

# 知的好奇心の計算論モデルとアルゴリズムモデルの接続：ベイジアンネットワークを用いたACT-Rモデルの分析

## Connecting Computational and Algorithmic Models of Intellectual Curiosity: Analysis of the ACT-R Model Using Bayesian Network

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

Kazuma Nagashima, Junya Morita & Yugo Takeuchi

<sup>†</sup> 静岡大学

Shizuoka University

nagashima.kazuma.16@shizuoka.ac.jp, j-morita@inf.shizuoka.ac.jp, takeuchi@inf.shizuoka.ac.jp

### 概要

Marr によれば認知モデルを表す階層として計算論とアルゴリズムの水準が区別される。これらの水準は相互に関連しているが、そのつながりは必ずしも明確ではない。そこで本研究では、アルゴリズムの水準の認知モデルを、計算論の観点から検討し、妥当性を付与することを目指す。本報告では、知的好奇心のACT-Rモデルの振る舞いを計算論に基づくベイジアンネットワークを用いて分析した。その結果、知的好奇心の仮説とACT-Rモデル間で整合する特徴が現れた。

**キーワード：**認知モデル, ACT-R, 内発的動機づけ, 知的好奇心

### 1. 背景

好奇心は、教育やエンターテイメントなどの、幅広い分野において、取り組む活動に関心を持たせ、その活動を継続させる動因となる。好奇心を持つことで人間は幅広い環境を学習することができる。この好奇心を応用したモデルの研究は、古くから様々なアプローチで行われている [3, 5, 17, 20]。

こういった人間の認知機能に関わるモデルを整理する際に、Marrの3水準は有効である [12]。これらは、数理によって表現される計算論の水準、情報処理・認知アーキテクチャによって表現されるアルゴリズムの水準、そしてハードウェアなどへの実装の水準からなる。これらの各水準は、相互に制約条件になる関係を持つが、その関係は弱いとされる。しかし、認知機能理解のために、各水準の独立した事項だけの理解だけでなくこの関係を理解する必要があるとされる。

そこで、本研究では、水準間の関係の理解に着目し、計算論の水準とアルゴリズムの水準を接続することを目指す。これらを接続することで、アルゴリズムの水

準は、計算論の水準に対して実現方法 (how) を与えることができ、計算論の水準はアルゴリズムの水準に対して妥当性 (why, what) を付与できると考える。本報告では、計算論とアルゴリズムの水準間の接続を行うという目的に向けて、著者らが過去に実装した知的好奇心のACT-Rモデルを説明するベイジアンネットワークモデルを示す。ACT-Rは記号的表現をベースとした情報処理モデルであり、アルゴリズムの水準に位置づけられる。しかし、その処理には多くのルールが介在し、計算論の水準に基づく議論が困難である。それを抽象化された数式によって説明することで、複雑なACT-Rモデルの振る舞いを計算論に基づいて議論できるようになると考える。

### 2. 関連研究

本研究は、知的好奇心の計算論の水準のモデルとアルゴリズムの水準のモデルの接続を目指すものである。この目的と関連した研究として、(1) 計算論の水準の好奇心のモデルの研究、(2) アルゴリズムの水準のACT-Rによる知的好奇心のモデルの研究を紹介する。

#### 2.1 計算理論に基づく好奇心のモデル

好奇心に関しては、近年、様々な人工エージェントの研究において扱われている。好奇心の計算論モデルとして、Fristonによるフリーエナジーの理論 [5] を挙げることができる。この理論では、好奇心のきっかけとなる驚きや興味、楽しさなどの感情は、外界の認識と経験から得られる予測との不一致によって引き起こされると説明されている。この考えを援用し、様々な研究者が、予測誤差の観点から、好奇心に基づく環

境学習における自律エージェントの開発を行ってきた [19, 20].

