

共進化を利用した共同作業における 知識の発展のモデル化に関する研究

A Study on the Modeling of the Development of Knowledge in Collaboration based on the Co-Evolution

研究代表者 大阪大学大学院基礎工学研究科 助手 吉田哲也
Research Associate, Graduate School of Engineering Science
Osaka Univ.
Tetsuya YOSHIDA

Abstract: The development of knowledge in collaboration was modeled based on the idea of "co-evolution" in biology by focusing on the diagnostic knowledge. The diagnostic knowledge held by each member in a group was structured into the decision trees and algorithms were designed so that the difference in the usage of symbol (word) between the members in a group could be detected based on the difference in the structure of decision tree. Experiments on motor diagnosis cases were carried out and the result indicated the effectiveness of the proposed method.

1 研究目的

近年では専門分野の異なる人間が共同で作業をする場合が多くなっているが、共同作業の初期の段階においては、個人の背景知識や専門の違いから互いに言っていることが理解できないために誤解が生じて相互理解が困難となることが多い。しかし、共同作業を進めるにつれて徐々に相互理解が深まっていくことが見られる。これは、共同作業に参加するメンバーが経験を通じて少しずつ共同作業が行いやすいように各自の知識体系を発展させていくためであると考えられる。

本研究では、上記のように共同作業を通じてメンバー間の相互理解が少しずつ深まっていくという現象に対して、共進化を利用して共同作業における知識の発展のモデル化に取り組んだ。ここでは特に意思決定における判断を表す診断知識を対象として、診断において用いられるシンボル（言葉）により表現される意味がどのようにメンバー間で相違が存在するのかを検出し、またそのすりあわせを行うための方法論の開発に取り組んだ。

2 研究経過

2.1 方法

一般的には、人間はある事象を捉える場合、何らかのイメージをもって認識すると考えられる。認識されたイメージには個人差があるため、言葉（シンボル）を用いて表現してコミュニケーションする場合においても、そのシンボルで指し示されるイメージには相違が存在すると考えられる。本研究はこのようなシンボルレベルでの相互理解がどのように発展していくかのモデル化に焦点をあて、次の2種類の相違を扱うこととした。

相違1 違う言葉を同じ意味で使っている場合

相違2 同じ言葉を違う意味で使っている場合

事象を表現するシンボルとしてクラス・属性・属性値を想定した。それぞれのシンボルに相違の可能性があると考えられるが、属性値に対してはこの2種類の相違が同時に発生することになるため、本研究では合計5種類の相違を扱うこととした。図1に実装したシステムのアーキテクチャを示す。

知識体系における相関関係

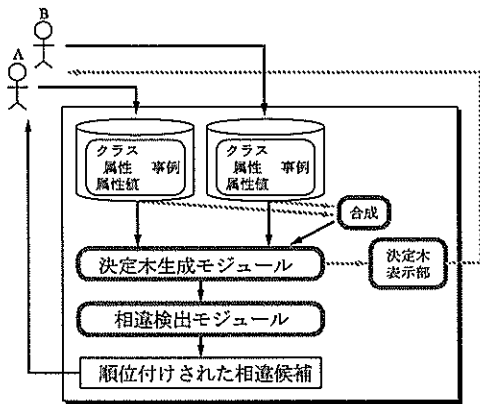


図 1: システムアーキテクチャ

各人が内的に持つ意思決定に対する知識体系を決定木として表現することとした。決定木とは、それぞれのノードに意思決定の判断材料となる属性を有し、属性の値に応じて木の枝に沿って進んでいくことにより最終的な意思決定を下すものである。なお、最終的な意思決定が下されるノードはリーフと呼ばれる。

本研究では共同作業中の相互理解促進が必要になる領域を対象としており、相互にある程度共有している知識があると考えられる。共有されている知識を表現する事例は相互に同様の意思決定を表現するものであるはずなため、内的な知識の外化手段として表現された事例を互いの決定木によって診断することにより、複数の決定木間におけるリーフ同士の相関関係を明確化することとした。この結果構築されるリーフ同士の相関関係をペアと呼ぶこととした。

ペアに基づいた相違検出を下記のように行うこととした。

- Step.1 自分の事例を相手の決定木に適用してクラス決定を行う。
- Step.2 リーフ同士に作成された関係をペアと呼び、事例の数だけペアを生成し、ペアの集合を生成する。
- Step.3 決定木間の関係（ペア）の集合を相違の典型例と比較する。

