人工知能学会第二種研究会資料
汎用人工知能研究会
SIG-AGI-028-04

複雑な知能を生み出す極限まで簡素化された環境の設計

Designing an Extremely Simplified Environment to Generate Complex Intelligence

一杉裕志[1*] 高橋直人[1] 竹内泉[1] 佐野崇[2] 中田秀基[3,1]
Yuuji Ichisugi[1] Naoto Takahashi[1] Izumi Takeuti[1] Takashi Sano[2] Hidemoto Nakada[3,1]

[1] 産業技術総合研究所 [1] National Institute of Advanced Industrial Science and Technology
[2] 東洋大学 [2] Toyo University [3] 順天堂大学 [3] Juntendo University

Abstract: The prefrontal cortex is the part of the brain that realizes complex human-like intelligence. This paper reports on the current status of the design and implementation of an agent environment with the goal of facilitating the experimentation and development of AI that reproduces the functions of the prefrontal cortex. In this environment, time and space are discretized and the environment is simplified to the maximum extent possible by eliminating the concepts of coordinates and distance. At the same time, agents can communicate with each other using logical expressions, making it possible to experiment and develop complex intelligence.

```
Room1: Alice
Room2: Key
Room3: Box
    ↓        Key の場所に移動
Room1:
Room2: Key Alice
Room3: Box
    ↓        Key を Box の場所に持っていく
Room1:
Room2:
Room3: Key Box Alice
    ↓        Key を Box に適用して Gold を得る
Room1:
Room2:
Room3: Gold Alice
```

図 1: テスト用タスクのシナリオ

1 はじめに

自律 AI エージェントが複雑な知能を獲得するためには、エージェントが活動する**環境**が十分に複雑である必要があるだろう。しかし複雑すぎる環境は AI アーキテクチャの設計・開発・評価を著しく困難にする。我々の研究の目的は、ヒトに似た複雑な知能に関する実験を可能にしつつも、エージェントにとっての「世界」を極限まで簡素化した実験・開発フレームワークを設計することである。本稿では、このフレームワークにおけるエージェント環境の設計方針と、現在の実装状況について報告する。

*連絡先：産業技術総合研究所
茨城県つくば市東 1 − 1 − 1 中央事業所
E-mail: y-ichisugi@aist.go.jp

このフレームワークは、ヒトの脳の前頭前野の本質的な機能の再現という目標に特化し、不要な要素を極力排除して簡素化している。エージェントが活動する環境では、時間・空間は離散化され、物体は記号化されているが、世界の大きさや物体の種類の数には制限はなく、ヒトが生活する実世界の本質的特徴を残している。エージェントは他のエージェントと意思疎通を図ることも可能で、そのことが特に複雑な知能を生み出す重要な要因となり得る。

2 関連研究

強化学習の汎用的な実験・開発フレームワークとして、OpenAI Gym [1] がかつて広く使われ、様々なエージェント環境が提供されている。本研究ではこのような汎用的なフレームワークではなく、前頭前野の機能に関する実験・開発に特化したものを目指している。

本稿で提案するエージェント環境に近いものを持つ研究としては Generative Agents [2] がある。ヒトのようにエピソード記憶、熟考、内省の機構を持つエージェントがレトロゲームのように簡略化された世界の中で生活し、お互いに対話を行いながら、「パーティーを開く」といった与えられた目標を達成する。ただしこのエージェントが持つ知識は大規模言語モデルから得ているものであり、本研究のように自律的な知識獲得の実験を行う目的のためには環境が複雑すぎるように思われる。

簡素化されたエージェント環境が知能の本質の理解に

役立つよい例として、関連性理論をモデル化した例 [3] がある。ここでは古典的な AI タスク Wumpus world [4] を拡張した世界の中でエージェント同士が簡単な意思疎通をする。ヒトの発話と言語理解の本質的目的を POMDP の枠組みを用いて明確化することに成功している。

我々の以前の研究 [5] では、環境の状態を「成り立つ命題の集合」として定義した。これは我々が提案している、推論規則を自律的に獲得する機構を説明する目的に適した極めて単純なエージェント環境である。しかし単純すぎるため言語活動に関する実験・開発には適さない。

