

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 24 年 6 月 6 日現在

機関番号：62615

研究種目：基盤研究（A）

研究期間：2008 ～ 2011

課題番号：20240016

研究課題名（和文）推論による仮説発見とシステム生物学への応用

研究課題名（英文）Inference-based Hypothesis-finding and its Application to Systems Biology

研究代表者

井上 克巳（INOUE KATSUMI）

国立情報学研究所・情報学プリンシプル研究系・教授

研究者番号：10252321

研究成果の概要（和文）：本研究では、システム生物学において、未知で有用であるような知識を発見するために、論理的な仮説発見手法であるアブダクションとインダクションに着目し、これらに共通する推論の基盤技術を開発している。仮説発見のエンジンとして、SOL 導出に基づく演繹手法である結論発見手続き SOLAR を使用するが、推論効率の改善と仮説選択手段の開発が重要課題として挙げられる。そこで、本研究では次の 3 テーマを設定して取り組んだ。

- (1) 仮説発見のために、SOLAR と関連する仮説生成手法の効率化と高機能化を図る。
- (2) 仮説選択のために、確率的・非確率的手法に基づいた有効な選択方法を開発する。
- (3) システム生物学の問題に応用し評価する。

研究成果の概要（英文）：Systems biology is an emergent field that aims to understand living organisms as biological systems. To find useful information from biological data, we focus on inference methods, *induction* and *abduction*, based on logic-based Artificial Intelligence. Both abduction and induction are used to infer hypotheses in inductive logic programming (ILP), and can be implemented using the consequence finding procedure SOLAR. However, in the current inference methods, efficiency of inference needs to be improved, and mechanisms for hypotheses selection are necessarily developed. Hence, the following three subtasks are set in this project:

- (1) Improving hypothesis finding methods based on SOLAR and related techniques;
- (2) Developing effective hypothesis selection methods based on both statistical and non-statistical approaches;
- (3) Applying those methods to problems in systems biology.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2008 年度	7,000,000	2,100,000	9,100,000
2009 年度	6,200,000	1,860,000	8,060,000
2010 年度	6,300,000	1,890,000	8,190,000
2011 年度	6,700,000	2,010,000	8,710,000
総計	26,200,000	7,860,000	34,060,000

研究分野：知能情報学

科研費の分科・細目：情報学・知能情報学

キーワード：人工知能，推論，システム生物学，仮説発見，結論発見，確率推論，帰納論理プログラミング，アブダクション

- | | |
|------------------------------------|--|
| 1. 研究開始当初の背景
生命・細胞・生体をシステムとして理解 | するのためのシステム生物学が注目されている。この中でも、代謝経路においてある代謝 |
|------------------------------------|--|

物の生成がどの反応により促進または阻害され、どの経路が支配的に使われたかを知ることが非常に難しい問題である。これらの情報が仮説として得られれば微分方程式系によるシステム解析が可能になり、特定の代謝物の増殖に関する系全体での巨視的な目標を達成する有効な手段となる。こうした問題を解くためには、未知で有用であると考えられる仮説を効率良く発見する必要がある。

このために、論理的な発見手法であるアブダクションとインダクションに着目し、両者に共通するツールとして、演繹手法の SOL 導出を効率的に実現した結論発見手続き SOLAR を開発してきた。アブダクションは SOLAR でそのまま実現できるが、インダクションは CF 帰納法と呼ばれる帰納仮説発見方法を用いる必要がある。しかしながら、いずれも仮説探索の領域が大きいため効率の改善が必要なこと、および得られた多くの仮説からより良い仮説を選ぶ手段が弱いこと、の2点が課題として挙げられていた。

2. 研究の目的

上記の背景の下で、本研究ではこれまでの仮説発見に関する研究成果を発展させ、背景知識と観測データから、アブダクションおよびインダクションにより観測を説明するための仮説を効率的に推論し、これらの仮説が多数得られた場合には確率推論等により尤もらしい仮説を選択する理論を構築・実現することを研究目的とした。すなわち、仮説発見システムの効率化と仮説選択方式の実現を行い、これらを組み合わせることで有用な仮説を得るための技術基盤が整理される。さらに、実際にシステム生物学の諸問題にこのシステムを適用し、その有効性を検証する。

