

動的環境下における不完全な領域知識からの説明と失敗に基づく知識獲得†

4M-2

宮井将之 上原邦昭††

神戸大学 工学部 情報知能工学科

1. はじめに

多くの学習システムでは、与えられた状態が変化しない、静的な事例からの学習を目的としている。これに対して、本稿では、競合する相手からのインタラクションにより状態が変化する、動的な事例からの学習について述べる。本システムでは、学習対象として連珠をとりあげ、詰め連珠を解決するための戦略知識を獲得することを目的としている。

2. 学習システム EbSL

本研究では、時間系列を伴う問題領域において、問題解決時に生じる時間的因果関係を利用して、新たな知識を獲得するシステム EbSL (Explanation based Strategy Learning system) を提案する。EbSL の学習は「説明に基づく一般化」に基いている。つまり、EbSL は目標概念の概念定義を事例中の行動がどのように満たしているかを説明し、それを一般化しながら知識を獲得するというアプローチ [1] を採用している。また、獲得した知識を新しい問題に適用し、目標を達成できないという失敗に対しては、失敗からの学習によって制御知識を獲得するようにしている。

3. 戦略知識の適用と失敗

EbSL の適用領域の例として、詰め連珠の問題をとりあげる。まず、5 問ほどの詰め連珠の問題解決事例から戦略知識を学習し、新たな詰め問題の解決を行なうと表 1 に示す結果が求められる。表 1 に示すように、EbSL は詰め連珠の問題の最良手を示す模範解答と一致している割合は多くないが、問題解決を行なう戦略知識を少ない事例数から獲得できていることがわかる。解決できなかった問題について検討すると、適用できる戦略知識が不足している場合と、連珠のよう

表 1: 詰め連珠に対する結果

戦略知識の数	34
問題の数	14
解決した数	7
模範解答と一致した数	3

な他者が存在する問題領域を対象としているため、図 1（自分が黒、相手が白）に示すように、自分は 15 手目の戦略知識で三を作り、次に四三を作る状態であったものが、相手が四を連続的に作る戦略知識を適用した結果、負けてしまうといった場合に分けられる。

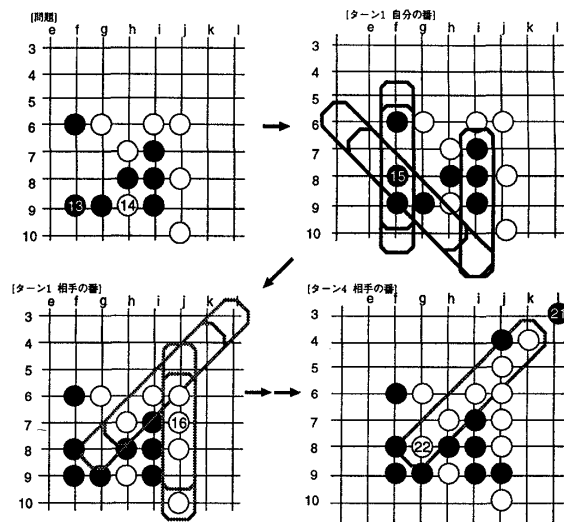


図 1: 失敗の例

前者の失敗は、多くの問題解決事例から学習して多様な知識を獲得すれば解決可能である。後者の失敗は、相手が選択する戦略知識よりも弱い戦略知識を適用したため生じたものである。そこで、ある状態とそれに時間的に連続する状態から得られる戦略知識の間には、図 2 に示すような強弱関係があることに注目し、相手が適用する可能性のある戦略知識を考慮して、戦略知識の適用を制御することを考える。つまり、失敗し

†Explanation and Failure based Learning from Imperfect Domain Knowledge on Dynamic Environment

††Masayuki Miyai and Kuniaki Uehara
Department of Computer and Systems Engineering, Faculty of Engineering, Kobe University
1-1 Rokkoudai, Nada, Kobe 657, Japan

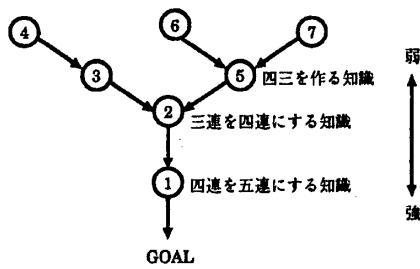


図 2: 戦略知識の関係

た状況でお互いに適用した戦略知識間の強弱関係を記録しておき、同様の状況に直面した場合には、他の戦略知識を適用して同じ失敗を回避するようなアプローチを考える。その結果、表 1 での段階では解決できなかった問題の中でも、再試行では失敗を回避して問題を解決できるものがあり、表 2 に示すように、より多くの問題解決が可能となった。

表 2: 失敗を考慮した結果

戦略知識の数	34
問題の数	14
失敗した後に解決した数	2

なお、表 2 の結果のうち、失敗を引き起こした後の行動を検討すると、

- 失敗したあとの再試行で問題を解決する。
- 同じ失敗を繰り返す。

という場合に分けられる。

このうち、同じ失敗を繰り返す状況を図 3 を例にして説明する。図 3 では、ある問題解決のために、戦略知識 8、戦略知識 5 を順に適用したとき、相手が戦略知識 2 を適用したために、先にゴールを達成してしまうという失敗が生じている。このとき、戦略知識 2 に対する戦略知識 5 の強弱関係を残すだけでは、再試行でも最初に戦略知識 8 を用いてしまい、同じ失敗を繰り返すことになる。したがって、戦略知識の強弱関係に加えて、状況に応じて戦略知識の適用を制限する知識を獲得する必要がある。

この場合、自分は戦略知識 8 と戦略知識 5 を順に適用して、相手はそれぞれに対して防御を行なう間に戦略知識 2 が適用可能な状況になってしまっている。これは、戦略知識 8 と戦略知識 5 を適用した箇所に、相手の戦略知識 2 を構成する要素が存在するためである。そこで、戦略知識 8 を適用する際に、その後の自分の行動（この場合は戦略知識 5）と相手の状態の中で戦略知識 8 の適用箇所に重なっている部分を考慮して、相手が戦略知識 2 を適用可能かどうかという、状況に応じた知識を制約として獲得すればよいことになる。その結果、最初に戦略知識 10 を適用して、相手は戦略知識 2 を用いることができないまま、自分がゴールを達成できるので、同じ失敗を繰り返すことを防ぐことが可能になる。

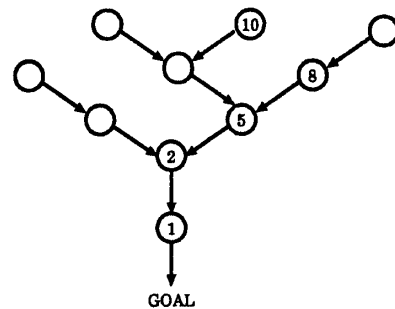


図 3: 失敗を繰り返す状況

4. 結論

本稿で提案した EbSL は、時間系列を伴う事例からの戦略知識獲得だけでなく、問題解決を行なったときの失敗の経験を通して、戦略知識の選択や行動の決定を制御するための制約条件を獲得することが可能である。その結果、戦略知識を抽出する時点では考慮できなかった、外部のインタラクションを考慮した行動決定が行なえるようになっている。

参考文献

[1] Epstein, S. L. : Learning Plans for Competitive Domains, *Proc. of the 8th Int. Conf. on Machine Learning*, pp.190-197 (1990).