

説明と予測: 科学的発見からレジリエンスへ

Explanation and Prediction: From Discovery to Recovery

井上 克巳*1

Katsumi Inoue

*1 国立情報学研究所

National Institute of Informatics

In the history of science, there exist the philosophy and methodology for scientific discovery. If this methodology of discovery processes were modeled and automated, scientific discovery could be realized on computers. Already it has become that scientists cannot work without computers in the era of information technology with a huge amount of information and big data, then it is important to think anew about the automation of discovery. Not only in scientific discovery, but such a methodology can be utilized in many areas of daily-life and business applications. In this talk, I will first review the history of science and the automation of discovery by Artificial Intelligence, and will introduce my past works on inference-based hypothesis-finding, stochastic hypothesis evaluation, meta-level abduction, and application to systems biology. In the science of discovery, for an event that cannot be explained by the existing theory, a hypothesis to explain it is produced by abduction and is then attempted to connect to a true discovery by evaluating it. In this case, the event has not been expected in the current situation, and the hypothesis is a cause that happened in the past or is a universal rule. To the contrary, if this reasoning is performed forward from the present to the future, we can make a prediction of an event. Then an event is unpredictable if it cannot be explained by abduction, and an unpredictable event corresponds to an unreachable state in the possible world semantics. Related to prediction, I have been involved in the new research area called “Systems Resilience” as a science to deal with unexpected events such as natural disasters. As a scientist who lives after the 3.11 earthquake disaster, along with the science of *discovery*, I now believe that the science of *recovery* is important as well.

1. 発見の科学

科学の歴史は人類の発見の歴史である。科学的発見について詳しく分析すると、発見のための方法論というものが存在している。仮説を用いた思考法をモデル化し、計算機上で実現してやることで仮説発見の自動化が期待できる。情報技術が発達し、膨大な情報を扱わなければならない今日においては、発見の作業もコンピューター抜きにはすでに難しくなっており、発見の自動化について改めて考えることは重要である。この技術は、科学的発見のみならず、ビジネスアプリケーションや日常生活の多くの場面でも活用できる。

科学的思考法は、Francis Bacon (1561–1626) による帰納法による法則形成とその実証の導入に始まり、Claude Bernard (1813–1878) が実験の重要性を説いて整備されたが、観測、モデリング、仮説、予測、実験、観測、... のサイクルから形成される [2]。予想外の観測は科学の出発点であり、説明される必要があるが、そのためにはその観測が起こった世界と状況をモデル化し、その上でさらに仮説を導入する。仮説は新たな予測を導き、それは実験によって確かめられる。その際に新しい観測によって、仮説は採択されるか、修正されるか、棄却される。Karl Popper (1902–1994) のように、仮説は反証可能 (falsifiable) でなければならないとする立場もある。仮説の強さは予測性能 (predictive power) によって決まる。仮説を生成する思考法はアブダクション (abduction) として知られている [1]。一方、多くの観測事例から一般的規則を導く推論はインダクション (induction, 帰納) と呼ばれ、帰納的論理プロ

グラミング (ILP) を中心に研究されている [10]。

以下では、筆者がこれまでに関わってきた仮説発見に関する研究として、与えられた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システムの構築・実現、およびシステム生物学への応用に関する概要を述べる。

1.1 仮説発見システム

筆者はこれまでに岩沼・鍋島らと共同で、仮説を効率良く発見する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 [3] を効率的に実現した結論発見手続き SOLAR [11] を開発してきた。アブダクションは SOLAR を用いて自然に実現できる。一方、インダクションを完全に行うために、CF 帰納法 [4] と呼ばれる帰納仮説発見方法を提案した。これらの推論は、背景理論を伴った場合の機械学習方式であり、モデリングや知識発展を前提とする発見過程の自動化に適している。

SOLAR の改良を重ねた結果、現在 SOLAR の仮説発見能力は、一階述語論理上の仮説推論システムとしては世界最高レベル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。さらに SOLAR を用いて、メタレベル・アブダクションと呼ばれる、規則を生成することができる新たなアブダクション方式を提案した [7]。これはメタ推論においてアブダクションを実現しており、領域に関するメタ公理を与えれば、規則のみならず新たなオブジェクトも自動生成することができる。これを用いて、正負の効果を持つリンクを有するネットワークの補完方法を開発した [6]。またインダクションにおいては、CF 手続きの論理的簡略化と逆包摂による完全な手続きの設計 [19]、双対化計算の効率化の実現等を行った。

連絡先: 井上 克巳, 国立情報学研究所 情報学プリンシプル研究系, 〒101-8430 東京都千代田区一ツ橋 2-1-2, Phone & Fax: 03-4212-2642, Email: inoue@nii.ac.jp URL: <http://research.nii.ac.jp/~inoue>

1.2 仮説選択

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ために 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 [9]。これはさらに、条件付き確率に基づいたランキング方法に拡張された [15]。また、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新たな仮説知識の発見可能性を示した [12]。また帰納推論の下での異なる知識ベース間での予測性能の比較について検討した [13]。

1.3 システム生物学

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 [18]。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 [17]。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レベル・アブダクションによる解法を検討した [6]。この問題においては、パスウェイが何らかの理由で寸断された場合に、バイパスを計算することや、リンクが切断された物質とある種の類似構造をもつ化合物との間に p53 遺伝子から新たなリンクが形成可能であることを発見しており、生物のロバスト性を再現するための計算機実験にもなっている。

