# 強化学習で行動・推論・対話の制御プログラムを獲得する脳型AGIアーキテクチャの設計

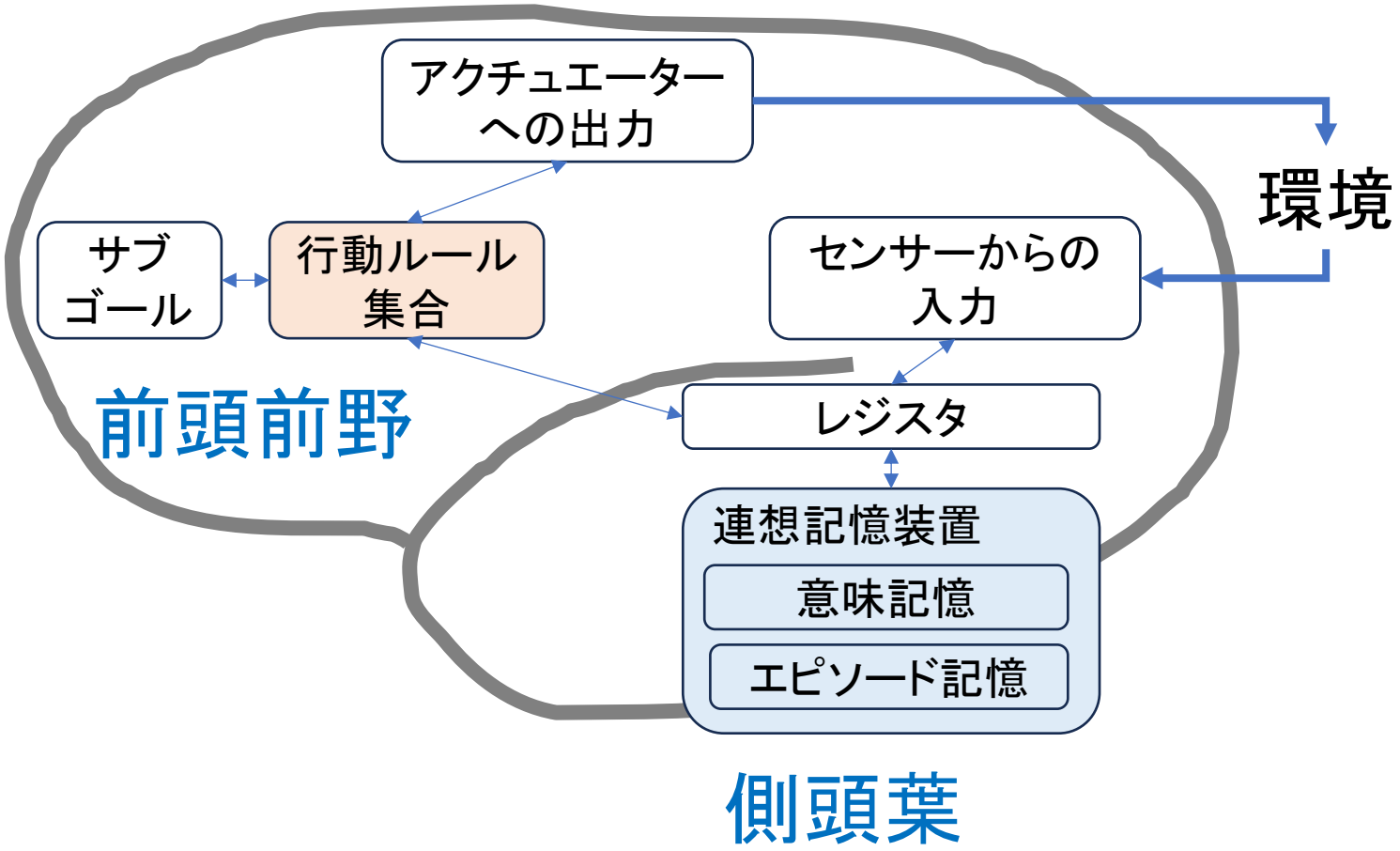
2025年度 人工知能学会全国大会（第39回）

4E1-OS-12a-01

一杉裕志,高橋直人,竹内泉（産総研）,
佐野崇（東洋大学）,中田秀基(産総研,順天堂大学)

2025-5-30

# 提案する脳型AGIアーキテクチャの全体像



# これまでの研究経緯と発表の内容

- 1993年　電子技術総合研究所（現　産総研）入所
  - 並列処理、プログラミング言語に関する研究
- 2005年より計算論的神経科学・人工知能を研究
  - 目標：脳の情報処理の理解と脳型AGIの実現
  - 大脳皮質、視覚野、運動野、言語野の計算モデル
- 最近約５年間の研究内容：
  - 前頭前野の計算モデル
    「前頭前野はヒトをヒトたらしめ,
    思考や創造性　を担う脳の最高中枢であると考えられている。」
    前頭前野 － 脳科学辞典
- 2025年度いっぱいで引退予定

知りたいこと、わからないことなどコメントを歓迎いたします

**6.　まとめ**

ヒトの脳の情報処理を模倣した脳型 AGI アーキテクチャ設計の取り組みの現状について報告した。個々の要素技術の動作確認はすんでおり、全体の統合も可能であるという見通しを得ている。

この方向で研究開発を続ける研究者がもし増えれば、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。

**参考文献**

[Anderson 07] Anderson, J. R.: *How Can the Human Mind Occur in the Physical Universe?*, Oxford University Press (2007)

[Baars 07] Baars, B. J. and Franklin, S.: An architectural model of conscious and unconscious brain functions: Global Workspace Theory and IDA, *Neural Networks*, Vol. 20, No. 9, pp. 955–961 (2007), Brain and Consciousness

[Ichisugi 07] Ichisugi, Y.: A Cerebral Cortex Model that Self-Organizes Conditional Probability Tables and Executes Belief Propagation, in *2007 International Joint Conference on Neural Networks*, pp. 178–183 (2007)

[Ichisugi 19a] Ichisugi, Y. and Takahashi, N.: A Formal Model of the Mechanism of Semantic Analysis in the Brain, in Samsonovich, A. V. ed., *Biologically Inspired Cognitive Architectures 2018*, pp. 128–137, Cham (2019), Springer International Publishing

[Ichisugi 19b] Ichisugi, Y., Takahashi, N., Nakada, H., and Sano, T.: Hierarchical Reinforcement Learning with Unlimited Recursive Subroutine Calls, in *Artificial Neural Networks and Machine Learning – ICANN 2019: Deep Learning*, pp. 103–114, Cham (2019)

[Kaelbling 98] Kaelbling, L. P., Littman, M. L., and Cassandra, A. R.: Planning and acting in partially observable stochastic domains, *Artificial Intelligence*, Vol. 101, pp. 99–134 (1998)

[Yamakawa 21] Yamakawa, H.: The whole brain architecture approach: Accelerating the development of artificial general intelligence by referring to the brain, *Neural Networks*, Vol. 144, pp. 478–495 (2021)

[一杉 16] 一杉裕志：疑似ベイジアンネットを用いた認知モデルのプロトタイピング手法の提案, 第 4 回 人工知能学会 汎用人工知能研究会 (SIG-AGI) (2016)

[一杉 18a] 一杉裕志, 高橋直人：脳における文の意味解析機構のモデル, 第 8 回 人工知能学会 汎用人工知能研究会 (SIG-AGI) (2018)

[一杉 18b] 一杉裕志, 高橋直人, 中田秀基, 佐野崇：単一化の機構を利用した階層型強化学習のテーブル圧縮手法の検討, 第 10 回 人工知能学会 汎用人工知能研究会 (SIG-AGI) (2018)

[一杉 19] 一杉裕志, 中田秀基, 高橋直人, 佐野崇：階層型強化学習 RGoal を用いた記号推論の実現手法の検討, 第 12 回 人工知能学会 汎用人工知能研究会 (SIG-AGI) (2019)

[一杉 20a] 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)

[一杉 20b] 一杉裕志, 中田秀基, 高橋直人, 佐野崇：脳の自律的プログラム合成機構のモデルに向けて： 2 層ベイジアンネットによる記号処理命令の獲得・実行機構, 第 15 回 人工知能学会 汎用人工知能研究会 (SIG-AGI) (2020)

[一杉 20c] 一杉裕志, 中田秀基, 高橋直人, 佐野崇：物体操作に適したワーキングメモリを持つ汎用人工知能アーキテクチャの検討, 第 16 回 人工知能学会 汎用人工知能研究会 (SIG-AGI) (2020)

[一杉 21] 一杉裕志：報酬最大化原理にもとづく脳型ＡＧＩアーキテクチャの構想, 第 18 回 人工知能学会 汎用人工知能研究会 (SIG-AGI) (2021)

[一杉 22a] 一杉裕志, 中田秀基, 高橋直人, 竹内泉, 佐野崇：プログラム合成対象言語 Pro5Lang のための行動価値関数圧縮アルゴリズム, 第 22 回 人工知能学会 汎用人工知能研究会 (SIG-AGI) (2022)

[一杉 22b] 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)

[一杉 22c] 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)

[一杉 23a] 一杉裕志, 中田秀基, 高橋直人, 竹内泉, 佐野崇：再帰的な階層型強化学習 RGoal へのサブルーチン例外終了機能の導入, 第 25 回 人工知能学会 汎用人工知能研究会 (SIG-AGI) (2023)

