

# 認知モデルをヴァーチャル世界に接地させるエージェント開発手法 Method of Developping Virtual Agent Grounding Cognitive Models

森田 純哉<sup>†</sup>, 長島 一真<sup>†</sup>, 竹内 勇剛<sup>†</sup>

Junya Morita, Kazuma Nagashima, Yugo Takeuchi

<sup>†</sup> 静岡大学情報学部

Faculty of Informatics, Shizuoka University

j-morita@inf.shizuoka.ac.jp

## Abstract

In this research, we attempted to integrate the cognitive architecture ACT-R and 3D game engine. We built a hierarchical architecture in which ACT-R and the game engine are connected through a black board server, and constructed a model for searching the 3D environment. The constructed model reproduced behavioral differences by parameters of the cognitive model. We also made interesting errors related to the brain-body connection. From these results, it is suggested that cognitive modeling method is useful to construct agents imitating humans behaviors in 3D space.

キーワード: ACT-R, virtual agent, game engine

## 1. はじめに

認知科学では、伝統的に、認知モデリングの手法によって、人工知能研究と心理学の融合が進められてきた。ここでいう認知モデルは、コンピュータ上に実装された人間の内部処理に関わる仮説である。通常の人工知能の研究に対し、認知モデルの研究は、心理学において見出された人間のエラーやバイアス（限定合理性）を再現することに重点が置かれ、心理実験の結果に対するシミュレーションによって評価される。

このような伝統的な認知科学のアプローチの他に、HAI (Human-Agent Interaction) と呼ばれる分野においても人工知能と心理学の融合が進められてきた。HAIにおいては、人間らしさ、あるいは生物らしさを感じさせるエージェントがコンピュータ上に実装され、それと相互作用する人間の反応が心理実験によって検討される。

近いトピックを扱っているにも関わらず、認知モデリングと HAI の間で、積極的な知見の交換は行われてこなかった。HAI で扱われる対人エージェントは、外見や社会的関係に焦点を当てるものが主で、人間の

内部処理を模倣しようとするものは稀であった。エージェントに対する人間の反応を主なデータとする HAI において、内部処理に深く踏み込むことは、研究戦略として合理的ではないのかもしれない。しかし、今後、HAI が従来よりも、複雑で長期間にわたる相互作用の系列を扱う際には、認知モデリングで扱われてきたような内部処理を含むエージェントの開発が求められていくはずである。

上記の認識から、著者らは、認知モデリングの手法を流用した対人エージェントの開発手法を検討する。本発表では、この目標に向けたアプローチ、および開発される基盤技術から導かれる研究課題を議論する。

## 2. HAI と認知モデリングの統合アプローチ

### 2.1 認知アーキテクチャの利用

近年の認知モデリングでは、認知アーキテクチャの役割が重要視される。認知アーキテクチャとは、個別の認知モデルの研究において開発された手法を統合する基盤である。共通の認知アーキテクチャを利用したモデルの開発を積み重ねることで、人間に普遍的な認知システムの構造に近づくと考えられている。これまでに複数の認知アーキテクチャが開発されてきた。本研究では ACT-R (Adaptive Control of Thought-Rational [1]) に着目する。ACT-R は、多くの研究者が参加するコミュニティにおいて開発されている。また、そのモジュールとパラメータを、脳の構造と生理機能に対応づける検討も進められている。また、オリジナルの ACT-R は Lisp で記述されているものの、Java[6] や Python[7] を含む複数のプログラミング言語による実装もあり、環境に応じた柔軟な開発が行えるようになっている。

## 2.2 ヴァーチャル世界との接続

ACT-R は、内部処理に関わるモジュールだけでなく、知覚や運動など環境との相互作用に利用されるモジュールを有している。しかし、それらのモジュールには物理的信号を取得するセンサ、あるいは物理的な世界に直接作用するアクチュエータは含まれない。つまり、ACT-R を利用した対人エージェントを構築するためには、ACT-R と接続する身体を別に用意する必要がある。過去、ACT-R をヒューマノイドロボットに実装し、実世界において人間と相互作用をさせる研究が行われている [2]。しかし、センサやアクチュエータなど現実環境とのインタフェースにおけるハードウェア上の限界もあり、その動作は限定的である。

