

## 多重ゴール環境下での行動選択の振舞いについて

栗原 聡\*      岡田 美智男  
NTT 基礎研究所

### 概要

Knowledge-Based AI では、プランニングや推論、学習は中心的で計画性のある処理として扱われてきた。これに対し Behavior-Based AI では、中心的な処理を行なうモジュールが存在しないかわりに、分散されたエージェントの集合によってシステムが構成され、システム全体として理解、学習、などの能力を発揮する。現在、我々は自然な発話の生成、理解機構をモデル化する研究を進めている。モデル化するにあたり、人間の知的活動の本質を、「多数の自律的なエージェントの集合が創発的に振舞うようなシステムの挙動」と捉え、Behavior-Based AI のアプローチを適用することにした。

### 1 背景

現在、我々は自然な発話 (spontaneous speech) の生成、理解機構に関する研究を進めている [3]。これは、自然な発話、理解というものがある人間の認知活動 (知的活動) を解明する一つの手がかりになると思われるからである。我々は、人間の知能に対する認識を次のような枠組で捉えている。「知能は人間の頭の中にあるというよりは、人間と人間をとりまく環境とのインタラクションの中にあるのではないか」。人間がうまく生きていられるのは、人間をとりまく環境とうまく関係し合っているからである。これは、人間が環境に対して受動的な存在であると言っているのではなく、環境に対して能動的な存在であるところに知的な振舞いがあるのではないかとすることを意味している。

自然な発話の生成、理解機構をモデル化することを考えたときに、人間の知的活動の本質を環境とのインタラクションの中に見出すと捉え、Behavior-Based AI のアプローチが適していると考えられる。人間の知的活動の本質を、多数の自律的なエージェントの集合が創発的に振舞うようなシステムの挙動であると捉えるのである。

### 2 実装について

今回本研究を開始するにあたり、行動選択ネットワークのモデルとして Maes のモデルを参考にすることにした [1]。このモデルは、一つの Behavior (振舞い) を一つのエージェントに割り当てたマルチエージェントモデルである。エージェントの集合が協調し合うことで目的を達成する。

エージェントの内部動作は隠蔽されている。エージェント同士は predecessor, successor そして conflictor の 3 種類のネットワークリンクによって接続されている。ネットワークリンクは activation を拡散するために使われる。activation とは、この行動選択ネットワークを動作させるためのエネルギーである。

activation を拡散するダイナミクスは以下の 3 種類である。(1) ゴールは、ゴールを達成するのに有効なエージェン

トの activation を増加させる。(2) ある時点での環境からの入力 activation は、その時点でのシステムをとりまく環境とマッチするエージェントの activation を増加させる。(3) すでに達成されて、そのまま達成され続けられなくてはならないゴール (protected goal) は、そのゴールから activation を奪おうとするエージェントの activation を減少させようとする。この activation の流れは、エージェント全体のグローバルなゴールが達成されるまで続けられる。

以上を参考に実際に行動選択ネットワークを実装した。現在 HP9000/755 上で動作している。個々のエージェントの挙動をモニタするためのワークベンチも実装した。このワークベンチでは、エージェントのモニタ機能だけでなく、ネットワークの局所的な変更や、各パラメータの動的な変更ができるなど、種々のシミュレートができるようになっていく。

### 3 考察

#### 3.1 プランニングについて

行動選択ネットワーク内には、プランニングに相当するモジュールないしはエージェントは存在しない。個々のエージェントは、自分の近傍での環境との関係しか保持していない。それでもエージェントネットワークは、あたかもプランニングされたかのようにゴールに向かって挙動を示した。また、ネットワークは、動的な環境の変化にも柔軟に対応できることを確認した。

#### 3.2 多重ゴール環境での振舞い

今回、数種類の例題をシミュレートした。その結果、ネットワークは十分に多重目的を達成することが確認できた。これはすべての例題に共通した結果である。多重ゴールを達成するために重要なパラメータとしては、上述した中の「activation を減少させる比率」がその役割を担っている。多重ゴールは、複数のゴールが協調、競合することによって達成される。この時、複数のゴールが協調し合うように機能する。あるゴールが達成されても、最終的なゴール達成のために一度そのゴール達成を壊す場合もある。「activation を減少させる比率」は、一度達成されたゴールが壊されないようにするため比率であり、このパラメータ値が大きすぎると最終的なゴール達成ができず、activation 拡散が発振、発散してしまう。逆に小さすぎると、ゴールが達成されてもその状態を維持できず、逆に最終的なゴールも達成されない結果となった。互いに強く競合するゴールがない場合には、「activation を減少させる比率」はほとんど影響を及ぼさない。

ゴールからの activation の流入比率は、環境からの activation の流入比率よりも常に大きい方が全般的によい結果が得られた。ゴールからの activation の流入比率がゼロで

\*kurihara@atom.ntt.jp

あると、当然ゴール達成はされなかった。目的がなければエージェントは正しい挙動は示さないからである。

環境からの activation の流入比率は、ゴールからの activation の流入比率ほどはゴール達成には影響を及ぼさないという結果が得られた。これは、環境からの後押しがなくともエージェントが能動的に環境に対して働きかけることを確認したものである。

パラメータの値は、エージェントにどのようなゴールを達成させるかによって左右される。このことは、学習や記憶の問題を考える際にパラメータ設定の問題が大きく関係することを示すものである。