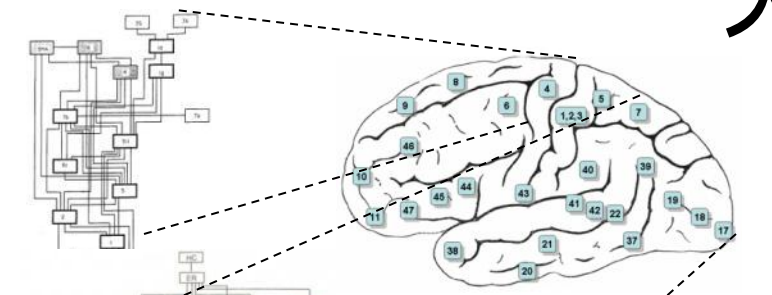
[一杉 23b] 一杉裕志, 中田秀基, 高橋直人, 竹内泉, 佐野崇：報酬最大化を目的とする行動計画・実行・対話・推論の統一的制御機構, 第 37 回 人工知能学会全国大会 (2023)

[一杉 24a] 一杉裕志, 高橋直人, 竹内泉, 佐野崇, 中田秀基：プログラム合成対象言語 Pro5Lang における知識表現形式, 第 27 回 人工知能学会 汎用人工知能研究会 (SIG-AGI) (2024)

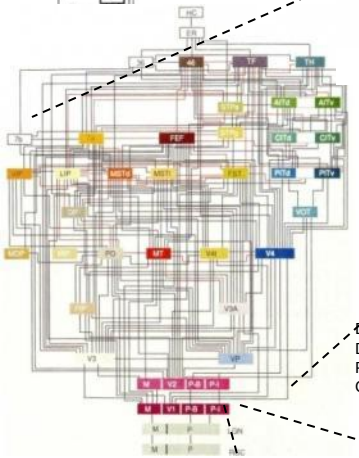
[一杉 24b] 一杉裕志, 高橋直人, 竹内泉, 佐野崇, 中田秀基：複雑な知能を生み出す極限まで簡素化された環境の設計, 第 28 回 人工知能学会 汎用人工知能研究会 (SIG-AGI) (2024)

[一杉 24c] 一杉裕志, 中田秀基, 高橋直人, 竹内泉, 佐野崇：モンテカルロ版 RGoal アルゴリズムの改良, 第 26 回 人工知能学会 汎用人工知能研究会 (SIG-AGI) (2024)

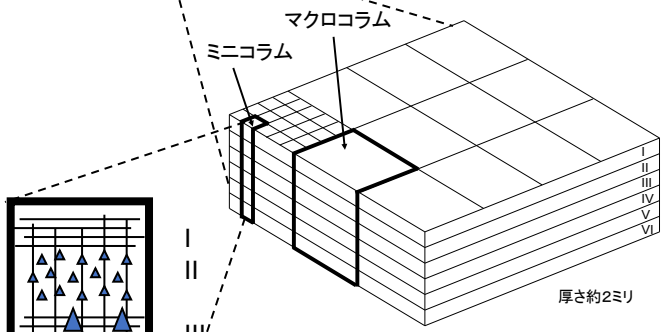
# 大脳皮質

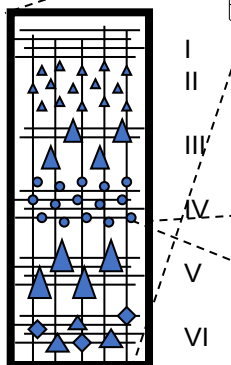


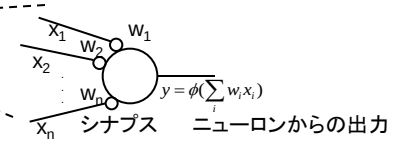
Wikipedia より



Daniel J. Felleman and David C. Van Essen
Distributed Hierarchical Processing in the
Primate Cerebral Cortex
Cerebral Cortex 1991 1: 1-47



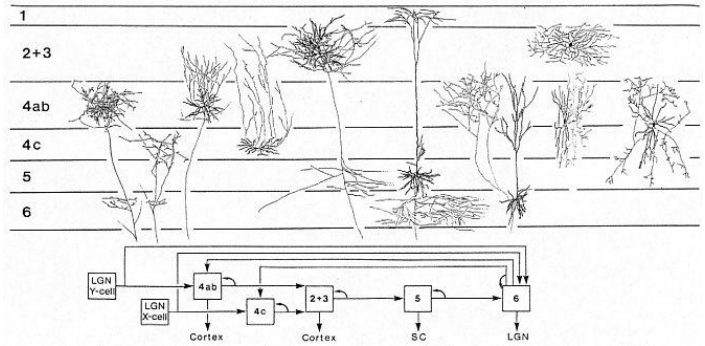



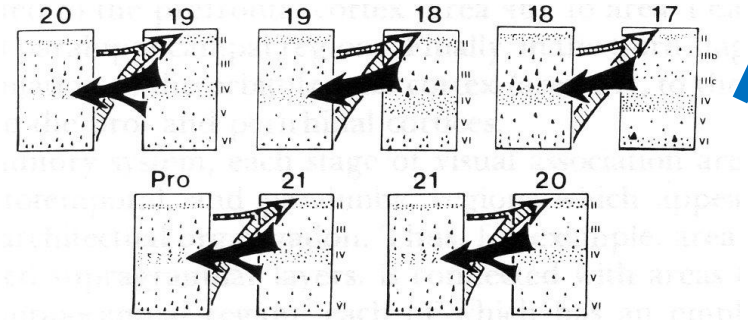

- 脳の様々な高次機能（認識、意思決定、運動制御、行動計画、推論、言語理解など）が、**たった５０個程度**の領野のネットワークで実現されている

**大脳皮質の動作原理がわかれば**
**脳のアーキテクチャ解明の**
**重要な手掛かりに**

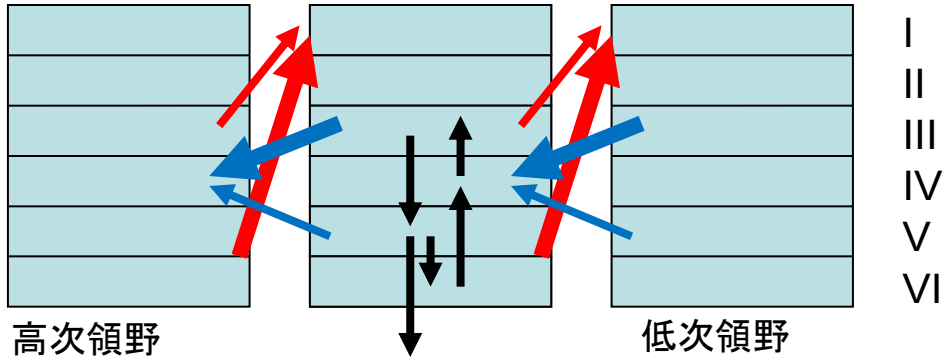
# BESOM モデル [Ichisugi IJCNN 2007] と 大脳皮質のコラム構造・６層構造との一致



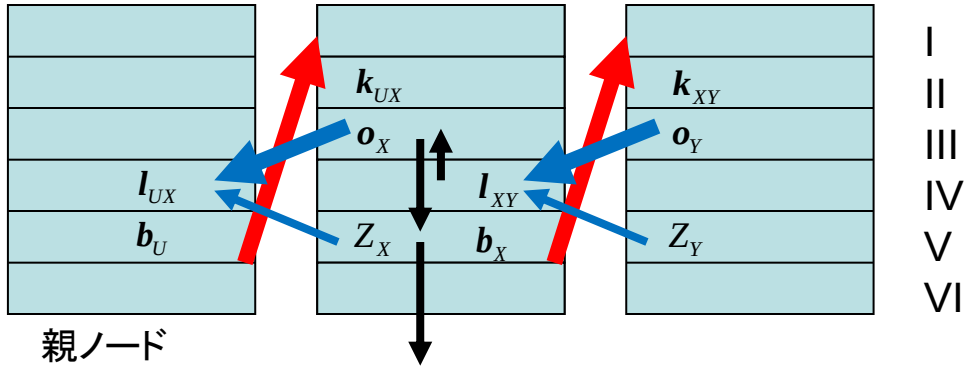

[Gilbert 1983]




[Pandya and Yeterian 1985]




解剖学的構造

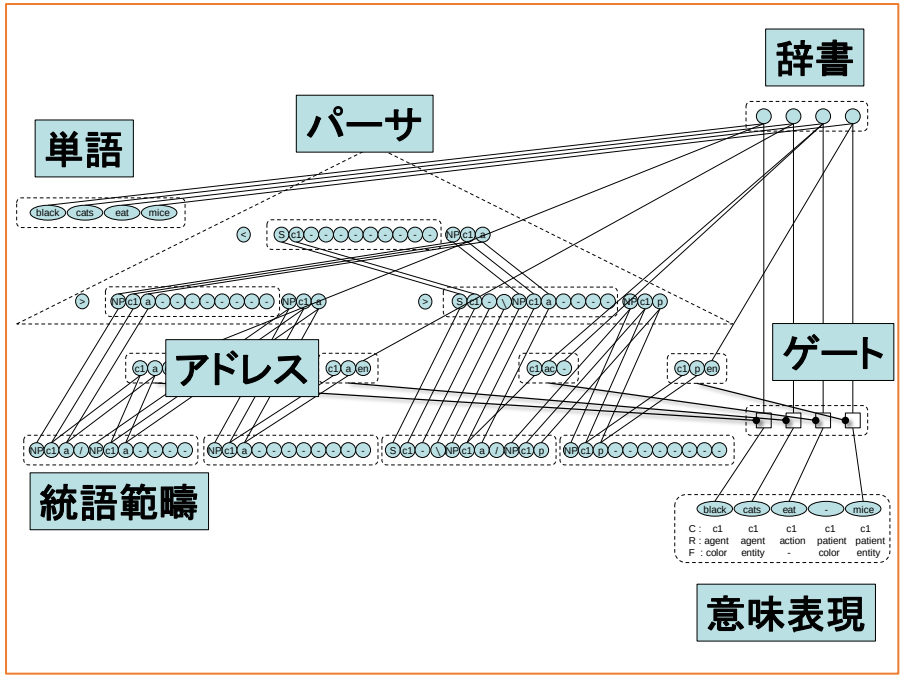



モデル


Pandya, D.N. and Yeterian, E.H., Architecture and connections of cortical association areas. In: Peters A, Jones EG, eds. Cerebral Cortex (Vol. 4): Association and Auditory Cortices. New York: Plenum Press, 3-61, 1985.

