

## 〔研究〕

# 理論・アルゴリズム基盤の研究開発状況

## Research and Development of Theoretical and Algorithmic Foundation

麻生 英樹

H. ASOH

An overview of research activities in the RWI theoretical and algorithmic foundation area is given. Various research issues have been investigated to develop, analyze, and implement novel learning and information integration algorithms based on statistics and probability theory. These issues are classified into five categories and explained.

### § 1 はじめに

われわれ人間は、さまざまな感覚信号から言語情報にいたるまでの多様なモダリティの情報を、何の苦勞もなく日常的に統合し、処理している。あらかじめ状況がきちんと設計された閉じた環境ではなく、家庭やオフィスなどの変化に富んだ開いた環境において、人間と協調して機能できるような「実世界知能システム」には、人間が日常的に扱っているのと同じような多種多様な情報(これらを「実世界情報」と呼ぶ)を統合的に処理できる能力が求められる。

こうした統合的な情報処理の難しさは、実世界情報の複雑さ・多様さ・不確かさに由来する。視覚情報だけをとってみても、われわれの眼が捕らえる情報は時々刻々と変化し、一生の間に「厳密に同じ」情報を二度見ることはありえない。にもかかわらず、われわれ人間はいとも簡単にそうした多様かつ複雑な情報を適切に分節し、構造化し、効率よく処理し、それぞれの入力情報や状況に対して適切な行動をすることができる。

多様な情報のすべてをあらかじめ想定し、処理をプログラムしておくことは不可能であるため、生物の進化の過程の中で、適応・学習の能力が発達してきた。学習とは一種のメタプログラムであり、一つのプログラムを書く代わりに、外界との相互作用を通じて、多様なプログラムの中から特定の環境にあったものを生

成するための「プログラム生成プログラム」を書いておくというシステム実現戦略である。「実世界知能システム」の実現のためには、システムに学習能力を持たせることが不可欠であると考えられている。

理論・アルゴリズム基盤領域の研究の目的は、「実世界情報」に対する「情報統合」および「学習」という二つのキーテクノロジーの理論的な基盤を確立し、それに基づいて実際的なアルゴリズムを提案し、その性能の理論的な評価を与えることにある。実世界情報の不確かさや曖昧さに対処するための方法論には、ファジー論理、非古典論理、確率論理などさまざまな手法が提案されているが、われわれは主に、確率論と統計学に基づくアプローチを採用している。

### § 2 研究の概要

理論・アルゴリズム基盤領域では、当所の実世界知能研究推進センター学習統合基礎ラボ(ETL)に加えて2000年1月現在で、技術研究組合新情報処理開発機構の5つの分散研究室(NEC、東芝、GMD:ドイツ国立情報技術研究センター、SNN:オランダニューラルネットワーク協会、SICS:スウェーデン国立情報処理研究所)および5つの大学再委託先(京都大学統計数理研究所、北陸先端科学技術大学院大学、千葉大学、九州工業大学)が研究を行っている。他の領域に比べて、外国からの参加が多いことが一つの特色である。研究テーマ

**KEY WORDS** : 実世界知能, 機械学習, 情報統合

は、学習や情報統合の基礎理論から評価用の応用システムまで幅広い。以下では、それらの研究拠点でのこれまでの研究内容を次のような5つのカテゴリに分けて概観しておく。

1. 情報や知識の表現と評価に関する基礎理論
2. 確率的な制約・知識を用いた推論のアルゴリズム
3. 確率的な制約・知識の学習アルゴリズム
4. 情報統合のためのソフトウェアアーキテクチャ
5. 評価用課題への応用

## 2.1 情報や知識の表現と評価に関する基礎理論

確率論的な観点に立てば、「情報」とは確率変数の値の組み合わせであり、「知識」とはそれらの変数の間に成り立つ確率的な制約である。モデルとはこの制約を、ある程度一般的な記述力を持つ言語によって記述したものであるが、記述言語自体をモデルと呼ぶこともある(日本語では単数と複数を区別できないためもある)。確率的な制約は、通常、確率分布関数によって表現される。歴史的にも多項分布や正規分布をはじめとして多くの確率分布関数が利用され、その性質が解析されてきた。しかしながら、実世界の複雑な確率的制約を記述するためには、通常、確率分布関数よりも複雑な組み合わせ構造を含む確率分布モデルを用いる必要がある。これらの確率分布モデルは総称的に Graphical Models などと呼ばれて近年研究が盛んになってきているが、本領域では、特に Bayesian Networks<sup>1,2,3)</sup>や有限混合分布<sup>4)</sup>についての研究を主として進めている。また、新奇なモデルの提案として、ETLでは、より高い記述力と操作性を持つ確率制約プログラムを提案している<sup>5,6)</sup>。