その中でも Singh は、好奇心をモデル化した IMRL (Intrinsically Motivated Reinforcement Learning) という手法を用いて強化学習エージェントにより広く環境を探索させる方法を検討した。近年、この考えは、深層強化学習の枠組みと共に発展している [3, 13, 17]. これらの研究は、トイワールドあるいは限定条件下のタスクにおいて高いパフォーマンスを発揮する。こういった研究から計算論の水準と実装の水準との接続についてはすでに高水準でなされているといえる。

しかし、深層強化学習におけるモデルの実装は end-to-end に行われるため、モデルの内部の処理は明確とは言いがたい。加えて、このような研究は、同一の計算論を背景にしているにも関わらず、同様のトイワールドなどのタスクにおいて、複数の実装方法が存在する [2]. 例えば、Pathak は深層強化学習を援用し、環境から得られる外部報酬と、ピクセル情報から得られるエージェントの次状態との予測誤差から得られる内部報酬を好奇心とし、好奇心のモデルを実装した [17]. この考えを援用し、Burda は内部報酬、つまり好奇心のみを用いた好奇心のモデルを実装した [3]. これらのモデルの計算論的背景は、いずれもフリーエナジーの理論 [5] に基づく。しかし、最適なアルゴリズムを特定の計算論に対して提供できていないため、計算論とアルゴリズムの水準の接続はいまだ不十分である。

## 2.2 アルゴリズムの水準による知的好奇心のモデル

これらの計算論の水準の研究に対し、筆者らはアルゴリズムの水準の好奇心のモデルの研究を行ってきた。ここでは、好奇心の一種である知的好奇心 [11] の認知モデルの実装を目指した。このモデルの実装には、現存するアーキテクチャのなかで、最も多くの機能を有し、最も多くの先行研究に利用されている認知アーキテクチャである ACT-R (Adaptive Control of Thought-Rational[1]) を用いた ([9] による網羅的な認知アーキテクチャのレビューを参照)。この豊富な ACT-R のモジュールとルールを用いてアルゴリズムを記述することで、多様な心理学的な知見に裏付けされたモデルを作成できる。

### 2.2.1 ACT-R のモデルの概要

本項では、先行研究 [16, 21] において構築した ACT-R モデルの設計を述べる。具体的なモデルの処理は付録に記載した。

2.1 節にて述べたように、好奇心は外界の認識と経験から得られる予測との差分によって生じるものとして数理的に説明できる。この予測からの差分が驚き(好奇心)を生じさせ、そのうちの一部は、楽しさなどの感情的反応を引き起こし、人間に興味や関心を持たせ、その活動を継続する動因となる。

楽しさに関しては複数の研究者が、環境における新規なパターンの発見と関連することを述べている [4, 7, 8, 19]. たとえば、ゲームデザイナーである Koster は、著書の “Theory of Fun for Game Design [8]” において、人間がゲームにおいて感じる楽しさは、ゲーム内での新しいパターンを発見することによって引き起こされると述べる。また Koster と同様のより形式的な楽しさの理論化は、Schmidhuber による “Formal Theory of Creativity, Fun, and Intrinsic Motivation [19]” においてなされている。この理論において、パターンの発見はデータの中で反復される定型的なパターンを発見し、圧縮することと定義される。Schmidhuber は、データを圧縮すること、あるいは圧縮可能なデータを取得することを楽しさと対応づけた。

以上のことより筆者らは、前述した知的好奇心が楽しさを引き起こすということと、人間が環境のパターンを発見することで得られる楽しさが対応するのではないかと考えた。つまり、人間の環境におけるパターンの発見を知的好奇心として捉え、それを ACT-R のパターンマッチングと対応させ知的好奇心のモデルを実装した。この実装された迷路の継続課題を行う知的好奇心のモデルは、課題中にパターンマッチングを発生させると楽しさを感じるという仮定に基づいている。さらに、ACT-R における手続き学習 (コンパイル) は、経験を重ねることによるパターンマッチの機会の減少 (ルールと知識の圧縮) を導く。著者らのモデルはこの学習のプロセスを含めることで、モデルの楽しさを感じる頻度の減衰により、飽きのプロセスを表現した。