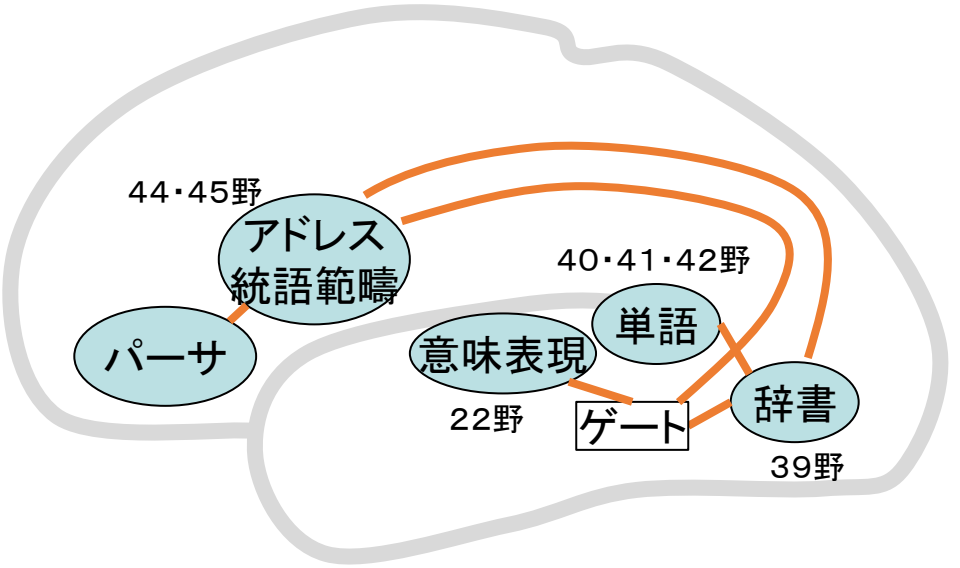
Gilbert, C.D., Microcircuitry of the visual-cortex, Annual review of neuroscience, 6: 217-247, 1983.

# 組み合わせ範疇文法パーサと言語野の対応




構文解析・意味解析機構の全体像

[一杉、高橋 第8回汎用人工知能研究会 2018]
[Ichisugi+ BICA 2018]




脳の領野との対応

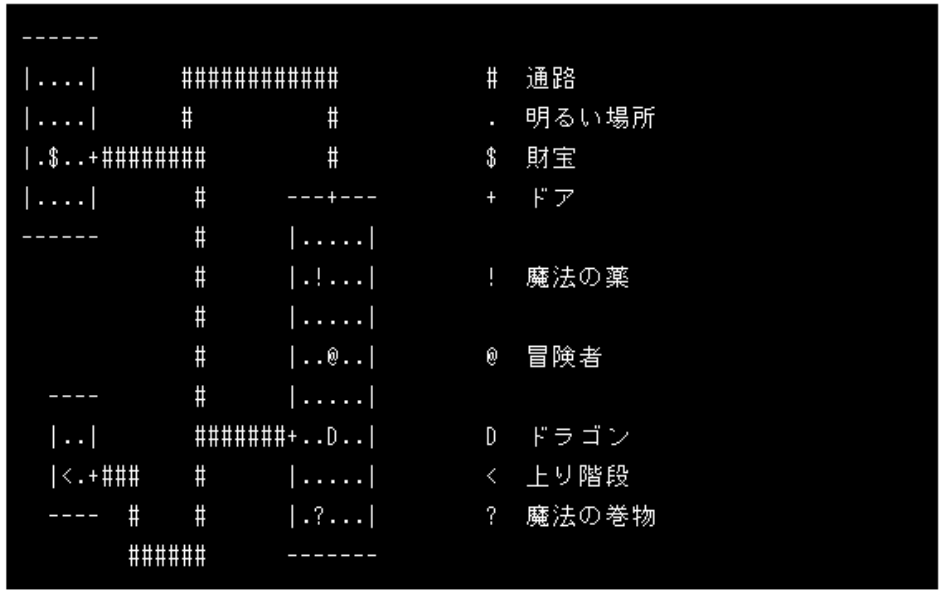
- 単語列と意味表現の間の相互変換を行う数理モデルを構築
- 文の意味を表す論理式を固定長ベクトルで表現
- Prolog で実装、おそらくベイジアンネットの固定回路で実現可能
- 破壊によりウエルニッケ失語・ブローカ失語に似た発話を再現

# 私の研究の中期的目標

[一杉+ 第18回 汎用人工知能研究会(SIG-AGI), 2021]

- ローグライクゲームのような、**実世界を極力単純化した環境**で脳型AGIのデモを動かす
- 前頭前野を中心とした情報処理の計算モデルを作る
- エージェントどうしが協調・競争しながら文化、法律、経済、宗教などを創発させるデモが目標

→ 最低限の道具立てはそろえた




「ローグライクゲーム - Wikipedia」
https://ja.wikipedia.org/wiki/ローグライクゲーム

Generative Agents [Park+ 2023]

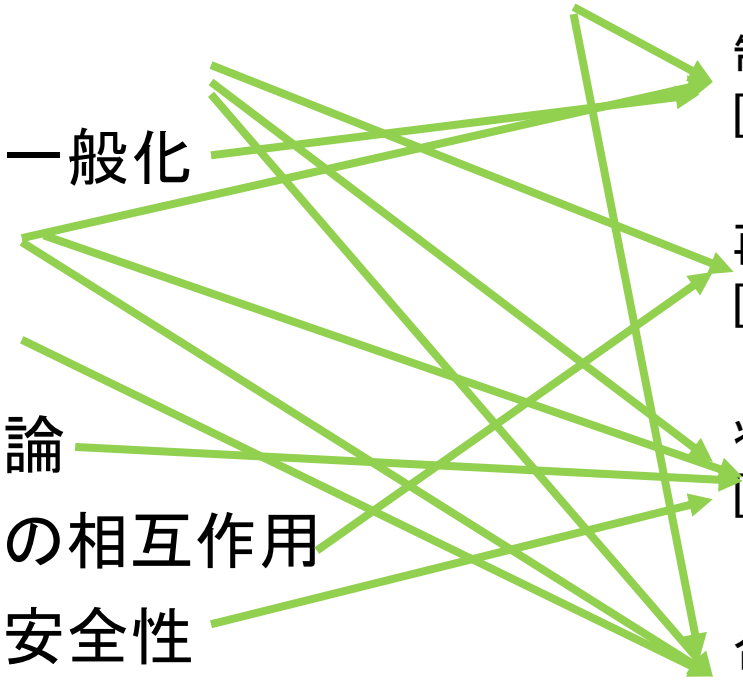
# AAAI 2025 PRESIDENTIAL PANEL ON THE Future of AI Research

- 回答者の大多数（76%）は、「現在のAIアプローチをスケールアップしてAGIを実現する」ことに否定的
- AGIに関連する９個の研究課題：
  - トランスフォーマーを超えるアーキテクチャ
  - 長期的な計画と推論
  - 学習データを超えた一般化
  - 継続学習
  - 記憶と想起
  - 因果推論と反事実推論
  - 身体性と現実世界との相互作用
  - 整合、解釈可能性、安全性
  - 社会的影響の理解と指導


https://aaai.org/wp-content/uploads/2025/03/AAAI-2025-PresPanel-Report-FINAL.pdf

# AAAI 2025 PRESIDENTIAL PANEL ON THE Future of AI Research が挙げる９個の研究課題

- トランスフォーマーを超えるアーキテクチャ
- 長期的な計画と推論
- 学習データを超えた一般化
- 継続学習
- 記憶と想起
- 因果推論と反事実推論
- 身体性と現実世界との相互作用
- 整合、解釈可能性、安全性
- 社会的影響の理解と指導



制限付きベイジアンネット BESOM
[一杉+ 第15回 汎用人工知能研究会, 2020]