2. 回復の科学

上記最後の実験は、生物の冗長性とそれによるロバスト性を示しているが、これはシステム生物学というシステムの捉え方により得られた知見であった。それでは、システム一般について、擾乱に対するロバスト性や障害が起きても機能を回復できる能力について、科学的知見が何か得られないだろうか？このような問いは、自然災害などの想定外事象に対して、いかにしてシステムが新しい状態に遷移し機能を回復できるかという研究につながる。このための「システムズ・レジリエンス」[16]と呼ばれる新しい研究に携わっており、震災後の科学者のあり方として、新しい発見 (discovery) だけではなく、望ましい状態への回復 (recovery) のための科学が重要であると考えている。レジリエンスという概念はエコロジーに端を発するが、生命科学、物理学、工学、政治・経済学、社会科学など多くの領域で系の耐久性・ロバスト性を理解する概念となり得る。

発見過程では、既存の理論では説明できない事象に対して、それを説明する仮説を導入することで発見に結び付けた。この場合の事象は予期していなかったものであり、時間的には事象が現在のもので、仮説は過去に起こった事象または時間的に普遍的な規則になる。この推論を、逆に現在から未来へと進めると事象の予測に使える。予測不可能な事象はアブダクションでは説明できないものであり、セルオートマトン (CA) では到達不可能な「エデンの園」となる [14]。さらに、状態遷移の論理 [5] に基づき、遷移軌跡から状態遷移規則を学習するアルゴリズムを開発しブーリアンネットワークや CA の学習に応用した [8]。このように、説明と予測には対称性があり、今後は両面からシステムのダイナミクスを理解したいと考えている。

References

[1] 井上 克巳: アブダクションの原理, 人工知能学会誌, 7(1):48-59 (1992).
 [2] 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).

[3] Inoue, K.: Linear resolution for consequence finding, *Artificial Intelligence*, 56(2,3):301-353 (1992).
 [4] Inoue, K.: Induction as consequence finding, *Machine Learning*, 55(2): 109-135 (2004).
 [5] Inoue, K.: Logic programming for Boolean networks, *Proc. IJCAI-11*, pp.924-930, AAAI Press (2011).
 [6] Inoue, K., Doncescu, A. and Nabeshima, H.: Completing causal networks by meta-level abduction, *Machine Learning*, 91(2):239-277 (2013).
 [7] Inoue, K., Furukawa, K., Kobayashi, I. and Nabeshima, H.: Discovering rules by meta-level abduction, *Revised Papers from ILP 2009*, LNAI, Vol.5989, pp.49-64, Springer (2010).
 [8] Inoue, K., Ribeiro, T. and Sakama, C.: Learning from interpretation transition. *Machine Learning*, doi: 10.1007/s10994-013-5353-8 (2013).
 [9] Inoue, K., Sato, T., Ishihata, M., Kameya, Y. and Nabeshima, H.: Evaluating abductive hypotheses using an EM algorithm on BDDs, *Proc. IJCAI-09*, pp.810-815, AAAI Press (2009).
 [10] Muggleton, S., De Raedt, L., Poole, D., Bratko, I., Flach, P., Inoue, K. and Srinivasan, A.: ILP turns 20: biography and future challenges, *Machine Learning*, 86(1):3-23 (2012).
 [11] Nabeshima, H., Iwanuma, K., Inoue, K. and Ray, O.: SOLAR: an automated deduction system for consequence finding, *AI Communications*, 23(2-3):183-203 (2010).
 [12] Sakama, C. and Inoue, K.: Brave induction: a logical framework for learning from incomplete information, *Machine Learning*, 76(1):3-35 (2009).
 [13] Sakama, C. and Inoue, K.: Inductive equivalence in clausal logic and nonmonotonic logic programming, *Machine Learning*, 83(1):1-29 (2011).
 [14] Sakama, C. and Inoue, K.: Abduction, unpredictability and Garden of Eden, *Logic Journal of the IGPL*, doi: 10.1093/jigpal/jzt015 (2013).
 [15] Sato, T., Ishihata, M. and Inoue, K.: Constraint-based probabilistic modeling for statistical abduction, *Machine Learning*, 83(2):241-264 (2011).
 [16] Schwind, N., Okimoto, T., Inoue, K., Chan, H., Ribeiro, T., Minami, K. and Maruyama, H.: Systems resilience: a challenge problem for dynamic constraint-based agent systems, *Proc. AAMAS 2013*, pp.785-788, IFAAMAS (2013).
 [17] Synnaeve, G., Inoue, K., Doncescu, A., Nabeshima, H., Kameya, Y., Ishihata, M. and Sato, T.: Discretized kinetic models for abductive reasoning in systems biology, *Revised Selected Papers from BIOSTEC 2011*, CCIS, Vol.273, pp.141-154, Springer (2013).
 [18] Yamamoto, Y., Inoue, K. and Doncescu, A.: Integrating abduction and induction in biological inference using CF-induction, in: *Elements of Computational Systems Biology*, pp.213-234, John Wiley & Sons (2010).
 [19] Yamamoto, Y., Inoue, K. and Iwanuma, K.: Inverse subsumption for complete explanatory induction, *Machine Learning*, 86(1):115-139 (2012).