### 2.2.2 シミュレーション

筆者らの過去の研究では、2.2.1 節の設計に基づくモデルを構築し、迷路の継続課題のシミュレーションを実施した [16]. シミュレーションの目的は提案する

知的好奇心のメカニズムの性質を検討することであった。詳細に知的好奇心のメカニズムを検討するため、それぞれ異なるサイズのマップと、複数の思考水準のモデルを実装した。それぞれのモデルは課題中に、パターンマッチングが発生させると、知的好奇心が刺激される。そこで筆者らは、知的好奇心の増加がモデルの課題の継続につながると仮説をたてた。そして、それに伴い、課題のゴール達成率、マップを広範囲に探索した際に増加するエントロピー、スキル生成数が増加すると考え、これらをシミュレーションの指標とした。

図1は、シミュレーション結果である。シミュレーションの前提として、グリッド状の迷路のマップのサイズを3段階用意し、それぞれの大きさに対し、異なるマップを10用意した。1つのマップに対して、パターンマッチングに付随する知的好奇心の大きさ（楽しさ）を20段階変化させた。モデルはこれらの条件の下、マップのスタートからゴールまで探索を行った。このスタートからゴールまでの探索（ゴール未達成も含む）を1ラウンドとし、ラウンド開始時に課題を続ける、やめるかの判断を行うルールを競合させた。シミュレーション開始時には、課題継続の動機づけが高いと設定し、ラウンド中にパターンマッチングが発生したら課題継続の知的好奇心の大きさに伴う動機づけが高まり（楽しさ）、パターンマッチングが生じないままラウンドを終了した場合、課題継続の動機づけが減少する（飽き）と設定した。1回のシミュレーション中にモデルは飽きる（課題継続の動機づけが課題をやめる動機づけを下回る）まで、これを繰り返す。そしてモデルが飽きたら繰り返したラウンド数を課題の継続回数として記録した。

グラフに示した結果は、それぞれのサイズに対して異なる10のマップの結果を平均したものになる。また、左側のモデル（DFS+IBL, DFS）の思考水準が高く、右側のモデル（Random）の思考水準が低い。ここでの思考水準は、モデルが行動を起こすまでに行う処理の量（パターンマッチングの頻度）を意味している。より端的に言えば、右に配置されたモデルに比べ、左に配置されたモデルはより多くのルールを有しているため、パターンマッチングが発生する頻度が多い。

グラフより、思考水準の高いDFS+IBLとDFSモデルは、知的好奇心に伴う報酬を与えると課題の継続数が増え、その他の指標も増加傾向にあることがわかる。この結果は、思考水準の高いモデルは、筆者らの仮定と整合する振る舞いを生じさせることが示される。しかし、思考水準の低いRandomモデルは、報酬