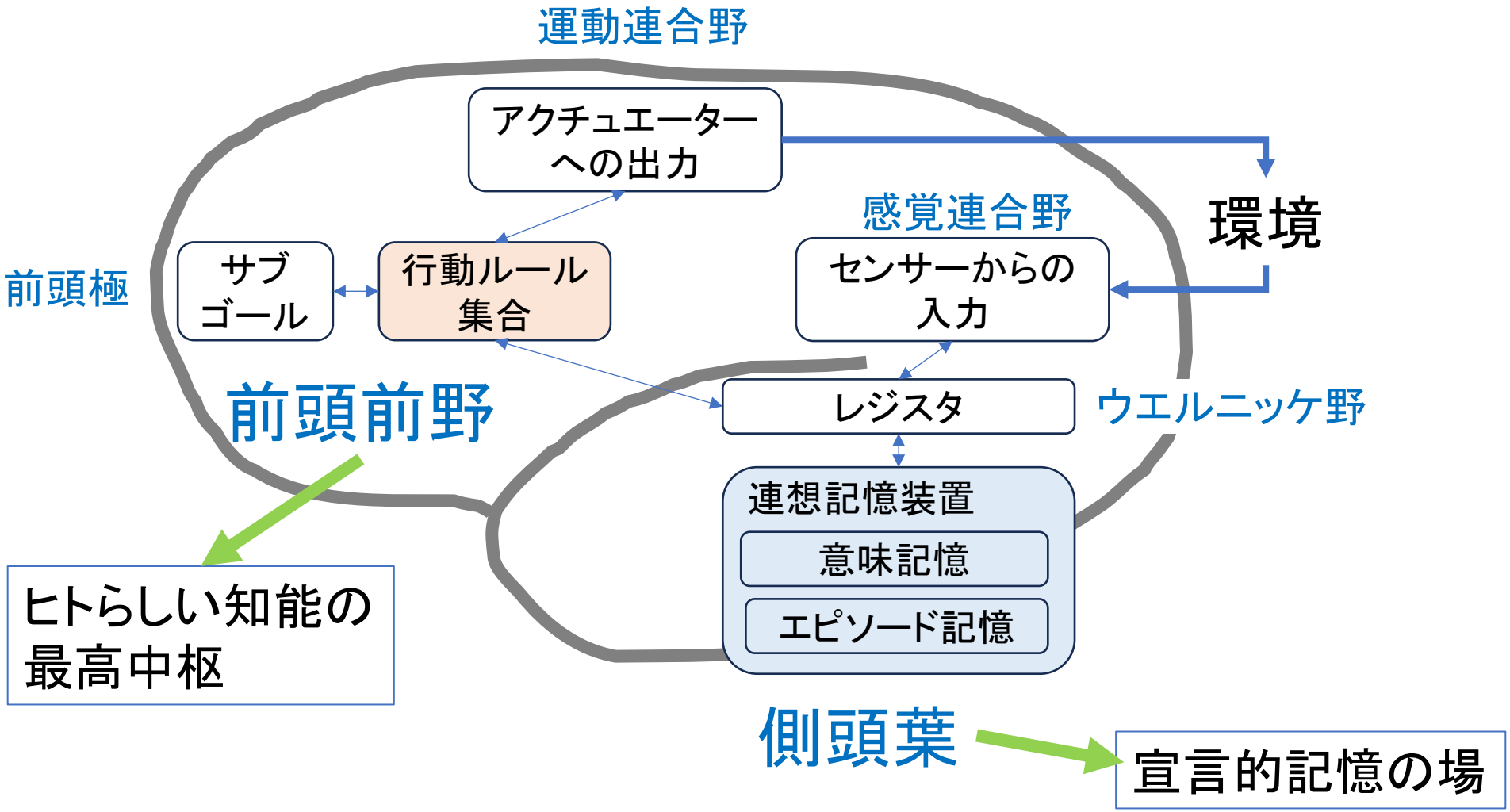
再帰的強化学習 RGoal
[一杉+ 第26回 汎用人工知能研究会, 2024]

状況を扱う知識表現形式
[一杉+ 第27回 汎用人工知能研究会, 2024]

合成対象言語 Pro5Lang と連想記憶装置
[一杉+ 第20回 汎用人工知能研究会, 2022]

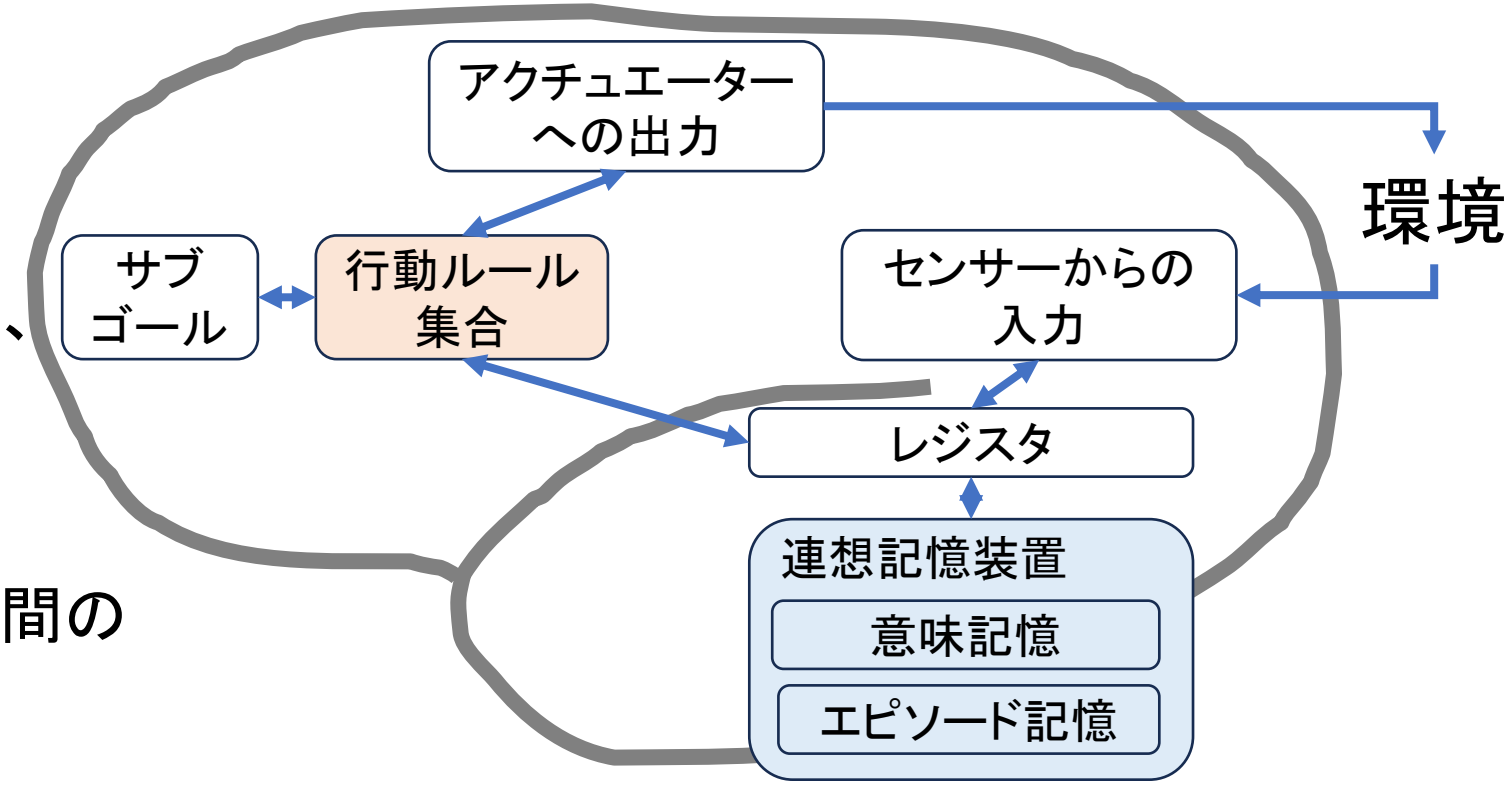
脳をヒントにしたこれまでの取り組み

# 提案する脳型AGIアーキテクチャの全体像

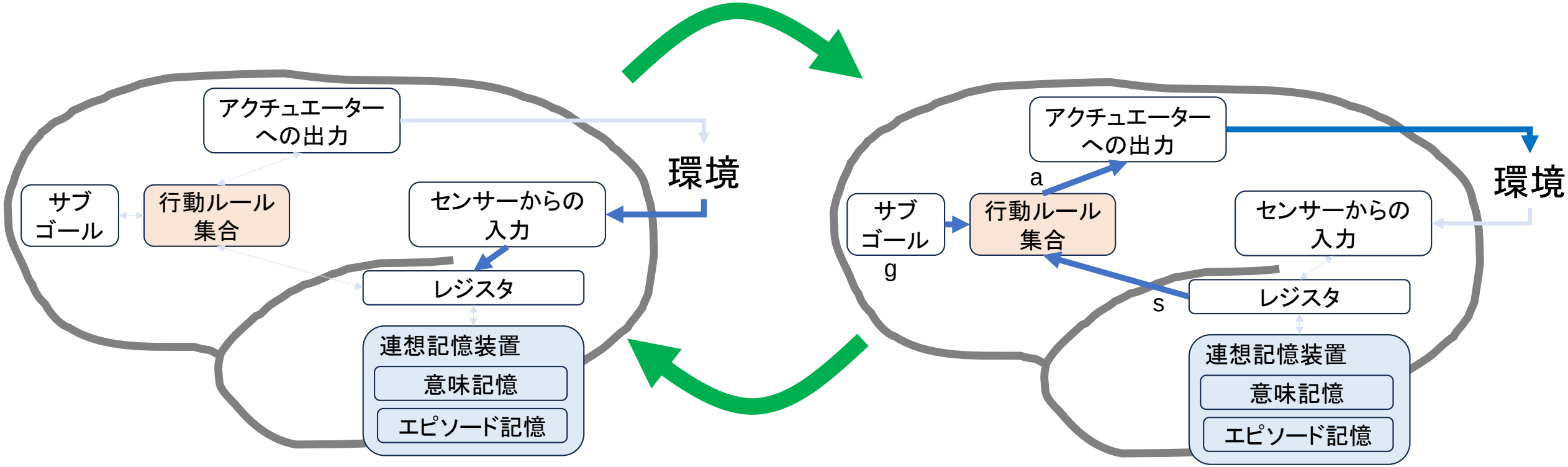


# 行動ルール集合がアーキテクチャの中心

- 行動ルール集合
  ＝ プログラム
  ＝ 行動価値関数 Q(s, g, a) を
  圧縮表現したもの
- 行動、推論、発話、言語理解、
  記憶の想起などの
  「意識的行動」は
  すべて行動ルールに従い実行
- 行動ルールはコンポーネント間の
  ゲートの開閉を制御



