究は多端子通信路や変復調など情報理論の中の研究分野で広がりを見せているだけでなく、周辺分野とも関連が深い。

例えば、ターボ符号、LDPC 符号、ポーラ符号の復号法では、最適な復号である事後確率最大の符号系列を求めるために、符号と通信路の確率構造を表現したグラフ上でメッセージを伝搬させるアルゴリズム（メッセージ伝搬法）を用いて、事後確率を効率的に求められたことが、ブレークスルーの重要な要因となった。データ解析や学習理論等の分野においても、問題の本質が事後確率計算に帰着される問題は多く存在し、それらの問題の解法アルゴリズムとしてもメッセージ伝搬法は有用で、これらの分野とも相互に関連をもって研究が進められている。

［スパース数理モデル問題］

圧縮センシング問題、スパース構造学習問題等，類似の問題がいろいろな呼ばれ方をしているが、推定や符号化する対象が疎な（スパースな）構造であると仮定されたり、ある仮定からスパースな構造が現れるためこのように呼ばれている。自然界の情報や生物情報にはこのような仮定が成り立つと考えられる問題も多く存在し、近年、スパース数理モデルの研究が盛んになっている。

誤り訂正の線形符号において、正しく復号可能な誤りの個数の限界や、効率よい復号アルゴリズムの研究は長い歴史がある。圧縮センシングで扱う、観測された疎なベクトル情報を線形変換で効率よく圧縮する問題と、復号の問題は本質的に同類の問題とみなせ、その限界の導出などで密接な関係をもっている。

また、効率的なアルゴリズムとして、L1 ノルム（ベクトルの各成分の絶対値和）最小化等の正規化項を付加した最適化アルゴリズムが用いられているが、統計学の統計モデル選択問題や正則化項付き学習問題とも密接な関係がある。例えば、代表的統計モデル選択問題である多重回帰分析の変数選択問題（データからどの説明変数を統計モデルに取り込むべきか判断をおこなう問題）において、AIC（赤池情報量規準）や MDL（最小記述長）等のモデル選択基準では、少ない変数で構成されるスパースなモデルが選ばれる傾向（オッカムの剃刀と呼ばれる）がある。さらに L1 ノルム最小化に対応する Lasso や Elastic Net 等の正規

化項付き最適化手法も従来から研究されており、スパース数理モデル推定と密接に関係している。

このようにスパース数理モデル問題は、統計理論、学習理論等、従来から様々な分野で研究されてきた応用的にも重要な問題であると同時に、情報変換（処理）の本質的問題の一つと考えられ、最近、各分野の成果も含め統一的な視点から研究が展開されてきている。今後、周辺分野とさらに密接な関係を保ちつつ研究の進展が期待されている。

［情報変換の限界の有限長解析］

情報理論における解析は、(4)で述べたように情報系列や符号系列を無限長まで延ばした場合の漸近的な解析が多く用いられてきたが、最近、有限長で解析する理論が提案され、情報理論において、今まで漸近的解析しか行なわれていなかった様々な問題に対して、有限長での解析の研究が精力的に行われている。有限長の情報列を扱う、実用的な符号、復号法や情報処理アルゴリズムの性能評価や高性能アルゴリズム構築のためには、この有限長解析の研究がさらに進展していくことが期待されている。

（６）キーワード

エントロピー、条件付きエントロピー、相互情報量、相対エントロピー、符号化、復号化、通信路容量、符号長、圧縮率、符号化レート、歪み、復号誤り率、シャノン理論、情報源符号化、データ圧縮、確率過程、符号理論、通信路符号化、符号化変調、系列、無線通信、通信方式、光通信理論、検定と推定、暗号、情報セキュリティ、ネットワーク情報理論、情報ネットワーク、ネットワーク符号化、量子情報理論、画像・音声処理、信号処理、圧縮センシング、パターン認識と学習、量子符号・暗号、記録素子用の符号化・信号処理、情報理論応用

（７）国際比較

国・地域	フェーズ	現状	トレンド	各国の状況、評価の際に参考にした根拠など
日本	基礎研究	◎	→	情報理論研究の歴史は欧米についで古く、通信路容量の効率的計算法、ユークリッド復号法、嵩符号、岩垂符号など歴史的に重要な研究成果をあげ、近年でも、情報理論の成果を一般情報源にまで拡張した情報スペクトル的方法等の研究成果があり、独自性の高いトップレベルの研究を行ってきている。情報通信や情報理論における最高の栄誉であるシャノン賞を欧米以外で受賞しているのは日本のみで、これまでに２名の受賞者を輩出している。また、最大の国際会議であるIEEE ISITはアメリカとヨーロッパで毎年基本的に交互に開催されているが、例外として日本で2回開催されている（韓国で1回開催され，香港では2015年6月開催予定）。国内の情報理論研究者のほとんどが属する電子情報通信学会の情報理論とその応用サブソサエティは、米国のIEEE Information Theory Societyには規模的に遠く及ばないものの、これに次ぐ規模の情報理論コミュニティであり、年次大会である情報理論とその応用シンポジウム（SITA）、隔年ごとの国際シンポジウム（ISITA）を開催する等，活発な活動を行っている。 しかし、(4)の課題でも述べたように、欧米での研究の進展の加速化、中国、インド、韓国の研究者の増加等、相対的な研究レベルの低下は避けられない状況となっている。それは、国際会議への参加者数や、IEEE Trans. Information Theory等の主要論文誌への掲載数の相対的比率の減少などにも見て取れる。
米国	基礎研究	◎	↗	シャノンによって創成された情報理論は、その弟子達をはじめ米国の研究者を中心に発展し、現在でもその層の厚さ、レベルの高さは他国の追従を許していない。MIT、プリンストン大、スタンフォード大等、米国の主要大学には、海外からも優秀な留学生、研究者が集まり、最先端の研究が行なわれており、情報理論の全ての分野で次々と新しい成果が出ている。米国のIEEE Information Theory Societyは世界の主要な情報理論研究者ほとんどが属する情報理論最大の学会で一極集中となっている。最高賞のシャノン賞もほとんどがアメリカの研究者によって占められているのが実情である。また、最近は中国、インド、韓国から優秀な研究者が流入することで、研究の量と質両面でさらなる発展がみられ、米国中心の構造はますます強まりつつある。
欧州	基礎研究	◎	→	歴史的にも、研究者の層の厚さや研究成果の面でも、米国に次ぐのが欧州である。もともと数学が強い東欧では、米国の情報理論の主流の流れとは異なる数学理論を用いた独自の情報理論を発展させ、イスラエルでは、米国と密接な関係を保ちながら画期的な研究を行っている等、様々な分野で重要な研究成果を出している。近年でも、通信路容量を達成またはそれに迫る符号や復号法の提案等、情報理論の根本的発展に寄与する研究はヨーロッパから発信されているものも少なくない。シャノン賞も米国に次ぐ数の受賞者を輩出しているが、欧州を統合する学会がないため、米国のIEEE Information Theory Societyに対抗する極とはなり得ていない。
中国	基礎研究	○	↗	従来から中国系アメリカ人によって、重要な研究成果が発表されていたが、近年、中国から膨大な人数の留学生が、米国の主要大学の有名研究室に籍を置き、様々な分野で最新の研究を行なっている。国際会議での発表件数も恐らく中国系研究者が最多と思われ、情報理論研究の一大勢力となる勢いである。中国の大学自体ではこのような基礎研究は、まだそれほど盛んではないが、2015年6月、情報理論界最大の国際会議であるIEEE ISITが中国（香港）ではじめて開催される等、今後急速な発展をとげることが予想される。
韓国	基礎研究	○	↗	多くの研究者や留学生が、アメリカの主要大学や主要企業の研究所において研究を行なっており、優れた成果も出てきつつある。2014年まで、最大の国際学会であるIEEE ISITが日本以外のアジアで開催されたのは韓国（ソウル）のみであり、今後ますます研究が発展していくと思われる。

（註 1）フェーズ
基礎研究フェーズ ：大学・国研などでの基礎研究のレベル
応用研究・開発フェーズ ：研究・技術開発（プロトタイプの開発含む）のレベル
産業化フェーズ ：量産技術・製品展開力のレベル
（註 2）現状
※我が国の現状を基準にした相対評価ではなく、絶対評価である。
◎：他国に比べて顕著な活動・成果が見えている、 ○：ある程度の活動・成果が見えている、
△：他国に比べて顕著な活動・成果が見えていない、×：特筆すべき活動・成果が見えていない
（註 3）トレンド
↗ ：上昇傾向、 →：現状維持、 ↘：下降傾向

（８）引用資料

[論文]

C. E. Shannon, "A mathematical theory of communication," Bell System Technical Journal, vol. 27, pp. 379-423 and 623-656, (1948).

[主要論文誌，研究報告]

IEEE Transactions on Information Theory
http://www.comm.utoronto.ca/trans-it/
電子情報通信学会技術研究報告（情報理論）

[一部に情報理論の研究成果を含む論文誌，雑誌]

電子情報通信学会(IEICE)論文誌 A
電子情報通信学会誌
IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences
IEICE Fundamentals Review

[主要会議]

IEEE ISIT (International Symposium of Information Theory)
IEEE ITW (Information Theory Workshop)
ISITA (International Symposium of Information Theory and It's Applications)
情報理論とその応用シンポジウム (SITA)

[著名な書籍]

- T. M. Cover and J. A. Thomas, "Elements of Information Theory", (John Wiley and Sons, 1991).
- 有本卓, "情報理論", (共立出版, 1976).
- 韓太舜, "情報理論における情報スペクトル的方法", (培風館, 1998).
- R. E. Blahut, "Thoery and Practice of Information Theory", (Addison-Wesley, 1987).
- I. Csiszár and J. Korner, "Information Theory：Coding Theorems for Discrete Memoryless Systems", (Academic Press, 1981).
- R. G. Gallager, "Information Theory and Reliable Communication", (John Wiley and Sons, 1968).
- 今井秀樹, "符号理論", (電子情報通信学会, 1990).
- W. W. Peterson and E. J. Weldon, "Error-Correcting Codes", (The MIT Press, 1972).

・ T. Richardson and R. Urbanke, “Modern Coding Theory”, (Cambridge University Press, 2008).
・ S. Lin and D. J. Costello, “Error Control Coding”, (Prentice Hall, 2004).
R. E. Blahut, “Algebraic Codes for Data Transmission”, (Cambridge University

３．１．２　暗号理論

（１）研究開発領域名

暗号理論

（２）研究開発領域の簡潔な説明

高度に安心・安全な通信を実現し、その通信の安全性を保証する理論

（３）研究開発領域の詳細な説明と国内外の動向

暗号技術の歴史は古く、シーザー暗号、エニグマをはじめ、当初は、軍事的な利用が主であった。1976 年の共通鍵暗号 DES（Data Encryption Standard）の制定、1976 年の公開鍵暗号の概念の提案、1977 年の RSA 暗号の提案を元に、商業的な利用に前進した。当初は、計算機リソースの不足などで、暗号技術は、広く社会に浸透することはなかったが、社会的ニーズや、方式の効率化、計算機能力の進展などにより、徐々に社会に浸透している。

暗号を安全に用いるためには、用いる暗号の安全性評価が必須である。しかしながら、暗号理論の黎明期は、新方式の「提案」と「解読」のイタチごっこの状況であった。現在、社会で用いられている暗号は、安全性が証明された暗号が用いられるべきであるし、実際、そのように理論整備が行われている。妥当な仮定なもとで、ある方式が安全であることを証明すること、および仮定の妥当性の理論的および実験による確認の研究が必要となる。理論的な研究においては、適応的選択暗号文攻撃の導入および安全な方式の提案も行われていたが、その重要性に関して懐疑的な研究者もいた。しかし、証明可能安全性の議論の重要性は、1998 年の Bleichenbacher による PKCS#1 v1.5 に対する攻撃（いわゆるベルコア・アタック）を契機に広く認識されるようになった。これにより、暗号は、単なる創意工夫から脱却し、確固たる理論研究へと進展している。実際に、電子政府推奨暗号リストでは、証明可能安全性を持つ方式が採用されている。

暗号技術が社会に浸透するにつれて、従来の想定よりも、多くの場面で用いられるようになっている。それに伴い、単に秘匿性を確保するだけでなく、より高機能な暗号の必要性が生じている。例えば、復号できる権限を属性の論理式により送信者が指定することができる「属性ベース暗号」や、暗号化したまま複雑な計算を行う「完全準同型暗号」などが、社会的ニーズに対応して提案されている。

その一方で、新たな暗号方式、暗号プロトコルが生み出される過程で、多くの場合、新たな数学的な問題の提起や計算量理論における困難な問題の提起も同時に行われている。このように、暗号理論は、他の基礎理論と密接に関係を持ちながら進展を続けている。

高機能な暗号が必要となる一方で、安価で省電力なデバイス上で暗号技術を用いる状況も増えている。例えば、大量のセンサーを用いてデータを収集し、収集したデータを用いて何らかのサービスを提供することが考えられている。その際、不正なセンサーを排除する目的で、認証を行う際や，取得したデータ自身の秘匿性を担保するために、暗号技術が使われる。このようなリソースの限られたデバイスに特化した軽量暗号技術が必要となる。

現在、広く用いられている暗号方式は、多くの場合、素因数分解、離散対数問題に安全性の根拠をおいている。特に制限のない素因数分解および離散対数問題の理論的な評価は一段落つき、これ以上の進展の可能性は少ないようである。その一方で、どの程度のコストを用

いれば、どの程度の素因数分解が可能となるのか？という評価が、今現在も大規模実験により行われている。さらに、特殊な構造を持った楕円曲線上での離散対数問題は、従来の想定よりも少ない計算量で解くことができることが徐々に明らかになっており、現在、重点的に研究が進行している。

近年考案されている暗号の多くは、LPN(Learning Parity with Noise)問題、LWE (Learning with Error)問題、Ring LWE 問題などの、比較的近年導入された問題の困難さに安全性の根拠をおいている。しかし、これらの問題の困難さは、正確に理解されていないため、実社会で用いるためには、さらなる困難さの理解が不可欠となる。特に、暗号理論においては、暗号方式の効率性自身が要請されるため、それほど大きなセキュリティマージンを取ることができない。そのため、セキュリティパラメターに応じた漸近的な評価と、実際の暗号で用いるパラメータでの評価という 2 種類の評価が不可欠である。

攻撃者の攻撃技術が高度化している現在、公開鍵などの公開情報や、暗号文などの通信路を盗聴することにより得た情報を元にした攻撃だけでなく、実際の暗号化装置から漏洩する情報（電力や電磁波など）を元にした攻撃も考慮に入れなくてはならない。近年でも、（コンピュータの動作中またはシャットダウン後数分以内にコンピュータに物理的にアクセスする攻撃者によって引き起こされる）Cold Boot 攻撃、Heartbleed バグ（オープンソース暗号化ライブラリ「OpenSSL」に混入したソフトウェア・バグ）等に起因する製品化（実装）の不備を元にした攻撃などが発見されており、依然、攻撃者の高度化がもたらす現実的な脅威が続いている。そのためには、事後対策のみならず、事前に脅威を防ぐ技術が必要となっている。

［他の基礎理論分野との連携による進展］

暗号理論は、他の基礎理論分野と連携を取ることにより進展している。1949 年のシャノンによる情報論的な暗号理論の創成は、暗号理論が情報理論とほぼ同時期に学問の道を歩み始めたことを意味している。また、1977 年の RSA 暗号の提案を契機とした「整数論の活用」や、1985 年の楕円曲線暗号の提案により生まれた「より複雑な代数構造の活用」は，数学分野、特に、代数学との密接な連携により生じており、今現在も精力的に研究が行われている。

1985 年のゼロ知識証明の提案は、計算複雑度理論分野からその概念がもたらされている。また、暗号の安全性の根拠となる困難な問題の多くが計算複雑度理論分野からもたらされていると同時に、暗号理論研究から提起された問題が計算複雑度理論分野へ提供されている。

［国内外の動向］

暗号技術は、全般的に、米国、欧州を中心に研究が進んでいる。暗号技術の特徴的な点として、イスラエルでも理論研究、応用研究とも研究が進んでいることが挙げられる。また、アジアでも、シンガポールで研究が進んでおり、優秀な研究者がシンガポールに集まりつつある。