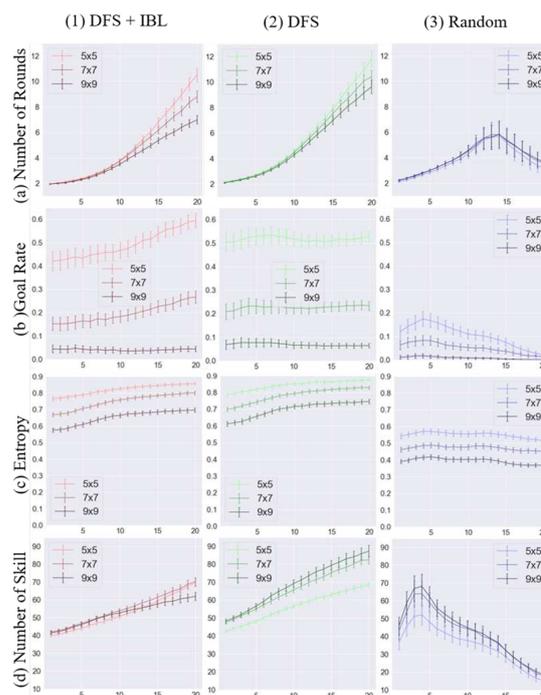


図1 シミュレーション結果。列方向によってモデルを区別し（1: DFS+IBL モデル, 2: DFS モデル, 3: Random モデル）縦方向のアルファベットにより指標を区別する（a: ラウンド継続数, b: ゴール達成率, c: エントロピー, ACT-R のモジュールによって生成されたルール数（d））。それぞれのグラフのエラーバーは各マップにおいて得られた標準偏差（ $n = 1000$ ）の平均（ $n = 10$ ）に1/10を乗じた値を示す。なお、それぞれのグラフの横軸は知的好奇心の強さである。

を増やすと、課題継続数が逆U字の増減を見せ、知的好奇心の仮説に沿う結果にならなかった。したがって、これらのモデルの振る舞いが、どのような計算論に基づいているのか検討できず、モデルの振る舞いに妥当性を得られなかった。

### 3. 提案手法

図1で示されたように、これらアルゴリズムの水準のモデルは、モデルごとに異なるルールを持ち、変数同士が相互に作用しながら課題を遂行するため複雑である。したがって、上記の研究 [15, 21] において実装された思考水準の異なるモデルについて、簡潔な説明を付与することは困難である。この複雑さを軽減し、より一般的な説明を付与するために、本報告では計算論の水準とアルゴリズムの水準を接続することを目指す。

この目的を達成するため、筆者らはアルゴリズムの水準のモデルに計算論的妥当性を付与するため、計算

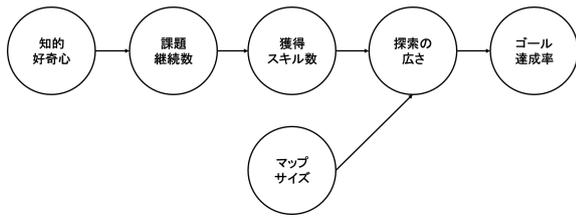


図2 シミュレーションで用いた各変数のネットワーク構造

論の水準のベイジアンネットワークを構築し、各変数間でどのような作用を持つのかという因果関係を推定することを提案する。これによって、アルゴリズム水準のモデルは目指すべきネットワーク構造（計算論）を持ち、シミュレーションで用いた操作変数（知的好奇心の大きさ、マップの広さ）や指標間の因果関係が判明する。その後、知的好奇心のモデルを、そのネットワーク構造にフィッティングするまで修正し、分析を繰り返すことで、計算論の水準とアルゴリズムの水準を接続できるのではないかと考える。

### 3.1 知的好奇心のモデルの分析

この提案手法を実現するための第一歩として、2.2.2節のシミュレーション結果を対象に、ベイジアンネットワークを用いて分析を行った。図2は、筆者らが仮定していた各変数間の因果関係をネットワーク図にしたものである。このネットワークは、

- (a) 知的好奇心の強さの増加に伴い課題の継続数が増加
- (b) 継続数の増加とともにマップを探索する機会が増えスキルが増加
- (c) スキルが増加すると、エージェントは広範囲にマップを探索することができ、エントロピーが増加
- (d) エントロピーの増加に伴いゴール達成率が上昇

という因果関係を表している。なお、マップサイズが大きくなると相対的にモデルの行動範囲が局所化すると考えられる。そして行動範囲の局在化は情報エントロピーとして表現できる。そのためマップサイズは探索の広さに関係するのではないかと考えた。

### 3.2 分析結果

表1は、図2のネットワーク構造に対して、図1の結果をフィッティングした結果である。得られた因

果関係を回帰式として表現する際の係数（2行から6行）、各モデルとネットワークの適合度（7行）が示される。回帰式における操作変数を Reward（知的好奇心の強さに伴う報酬）と Map Size（環境の大きさ）とし、目的変数を Goal Rate（ゴール達成率）とした。列は図1であげたそれぞれのモデルに対応する。

