

第5回 全脳アーキテクチャ勉強会

「意思決定 – 深いゴール探索と深い強化学習の技術をヒントにして、前頭前野の機構の解明を目指す – 」

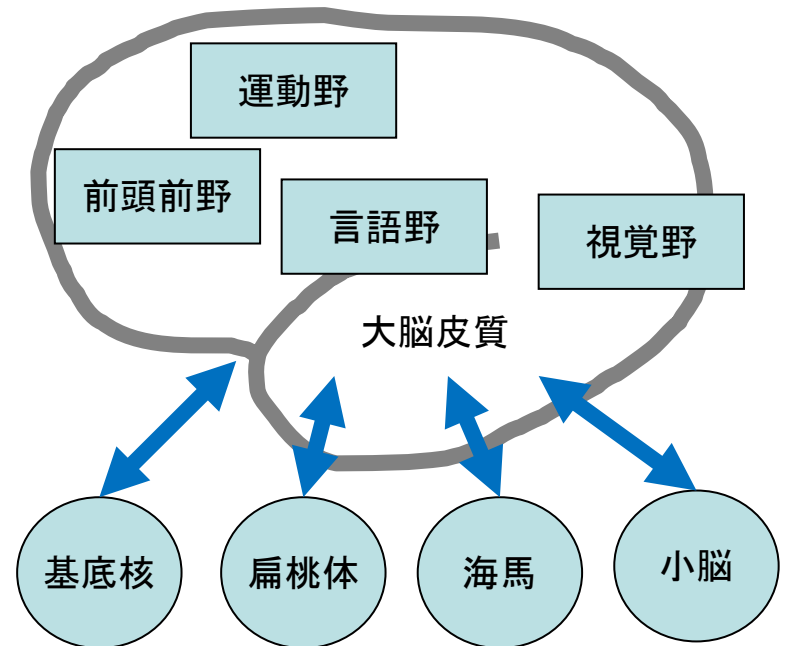
- 18:30-19:00 「Deep Learning とベイジアンネットと強化学習を組み合わせた機構による、前頭前野周辺の計算論的モデルの構想」(産総研 一杉裕志)
- 19:00-19:40 「BDI –モデル、アーキテクチャ、論理–」(奈良女子大 新出尚之)
- 19:40-19:50 休憩
- 19:50-20:30 「強化学習から見た意思決定の階層」(グーグルジャパン 牧野貴樹)

全脳アーキテクチャ(WBA)アプローチの中心仮説

- 「脳の機能はいくつかの機械学習装置が連携することで実現されている。それを模倣することで、脳の機能を持つ汎用の知能ロボットが実現可能である。」

神経科学と機械学習の最低限の基礎知識を持った研究者を増やすことが最優先課題。

そういう研究者を増やす活動をする人を増やすことが最優先課題。



社会へのインパクト

- **ロボットによる労働支援により、人間の労働生産性が限りなく増大。**
 - 富の再配分が正しく行われ、かつ資源制約の問題が解決されれば、人類は限りなく豊かになる。
 - 1人1人すべての人間が貴族のような生活。
 - すべての人に主治医と家庭教師と専属弁護士。
- **核融合研究等と同様に、実現すれば人類に莫大な利益をもたらす研究分野。**

参加者のみなさまにお願いしたいこと

- 質問大歓迎。できれば手短かに。1人でも多く。
- 懇親会やネットでもどんどん質問してください。
 - Facebook グループ「全脳アーキテクチャ」
<https://www.facebook.com/groups/whole.brain.architecture/>
 - 私のブログ「BESOM(ビーソム)ブログ」
<http://besom1.blog85.fc2.com/>
- まだほとんどの人が傍観者状態。他の人から見える行動をぜひともお願いします。
 - 「BESOM(ビーソム)ブログ 傍観者効果」
<http://besom1.blog85.fc2.com/blog-entry-156.html>
 - 批判も歓迎。