欧州では、2004 年に ECRYPT（European Network of Excellence in Cryptology）project が始まり、欧州全体で仮想的な研究室を構成して研究が進展している。さらに、2008 年に後継プロジェクトである ECRYPT II が始まり、2013 年に終了している。また、世界的

な公募暗号のコンペティションとして、AES、SHA3、CEASER などが行われており、国内からも多くの暗号が提案されている。このコンペティションを通じて、多くの研究成果が生み出されている。

一方、国内では、産学官の研究機関が、必要に応じて連携を取りながら、理論研究から応用研究にいたるまで幅広く進めている。また、日本政府によるプロジェクトとして、暗号技術検討会 CRYPTREC（Cryptography Research and Evaluation Committees）による暗号技術の重点的な調査が行われている。

（４）科学技術的・政策的課題

・属性ベース暗号、関数暗号などの復号者制御可能な高機能暗号：
社会的ニーズの増加に伴い、アクセス制御などのアドホックな技術に頼らず復号権限を柔軟に設定することが可能である暗号（属性ベース暗号，関数暗号など）が提案されている。提案時では、現実的な計算量ではないものや、機能が制限されたものであった方式が、多くの研究者の改良により効率化が行われ、幅広いクラスに対応できるようになっている。しかし、依然、実用に耐えうる効率性や汎用性を有しておらず、さらなる改良や革新的なアイデアに基づくブレークスルーが必要である。

・完全準同型暗号：
社会的ニーズの増加に伴い、暗号化したまま様々な計算を行うことができる暗号（完全準同型暗号）が提案されている。属性ベース暗号、関数暗号と同様に、提案時では現実的な計算量ではなかったものの、多くの研究者の改良により効率化が行われている。しかし、依然、実用に耐えうる効率性を有しておらず、さらなる改良や革新的なアイデアに基づくブレークスルーが必要である。

・素因数分解、離散対数問題などの困難性評価：
現実社会で用いられている暗号技術は、多くの場合は、素因数分解、離散対数問題、および楕円離散対数問題の困難さに安全性の根拠をおいている。RSA 暗号の提案以後、素因数分解アルゴリズムの理論的な改良が継続的に行われている。同様に、世界規模での大規模な数値実験により、より大きな合成数の素因数分解に成功している。特定のビット数（例えば，1024 ビット）の RSA 暗号は、いつまで安全に用いることができるか？という評価を正確に行うことが本研究の目的である。その評価のためには、さらなる大規模な数値実験が必要である。

・LPN、LWE、Ring LWE 問題の安全性評価：
LPN 問題や LWE 問題は、符号理論における復号問題と類似した問題であるが、この問題の困難さを安全性の根拠とした暗号が提案されている。これらの提案により、問題の困難さに関する研究が重点的に行われている。この問題は、計算複雑度理論、情報理論との境界領域に属しており、活発に研究されている。個別の暗号の厳密な安全性を理解するためには、漸近的な計算量の評価だけでなく、現実の利用を想定した安全性評価が必要となる。

・耐量子計算機暗号：
量子計算機が実現すると、現在広く使われている素因数分解や離散対数問題の困難さに安全性の根拠をおく暗号の多くは破られてしまうことになる。そのためには、これらの問題の困難さに依存しない暗号の構成が必要となる。前述の LPN 問題、LWE 問題、Ring

LWE 問題などが、耐量子計算機対策として有望な問題の候補であり、これらの問題の困難さに安全性の根拠をおいた暗号方式の構成が急務である。

・サイドチャネル攻撃に対する対策：
公開情報、暗号文だけでなく、暗号デバイスから漏洩する情報（電力や電磁波など）を元にして、秘密鍵をあばき出す攻撃を総称して、サイドチャネル攻撃と呼ぶ。ハードウェアによる対策と並行して、漏洩情報が秘密鍵に依存しないように暗号化アルゴリズムを改良する研究も行われている。本研究はハードウェア研究とアルゴリズム研究の境界領域にあり、両者の協調が必要となる。

・軽量暗号：
IC チップなどのリソースが制限されたいデバイス向けの暗号技術が必要となる。軽量暗号では、電力などの制限により、高度な暗号プリミティブを用いることができない。そのため、安全なプロトコルを構成することが困難であり、攻撃者のモデルおよび効率性のバランスがとれたプロトコルの構成する技術が必要となる。

（５）注目動向（新たな知見や新技術の創出、大規模プロジェクトの動向など）

・大規模プロジェクトとして、国内では、JST CREST において、暗号理論に関連したプロジェクトが複数動いている。そのれらのプロジェクトの研究領域名は、「ビッグデータ統合利活用のための次世代基盤技術の創出・体系化」および「現代の数理科学と連携するモデリング手法の構築」であり、精力的に研究が進められている。

・新技術として、「難読化」が近年注目を浴びている。ソフトウェアの難読化は、知的財産を確保する上で極めて重要な技術であるが、暗号理論を最大限に活用して難読化を実現する研究が急速に進展している。「難読化」は、さらなる研究の進展と実用化が期待される。

・暗号学的に安全な多重線形写像が、ごく最近に提案され、多くの研究が始まっている。1999 年の双線形写像の暗号構成での活用を契機として、有用な暗号プロトコルの構築が促進されたのと同様に、多重線形写像を用いた暗号プロトコルの構築が盛んである。これまで想定していなかった暗号プロトコルの構成が実現する可能性があり、今後の研究動向に注視する必要がある。しかし、多重線形写像を用いた暗号プロトコルにおける暗号の安全性に関する数学的証明（安全性の根拠となる問題の数学的探求）は、まだ十分になされておらず、証明可能安全性に関して今後の研究が必要となる。

（６）キーワード

公開鍵暗号，証明可能安全性，安全性解析，高機能暗号

（７）国際比較

国・地域	フェーズ	現状	トレンド	各国の状況、評価の際に参考にした根拠など
日本	基礎研究	○	↗	大学、企業、国研、いずれにおいても精力的に研究を進めている。また、連携もとれており、共同研究も活発である。専門の研究会（情報セキュリティ研究会）および大規模なシンポジウム(SCIS)を通じて、研究者の交流が盛んに行われている。世界的に主要な国際会議（CRYPTO、Eurocrypt、Asiacrypt、PKC、TCC、CHES、FSE）に多くの研究成果を発表しているとともに、会議の誘致を積極的行っている。
	応用研究・開発	○	↗	産業界からのニーズに基づき、高機能暗号の研究が進展している。近い将来、実用化が期待されている。
	産業化	○	→	様々な状況で、暗号技術が使われている。しかし、最先端の理論が、実際の場面で用いられるまでには、依然、タイムラグが存在する。
米国	基礎研究	◎	→	多くの大学の研究者が参画している。また、IBM Watson研究所、Microsoft Researchなどの民間の研究所でも多くの研究が行われている。
	応用研究・開発	◎	→	完全準同型暗号の提案や実現に向けての取り組みで、IBM Watson研究所が先行している。
	産業化	◎	→	IBM、マイクロソフトが先行している。
欧州	基礎研究	◎	→	ECRYPT IIプロジェクトを通じて、多くの研究機関が基礎研究に参画しているおり、終了後も、その状況は続いている。フランスが、特殊なクラスの離散対数問題のアルゴリズムに関して、著しい成果を出している。
	応用研究・開発	○	→	企業の研究者が大学等と連携を取りながら、研究を進めている。
	産業化	◎	→	ICカードなどの物理デバイスに関するセキュリティでは、先行している。フィリップス、フランステレコム、ジーメンス、ノキアなどが先行している。
中国	基礎研究	○	↗	中国科学院（Chinese Academy of Sciences)、清華大、上海交通大などで積極的に研究が行われている。
	応用研究・開発	△	→	1999年に商用暗号管理条例が制定されており、中国国内における商用暗号化技術及び商用暗号化製品の科学研究、開発と生産、販売及び使用は、中国政府によって包括的に管理規制されている。応用研究・開発の動向は、外部から状況を伺い知ることができないが、企業による自由な研究・開発は、それほど活発ではないと推測される。
	産業化	△	↘	Huaweiが暗号技術の産業化に向けて先行している。
韓国	基礎研究	○	→	ソウル大、KAISTなどで盛んに研究が行われている。
	応用研究・開発	△	→	目立った活動は、外部からは見受けられない。
	産業化	△	→	サムスン電子が暗号技術の産業化に向けて先行している。

（註 1）フェーズ
基礎研究フェーズ ：大学・国研などでの基礎研究のレベル
応用研究・開発フェーズ ：研究・技術開発（プロトタイプの開発含む）のレベル
産業化フェーズ ：量産技術・製品展開力のレベル

（註 2）現状
※我が国の現状を基準にした相対評価ではなく、絶対評価である。
◎：他国に比べて顕著な活動・成果が見えている、 ○：ある程度の活動・成果が見えている、
△：他国に比べて顕著な活動・成果が見えていない、×：特筆すべき活動・成果が見えていない

（註 3）トレンド
↗ ：上昇傾向、 →：現状維持、 ↘：下降傾向

（８）引用資料

1) ISEC （情報セキュリティ研究会）
http://www.ieice.org/~isec/
2) IACR (International Association for Cryptologic Research)
http://www.iacr.org/
3) NIST (National Institute of Standards and Technology)
http://www.nist.gov/
4) ECRYPT II
http://www.ecrypt.eu.org/
5) CRYPTREC
http://www.cryptrec.go.jp/
6) インテリジェント暗号
http://www.seclab.ecl.ntt.co.jp/project/information-security/intelligent-crypto.html
7) Fully Homomorphic Encryption –IBM
http://researcher.watson.ibm.com/researcher/view_group.php?id=1548
8) CRYPTREC Report 2013 暗号技術評価委員会報告書の付録 7
http://www.cryptrec.go.jp/report/c13_eval_web_final.pdf
9) 平成 21 年度特許出願技術動向調査報告書．平成 22 年 4 月，特許庁発行

３．１．３　離散構造と組合せ論

（１）研究開発領域名

離散構造と組合せ論

（２）研究開発領域の簡潔な説明

離散構造は情報構造の数理定式化に広く用いられてきた手法である。特に、グラフ探索やネットワークフローなどの組合せ論的アルゴリズム分野の研究成果は，計算機の発祥期から計算機の設計や計算機による問題解決に広く利用されてきた。例えば、IT ビジネスの発展を語る際、Microsoft、IBM、Google、Yahoo、AT&T などの巨大企業で、多くの著名な離散数学者、組合せ論研究者が斬新なソフト開発と様々な問題点の解決に貢献してきた歴史を欠かすことはできない。これらネットワーク解析における離散構造と組合せ論に関する研究の重要性は近年ますます増加し続けており、上記の IT 企業は、現在でも常時 10～20 名の理論計算機科学・アルゴリズム・離散数学を専門とする研究者を抱えている。これは製品開発やシステム構築においては、理論的保障が常に必要不可欠な存在であるからである。また、Web 検索における PageRank や Hits 法などの IT ビジネスにおける発明は、主に理論研究に基づいて生み出されてきたという事実がある [1)]。しかしながら、日本では、理論研究者が日本企業に貢献した例が極端に少ないだけでなく、理論研究者が企業研究所に雇われている例もほとんどない。

（３）研究開発領域の詳細な説明と国内外の動向

【背景と意義】

離散構造と組合せ論は、現代の数学および計算機科学両分野において中核をなしている。これは数学分野では、例えば米国プリンストン大学の Institute of Advanced Study（高等研究所）の数学科において「Theoretical Computer Science and Discrete Mathematics (CSDM)」というプログラムが 2000 年より 15 年近く続いているという事実によく現れている。さらに計算機科学分野でも、歴代ネバンリンナ賞受賞者のほぼ全員が「離散構造と組合せ論」で仕事をしていることを見ても明らかである。

離散構造、組合せ論的解析が実世界のネットワーク解析に大きな役割を果たすことは、かつてより多くの研究者によって指摘されてきた。しかしながら、離散構造で記述されたデータを利用した意思決定や知識抽出をする際に必要な多くの重要な組合せ最適化問題は、計算階層理論的な計算困難性を持ち、巨大データに対しては理論保証を持つアルゴリズム設計が難しいという問題がある。

この困難性を打ち破るには、情報構造を表現する離散構造の特徴を利用した問題解決が必要不可欠である。チューリング賞を受賞した Hopcroft と Tarjan らによる平面性判定や平面グラフ同型問題の解決、またネバリンナ賞を受賞した Kleinberg らによる Web のグラフ理論的特徴を利用した検索エンジンの設計は、これらの困難性に対する解決を試みた新たな挑戦であり、後者は現在インターネット検索の基盤となっている。これらの例からも分かる通り、離散構造、組合せ論的解析が計算機科学において現在、決定的な役割を果たしている。

【これまでの動向】

離散数学が飛躍的に発展した 80 年代には、線形計画法（Linear Programming）に対する単体法、そして Khachiyan による多項式時間アルゴリズムの開発など、多面体的組合せ論（Polyhedral Combinatorics）が急速に発展した。また 1970 年代に開発された Szemeredi の Regularity Lemma が、整数論のみならず、グラフ理論の問題（とくに極値論的グラフ理論）に応用されるようになった。さらに 1980 年代には、群論における「単純群分類の完成」の大事業が終了したことによって、多くの群論研究者が組合せ論的群論、代数的組合せ論に転入してきた時期でもある。

90 年代に入ると、多くの組合せ論的近似アルゴリズムが開発されてきた。これらのアルゴリズムは、70 年代後半から活発に研究されてきた線形計画法の深遠な理解に基づくものであり、確率論的組合せ論が花開いたのもこの時期である。この確率論的解析は、90 年代以降、アルゴリズム分野のハイライトであるランダマイズ（乱択）アルゴリズムの大発展につながっていく。

90 年代以降、急激に進展したグラフアルゴリズムとグラフ理論の最先端の研究成果は、グラフマイナー理論の発展によるところが大きい。グラフマイナー理論は、90 年代、2000 年代における離散数学の最大のハイライトの一つである。また、グラフマイナー理論で開発された「構造的グラフ理論」的手法により、Robertson、Seymour、Thomas が多くの成功を収めた。特にパーフェクトグラフと呼ばれるグラフ特性とその構造に関する予想の解決は、現在までの離散数学における最大の成果の一つである。

2000 年代以降は加法的組合せ論（Additive Combinatorics）が、エクスパンダーグラフの発展とともに、現代数学そして理論計算機科学の中心となっていく。実際、数学からは Terry Tao、Ben Green、Tim Gowers、Van Vu、そして理論計算機科学からは Avi Wigderson などの指導的に立場にある研究者が、数学・計算機科学の垣根を越えて多くの共同研究を行っている。

また 2000 年代に入り、Kleinberg, Page らが Web グラフのグラフ理論的な解析を行い、その研究成果が Google などの検索エンジンに利用されている [3)]。

【日本国内での動向】

組合せ最適化分野において、日本人研究者は多くの重要な仕事を行ってきた。藤重悟（元京都大学数理解析研所長）、岩田覚（東大教授）の劣モジュラー関数最小化に関する研究成果は、2003 年に Fulkerson 賞を受賞するなど世界的に高い評価を受けている。また、室田一雄（東大教授）の離散凸解析理論に関する仕事は、SIAM から出版された本「Discrete Convex Analysis」が当該分野の指導書になるなど、世界的評価も高い。

グラフ理論、グラフアルゴリズム分野では、80 年代より西関隆夫（元東北大教授）、茨木俊秀（元京大教授）らの業績が世界的に高く評価され、両名とも ACM のフェローに選ばれている（日本人の ACM フェローは他分野も含め，現在までのところ 10 人にも満たない）。また 2000 年以降は、河原林健一（NII 教授）のアルゴリズム的グラフマイナー理論に関する研究が、離散アルゴリズム最高峰の国際会議 SODA の論文賞を受賞するなど、当該分野を世界的にリードしている。

また日本人研究者は、現在当該分野のトップジャーナルの Editor も務めている。一例を

挙げると、SIAM Journal on Discrete Mathematics (SIDMA)の岩田・河原林、Algorithmica、Journal of Graph Theory の河原林、Mathematical Programming の岩田などが世界トップクラスの学術誌で中心的な役割を担っている。

さらに若手研究者も数多く育っており、実際に日本のトップクラスの研究機関で准教授、助教として活躍している。一例を挙げると東京大学で三名、九州大学で二名、京都大学数理解析研究所で二名、国立情報学研究所で二名、中央大学で一名の当該分野の若手研究者が研究・教育に取り組んでいる。