### 3.3 学習と記憶について

行動選択ネットワークにおいては、学習や記憶に関してどのように考えればよいのだろうか。学習の問題は、ネットワークが新しいゴールを達成する際に、一度成功した activation 拡散を、次回再び同じゴールを達成しようとする時に、いかに効率良く達成するかという問題に対応する。また記憶の問題は、一度経験した activation 拡散をどのように再利用できるようにするかという問題と対応する。従来の AI 研究では学習と記憶の問題は別々に考えられてきた傾向があったが、行動ベース AI では学習と記憶の問題は同時に扱う必要がある。我々の考え方では、記憶は、データベースと考えるのではなく、「エージェントたちの一部に対して、過去に何度か働いたのと同じように働かせるプロセス」と捉える。つまりこの行動選択ネットワークにおける記憶とは、ある経験をしたときに、その経験によって活性化されたエージェントと、活性化した順番などを記憶することであると考える。そして、記憶を取り出す(思い出す)という作業は、データベースからデータと取り出すのではなく、記憶した一連のエージェントの集合に対して activation を適切に与えてやることによって、ネットワークの状態を「過去に経験したのと同じような状態にすること」と考えるのである。

学習という作業は、この記憶のメカニズムを実際に働かせるための手段である。この手段には 2 種類の方法が考えられる。1 つは、一度経験した一連の activation の流れを効率良くするための手段である。これは、個々のエージェントが保持している閾値をうまく操作することで実現できそうである。このときに、この操作をメタレベルで行なうのか、それとも個々のエージェントが行なうのかなどの問題がある。2 つめの手段は、ある経験によって活性化された一連のエージェントの集合をエージェンシーとしてユニット化することである。この際に、ユニット化するものは、エージェントの他にパラメータなどの値も含まれるだろう。この考え方を実現することで、ネットワークを構成する単位としてエージェンシーとは別のエージェンシーを生成することができる。このエージェンシーとエージェンシーとの関係は Minsky の考え方に則したものである [2]。また、ユニット化することは、経験によって活性化されたエージェントと活性化した順番などを記憶するための手段であることであるから、これは K-line に相当すると考えられる。

## 4 課題

このモデルの問題点は、エージェントの挙動を左右するパラメータ設定の理想的な設定方法がないということである。パラメータ設定は、上述した学習のメカニズムと関連した問題であり解決しなくてはならない。初期のパラメータ設定は人間が行なうが、エージェントネットワークが経験を重ねていく過程で、パラメータに関しても学習することができ

ば、解決できる可能性がある。エージェントの仕様や activation 拡散方法などにおいてもさらに拡張を行なう必要がある。

#### [◆多重ゴール環境]

今回の例では、最終的なゴールは一つしかないが、我々の目標とするアーキテクチャには、この最終的なゴールに相当するものも複数存在する。普段は定常状態のようになっているが、環境などからの働きかけがトリガーとなって、activation 拡散が開始され、その時点での環境やシステムの状態において最適なゴールが達成される。システムの状態は動的に変化し、システムが選択するゴールも動的に変化していく。最終的には、多重ゴール環境とはこのような意味として使うことになる。

#### [◆信念エージェント]

我々が本来の多重ゴール環境を作り上げるためには、信念エージェントが必要である(従来のエージェントは行動エージェントと呼ぶことにする)。信念エージェントは、物理的環境と関係するのではなく、他のエージェントと関係する。行動エージェント系の動きに対してゴールを持つようなエージェントである。例えば、「喉が乾いた」という環境で、「急げ」というエージェントがゴールを達成すれば、「コップを使わずに水を飲む」ことになり、「行儀良くしろ」というエージェントがゴールを達成すれば、「コップを使って水を飲む」であろう。物理的環境と対するエージェントと、エージェントと対するエージェントがうまく協調することで、より柔軟な多重ゴール環境の構築が可能となると考えられる。

#### [◆ネットワークの双方向性の可能性の確認]

我々はこの行動選択モデルを、自然な発話の生成 / 理解機構モデルに応用することを考えているわけだが、現在生成系のモデルについて研究が主に進められている。ここで、行動選択ネットワークに双方向性の性質が認められれば生成系に用いたネットワークを逆向きに動作させることで、解釈系にそのまま適用できる可能性がある。

## 5 おわりに

現在、自然な発話の生成、理解機構をモデル化する研究を進めている。我々は人間の知的活動の本質を、多数の自律的なエージェントの集合が創発的に振舞うようなシステムの挙動であると捉え、モデル化には Behavior-Based AI のアプローチを適用することにした。今回は、Maes の行動選択ネットワークを参考に、実際にモデルを実装し、ネットワークの振舞い、多重ゴールでの行動選択の成り立ち、状況依存性などについて検証した。そしてプランニングや学習、そしてモデル化するために必要な拡張の問題などについて考察した。

## 参考文献

- [1] Maes, P. : How to Learn to Do the Right Thing, *The European ALIFE Conference* 1991.
- [2] Minsky, M. : *The Society of the Mind*. Simon and Schuster, New York, New York, 1986.
- [3] 岡田: 「聞き耳をたてるコンピュータ」, 竹内都雄編, 「AI 奇想曲」 NTT 出版 1992.