3 提案するエージェント環境

3.1 テスト用タスク

提案する環境の特徴を説明するためのテスト用タスク（図 1）についてまず説明する。AI エージェントである Alice が、ある部屋にある Key を別の部屋にある Box に適用することで、中から Gold を取り出すことがこのタスクのゴールである。

このタスクの前提として、 Alice は Key と Box がどの部屋にあるかを推論できるものとする。（そのために例えば自分自身の過去の記憶を想起するとか、他者に聞くといった手段を取り得る。）Alice は環境と相互作用するための**プリミティブ行動**として、部屋の中にある物体を探す (FindItem)、他の部屋に移動する (GoToRoom)、部屋内にある物体を他の部屋に持っていく (Bring)、部屋内の物体をもう 1 つの物体に適用する (Apply)、が実行できる。

このタスクの世界における**物理法則**では、 Key を Box に適用すると部屋内に Gold が発生し、Key と Box は消滅するものとする。

実装中の実験・開発フレームワークではより複雑なタスクも実現可能であるが、以下の節では説明の具体例として適宜このタスクを用いる。

3.2 部屋と物体

エージェントがいる世界は、複数の**部屋**から構成され、各部屋には 0 個以上の**物体**があるものとする。エージェントもまた物体の一種である。図 1 のシナリオの最後の状態は、Room1 と Room2 に物体がなく、Room3 には Gold と Alice が存在するという状態を表している。

このフレームワークでは、部屋内や部屋間の座標・距離の概念を排除することで、世界を極力簡素化している。（前頭前野は座標・距離を直接的には扱わず、視野内の座標は頭頂葉・前運動野・前頭眼野が処理するのではないかと考えたため、このような設計とした。）

```
1: rule(_,
     「Alice は PLS で PLS に注視している」, FindItem)
2: rule(「Alice は r で x に注視している」,
     「Alice は r で x と PLS に注視している」, FindSecondItem)
3: rule(「Alice は r で x と y に注視している」,
     「Alice は r で x を y に適用した」, Apply)
4: rule(_,
     「Alice は PLS から PLS に移動した」, GoToRoom)
5: rule(「Alice は r1 で x に注視している」,
     「Alice は x を r1 から r2 に持ってきた」, Bring)
```

図 2: プリミティブ行動を呼び出す行動ルールの例を疑似コードで示したもの。“_” は任意の状態にマッチするワイルドカード、小文字のシンボルは行動選択時のパターンマッチで値がセットされる変数、 “PLS” はサブルーチン終了時までに値がセットされるスペースホルダー。

一般には、エージェントは世界にあるすべての部屋を最初から知っているとは限らない。また、どの部屋からどの部屋に直接移動可能なのか、各部屋内でどのような物理法則が成り立つのかは、タスクの設定ごとに変わるものとする。

未実装であるが、部屋の中には他の AI エージェントが存在することもあり、相互に対話することができる。（プロトタイプ実装については [6] を参照。）対話は自然言語ではなく文の内容を表す論理式を直接やり取りする。これにより、脳において言語野が担う機能をシステムから排除し、前頭前野の機能の開発に集中できるようにしている。エージェントは、他のエージェントの行動を観測することもでき、模倣を通じた知識伝達の再現も容易に行えるはずである。

一般には環境の状態も自律的に変化し得る。定期的に食物が発生する農場、時間の経過による食物の腐敗、自然災害など、さまざまな自然現象を模した事象が、この簡素化された環境内でも表現可能である。自律的に知識獲得する AI エージェントがそのような環境内で自身の報酬を最大化するためには、複雑な知能を獲得する必要があるだろう。例えば、効率的な収穫戦略、将来に備えた備蓄、さらには他の地域にいるエージェントとの物々交換による経済活動などを発展させる必要がある。

3.3 Pro5Lang とプリミティブ行動

エージェントの脳内にある Pro5Lang プログラムがエージェントの行動を制御する。（Pro5Lang の詳細については [7] を参照されたい。）Pro5Lang プログラムは**行動ルール**の集合として定義される。Pro5Lang のプロトタイプ実装では、行動ルールは rule(s,g,a) とい

う構文で定義される。エージェントは、現在のレジスタの状態 s およびサブゴールの値 g とマッチする s,g のパターンを持つ行動ルールを、その価値の高さに応じて確率的に１つ選択し、その行動 a を実行する。

