

# 一般化状態における協調と競合のためのエージェント内部機構

山川宏<sup>\*1</sup>

yamakawa@flab.fujitsu.co.jp

岡田浩之<sup>\*1</sup>

okada@flab.fujitsu.co.jp

渡部信雄<sup>\*1</sup>

aiwata@flab.fujitsu.co.jp

<sup>\*1</sup>新情報処理開発機構 自律学習機能富士通研究室

**Abstract:** CITTA is developed as a pattern based distributed intelligent architecture suitable for learning. CITTA uses generalized state vector, which integrates sensory input and action output. Proposing cooperation and negotiation mechanism generate action goal based on utility function each of which is derived from declarative knowledge and procedural knowledge in the agent.

## 1. はじめに

パターンベース推論を目指したエージェントネットワーク型のアーキテクチャの一種である Cognition based Intelligent Transaction Architecture (CITTA: 認識に基づく知的処理アーキテクチャ)はエージェント内学習のみならずエージェント間学習を可能とすることを目指し開発を進めている。そのため、一つにはパターン処理をベースとすることでセマンティクスを伴うシンボルを主体の情報処理における自律的学習能力の限界を回避する、一つには後述するようにセンサ入力と動作出力を区別せずに一元的に扱う一般化状態を用い[1][2]。そして現在は、エージェント間学習(例えばMatchability基準による状況分解[3])によって新たに生成されたエージェントを、エージェント内の学習により動作可能とするために、エージェント内部の記述の統一化を図りつつある。

今回はその一環として、既に研究を進めていた協調機構[1]に対して競合解消機構を追加する。この際に後述するゴール伝播の双方向性問題が現れる。提案する協調競合機構はエージェント内の手続的知識と宣言的知識の両方を効用関数として表現し、それに基づき確率的に動作することで、ゴール伝播の双方向性問題を解決する。本報告ではまず、効用関数を基に協調競合のためのゴール生成を行う機構を提案する。

主な関連技術との比較を行う[4]。黑板モデルが扱う課題では複数のエージェントが共通のあるいは両立し得るゴールをもつことで何ら制約もなく協調できる前提が存在するが、我々の扱う課題ではそのような前提は存在しない。契約ネットプロトコルとはタスク分解やエージェント毎の効用関数を持つ点で類似しているが、ゴール伝播の双方向性問題が無い点や入札機構に相違点がある。ゲームと交渉の研究とはエージェント毎に効用関数がある点で類似するが、多数エージェントによるタスク分解などの議論は少なく、ゴール伝播の双方向性問題が考慮されない点で相違する。一方、制約伝播の

研究は、宣言的知識によるゴールの伝播の点で類似しているが、エージェント毎の効用関数が存在しない点で相違する。

## 2. CITTAと一般化状態

### 2.1 CITTAの構成

本アーキテクチャにおいてエージェント間で交換される情報は実数値のベクトルのみであり、各要素毎に値の存在を表示する有効値(Valid/Invalid)が付加されている。エージェント間の接続関係は短時間では固定的であるが学習により比較的緩やかに変化することが想定されている。

エージェント間の接続を通じた情報交換はレイヤ構造により多線化された3種類のノードを通して行われる。これによりエージェントは自身が接続しているB-node上のゴールの書き込み主や、入札状態を知ることができる。

- (1) B-node: 自身で生成管理する情報を発信する
  - (2) L-node: 他エージェントのB-nodeに接続して参照する
  - (3) S-node: 活性度などの特定の意味付けされた情報発信
- ノード毎の3種類のレイヤは以下の通りである。

- (1) State layer: 現在の状態を表わす
- (2) Goal layer: 実現したい状態を表わす
- (3) Contract layer: 入札を表わす(有効値のみ)

各エージェントの基本動作は外部から得られた状態とゴールに応じて、自身のノードの一部にゴールを出力することである。そのためゴールレイヤのノードではしばしば値が存在しないInvalidの状態になる。

### 学習

に適したパターンベースの分散知能アーキテクチャであるCITTAにおいては、以下のような点が考慮されている。第一に、明示的なセマンティクスを用いないので、情報の対応関係は接続構造に埋め込まれる。そして同一のノードを複製できないので、B-nodeの表現はすべてユニークである。このため、ある表現の利用は必ずL-nodeによるB-nodeの参照となり、B-node同士の接続は不可能である。第二にレイヤ構造が与える意味付けはB-nodeの表現とは直交して与えられ、同じノード上の異なるレイヤの情報は対応関係を持つ。第三に人為的な命令信号の利用によりセマンティクスが導入されることを防ぐため、動作命令にはすべてゴールを利用する。