本研究では、より簡便に、3次元空間として構築されたヴァーチャル世界と ACT-R を接続する。ヴァーチャル世界の構築には、ゲームエンジンを利用する。近年開発されている多くのゲームエンジンには、精巧な物理エンジンや身体モデルが含まれ、現実感の高い世界を構築できる。最近になって、それらの 3D 環境と ACT-R を繋げる複数の研究が現れ始めている。環境からの知覚に応じて単純な動作（回転など）を決定するロボットエージェント [3]、あるいはヴァーチャル世界におけるロボットの環境探索 [4] などが課題として取り扱われている。本研究はそれらの知見を援用しつつ適用範囲を拡張させる。

## 2.3 認知アーキテクチャとヴァーチャル身体の統合

ACT-R が接続するヴァーチャル世界では、複数の独立のイベントがリアルタイムに進行する。それに対して、ACT-R の内部で生じるプロセスは基本的には系列的である。よって、ACT-R とヴァーチャルな世界の統合には、異なる階層のプロセスを並行して動作させるサブサンクションアーキテクチャ [5] 的な枠組みが必要になる。つまり、ヴァーチャル世界における身体動作の制御を下位層、ACT-R による知識ベースの意思決定を上位層に対応させる。両者は定期的に通信を行いつつ、並行して動作する。上位層は、下位レイヤにて取得される知覚情報を入力としつつ、知識ベースの意思決定を行う。下位層は上位層の意思決定の結果をコマンドとして受け、低レベルの身体動作（歩く、姿勢を変える、回るなど）を遂行する。

## 3. プロトタイプシステム

### 3.1 アーキテクチャ

これまで、ACT-R (Python ACT-R) とゲームエンジン (Unreal Engine 4) をサーバを介して接続するプロトタイプシステムを実装している (図 1)。サーバは C 言語によって実装され、エージェントの動作コマンドを格納するスロット、および視界情報を格納するスロットをもつ。スロットの値は、ゲームエンジン、もしくは ACT-R からの定期的なソケット通信を介して更新される。過去の研究 [4] を参考に、通信におけるデータのフォーマットを JSON (JavaScript Object Notation) に統一している。



図 1 ゲームエンジンと ACT-R の接続

サーバ上のスロットに対する読み込みと書き込みは、ACT-R とゲームエンジンで対称的に行われる。3D 環境内におけるエージェントの一人称視点に基づいて、ゲームエンジンが、視界情報を更新し、その値を ACT-R が定期的に読みこむ。また、ACT-R の実行結果によって、サーバ上の動作コマンドが更新され、その情報を読みこむことによって 3D 環境内のエージェントの動作状態が更新される。ACT-R とゲームエンジンのサーバへのアクセスは非同期である。つ

まり，サーバを介在して，ACT-Rによる認知プロセスの実行と3D環境内における身体動作が，並列的に遂行される．このような高次の意思決定機構（脳）と低次の身体運動の関係は，階層的なサブサンクションアーキテクチャの枠組みと整合的である．

## 3.2 オブジェクト回収課題

### 3.2.1 課題

構築されたアーキテクチャを用い，単純な環境探索を行うエージェントを実装した．図2はエージェントが配置された3D環境を示している．左上に俯瞰的な映像が示され，右下にエージェントの視界情報が示される．この環境におけるエージェントの課題は，可能な限り短い時間の中で，環境内の全ての青色のオブジェクトを回収することである．ただし，今回のエージェントにおいては，オブジェクトの位置を結ぶ最短経路の探索は目指さなかった．エージェントは，各時点において，最も近いオブジェクトに向かう前向きの探索を繰り返した．