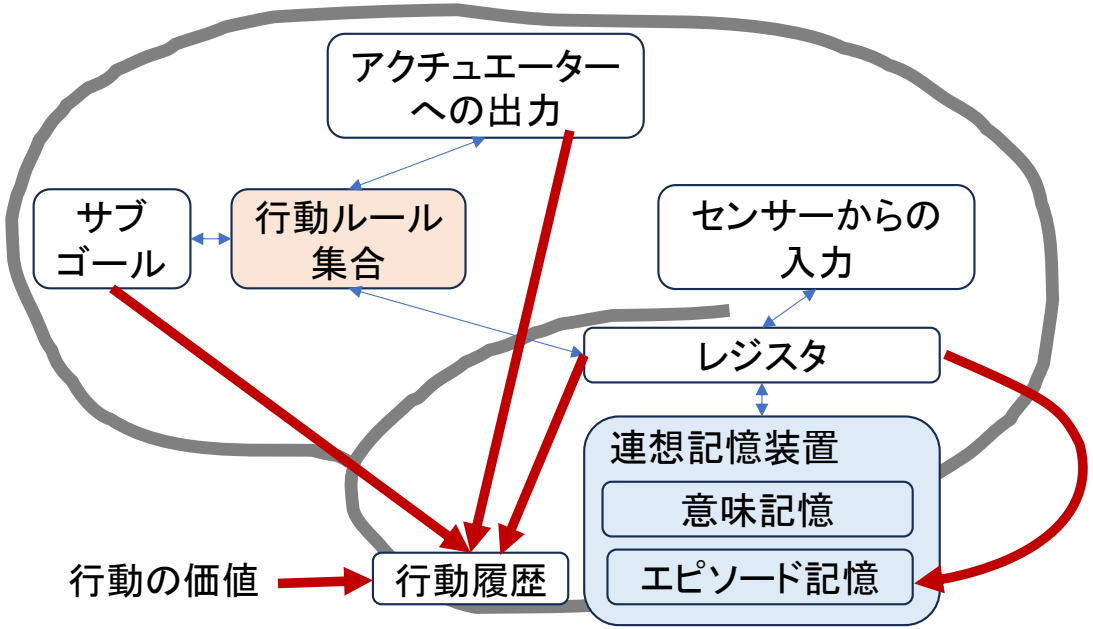
# 実行の１サイクル（200msec?）：ベイジアンネットの推論



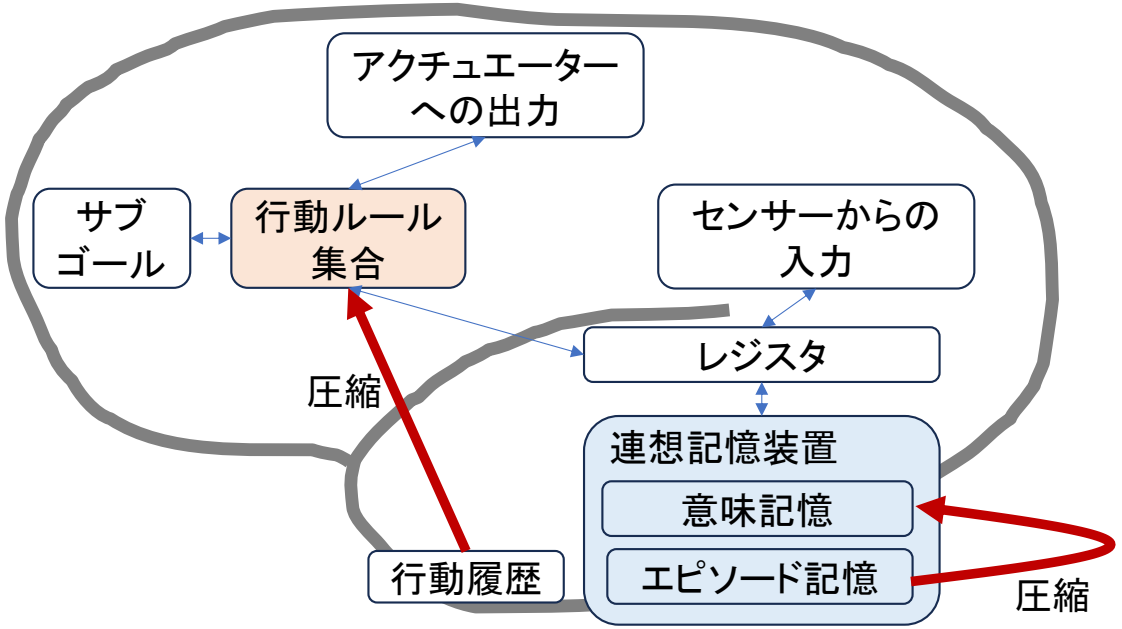

外界を認識し、認識
結果をレジスタに入れる

現在の状態 s, g にマッチする行動ルール Q(s,g,a) のうち
価値が最も高いものを1つ選択し a を実行

# 知識獲得：ベイジアンネットの学習




行動履歴と認識履歴を記憶



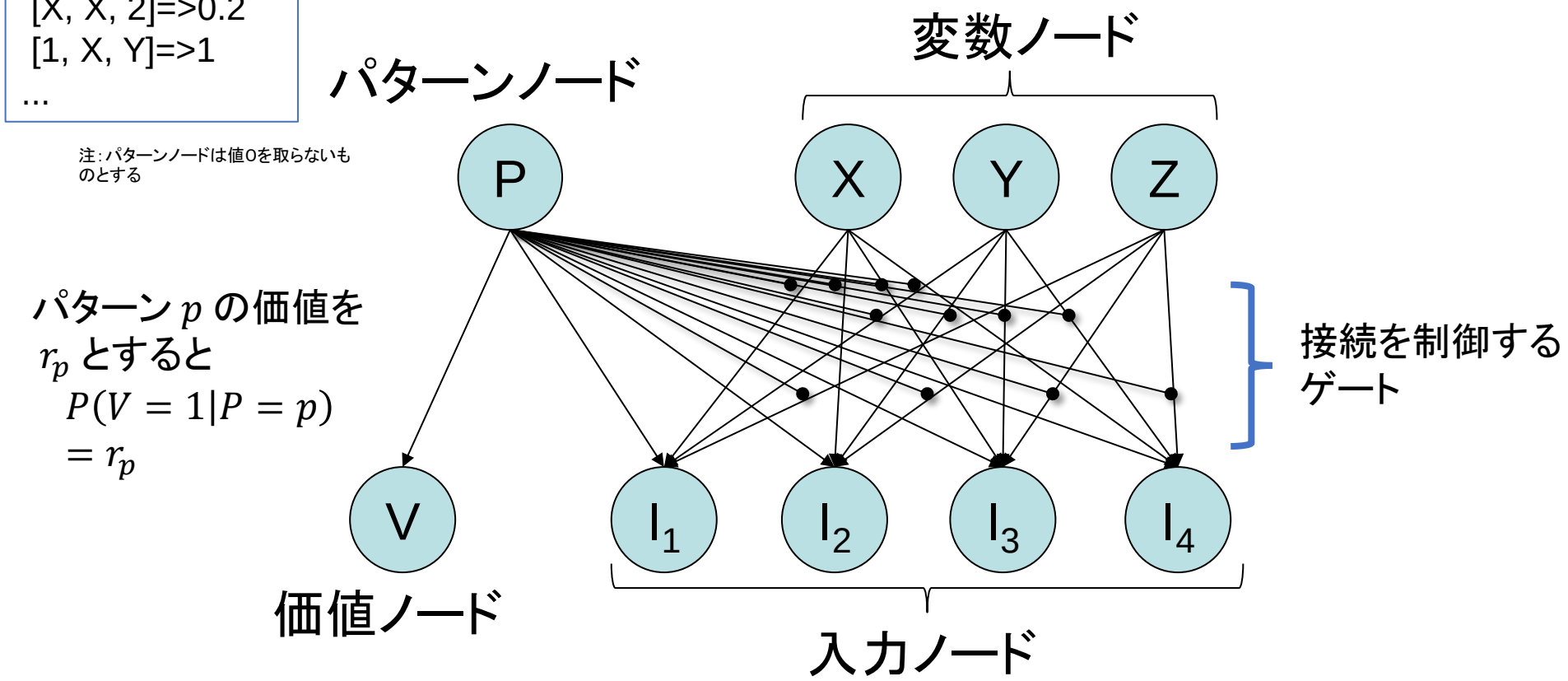

定期的に
行動履歴と認識履歴を圧縮・抽象化し、
手続き的知識と宣言的知識を獲得

[一杉+ 第15回 汎用人工知能研究会, 2020]

# 記号的な手続き的知識、宣言的知識の獲得は特殊な２層ベイジアンネットで実現可能

[X, 1, 2]=>0.1
[X, X, 2]=>0.2
[1, X, Y]=>1
...

