

寄稿 技報 90 号の発刊によせて

宮 地 恵 美

本誌は、特定の技術テーマに沿って編集された特集号ではない。日本ユニシスグループでは、今すぐ役に立つ技術から将来的な技術まで、幅広く研究開発に取り組んでいる。本誌は、そのような取り組みの成果の一端を編集したものである。

各論文はそれぞれ独立して読むことができるが、将来技術の理解のために、第 1 章で高速コンピュータ処理について、第 2 章でデータ解析について解説する。第 3 章では、日本ユニシスグループのイノベーションに向けた取り組みについて述べる。

1. 高速処理コンピュータ

1.1 高速処理コンピュータの種類

大規模な科学計算処理、データ処理を高速に行うために使われるコンピュータは、従来の素子を利用するものと、新しい素子を利用するものに大別される（図 1）。

従来素子を利用した高速処理コンピュータは、一般的にはスーパーコンピュータと呼ばれ、並列処理専用の演算プロセッサを用いたベクトル型と、汎用プロセッサを並列に結合したスカラ型がある。また近年では PC の計算性能が向上したため、スーパーコンピュータを使わず、低価格な PC を大量にリンクして高速処理を実現する PC クラスタが商用で利用されている。PC クラスタは、スーパーコンピュータに計算速度では劣るが、コストパフォーマンスが非常に良く、ゲノム情報処理、Web 検索エンジンなどでの利用が広がってきている。

神経細胞構造を真似たバイオコンピュータには、従来素子を利用したものと、従来素子を利用せずに DNA を用いたものがある。従来のコンピュータでは莫大に処理時間かかる NP 問題など、バイオコンピュータの並列処理に適した問題であれば、スーパーコンピュータよりも高速に計算を行うことができると言われている。

従来素子は、その集積に物理的限界があるため、新しい素子として、DNA を構成する塩基分子を利用する DNA コンピュータや、量子的振る舞いをする粒子を利用する量子コンピュータの研究が行われている。

新しい素子を用いたコンピュータの実用化には、ハードウェアだけでなく、ハードウェアに適した専用のアルゴリズムの開発が必要である。現状では、ハードウェアがまだ実用レベルに達していない。また、専用のアルゴリズムも種類が少ないため、有効な適用分野も限られている。

1.2 量子コンピュータ

量子コンピュータのハードウェアは、量子的振る舞いをする新しい粒子を利用することで、従来のコンピュータとは比べものにならないほど高速に計算を行うことができる。

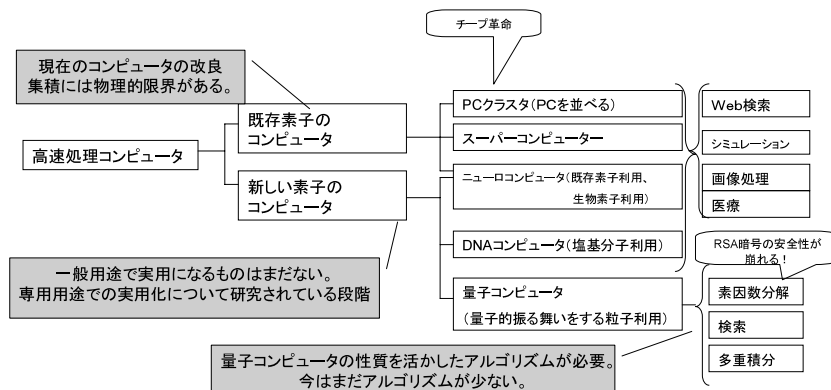


図1 高速処理コンピュータ

2000年に米国のクリントン大統領によって出された「国家ナノテクノロジー優先施策 (National Nanotechnology Initiative)」の具体的目標の一つに、現行スーパーコンピュータの数千倍の計算速度の新しい素子のコンピュータが掲げられている。

量子コンピュータのハードウェアの実現時期については諸説があるが、10数年後から20数年後と言われている。もし、ハードウェアの実現までの間に、量子コンピュータの性能を引き出す各種のアルゴリズムが開発されていれば、ハードウェアが実用化されると同時に、量子コンピュータの高速処理の恩恵を様々な分野で受けられるはずである。