表に示される適合度（BIC: Bayesian Information Criterion）は、

$$BGe = \sum_{i=1}^k \log f_{X_i} \left( X_i \mid \prod X_i \right) - \frac{d}{2} \log n \quad (1)$$

によって計算された。この式は、Rパッケージである bnlearn[14] において定義されるものである。 $k$  は変数の数を表し、 $X$  は変数自身を表す。また、 $d$  はパラメータ数を表し、 $n$  はサンプル数を表す。左項が精度、右項がペナルティを表しているため、つまり、このスコアの値が大きいほどデータとネットワークが適合していると言える。したがって、IBL+DFSモデル、DFSモデル、ランダムモデルの順で値が大きいため、この順番でモデルとネットワークが適合していると言える。

これらの結果からそれぞれのモデルの特徴が示される。例えば操作変数である知的好奇心の大きさに伴う報酬は、パターンマッチングの頻度が多い思考水準の高い IBL+DFS、DFSモデルに対して強く作用するが、思考水準の Random モデルにはあまり作用しない。それを裏付けるように、思考水準の高いモデルはネットワークとの適合度が高いことを示し、思考水準の低いモデルはネットワークとの適合度が低い結果となった。

### 4. 今後の計画とまとめ

本報告では ACT-R によって実装された知的好奇心のモデルに対し、一般的な説明を付与するために、その出力からベイジアンネットワークを構築した。そのベイジアンネットワークを用いて分析を行い、モデルごとの変数間の関係を回帰式の係数、およびネットワー

表1 各々の変数に対する回帰式の係数

|                  | IBL+DFS              | DFS                  | Random               |
|------------------|----------------------|----------------------|----------------------|
| GoalRate~Entropy | 0.547                | 0.559                | 0.448                |
| Entropy~MapSize  | -0.446               | -0.563               | -0.386               |
| Entropy~Skill    | 0.591                | 0.657                | 0.041                |
| Skill~Round      | 0.796                | 0.741                | 0.394                |
| Round~Reward     | 0.651                | 0.655                | 0.117                |
| BIC              | $-4.295 \times 10^6$ | $-4.331 \times 10^6$ | $-4.938 \times 10^6$ |

クの適合度を表すことができた(表1)。これらの結果から、思考水準の高いモデルは知的好奇心の強さの増加と共にパフォーマンスが向上し、一方で思考水準の低いモデルは知的好奇心からの影響は少ないことが示される。この分析は、アルゴリズム水準のモデルによって生成された振る舞いを、計算論の水準(回帰式)に変換するものである。アルゴリズムによって生成される複雑な振る舞いを、より簡潔な表現に置き換えることで、回帰式の係数やネットワークの適合度からその妥当性を検討することができるようになったといえる。

しかし、本研究には、表1で得られた変数間の関係を詳細に分析できていないという限界もある。そのため、知的好奇心のモデルを計算論と接続するという目標の達成は、未だ十分ではない。今後は、さらに変数間の関係を詳細に分析する必要がある。前節の通り、この表は図2のネットワーク構造に対して図1の結果をフィッティングしたものである。したがって、各変数間の関係を表している。モデルごとの特徴が各変数間の回帰式の係数として現れているため、モデルのアルゴリズムがネットワーク構造に合うまでモデルを更新し、モデルと計算論の水準の接続ができると考える。

また、シミュレーション結果からボトムアップにネットワーク推定を行い、仮説のネットワークと各モデルのネットワークとの相違点および変数間の関係の相違点を比較分析する必要がある。今回行った推定では、因果構造のモデルを仮説としてトップダウンに構築し、それをもとに変数間の分析を行った。これによって、仮説に合うモデルの評価ができるが、合わないモデルに対して評価が行えない。トップダウンのネットワーク、ボトムアップに推定した各モデルのネットワークを比較分析することで、仮説に合わなかったモデルに対して解釈を与えられると考える。