Deep Learning とベイジアンネットと強 化学習を組み合わせた機構による、 前頭前野周辺の計算論的モデルの 構想

第5回 全脳アーキテクチャ勉強会

2014-07-01

産業技術総合研究所

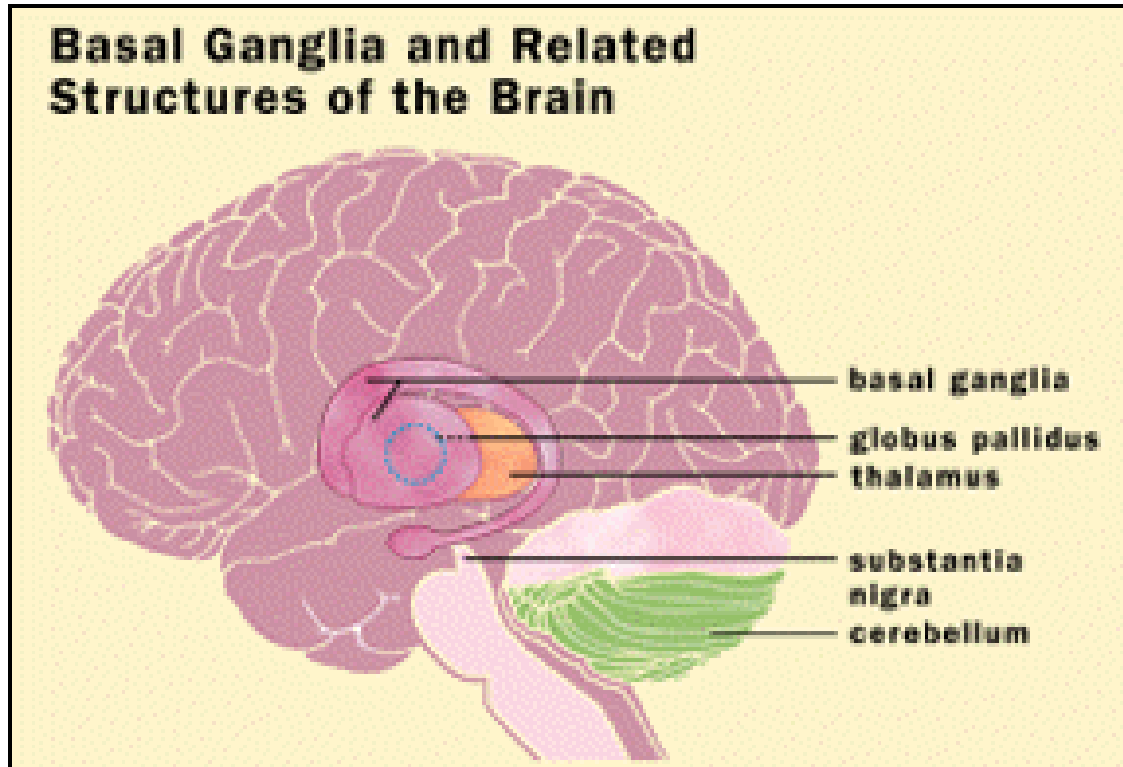
一杉裕志

アウトライン

- 前頭前野周辺に関する神経科学的知見
- 前頭前野周辺の計算論的モデル構築の構想

前頭前野周辺に関する知見

大腦基底核(basal ganglia)



Wikipedia

http://upload.wikimedia.org/wikipedia/commons/e/ee/Brain_structure.gif

原典 Henkel, John (July-August 1998). "Parkinson's Disease: New Treatments Slow Onslaught of Symptoms". FDA Consumer: p. 17. ISSN 0362-1332. Retrieved on 2009-07-23.

大脳皮質－基底核ループは 強化学習に関与

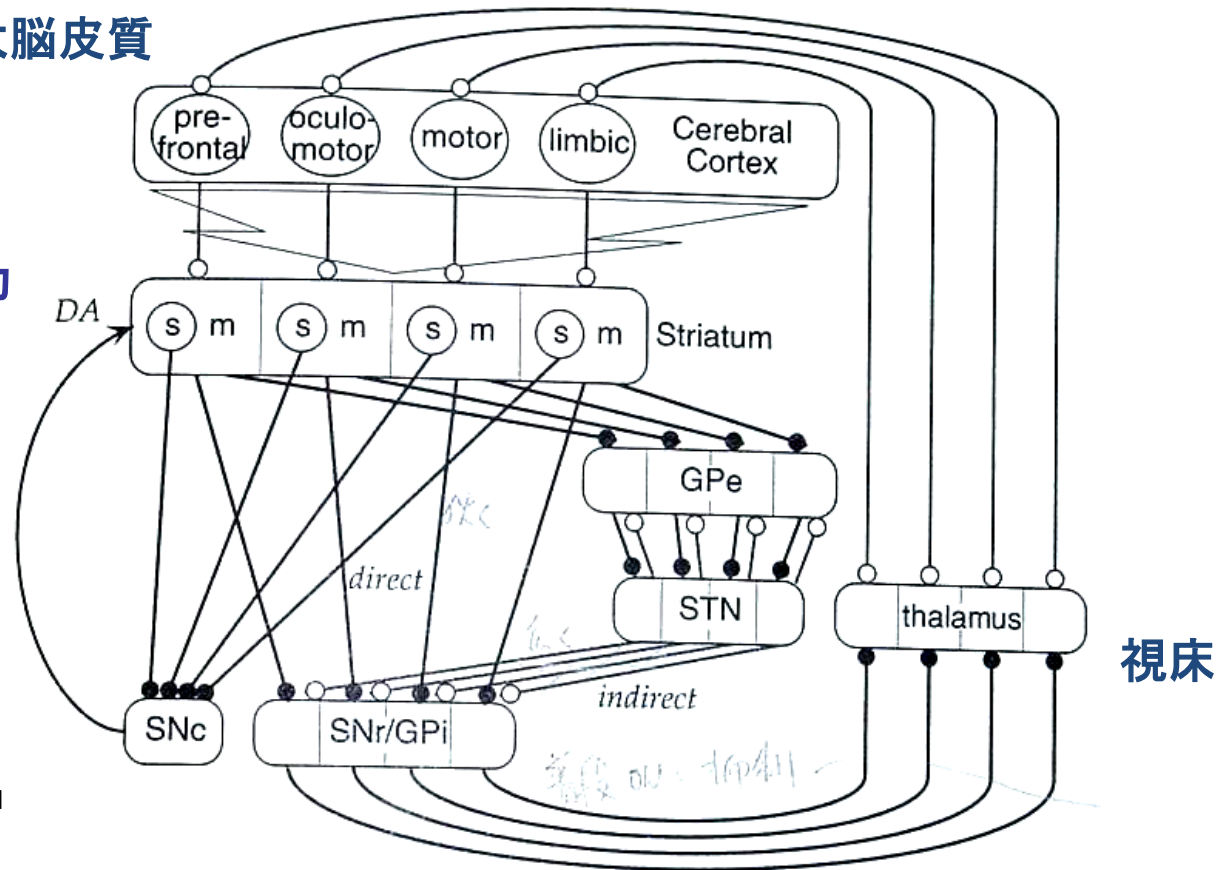
大脳皮質

線条体：
行動価値を出力

K. Samejima, K. Ueda, K. Doya, and M. Kimura.
Representation of action-specific
reward values in the striatum. Science, 2005.

黒質緻密部：
TD誤差を出力

W. Schultz, P. Dayan, and P. R. Montague. A neural
substrate of prediction and
reward. Science, Vol. 275, pp. 1593–1599, 1997.