以上のように、世界をリードする数名の研究者と有望な若手研究者を多数抱える日本の離散構造、組合せ論分野は、世界的にもその存在感を増しつつある。

【今後必要となる取り組み】

Web 構造や Facebook, Twitter など、我々の身近にある複雑ネットワークは、日々膨張を続けている。これらのネットワークは、年々急速に発達し、数年後には 10 倍のサイズになると予想されている。そのためこれらを解析するアルゴリズムの高速化が現在急務となっている。

この様な巨大なネットワークは、ユーザーに相当する「頂点」と、各ユーザーを結ぶ「辺」からなる離散構造「グラフ」をなしている。したがって、上記のネットワークに対するアルゴリズム的課題は、巨大グラフに対して動作する高速アルゴリズム開発とみなすことが可能である。

実際、このアルゴリズム的課題は、世界的にも注目されている研究テーマである。一例を挙げると、インターネットの原型である ARPANET や全地球測位システムの GPS を開発したことで知られているアメリカ国防省の研究機関である国防高等研究計画局（DARPA）が、2011 - 2012 年の研究テーマ（「Graph-theoretic Research in Algorithms and the Phenomenology of Social networks (GRAPHS)」）として、公募を行ったのはまさにこのテーマである [2)]。この研究公募の主な目的は、Twitter、Facebook 等の超巨大ソーシャルネットワークに対するモデル化、およびそれらのグラフに対する高速アルゴリズムの開発にある。

以上のように、理論研究の視点から実際の巨大グラフを解析する研究は、計算機科学における最先端の次世代チャンレンジの一つであると認識されている。

（4）科学技術的・政策的課題

【科学技術的課題】

10 万「辺」程度のグラフを扱っている場合、計算量が頂点数の 2 乗も必要となるアルゴリズムでも、通常のサーバーで辛うじて計算が可能である。しかし、1 億「辺」以上のグラフを扱う場合、計算量は O(n log n)程度のオーダーが限界である。そのため膨張し続ける現在の巨大グラフに対しては、新しい高速アルゴリズムの開発が急務となっている。現在までのところ、ほぼ線形時間で動くアルゴリズム開発は、ダイクストラアルゴリズムなどの特殊なアルゴリズムを除き、理論的にも困難（不可能）であると考えられてきた。これらの問題を乗り越えるためには、ソーシャルネットワーク、Web グラフ、道路網などのグラフの「特徴」を深く理解し、それらの情報構造に基づいて離散構造の手法を最大限生かした高速アルゴリズムの開発が必要不可欠である。

【政策的課題】

グラフ理論、組合せ論に基づくアルゴリズムは、HITS や PageRank など、最先端の IT 企業で極めて重要な役割を果たしている。これら巨大ネットワークとそれに対する高速アルゴリズムの開発は、現在、全世界が共通して抱える最優先課題であり、その解決にあたっては各国の研究者が理論研究の成果を常に発信し、情報を共有する必要がある。情報学の世界では、これらの研究の担い手である研究者は、すでに 30 代で世界を牽引する業績をあげていることが多い。現代の我々が抱える複雑ネットワークの問題においても、これら 30 代・40 代の研究者が業界をリードし、研究者のみならず IT 企業とともに理論面・実用面両面から解決に取り組むことが不可欠である。

（５）注目動向（新たな知見や新技術の創出、大規模プロジェクトの動向など）

前述の通り、ARPANET、GPS を開発したことで知られているアメリカ国防省の研究機関である国防高等研究計画局（DARPA）が、2011 - 2012 年の研究テーマとして、「Graph-theoretic Research in Algorithms and the Phenomenology of Social networks (GRAPHS)」を公募した [2)]。この研究公募の主な目的は、Twitter、Facebook 等の超巨大ソーシャルネットワークに対するモデル化、およびそれらのグラフに対する高速アルゴリズムの開発にある。

さらにドイツでは、組合せ最適化分野におけるトップ研究者 M.Grotchel 氏を代表に据え、「最適化」研究に特化した ZIB（Zuse Institute Berlin）という研究組織が設立されている [4)]。ZIB は、ドイツ鉄道（DFB）やルフトハンザ航空のスケジュールの効率化を達成するなど、最適化技術を実用的問題に応用し多くの成果をあげている [4)]。

（６）キーワード

グラフ理論、組合せ論、組合せ最適化、巨大グラフ、ソーシャルネットワーク、Web グラフ、高速アルゴリズム

（７）国際比較

国・地域	フェーズ	現状	トレンド	各国の状況、評価の際に参考にした根拠など
日本	基礎研究	◎	↗	上記の通り、世界をリードする研究者を数人輩出し、またそれに続く若手も育ってきている。今後も日本人研究者が、当該分野で非常に大きな役割を担っていくであろう。
	応用研究・開発	○	↗	近年、基礎研究者と応用研究者の共同研究が始まっているものの、大きな成果はまだ出ていない。今後の発展が期待される。
	産業化	△	→	基礎研究が産業化につながっていない。
米国	基礎研究	◎	↗	世界の主要の研究者を抱えている。今後も世界の中心地として当該分野をリードしていくと思われる。
	応用研究・開発	◎	↗	応用数学科、計算機科学科に所属している離散構造と組合せ論の研究者が多くのIT企業と共同研究を行っている。また、MICROSOFT、 GOOGLE、YAHOO、FACEBOOKなどの巨大IT企業は、多くの当該分野の研究者を抱えている。彼らは常に応用研究開発に携わっている。
	産業化	◎	↗	上記の共同研究から、巨大IT企業でいくつかの製品化、サービス化が生まれてきている。Googleの検索エンジンも、基礎研究の技術が決定的な役割を果たした。
欧州	基礎研究	◎	↗	イギリス、フランス、ドイツなどの欧州の大国は、自然科学、特に数学に力を入れてきた。これらの国々では、離散構造、組合せ論も数学分野の一つとして発展してきた。実際多くの数学者がヨーロッパにおいて、当該分野の研究をけん引している。
	応用研究・開発	○	↗	ビッグデータ時代に、巨大企業（主に自動車，携帯電話会社）との共同研究が始まってきている。また、ドイツのZIBなど、最適化を使った応用開発も始まってきている。
	産業化	○	→	企業等共同研究が始まったばかりであるので、まだ目立った成果は出ていない。
中国	基礎研究	△	↘	ほとんどの世界的研究者は、中国国内ではなく米国、欧州で研究している。
	応用研究・開発	△	↘	基礎研究で終わってしまう場合が多い。
	産業化	×	↘	上記と同様。
韓国	基礎研究	○	→	何人かの世界的に著名な研究者を抱えるものの、若手研究者がそれほど育っていない。
	応用研究・開発	△	↘	基礎研究で終わってしまう場合が多い。
	産業化	×	↘	上記と同様。

（註 1）フェーズ
基礎研究フェーズ ：大学・国研などでの基礎研究のレベル
応用研究・開発フェーズ ：研究・技術開発（プロトタイプの開発含む）のレベル
産業化フェーズ ：量産技術・製品展開力のレベル

（註 2）現状
※我が国の現状を基準にした相対評価ではなく、絶対評価である。
◎：他国に比べて顕著な活動・成果が見えている、 ○：ある程度の活動・成果が見えている、
△：他国に比べて顕著な活動・成果が見えていない、×：特筆すべき活動・成果が見えていない

（註 3）トレンド
↗ ：上昇傾向、 →：現状維持、 ↘：下降傾向

（８）引用資料

1) PageRank について
http://ja.wikipedia.org/wiki/%E3%83%9A%E3%83%BC%E3%82%B8%E3%83%A9%E3%83%B3%E3%82%AF
2) DARPA の Graph-theoretic Research in Algorithms and the Phenomenology of Social networks (GRAPHS)のオープンコール
http://www.darpa.mil/Our_Work/DSO/Programs/Graph-theoretic_Research_in_Algorithms_and_the_Phenomenology_of_Social_Networks_(GRAPHS).aspx
3) HITS について
http://en.wikipedia.org/wiki/HITS_algorithm
4) ZIB（Zuse Institute Berlin）について
http://www.zib.de/de/home.html

３．１．４　計算複雑度理論

（１）研究開発領域名

計算複雑度理論

（２）研究開発領域の簡潔な説明

与えられた計算問題に対し、どこまで効率の良いアルゴリズムを作ることができるのか、その限界を追求する研究領域である。計算の本質に最も直接的に迫る研究であり、その成果は、革新的なアルゴリズムや新たな計算機構の創出にもつながり、情報セキュリティの安全性の拠り所となる基礎理論も提供する。

（３）研究開発領域の詳細な説明と国内外の動向

【背景と意義】

情報科学技術の重要なテーマは、社会の様々な要求にコンピュータ・システムを用いることでいかに応えていくか、という点である。そのようなコンピュータの利用を設計する課程で、必要とされる仕事が具体的な計算問題として明確になったときに、それを効率的に処理する計算方法（アルゴリズム）が、どこまで効率的に成り得るか、その限界を追求する研究領域が計算複雑度理論である。

コンピュータ発祥の元とも言われているのが、ゲーデルやチューリングらによる計算可能性・不可能性の概念の定式化とその探究であった。こうした研究では、与えられた計算問題に対する純粋な計算可能性が主題であった。それに対し、計算時間などの計算資源に対する課せられた制約の下で、与えられた計算問題が計算可能か否かを議論するのが計算複雑度理論である。つまり、より定量的な形での計算限界を探究する分野である。

同じ計算限界に挑む研究領域にアルゴリズム理論（３.１.５）がある。アルゴリズム理論では、与えられた計算問題に対して、より効率的なアルゴリズムの開発を目指す立場から計算限界に迫ろうとしている。それに対し、計算複雑度理論では、効率的な計算の可能性よりも、その不可能性を追求し、そのため、計算の困難さの要因や特徴づけを明らかにしようとしている。その意味で、アルゴリズム理論と対をなす研究分野である。

古来より、角の三等分の作図不可能性や 5 次以上の方程式の解公式の不可能性などの例が示すように、困難性や不可能性の追求から、対象となる概念（たとえば，作図や解公式）の本質が明らかにされてきた。それが、新たな学問分野の誕生や、さらには人類の考え方に関わる変革をもたらす場合もあった。実際、計算不可能性の研究から「計算は単純な処理ステップから構成できる」ことが明らかになり、それが現在のコンピュータの設計思想のもとにもなった。計算複雑度理論でも、効率的計算の重要な側面や、計算の困難さを特徴付けている要因が数多く見出されており、その中から、情報科学技術の新たな分野も生まれてきた。

その重要な例の 1 つが公開鍵暗号プロトコルである。暗号理論（３.１.２）を変貌させ、現代の情報セキュリティ技術の核となった考え方は、NP 問題の困難さの特徴付けの 1 つとして見出された計算の一方向性についての研究から生まれた成果である。また、データアナリシス（３.１.８）の最初の原動力となった規則発見のためのアルゴリズム技法として著名なブースティング技法は、弱学習と強学習の計算論的等価性の研究から生まれたものである。

計算複雑度理論は、情報科学技術の土台であり、様々な分野の基礎となる領域で鍵となる

考え方や技術、そして重要な研究課題を与えてきた。それとともに、そうした各分野の基礎領域で活躍する人材も供給してきた。「計算論的○○学」と呼ばれるような分野の牽引者の中の多くが、計算複雑度理論の研究者からの転進者である。つまり、計算複雑度理論は、人材の育成・輩出という観点からも情報科学技術の土台としての役割を果たしてきたのである。

【これまでの取組み】

効率的なアルゴリズムの開発は、コンピュータの登場の当初から重要な課題であった。チューリング機械などの抽象的な計算モデルを用い、アルゴリズムの計算効率を数学的に議論するための枠組みは、1960 年代の初頭に複数の研究者[1]により整えられていった。そうした研究の上に立ち、1965 年、ハートマニスとスターンズは計算量というアルゴリズムの効率を測る尺度を提案し、さらにそれを用いて計算問題の困難さを議論した。彼らは計算量の異なる計算問題が無数に存在することを示し、計算量クラスという考え方を導入した。この研究から計算複雑度理論が始まったと言ってもよいだろう。コッブハムとエドモンズが多項式時間計算量を現実的に解ける計算の第一次近似的な基準として提唱したのも、この頃であった。なお、多項式時間計算量を持つ計算問題のクラスは現在ではクラス P と呼ばれている。

1960 年代の後半になると、現実問題に対処する中で登場してくる計算問題の中に、似たような特徴を持つものが多数あることが認識されるようになってきた。計算問題を解の探索問題として見たとき、「その解候補の検証は比較的容易」という特徴である。このような計算問題を NP 問題と呼び、その全体をクラス NP という。NP 問題では、解の検証は容易なので、理論上はすべての解の候補を調べれば解は発見できる。しかしながら、その候補数が非常に多いため、単純なしらみつぶし的な計算は実現不可能である。果たして、解の発見を、しらみつぶし的な探索より効率的に、たとえば多項式時間で行う計算方法が存在するか？（記号的に言えば，P＝NP か？）という点が重要である。

クックは 1971 年に、NP 問題の困難さを特徴付ける性質（NP 完全性）を提唱し、NP 完全性を持つ計算問題―― NP 完全問題 ―― の 1 つを具体的に示した。NP 完全問題は、その問題さえ効率的に解く（正確には多項式時間で解く）ことができれば、すべての NP 問題を効率的に解くことができる、といった性質を持つ計算問題である。つまり、P＝NP なのか否かの鍵となる問題であり、NP の困難さを特徴付ける問題である。このクックの発表の翌年にカープが、いろいろな分野で鍵となっている計算問題が、実は NP 完全性を持つことを示した。この 2 つの研究が契機[2]となり、数多くの NP 完全問題が情報処理の様々な分野で見出された。以来、こうした問題群に対する効率的な解法が探究されてきたが、多項式時間で解ける見込みは得られておらず、多くの研究者は、クラス NP の中にはクラス P に入らない問題、すなわち多項式時間では解けない問題があると予想している。これが「P≠NP 予想」である。

この P≠NP 予想の証明（もしくは反証）は計算複雑度理論の最も重要な研究課題である。その根拠として、(i) NP 問題が情報処理の多くの分野で鍵となるため、(ii) 革新的アルゴリズムの発見と計算論的な定式化に結び付くため、さらには、(iii) この予想の証明が非常に

1 その中で日本人研究者（山田尚夫）も重要な成果を出している。
2 実は、当時のソ連でも同じ発見がレビンにより同時期になされていた。クックの場合と異なるのは、レビンの結果に対してカープのように、その重要性を認識して、多くの実例を示す人物がいなかった点である。そのためソ連では、NP 問題の重要性の認識は広まらなかったのである。

難しいため、など複数の理由を挙げることができる。もちろん、P≠NP 予想が正しかったとすると、NP 問題群を完全に、かつ統一的な方法で効率的に解く計算方法は存在しない。そのような否定的なことであっても、それを証明することは情報科学技術の発展に重要である。P≠NP の証明が得られたとすれば、その成果として計算困難さの本質に対して数多くの事実が明らかになる。そうした成果は計算に対する新たな理解であり、それらをもとに革新的なアルゴリズムの構築法も生まれてくるはずだからである。たとえば、ある種の NP 問題に対しては、完璧ではないが十分利用可能な解法の発見の可能性も大きい。実際、これまでにも、NP 問題の計算困難さの特徴付けから、情報セキュリティ技術の基礎となった一方向関数の理論や、後で述べるランダムネスの解析理論が産み出されている。

計算複雑度理論では、この P≠NP 予想だけでなく、非常に多くの計算量とそれに基づく計算量クラスが定義され、その関係について様々な考察がなされてきた。計算量として代表的なものだけでも、領域計算量、乱択計算量、量子計算量、対話型計算モデルでの計算量、論理式計算量、回路計算量、算術回路計算量、通信計算量、証明複雑さ、記述複雑度などがある。これらの中には、通常のコンピュータの計算とは大きく異なる計算モデルに基づくものもある。そうした研究の全貌を簡潔には紹介できない。ただ重要なのは、こうした計算量や計算量クラスは、すべて計算のある特定の本質を理解するために導入され研究されてきた、という点である。たとえば、論理式と論理回路の計算能力の差に関する研究は、「計算の途中結果の再利用が計算効率の向上に本質的か？」という点を明らかにする研究と見なすことができる。こうした研究の成果、あるいは、その研究のために開発された解析手法や理論などが、新たなアルゴリズムや計算の新しい利用方法の基礎となってきたのである。