### 2.2 一般化状態とその意義

上記のB-nodeとL-nodeは一般化状態でありセンサ入力と行動出力を区別せず扱う。その代わりB-nodeやL-nodeの各ノード毎のレイヤ構造が状態やゴールの意味を与える。

一般化状態を利用する意義は:(1)センサ入力はその操作性によって動作出力から区別されるが、CITTAの協調機構はすべてのゴールに解決の可能性を与えるのでその区別は無意味である。(2)情報構造の獲得は動作出力情報に比べて圧倒的に量が多いセンサ入力を利用し、動作生成時にゴールレイヤ

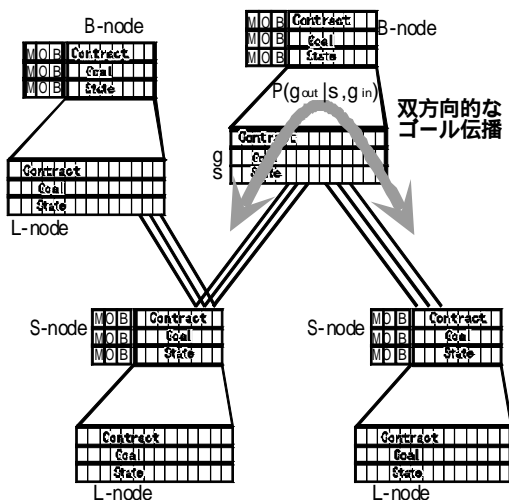


図1 CITTAの構成図

を通じてその情報構造を流用することにある。ここで情報構造はエージェント間接続として保持される。

### 3. 協調競合機構への要請

#### 3.1 競合解消とエージェント毎の効用関数

エージェント間でのゴール出力の競合を解消するには、関連するエージェントがゴールを評価する必要がある。

行動により獲得される手続的知識がエージェント内に貯えられている場合、手続的知識は状態からゴールへのマッピングであるから、これを状態から効用関数へのマッピングと効用関数からゴールへのマッピングと見なすことは可能である(強化学習で良く見られる定式化)。つまりエージェント内の手続的知識はエージェント内の効用関数に結びつく。

次に観察を通して獲得し得る宣言的知識がエージェント内に貯えられている場合について検討する。他のエージェントから得られたゴールを変換する宣言的知識は、状態の関係を記述しているが効用を保持していない。ここで効用関数をエージェント間で伝播させることは難しく、CITTAにおいても伝播されるゴールに効用は付加されていない。このため宣言的知識により生成されたゴールには評価が与えられず、ゴールの不整合による競合解消を実行できない。そこで4節では宣言的知識によるゴールの変換もエージェント毎の効用関数に基づくように定式化することを試みる。

宣言的知識に対して効用関数を設定することは単純なルールを置換えることを考えると不自然に思えるが、宣言的知識を経験から学習する場合には、経験の量や確率的な挙動などが効用関数の存在を十分に意味あるものとするであろう。また手続的知識と宣言的知識の両方をエージェント毎の効用関数を介して一元的に実現すれば、宣言的知識に基づく行動によって起こる知識の手続化の過程を自然に取り扱えることが期待される。

#### 3.2 ゴール伝播の双方向性問題と確率的な協調競合機構

あるエージェントの協調機構は、接続先のB-nodeに他のエージェントが書き込んだ入札されていないゴールが存在すると、そのゴールを入札するように動作を試みる。一方、競合の発生は既に他のエージェントがゴールを書き込んだノードへのゴールの書き込みである。つまり何れの動作も他のエージェントが書き込んだゴールが起動条件になる。

そのため、協調機構と競合機構を混在させると、入札相手と競合相手を一意に決定できないゴール伝播の双方向性問題が発生する。この問題は主にエージェント内の宣言的知識によってゴールが生成される場合に相当し、入札の方向を明言する契約ネットプロトコルでは発生しない。4節ではこの問題を解決するエージェント内部機構として、確率的にゴール出力を行う動作機構を提案する。

### 4. エージェント毎の協調競合機構

各エージェントのL-nodeにおいて、協調と競合のために確率的にゴールを出力する機構を提案する。あるエージェントの、無出力状態(Invalid)も含むL-nodeの状態空間を $S$ としノード毎の要素は離散的な値を取るものとする。状態レイヤのノード値によるベクトルを $s \in S$ とし、 $i$ 番目のノード値を $s_i$ とする。同様にゴールレイヤについても、ベクトルを $g \in S$ としてその要素を $g_i$ とする。なお、 $s_i, g_i$ にはValidな値が書き込まれている場合と、Invalidな状態にある場合の両方がある。そして、他のエージェントから得られたゴール $g^{in}$ と自身で書き出したゴール $g^{out}$ を区別すればエージェントの動作は $s$ と $g^{in}$ から $g^{out}$ へのマッピングであるから、効用関数 $U(s, g^{in}, g^{out})$ を利用してゴール出力、 $P(g^{out} | s, g^{in})$ を決定するのが