甘利 俊一, 外山 敬介 (編集) 出版社: 朝倉書店
脳科学大事典 p.607 大脳基底核のモデル (銅谷賢治)

4つの並行したループ

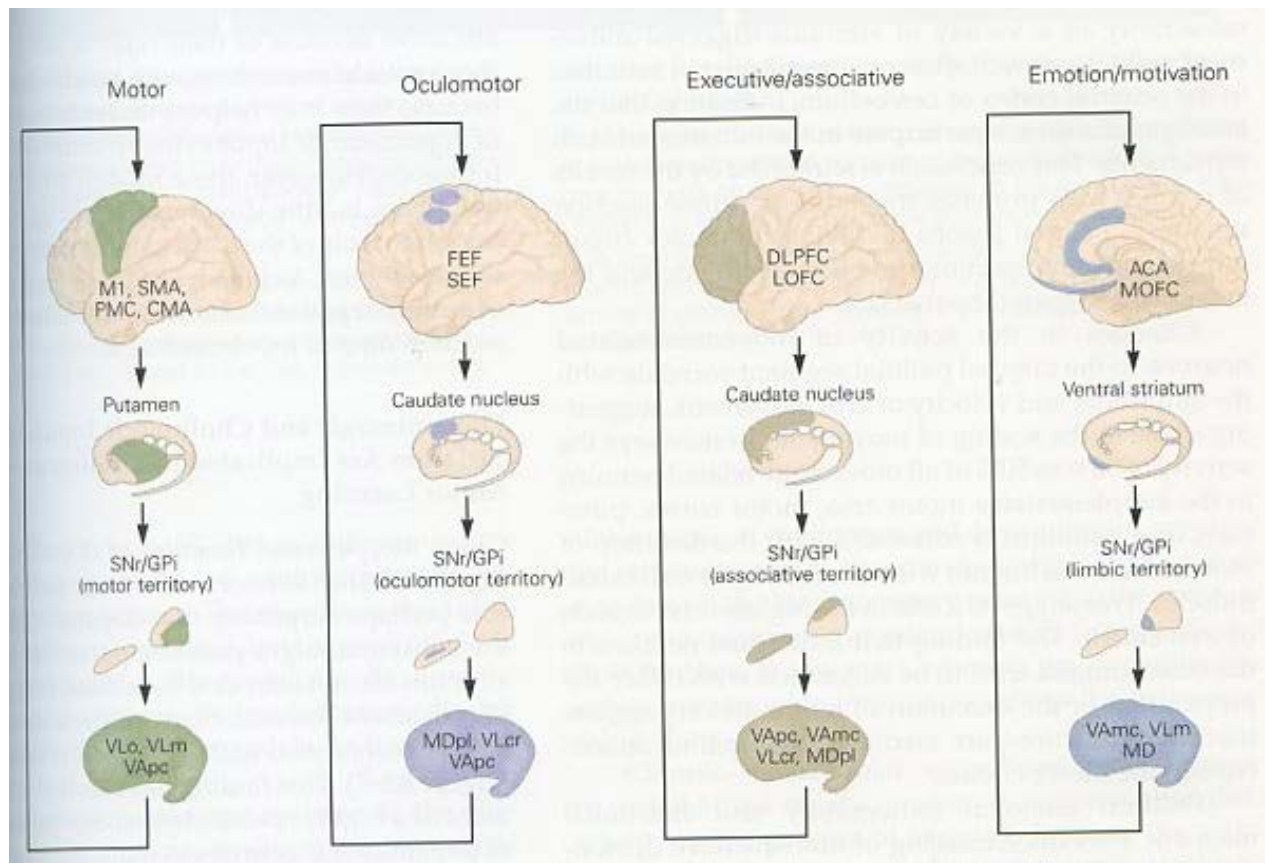
関係する主な
領野と機能:

M1・PM
運動

FEF
眼球運動

DLPFC
実行機能・連合

ACC
情動・動機



大脳皮質

線条体

黒質網様体・
淡蒼球内節

視床

Principles of Neural Science 5th ed.

Eric R. Kandel et al. , McGraw-Hill

p. 987

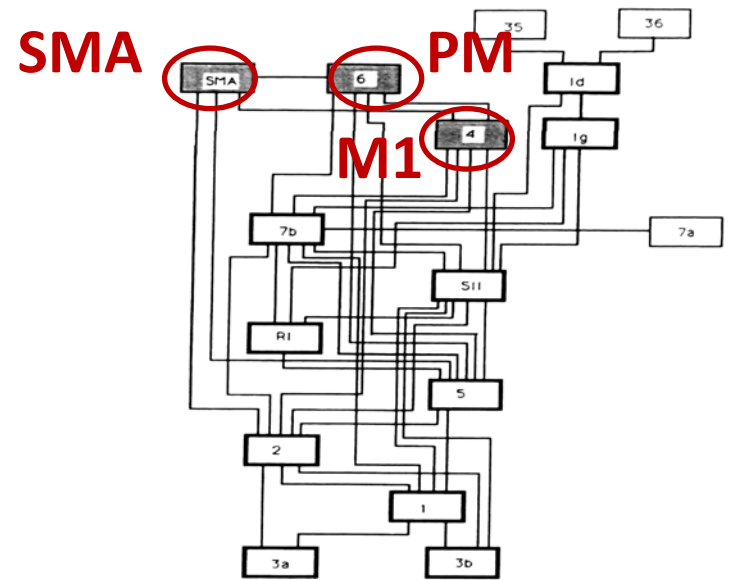
運動野

- **M1** (一次運動野、4野)
 - 運動出力との結びつきが強い。
- **PM** (運動前野、6野)
 - ネクタイをしめる、など習熟した動作に関与。



運動野のホムンクルス

http://ja.wikipedia.org/wiki/%E3%83%95%E3%82%A1%E3%82%A4%E3%83%AB:Motor_homunculus-ja.png