エージェントが環境と相互作用するためにはプリミティブ行動を実行する。図２は、 Pro5Lang プログラムがプリミティブ行動を呼び出すための行動ルールの例である。

プリミティブ行動は一般に次のように動作する。

1. 呼び出し時のサブゴールから必要なパラメタを取得する。
2. 環境を操作する。
3. 行動の結果をエージェントの a レジスタにセットする。

例えば図２の１行目の行動ルールは、「Alice は Room1 で Key に注視している」という命題をサブゴールとするサブルーチン呼び出しを行ったときに選択され得る。このサブゴールを見れば、 FindItem で探すべき物体が Key であることがわかる。FindItem が実行されると、エージェントの「身体」は現在の部屋 Room1 内に Key が存在するかどうかを調べる。存在するならば「Alice は Room1 で Key に注視している」、しないなら「Alice は Room1 で Key に注視していない」という命題が a レジスタにセットされる。

プリミティブ行動は様々な原因で異常終了することがある。例えば現在いる部屋は Room1 ではなく Room2 であるとか、部屋が真っ暗で何も探せない、といったことが起こり得るが、その状況を適切に表現した命題が a レジスタにセットされる。プリミティブ行動を呼び出したプログラムは、その値を見て例外状態に適宜対処できる。

一般に、プリミティブ行動により環境が変化するが、どのように変化したかという情報が a レジスタにすべて入っているとは限らない。例えば図２の３行目のプリミティブ行動 Apply の呼び出しは、a レジスタに「Alice は r で x を y に適用した」という情報をセットするが、適用した結果何が生じたかという情報はない。一般にはエージェントは適宜環境を観測し、必要な情報を能動的に取得する必要がある。

なお、 Pro5Lang を脳の計算論的モデルとみなす場合、プリミティブ行動は前運動野（視覚情報などを用いた習熟した運動の実行に関与する部位）が記憶しており、それを前頭前野（行動計画・推論などを実行する部位）のプログラムが呼び出すと解釈する。

3.4　テスト用タスクを解くプログラム

図３はテスト用タスクを解くフローチャートであり、図４と図５ほぼそれに相当する Pro5Lang プログラムの疑似コードである。

図４のメインルーチンは、まず「Key と Box を得る」という目的のために図５のサブルーチンを呼び出す。その後、 Key を Box に適用する。その直後はエージェントは Gold をまだ認識していないが、図２の１行目の行動ルールが実行されることで Gold が認識され、それによりメインルーチンのゴールが達成される。

図５の「Key と Box を得る」ためのサブルーチンの中には２つの条件分岐がある。１行目、２行目のサブルーチン呼び出しはそれぞれプリミティブ行動 FindItem, FindSecondItem を実行することになるが、その結果は目的の物体が部屋内に存在する場合としない場合とで異なる命題を a レジスタにセットする。その結果、１行目の次は２行目または６行目に、２行目の次は “ サブルーチンの終了” または３行目に分岐する。環境の初期状態が図１の先頭であった場合は、図５のプログラムは１→６→７→１→２→３→４→５→１→２→終了、というふうに実行が進む。

4　まとめと今後

前頭前野はヒトらしい複雑な知能を実現する脳の部位である。本稿では、前頭前野の機能を再現する AI の実験・開発を容易にすることを目標とした、エージェント環境の設計と実装の現状について報告した。この環境では、時間・空間は離散化されており、座標・距離の概念を排除することで極限まで簡素化されている。一方で、他のエージェントと論理式を使った対話が可能であり、それにより複雑な知能に関する実験・開発を可能にしている。現状の実装では対話、イベント時系列に関する推論、物体ＩＤと物体の内部状態に関する推論などの扱いが不完全であり、これらについては今後適宜拡張を行っていく。

本稿で述べた設計方針を推し進めることで、前頭前野の機能に特化した AI の実験・開発を容易にし、現状の大規模言語モデルの延長では到達し得ないような、ヒトのような知能が実現可能になると考えている。