## 文献

- [1] J. R. Anderson. *How Can the Human Mind Occur in the Physical Universe*. Oxford Press, 2007.
- [2] Arthur Aubret, Laëtitia Matignon, and Salima Has-sas. A survey on intrinsic motivation in reinforcement learning. *Corr*, 2019.
- [3] Yuri Burda, Harri Edwards, Deepak Pathak, Amos Storkey, Trevor Darrell, and Alexei A Efros. Large-scale study of curiosity-driven learning. *arXiv preprint arXiv:1808.04355*, 2018.
- [4] Roger Caillois. *Les jeux et les hommes*. 1958.
- [5] Karl Friston. The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, Vol. 11, No. 2, pp. 127–138, 2010.
- [6] Cleotilde Gonzalez, Javier F Lerch, and Christian Lebiere. Instance-based learning in dynamic decision making. *Cognitive Science*, Vol. 27, No. 4, pp. 591–635, 2003.
- [7] Johan Huizinga. *Homo ludens versuch einer bestimmung des spielementes der kultur*. 1939.
- [8] Raph Koster. *Theory of fun for game design*. ” O’Reilly Media, Inc.”, 2013.
- [9] Iuliia Kotseruba and John K. Tsotsos. 40 years of cognitive architectures: core cognitive abilities and practical applications. *Artificial Intelligence Review*, Jul 2018.
- [10] Christian Lebiere, Cleotilde Gonzalez, and Michael Martin. Instance-based decision making model of repeated binary choice. 2007.
- [11] Thomas W Malone. Toward a theory of intrinsically motivating instruction. *Cognitive Science*, Vol. 5, No. 4, pp. 333–369, 1981.
- [12] David Marr. Vision: A computational investigation into the human representation and processing of visual information. 1982.
- [13] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. Human-level control through deep reinforcement learning. *Nature*, Vol. 518, No. 7540, pp. 529–533, 2015.
- [14] Radhakrishnan Nagarajan, Marco Scutari, and Sophie Lèbre. Bayesian networks in r. *Springer*, Vol. 122, pp. 125–127, 2013.
- [15] K. Nagashima, J. Morita, and Y. Takeuchi. Modeling intrinsic motivation in act-r: Focusing on the relation between pattern matching and intellectual curiosity. In *Iccm2020: 18th International Conference on Cognitive Modeling*, 2020.
- [16] K. Nagashima, J. Morita, and Y. Takeuchi. Curiosity as pattern matching: Simulating the effects of intrinsic rewards on the levels of processing. In *ICCM 2021: 19th International Conference on Cognitive Modeling*, 2021.
- [17] Deepak Pathak, Pulkit Agrawal, Alexei A Efros, and Trevor Darrell. Curiosity-driven exploration by self-supervised prediction. In *Proceedings of the Ieee Conference on Computer Vision and Pattern Recognition Workshops*, pp. 16–17, 2017.
- [18] D. Reitter and C. Lebiere. A cognitive model of spatial path-planning. *Computational and Mathematical Organization Theory*, Vol. 16, No. 3, pp. 220–245, 2010.
- [19] Jürgen Schmidhuber. Formal theory of creativity, fun, and intrinsic motivation (1990–2010). *Ieee Transactions on Autonomous Mental Development*, Vol. 2, No. 3, pp. 230–247, 2010.
- [20] S. Singh, A. G. Barto, and N. Chentanez. Intrinsically motivated reinforcement learning. In L. K. Saul, Y. Weiss, and L. Bottou, editors, *Advances in Neural Information Processing Systems 17*, pp. 1281–1288. Mit Press, 2005.
- [21] 長島一真, 森田純哉, 竹内勇剛. 多様な環境の学習における ACT-R を用いた内発的動機づけのモデル. No. O3-2, pp. 62–71. 日本認知科学会第 37 回大会, 2020.