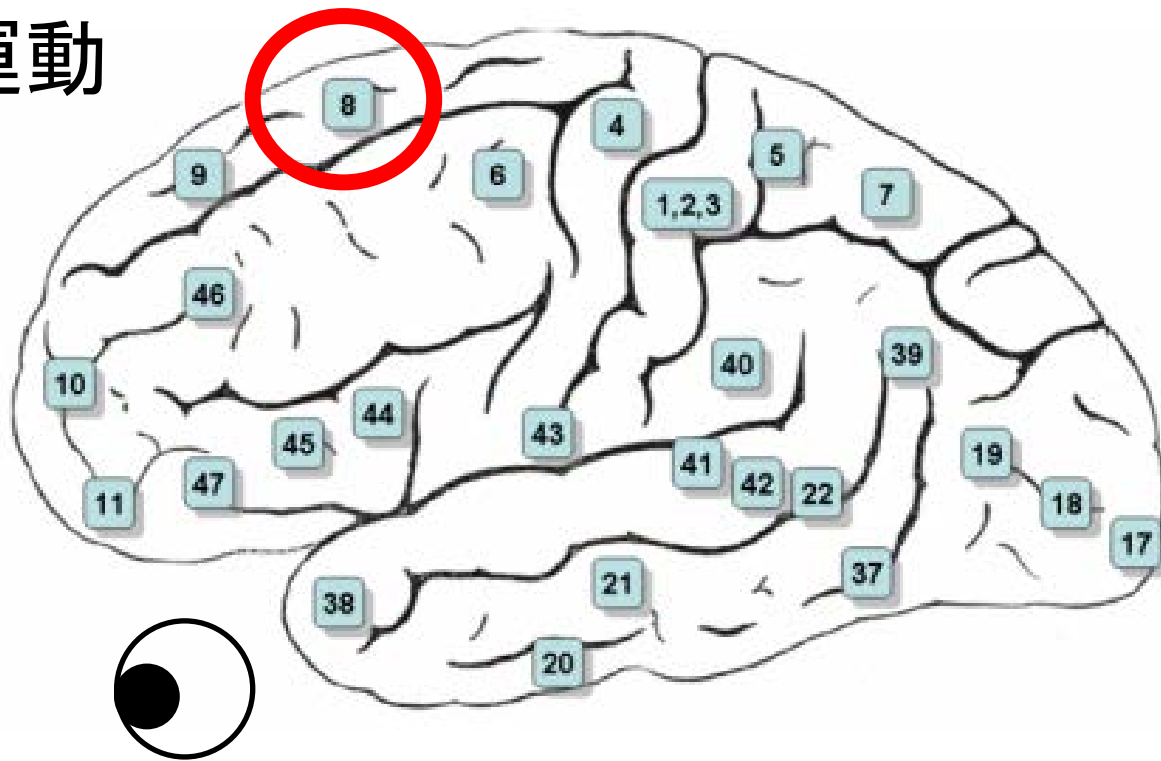


参考:「脳と運動—アクションを実行させる脳」
丹治 順 共立出版 (1999/11)
ASIN: 432005394X

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

前頭眼野(**FEF**, frontal eye field)

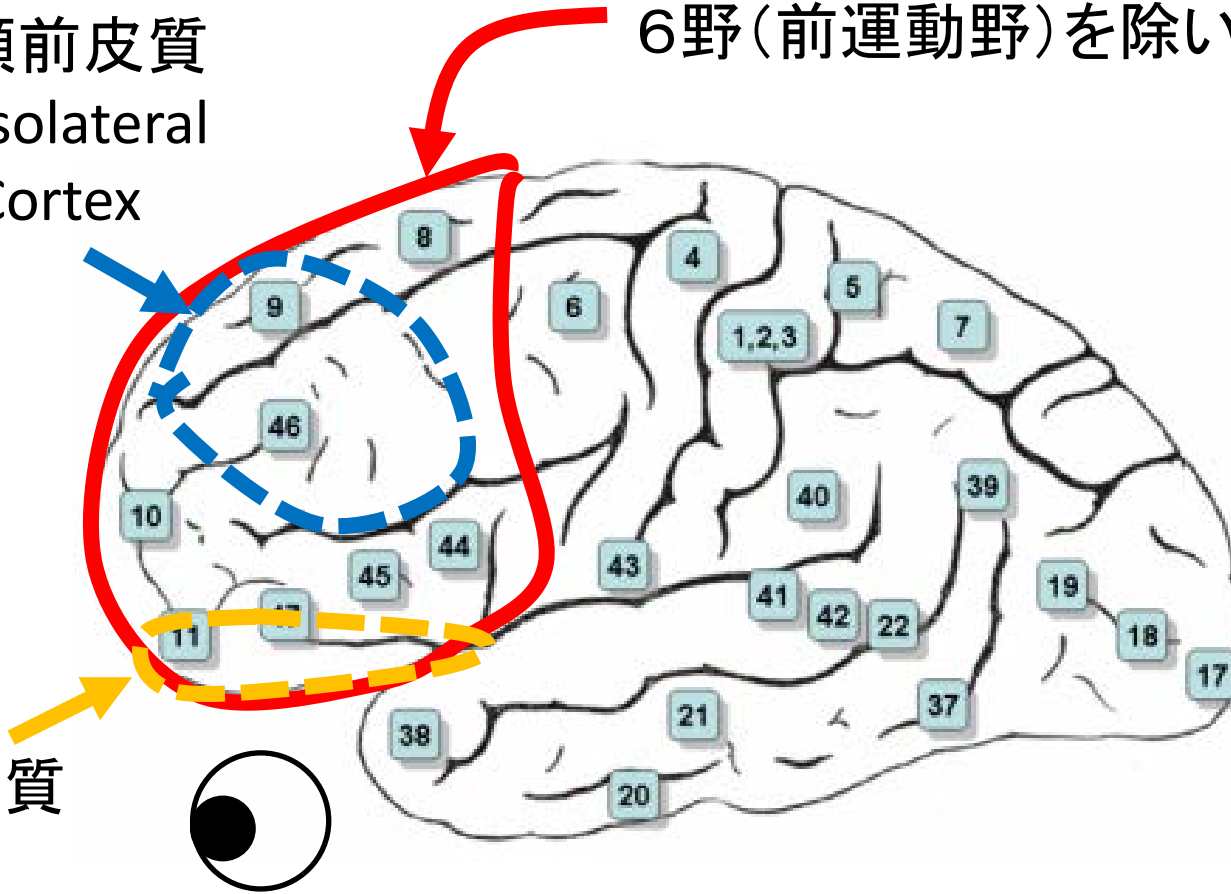
- ブロードマン8野
- 眼球運動



前頭前野(PFC, Prefrontal cortex)

