

氏 名 山本 泰生

学位（専攻分野） 博士（情報学）

学位記番号 総研大甲第 1385 号

学位授与の日付 平成 22 年 9 月 30 日

学位授与の要件 複合科学研究科 情報学専攻
学位規則第 6 条第 1 項該当

学位論文題目 Research on Logic and Computation in Hypothesis
Finding

論文審査委員 主 査 教授 井上 克巳
教授 佐藤 健
教授 龍田 真
教授 山本 章博 (京都大学)
教授 佐藤 泰介 (東京工業大学)

論文内容の要旨

The thesis studies the logical mechanism and its computational procedures in hypothesis finding. Given a background theory and an observation that is not logically derived by the prior theory, we try to find a hypothesis that explains the observation with respect to the background theory. The hypothesis may contradict with a newly observed fact. That is why the logic in hypothesis finding is often regarded as ampliative inference.

In first-order logic, the principle of inverse entailment (IE) has been actively used to find hypotheses. Previously, many IE-based hypothesis finding methods have been proposed, and several of them are now being applied to practical problems in life sciences concerned with the study of living organisms, like biology. However, those state-of-the-art methods have two kinds of fundamental limitations. The first limitation is related to the incompleteness for finding hypotheses. Most of the previously proposed methods restrict their search spaces due to computational efficiency. However, this fact brings an inherent possibility that they may fail to find such sufficient hypotheses that are worth examining. On the other hand, complete methods, which enable to find every hypothesis, have to deal with a huge search space. Usually, those complete methods are achieved by consisting of many non-deterministic procedures, each of which imposes users on several selections for finding relevant hypotheses. The second limitation is related to non-determinisms in those complete methods.

In terms of the first problem, the thesis actually provides a practical problem, where the incomplete methods cannot work well, to show the necessity for preserving the completeness in hypothesis finding. In terms of the second problem, the thesis proposes an alternative approach for finding hypotheses, which is based on the inverse relation of subsumption, instead of inverse entailment. The proposed approach can be used to simplify every IE-based method by reducing their non-determinisms without losing completeness in hypothesis finding. The thesis then focuses on CF-induction, which is an IE-based complete method in full clausal theories. Based on our proposed reduction way, the thesis logically reconstructs the current procedure of CF-induction into a more simplified one, in order to reduce the non-determinisms in CF-induction.

The thesis is organized as follows. In Chapter 1, we first review the historical background on explanatory induction as well as inductive logic programming (ILP). We then introduce recent application problems in life sciences, which have been actively studied in ILP. We also point out again two problems above that the modern ILP methods have. Chapter 2 reviews the notions and terminologies in thesis, which include the syntax and semantics in first-order logic, clausal forms, and consequence finding as well as the dualization problem. Chapter 3 reviews the principle of inverse entailment and introduces each previously proposed method based on inverse entailment including CF-induction. Chapter 4 provides a new practical application in systems biology. Its task is to find dominant enzyme reactions in metabolic pathways. Together with a

logical model on enzyme reactions, we estimate possible states of those chemical reactions based on abductive inference, and also complete some causal relations that are lacked in the prior model. We show how this advanced inference integrating abduction and induction can be realized using CF-induction. Chapter 5 shows that generalization based on inverse entailment can be logically reduced into generalization based on inverse subsumption while it ensures the completeness for finding hypotheses. In Chapter 6, we propose two approaches for generalization of CF-induction. The first one is based on a new deductive operator, which can be regarded as simplifying the existing operators of subsumption, resolution and weakening. The second one is based on inverse subsumption that is introduced in Chapter 5. We then show both approaches preserve the completeness of CF-induction and study their differences using several examples. Chapter 7 concludes and describes future works

Through the thesis, we will see underlying nature and insights to overcome limitations in the current IE-based hypothesis finding procedures. Especially, the logical reduction from inverse entailment to inverse subsumption contributes to lightening a huge search space and non-determinisms that the IE-based complete methods have to deal with. It would be significant towards hypothesis enumeration that we logically reconstruct the current procedure of CF-induction into a more simplified one. Through the case study, we also show new applicability of hypothesis finding methods.

博士論文の審査結果の要旨

博士論文は、説明的帰納推論による仮説発見に関する論理と計算に関するものであり、逆伴意法と呼ばれる計算方式の非決定性と探索空間の削減を目的としている。逆伴意法において仮説発見に対する完全性を保証するためには多くの非決定的手続きを扱う必要があり、本論文では、逆包摂法を用いる方法と単一の演繹操作を用いる方法を提案し、それぞれが仮説発見に対して完全であることを証明している。

ある背景理論と観測事例が与えられたとき、説明的帰納推論では、背景理論を伴って観測を論理的に説明する仮説を発見することがタスクとなる。この仮説発見を実現するために、逆伴意法に基づく手続きが従来数多く提案されており、近年、システム生物学などの生命科学分野の実問題に応用され、生化学の知識ベースに欠落している理論補完に利用されている。