注：パターンノードは値0を取らないものとする




パターン $p$ の価値を $r_p$ とすると

$$P(V = 1|P = p) = r_p$$

# プログラム合成と合成対象言語

- プログラム合成は AGI 実現への有望なアプローチ
  - AIXI [Hutter 2000], MagicHaskeller [Katayama 2008, 2018], DNC(Differentiable Neural Computers) [Graves et al. 2016], DreamCoder [Ellis+ 2020]
- どういう「プログラミング言語」を合成対象にするかが性能に大きく影響
  - ヒトが解くべきタスク（物体操作、推論、対話）に特化した言語仕様
  - 強化学習で合成しやすい言語仕様
    - シンタックスエラーになりにくい
    - 無意味なプログラムになりにくい
    - 逐次的な改善が行いやすい
    - …
  - 制御構造は？データ構造は？メモリ管理は？

[一杉+ 第20回 汎用人工知能研究会, 2022]

# 合成対象言語 Pro5Lang

- 機械語と論理型言語の特徴を合わせ持った**手続き型言語**
- **機械語に似た特徴**：
  - 固定長のデータ、固定長の命令
  - 少数のレジスタと大容量のメモリ
  - 神経回路での実現が容易
- **論理型言語に似た特徴**：
  - パターンマッチングによる行動ルール選択
  - （学習が理想的に進めば）正しい推論をする
- プログラム（行動ルール集合）は
  教師なし学習と強化学習で獲得

```
 1: k(e(0, 0, Socrates, Is, AMan, 0));      // ソクラテスは人間である
 2: k(e(If, c(0, 0, x, Is, AMan, 0), c(0, 0, x, Is, Mortal, 0))); // 人間は死ぬ

 3: eifxy = e(If, c(x1,x2,x3,x4,x5,x6), c(y1,y2,y3,y4,y5,y6));
 4: eif0y = e(If, c(__,__,__,__,__,__), c(y1,y2,y3,y4,y5,y6));
 5: ax = a(x1,x2,x3,x4,x5,x6);
 6: ay = a(y1,y2,y3,y4,y5,y6);
 7: rule(w(),       w(ay), recall(eif0y));  // y を証明するために x → y を想起
 8: rule(w(eifxy),  w(ay), call(ax));       // x → y なら y を証明するために x を証明
 9: rule(w(ax),     w(ay), recall(eifxy));  // x なら y を証明するために x → y を想起
10: rule(w(ax, eifxy), w(ay), set(ay));     // x, x → y なら y が成り立つ

11: ex = e(x1,x2,x3,x4,x5,x6);
12: rule(w(),   w(ax), recall(ex));
13: rule(w(ex), w(ax), set(ax));
```

プロトタイプ版 Pro5Lang のプログラム例

# Pro5Lang と強化学習

- 行動価値関数 Q(s,g,a) を圧縮抽象化したものが
  Pro5Lang プログラム ＝ 行動ルール集合
- 価値の高い行動ルールを選択し実行 → 報酬最大化
- サブルーチンを再帰的に呼び出すことが可能
  - 再帰的強化学習アルゴリズム RGoal [一杉+ 第26回 汎用人工知能研究会, 2024]
- サブルーチンの価値の定義： MAXQ 価値観数分解[Dietterich 2000]の拡張
- **推論規則の「正しさ」を「価値」で代用**[一杉+ 第14回 汎用人工知能研究会, 2020]
  → 経験から「正しい知識」を獲得可能に

# 再帰的強化学習の枠組みの中で<br>記号推論システムを近似的に実現

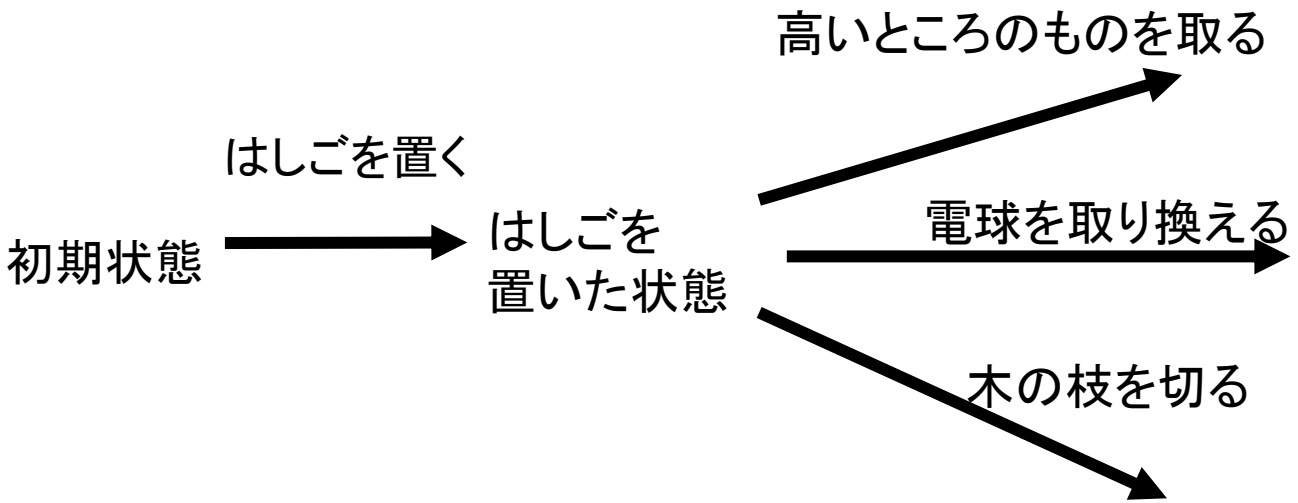
- 論理体系は推論規則の集合<br>**前提１，前提２　⊢ 結論**
- **推論規則の「正しさ」を「価値」で代用**[一杉+ 第14回 汎用人工知能研究会，2020]<br>推論の結果のちのち悪いことが起きる → 低い価値<br>推論の結果のちのち良いことが起きる → 高い価値

記号の意味が推論規則集合を通じて間接的に環境に接地

- 学習する論理体系はあくまで近似
  - 完全性や無矛盾性にしばられない**ゆるい**推論システム
  - 高階論理や非古典論理の多くを取り込んだ非常に高い表現力

→ **反実仮想的状況も含む世界の森羅万象についての推論が可能**

# サブルーチンの例




最終目的は違っていても、
「はしごを置く」というサブルーチンは共通

サブタスク共有

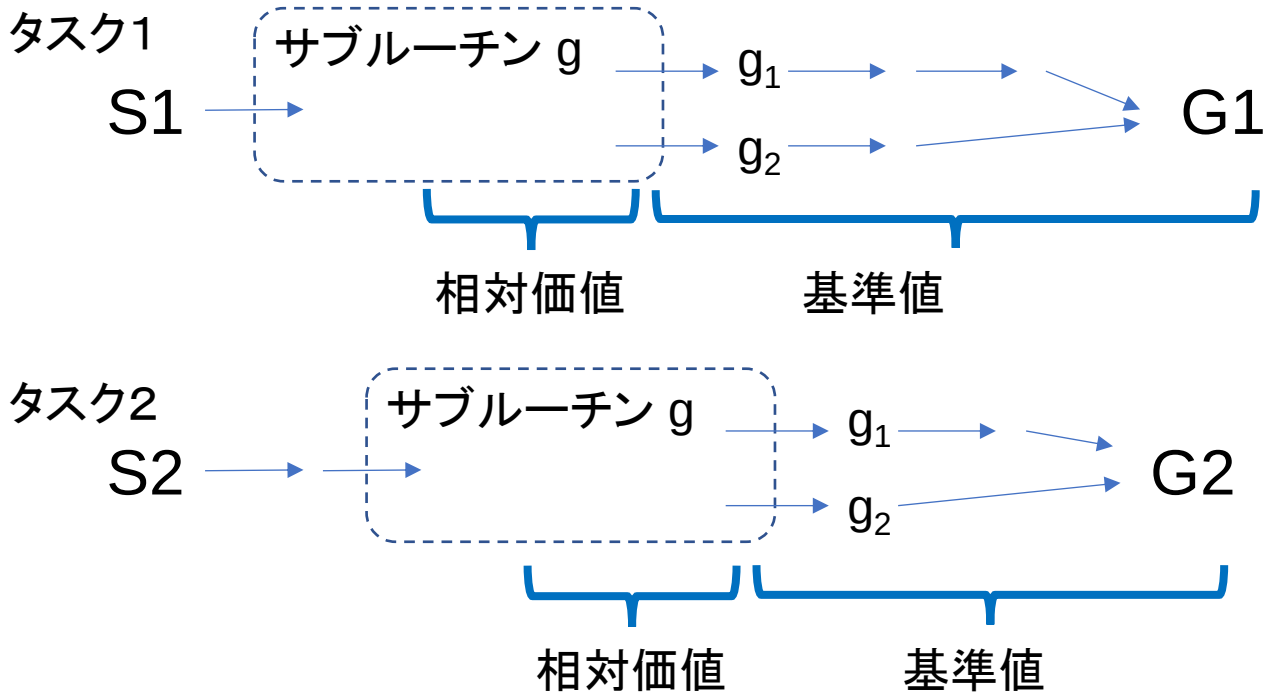
「はしごを置く」サブルーチンを実行中は
はしごの周辺以外の情報をすべて無視

状態抽象

[一杉+ 第26回 汎用人工知能研究会, 2024]