【今後必要となる取組み】

P≠NP 予想は計算複雑度理論の最も重要な課題ではあるが、情報科学技術の重要な基盤と成り得る計算複雑度理論の研究は、それだけではない。また、計算困難さとは直接関係なさそうに見える研究の中にも計算複雑度理論が深く関わっているものも少なくない。その中でも、今後 5 年間の研究次第で重要な基礎理論となり得る研究課題について述べる。

(a) 計算論的情報理論：計算論的意味を持ったデータ符号化に関する理論

現在、「計算論的○○学」という分野が多くあるが、計算複雑度理論の研究に端を発するものが多い。それにさらにもう 1 つ、計算論的情報理論とでも呼ぶことのできる分野が加わるかもしれない。そのような理論に発展し得る研究テーマがある。意味を持ったデータの符号化に関する研究である。従来の符号理論、情報理論は、どのようなデータも符号化できる技法についてのものであった。それに対し、意味を持ったデータは、その背後にある意味を利用すれば、より高性能でより効率的な符号化・復号化が可能と思われる。「背後にある意味」を定式化する 1 つの方法は、意味を解釈する計算を用いる方法である。たとえば、NP を特徴付ける「解の検証」がそのような計算に対応する。つまり、NP 問題の解のようなものは高度に符号化できる、という考え方である。実は、後述するように、そのような高度な符号化が NP 問題の解に対しては発見されている（確率的検証法）。この考え方を発展させた手法は、局所性質検査というテーマで研究されている。こうした研究を深めれば、背後に意味のあるデータに関する符号化法や、そのようなデータの情報量を解析するための基礎理

論が得られるだろう。それが計算論的情報理論である。背後に意味があるデータというのは、多くの巨大データにも共通する性質である。計算論的情報理論は、そのような巨大データの解析の基盤になることも期待できる。

(b) ランダムネスの計算論的な意味

乱数を使った計算は、シミュレーションなどで非常によく用いられている。それだけではない。1980 年代の中ごろから、乱数（ランダムネス）を使うことで（場合によっては誤ることもあるが）通常のアルゴリズムでは実現できない高速計算を行うアルゴリズムが続々と発見された。それらは乱択アルゴリズムと呼ばれている。一方で、「そうしたランダムネスは計算の効率化に本質的か？」という研究も計算複雑度理論の重要なテーマとなった。そもそも、ランダムネスとは何か？より具体的には、乱数を数学的にはどのように定義したらよいのか？というテーマは、確率論の創始者と言われるコルモゴロフなども研究したテーマであり、計算可能性理論の枠組みでも研究されてきた。その計算複雑度版の研究は、1980 年代の後半から始まり、2000 年には、たとえば、NP の困難さを仮定すれば、計算上のランダムネスを規則化できることが示された。これは脱乱化と呼ばれている。たとえば、乱数不要のシミュレーションが可能なのである。しかし、脱乱化の効率がどの程度期待できるのか、また、どのような計算に脱乱化が可能なのか、など実際への応用の基盤となるには不明な点が多い。さらには、計算の中のランダムネスの脱乱化が、データの中のランダムネスの解析にどのように利用できるか、という研究まで進めば、(a) と合せて、情報を解析するための新たな理論基盤になることが期待できる。

以上が情報科学技術分野の基礎にも成り得る研究テーマであるが、計算複雑度理論そのものを深めるためのテーマも挙げておく。

(c) 計算限界証明技法の統括的な研究

計算複雑度理論では、非常に多くの計算量とそれに基づく計算量クラスが定義され、その関係について様々な考察がなされてきた。その中で計算限界を証明するための重要な解析技法も数多く発見され、洗練されてきた。ただし、それらは、たとえば「P≠NP の証明には、まだまだ不十分である」と思われている。実際、主要な各解析技法に対し、その技法では、P≠NP の証明は不可能である、という技法の限界を示す結果も得られている。一方で、解析技法の解釈や関係については不明な点が多い。場合によっては、複数の解析技法を組合せることにより、より強力な技法に発展させられるかもしれない。あるいは、どのように組み合わせても、この点を解析しきれない、という穴があるのかもしれない。数多くの解析技法の関連を見出せるような統一的な研究が、計算複雑度理論を次のステップへ進めるためには重要と思われる。

（4）キーワード

計算量、アルゴリズム、計算量クラス、多項式時間計算可能性、P≠NP 予想、計算量下界、一方向関数、乱択計算、擬似乱数生成器、近似可能性・不可能性、確率的検査可能証明、局所性質検査

（５）国際比較

国・地域	フェーズ	現状	トレンド	各国の状況、評価の際に参考にした根拠など
日本	基礎研究	○	↗	新学術領域「計算限界解明」[1]（2012年度～2016年度）を中心に、共同研究が活発に行われるようになるとともに、重要な成果も挙がり、世界的にも注目されている。ただし、理論計算機科学全般で考えても、中国や韓国と比べ、米国で活躍している研究者の数が少ない点が気がかりである。
	応用研究・開発	◎	↗	アルゴリズム理論の大型プロジェクト[2]と上記の計算限界解明プロジェクトの緊密な連携のもとに、上記の研究から得られるアルゴリズム理論の新たな技法や基礎理論が、革新的なアルゴリズムの開発から、その実システムへの展開研究まで切れ目の無い共同研究が行われている。
米国	基礎研究	◎	↗	NSFのプロジェクトとして始まったIntractability Center[3]（2009～2014）では、ネヴァリナ賞（フィールズ賞の離散数学版）の2014年受賞者を出すような重要な研究を数多く行ってきた。それをSimons Foundationの6千万ドルの資金をもとにしてカリフォルニア大に創設されたSimons Institute[4]（2012～）が発展的に受け継ぐ形で、世界の基礎研究をリードする拠点を作ろうとしている。
	応用研究・開発	◎	→	基礎理論の研究からアルゴリズム理論まで、非常に厚みのある研究者層を持っている。また、この2つの領域で最も重要と見なされている3つの国際会議（略称：STOC，FOCS, SODA）の運営母体が米国の学会もしくは米国を拠点としている。また、Google、Microsoftなどが研究拠点を持ち、基礎研究を展開研究に結び付ける活動も活発である。（逆に、展開研究を企業に依存しすぎている点に危うさもある。）
欧州	基礎研究	○	→	基礎研究の歴史があり、世界で唯一の理論計算機科学の学会も欧州を拠点としている（European Associateion of Theoretical Computer Science[5]）。その意味で研究者層は厚い。しかしながら、計算複雑度理論に関しては、個別の研究以上のプロジェクトは現在走っていない。
	応用研究・開発	△	→	研究が個別的であるため、計算複雑度理論の研究成果が、うまく活かされる仕組みができていないように思われる。
中国	基礎研究	◎	→	ヤオ（200年チューリング賞）をプリンストン大学から清華大学に招聘し、基礎理論の強化をはかっている。実際、ヤオは、2010年にInstitute for Theoretical Computer Science（ITCS，理論計算機科学研究所）を、さらにそれを核に2013年には、Institute for Interdisciplinary Information Sciences[6]（IIIS: 学際情報科学院）という高等教育課程を構築し（現在、教員22名、博士学生60名）、優秀な人材を輩出している。ただし、ヤオに続く世代が欠けているため、今後の進展は不明確である。
	応用研究・開発	○	→	Microsoftが中国の豊富な人材を想定して、アジアの研究の拠点として、10年以上も前に研究所[7]を開設している。その意味では、展開研究への連携体制も構築されている。しかしながら、基礎理論は、清華大学のみ、その展開研究はMicrosoftのみ、といった点に少し危うさを感じる。
韓国	基礎研究	△	→	計算複雑度理論の研究はあまり活発には行われていない。ただし、アルゴリズム論など理論計算機科の中の関連分野には優秀な研究者が多い。米国で活躍している研究者もいる。
	応用研究・開発	○	→	アルゴリズム開発の分野の研究者層は比較的厚く、活躍している研究者も多い。

（註 1）フェーズ
基礎研究フェーズ ：大学・国研などでの基礎研究のレベル
応用研究・開発フェーズ ：研究・技術開発（プロトタイプの開発含む）のレベル
産業化フェーズ ：量産技術・製品展開力のレベル

（註 2）現状
※我が国の現状を基準にした相対評価ではなく、絶対評価である。
◎：他国に比べて顕著な活動・成果が見えている、 ○：ある程度の活動・成果が見えている、
△：他国に比べて顕著な活動・成果が見えていない、×：特筆すべき活動・成果が見えていない

（註 3）トレンド
↗ ：上昇傾向、 →：現状維持、 ↘：下降傾向

（６）引用資料

1) http://www.al.ics.saitama-u.ac.jp/elc/
2) JST ERATO 湊離散構造処理系プロジェクト
http://www.jst.go.jp/erato/minato/
JST ERATO 河原林巨大グラフプロジェクト
http://www.jst.go.jp/erato/kawarabayashi/
3) http://intractability.princeton.edu/
4) http://simons.berkeley.edu/
5) http://www.eatcs.org/
6) http://iiis.tsinghua.edu.cn/
7) http://research.microsoft.com/en-us/labs/asia/

３．１．５　アルゴリズム理論

（１）研究開発領域名

アルゴリズム理論

（２）研究開発領域の簡潔な説明

プログラムよりは高位の概念であるアルゴリズム（計算機の利用を前提とした効率的な「計算法」）に関する理論的研究領域。アルゴリズムの「良さ」の評価は初期には計算の速度の大きさ、つまり計算時間の少なさという観点からなされてきたが、20 世紀の終頃から大きく様変りし、様々な異った断面から評価されるようになってきた。本稿では、そのような「様々な異った断面」に関して、その現状と今後の課題を述べる。

（３）研究開発領域の詳細な説明と国内外の動向

計算機の黎明期の利用は、「手計算の機械化」に続いて「具体的問題をかなり複雑な手順で自動的に解くプログラムの実行」へと進化した。そこでは、プログラムの正確さが勿論重要であったが、正しいプログラムでも、上手なプログラムと下手なプログラムがあることが分かってきた。つまり、速いプログラムと遅いプログラムである。特に戦時中には、暗号の解読や兵站のためのオペレーションズ・リサーチをいかに高速に実行するかの研究・開発が欧米において大いに発展し、後に、Danzig のシンプレックス法や高速フーリェ変換、最短経路計算等々の画期的な計算法を生み出した。このように計算機の利用を前提とはするが、プログラムよりは高位の概念である「計算法」をアルゴリズムと呼ぶようになった。

初期のアルゴリズムの研究は、ほとんどその速度を議論するものであったと言ってよい。さらには最も都合の悪い入力に対する速度、つまり最悪の場合の速度を議論していた。まもなく、必要記憶容量、つまり領域量も評価尺度として利用されるようになった。この様な尺度の元で、画期的と言ってよい素晴らしいアルゴリズムが数多く発見されてはきたが、評価尺度自体には 1970 年代くらいまで大きな変化は無かった。しかし、1970 年代後半から 1980 年代にかけて、大きな動きが生じた。つまり様々な新しい尺度が現れてきたのである。

その理由を一言で言ってしまえば、計算機の能力の進展とともに扱う問題の範囲が一気に拡大したからである。典型的な例の一つが、経済活動である。従来人間による経験と勘の世界を経済学のモデルで論じてきたが、計算機が急速に進出してくると、アルゴリズム論のモデルによる解析、つまりどのようなアルゴリズムが、どのような理由で勝っているかを議論せざるを得なくなってしまった。そうした要求が自然に数学的に頑健な新しいモデルの提案とその利用による新たな知識の獲得につながっていった。つまり、計算機が進出する多くの場所で、その場所に適合し様々な要求に応えることができるアルゴリズム理論を構築する必要に迫られたのである。この動きは 20 世紀の終わりから 21 世紀にかけて一気に盛んになり現在も発展を続けている。例えばビッグデータでは、アルゴリズムが扱うデータ量に関する従来の仮定、つまりすべてのデータがアクセスできるという仮定を根本から考え直した上でのアルゴリズム理論が求められている。

以下ではアルゴリズム理論のこうした新しい流れに関して、その全体像を整理するとともに、重要なものをいくつか選んで、それらに対する研究の意義、現状、課題等々について述べる。アルゴリズムは、比較的新しい学問分野であるとはいえ既に数十年以上の歴史を有している。従って、多くの基本的なアイデアやテクニックが蓄積されている。上記のような新しい流れに比較的上手く対応して来られたのも正にその蓄積の賜物であって、今後の更なる発展に関しても大きな心配はないと確信している。

以下では、アルゴリズム理論のタイプに関して、いくつかのカテゴリを設けそれに沿って説明していく。

（3-1）最初のカテゴリは、あくまで正解（厳密解）を求めるアルゴリズムのための理論やモデルのカテゴリである。前述の様に、最悪時間計算量の解析がアルゴリズム理論の幕開けになった。ただ、かなり早い時点から「最悪の場合」だけ考えるのは悲観的過ぎるという指摘があり、「平均的時間計算量」や「確率的時間計算量」の研究が活発になった。前者は入力に適当な確率分布を仮定してその分布に対する平均的な計算時間のことで、直観的には、実用的に良く現れる入力に対しては速いアルゴリズムがあることを示すのが目的である。実際頂点間に枝をランダムに引くランダムグラフに対しては多くの問題（彩色問題、ハミルトン閉路問題等）が格段に易しくなることが知られている。

一方で，アルゴリズムの側に乱数を利用させることによっても（平均的な）計算時間が短縮できる数多くの例があることが分かっている。単純な例として、裏向きのトランプのカードが 100 枚あってその内 10 枚が黒だとする。黒いカードを見つけ出すには、最悪の場合 91 枚を調べなければならないが、ランダムに調べることを許せば 10 回ほどで十分である。これらは、共に領域計算量に対しても同様に重要な尺度になっている。

これら平均的計算量と確率的計算量は共に長い歴史があって、最悪計算量と共に現在でもアルゴリズム理論の核になっている。近年このカテゴリに対して、2 つの新しい流れが出てきた。

一つは「穏健指数時間アルゴリズム」である。多くの組合せ的問題は問題のサイズの増大と共に計算時間が指数的に増加してしまう。例えば、頂点数 n のグラフ問題において、単純なアルゴリズムでは計算時間がおよそ n の階乗（n!）や 2 の n 乗（2^n）になってしまう例が数多く存在する。これでは n の値が 30 程度で既に苦しくなって、実用的問題が解けるとはいえない。しかし、アルゴリズム的工夫によって、この計算時間を 2 の(n/3)乗（$2^{n/3}$）まで短縮できたらどうであろうか。n の限界が 30 くらいだった計算が 100 くらいまでのびるので実用的価値は大きい。さらに、この分野は、データ構造に強く依存した従来型のテクニックと本質的に異った様々なアルゴリズム的な工夫が役立つことも知られていて今後も安定的に発展していくと思われる。

もう一つは「パラメータ計算量モデル」である。上と同様に、n 頂点のグラフが与えられて、そこから互いに素（枝が無い）な 10 頂点を（独立集合）見付けたいとする。すべての

10 頂点の組を調べれば良いので，n^{10} の計算時間のアルゴリズムなら自明である。しかし、これも n の限界は 2 桁程度ですぐにきてしまう。しかし、工夫によって、計算量を$(2^{10})n$ に出来たらどうであろうか。2^{10}は、たった 1000 であり、今の計算機なら億レベルの n の値まで全く問題ないので、実用性という面では段違いである。パラメータ計算量の基本は、問題のサイズ n の他に別のパラメータ k を導入することである。上の問題であるなら、サイズ n のグラフのサイズ k の独立集合を見付ける問題ということができ，計算時間が 2^n の自明なアルゴリズムにかえて，計算時間が$(2^k)n$ のアルゴリズムを発見しようというのがゴールである。従来のサイズパラメータ n に関しては極めて高速（線形時間）である。

このモデルの良いところはパラメータ k として多様な選択があることである。上の k は答えのサイズであり、最も標準的ではあるが、様々な他の可能性もある。実社会の重要問題においては、パラメータ計算量が小さい k を探せることは非常に多く、たとえば Web 解析では、通常のモデルで求解困難な問題が、パラメータ計算量モデルでは効率よく解けることが知られている。多様なパラメータから適切なものを選ぶことにより、求解できる課題の範囲は格段に広がりつつあり、将来の発展の余地は大きい。

