

書 評

古川康一、尾崎知伸、植野 研 著：帰納論理プログラム、共立出版 (2001)

筆者の一人古川氏は第五世代コンピュータの研究リーダーとして、また我が国における帰納論理プログラム研究のリーダーとして世界にその名を知られている。本書には古川氏の帰納論理プログラムを広く世に知らしめようとする思いが強く込められている。難解な論理学から攻めるのではなく、一般の研究者にとってなじみやすい決定木の学習が命題論理による分類問題の学習と同じであり、帰納論理プログラムはその自然な拡張になっていることを基調とする本書の流れは、読んでいて非常に自然である。

本書は導入、演繹推論、帰納推論、付録の4部から構成されている。本書のタイトルでもある帰納論理プログラムは第8章(第3部帰納推論)になって初めて出てきており、それまでに必要最小限の知識が自然に得られるように配慮されている。このことが逆に、帰納論理プログラムを機械学習全体の枠組みの中で、また論理を用いた推論の枠組みの中で、適確に捉えることを可能にしている。

第1部は導入部であり上に述べた著者らの姿勢がまとめられている。第2部の演繹推論では、最初に、命題論理で、命題文の意味とは何かを、モデルや解釈の説明を通して理解させ、エルブランモデルを考えればよいことを述べている。さらに、選言標準形、連言標準形、命題文の分類(恒真、恒偽、妥当、矛盾)、融合演算、演繹定理などの基礎的な知識を整理して、述語論理への橋渡しをしている。次に、述語論理が命題論理の拡張であり、命題論理で説明して来たことがそのまま成り立つことを述べ、限量化、単一化、融合法の健全性と反駁完全性、包摂の概念など述語論理解に必要最小限の項目を述べ、最後に論理プログラミング言語 Prolog を解説している。簡単な例題を用い、実際に Prolog により演繹推論がどのように実行されるかを要領良く説明している。再帰呼出し、差分リスト、メタプログラミングなどの考えもここでわかる。

第3部が本論の帰納推論である。常識的には、第2部の述語論理を受けて、すぐに帰納論理プログラムの説明に入るところであるが、決定木の説明から始まる。決定木の処理が命題論理の分類問題を解いていることと等価であることを説明し、トップダウンの分割統治法、情報量基準に基づく属性選択の指標、過学習回避のための枝

刈り、そのための真の誤差の悲観的推定法、bagging や boosting のコミッティ学習と、主要な話題を概説する。ついで、計算論的な学習理論で、極限における同定、文法推論、Shapiro のモデル推論、PAC 学習に触れ、帰納推論プログラムの説明に入る。この時点では読者はすでに必要な概念は習得しており、目標概念、事例、背景知識、仮説、被覆、伴意、包摂、相対包摂、相対最小汎化などの諸概念がすなおに受け入れられ、集合被覆アルゴリズムで仮説空間を探索する仕組みが理解できる。さらに、仮説の評価に関連し、分割表を用いた χ^2 検定や交差検定、正規分布の点推定と区間推定を説明した後、帰納論理プログラムの応用例として、電子メール分類、突然変異誘発化合物の部分構造推定、連水管の定性モデルの構築の三つの具体的な問題を取りあげている。最後の今後の展開では、記号を主体とする帰納論理プログラムの中での数値データの取扱い、正例だけからの学習、多述語学習問題、発想的帰納推論、伴意からの学習とは別の解釈からの学習に基づく帰納論理プログラムの手法、帰納論理プログラムの計算量、確率論理プログラムの学習など最新の研究成果にも、簡潔ではあるが、言及しており、これらに対する参考文献へのポイントもしっかりしている。

第4部の付録には Progol の具体的な使い方と、Progol における最弱仮説の生成法と仮説空間の探索法、Progol における仮説の評価法と正例からの学習などがまとめられている。

本書は事前の深い知識なしでも読める。難しい証明はいっさいない。といって、重要なものをカットしているというのではない。難しい証明を要するものは、簡単な例を示すことにより、読者が直感的に理解できればよいとの姿勢を貫いている。帰納論理プログラムは、表現力は豊富であるが、時間がかかりすぎる、調整するパラメータが多く使いにくいなどの批判をよく聞くが、本書を読めば、どの程度のことができそうかが正しく評価できるようになる。文系の学生や帰納論理プログラムという言葉にしきいの高さを感じている理工系の学生には入門書として、帰納論理プログラムを一応心得ている学生や研究者にも頭の整理に、一読を推薦したい本である。

[元田 浩 (大阪大学産業科学研究所
知能システム科学研究部門)]