# RGoal における「価値」

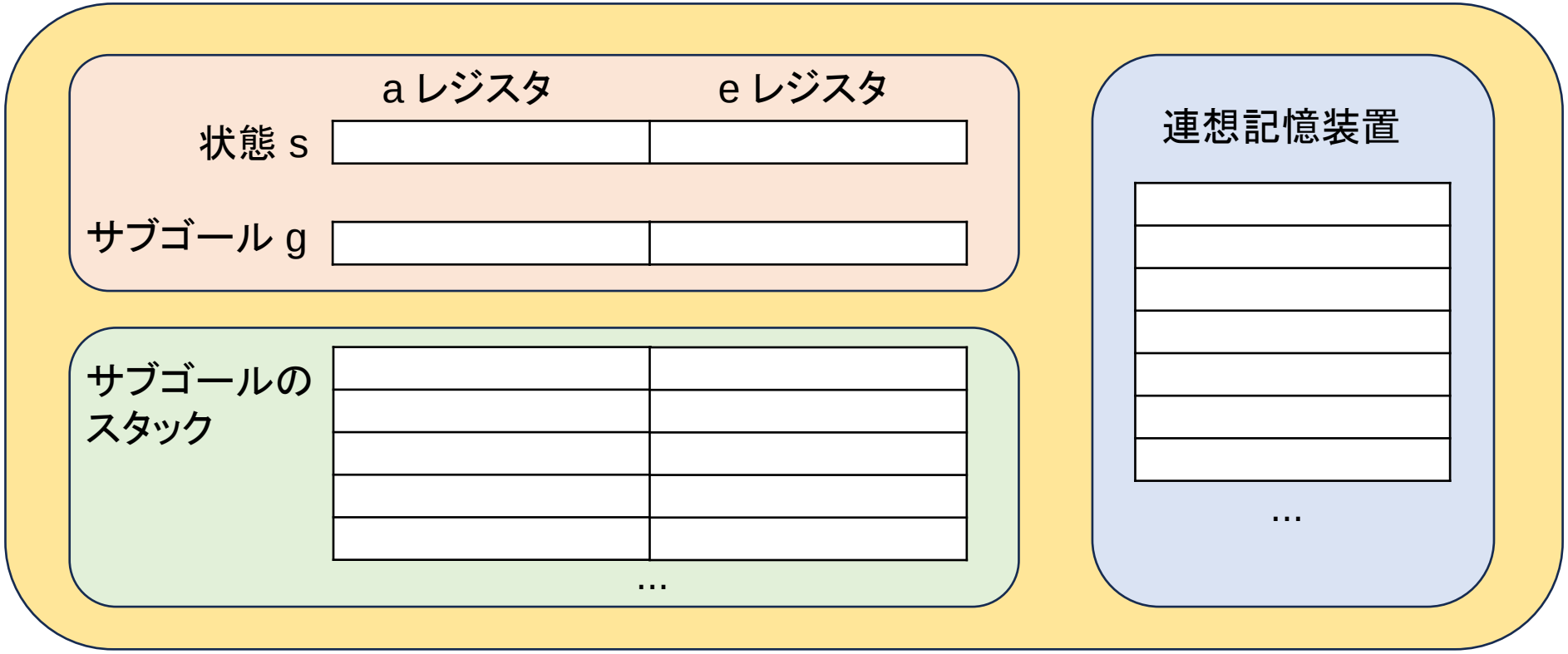
MAXQ 価値観数分解[Dietterich 2000]の拡張




- サブルーチン終了時の価値を基準値とした相対価値を学習
  → 異なるタスク間でサブルーチンの最適方策を共有可能
- のちのち高い報酬が得られるような終わり方でサブルーチンを終わるように学習

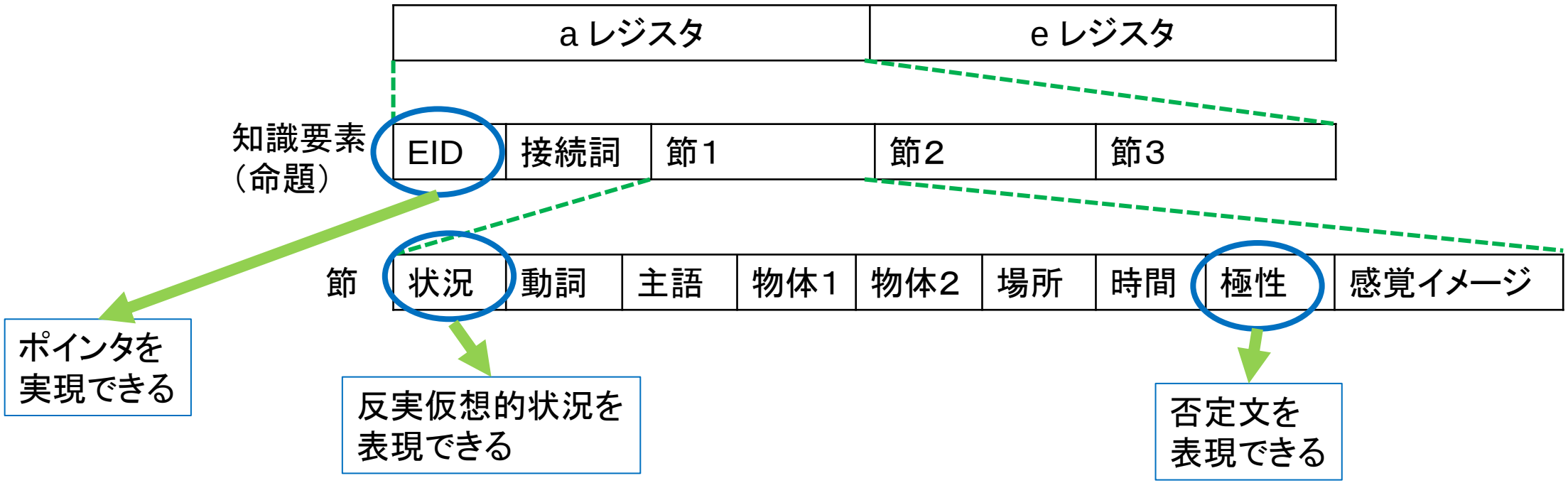
# Pro5Lang のメモリ機構の特徴

- 情報処理の最小単位は固定長の「命題」、命題集合で世界を表現
- 各ステップの認識・推論の結果は自動的に大容量の連想記憶装置に追加
- パターンマッチングにより想起



[一杉+ 第27回 汎用人工知能研究会, 2024]

# 情報処理の基本単位：知識要素(Knowledge element)

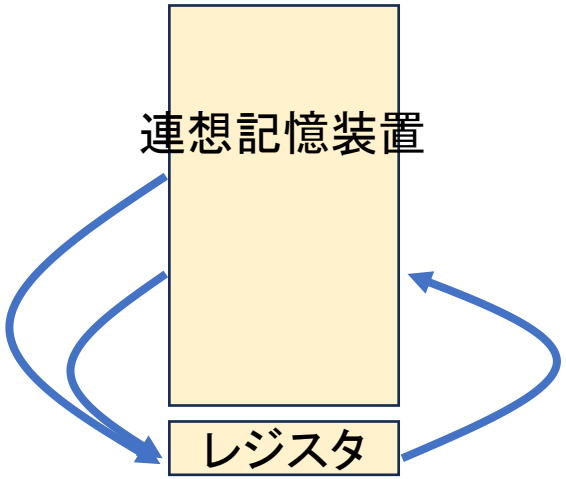



unknown を表現できる

# Pro5Lang プログラムと LLM の動作の類似点

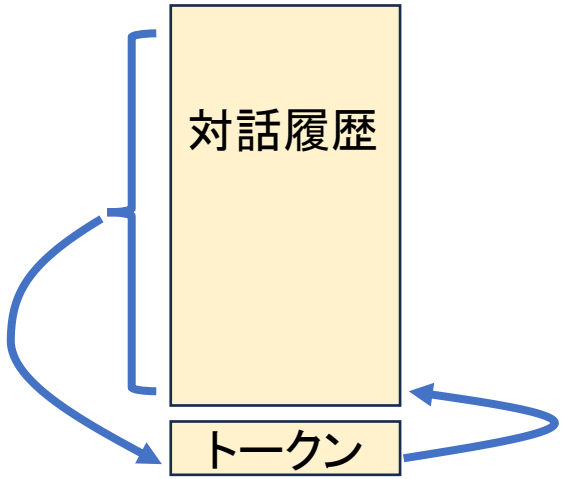
| | Pro5Lang プログラム | LLM |
|---|---|---|
| 各ステップの動作 | 新たな命題を推論 | 次のトークンを推論 |
| プログラムの実体 | 学習で獲得した行動ルール集合 | パラメタの値 |
| メモリへの書き込み | 推論結果の追加 | 推論結果の追加 |
| メモリの読み出し | クエリによる検索 | self-attention による検索 |

**メモリの読み書きに制約があるため解の探索空間が狭くプログラム合成が容易**




連想記憶装置内の命題から新たな命題を推論

推論した命題を連想記憶装置に追加




対話履歴からトークンを推論

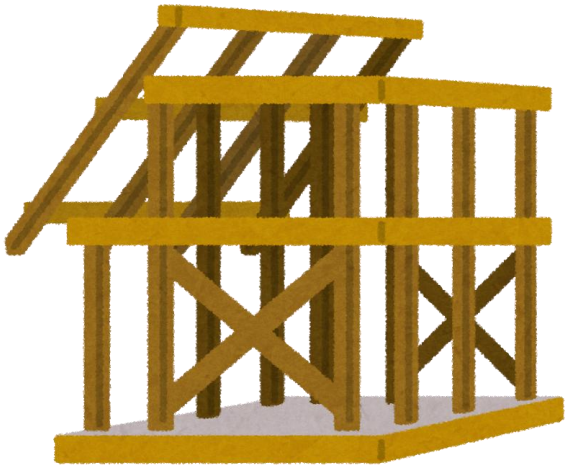
推論したトークンを対話履歴に追加

# 未解決の課題

- 設計したアーキテクチャ全体を実装してデモ
- 大脳皮質モデルBESOMの大規模化
- 再帰的強化学習アルゴリズム RGoal の学習の効率化・安定化
- Pro5Lang の推論機構の効率化
- メタ認知の機構
- タイムアウト、割り込みの機構
- …