前頭葉のうち4野(一次運動野)と6野(前運動野)を除いたところ。

背外側前頭前皮質
DLPFC, Dorsolateral
Prefrontal Cortex



眼窩前頭皮質
OFC,
Orbitofrontal cortex

前頭前野(PFC)の機能

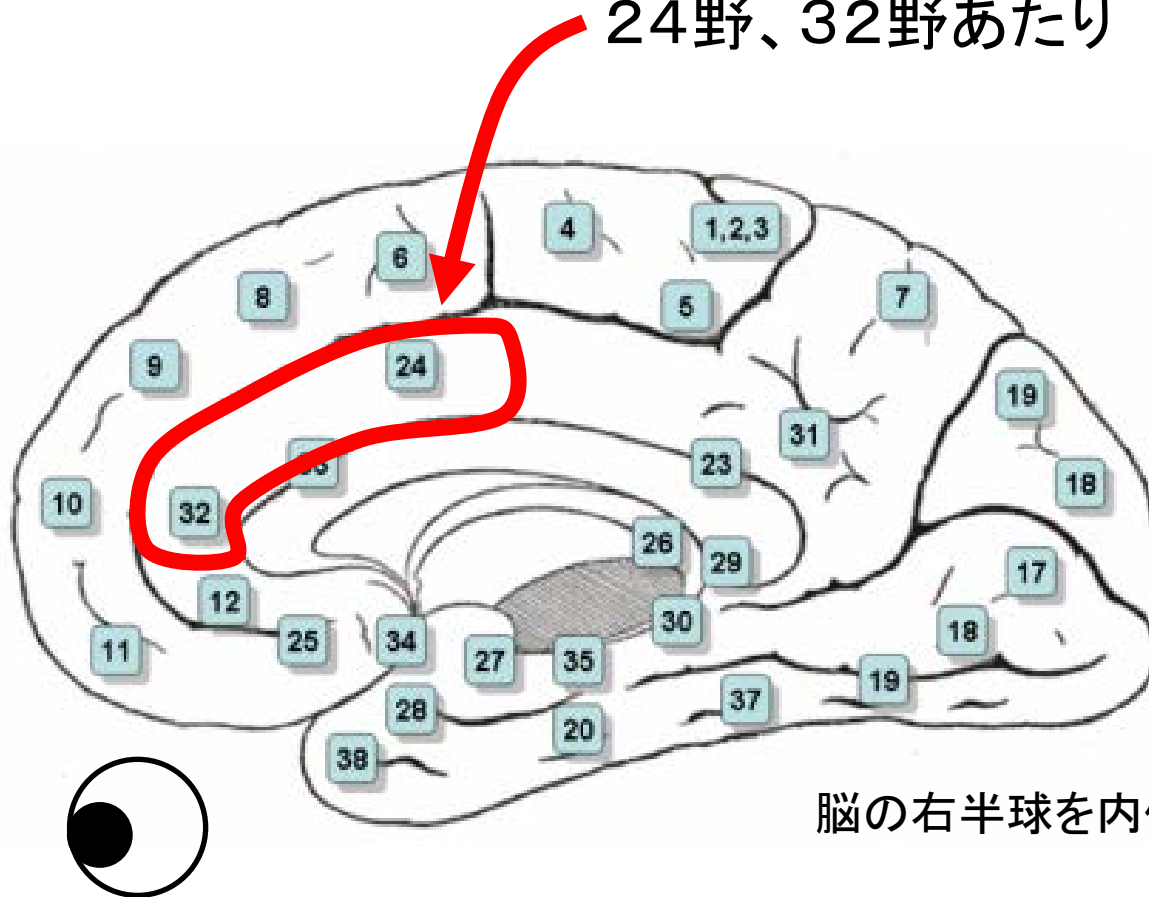
- 「前頭前野はヒトをヒトたらしめ、思考や創造性を担う脳の最高中枢であると考えられている。」
- 脳科学者ペンフィールドの姉の例：
「彼の姉は前頭前野に脳腫瘍ができたため、その切除手術を受けた結果、例えば「料理」のような行動が困難になったことが報告されている。」
「前頭前野に損傷を受けると、このような順序だった行動の組立をする、つまり段取りをうまくとる事ができなくなってしまうのである。」

「前頭前野 - 脳科学辞典」

<http://bsd.neuroinf.jp/wiki/%E5%89%8D%E9%A0%AD%E5%89%8D%E9%87%8E>

前帯状皮質(ACC, Anterior cingulate cortex)

24野、32野あたり

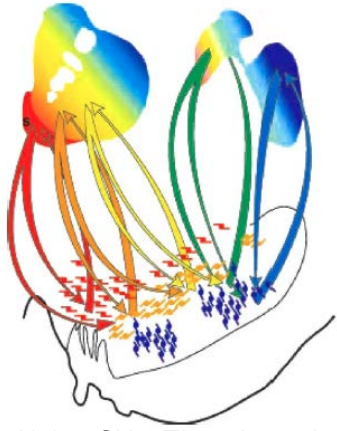


脳の右半球を内側から見たところ

前帯状皮質(ACC)の機能

- **非常に多くのことに関与。複雑でわかりにくい。**
 - 「刺激のトップダウンとボトムアップの処理や他の脳領域への適切な制御の割り当ての中心的役割を担っている。前帯状皮質は学習の初期や問題解決のような、実行に特別な努力を必要とする課題に特に関係していると考えられている。エラー検出 (error detection)、課題の予測、**動機付け、情動**反応の調節といった機能を前帯状皮質によるものとする多くの研究がある」
 - 「前帯状皮質 - Wikipedia」
<http://ja.wikipedia.org/wiki/%E5%89%8D%E5%B8%AF%E7%8A%B6%E7%9A%AE%E8%B3%AA>

線条体と中脳のための スパイラル構造 [Haber 2003]



Haber SN. , The primate basal ganglia: parallel and integrative networks. J Chem Neuroanat. 2003 Dec;26(4):317-30.