（3-2） 2 つ目のカテゴリは広い意味での近似計算である。厳密解を求めるのに時間がかかりすぎる（指数時間）時に、厳密解にできるだけ近い近似解で我慢する代わりに計算時間を画期的に改善（多項式時間）しようというのが目的である。例えば、前述した独立集合なら、最適解（厳密解）のサイズは 50 であるが、30 頂点の独立集合で我慢する。この場合の解の近似度は 50/30=1.67 であるという。この近似度を保証する（例えば近似度 2 以内）形でアルゴリズム設計をするのが近似アルゴリズムである。1970 年代に NP 完全性の概念が導入されて、多くの重要な問題が多項式時間で計算するのが不可能に違いないと考えられて以来、近似計算はアルゴリズム理論で最も人気のある分野になった。したがって、既に多くの結果が得られているが。いまだ解決されていない重要な問題（例えば平面上の巡回セールスマン問題の近似度）も相当数残っており、水面下での熾烈な競争が進んでいると言われている。

別の言葉で言えば、このカテゴリでは、要求される解がユニークではないという環境の元で、できるだけ良い解を求めることを目的にしていて、上記の近似計算モデル以外にも 2 つの興味深いモデルがある。 1 つはオンライン計算モデルである。これは、株や通貨の取引きのように、現在の行動（例えば保有の株式を売却する）の良さが将来の入力（その株価が上がる・下がる）によって左右される問題に対するアルゴリズムを議論するためのモデルである。将来の入力は原理的に分からないので推測するのが自然なアプローチではあるが、この「推測」に対する数学的評価が容易でないことは明らかである。しかし、1980 年代に、「競合比」という極めて合理的かつ数学的に扱い易い尺度が導入されて一気にメジャーな研究分野に昇格した。

オンラインアルゴリズムは、将来の入力を見ずに（原理的に見られないので）現在までの入力によって現在の行動を決定する。つまり、株取引なら現在までの株価の情報は全て使えるが、将来の情報はゼロで現在買うか売るかまたその量を決定しなければならない。その結果利得（損失）が得られるがその量 x の多いアルゴリズムが良いアルゴリズムである。その良さを評価するのが競合比で、x を「神様の利得 y」と比較（y/x）するのである。ここで神

様（オフラインアルゴリズム）は将来の入力も全て分かっているとする。つまり、n ステップの動作を考えるなら、株価 $v_1, v_2, \ldots, v_n$ が与えられて、オンラインアルゴリズムはステップ i では $v_1, \ldots, v_i$ までしか見られないが、オフラインアルゴリズムは n 個全ての価格を見て最適な売り買いが出来る。そうして得られた双方の最終的な利得を比較する。オフラインアルゴリズムの方が格段に有利なので、意味のある競合比が得られないのではないかと思うかもしれないが、多くの問題に対して、例えば 2.0 以下という予想以上に良い競合比が得られる。多くの実際の問題がこのモデルの元で議論できることも分かっている。最近は競合比に優れているアルゴリズムを実際の価格変動を使用してシミュレーションを行い、その有用性を確かめるという動きも活発になっている。

より最近（1990 年代終わり）導入された準線形時間計算モデルも大変興味深い。オンラインの場合と同様、全ての入力を見ないで（ただし、オンラインの様に時間的順序の概念はなく、レジスタの中の入力アイテムをそのインデックスを指定して見ることが出来る）線形時間よりも短い（例えば入力長 n に対してその平方根くらい）時間で答えを出さねばいけない。例えば、100 頂点のグラフに対して、そのグラフが連結かどうかの答えを 20 頂点くらい見て出さなければいけない。そのようなことが可能かと思われるが、以下の様に近似の概念を少し強めると多くの非自明な結果が得られる。答えは普通イエスかノーの二者択一を要求されるが（ある性質を満たすかどうかを問われる場合が多く、その場合は性質検査問題と呼ばれる)、イエスか「極端に」ノーの例題しか与えられないと仮定するのである。上のグラフの連結性を判定する問題では、連結でないものを連結にするのには多くの枝を追加しなければならないような例題しか入力されないと仮定するのである（逆にそうでない微妙な例題が与えられたなら答えを間違っても良い)。

この様な設定によって、上のグラフの連結性判定問題を含む多くの性質検査問題が準線系どころか入力のサイズに依存しない定数時間で解ける。ビッグデータの時代にマッチした設定であると言えよう。容易に想像できる様に、上記の入力に対するアクセス法の仮定に様々な自由度があり、それはすなわちモデルの多様性も意味する。したがって、現時点でも未だ発展途上と言っても良いくらい興味深い問題や課題に富んでいる。

ビッグデータとなれば、たとえ全ての入力が得られるとしても、記憶装置の容量をはるかに越えてしまう。その場合は、入力を記憶装置に格納された形で自由に読めるという通常の設定は現実的でない。例えば、大量のデータがネット上を流れていく場合を考えるなら、個々の入力アイテム（例えば数値）を 1 回だけリアルタイムで読むだけで、記憶装置内に取り込むこと無く計算することが要求される。これをストリーム計算と呼び、そのためのアルゴリズムがストリームアルゴリズムである。例えば、数値データの平均を求めるくらいの簡単な問題であれば、記憶装置をほとんど使わずに正しい答えがえられるかもしれないが、一般には（例えば平均ではなくて中央値なら）近似解しか得られない場合が多いのは当然である。実用上の重要性は極めて高く、近似度を上げるために様々なアルゴリズム的手法が使えることが分かってきている。

（3-3）第三のカテゴリは並列分散計算に対するアルゴリズム理論である。並列計算は高速化の切札であるから、その理論的解析も当然活発であった（主に1980年代)。特に、プロセッサ数と計算時間の積を直列計算の計算時間に一致させる並列計算の最適化に関して、多くの興味深い結果が得られた。さらには、この様な最適性をある程度保った上で、計算時間をどこまで速くできるかどうか、出来る問題の種類と出来ない問題の種類に関しても、ある程度網羅的な結果が得られた。つまり、理論的成果を得るための枠組に関しては何の不満もない状態であったが、比較的短時間で研究が下火になってしまった。その原因は並列計算機のモデルである。当時主に使われていたモデルは並列 RAM と言って、入力のサイズにまで匹敵する多数の計算機からこれも同様に大きな共有メモリへのアクセスを自由に許すというモデルである。容易に想像できる様に、このモデルを実機で実現するのは至難である。その意味では、上の様な研究は実際の並列計算とは独立の理論であって、その目的は何かという批判が強くなった。したがって、メッセージ転送を基本とした様々な現実的モデルが各種提唱され、実際の並列計算機との乖離を避ける努力が続いている。

分散計算はどうであろうか。並列計算との違いは、その目的が高速化ではなく、貧弱な通信機能を如何に上手く使って空間的に広く分布した計算機をまとめ上げることである。例えば、素なネットワークで接続された計算機群の中からリーダーを１台決めるという問題である。簡単に見えるかもしれないが、通信が非同期の１対１であったり、各計算機は全体のネットワークの構造が全く分からないといった様々な制約の元では簡単ではない。もっと基本的なデータを多くの計算機から計算機に並列的に移動させる（並列ラウティング）ことさえ限定されたネットワーク上では簡単ではない。一部の計算機が故障していたり、スリープ状態であったり、さらには悪意をもって計算を妨害するという設定もある。このように、この分野は問題と設定の範囲が極めて広範囲であり、利用するテクニックも他のアルゴリズム研究の場合と少し傾向が違うこともあってか、長期間継続して興味深い結果が現れている。例えば、最近になって、グラフの最大次数のみに依存し全体のサイズに(ほとんど）依存しない分散アルゴリズムが開発されており、Web 時代にマッチした方向と言えよう。

（3-4）第四のカテゴリは、計算よりは通信に重きをおいて、如何に通信の手順（プロトコル）を工夫して目的を達成するかを追求する。目的としては、効率を上げる（通信量を減らす）ことや漏洩する情報量を少なくすることなどがある。IT 社会は通信の方がむしろ重要であるという考えもあるくらいであるから、プロトコルの設計はアルゴリズムの設計と同等あるいはそれ以上の重要性を持っているとも考えられる。

数多くの計算機間の通信も対象にすることもあるが、通常は２台の計算機（というよりも２人のプレーヤーと呼ばれることが多い）の間の通信を対象にすることが多く、対話型計算と呼ばれることもある。典型的な問題としては、２台の計算機がそれぞれ整数の集合 X と Y をもっていて、それらの和集合に関して何かを計算したい時、何ビットの通信量が必要かを議論する問題である。これも、「何か」が平均なら易しいが中央値となると難しい（勿論 X あるいはＹの全体を相手に送ってしまえば何でも計算できるので、それよりも本質的に少ない通信量で実現することが大前提である)。

別の目的としては、情報の漏洩を防ぐことがある。2 人のプレーヤーの内、A が自分の偉大さ（計算能力の高さ）をもう 1 人のプレーヤーB に納得させるという問題が有名である（少し奇異な感じがするかもしれないが、実用上の応用もある）。例えば、素因数分解は計算困難な問題として良く知られている。万能の A は、A と B に共通に与えられた整数 N の素因数分解の答え（つまり N=PQ となる整数 P と Q）を知っている。そのことを B に納得させるには P と Q を送ってしまえば良い（それらを掛けて検算することは簡単であるから B にも出来る）が、P と Q を秘密にしておきたいという設定である。この様に、その情報を公開してしまえば簡単であるが、公開しないで何かの目的を達成したいのである。B の計算力の弱さ（多項式計算しか出来ない）と乱数を上手く利用して、P と Q の情報を全く漏らすこと無く目的が実現できることが 1980 年代に初めて証明された時は大きな話題となった。

（3-5）第五のカテゴリはメカニズムデザインやアルゴリズム的ゲーム理論である。メカニズムデザインとは、アルゴリズムだけでなく、あるシステムや枠組を総合的に設計することを目的にする。例えば、オークションである。オークションの目的は勿論オークショニア（競売人）の利得を最大にすることである。ただし、競売の対象の品物は様々であり、例えば希少品 1 点の場合もあれば、限りなくコピー出来る音楽・動画ソフトなどの場合もある。さらには競売人のアルゴリズム（勿論出来るだけ多くの利益を得られるように設計する）は公開されているという前提なので、ビッダー（入札者）に悪用されない様に注意しないといけない。例えば、映画ソフトの競売では入札価格（の集合）から各入札者に売る・売らない、売るならその価格（入札額以下）を提示するわけである。簡単に考えると、全員に入札価格そのもので売るという単純なアルゴリズムが最も売上げが多くなりそうに見えるが、提示価格が入札価格に等しいということが入札者に知られれば、皆極めて低い価格で入札するに決まっている。したがって、競売人はそのアルゴリズムを設計するときに、入札者が「正直に」自分が思っている品物に対する価値に対応する値段で入札するように（そうしないと入札者が損をする）誘導できなければならない。これを「正直さを誘導するアルゴリズム」と呼び、オークションアルゴリズムの基本である。Web 上では、アマチュアのみならずプロが参加する様々な種類のオークションが出現しており、また理論的結果の実用性が比較的高いとも言われているので、将来的にも有望である。

ゲーム理論は長い歴史があるが、様々な場面でアルゴリズム理論が重宝される。例えば、上述のオークションも競売人と入札者の間のゲームであるが、正直さを誘導するアルゴリズムはその様な要求の無いアルゴリズムに比べれば（同じ入力に対しては）もちろん利益が少ない。したがって、アルゴリズムの良さを評価するのに 2 つの利益の比をとることが自然な方法である。これは正に競合比の考え方である。また、あるゲームでナッシュ均衡があるのか無いのか、あるとすればどんな値で均衡するのかは是非知りたいことである。しかし、多くの場合にその計算は困難であることが分かっており、これもアルゴリズム理論のお蔭である。

中央制御が存在しない分散的な環境でのアルゴリズムということで、広い意味では分散アルゴリズムの一種かもしれない。しかし、伝統的な分散アルゴリズムでは全ての参加者が同

じアルゴリズムを使って何かを全体的に実現することを目的にしており、敵対的なゲームの要素がない。この意味で新鮮な分野であり、今後も重要性を増すと考えられる。

（3-6）最後のカテゴリは未だ実現出来ていない計算メカニズムに対するアルゴリズム理論であり、量子アルゴリズムがその代表である。量子コンピュータによって多項式時間で高い確率で解ける量子アルゴリズムが存在する決定問題の複雑性クラスを BQP （Bounded-error Quantum Polynomial time）と呼ぶが，素因数分解がこのクラスに属することが証明されたときは極めて大きな反響があった。素因数分解の通常の計算機上での困難さは、証明されていないとはいえ、強く信じられていて、実際その計算困難性が現在広く利用されている暗号系の原点になっているからである。この結果が出たときは、量子メカニズムが他の計算困難問題に対しても有効ではないかという予想を生み、多くの期待を抱かせた。しかし、それはかなり困難で、逆に量子計算の限界に関するいくつかの興味深い結果が証明されるに至った。

同時期に発見され同様に重要視された探索問題に対する準線系時間量子アルゴリズムは、極めて広い応用があることが分かったが、元々が古典計算機でも高速に解ける問題を対象としていることから、その実用上のインパクトは限定されてくる。最近は量子アニーリング（焼き鈍し）法といった、理論よりもむしろ実験やヒューリスティックを重視した研究も現れてきており、今後の動向に関しては注意深く見守る必要がある。

生物計算も一時期脚光を浴び、最近は計算という観点からよりもむしろ生物学的側面から実験を重視した研究が盛んである。興味深い面はいくつかあるが、アルゴリズム理論という枠からは距離がある。

最後に番外として、学習と暗号に関するアルゴリズム理論を挙げる。これに関しては、それ自体の領域があるので、そこで詳しく議論されるはずであるが、「実用に直結する計算機の理論」という観点からは極めて重要である。例えば、最近急速にその人気を高めているビットコインであるが、その中にはアルゴリズム理論の成果が宝石の様にちりばめられていると言ってよい。

（４）科学技術的・政策的課題

最初の課題が人材確保である。アルゴリズム理論全体では、我が国のレベルは決して低くない。特に、その骨格をなす従来型の時間計算量や近似理論の分野では、世界的研究者を数多く排出している。また、最近では、学部で数学や物理を専門とした優秀な学生を引きつける分野としての力も備わってきている。歴史的に見て、我が国の大学の情報関連学科が主に工学部から派生していること（諸外国では理学部数学科から派生している場合の方が多い）はメリットとデメリットがあるが、デメリットの一つがこのような動きによって確実に解消されつつある。さらに、この動きを加速する手段が高校教育であると言われている。例えば、スイスでは最近小学校高学年から情報教育がシステム化され、中学・高校向けの教科書が他の従来科目と同様の規模で出回っている。その内容はやはりアルゴリズム中心であって、実際この教育改革には一人の著名なアルゴリズム理論の研究者の 10 年以上にわたる献身的努

力が深くかかわっている。

もう一つの重要な政策的課題は研究の新しい流れをタイムリーにプロモートすることである。先に述べた様に、アルゴリズム理論は他の分野に比べて研究テーマの機動性が高い。前節で述べた様に、20 世紀の終わりごろからメカニズムデザイン、アルゴリズム的ゲーム理論、量子計算等々多くの新しい潮流が出たが、その初期における我が国の研究者の貢献は残念ながら大きいとはいえない。原因はいくつかあると思うが、その一つが研究者間の国際的交流の問題である。インターネット時代ということで、情報そのものはどこにいても得られるが、やはり研究者同士の生の交流は何者にもかえがたい。実際、中国・インドの研究者はこの面では非常に積極的であって、その効果は確実に出ていると言われている。もちろん、我が国の研究者の国際性は近年著しく改善されてはいるが、一層の努力が必要と考える。また、長い目でみれば、我が国の留学生数が減少傾向にあることなども改善が必要である。

異分野交流も重要な政策的課題である。例えば、新しい流れという意味では「今の旬」はビッグデータである。アルゴリズム理論の研究者が今まで頼りにしてきたのは離散的・組合せ的な数学（考え方）である。しかし、ビッグデータに対応するにはそれでは不十分で、連続系の数学やそうしたデータを必要とする学問分野との強い融合が必要である。一つの方向として考えられるのは、こうした数学や、経済学、生物学等を専門とする研究者をアルゴリズム理論の世界に引き込むことである。実は、この動きは既に具体化しつつあって、大学における情報関連学科の人気を高める切り札の一つであると言われている。アルゴリズム理論が現実の社会に貢献できる場面として最も期待されている（あるいは既にかなり貢献している）のが金融や経済の分野である。2013 年のノーベル経済学賞は、アルゴリズム研究者が受賞しているし、この流れは今後ますます強くなっていくと思われる.