# まとめと今後

- 脳の前頭前野周辺を極力単純化した脳型AGIアーキテクチャを提案
  - 強化学習、ベイジアンネット、数理論理学が主な理論的基盤
  - 神経科学、言語学などの知見も踏まえて設計
- アーキテクチャ全体の動作デモは未実装、
  個々の要素技術については最低限の動作を確認済み
  - この骨組みに肉付けしていく研究者が増えれば脳型AGIが実現可能では
- ２０２５年度末で引退予定
- 今後は情報発信・アーキテクチャの解説に注力
- これまでの発表資料：
  https://staff.aist.go.jp/y-ichisugi/j-index.html
  質問・コメント等を歓迎いたします

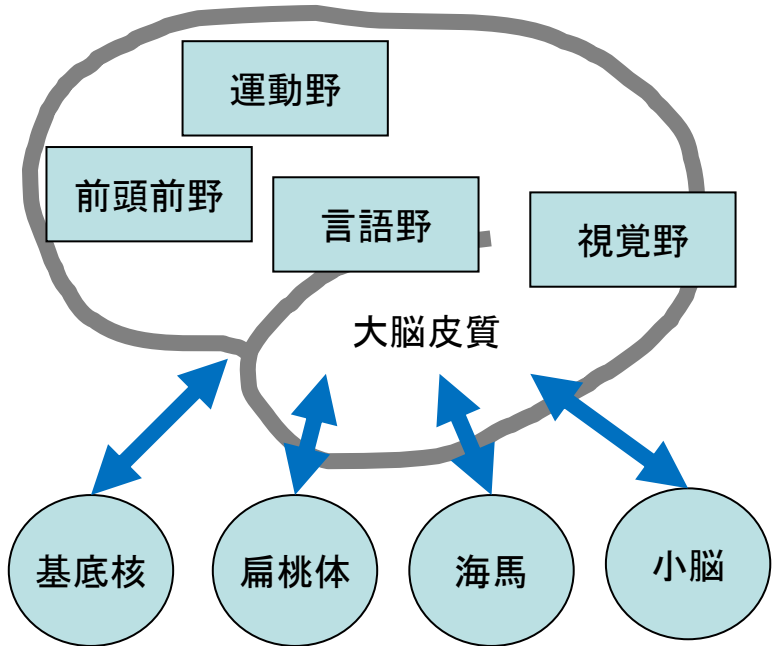


建築中の家のイラスト | かわいいフリー素材集 いらすとや

以上

# 私の研究の長期的目標

- 人類を豊かにするためには人間の仕事を低コストで代替できる機械が必要
- 脳のアーキテクチャをヒントにするのが近道
  → 脳型AGI
- AGI研究に携わる研究者を増やす



# 現状の大規模言語モデル(LLM)の問題点

- **数学的推論能力が弱い** [Mirzadeh+ arXiv:2410.05229]
  - GPT-4o や o1-preview ですら、
    問題に無関係な情報を追加すると正解率低下
  - 論理的推論ではなく主にパターンマッチング
- **ハルシネーション、ジェイルブレイクの問題**
  - 何が「事実」なのか、何が「悪いこと」なのかを、
    正確に学習できていない・正確に推論できない
- パラメタサイズとデータを増やせば解決？ → **すでに限界**
- 思考時間を増やせば解決？ → **不正確な推論を積み重ねても無意味**

# 私が考える現状の LLM の問題点

- **コンテキストウインドウが広い＝パラメタサイズが多い＝汎化性能悪化**
  - 文脈に依存した「それっぽい答え」を出すには向いているが、論理的な推論規則を獲得するには向かない
- **「内部状態」をテキスト表現の対話履歴の形でしか持てない**
  - テキスト表現はあいまい、知識の追加と検索が非効率的かつ不正確
  - トークン列で表現できない感覚イメージなどを直接表現できない
- **プログラムとデータの表現が全く違う**
  - （プログラム＝パラメタの値、 データ＝トークン列）
  - 「自己書き換えコード」のようなことがやりにくい
  - 対話・読書等を通じた継続的な知識獲得がやりにくい

これらの問題が、ヒトレベルの知能への到達の障害になるのではないか？

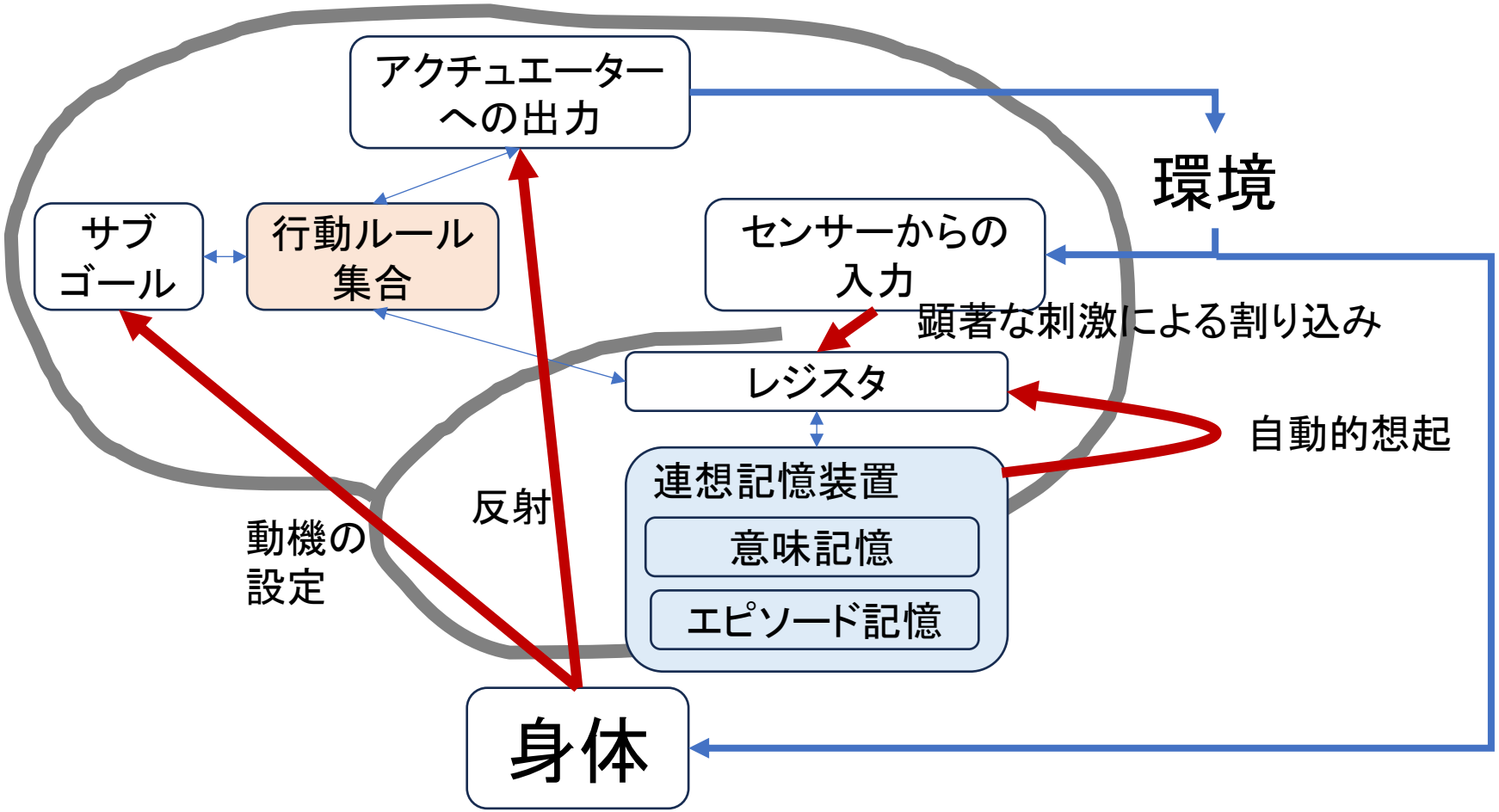
# 研究の前提

- AGIがいまだにできてない理由は汎用性が足りないのでなく、
  むしろ人間が得意とするタスクへの特殊化が足りない。
  （Cf. ノーフリーランチ定理[Wolpert+ 1997]）

- 計算パワーだけに頼ってAGIが知識獲得することは不可能。
  事前知識の作り込みや他者からの知識伝達が重要。

- 人類がまず最初に必要とするＡＧＩは、人間に解けない問題を解く機械
  ではなく、人間が仕事のやり方を教えればそれを人間程度の上手さで実
  行してくれる機械。

# 行動ルールの制御下にない自発的な情報の流れの例



[一杉+ 第28回 汎用人工知能研究会, 2024]

# テスト用タスクのシナリオ：Key を Box に適用して Gold を得るタスク

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

「宝箱・宝物のイラスト | かわいいフリー素材集 いらすとや」
https://www.irasutoya.com/2013/05/blog-post_2133.html