謝辞

本研究は JSPS 科研費 JP22K12188 の助成を受けたものです。

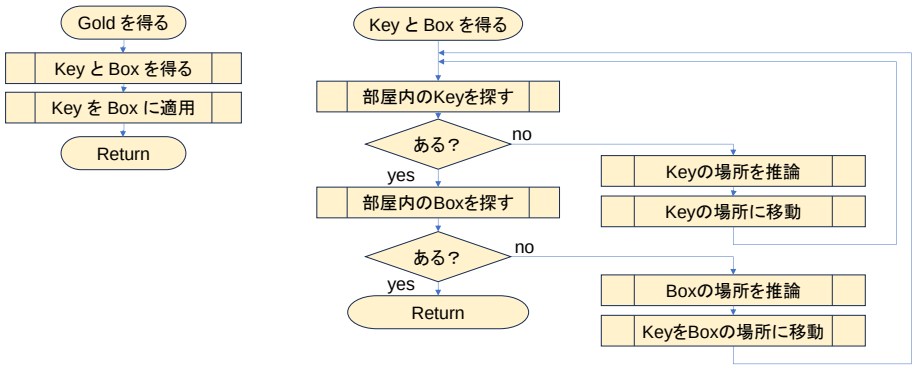

図 3: テスト用タスク「Gold を得る」を解くフローチャート

```
   g = 「Alice は PLS で Gold に注視している」
1: rule(_, g,
      call(「Alice は PLS で Key と Box に注視している」));
2: rule(「Alice は r1 で Key と Box に注視している」, g,
      call(「Alice は r1 で Key を Box に適用した」));
```

図 4: メインルーチンの疑似コード。「Gold を得る」、「Key と Box を得る」というサブゴールを、「Gold に注視している」、「Key と Box に注視している」という、より具体的な命題で表現している。

```
   l2g = 「Alice は PLS で x と y に注視している」
1: rule(_, l2g,
      call(「Alice は PLS で x に注視している」))
2: rule(「Alice は r1 で x に注視している」, l2g,
      call(「Alice は r1 で x と y に注視している」))
3: rule(「Alice は r1 で x と y に注視していない」, l2g,
      call(「y が PLS にある」))
4: rule(「y が r2 にある」, l2g,
      recall(「Alice は PLS で x に注視している」))
5: rule({「y が r2 にある」, 「Alice は r1 で x に注視している」}, l2g,
      call(「Alice は x を r1 から r2 に持ってきた」))
6: rule(「Alice は r1 で x に注視していない」, l2g,
      call(「x が PLS にある」))
7: rule(「x が r1 にある」, l2g,
      call(「Alice は _ から r1 に移動した」))
```

図 5: 「Key と Box を得る」という目的のために呼び出されるサブルーチンの疑似コード。 Key と Box は変数 x と y で表現し、汎用性を持たせている。

参考文献

[1] Greg Brockman, Vicki Cheung, Ludwig Pettersson, Jonas Schneider, John Schulman, Jie Tang, and Wojciech Zaremba. OpenAI Gym. *arXiv:1606.01540*, 2016.

[2] Joon Sung Park, Joseph O'Brien, Carrie Jun Cai, Meredith Ringel Morris, Percy Liang, and Michael S. Bernstein. Generative agents: Interactive simulacra of human behavior. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*, UIST '23, New York, NY, USA, 2023. Association for Computing Machinery.

[3] Kaiwen Jiang, Stephanie Stacy, Annya L Dahmani, Boxuan Jiang, Federico Rossano, Yixin Zhu, and Tao Gao. What is the point? a theory of mind model of relevance. In *Proceedings of the annual meeting of the cognitive science society*, Vol. 44, 2022.

[4] Stuart J Russell and Peter Norvig. *Artificial intelligence: A modern approach*. Prentice Hall, 2003.

[5] 一杉裕志, 中田秀基, 高橋直人, 佐野崇. 推論規則の価値を階層型強化学習 RGoal を用いて学習する手法の提案. 第 14 回 人工知能学会 汎用人工知能研究会 (SIG-AGI), 2020.

[6] 一杉裕志, 中田秀基, 高橋直人, 竹内泉, 佐野崇. 報酬最大化ＡＧＩのための意思疎通機構の設計とプロトタイプ実装. 第 21 回 人工知能学会 汎用人工知能研究会 (SIG-AGI), 2022.

[7] 一杉裕志, 中田秀基, 高橋直人, 竹内泉, 佐野崇. 汎用人工知能のためのプログラム合成対象言語 Pro5Lang のエピソード記憶機構. 第 20 回 人工知能学会 汎用人工知能研究会 (SIG-AGI), 2022.