

## (31) 大規模データベースからの知識マイニングに関する研究

研究代表者 工学部 鍾 寧 (Ning Zhong)

### 研究の目的

コンピュータネットワークの進展により、大規模な情報（知識とデータ）の共有や再利用が可能となった。データベースに対するニーズが増加するにつれ、それらに蓄積されているデータを有効に活用してデータ群の背後にある法則や因果関係を分析/抽出する専門家の不足が指摘されている。このような動きの中で、データベースからルールや因果関係などの知識を発見/マイニングするシステムの必要性が高まっている。しかし、知識発見/データマイニング (Knowledge Discovery and Data Mining, KDD と略記) に関する研究は開始されてから日が浅く、実世界のデータベースの“複雑性や大規模性、不完全性と曖昧性、動的性、多様性”などの特性に体系的に対応できるシステムはまだ開発されていない。本研究は、実世界のデータベースの複雑性や大規模性などの特性を考慮した並列分散協調型マルチデータベースマイニングシステムの研究・開発を目的とする。

### 研究の成果

本研究は、順調に進行しており、既にデータベースからの知識発見のための多戦略学習・分散協調方式のモデルを構築した。これに基づいて、以下の四つの面から具体的な結果が出始めている。

- 自律知識発見エージェントの社会のマルチレベルの構造。  
多種学習エージェントの協調利用や多学習段階で連続、及び並列分散協調処理できる知識発見プロセスを動的に構成できる発見システムのマルチレベルの構造を確立した。これにより、自律性及び汎用性をともに増大させるようになる。更に自律的発見とインタラクション発見を統合するシステムが構築できる。
- 知識発見プロセスのプランニング・組織化の方式。  
知識発見の目的に応じて環境の認識またはヒュー

マン・コンピュータインタラクションによる自律的知識発見プロセスをプランニング・組織化する方式を開発した。ことにより、ネットワークで結合された任意の計算機上に多種学習エージェントを動的に生成でき、更に動的に生成されたマルチエージェント間の通信交渉・分散協調利用などを制御し、複雑な発見目標を達成できる。

- ハイブリッド知能モデルによるルール発見法。  
演繹、帰納、アブダクションの推論過程と確率過程との対応関係を見つけ、遷移確率行列及びそのコネクションズネットワークの表現を通して、記号処理と非記号処理の手法と機能を一体化させる「ハイブリッド知能モデル」を開発した。これにより、大規模データベースからの知識発見のための新しいパラダイムを確立し、知識マイニングシステムを構築している。また、ラフ集合理論やファジィ論理などのソフトコンピューティング技術、帰納論理プログラミング技術を利用することによって、提案された「ハイブリッド知能モデル」を制約し、効率化する方法を開発している。
- マルチデータベースマイニングの方式。  
ある発見の目的に応じて、データベースリバースエンジニアリング (Database Reverse Engineering)、グラニューラコンピューティング (Granular Computing)、知識指向相関性分析などの技術により、マルチデータベース間の相関性を見つけ、マルチデータベースに隠れている概念モデルを発見し、拡張した ER モデルによる表現する方式を開発した。また、拡張した ER モデルから「相関ルール」の表現へ変換する方法を開発している。

### 産業技術への貢献

新しい多戦略学習・分散協調方式、ハイブリッドモデル、マルチデータベースの構造化の技術を基に、

大規模データベースからの知識マイニング・科学的発見システムを構築することにより、実世界の問題の取扱いに向かって、現状のデータ処理および知識処理の技術をブレイクスルーするような、大規模・複雑な処理を可能とする技術、更に各機械学習の要素技術がインフラとして使用されるために必要な相互運用性の向上に関する技術の発展に貢献する。