知識を記述するための確率分布モデルだけではなく、実世界のさまざまな情報を統一的に表現し、操作するための情報表現についても研究を進めている。SICS研究室では、疎で分散的な情報表現である Spatter Code および Sparchunk Code を提案している<sup>7,8)</sup>。また、東芝研究室ではニューラルネットワークを論理命題とする情報表現について非古典論理の枠組みを用いて検討している<sup>9)</sup>。

ある知識がどれくらいよいものであるかを評価する基準が与えられれば、知識獲得の問題は、その評価基準を最適化する最適探索問題に帰着される。知識を確率分布によって記述する場合には、学習とは確率分

布の推定になる。そのためのよく知られている基準は、尤度関数であるが、この基準は、複数の組み合わせ構造のうちどちらが適切かを選択するためには使えない。NEC研究室では、Rissanenの提案するMDL(Minimum Description Length)基準のもとになっている Stochastic Complexity を Bayes 決定理論の枠組みへと拡張した ESC (Extended Stochastic Complexity) を提案している<sup>10)</sup>。

さらに、情報幾何学の Bayes 的な拡張 (NEC研究室)、遺伝的アルゴリズムの動的振る舞いの解析 (GMD 研究室<sup>2)</sup>、ETL<sup>11)</sup>、有限混合分布の学習過程のダイナミクス解析 (SNN 研究室<sup>4)</sup>) などの基礎理論研究も行ってきている。

## 2.2 確率的な制約・知識を用いた推論と統合のアルゴリズム

組み合わせ構造を持つ確率分布を使った確率計算は一般に計算量が多いため、実時間で応答するシステムを実現するためには、効率よい近似アルゴリズムが必要とされる。特に、計算の途中に、任意のタイミングでそれなりの精度の答えを返せるような anytime アルゴリズムが望ましい。モンテカルロ型のアルゴリズムはこの意味で有望である。NEC研究室では、統計数理研究所と協力してモンテカルロ型の確率計算アルゴリズムについて研究を行っている<sup>12)</sup>。SNN研究室では、複雑なグラフ構造を簡単な構造で近似するアルゴリズムを開発しているほか<sup>13)</sup>、確率計算過程の能動的な制御についても研究を行っている<sup>14)</sup>。

情報統合のアルゴリズムとしては、SICS研究室で、分散的に表現された情報の間の連想アルゴリズムについて研究を行っている<sup>8)</sup>。また、東芝研究室では、ニューラルネットワークを論理命題とみなした場合の推論方式について研究を行っている<sup>9)</sup>。

## 2.3 確率的な制約・知識の学習アルゴリズム

機械学習の分野における最近の話題に、分散学習と能動学習がある。分散学習とは、複数の学習者に、それぞれ、学習用データの一部を使って学習させ、それらの結果を統合する学習方式である。一方、能動学習とは、あらかじめ学習用データが与えられるのではなく、学習者自身が、情報量の多い学習用データを能動的に要求・探索してゆく学習方式である。NEC研究室および SNN研究室ではこうした新しい学習方式の計

算論的学習理論について研究を行っている<sup>15,16)</sup>。また、GMD 研究室では、遺伝的アルゴリズムを分散的な学習方式とみなして、FDA(Factorized Distributed Algorithm)と呼ばれる最適化アルゴリズムを提案している<sup>2)</sup>。また、さらに、各変数間の関係を学習しながら最適化を行う LFDA(Learning Factorized Distributed Algorithm)への拡張も行っている<sup>17)</sup>。SNN 研究室では、複雑な確率分布モデルの学習に必要な統計量の計算に、統計力学で使われている中間場近似を用いる方法を提案している<sup>3)</sup>。ETL では、Bayesian Network とニューラルネットワークを組み合わせ、Bayesian Network の条件付き確率を効率よく学習する方法を提案し、BAYONET と呼ぶ Bayesian Network シミュレータを Java 言語によって実装している。このシミュレータは、Java によるものとしては世界で最も早く実装されたものの一つで、エレガントな GUI によるモデル選択支援機構や外部データベース接続機能を持っている<sup>1)</sup>。