量子コンピュータのアルゴリズムの一つに、1994年に発表されたShorの素因数分解のアルゴリズムがある。このアルゴリズムを用いれば、従来の計算機では現実的な時間で解くことができないとされる素因数分解を、極めて短い時間で計算できるため、素因数分解の困難性を利用したRSA暗号の安全性が崩れることになる。このアルゴリズムの発表で、量子コンピュータは一躍世間の注目を集めるようになった。

その後、積分計算を行う量子コンピュータのアルゴリズムが発表され、素因数分解や検索といった離散問題だけでなく、連続的な問題での量子コンピュータ適用の効果が示された。

本誌の「量子コンピュータを用いた数値積分計算について」では、量子コンピュータを利用した多重積分の計算方法と計算量についての考察を行っている。多重積分は、金融工学、物理学等様々な分野の解析、予測などのシミュレーションで使われる基本的な計算である。

素因数分解、積分などの基本的計算アルゴリズムに加えて、様々な分野での応用アルゴリズムが研究開発されることによって、今まで計算不能であった気候モデリング、生態系のフルスケールシミュレーション、宇宙力学シミュレーションや、たんぱく質解析の高速化による創薬などが量子コンピュータで可能になるかもしれない。

2. データ解析

データ解析には古くから、線形回帰モデルや相関分析などの統計的方法が利用されている。統計的方法を用いれば、実験や調査から得られたバラツキのあるデータから、データの間の規則性や不規則性を見つけることができるので、社会現象の分析や経済動向の予測など、様々な分野で使われている。

データ解析にコンピュータが初めて使われたのは、1952年のアメリカ大統領選挙のときで

ある。CBS テレビが、世界初のプログラム内蔵方式商用コンピュータ UNIVAC I を使って開票予測を行い、開票わずかの段階で当選結果を正しく予測したことにより、コンピュータによるデータ解析の威力が認知された^[1]。

近年では、確率モデルを用いた学習による解析技術が、不確実さやルールの複雑さをもつ大量データの解析に有効であることが示され、注目を集めている。

2.1 確率モデルを用いた学習による解析

確率モデルを用いた学習の代表的なモデルには、ニューラルネット、ベイジアンネットワーク、確率文法の三つがある。これらのモデルは、学習データによってデータ解析の精度が良くなるという特徴を持つ。図 2 に代表的な確率モデルを用いた学習による解析技術とその適用分野を示した。以下で三つの代表的な確率モデルを簡単に説明する。

1) ニューラルネット

ニューラルネットは、人の脳神経細胞を使った情報処理の特性をシミュレーションした数学モデルである。脳の情報処理は、神経細胞（ニューロン）をつなぐ複雑な神経回路網の上で、ニューロン間の情報のやり取りによって行われている。学習によってニューロン間の結合の強さが変わることによって、新しい環境への適応能力や学習能力を持つ。

ニューラルネットは、データ構造が決まっていないデータの解析に適している。データマイニング、画像認識などの実用システムで利用されている。

2) ベイジアンネットワーク

ベイジアンネットワークは、何か事象が起きる確率（確率変数）をノードで表し、因果関係や相関関係の間の向きを一方向のリンクで結んだグラフ構造のモデルである。学習によって、適切なグラフ構造や確率変数の条件付き確率を決めることができる。観測された事象から知りたい対象を予測する計算モデルとして利用されることが多い。

ベイジアンネットワークは、「ゆるい」構造である因果関係や共起関係を表現する確率モデルで、意思決定支援、故障診断などで利用されている。

3) 確率文法

確率文法は、形式文法の生成規則を確率化したものである。生成規則は、まずひとつの開始文字だけから成る文字列から始めて、文字列を書き換えていく文字列変換の規則である。この書き換えの規則の確率（例えば、A という文字の次に B という文字がくる確率）は、学習によって精度をあげることができる。

確率文法は、自然言語の文字、単語の並び、あるいは RNA の塩基配列のように、あらかじめ「かたい」構造が決まってはいるが、ルールがわからないデータ解析に適している。

確率文法には、隠れマルコフモデルと確率文脈自由文法の二つのモデルがあり、それぞれ解析対象に向いているデータの性質が違う。

隠れマルコフモデルは、単純なデータの解析に向き、音声認識や、英語や日本語などの自然言語の文から単語を切り出す処理（形態素解析）で、標準的手法として利用されている。