（５）注目動向（新たな知見や新技術の創出、大規模プロジェクトの動向など）

理論研究はアカデミアという先入観があるが、アルゴリズム理論は必ずしもそうではない。豊富な資金力を有する、その時々の IT トップ企業（20 世紀は IBM や AT&T、21 世紀に入って Google やマイクロソフト等）は常に強力な研究組織を有しており、そこでは多くのアルゴリズム研究者が採用され活躍している。このことは、２つの面で重要である。一つは、アカデミアと産業界の研究部門との交流である。たとえ理論研究者の最終的な目標がアカデミアであったとしても、途中段階で経済的に安心感のあるポジションを供給してくれる企業の存在は何者にもかえがたいし、長期的にはそれが人材確保の切札になる。また、企業側のメリットも大きい。その理由は、どんなことであれ新しい動きは抽象的なものからスタートすることが多いし、その意味でアルゴリズムやメカニズムデザインの研究者が新しいビジネスモデルの端著を開くことが期待されているからである。具体的事例（Google のスポンサーサイト等）も豊富である。

現在、アルゴリズム理論を主要テーマにした国家主導の大型プロジェクトが、我が国や米国で複数走っている。最近は、プロジェクトの構成員は海外にも広く門戸が開かれているが、この傾向はさらに押し進めてよい。こうした大規模プロジェクトが予算に見合う成果を上げ

ることができるためには、国外の優秀な人材を活用するべきだからである。アルゴリズム理論は他の分野からの人材を取り込むことが必要であると述べたが、逆もまた真である。幅広い人材を集めた大規模プロジェクトでは既に普通のことになっていて、アルゴリズム理論の研究者が従来の狭い理論コミュニティに閉じることなく、周辺部の応用的要素の強いコミュニティ（AI やネットワーク）に活発に進出している例が目に付く。これは、大変良い流れであると言える。

（６）キーワード

アルゴリズムの設計と解析、計算量理論、近似アルゴリズム、確率アルゴリズム、オンラインアルゴリズム、並列アルゴリズム、分散アルゴリズム、アルゴリズム的ゲーム理論、メカニズムデザイン、量子アルゴリズム

（７）国際比較

個々の事例に関しては既に述べた。総合的な順位としては、米国（＋カナダ)、欧州（＋イスラエル）がほぼ拮抗し、次のグループとして、中国（＋香港)、インド、日本、オーストラリア、韓国、シンガポールといった感じである。

３．１．６　最適化理論

（１）研究開発領域名

最適化理論

（２）研究開発領域の簡潔な説明

与えられた制約の下で与えられた関数を最大化、または、最小化するための手法を開発する。現実世界で生じる多くの問題がこの枠組みで記述できるが、対象とする問題に応じて、連続最適化と離散最適化に大別され、また、必要とする手法に応じて、そのそれぞれが大域最適化と局所最適化に分類される。効率性（計算複雑度）や精度に対する数学的保証を与えることが主な研究課題である。

（３）研究開発領域の詳細な説明と国内外の動向

与えられた制約の下で与えられた関数を最大化、または、最小化するための計算手法を数理的に究明する分野が最適化理論である。現実世界で生じる多くの問題が最適化理論の枠組みで記述される。例えば、医療施設の配置、船舶の航路設計、構造物の部材組合せ、データベースにおける構造発見、金融商品の価格付け等、枚挙に暇がない。このような問題を最適化問題として記述するための方法論が最適化モデリングである。しかし、最適化モデリング自体が研究対象となってきてはいない。実際、現状の最適化理論は、最適化モデリングによって記述された具体的な問題、あるいは、問題群に対する研究に終始している。

最適化理論は決定すべき変数が連続である場合と離散である場合に大きく二分され、前者は連続最適化、後者は離散最適化と呼ばれる。後で述べるように、この分類は現代的な見地から重要性を失いつつあり、従来の連続最適化と離散最適化を渡り歩ける人材が必要とされてきている。

理想を言うと、最適化理論が目指す計算手法は以下の 2 つを兼ね備えているとよい。1 つは「効率性」であり、これは、計算自体が高速に終了すること、そして、それが数学的に保証されていることである。もう 1 つは「最適性」であり、これは、計算が終了したときに最適解が必ず得られること、そして、それが数学的に保証されていることである。連続最適化においては、浮動小数点数を用いて計算を行うため、最適性を厳密に保証することは難しい。そのため、最適解と任意に近い許容解を必ず得ることとして、最適性を述べ直すことが多い。効率性は、現在の最適化理論において「多項式時間アルゴリズムの存在性」として定式化される。計算複雑度理論において P≠NP 予想が未解決であるという現状を鑑みると、多項式時間アルゴリズムの非存在性を証明することは現状の理論において難しいと考えざるを得ないが、その代わりに、問題が NP 困難であることを証明して、多項式時間アルゴリズムが存在しないことの傍証とすることが多い。一方、多項式時間アルゴリズムの存在性を証明するためには、そのようなアルゴリズムを 1 つ設計し、実際に多項式時間アルゴリズムで計算が終了することを証明すればよい。

以下、連続最適化と離散最適化に分けて説明を行う。

【連続最適化における動向】

連続最適化における最も基礎的な研究対象は線形制約のもとで線形関数を最適化する線形

計画問題である。線形計画法に対して、最適性保証を持つ多項式時間アルゴリズムの存在性は昔から未解決問題として認識されていた。それを解決したのは Khachiyan であり、論文が出版されたのは 1979 年のことであった。その後、効率性を大幅に向上させた内点法と呼ばれるアルゴリズムが設計され、現在では、多くの最適化ソルバーに実装されている。その開発に日本人研究者が多く携わり、1992 年に日本人を含む研究グループが ORSA（米国オペレーションズ・リサーチ学会）の Lanchester 賞を受賞した。

21 世紀初頭までの連続最適化研究は以下のように大別できる。(C1) 線形計画法の拡張としての錐線形計画問題に関する研究、(C2) 非線形最適化問題の局所的最適化に関する研究、(C3) 非線形最適化問題の大域的最適化に関する研究。C1 は、1990 年代から興った、線形計画問題に対する内点法を他の凸最適化問題に適用しようとする研究潮流を指し、実際、二次錐計画問題、半正定値計画問題といった問題に拡張できると分かった。究極的には、「Euclid 的 Jordan 代数における対称錐上の線形最適化」にまで拡張可能であることが今では分かっており、現在は、他の錐上の線形最適化問題に対する研究に主眼が移ってきている。その例として、最近注目されているのが、完全正値錐上の線形最適化、共正値錐上の線形最適化である。しかし、この 2 問題は NP 困難であることが知られているため、次に述べる大域的最適化の手法を用いて解く手法が研究されている。C2 は、大域的な最適性は保証しないが局所最適解を高速に得るアルゴリズムに関する研究を指す。多くのアルゴリズムが提案され、理論の成熟を見ている。そのため、現在では、均衡問題や相補性問題といった、最適化問題を含む、さらに広い枠組みにおいて、同様のアルゴリズムを設計・解析する研究が行われている。C3 は効率性を保証しないが、最適性を保証するアルゴリズムに関する研究を指す。分枝限定法のように問題を分割することによって上手に解く手法の他に、緩和法という巧妙な方法が研究されている。緩和法では、問題の制約や目的関数を緩めることで解きやすい問題に変換し、それを解く。そのとき、緩める度合いを調整することで、大域的最適性を保証するのである。緩和法による大域的最適性の保証に関する方法論が急速に進歩しており、特に、変換後の解きやすい問題が C1 で取り扱った対称錐上の線形最適化問題となる場合が現在活発に研究されている話題である。

【離散最適化における動向】

離散最適化における最も基礎的な研究対象は、線形制約と整数制約のもとで線形関数を最適化するもので、通常これを整数計画問題と呼んでいる。整数計画問題は NP 困難であるため、効率性と最適性を共に満たすアルゴリズムの存在を期待できない。そのため、最適性は保証するが、効率性を犠牲にするアルゴリズムの設計と解析が昔から行われてきた。基本的な設計技法は分枝限定法と切除平面法である。また、それらに緩和法や分解法を組み合わせて用いることで、最先端の整数計画アルゴリズムが設計されている。理論的に解析されている側面は最適性がほとんどで、効率性に関する理論的解析はほとんどされていない。

一般の整数計画問題は NP 困難であるが、特殊な構造を持つ整数計画問題は多項式時間で解けることが知られている。その典型例はネットワーク流に関する問題で、特に、最大流問題と最小費用流問題である。これらは、「係数行列の完全ユニモジュラ性」という視点で一般化され、整多面体に対する研究潮流を作った。その一方で、最大マッチング問題や最小全域木問題のように、係数行列の完全ユニモジュラ性がないにも関わらず、効率性と最適性を

兼ね備えたアルゴリズムが設計できる場合も知られている。これらを包括する枠組みとして、マトロイドや劣モジュラ関数といった離散構造を経て、離散凸解析が日本発の理論として発展している。離散凸解析の射程は線形最適化だけではなく、凸最適化も含んでおり、その中で重要な役割を果たす劣モジュラ関数最小化多項式時間アルゴリズムに関する研究に対して、日本人を含む研究グループが Fulkerson 賞を 2003 年に受賞した。

21 世紀初頭までの離散最適化研究のテーマは、今述べたような（D1）一般の整数計画問題に対する最適性保証を持ったアルゴリズムの設計と解析、(D2) ネットワーク流問題に代表される、特殊な構造を持った整数計画問題に対するアルゴリズムの開発とその整多面体理論、離散凸解析への展開、の他に、(D3) 理論的精度保証を持つ近似アルゴリズムの設計と解析、がある。これは、1990 年代から急速に進展を見せたテーマであり、NP 困難である組合せ最適化問題に対して、最適性は犠牲にするが、その逸脱度が抑えられるようなアルゴリズムを設計することを目標としている。特に、急速な進展を見せた背景には、アルゴリズム設計において線形計画法や半正定値計画法という連続最適化の知見を多く用いる方法論が確立したことが挙げられる。近似アルゴリズム設計には、その困難性を導く確率的検査可能証明（PCP: probabilistically checkable proof）定理と呼ばれる計算複雑度理論における大定理が同時期に証明されたこともあって、多くの研究者が魅せられた。また、連続最適化における局所最適化 C2 のように、(D4) 離散最適化における局所探索とその変種が研究されている。その効率性に関して、通常の多項式時間アルゴリズムに関する NP 完全性の理論の類似として、多項式時間局所探索アルゴリズムに関する PLS 完全性の理論が存在する。実際、多くの問題に対する自然な局所最適解探索が PLS 完全であるという結果が示されている。その一方で、アニーリング（焼き鈍し）法と呼ばれる局所探索法の変種に対する漸近的大域収束性のような肯定的理論保証も得られている。しかし、局所探索法の変種であるメタヒューリスティクスや進化的手法については、その理論的解析があまり進んでいないのが現状である。

ここで挙げた事項以外にも、不確実性を考慮した確率計画法やロバスト最適化、入力が漸次的に与えられるオンライン最適化、多主体が関わるゲーム理論や多目的最適化、変数の次元が無限である無限計画や半無限計画や変分問題、より複雑な制約を扱う均衡制約付き数理計画や微分方程式制約付き数理計画など、広範な問題を最適化理論は取り扱っている。

【最適化理論に関わる学術団体】

日本においては、日本オペレーションズ・リサーチ学会に多くの最適化理論研究者が所属している。年に二度行われる研究発表会以外にも、数理計画研究部会（通称 RAMP）が年に一度シンポジウムを主催し、最適化の理論と応用に関する最先端の研究成果について情報交換を行っている。また、情報処理学会のアルゴリズム研究会、数理モデル化と問題解決研究会、電子情報通信学会のコンピュテーション研究会、日本応用数理学会においても、最適化の理論と応用に関する発表が行われている。海外においては、全世界的な学術団体として Mathematical Optimization Society（通称 MOS、旧名 Mathematical Programming Society）が存在し、3 年に一度、International Symposium on Mathematical Programming (ISMP) を開催している。MOS は、ISMP が開催されない年に、連続最適化に関するシン

ポジウムである ICCOPT や、離散最適化に関するシンポジウムである IPCO を開催している。また、米国の SIAM や INFORMS をはじめ、各国の応用数学およびオペレーションズ・リサーチに関わる団体に最適化理論の研究者が多く所属している。その他にも、特に離散最適化については顕著であるが、理論計算機科学に関わる学術団体に最適化理論の研究者が所属していることも多い。具体的には、米国の ACM、IEEE、欧州の EATCS などである。そのため、理論計算機科学に関わる国際会議である STOC、FOCS、SODA、ICALP、ESA などでも最適化理論に関する研究が発表されており、それらの国際会議に参加する日本人研究者も少なくない。

（４）科学技術的・政策的課題

従来のような「連続最適化と離散最適化」という大別が意味を成さなくなってきているという現状があるが、連続最適化と離散最適化の双方が高度な専門性を必要とするために、その 2 分野を渡り歩ける人材の育成も難しくなっている。そのような最適化理論の専門家を育成するためには、専門の教育プログラムや学科を設立するということも考慮する必要がある。実際、米国カーネギーメロン大学では、「Algorithms, Combinatorics, and Optimization Program」という研究教育プログラムが 1990 年代から学部を横断する形で進行しており、多数の優秀な研究者を輩出している。また、カナダのウォータールー大学には、「Department of Combinatorics & Optimization」が 1960 年代から存在し、最適化理論の研究における世界的拠点となっている。また、最適化理論の産業応用が日本において欧米ほど広く行われていないという現状がある。最適化理論の重要性と有用性を広報し、最適化理論の社会的認知度を向上させることも、最適化研究の底上げに資するであろう。

（５）注目動向（新たな知見や新技術の創出、大規模プロジェクトの動向など）

先にも述べた通り、連続最適化と離散最適化の間の垣根が低くなり、最先端研究では、両者を融合した視点が重要であるという認識が醸成されてきている。例えば、線形方程式の解の中で、非零成分の数が最も少ないものを発見したいという要求が、圧縮センシングなどの応用から生まれている。これは非零成分数の最小化であるが、非零成分の数は離散的であり、NP 困難な問題として知られている。そのため、基底追跡と呼ばれる発見的解法が使われているが、そこでは、非零成分数の最小化をする代わりに、ℓ_1ノルムの最小化を行う。このように変更した問題は線形計画問題として定式化することができ、効率よく解ける。最近の研究では、その係数行列がある種のランダム行列である場合には、このℓ_1ノルム最小化によって、非零成分数最小化問題が高確率で解けることが証明されている。また、同様な設定として、部分的に成分が欠損している行列を補完する際に、結果として生成される行列の階数を最小化したいという問題がある。これも階数最小化という離散最適化問題であるが、行列の核ノルムを用いた発見的解法では半正定値計画法が使われ、現れる問題の係数行列がある種のランダム行列である場合には、これによって階数最小化問題が高確率で解けることが分かっている。その証明では、ランダム行列の理論など、高度な線形代数の知見が使われている。

また、二次錐計画問題や半正定値計画問題といった連続最適化問題が効率よく解け、ソフトウェアも整備されてきていることを背景として、それらの問題の変数を整数変数とした離散版に関する研究も増えてきている。特に、整数二次錐計画問題に対するアルゴリズムは

CPLEX や Gurobi Optimizer といった有名な最適化ソフトウェアに実装され、実際に使われるようになっている。理論面では、切除平面法や緩和法に関する新たな方法論が提案されてきている。これは C1 と D1 の融合と見ることができる。

一方、C1 と D3 の融合として、精度保証を持つ近似アルゴリズム設計において半正定値計画法を用いる研究潮流が精緻化されてきている。特に、Lasserre 階層や二乗和最適化の枠組みを用いて精度保証の上界を改善するという流れとその整数性ギャップを通して精度保証の下界を導出するという流れがあり、計算複雑性理論における一意ゲーム予想（2014 年のネヴァンリンナ賞）との関連から、多くの研究者が活発に議論している。