自然である。しかし3種類の情報に基づく効用関数は大規模な学習や設計のコストが大きく不利である。そこで、2種類の情報による効用関数 $U(s, g)$ による動作機構を提案する。

効用が均一である場合の初期確率を $P_0(g^{out})$ とし、効用から確率への影響は、温度定数 $T$ を含む指数関数を用いて $M(u) = \exp(u/T)$ とする。さらに入力ゴールをアクセス (式1) 初期確率を $P_{0A}(g^{result})$ とればゴール生成確率は、

$$R(g^{out} | s, g^{in}) = P_0(g^{out}) \frac{M(U(s, g^{out}))}{\sum_{g^{out} \in S} M(U(s, g^{out}))} + a(g^{in}) \frac{P_{0A}(g^{result}) M(U(s, g^{result}))}{\sum_{g^{result} \in S} M(U(s, g^{result}))} \quad (式2)$$

$$N(s, g^{in}) = \sum_{g^{out} \in S} R(g^{out} | s, g^{in}) \quad (式3)$$

$$P(g^{out} | s, g^{in}) = \frac{R(g^{out} | s, g^{in})}{N(s, g^{in})}$$

となり、ここで、 $(g^{in}), g^{result}$ は全ての次式により求められる。

$$a(g^{in}) = \begin{cases} 0 & \text{if } \forall g_i^{in} = \text{Invalid} \\ 1 & \text{else} \end{cases} \quad (式4)$$

$$g_i^{result} = \begin{cases} g_i^{out} & \text{if } (g_i^{out} = \text{Valid}) \\ g_i^{in} & \text{if } (g_i^{out} = \text{Invalid}) \end{cases} \quad (式5)$$

本手法はゴールの入力と出力を区別せずに連想的な処理を行うので、あらかじめゴールの伝播方向を指定しなくても動作できる。このため、ゴール伝播の双方向性に対応できる。また2種類の情報による効用関数 $U(s, g)$ と動作した結果予測されるゴール状態 $g^{result}$ を用いてゴール出力を決定できる。

本手法では、他のエージェントにより書き込まれたゴールが、自身の評価から見ても良いほどゴールを出力しない確率が大きくなるので、結果として競合解消が実現される。

### 5. まとめ

エージェント間の情報交換にセンサ入力と動作出力を一元的に扱う一般化状態を用いるCITTAでは、各エージェントの処理に協調機構と競合機構を混在させた場合に、ゴール伝播の双方向性問題が発生する。そこで、これを克服するために、エージェント毎の効用関数に基づいて確率的にゴールを生成する方法を提案した。

確率的な機構を採用した理由の一つには、CITTAが移動ロボットの制御などの実時間問題に適用を重視していることに起因する。常に変化し続ける交渉条件へ対応するには、均衡点まで交渉を進めることや、一度提案したゴールをそれ以降は提案しないアルゴリズムは不向きである。また提案した確率機構は強化学習(特にQ-Learning) [5]に用いられるアルゴリズムに近いので、少なくともエージェント毎にマルコフ決定過程が成り立つ環境下では動作が安定すると思われる。

今後は、提案した協調競合機構の実験をCITTA上で行い改良を進める、さらに、競合解消の交渉中にそのゴールを利用する他のエージェントがいかに動作を決定するかもという問題などについても検討したい。

### 参考文献

- [1] 山川 他, "実世界において自律学習する分散知能アーキテクチャ," RWC'98, pp.253-258. 1998.
- [2] 末広, 高橋, 山川, "エージェント・ネットワークによる手渡しロボットシステムの構築," 第15回日本ロボット学会, Vol.2, pp.373-374. 1997.
- [3] H.Yamakawa, "Proposing Matchability Criterion for Situation Decomposition - Extracting situations each of which contains a rule -," ICONIP'98, To be appeared, 1998.

- [4] 石田 他, "分散人工知能", コロナ社, 1996.
- [5] L.P. Kaelbling et al, "Reinforcement Learning: A Survey," *J. Artificial Intelligence Research*, Vol.4, pp.237-285, 1996.  
参考URL: <http://www.rwcp.or.jp/people/yamakawa/>