### 研究発表

- 1) Zhong N. and Yamashita S. "A Way of Multi-Database Mining", *Proc. the IASTED International Conference on Artificial Intelligence and Soft Computing (ASC'98)* (1998) (in press).
- 2) Zhong, N., Dong, J.Z., Ohsuga, S., and Lin, T.Y. "An Incremental, Probabilistic Rough Set Approach to Rule Discovery", *Proc. 1998 IEEE World Congress on Computational Intelligence (WCCI'98)* (in press).
- 3) Zhong, N., Dong, J.Z., and Ohsuga, S. "Data Mining based on the Generalization Distribution Table and Rough Sets", X. Wu et al. (eds.) *Research and Development in Knowledge Discovery and Data Mining*, Lecture Notes in AI 1394, Springer (1998) (in press).
- 4) Zhong, N., Dong, J.Z., and Ohsuga, S. "Data Mining : A Probabilistic Rough Set Approach", A. Skowron and L. Polkowski (eds.) *Rough Sets in Knowledge Discovery*, Physica Verlag (1998) (in press).
- 5) Zhong, N., Dong, J. Z., Ohsuga, S., and Lin, T.Y. "A Hybrid Approach to Rule Discovery in Databases", *Information Sciences, An International Journal*, the special issue on Rough Sets and Data Mining, North-Holland (1998) (in press).
- 6) Zhong, N. and Ohsuga, S. "Techniques and Applications of Automatic Knowledge Discovery in Larger Scale Knowledge-Data Bases", Leonides, C.T. (ed.) *Expert Systems Techniques and Applications*, GORDON and BREACH Science Publishers (1998) (in press).
- 7) Zhong, N., Liu, C., and Ohsuga, S. "A Way of Increasing both Autonomy and Versatility of a KDD System", Z.W. Ras and A. Skowron (eds.) *Foundations of Intelligent Systems*. Lecture Notes in AI 1325, Springer (1997) 94-105.
- 8) Zhong, N., Dong, J.Z., and Ohsuga, S. "Soft Techniques to Rule Discovery in Data", *Proc. the Fifth Congress on Intelligent Techniques and Soft Computing (EUFUT-97)* edited in the Invited Session on Soft Techniques in Knowledge Discovery (1997) 212-217.
- 9) Zhong, N., Liu, C., Kakemoto, Y., and Ohsuga, S. "KDD Process Planning", *Proc. Third International Conference on Knowledge Discovery and Data Mining (KDD-97)*, AAAI Press (1997) 291-294.
- 10) Zhong, N., Ohsuga, S., Liu, C., Kakemoto, Y., and Zhang, X. "On Meta Levels of an Organized Society of KDD Agents.", J. Komorowski and J. Zytkow (eds.) *Principles of Data Mining and Knowledge Discovery*. Lecture Notes in AI 1263, Springer (1997) 367-375.
- 11) Zhong, N., Dong, J.Z., and Ohsuga, S. "Discovering Rules in the Environment with Noise and Incompleteness", *Proc. 10th International Florida AI Research Symposium (FLAIRS-97)* edited in the *Special Track on Uncertainty in AI* (1997) 186-191.
- 12) Zhong, N. and Ohsuga, S. "On Information of Logical Expression and Knowledge Refinement", *Transactions of Information Processing Society of Japan*, Vol.38, No.4 (1997) 687-697.
- 13) Zhong, N. and Ohsuga, S. "A Multi-Phase Process for Discovering, Managing and Refining Strong Functional Relationships Hidden in Databases", *Transactions of Information Processing Society of Japan*, Vol.38, No.4 (1997) 698-706.
- 14) Zhong, N., Kakemoto, Y., and Ohsuga, S. "An Organized Society of Autonomous Knowledge Discovery Agents", Peter Kandzia and Matthias Klusch (eds.) *Cooperative Information Agents*. Lecture Notes in AI 1202, Springer (1997) 183-194.
- 15) Zhong, N., Fujitsu, S., and Ohsuga, S. "The Connectionist Networks Representation of a GDT for Mining Rules from Databases", *Proc. Third Joint Conference on Information Sciences (JCIS'97)* (1997) Vol. 3, 375-378.
- 16) Zhong, N., Fujitsu, S., and Ohsuga, S. "Generalization based on the Connectionist Networks Representation of a Generalization Distribution Table", H. Lu, H. Motoda, and H. Liu (eds.) *KDD : TECHNIQUES AND APPLICATIONS*. World Scientific (1997) 183-197.

### グループメンバー

氏名	所属	職(学年)
鍾 寧	工・知能情報	助教授
董 菊珍	理工・システム工学	D 1
山下 悟司	理工・知能情報	M 1

### 連絡先

TEL&FAX : 0836-35-9949

E-mail : zhong@ai.csse.yamaguchi-u.ac.jp