クラスシンボルにおける相違検出

この場合、クラスにおいて「互いに違うシンボルを用いているが、同じ意味を示している」シンボル

が使用されていることになり、この相違は決定木においてリーフの部分に現れることになる。具体的な事例と決定木の例を図2に示す。クラスCcとクラスCdにこの相違が発生している場合、両者の決定木構造に表れるのはリーフでのシンボルの違いとして浮かび上がってくる。両者の事例を互いに適用すると、クラスCcに決定される事例の集合は全てCdに対応し、クラスd4に決定される事例の集合は全てCcに対応することになるため、このような状態を検出すればクラスシンボルにおける相違が検出できることが明らかとなった。

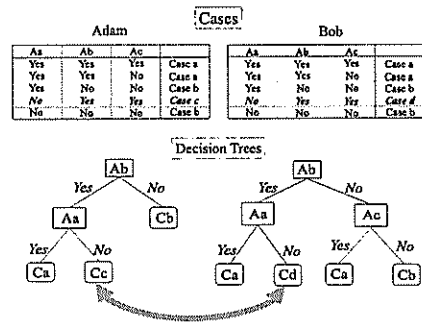


図 2: クラスシンボルにおける相違がある場合のペア

上記のアイデアに基づいて検出アルゴリズムを以下のように設計した。

- Step.1 あるペアを構成するクラスXとペアの対のクラスYが異なっているものを調べる。
- Step.2 そのクラスYに対応するペア先のクラスX'を調べる。
- Step.3 全てのペアに対してX=X'であれば、相違の可能性があると考えられる。

Step.3において、全ての場合に成立するのは典型的な例のみであるので、ロバスト性を持たせるために、全体の一致度として、全クラスに対して

$$\frac{\text{アルゴリズムに適合したペア数}}{\text{そのクラスに存在する総ペア数}}$$

を計算し、値が高いほど相違の可能性が高いと解釈することとした。

他の相違に対しても同様に、決定木の構造に着目して検出アルゴリズムを設計した。

表 1: 相違が単独で存在する場合の再現率と適合率

種類	再現率	適合率
クラスシンボルにおける相違 1(C1)	100%	52.6%
クラスシンボルにおける相違 2(C2)	100%	40.0%
属性シンボルにおける相違 1(A1)	100%	92.3%
属性シンボルにおける相違 2(A2)	100%	8.3%
属性値シンボルにおける相違 (V)	100%	29.3%

2.2 結果および考察

UNIX 上で C++ 言語を用いたプロトタイプを実装し、これを用いて提案した共同作業におけるユーザの知識の発展モデルにおいて、ユーザ間に存在する相違がどのように変化していくかに対する実験を行った。ここでは

- 概念相違が単独で存在する場合
- 概念相違が複数で存在する場合

に対して実験を行った。まず相違が単独で存在する場合を想定し、このような場合に相違検出アルゴリズムを適用することにより、単独でのアルゴリズムの検出能力と他の相違の種類を誤認識しない能力を測定した。次に 2 種類の相違が混在する場合に対して実験を行いアルゴリズムの検証を行い、互いの干渉が知識の発展プロセスに与える影響を調べた。

評価実験に用いる対象事例として図 3 に示すようなモータの故障診断例を用い、クラスシンボル 5 種類・属性シンボル 6 種類・属性値シンボルに 2-3 種類を用意し、事例数はそれぞれ 50 とした。この場合、相違の種類がシンボル別に 5 種類存在し、それぞれのシンボルで相違が発生することを考慮して、作成される事例集合として 28 種類¹を考慮した。そこから無作為に 2 種選択して混在する場合の評価実験を行ったが、決定木構造によって検出精度は若干変化する傾向があることを考慮して、実験では相違の種類を無作為に 1 種抽出することで決定木構造を変化させることにより知識形態を進化させることとした。

以下、クラスシンボルにおける相違 1、クラスシンボルにおける相違 2、属性シンボルにおける相違 1、属性シンボルにおける相違 2、属性値シンボルにおける相違をそれぞれ C1、C2、A1、A2、V と呼ぶこととする。