3. 研究の方法

上記の研究目的実現のために、3つのサブテーマを設定し、各サブテーマ毎に以下に記す方法によって研究を進めた。

(1) 仮説発見システムの開発

等号を含む仮説発見の効率化と高度化のために、SOL 結論発見手続きにおける等号推論の高速化について研究を行った。具体的には Ordred-Modification 法に基づく等号推論の高速化、再スタート型タブロー法における再構成、分割統治法の導入による効率改善を試み、SOLAR に実装し性能評価実験を行った。また帰納推論に基づく仮説生成手法である CF 帰納法の汎化手続きを簡略化した。さらに SOLAR を用いて、メタアブダクションと呼ばれる、規則を生成する新たなアブダクション方式を提案し、正負の効果を持つリンクを有するネットワーク補完方法を開発した。

(2) 仮説選択システムの開発

論理関数をコンパクトに表現する BDD を用

いて確率計算を行う研究を進め、EM アルゴリズムを実装し、アブダクションの仮説をランキングする方法を提案した。確率モデリング言語 PRISM を拡張し、最尤推定・MAP 推定・変分ベイズ法・MCMC 法など豊富な推論機能を提供する確率モデリング言語を開発した。また実数値データの重要性に鑑み、実数値データをガウシアン隠れマルコフモデル等により離散データに変換するアルゴリズムを導出した。非確率的な仮説選択手法として、敢為的帰納推論を導入し、生化学ネットワークからの新たな仮説知識の発見可能性を示した。また帰納推論の下での異なる生物知識ベースの比較について検討した。

(3) システム生物学における仮説発見

酵母菌の解糖系に関する論理モデルに対して、モデルに欠落している一般規則を CF 帰納法により推定する研究を行い、従来手法では発見困難な仮説の生成に成功した。また、大腸菌の論理モデルを構築し、代謝物質の変化に関する規則を方程式系から導き、SOLAR によるアブダクション結果を確率的にランク付けして状態推定を行う閉ループ学習システムを開発した。さらに、癌の遺伝子制御ネットワークにおける補完問題のメタアブダクションによる解法を検討した。

4. 研究成果

各サブテーマに対応した主な研究成果とを3点挙げるが、他にも多くの進展があった。

(1) SOLAR の改良を重ねた結果、現在 SOLAR の仮説発見能力は、仮説推論システムとしては世界最高の性能に達しており、仮説発見の特殊形である定理証明に適用した場合でも最先端の定理証明専用システムに準じる性能を有するに至った。仮説生成に関しては、当初の研究課題である CF 手続きの論理的簡略化、解仮説の枚挙、演繹計算や双対化計算の効率化の実現について、ILP 2010

で論文賞を受賞し Machine Learning 誌に論文が掲載されるなど優れた成果を得た。

(2) アブダクション仮説の確率的評価方法は、IJCAI-09 にて発表し、条件付き確率に基づいた方法を Machine Learning 誌で提案するなど優れた成果を得た。非確率的な方式についても、Machine Learning 誌に2本の研究論文が掲載され優れた業績を得た。さらに ILP 総括に関する論文が Machine Learning 誌に掲載され、4年の研究期間で Machine Learning 誌に合計5本の論文が掲載された。

(3) 仮説発見・仮説選択の研究成果をシステム生物学に総合的に適用した閉ループ学習に関する研究は、BIOINFORMATICS 2011 にて(学生)最優秀論文賞を受賞するなど国際的にも高く評価された。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 55 件)