最適化問題を定式化する際に、変数の数（次元）を大きくすることでそのサイズを小さくすることができる。これは拡張定式化と呼ばれる概念で、そのサイズに対する研究が現在盛んに行われている。特に、巡回セールスマン問題を線形計画問題として定式化する際のサイズが必ず指数関数的になってしまうことが 2012 年に証明され、完全マッチング問題に対しても同様な結果が 2014 年に証明された。今後は、線形計画問題としての定式化以外に、他の錐線形計画問題、特に、半正定値計画問題としての定式化に関するサイズが研究対象となっていくだろう。

線形計画問題は多項式時間で解けるが、強多項式時間で解けるか？（すなわち、変数の数と制約の数に関する多項式時間で解けるか？）という問題は未解決であり、大きな課題であると認識されている。そのようなアルゴリズムの有力候補は単体法であり、その動きはピボット規則によって定められる。多くのピボット規則に対して単体法が多項式時間アルゴリズムではないことが知られていたが、最近、乱数を用いたピボット規則や探索の履歴を使用するピボット規則に対しても指数時間下界が証明された。

一般化最大流問題と呼ばれるネットワーク最適化問題があり、それは線形計画問題としてモデル化できるため、多項式時間で解けることは知られていたが、それに依らず、組合せ的な手法によって多項式時間（特に，強多項式時間）で解けるかどうかは知られていなかった。しかし、2014 年にこの問題は肯定的に解決された。

最適化理論における効率性は「最悪の場合」に対する効率性として解析されることが多い。しかし、線形計画問題に対する単体法は実際に計算機上では高速に動作することが確認されている。そのような理論と実践のギャップに対する認識は昔から存在し、確率的解析などが提案されてきている。その中で、最近注目されているのが「平滑化解析」である。これは最悪の場合の摂動を考えて解析をする手法であり、これにより、単体法をはじめ、最悪の場合に指数時間かかってしまうアルゴリズムが平滑化解析の意味で多項式時間アルゴリズムとなることが分かってきた。

進化的計算に対する理論的解析の方法論が開発されてきている。特に、ドリフト解析という新しい手法により、厳密な確率的解析が可能になってきている。また、アニーリング（焼き鈍し）法の量子計算版と言える量子アニーリング法が日本発の技術として広く知られている。これに関して、カナダの D-Wave Systems 社が物理的実装を 2011 年に初めて発売し、注目を浴びている。

また、最適化理論は、機械学習、データマイニング、符号理論、情報理論、暗号理論、代数幾何学など、周辺の数理科学分野を巻き込んで新たな展開を見せている。

（６）キーワード

連続最適化、線形計画法、錐線形計画法、半正定値計画法、二次錐計画法、内点法、単体法、非線形計画法、多項式時間アルゴリズム、NP 完全性、NP 困難性、離散最適化、組合せ最適化、整数計画法、分枝限定法、切除平面法、離散凸解析、劣モジュラ関数、整多面体、ネットワーク流、近似アルゴリズム、精度保証、局所探索法、PLS 完全性、基底追跡、拡張定式化、平滑化解析

（７）国際比較

国・地域	フェーズ	現状	トレンド	各国の状況、評価の際に参考にした根拠など
日本	基礎研究	◎	↗	連続最適化、離散最適化ともに基礎的な成果を着実に挙げている。特に、若い世代の成果が顕著であり、世界的に見ても層が厚い。
	応用研究・開発	○	→	半正定値計画法に対する内点法や基礎的なグラフアルゴリズムの並列計算など、特筆すべき成果がある。
	産業化	△	↗	最適化理論の認知度の低さゆえか、企業において最適化理論に関する際立った活動が見えなかったが、最近、機械学習との関連から最適化理論に対する認識に変化が生まれてきたように思える。
米国	基礎研究	◎	↗	MIT、カーネギーメロン大学、ジョージア工科大学など、最適化理論の拠点が多く、理論的成果の質も高い。
	応用研究・開発	◎	→	最適化理論の成果を活かした商用ソフトウェアとして、CPLEXやGurobi Optimizerが開発され、世界の標準として用いられている。
	産業化	○	→	企業研究所においても、最適化理論を専門とする研究者が多く在籍し、重要な成果を挙げている。しかし、最近、MicrosoftがSilicon Valleyの研究所を閉鎖したことを受け、今後、下降傾向に転じる可能性がある。
欧州	基礎研究	◎	↗	ベルリン、ブタペスト、アムステルダム、グルノーブルなど、各国に最適化理論の拠点があり、顕著な成果を挙げている。
	応用研究・開発	◎	↗	ドイツやフランスを中心として、「アルゴリズム工学」に関するプロジェクトが1990年代から興り、理論と実践のギャップをつなぐ研究が新たな潮流となっている。
	産業化	◎	→	ドイツなどでは、産業や公共政策にて最適化理論の成果を用いている事例が数多く存在する。
中国	基礎研究	△	→	中国出身の研究者が欧米の大学および研究所に所属し、顕著な成果を挙げているが、国内の動向としては際立ったものが見えてこない。
	応用研究・開発	×	→	際立った活動が見えてこない。
	産業化	×	→	際立った活動が見えてこない。
韓国	基礎研究	△	→	離散数学分野ではFulkerson賞など顕著な成果があるが、最適化理論における大きな潮流は生み出されていない。
	応用研究・開発	×	→	際立った活動が見えてこない。
	産業化	×	→	際立った活動が見えてこない。

（註 1）フェーズ
基礎研究フェーズ ：大学・国研などでの基礎研究のレベル
応用研究・開発フェーズ ：研究・技術開発（プロトタイプの開発含む）のレベル
産業化フェーズ ：量産技術・製品展開力のレベル

（註 2）現状
※我が国の現状を基準にした相対評価ではなく、絶対評価である。
◎：他国に比べて顕著な活動・成果が見えている、 ○：ある程度の活動・成果が見えている、
△：他国に比べて顕著な活動・成果が見えていない、×：特筆すべき活動・成果が見えていない

（註 3）トレンド
↗ ：上昇傾向、 →：現状維持、 ↘：下降傾向

３．１．７　プログラム基礎理論

（１）研究開発領域名

プログラム基礎理論

（２）研究開発領域の簡潔な説明

高品質ソフトウェアの構築のための基礎となる理論。プログラムの意味や動作を数学的に厳密に定義し、さらにそれに基づいてプログラミング言語の設計、コンパイラの作成、プログラムの検証・変換・合成などに関する研究開発を行う。高信頼、高性能なソフトウェアの構築のために不可欠な研究開発領域である。

（３）研究開発領域の詳細な説明と国内外の動向

本領域は、アルゴリズム理論とならんで理論計算機科学の二大分野の一つに位置付けられ、「計算とは何か」という根源的な問いに関わる、学問的に深淵な分野であると同時に、技術的には高品質ソフトウェア、特にソフトウェア（ひいてはそれによって制御されている現代の社会システム）の信頼性に直接関わる、社会的に極めて重要な研究開発領域である。プログラム基礎理論のルーツは電子計算機が作られる以前の 20 世紀前半のチャーチによるラムダ計算の理論やゲーデルの不完全性定理などにあり、ラムダ計算は関数型言語に代表される近代的なプログラミング言語の基礎となる計算体系を与え、不完全性定理は、ソフトウェア検証などの形式手法の理論的な限界を与えている。また、不完全性定理に限らず、本領域は論理学と深いつながりがあり、カリーハワード同型に代表されるように、プログラムと証明の間に深い対応関係があることが知られ、しばしば相互に影響を与えあって発展している学問分野である。

プログラム基礎理論の初期においてのホットなトピックは、プログラムに対する数学的な意味を与え、プログラミング言語の設計の指針とすること、また効率のよいコードを生成するためのコンパイラの構成法の確立であった。プログラム意味論およびコンパイラ構成法がある程度確立し、計算機の性能が向上した今日においては、主要なトピックはソフトウェアの信頼性や生産性の向上にシフトしてきており、特にプログラム検証（またはソフトウェアを含むシステム検証）の理論・技術が盛んに研究されている。この背景には、現代社会で交通システム、発電所、金融システム、電子政府など様々な社会インフラがコンピュータによって制御されるようになっており、ソフトウェアの欠陥が重大な問題を起こしうる（また実際に起こしている）ことがある。また、検証技術を応用したプログラム合成技術の研究も盛んになりつつある。

現在研究されているプログラム検証手法は２つに大別され、一つは定理証明支援システムを用いて検証する手法であり、これは膨大な人的コストと専門知識が必要ではあるが、詳細な仕様を検証可能であるため、安全性が極めて重要なソフトウェア（例えば原子力発電所の制御ソフトウェア、飛行機や自動車のエンジン制御ソフトウェア）の検証に有効である。この手法およびそれに用いる定理証明支援システムの研究は欧州を中心に近年盛んに研究され、C 言語のサブセットのコンパイラの検証（フランス）、オペレーティングシステムのカーネルの検証（オーストラリア）などが行われている。定理証明支援システムとして現在最も広まっているのはフランスで開発された Coq であるが、日本では独立行政法人産業技術総合

研究所（産総研）のグループがスウェーデンのグループと共同で Agda の開発に携わっている。また、北陸先端科学技術大学院大学(JAIST)のグループが CafeObj と呼ばれる代数的仕様に基づく検証ツールの開発を行っている。

もう一つの検証手法は、モデル検査や、型推論・抽象解釈などのプログラム解析技法、自動定理証明器などを用いて全自動または半自動で検証する手法である。これは定理証明支援システムを使う手法と比べ、ほとんど人の手を必要としないという点で大きな強みがある。中でもモデル検査は、1970 年代に提唱された技術だが、近年では広くシステム検証に用いられるようになってきており、2006 年には提唱者の Edmund Clarke（エドムンド・クラーク）らが「コンピュータサイエンスにおけるノーベル賞」とされるチューリング賞を受賞している。また、近年では自動定理証明器の進歩もあいまって、検証可能なプログラムのサイズも増大している。本分野においても欧米がリードし、デバイスドライバ用の自動検証器などが開発されている。日本ではこの分野で長らく遅れをとっていたが、最近になって東大のグループがモデル検査の一般化である高階モデル検査についてのブレークスルーを達成し、関数型プログラムの全自動検証ツールを世界に先駆けて構築するなど、高レベルプログラムの自動検証手法の研究の一大拠点になりつつある。

本研究開発全体についての世界的な動向としては、欧州は伝統的に基礎理論に強く、上記の定理証明支援システムやモデル検査などの発展に基礎的な貢献をしている。それに対し、米国は基礎理論を応用し、産業化につなげる能力に長けており、上記のデバイスドライバ用の自動検証ツールの構築、システム検証を専門に請け負うベンチャー企業の創出などの実績がある。日本では、欧州同様、伝統的にプログラム意味論など基礎理論に強く、東大や京大数理解析研究所などに拠点がある。一方、プログラム検証などへの応用・産業化という点では遅れをとってきた。アジアの他の国では、韓国・シンガポールは最近ではプログラム検証などへの応用分野で一定の成果を挙げている一方、基礎理論はあまり強くない。中国はこの分野では全般的に遅れている。

（4）科学技術的・政策的課題

ソフトウェアの信頼性向上は、社会基盤全体の信頼性向上に直結する待ったなしの課題であり、本研究開発分野において、基礎理論から応用、産業化までが有機的に結合することが極めて重要である。

日本では、米国のマイクロソフト、グーグルのように基盤ソフトウェアの開発に直接携わる大企業が少ないこともあり、基礎理論の強みを応用、産業化に十分に生かせていない。このままでは日本のソフトウェア産業の衰退のみならず、自動車などソフトウェアの比重が増している産業への影響も懸念され、産学官を連携させる仕組みの構築が急務である。また、基礎理論に関しても個人ベースの研究が多く、プログラム検証など喫緊の大きなテーマを旗頭に知を結集するための大きな予算措置が必要であると思われる。産学の連携の試みとしては、産総研に平成 16 年から 6 年間にわたってシステム検証研究センターが設置され、一定の成果は挙げたが、学問的に新しい技術の創出や産業化という観点からは道半ばで終わっており、この種の試みをより大きな規模で継続して行うことが重要であると考えられる。また、複数の大学に散らばっている基礎理論研究者や企業の研究所・技術者の連携を促進するため、一センターの設置よりも国家規模のプロジェクトの創出が望ましい。

（５）注目動向（新たな知見や新技術の創出、大規模プロジェクトの動向など）

プログラム検証に関しては項目（３）で触れたように、大別して(i)定理証明支援器を用いた人手を介在した検証、(ii)モデル検査や型、抽象解釈などのプログラム解析技術、自動定理証明器などを駆使した全自動または半自動検証の研究にわけられる。前者については、フランスで開発された定理証明支援器 Coq が広く使われるようになっており、C のサブセットのコンパイラが検証されるなど、検証対象のシステムの規模も徐々に大きくなっている。後者については、2000 年以降、分離論理(separation logic)、高階モデル検査などのブレークスルーとなる理論、自動定理証明器（SMT ソルバ）の大幅な進歩が牽引役となって自動検証分野全体が大きな進歩を遂げている。自動検証の成果は徐々に産業化につながっており、マイクロソフトなどの大手企業の他に、米 Coverity 社、Facebook に買収された英 Monoidics 社など検証を専門に請け負うベンチャー会社が立ち上がっている。また、フランスの研究グループが Airbus 社と飛行機のエンジン制御プログラムの検証を行うなどの取り組みもある。

プログラム理論のコンピュータセキュリティへの応用も活発になってきており、フランス国立情報学自動制御研究所(INRIA)やスペインの IMDEA Software などが暗号通信プロトコルの全自動検証器などを研究・開発している。

INRIA を中心に開発されている OCaml や、マイクロソフトケンブリッジ研究所等で開発されている Haskell といった、数理論理学的基礎を持つ関数型プログラミング言語の実用・普及も、世界的に急速に進んでいる。日本でも学会等でいくつかの応用事例が報告されているだけでなく、IT 系経営誌で複数の特集が組まれている。また、草の根レベルの技術者らによる多数の勉強会が頻繁に開催され、プログラミング言語理論分野の著名な教科書の和訳が Amazon.co.jp において専門書としては異例の新刊 42 位（漫画や雑誌、小説等を含む）にランクインする等、技術的・商業的に注目が高まっている。

（６）キーワード

プログラム意味論、プログラム検証、モデル検査、定理証明支援器、自動定理証明器、コンピュータセキュリティ、関数型言語

（７）国際比較

「基礎研究」、「応用研究・開発」、「産業化」の３つのフェーズについて、現状およびトレンドを記す。

国・地域	フェーズ	現状	トレンド	各国の状況、評価の際に参考にした根拠など
日本	基礎研究	◎	→	プログラム意味論、型システムなど伝統的に強い研究グループが東京大学、京都大学、東北大学などを中心にいくつかあり、国際的な影響力を保っている。ただし、どちらかというと個人ベースの研究に頼っており、研究グループが組織化されていない。
	応用研究・開発	△	↗	基礎研究の成果が検証ツールなどになかなか結び付いていない。ただし本文で述べたように東京大学で実際にプログラム検証ツールの作成などの動きがある。
	産業化	△	→	本文で述べた産総研のシステム検証研究センターなどの取り組みもあるが、全体的に大企業のソフトウェア開発への形式手法の導入の動きが鈍い。ごく一部、フェリカネットワークス社他によるFelica電子マネーの形式的仕様記述や、（形式手法そのものではないが）NTTデータ社によるHaskellを用いたレガシーCOBOLプログラムの解析、（東北大学で開発されている）関数型言語SML#のNECグループによる利用といった、いくつかの応用事例がある。
米国	基礎研究	○	→	どちらかというと応用研究に重点があり、基礎研究の比率が小さい。
	応用研究・開発	◎	→	伝統的に、基礎研究の成果を応用に結び付けることに長けており、大規模プログラムの自動解析ツールなど大学の研究グループでもかなり実践的な研究を行っている。国際会議の発表も、PLDI, CAVなどの実践的なテーマの会議で多い。
	産業化	◎	→	マイクロソフト社がモデル検査に基づくデバイスドライバ検証ツールを作成したり、検証を専門に請け負うベンチャー会社Coverityが立ち上げられるなど、本分野の研究成果が直接産業化につなげられている。
欧州	基礎研究	◎	→	伝統的に基礎理論に強く、LICSやPOPLなど理論重視の国際会議での発表が多い。
	応用研究・開発	○	↗	マイクロソフトケンブリッジ研究所などが本分野のトップレベルの研究者を多く集めていることもあり、最近では応用研究も盛んであり、この傾向は今後も続くと思われる。フランスで開発された定理証明支援器Coqは最近では広く使われるようになってきている。
	産業化	○	↗	上記のマイクロソフトケンブリッジ研究所において研究部門と開発部門が有機的に結びつき、また分離論理に基づく検証を行うベンチャー会社MonoidicsがFacebookに買収されるなど、産業化の動きが活発になってきている。
中国	基礎研究	△	↗	本分野では目立った研究成果はこれまで見られなかった。ただし、最近になって米国などから帰国した研究者を中心に徐々に国際的に見える成果も出現しつつある。
	応用研究・開発	△	↗	本分野では目立った研究成果はこれまで見られなかった。ただし、最近になって米国などから帰国した研究者を中心に徐々に国際的に見える成果も出現しつつある。
	産業化	△	→	特に目立った動きはみられない。
韓国	基礎研究	△	→	あまり目立った成果は見られない。
	応用研究・開発	○	→	プログラム解析の研究など一部グループの成果は見えている。
	産業化	△	→	特に目立った動きはみられない。