ACC/OFC

DLPFC

PM

MI

線条体

TD誤差

行動価値

報酬信号



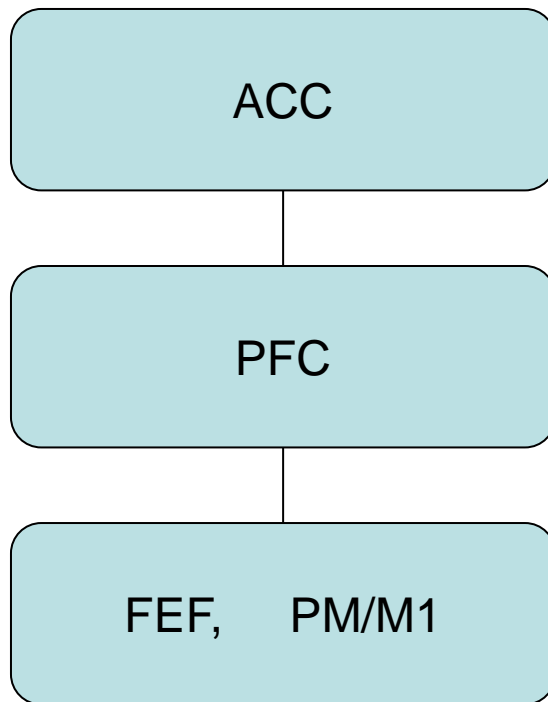
VTA (腹側被蓋野) SNc (黒質緻密部)

中脳

複数の皮質－基底核ループが
階層的な関係にあることを示唆

並行ループの階層

機能



とてつもなく複雑で
わかりにくい。
情動・動機などに関与。

複雑でわかりにくい。
行動計画などに関与。

比較的わかりやすい。
運動に関与。

すべての領野は
似た解剖学的構造

Deep Learning で上の層ほど獲得される表現が
わかりにくくなるのと似てませんか・・・？

大脳皮質ベイジアンネットモデルと はどう関係するのか？

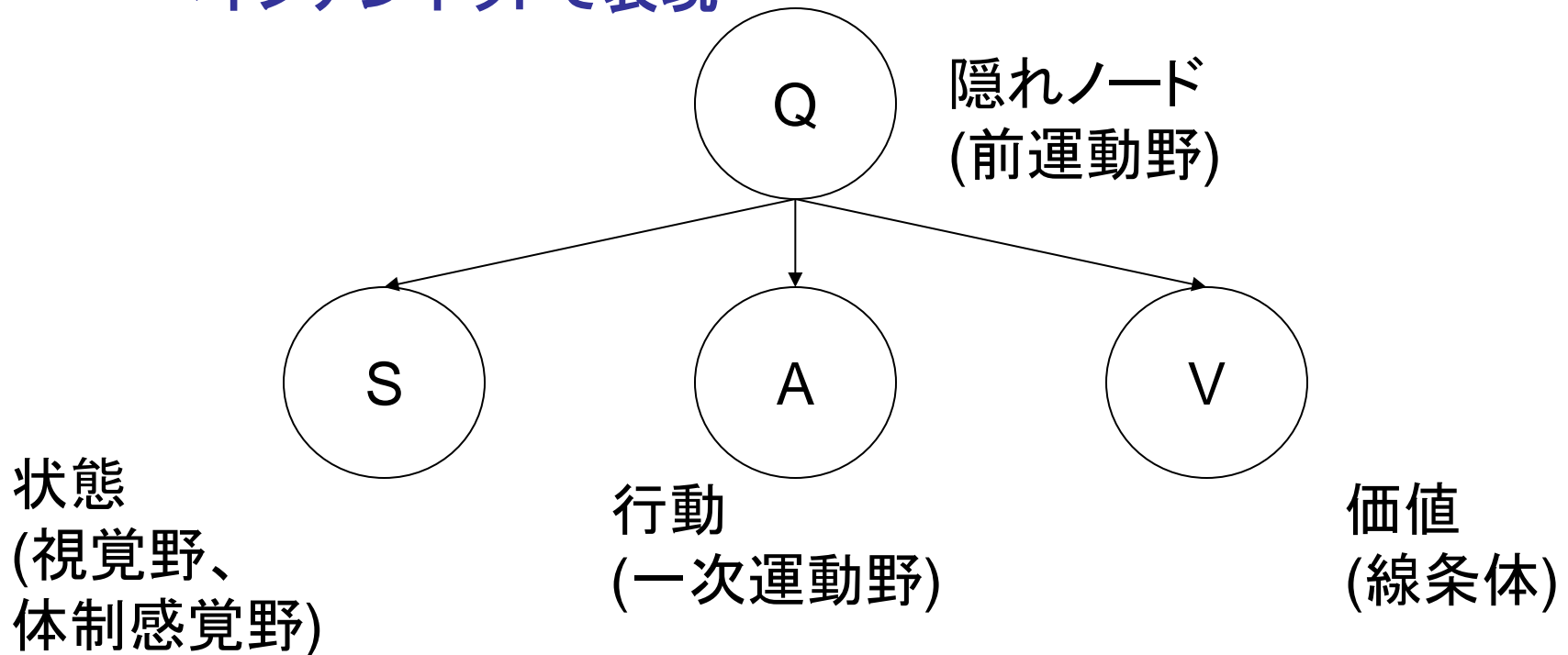
- 視覚野がある種のベイジアンネットであることを多くの研究が示唆。

一杉裕志, 解説: 大脳皮質とベイジアンネット、
日本ロボット学会誌 Vol.29 No.5, pp.412--415, 2011.
https://staff.aist.go.jp/y-ichisugi/besom/29_412.pdf