確率文脈自由文法は、複雑な入れ子構造のあるようなデータの解析に向き、自然言語の文

の解析や RNA 解析で利用されている。

2.2 確率文法推定の確かさ

確率文法を用いることで、自然言語や RNA など、データの構造が決まっているが、そのルールがわからないデータのルールを見つけることができる。ただしこれまでは、見つけたルールの正しさを保証することができなかった。

本誌の「確率文法の推定 ベイズ法によるアプローチ」は、執筆者が行った研究を基にして、データから導き出した仮説モデルの選択結果の正しさを数学的に保証できることについて述べている。

現在、RNA 解析では、確率文法モデルを使った解析結果が正しいかどうかの判断は、人手で行なわれているが、本論文の研究成果を用いれば、解析結果の正しさを保証することができるため、解析結果の検証作業の効率化が見込まれる。さらに、解析結果が正しいかどうかを判断できないために、これまで確率文法が利用されていなかった分野での、確率文法の利用も期待できる。

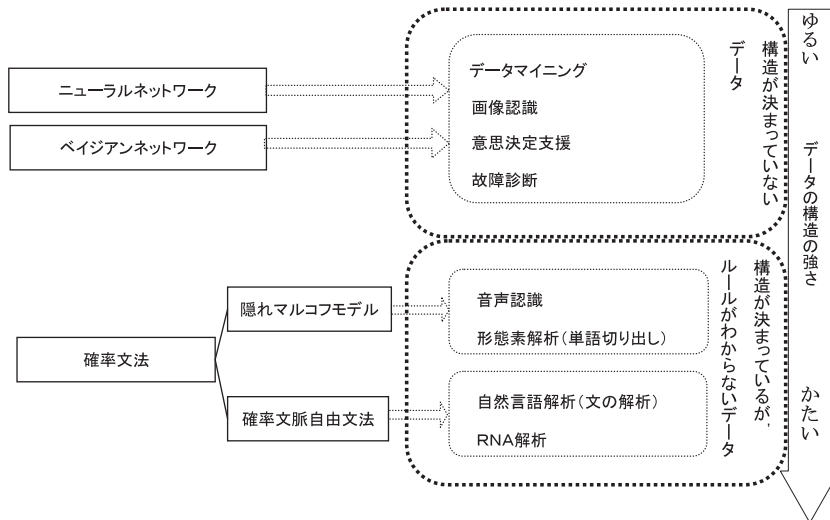


図2 確率モデルを用いた学習によるデータ解析の方法と適用分野

3. イノベーションに向けた日本ユニシスグループの取り組み

本誌の「量子コンピュータの積分」や「確率文法推定」のような基礎研究を、人々に新しい価値をもたらし、経済の新しい価値にするイノベーションにするためには、数学者だけでなく、ハードウェア、ソフトウェア、適用分野の研究者や技術者が、協力して技術開発を進めなければならない。基礎寄りの技術であればあるほど、実用化までの道のりが遠く、関係する研究者、技術者の数も種類も増える。

多種多様な研究者や技術者を集めて、どのような組織をつくり、どのようなプロセスにすれば、イノベーションを推進できるかについては、歴史、経済、社会、経営などの学門分野で論じられている。また技術経営(MOT)においても、イノベーションの推進は、最も関心の高いテーマの一つである。

社会経営学の最新の研究では、従来の階層型の組織よりも、自律的な水平連携型組織のほうが、イノベーション促進に有効であることが提唱されている。一方、社会経営学とは異なり、社会や組織を研究する社会ネットワーク解析、あるいは複雑ネットワークという方法が、それに基づいた google の検索エンジンや Mixi 等の SNS の出現によって、最近注目を浴びるようになってきた。

本誌の「イノベーション促進のためのネットワーク最適化の考察」では、社会経営学の組織についてのマクロ的解釈であるコーディネーション理論を、複雑系ネットワーク理論の分析視点で補い、最適化されたネットワーク、すなわち、最適化された人と人とのつながりがイノベーションの促進につながる可能性について示唆している。

組織を超えて最適化されたネットワークを作るためには、様々な分野の、様々なフェーズの研究者、技術者が、それぞれの研究内容や技術開発課題に関心を持つ必要がある。このために、日本ユニシスグループでは、大学への留学生の派遣、産学連携による研究技術開発を行い、その成果を現場の技術者や営業担当者と共有する取り組みをしている。

参考文献 [1] 米口肇, 「選挙の予想」, 数字セミナー, 日本評論社, 1962年11月

(総合技術研究所先端技術部)