## 付録

図1で用いた3つの思考水準のモデルの概要を記述する。これらは、筆者らが ICCM 2021 において発表したモデルと同様である [16]。モデルは段階的に機能を付与され、その順に課題を遂行する中で利用される宣言的記憶（パターンマッチング）が増大する。ACT-R の理論において、パターンマッチングの対象となる宣言的記憶は脳皮質（腹側前頭前野）を介して引き出されるとされる。対して、それが利用されない処理は、プロダクションモジュール（脳基底核と対応）の比重が大きい。人間の意識的な処理が前頭前野を介して行われ、自動的な処理が脳基底核を介して行われるとの想定 [1] に立てば、ACT-R は、パターンマッチングを伴わない低次の思考とパターンマッチングを伴う高次の思考に区別するといえる。したがって、図3の左から右に向かうに従って、深い思考水準を組み入れたモデルと仮定した。それぞれのモデルの詳細を下記に記述する。

## ランダムモデル

最も思考水準の低いモデルは、ランダムにゴールモジュールの現在位置を遷移させる。ラウンド内において、モデルはゴールに達するか制限時間に達するまで下記の処理を繰り返す。

1. 方向の決定: 4つの移動方向（東、西、南、北）を表すプロダクションを式(1)に基づいて、確率的に発火させる。なお、4つのプロダクションのユーティリティの初期値は等しく設定する。
2. 行き先の決定: 決定された方向への移動が可能か否かをパスに関わる宣言的記憶を検索することで判断する。この判断に利用するプロダクションには、ゴールモジュールに格納される現在位置と(1)にて決定された方向が、変数として含まれる。パターンマッチの結果、該当するパスの記憶が見つからなかった場合、(1)に戻る。パター

ンマッチに成功するパスが見つかった場合、ゴールモジュールの状態を検索されたパスに従って変更した後に(1)に戻る。一旦検索されたパスはプロダクションコンパイルされる。なお、プロダクションコンパイル後のプロダクションのユーティリティは、プロダクションコンパイル前のプロダクションのユーティリティよりも低く設定される。同じ宣言的記憶がプロダクションコンパイルされるごとにユーティリティが増大し、最終的にプロダクションコンパイルされる前のプロダクションのユーティリティと等しくなる。この設定は ACT-R のデフォルトの設定に従うものである。

## 確率的 DFS モデル

より高次の認知機能（宣言的記憶）をモデルに含めるために、確率的な DFS (depth-first search)、あるいはバックトラックによって環境を探索するモデルを構築した。その実装は、Reitter らの研究 [18] を参考にした。このモデルは、4に示すように、ACT-R のイマジナルモジュールで生成されたチャンクを用いたスタック構造を備える。スタック内のプッシュ機能は、ARG1 スロットに過去のチャンク名を格納することで、スタック内のポップ機能は、ARG1 スロットの値を過去のスロットの値に戻すことで実現している。これらの生成されたチャンクは宣言的モジュールに格納され、後から検索することでポップ機能が実現される。ACT-R は Lisp で記述されているため、Lisp の処理系にアクセスすることが容易である。しかし、これらの処理はすべて、Lisp など他のプログラミング言語で書かれた外部関数を定義することなく、ACT-R のプロダクションのみで実装した。

具体的なモデルの動作は、以下の通りである。

1. 方向の決定: ランダムモデルと同様に、まず移動方向をランダムに決定する。
2. 行き先の決定: ランダムモデルと同様に、現在位置と移動方向を結ぶパス検索に成功した場合、ゴールモジュールの現在位置を変更する。加えて、DFS モデルは、検索されたパスに検索済みのタグを付与した新たなチャンクをイマジナルモジュールを介して作成し、宣言的モジュールに格納する。モデルがパスの検索に失敗したとき（行き止まりになったとき）、モデルは検索済みのタグが付与されたチャンクを辿ることで以前の場所に戻り、(1)を繰り返す。ランダムモデルと同様に、確率的 DFS モデルは、パスに関する宣言的