しかしながら、従来の逆伴意法に基づく仮説発見技法は2つの問題点がある。第1の問題は仮説発見に対する不完全性に関するものである。多くの仮説発見技法は、計算効率のため仮説の探索空間を限定している。不完全な仮説発見手続きでは、検証に値するにもかかわらず原理的に発見困難な仮説を持つ可能性がある。その反面、完全な仮説発見手続きは、非決定的手続きから構成されているため、仮説を得るには多くのユーザーによる選択が必要となる。第2の問題はこの手続きの非決定性に起因するものである。

本論文では、第1の問題に関して、従来の不完全な仮説発見技法では解けない応用事例を実際に与え、実応用の観点から仮説発見に対して完全性を保証する必要性を示している。また第2の問題に関して、逆伴意法に基づく任意の仮説発見手続きが、その完全性を損なうことなく、より単純な逆包摂法を用いた手続きへと変形できることを示している。さらに、従来手続きの一つで仮説発見に対して完全である CF 帰納法に焦点をあて、包摂関係を用いた従来手続きの簡略化技法を提案し、CF 帰納法の非決定性を削減している。

第1章では、説明的帰納推論および帰納論理プログラミングの歴史的背景を述べ、近年活発に進められている生命科学分野への応用事例を紹介し、従来の仮説発見技法に関する上記の問題点を示している。

第2章では、本論文を通して用いられる帰納推論の基礎解説および用語の定義を与えている。

第3章では、従来までに提案されている逆伴意法に基づく仮説発見手続きを整理している。特に第6章で用いる CF 帰納法に関しては、いくつかの例を用いて手続きの仮説発見手順を具体的に示している。

第4章では、生化学ネットワークの主要反応経路を推定する応用問題を扱っている。この問題に対して、CF 帰納法が、状態推定だけでなく背景理論に欠落していた酵素反応に関する因果規則を導出できることを示している。反応状態推定と因果規則の補完の同時に実現することは従来の不完全な仮説発見手続きを用いることが難しく、この結果は CF 帰納法による理論補完の新しい応用可能性を示している。

第5章では、逆伴意法における一般化手続きが、仮説発見に対する完全性を損なうことなく、捜査がよりシンプルである逆包摂を用いた一般化手続きに変形できることを示している。また逆伴意法の逆融合操作がどのように逆包摂法に組込まれるかについて、いくつ

かの例を用いて解説している。

第6章では、CF 帰納法の一般化手続きが、逆包摂法を用いる方法と単一の演繹操作を用いる方法でそれぞれ簡略化できることを示し、それぞれの方法が発見する仮説の特徴を考察している。さらに CF 帰納法で用いる橋渡し理論の非決定的な構築法に関して、特徴節を漸次的に追加していく手続きを提案している。

第7章では、結論と将来研究について述べている。

上記のように本博士論文は、仮説発見に関する基礎理論の発展に貢献するものである。とくに、逆伴意法から逆包摂法への論理の変形は、逆伴意法に基づく完全な仮説発見手続きが直面する莫大な探索空間と非決定性の軽減に寄与する。またCF帰納法の論理的再構築は、これまで困難とされた仮説枚挙の実現可能性を示しており、当該研究分野の発展に大きく貢献する。さらに生命科学分野における新しい応用事例を通して、仮説発見技術の応用展開の方向性を示したことは意義があると認められる。

博士論文審査会においては、まず論文の概要について約50分の説明が行われた。博士論文の構成にしたがって、まず仮説発見法に関する従来研究の動向および問題点が示された後に、逆伴意法とCF帰納法の仮説発見手続きが説明された。続いて本研究のメインである、(1) CF 帰納法による生化学ネットワークの主要反応経路の推定、(2) 逆伴意法から逆包摂法への論理的簡略化、(3) 逆包摂法と演繹操作に基づくCF帰納法の論理的再構築、の3つのテーマについてそれぞれ説明があり、これらの研究を通して得られた知見についてのまとめと考察が示された。

公開質疑では、それぞれのテーマに関する質疑・コメントがあった。テーマ(1)については、生化学反応の因果関係のモデル化に関する一般的なコメントがあった。またテーマ(2)については、逆包摂法による仮説に関する質問と極小補における不動点定理の持つ意味に関するコメントがあった。またテーマ(3)では、逆包摂と演繹操作に基づく2つの一般化手法の比較考察に関する質疑・コメントがあった。それぞれの質問に対して、出願者から適切な回答がなされた。引き続き行なわれた非公開審査では、出願者より、予備審査における指摘事項についての対応と変更箇所について説明が行なわれた。

発表及び質疑応答と博士学位請求審査論文の内容に基づき、審査員全員で審議した結果、本論文は博士論文としての基準を満足しているとの結論になった。

出願者は論文中で、仮説発見に関わる論理的基礎、計算手続き及び応用に至る幅広い関連研究を俯瞰するとともに、仮説発見の新しい基礎理論を構築している。またシステム生物学における仮説発見のような萌芽的研究領域に対して今後の方向性を示す研究成果を残している。審査会の場における口頭試験を兼ねた質疑の内容からも、情報学分野の十分な専門的知識を有していると認められる。また博士論文および国際会議での発表などの実績から、研究者として相応しい英語の学力があるものと判断される。

以上から、出願者は、博士(情報学)の学位を受けるのに十分な学識と研究能力を有するものと認め、合格と判定した。