

4 Q-5 帰納と対話の併用による概念知識ベースの構築*

野沢 慎吾 遠藤 聡志 大内 東

北海道大学工学部

1 はじめに

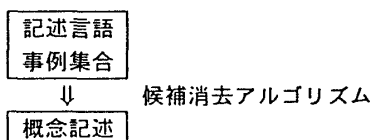
近年、エキスパートシステムをはじめとする知的システムの研究は理論的な実験段階から、実用段階に入ってきた。しかし、そのほとんどは、適用範囲が初期状態に与えられた知識ベースに依存してしまう。そのため、帰納学習の技術を取り込んだ、効率の良い知識獲得機構の確立に多くの期待が寄せられている。

本稿では、帰納学習によって概念を形成する Mitchell のバージョン空間法を取り上げ、これを拡張した実用的な知識獲得機構を提案する。

2 帰納的概念形成

バージョン空間法は Mitchell によって提案された概念形成技法で、候補消去アルゴリズムと呼ばれる仮説空間探索法により漸進的な帰納学習を行う [1]。

形式的には、バージョン空間法による概念形成の問題は次のように表すことができる。



記述言語とは、事例や概念を記述するために予め用意された、属性を一般性に基づいた木構造で表したものである。事例集合は、属性を要素に持つ特徴ベクトルで表現された対象概念の例の集まりで、対象概念に当てはまる正例と、当てはまらない負例が存在する。バージョン空間法は、これらを受け取り、候補消去アルゴリズムを用いて対象の概念記述、すなわち、すべての正例を満たし、すべての負例を排除する特徴ベクトルを形成する。

3 概念知識ベースの構築

現在、我々は診断型や分類型のエキスパートシステムを想定し、これらにおいて汎用的に利用可能な知識ベースを構築することを目的に、バージョン空間法を基にした知識獲得機構の実現を検討している。

3.1 従来法の問題点

バージョン空間法は、代表的な帰納学習技法であるが、理論的な側面が強く、実用的な知識獲得機構に取り入れるには以下のような問題点がある。

- 正しい学習を行うために、記述言語を用意する必要がある。しかし、知識獲得機構のためにこのような膨大で完全な背景知識を、予め内部に準備しておくことは困難である。
- 獲得可能な概念は単一の特徴ベクトルで表現できなければならないが、実世界における概念をこのような形式で表現するのは難しい。

これらの問題を解決するために、我々は以下のような拡張をバージョン空間法に施す。

- 背景知識である記述言語と学習の結果得られる概念記述を、概念知識ベースとして統合。
- 記述言語の不足を対話機構によって専門家から補充。
- 候補消去アルゴリズムを拡張し、選言的概念記述にも対応。

我々が提案する概念知識ベース構築システムは以下のように構成される。破線枠が従来のバージョン空間法を示している。

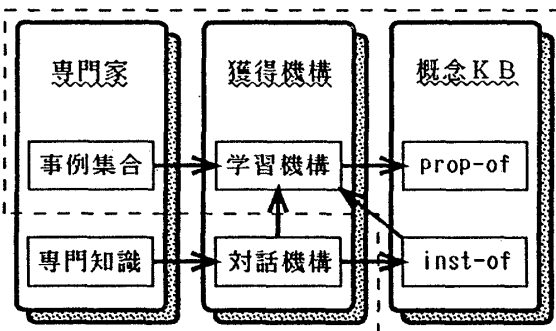


図1: 概念知識ベース構築システム

3.2 概念知識ベース

概念知識ベースは、個々の事象をノードとし、それらを instance-of 関係、property-of 関係という2つの階層的な

*A Method for Conceptual Knowledge Base Building
by Induction with Interaction
†Shingo NOZAWA, Satoshi ENDOH and Azuma OHUCHI
‡Faculty of Engineering, Hokkaido University

アークでつなぐことにより、複雑な概念ネットワークを構成する。

instance-of 関係とは、[クラス] $\xrightarrow{\text{inst-of}}$ [メンバ] 間の関係で、上位層の具体例が下位層となるような状態を表す。一方 property-of 関係とは、[クラス] $\xrightarrow{\text{prop-of}}$ [特性] 間の関係で、ある概念の持つ性質を表現する。この2つの関係が性質継承によって互いに絡み合い、複雑な知識構造を表現する。