- ① Yoshitaka Yamamoto, Katsumi Inoue, Koji Iwanuma. Inverse Subsumption for Complete Explanatory Induction. *Machine Learning*, 86(1):115-139, 2012, 査読有.
[DOI: 10.1007/s10994-011-5250-y](https://doi.org/10.1007/s10994-011-5250-y)
- ② Stephen Muggleton, Luc De Raedt, David Poole, Ivan Bratko, Peter Flach, Katsumi Inoue, Ashwin Srinivasan. ILP Turns 20: Biography and Future Challenges. *Machine Learning*, 86(1):3-23, 2012, 査読有.
[DOI: 10.1007/s10994-011-5259-2](https://doi.org/10.1007/s10994-011-5259-2)
- ③ Taisuke Sato, Masakazu Ishihata, Katsumi Inoue. Constraint-based Probabilistic Modeling for Statistical Abduction. *Machine Learning*, 83(2):241-264, 2011, 査読有.
[DOI: 10.1007/s10994-010-5206-7](https://doi.org/10.1007/s10994-010-5206-7)
- ④ Chiaki Sakama, Katsumi Inoue. Inductive Equivalence in Clausal Logic and Nonmonotonic Logic Programming. *Machine Learning*, 83(1):1-29, 2011, 査読有
[DOI: 10.1007/s10994-010-5189-4](https://doi.org/10.1007/s10994-010-5189-4)
- ⑤ 岩沼 宏治, 鍋島 英知, 井上 克巳. 一階論理上の等号推論: 理論と実際. コンピュータソフトウェア, 28(4):282-305, 2011, 査読有
[JST/JSTAGE/jssst/28.4 282](https://doi.org/10.1007/s10994-010-5189-4)
- ⑥ 井上 克巳. アブダクションとインダクション. 人工知能学会誌, 25(3):389-399, 2010, 査読有.
[URL: ci.nii.ac.jp/naid/110007610263/ja/](http://ci.nii.ac.jp/naid/110007610263/ja/)
- ⑦ Hidetomo Nabeshima, Koji Iwanuma, Katsumi Inoue, Oliver Ray. SOLAR: An Automated Deduction System for Consequence Finding. *AI Communications*, 23(2-3):183-203, 2010, 査読有.
[DOI: 10.3233/AIC-2010-0465](https://doi.org/10.3233/AIC-2010-0465)
- ⑧ Chiaki Sakama, Katsumi Inoue. Brave Induction: A Logical Framework for Learning from Incomplete Information. *Machine Learning*, 76(1):3-35, 2009, 査読有.
[DOI: 10.1007/s10994-009-5113-y](https://doi.org/10.1007/s10994-009-5113-y)
- ⑨ Chiaki Sakama, Katsumi Inoue. Equivalence Issues in Abduction and Induction. *Journal of Applied Logic*, 7(3):318-328, 2009, 査読有.
[DOI: 10.1016/j.jal.2008.10.006](https://doi.org/10.1016/j.jal.2008.10.006)
- ⑩ Koji Iwanuma, Katsumi Inoue,

Hidetomo Nabeshima. Reconsideration of Circumscriptive Induction with Pointwise Circumscription. *Journal of Applied Logic*, 7(3):307-317, 2009, 査読有.
[DOI: 10.1016/j.jal.2008.10.005](https://doi.org/10.1016/j.jal.2008.10.005)

[学会発表] (計 101 件)

- ① Gauvain Bourgne, Katsumi Inoue. Partition-Based Consequence Finding. *The 23rd IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2011)*, Palm Beach County, Florida, USA, November 9, 2011
- ② Yoshitaka Kameya. Time Series Discretization via MDL-Based Histogram Density Estimation. *The 23rd IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2011)*, Palm Beach County, Florida, USA, November 9, 2011
- ③ Yoshitaka Yamamoto, Katsumi Inoue, Koji Iwanuma. Comparison of Upward and Downward Generalization in CF-Induction. *The 21st International Conference on Inductive Logic Programming (ILP 2011)*, Windsor, Great Park, UK, August 1, 2011
- ④ Katsumi Inoue. DNF Hypotheses in Bottom-Directed ILP. *The 21st International Conference on Inductive Logic Programming (ILP 2011)*, Windsor, Great Park, UK, August 1, 2011
- ⑤ Katsumi Inoue. Logic Programming for Boolean Networks. *The 22nd International Joint Conference on Artificial Intelligence (IJCAI-11)*, Barcelona, Spain, July 22, 2011
- ⑥ Chiaki Sakama. Dishonest Reasoning by Abduction. *The 22nd International Joint Conference on Artificial Intelligence (IJCAI-11)*, Barcelona, Spain, July 22, 2011
- ⑦ Taisuke Sato. A General MCMC Method for Bayesian Inference in Logic-Based Probabilistic Modeling. *The 22nd International Joint Conference on Artificial Intelligence (IJCAI-11)*, Barcelona, Spain, July 22, 2011
- ⑧ Gabriel Synnaeve, Katsumi Inoue, Andrei Doncescu, Hidetomo Nabeshima, Yoshitaka Kameya, Masakazu Ishihata, Taisuke Sato. Kinetic Models and Qualitative Abstraction for Relational Learning in Systems Biology. *The 2011 International Conference on Bioinformatics Models, Methods and Algorithms (BIOINFORMATICS 2011)*, Rome, Italy, January 27, 2011