- 視覚野も他の領野も解剖学的にはよく似た6層構造。
- 前頭前野周辺の領野は強化学習に関与。
- ベイジアンネットモデルと強化学習は神経科学的知見と矛盾することなく統合可能か？

ベイジアンネットで表現した 運動野のモデル [Ichisugi 2012]

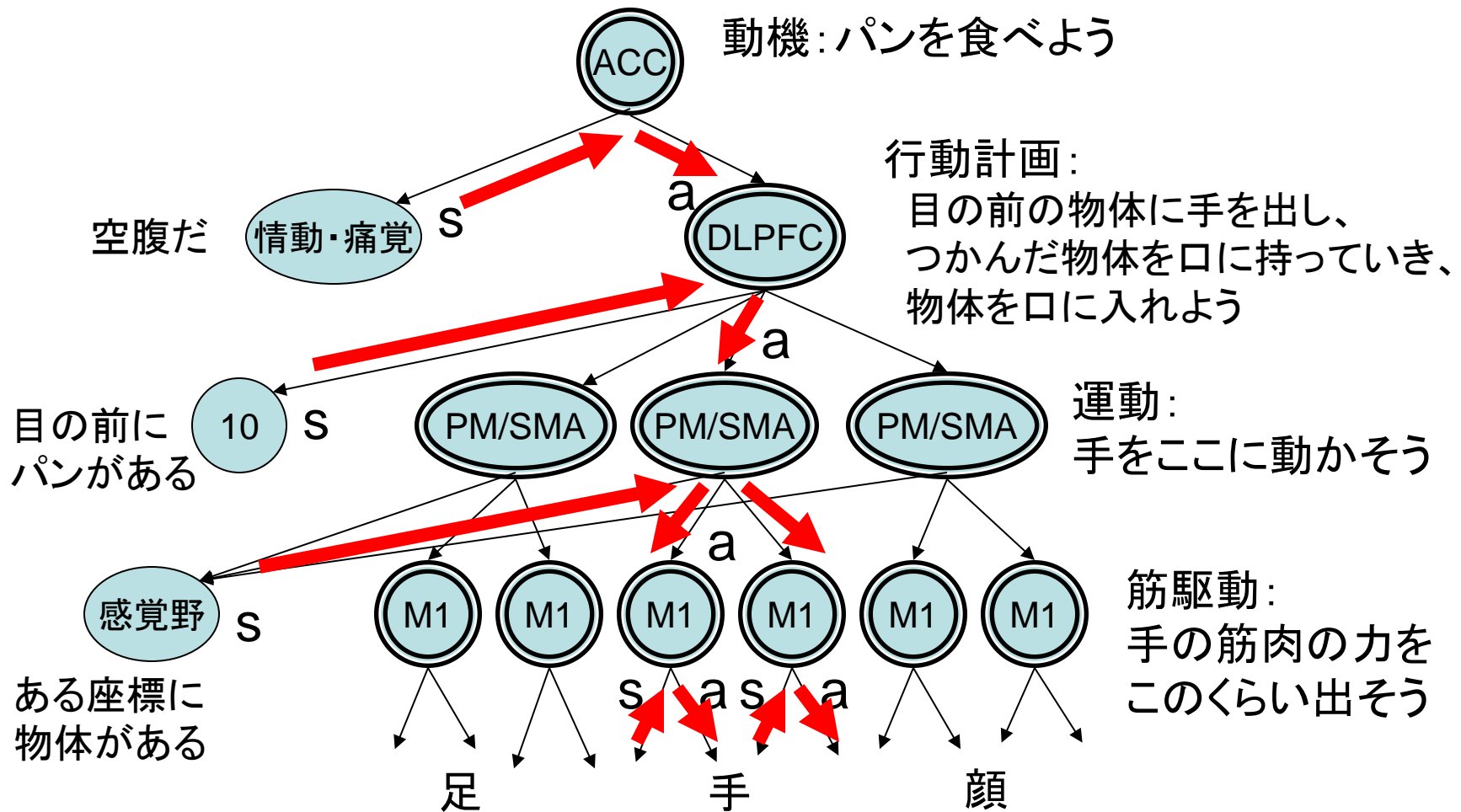
ノードQが競合学習で行動価値関数 $Q(s,a)$ を学習、
ベイジアンネットで表現