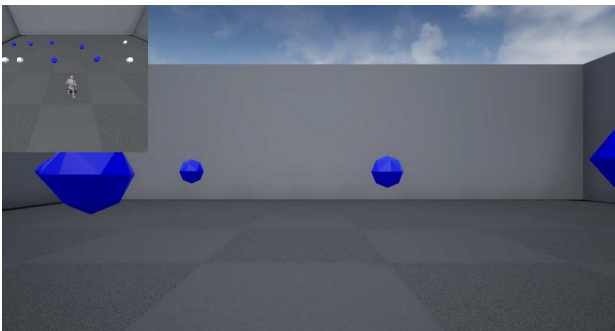


図2 オブジェクト回収課題の環境

### 3.2.2 エージェントの動作

図3はエージェントの動作をフローチャートで示している．各オブジェクトの回収に先立ち，エージェントは，身体を回転させ，環境中のオブジェクトを探索する．視界に入ったオブジェクトの一つに注意を向け，エージェントからの距離を知覚する．視界内に複数のオブジェクトが存在した時には，オブジェクトに設定された顕著性 (saliency) に従って注意を向けるオブジェクトを選択する．顕著性の設定は，今回のエージェントでは，視界に投影されるオブジェクトの大きさ（今回の課題は距離に対応）によって定まる．注意を向けたオブジェクトの距離に基づいて，エージェン

トはゴールバッファ内の「最短距離オブジェクト」を更新する．

視界内のオブジェクトを探索するルールと，現在までに取得された最短距離オブジェクトに移動先を設定し歩行するルールは，ACT-Rの中で競合する．つまり，最短ではないオブジェクトに誤って向かうエラーや，全てのオブジェクトの探索が終わった後にも環境内の探索を続けるなどのエラーが生起する．ACT-Rにおいて，これらのエラーの頻度は競合解消のノイズによってモデル化される．オブジェクト探索のルールのユーティリティ（優先度）をオブジェクトに向かうルールのユーティリティに比べて高く設定した際，ノイズが小さいエージェントは慎重に最短距離オブジェクトを見極め，ノイズの大きいエージェントは目についたオブジェクトに向かう短絡的な振る舞いを増やすことになる．

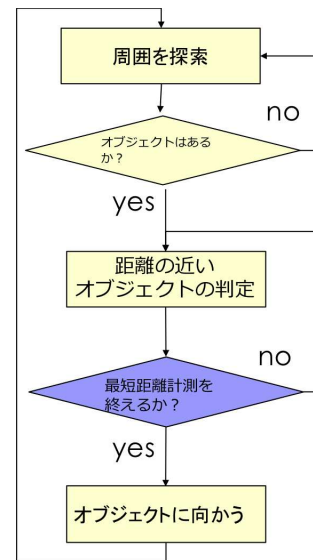


図3 エージェントの動作

### 3.2.3 予備的結果

図4は図3に示したエージェントの振る舞いから得られるシミュレーション結果を示している．ノイズの異なる熟考型エージェントと短絡型エージェントを用意し，オブジェクトの回収条件（オブジェクト数とオブジェクトの配置）を揃えた条件で実行に要する時間を比較した．その結果，設定された環境において，ノイズの少ない熟考型がノイズの大きい短絡型に比べ，課題を効率的に遂行することが示された．ただし，この結果はオブジェクトの回収条件やエージェントの動作速度などのパラメータによって変化する可能性がある