⑨ Yoshitaka Kameya, Gabriel Synnaeve, Andrei Doncescu, Katsumi Inoue, Taisuke Sato. A Bayesian Hybrid Approach to Unsupervised Time Series Discretization. *The 15th Conference on Technologies and Applications of Artificial Intelligence* (TAAI 2010), Hsinchu, Taiwan, November 20, 2010

⑩ Gauvain Bourgne, Katsumi Inoue, Nicolas Maudet. Abduction of Distributed Theories through Local Interactions. *The 19th European Conference on Artificial Intelligence* (ECAI 2010), Lisbon, Portugal, August 20, 2010

⑪ Takehide Soh, Katsumi Inoue. Identifying Necessary Reactions in Metabolic Pathways by Minimal Model Generation. *The 6th Conference on Prestigious Applications of Artificial Intelligence* (PAIS 2010), Lisbon, Portugal, August 18, 2010

⑫ Katsumi Inoue, Andrei Doncescu, Hidetomo Nabeshima. Hypothesizing about Networks by Meta-level Abduction. *The 20th International Conference on Inductive Logic Programming* (ILP 2010), Firenze, Italy, June 29, 2010

⑬ Katsumi Inoue, Taisuke Sato, Masakazu Ishihata, Yoshitaka Kameya, Hidetomo Nabeshima. Evaluating Abductive Hypotheses Using an EM Algorithm on BDDs. *The 21st International Joint Conference on Artificial Intelligence* (IJCAI-09), Pasadena, USA, July 14, 2009

⑭ Koji Iwanuma, Hidetomo Nabeshima, Katsumi Inoue. Toward an Efficient Equality Computation in Connection Tableaux: A Modification Method without Symmetry Transformation. *The 7th International Workshop on First-Order Theorem Proving* (FTP 2009), Oslo, Norway, July 6, 2009

⑮ Katsumi Inoue, Chiaki Sakama. Comparing Abductive Theories. *The 18th European Conference on Artificial Intelligence* (ECAI 2008), Patras, Greece, July 24, 2008

[図書] (計 2 件)

① Andrei Doncescu, Sebastien Regis, Katsumi Inoue, Nathalie Goma. Physiological Analysis of Yeast Cell by Intelligent Signal Processing. *Applied Biological Engineering—Principles and Practice*, Chapter 19, pages 435-460, 2012
[DOI: 10.5772/29776](https://doi.org/10.5772/29776)

② Yoshitaka Yamamoto, Katsumi Inoue,

Andrei Doncescu. Integrating Abduction and Induction in Biological Inference using CF-Induction. In: Huma Lodhi and Stephen Muggleton (eds.), *Elements of Computational Systems Biology*, Chapter 9, pages 213-234, John Wiley and Sons, Inc., 2009

[DOI: 10.1002/9780470556757.ch9](https://doi.org/10.1002/9780470556757.ch9)

[その他]

ホームページ

http://research.nii.ac.jp/il/event/project_a.html

6. 研究組織

(1) 研究代表者

井上 克巳 (INOUE KATSUMI)

国立情報学研究所・
情報学プリンシプル研究系・教授
研究者番号：10252321

(2) 研究分担者

佐藤 泰介 (SATO TAISUKE)

東京工業大学・大学院情報理工学研究科・教授
研究者番号：90272690

亀谷 由隆 (KAMEYA YOSHITAKA)

東京工業大学・大学院情報理工学研究科・助教
研究者番号：60361789

岩沼 宏治 (IWANUMA KOJI)

山梨大学・大学院医学工学総合研究部・教授
研究者番号：30176557

鍋島 英知 (NABESHIMA HIDETOMO)

山梨大学・大学院医学工学総合研究部・准教授
研究者番号：10334848

山本 泰生 (YAMAMOTO YOSHITAKA)

山梨大学・大学院医学工学総合研究部・助教
研究者番号：30550793

坂間 千秋 (SAKAMA CHIAKI)

和歌山大学・システム工学部・教授
研究者番号：20273873

(3) 連携研究者

藤山 秋佐夫 (FUJIYAMA ASAO)

国立情報学研究所・
情報学プリンシプル研究系・教授
研究者番号：60142311