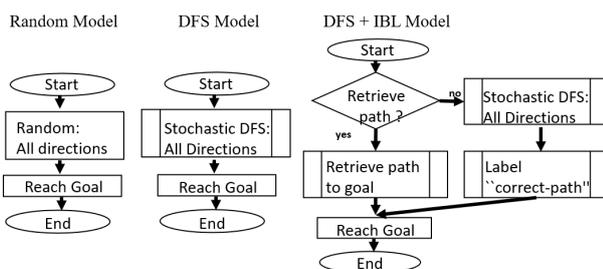


図3 思考水準の異なるモデル

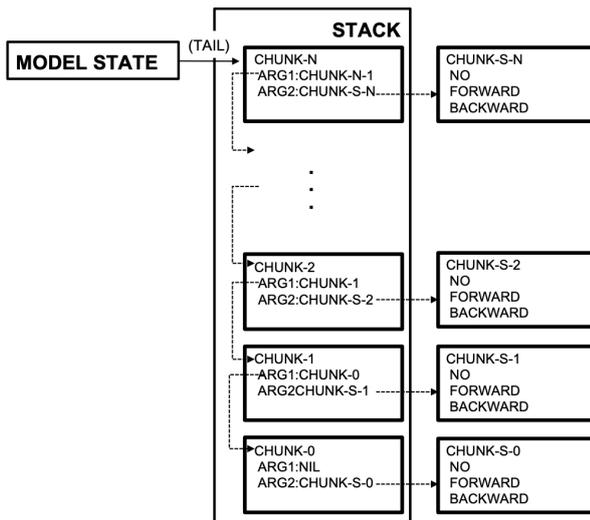


図4 ACT-Rのチャンクで構築したスタック構造。スタックの実装のためにACT-Rのイマジナルモジュールを使用している。

記憶をプロダクションコンパイルし、変数を含まない新しいプロダクションを学習する。

モデルはゴールに到達するか制限時間に達するまで(1)と(2)を繰り返す。パスの検索に失敗した場合、ランダムモデルが直ちにパスの検索を繰り返すのに対し、DFSモデルは宣言的記憶を検索しながらスタックをポップしていく。そのため、このモデルが効果的に環境を探索するためには、より多くのラウンドを繰り返し、パスの記憶をプロダクションコンパイルする必要が生じる。

## 確率的DFSとIBL組み合わせモデル

このモデルは確率的DFSとIBL(instance-based learning)を組み合わせる。IBLとは、現在の課題の解決に過去の記憶を用いる学習方法である[6, 10]。この課題では、各ラウンドの開始から、ゴールに到達するまで、経由したパスをスタックに蓄える。ゴールに到達した後に、スタック内のパスを辿り、正解ラベルの付与されたチャンクを生成する。ラウンドの進行は、以下の2つのステップをゴールに到達するか制限時間に達するまで繰り返す。

1. 方略の決定: DFS戦略を取るかIBL戦略を取るかを競合解決により決定する。
2. 方向と移動先の決定
  - (a) DFS戦略を採用した場合は、DFSモデルと同様に振る舞う。
  - (b) IBL戦略を採用した場合は正解ラベル付き

のパスの検索を試みる。失敗した場合は(1)に戻る。成功した場合は、正解ラベルの付与されたパスに従って現在位置を更新する。

課題の初期段階で、このモデルは確率的DFSモデルと同様に振る舞う。ラウンドが繰り返され、正解ラベル付きのパスの記憶が増加すると、効果的にゴールに到達できるようになる。IBLは、DFSに比べて、時間を要するコストの大きい処理である。正解ラベルをパスに付与するためには、ラウンドの終了時にスタック内のパスを検索しなければならない。また、パスの選択時にて、過去のラウンドの記憶を思い出すことも、他のモデルには存在しない時間的コストとなる。