機械学習や統計学の分野で研究されてきたこれまでの学習アルゴリズムは、あらかじめ学習用データが人手によって集められていることを前提にしているものが多い。しかしながら、実世界知能システムの学習を考えた場合に、そうした固定的なデータセットを大量に用意することはむずかしく、より環境やユーザと密にインタラクションしながら、少量のデータを用いて効率よく学習を進めることが望まれる。そのような、実世界知能システムに適した学習の枠組みとして、ETL では、ユーザとの対話から学習に必要な情報を得る、対話ベースの学習<sup>18)</sup>や、複数の情報モダリティの関係を学習するインターモーダル学習、多数の属性を同時並行的に学習する多属性同時学習<sup>19)</sup>などを提案している。

#### 2.4 情報統合のためのソフトウェアアーキテクチャ

多様なモダリティの情報を柔軟に統合するシステムを実装するためには、適切なソフトウェアアーキテクチャが用意されていることが望まれる。GMD 研究室では、Flip-Tick アーキテクチャという複数のソフトウェアモジュールの分散的な実行の実装と制御を容易に行うためのアーキテクチャを提案し、ロボットの制御や Bonn 市の交通シミュレーションなどに適用している<sup>20)</sup>。Flip-Tick アーキテクチャでは、多数の周期的なモジュールが、それぞれの異なる周期 (Tick) ごとに、

タグボードと呼ばれるメッセージボードにアクセスし、情報を交換することで処理を進めて行く。メッセージボードの内容のアップデート(これを、タグボードの Flip と呼ぶ)自体も、エージェントによって制御される。ETL では、イベント駆動型のソフトウェアアーキテクチャを提案し、オフィス移動ロボットに適用している<sup>21)</sup>。それぞれのモジュールは、過去の経験から学習する能力を持つ。

#### 2.5 評価用課題への応用

提案した情報・知識の表現様式、推論・統合・学習のアルゴリズムを評価するために、評価用の課題が必要である。その目的のために、自然言語処理、医療診断、自律学習ロボットなどの領域から課題が選ばれている。「評価用」という意味は、小規模なトイプロブレムということではなく、実世界の大規模なデータを扱う問題や、実世界知能システムのプロトタイプと密接に関係した課題である。NEC 研究室では、大規模な言語データを利用して単語のクラスタリングを行い、構文解析の曖昧性解消などに適用して、世界的水準の精度を示している<sup>22)</sup>。また、京都大学と協力して、免疫蛋白の同定実験に能動学習法を適用し、実験の効率化を試みている。SNN 研究室では、ユトレヒト大学の病院と協力して、大規模な確率ネットワークを貧血の医療診断に適用している。GMD 研究室では、ボン市と協力して Flip-Tick アーキテクチャによるマイクロシミュレーションを、大規模交通流シミュレーションに適用している。また、FDA による最適化手法を、フランスの電話会社と協力して、携帯電話用のアンテナの最適配置問題に適用している。ETL では、対話ベース学習を事情通ロボットの地図学習に適用している<sup>18)</sup>。

### § 3 おわりに

本稿では、実世界知能分野の理論・アルゴリズム領域における研究を概観した。基礎理論から評価用の応用に至る幅広い研究が、複数の研究拠点において、分散並行的に実施されているが、それらはまったく独立に進められているわけではなく、年 3 回程度開催されている領域のワークショップ等を通じて、相互に交流が図られている。中でも、Graphical Models、特に Bayesian Network に関連した研究は、SNN、GMD、ETL で行われ

ており、相互の研究者の交流も盛んである。今後も、より一層の交流と協力関係の構築を行い、領域の成果をまとめてゆきたい。本領域の研究成果は、理論的成果の論文発表等による学問的貢献のみならず、提案アルゴリズムのソフトウェアによる提供、評価用の応用システムの開発、実世界知能プロトタイプシステムへの貢献、などの形態がありえる。また、現在、領域の成果を教科書にとりまとめて出版する方向での検討も進めている。