る。重要なことは ACT-R のパラメータを操作することで、3D 空間におけるエージェントの振る舞いに多様性を持たせることができたことである。

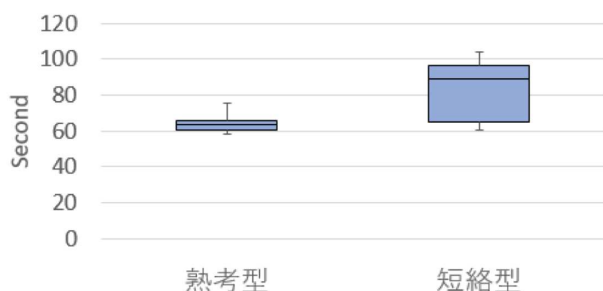


図4 エージェントの動作

上記に加え、予備的シミュレーションにおいては、アーキテクチャに固有のエラーも見出している。本アーキテクチャにおいては、ゲームエンジンと ACT-R の間で、サーバを介した定期的な通信が行われている (図1)。この通信において、環境の変化を適切なタイミングでサーバにアップロードすることに失敗、あるいはサーバ上の視覚情報が ACT-R が取得する前に更新されてしまったなどの原因によるエラーが生じる。こういった通信エラーが生じたとき、本研究のエージェントは、オブジェクトを取得したという自分の振る舞いに自覚できず、環境中のオブジェクトを探し続けるなどの振る舞いを示す。工学的には、このようなエラーは修正すべきバグと見做される。しかし、認知モデリングにおいては、生じたエラーと人間との対応を検討することが重要と考える。本件に関し、生態心理学の分野においては、本人が意識しない行為のエラーをマイクロスリップとして注目してきた。生態心理学において指摘される人間のエラーと本アーキテクチャによって生成されるエラーの類似は、本アーキテクチャの構造に認知モデルとしての一定の妥当性があることを示している。

#### 4. まとめと展望

本研究では、認知アーキテクチャである ACT-R とゲームエンジンによって構成された 3D 世界を統合する仕組みを構築した。ゲームエンジンと ACT-R を接続する試みは本研究において初めて行われたわけではない。ただし、従来の研究 [3, 4] は、ACT-R とゲームエンジンを Pear to Pear で結合し、図1のようにサーバを介して結合するものではなかった。前節の最後で述べたマイクロスリップは、ACT-R とゲームエンジンが並列的に動作するからこそ生じるエラーである。

このように本研究において構築したアーキテクチャ

は、従来の研究では見落としていた認知プロセスのモデル化を導く可能性がある。従来の認知モデルの多くは身体を持たず、基本的には脳内に閉じた記号系の問題を扱ってきた。認知モデルに対して、ヴァーチャル空間の身体を付与することで、身体と脳のインタラクションに関する重要な現象をシミュレーションできる可能性がある。加えて、本研究のアーキテクチャは、マルチエージェント環境への拡張という点でも利点を有している。この利点を活かし、今後ヴァーチャル世界に複数のエージェントを配置した、身体化された集団のインタラクションのモデル化を行う予定である。

さらに、ゲームエンジンによるヴァーチャル世界の可視化は、ACT-R によって稼働するエージェントと人間との相互作用を行えるようにするという利点も有している。こういった ACT-R による対人エージェント開発の利点は、認知モデリングの研究において蓄積されてきた研究知見の体系的な流用にある。さらに、ヴァーチャル世界における対人エージェントの振る舞いを多角的に可視化することは、認知モデリングに実装された内部処理の仮説に対する新たな評価にもつながる。このように、本研究が目指す統合アプローチは、あらたな認知モデリングの手法、HAI の手法を導き、両者を融合する認知科学の新たな研究領域を生み出す可能性がある。

#### 文献

- [1] Anderson, J. R. (2007). How Can the Human Mind Occur in the Physical Universe. Oxford University Press.
- [2] Trafton, J. G., Hiatt, L. M., Harrison, A. M., Tamborello II, F. P., Khemlani, S. S., and Schultz, A. C. (2012). ACT-R/E: An embodied cognitive architecture for Human-Robot Interaction. *Journal of Human-Robot Interaction*, 1(1), 78-95.
- [3] Somers, S (2016). ACT-R 3D: A 3D Simulation Environment for Python ACT-R. *Proceedings of the 14th International Conference on Cognitive Modeling (ICCM 2016)*, pp 107-112.
- [4] Smart, P. R., Scutt, T., Sycara, K., and Shadbolt, N. R. (2016) Integrating ACT-R Cognitive Models with the Unity Game Engine. In J. O. Turner, M. Nixon, U. Bernardet and S. DiPaola (Eds.), *Integrating Cognitive Architectures into Virtual Character Design*. IGI Global, Hershey, Pennsylvania, USA.
- [5] Brooks, R. A. (1987). Planning is just a way of avoiding figuring out what to do next. Technical report, MIT Artificial Intelligence Laboratory.
- [6] Harrison, A. (2002). jACT-R: Java ACT-R. *Proceedings of the 8th Annual ACT-R Workshop*.
- [7] Stewart, T. C. and West, R. L. (2005) Python ACT-R: A New implementation and a new syntax. *Proceedings of 12th Annual ACT-R Workshop*.