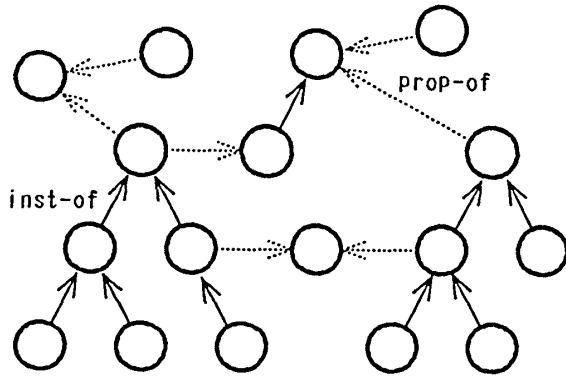


図2: 概念知識ベース

前述のようにバージョン空間法では、記述言語として属性の一般性に基づいた完全な木構造が必要である。この木構造は、instance-of 関係で構成されるものである。一方、学習の結果得られる概念記述は、同様に property-of 関係で表現されるものである。このように概念知識ベースは記述言語と学習の対象概念を統合管理できる。

また、従来記述言語は固定的なものであったが、本システムでは拡張性のある知識ベースの一部に過ぎない。その結果、記述言語の自由度を高め、広範な知識獲得を可能にする。

3.3 選言的概念記述

従来の方で得られる概念記述は、属性の連言で表現された1つの特徴ベクトルである。実世界において、複数の特徴ベクトルを論理和で結んだ選言的概念記述は必要不可欠である (Ex. "apple" \equiv [red,sweet] \vee [green,sour])。

単に選言的な表現を受け入れると、前述の候補消去アルゴリズムはまったく一般化を行わなくなる。S集合は今までに与えられたすべての正例を論理和で結んだ記述となってしまう。

この問題に対する解決策は、既にいくつか発表 [2,3] されているが、現在我々は次節で述べる対話機構と協調して働く選言的概念形成アルゴリズムを検討している。

3.4 対話機構の併用

完全な記述言語の木は、バージョン空間法が正しく学習するために必要な背景知識である。しかし、この膨大な背景知識を予め用意しておくことは大規模な知識獲得においては非現実的で、背景知識が不足する状態でも都度それを補って学習を進める機構が必要である。

我々は専門家との対話にこの問題の解決策を見出そうとしている。事例から仮説一般化、特殊化する際、専門家へ

の問い合わせを発行し背景知識の不足をカバーする。また、問い合わせから新たに知り得た知識も、都度、知識ベースに貯える。すなわち、帰納による property-of 関係の学習と、対話による instance-of 関係の獲得が効率良く同時に行われる。

例えば、下図のように、仮説 B に対し正例 C が与えられた場合、従来のバージョン空間法は、A へ一般化する。しかし、D が正例でない場合、この処理は過度な一般化である。すなわち以下のような場合に対処できない。

- 記述言語が不十分で、本来は B,C を受け入れ、D を排除するようなノード E が存在する場合。
- 目的の概念が選言的な表現を必要とする、つまり、BVC が本来の概念記述である場合。

本システムでは、与えられた正例に未確定な“兄弟”ノードが存在する時、一般化したノードを仮に追加する。対話機構によってこのノードが実在するものか専門家に問い合わせ、実在する場合、概念知識ベースは補充される。実在しない場合は、仮のノードのまま選言的な結合として扱われる。

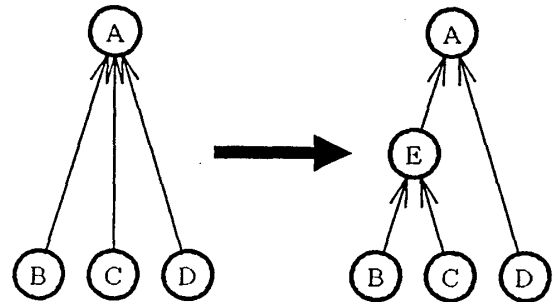


図3: 対話を併用した一般化

4 おわりに

現在、インプリメントに向けてアルゴリズムの詳細を具体化している。今後は、さらに実用性を追求し、知識表現の汎用化、事例に含まれるノイズへの対処などを検討する予定である。

参考文献

- [1] Mitchell, T.M., "Version Spaces: A Candidate Elimination Approach to Rule Learning," *Proceedings of the Fifth IJCAI*, 1977
- [2] Dietterich, T.G., "Learning and Inductive Inference," in *The Handbook of Artificial Intelligence*, Vol. 3, Cohen, P.R., and Feigenbaum, E.A. (Eds.), 1982
- [3] Murray, K.S., "Multiple convergence: An Approach to Disjunctive Concept Acquisition," *Proceedings of the Tenth IJCAI*, 1987