### 謝辞

プロジェクトの遂行を支援していただいている、大津展之実世界知能研究推進センター長、島田潤一RWCPつくば研究所長をはじめとする皆様方に感謝いたします。

### 参考文献

- 1) Y. Motomura, I. Hara, H. Asoh, and T. Matsui, "Bayesian network that learns conditional probabilities by neural networks", Proc. of 1997 Int. Conf. on Neural Information Processing (1997) 584-587.
- 2) H. Muehlenbein, Th. Mahning, and A. Ochoa, "Schemata, distributions and graphical models in evolutionary optimization", Journal of Heuristics **5** (1999) 215-247.
- 3) W. Wiegerinck and D. Barber, "Mean field theory based on belief networks for approximate inference", Proc. of 1998 Int. Conf. on Artificial Neural Networks (1998)
- 4) S. Akaho and B. Kappen, "Nonmonotonic generalization bias of Gaussian mixture models", Neural Computation (2000) (in press).
- 5) K. Hasida, "Constraint-based derivation of cognitive model on parsing", Proc. of the International Conference on Cognitive Science '97 (1997) 51-56.
- 6) K. Hasida, "Parsing and Generation with Tabulation and Compilation", Proc. of TAPD '98 (1998) 26-35.
- 7) G. Sjoedin, "The Sparchunk code: A method to build higher-level structures in a sparsely encoded SDM", Proc. of 1998 Int. Joint Conf. on Neural Networks (1998).
- 8) P. Kanerva, "Pattern completion with distributed representation", Proc. of 1998 Int. Joint Conf. on Neural Networks (1998).
- 9) 月本 洋, "ニューラルネットワークによる非単調推論", 人工知能学会研究会資料 SIG-J-9702-14 (1997) 83-88.
- 10) K. Yamanishi, "Minimax relative loss analysis for sequential prediction algorithms using parametric hypotheses", Proc. of the 11th Conf. on Computational Learning Theory (1998).
- 11) T. Niwa and M. Tanaka, "Analyses of simple and island model GAs", Proc. of 1997 Int. Conf. on Artificial Neural Networks and Genetic Algorithms (1997).
- 12) Y. Iba, "Markov chain Monte Carlo algorithms and their applications to statistics", Proc. of the Institute of Statistical Mathematics **44** (1996) 49-84.
- 13) W. Wiegerinck and B. Kappen, "Approximation of Bayesian networks through KL minimization", New Generation Computing **18** (2000) 167-175.
- 14) P. Laar, S. Gielen and T. Heskes, "Input selection with partial retraining", Proc. of 1997 Int. Conf. on Artificial Neural Networks (1997) 469-474.
- 15) N. Abe and H. Mamitsuka, "Query learning strategies using boosting and bagging", Proc. of the 15th Int. Conf. on Machine Learning (1998).
- 16) T. Heskes, "Balancing between bagging and bumping", Advances in Neural Information Processing Systems **9** (1997) 466-472.
- 17) H. Muehlenbein and T. Mahning, "Evolutionary optimization using graphical models", New Generation Computing **18** (2000) 157-166.
- 18) H. Asoh, S. Hayamizu, I. Hara, Y. Motomura, S. Akaho, and T. Matsui, "Socially embedded learning of the office-conversant robot Jijo-2", Proc. of the 15th Int. Joint Conf. on Artificial Intelligence (1997) 888-885.
- 19) 赤穂, 速水, 長谷川, 吉村, 麻生, "EM法を用いた複数情報源からの概念獲得", 電子情報通信学会論文誌, **J80-A** (1997) 1546-1553.
- 20) G. Richter, "Flip-Tick: a design paradigm for cycle-oriented distributed systems", GMD Report 84 (1999).
- 21) H. Asoh, I. Hara, and T. Matsui, "A structured dynamic multi-agent architecture for controlling mobile office-conversant robot", Proc. of 1998 IEEE Int. Conf. on Robotics and Automation (1998) 1552-1557.

- 22) H. Li and N. Abe, "Word clustering and disambiguation based on co-occurrence data", Proc. of Computational Linguistics/ACL-98 (1998).

(2000.5.1受付)