（註１）フェーズ
基礎研究フェーズ ：大学・国研などでの基礎研究のレベル
応用研究・開発フェーズ ：研究・技術開発（プロトタイプの開発含む）のレベル

産業化フェーズ ：量産技術・製品展開力のレベル
（註 2）現状
※我が国の現状を基準にした相対評価ではなく、絶対評価である。
◎：他国に比べて顕著な活動・成果が見えている、 ○：ある程度の活動・成果が見えている、
△：他国に比べて顕著な活動・成果が見えていない、×：特筆すべき活動・成果が見えていない
（註 3）トレンド
↗ ：上昇傾向、 →：現状維持、 ↘：下降傾向

（８）引用資料

1) Proceedings of Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages (POPL), http://dl.acm.org/event.cfm?id=RE180
2) Proceedings of ACM/IEEE Symposium on Logic in Computer Science (LICS), http://ieeexplore.ieee.org/servlet/opac?punumber=1000420
3) Proceedings of ACM SIGPLAN International Conference on Functional Programming (ICFP), http://dl.acm.org/event.cfm?id=RE307
4) Proceedings of International Conference on Computer Aided Verification (CAV), http://www.informatik.uni-trier.de/~Ley/db/conf/cav/index.html
5) Proceedings of ACM SIGPLAN Conference on Programming Language Design and Implementation (PLDI), http://dl.acm.org/event.cfm?id=RE200
6) 特集「フォーマルメソッドの新潮流」情報処理 Vol.49 No.5 May 2008
7) 「形式手法適用調査 調査報告書」独立行政法人 情報処理推進機構 ソフトウェア・エンジニアリング・センター 2010 年 7 月 29 日
8) 「型システム入門 プログラミング言語と型の理論」オーム社（原著 "Types and Programming Languages", MIT Press）
9) 特集「ソフトウエアは硬い」日経エレクトロニクス 2005 年 12 月 19 日号 http://techon.nikkeibp.co.jp/article/FEATURE/20090204/165183/
10) 特集『Java はもう古い！次の主流は「関数型」』日経コンピュータ 2012 年 9 月 27 日号 http://itpro.nikkeibp.co.jp/article/COLUMN/20130112/449222/
11) 特集『「オブジェクト指向」のウソ 「関数型」のウソ』日経ソフトウエア 2014 年 11 月号
12) 『ホーア論理を拡張しポインタを扱える「Separation Logic」、静的検証ツール「Infer」で Monoidics 社が実用化』英 Monoidics 取締役兼共同創業者 Peter O'Hearn 氏および CTO 兼共同創業者 Cristiano Calcagno 氏インタビュー、聞き手 日経コンピュータ 進藤 智則、2012/09/26, http://itpro.nikkeibp.co.jp/article/Interview/20120926/425324/
13)「静的検証ツールのコベリティ、Web の脆弱性検出ツールでエンタープライズ分野に本格参入」コベリティ共同創業者兼 CTO Andy Chou 氏インタビュー、聞き手 日経コンピュータ 進藤 智則、2012/10/10, http://itpro.nikkeibp.co.jp/article/Interview/20121009/428581/

３．１．８　データアナリシス

（１）研究開発領域名

データアナリシス

（２）研究開発領域の簡潔な説明

データ解析技術に対する社会からの要請は、ビッグデータの潮流を牽引力としてかつてないほどに高まっている。技術的にはベイズモデリング・スパースモデリングといった理論的な面からの進展や、推薦システム・ネット広告配信といった新しい応用から端を発した発展がみられる。最近特に注目すべき話題は高い性能を理由に爆発的にブームを巻き起こしている深層学習であるが、その理論的な解明は発展途上である。

（３）研究開発領域の詳細な説明と国内外の動向

数年前から始まるビッグデータの潮流は、ビジネス・社会・科学などあらゆる分野を巻き込み、データ解析に対する期待や需要はかつてないほどに高まっている。そして、その基盤技術である統計数理や機械学習・データ工学もまた大きな注目を集めている。当然のことながら、これらの分野はビッグデータが注目されるずっと以前から、それぞれの（統計数理は特に）長い研究の歴史を有しているが、上記の要請に応える形で理論と応用の両面においてますますその発展が加速している状況である。

ビッグデータ時代におけるデータ解析のもつ顕著な傾向のひとつは、仮説検定を中心とした仮説検証型の解析からデータからの帰納的推論を中心とする仮説発見型への重心の移行であろう。大量のデータから有用な知見を発見し、あるいは予測モデルを学習することによって（特に自動的な）意思決定に役立てていくという考え方が、Google や Amazon をはじめとする現在の主要なインターネット企業を支える基盤技術の背景となっている。

統計数理分野では主に理論的な側面からの研究や、小～中規模のデータを対象としたデータ解析技術が発展してきた。特にこの分野で発展してきたベイズモデリングは、現実の様々な制約や知識を柔軟にモデルに取り入れることのできる枠組みであり、マーケティングや生物学など様々な分野で成功裏に用いられている。機械学習分野においてもベイズモデリングは主要テーマのひとつであり、ビッグデータを対象とした効率的なアルゴリズム開発や、ウェブ等の新しい分析対象を扱うための技術が研究されている。

データ解析関連分野において近年特に精力的に研究が進められているテーマのひとつがスパースモデリングであり、疎な信号源を復元する圧縮センシング [1,2]や、予測モデルの疎なパラメータ推定を得るラッソ [2]などがこれら一連の研究の火付け役となった。スパースモデリングとは、スパース（疎）という名からもわかるように、観測されたデータを少数のパラメータだけをもちいて説明するための方法論であり、超高次元のデータからの推論だけではなく、モデルの可解釈性の向上という面からも有用である。また、これらは一定の条件のもと真のパラメータを復元できる保証を与えられるなど理論的な面からも興味深く、その学習理論的解析が進んでいる。

機械学習分野では情報技術の発展とともに現れた様々な形式のデータや、大量データの処理といった問題に対応する形で、新しい問題設定や効率的なアルゴリズムの開発が行われてきた。たとえばオンラインショッピングなど様々なインターネット上のサービスにおいて欠

かせない機能となっている推薦システムは、近年の機械学習キラーアプリケーションのひとつとして認識されている。1990 年代半ばに登場した GroupLens[4]とよばれる推薦システムに端を発する推薦システム研究は、近年米 Netflix 社によって開催されたコンペティションを火付け役に著しい発展を遂げ、現在は行列分解を中心とするアプローチが主流であり [5]、その一般化であるテンソル分解も、さらに高度な分析目的で研究が進んでいる。行列分解は固有空間のスパース性を仮定する一種のスパースモデリングともいえ、これらに対する学習理論的な解析も盛んである。最近ではソーシャルメディアの普及とともに、SNS 等の分析にもこれらの技術が用いられている。

インターネット広告配信も、実世界応用と理論が相互に刺激を与えながら近年大きく発展している領域として挙げられる。インターネット上の広告は多くのインターネット企業にとって依然主要なビジネスモデルであり、その配信精度は企業の収益に結びつく重要な問題である。古くは確率・統計分野で研究されてきた多腕バンティット問題 [6]と広告配信最適化問題との類似性から、これらのアプローチが注目され、現在盛んに研究と応用が進んでいる。

ところで、機械学習分野におけるここ十数年間は、データ解析の問題を凸最適化問題や固有値問題として定式化することによって最適解を保証するアプローチの研究が主流であり、これらを基盤とした学習理論が発展してきた。一方で、近年の応用上の成功を背景として、深層学習 [7]のように複雑な非線形モデルの利用が再び注目を集めている。ここ数年、深層学習の躍進は目覚ましく、コンピュータビジョン・音声認識・自然言語処理など様々な分野において記録を塗り替えつつある。さらに、Google や Facebook といった米国を代表する企業が深層学習に大きな投資を始めたというニュースは深層学習にさらなる注目を集める契機となった。深層学習の成功の背景にあるのは、大量データの出現と、Hadoop などデータを効率的に処理するための並列処理・実装技術、従来のニューラルネットワークで用いられていた誤差逆伝播法に代わる新しい最適化のテクニックなどである。非線形なモデルの複雑さも相まって、深層学習の学習アルゴリズムの多くはまだ理論的な解析が十分ではない状況ではあるが、たとえば深層学習の代表的なテクニックであるドロップアウトなどは従来の学習理論における正則化としての解釈が可能であることが示される [8]など、徐々にではあるが理解は進み始めている。

（４）科学技術的・政策的課題

ビッグデータに関連した種々の取り組みによって、データ解析技術はまさにいま社会へ大きく羽ばたこうとしている状況である。しかし、それは同時にこれらの技術が社会的な問題と無関係ではいられなくなることを意味する。たとえば、プライバシーへの配慮は重要な問題のひとつである。あらゆる種類のセンサーが普及し、これらがネットワークにつながって連携し、一方で多くの人がソーシャルメディアを通じてつながる現在、サービスの利用者がそうとは知らずに自身のプライベートな情報を意図せぬ範囲に晒しているケースが数多くみられ、中には社会的な問題に発展するケースも少なくない。個々のデータとしては一見それほど重要ではないと思えるものであっても組み合わせることで容易に個人情報にたどりつけることも多い。また、結局のところデータを利用したサービスの利便性とプライバシーの保護はトレードオフの関係にあるため、技術的な問題としてだけでは片付かず、扱いは極めて困難である。

個人情報を秘匿した形で公開するための技術は主に統計分野における個票開示[9]の問題として、一方データを秘匿した形でデータ解析を行ういわゆるプライバシー保護データマイニング[10]は機械学習・データマイニングの分野で盛んに研究されている。いずれも技術的にはさまざまな可能性が示されているものの、実際の社会利用における有用性や、制度面との兼ね合いなど、解決すべき課題はまだ多い。

技術的なトレンドとして最も注目されているものはやはり深層学習であろう。応用としては既に多くの成功を収めており、今後もありとあらゆる分野において深層学習の導入が検討されていくものと思われる。その一方で、前述したようにモデルの複雑な非線形性が災いし、その理論的理解が十分になされているとは言い難い。今後、重要な場面において利用される機会が増えれば増えるほど、その礎となる理論的な解析の重要性は高まっていくものと予想される。

（５）注目動向（新たな知見や新技術の創出、大規模プロジェクトの動向など）

やはり最も大きな話題は深層学習であろう。近年 Facebook や Google、Baidu などといった世界的なインターネット企業が深層学習に注力することを発表した。大学で取り組まれていた関連技術が企業に買収されたり、あるいは大学の研究者がこれらの企業に移籍して技術開発に取り組むケースも相次いでいる。

国内ではスパースモデリングを対象とした新学術領域プロジェクト「スパースモデリングの深化と高次元データ駆動科学の創成」が平成 25 年度より始まっており、基礎理論と応用の両面での進展が期待されている。

（６）キーワード

ベイズモデリング、スパースモデリング、推薦システム、ネット広告配信、深層学習、プライバシー

（７）国際比較

国・地域	フェーズ	現状	トレンド	各国の状況、評価の際に参考にした根拠など
日本	基礎研究	○	↗	統計数理研究所等をはじめとする国内機関ではベイズモデリング・スパースモデリング等の基礎研究が脈々と受け継がれてきており、また深層学習の源流であるネオコグニトロンは日本初の技術であるなど、研究の地力は高い。
	応用研究・開発	○	→	自然言語処理やマルチメディア処理の分野では日本は一定の地位を得ており、これらの分野では近年では統計的なデータ処理技術に基づくアプローチが主流である。
	産業化	△	↗	国内企業でもデータ解析を主軸に据える機運は高まっている。基礎研究と産業化の結びつきはまだそれほど強くない。
米国	基礎研究	◎	↗	深層学習の研究を牽引している。プライバシー保護データマイニングの研究も米国発祥である。
	応用研究・開発	◎	↗	あらゆる分野で深層学習の応用が進んでいる。
	産業化	◎	↗	研究者自身が産業化まで取り組む。インターネット企業が相次いで深層学習の産業化に取り組んでおり、現在の流れを牽引している。
欧州	基礎研究	◎	→	理論的な研究が強い。
	応用研究・開発	○	↗	米国に準ずる形で発展している。
	産業化	○	→	プライバシーに関する規制が厳しく、産業化の阻害する可能性もある。
中国	基礎研究	○	↗	コンスタントに論文が発表されている。研究者の数も増えており、今後大きな勢力になると予想される。
	応用研究・開発	○	↗	Baidu等の企業が深層学習に積極的に投資を行っており、早晩頭角を現すと予想される。
	産業化	△	↗	現時点では目立った動きは内が、上記理由等により今後産業化が急速に進むと予想される。
韓国	基礎研究	△	→	特に目立った動きはない。
	応用研究・開発	△	→	特に目立った動きはない。
	産業化	△	→	特に目立った動きはない。

（註 1）フェーズ
基礎研究フェーズ ：大学・国研などでの基礎研究のレベル
応用研究・開発フェーズ ：研究・技術開発（プロトタイプの開発含む）のレベル
産業化フェーズ ：量産技術・製品展開力のレベル
（註 2）現状
※我が国の現状を基準にした相対評価ではなく、絶対評価である。
◎：他国に比べて顕著な活動・成果が見えている、 ○：ある程度の活動・成果が見えている、
△：他国に比べて顕著な活動・成果が見えていない、×：特筆すべき活動・成果が見えていない
（註 3）トレンド
↗ ：上昇傾向、 →：現状維持、 ↘：下降傾向

（８）引用資料

1) Candes, E. & Tao, T. (2006). Near-Optimal Signal Recovery From Random Projections: Universal Encoding Strategies?. Information Theory, IEEE Transactions on, **52**, 5406-5425.
2) Donoho, D. L. (2006). Compressed sensing. IEEE Transactions on Information Theory, **52**, 1289-1306.
3) Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. Journal of the Royal Statistical Society (Series B), **58**, 267-288.
4) Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P. & Riedl, J. (1994). GroupLens: an open architecture for collaborative filtering of netnews. Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work (pp. 175-186).
5) Koren, Y., Bell, R. & Volinsky, C. (2009). Matrix Factorization Techniques for Recommender Systems. Computer, **42**, 30-37.
6) Auer, P., Cesa-Bianchi, N. & Fischer, P. (2002). Finite-time Analysis of the Multiarmed Bandit Problem. Machine Learning, **47**, 235-256.
7) Bengio, Y. (2009). Learning Deep Architectures for AI. Foundations and Trends in Machine Learning, **2**, 1-127.
8) Wager, S., Wang, S. & Liang, P. (2013). Dropout Training as Adaptive Regularization. In C. J. C. Burges, L. Bottou, Z. Ghahramani & K. Q. Weinberger (eds.), Advances in Neural Information Processing (pp. 351-359).
9) 竹村彰通. (2003). 個票開示問題の研究の現状と課題. 統計数理, **51**, 2, 241-260.
10) Aggarwal, C. C. & Yu, P. S. (2008). Privacy-preserving Data Mining: Models and Algorithms. Springer.