相違が単独で存在する場合

¹図 3 より、クラス 5 種類 × 2Type + 属性 6 種類 × 2Type + 属性値を含む属性で 6 種類 = 28 種類

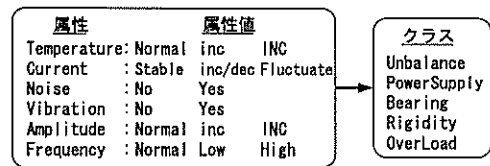


図 3: 分類型知識の事例

実験結果を表 1 に示す。この表より、相違は全て検出可能であるが、どこに相違が発生しているかの発見が困難である場合も存在することを確認した。

相違が複数混在する場合

実験結果を表 2, 3 に示す。用意した 28 種類の事例集合の中から各相違の種類別に一つを選択し、その事例と未選択の 27 種から一つずつ選択して混在させた場合の実験を行った。表中の数値は再現率（もれなく検出する能力）の場合には検出された候補に正しい相違が含まれている確率を示し、適合率（正確に検出する能力）の場合には検出された候補中に含む正しい相違の割合を示す。表 2 より、特に C1, C2, A1 などは他の相違と混同した場合においても検出精度が高いということを検証した。また、表 3 から A2 に対する検出アルゴリズムが特に他の相違に影響を受けるということを確認できた。

まとめ

表 1, 2 の再現率より、本研究で提案したアルゴリズムはほぼ正しく動作することを確認した。特に、クラスの相違に対しての検出精度が良いが、これは決定木間の相関関係を積極的に利用することの効果だと考えられる。また、属性や属性値などの決定木中のノードで発生する相違に対しても、一部検出精度が悪い部分はあるものの、概ね検出可能であることを確認した。

しかし、表 1, 2, 3 の適合率については、ほとんど低い値となっている。これは

1. 相違の候補数を多くあげている。

表 2: 相違が複数混在する場合の再現率

相違の種類	C1	C2	A1	A2	V
^{C1} Unbalance	100%	100%	87.5%	91.7%	54.2%
^{C2} PowerSupply	100%	100%	100%	100%	100%
^{A1} Amplitude	100%	100%	100%	50.0%	100%
^{A2} Noise	100%	100%	14.3%	50.0%	50.0%
^V Current	50.0%	50.0%	90.0%	40.0%	0.0%

表 3: 相違が複数混在する場合の適合率

相違の種類	C1	C2	A1	A2	V
^{C1} Unbalance	66.6%	57.1%	75.0%	15.9%	28.9%
^{C2} PowerSupply	50.0%	40.0%	63.6%	14.0%	50.0%
^{A1} Amplitude	100%	47.6%	100%	8.5%	38.5%
^{A2} Noise	25.6%	24.4%	14.3%	5.3%	9.4%
^V Current	33.3%	29.4%	56.3%	5.4%	0.0%

2. 検出目標以外の相違に影響を受けている。

によるものと考えられる。

3 研究成果

個人の知識を表現した決定木構造を保存して決定木間の相互関係を構築することにより、共同作業中における個性を残した相互理解のための知識の発展のモデル化を提案し、それを利用してメンバーで用いられるシンボル（言葉）の意味における相違を検出し、その相違を少しずつ取り除いてメンバー間の意思疎通を向上させていく手法を提案した。

4 今後の課題と発展

本研究で提案した知識の発展プロセスのモデル化は、各人の知識体系を保存して極力個性を残していくという利点はあるものの、構築した決定木間の相互関係を積極的に利用して更にダイナミックに共同作業に適した形に知識形態を変化させていくということに関しては十分なものではない。今後は、提案した決定木間の相互関係をより積極的に利用してダイナミックが知識形態が変化していくプロセスのモデル化を試みる。さらに、大量の知識を扱えるようにシステムを発展させていきたい。

5 発表論文リスト

1. Discovering Conceptual Differences among People from Cases, Yoshida, T., Kondo, T. and Nishida, S., The First International Conference on Discovery Science, pp. 162 - 173, December, 1998.
2. 複数の人間の間の概念相違発見手法の高速化の検討, 吉田哲也, 裏谷治, 近藤輝幸, 西田正吾, 人工知能学会基礎論研究会 SIG-FAI-9803, pp. 97 - 100, 1998年12月.
3. 決定木間のリンクを利用した概念相違発見手法, 大西健介, 吉田哲也, 西田正吾, 人工知能学会知識ベースシステム研究会 SIG-KBS-9803, pp. 69 - 74, 1999年3月.
4. Discovering Conceptual Differences among Different People via Diverse Structures, Yoshida, T., Kondo, T. and Nishida, S., The Third Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD-99), pp.494 - 498, April, 1999.
5. Evolving Granules for Classification for Discovering Difference in the Usage of Words, Yoshida, T., Kondo, T. and Nishida, S., The Seventh International Workshop on Rough Sets, Fuzzy Sets, Data Mining, and Granular-Soft Computing (RSFDGrC'99), November, 1999, accepted.