$$P(q, s, a, V = 1)$$

$$= P(s | q)P(a | q)P(V = 1 | q)P(q)$$

前頭前野周辺の階層型強化学習のイメージ図



ダイナミックベイジアンネット上での 思考の機構のモデルの提案

下記資料もご参照ください。

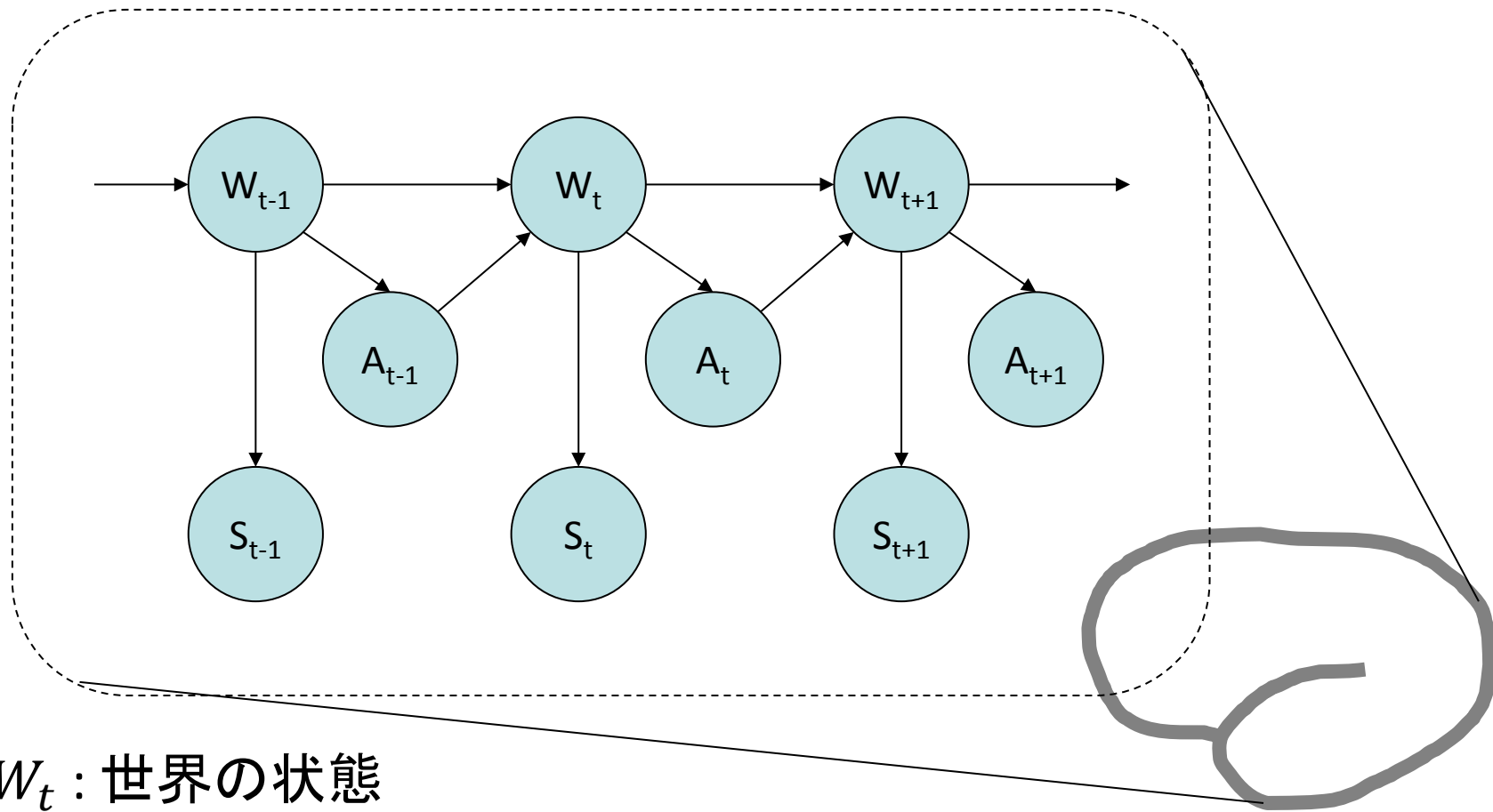
研究構想説明資料:「脳における思考の役割とその実現機構についての考察」:

<https://staff.aist.go.jp/y-ichisugi/besom/20130514serotonin.pdf>

作業仮説1：報酬期待値最大化原理

- 脳の推論・学習・行動の目的は報酬期待値最大化であると仮定。
 - 参考：「全脳アーキテクチャの見取り図」
<https://staff.aist.go.jp/y-ichisugi/brain-archi/wba-sketch.html>

作業仮説2: 脳はダイナミックベイジアンネットワーク



W_t : 世界の状態

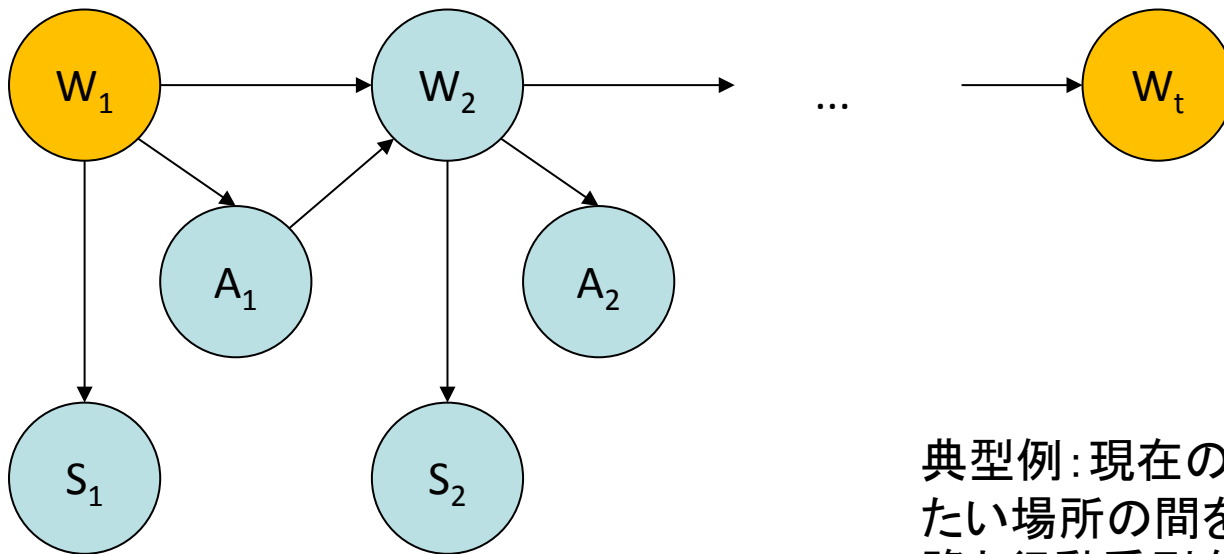
A_t : 運動出力

S_t : 感覚入力

本資料ではモデルの学習は扱わない。
自分自身と外界のモデルが学習済みと考える。

思考の目的の定式化： ゴールベースの場合

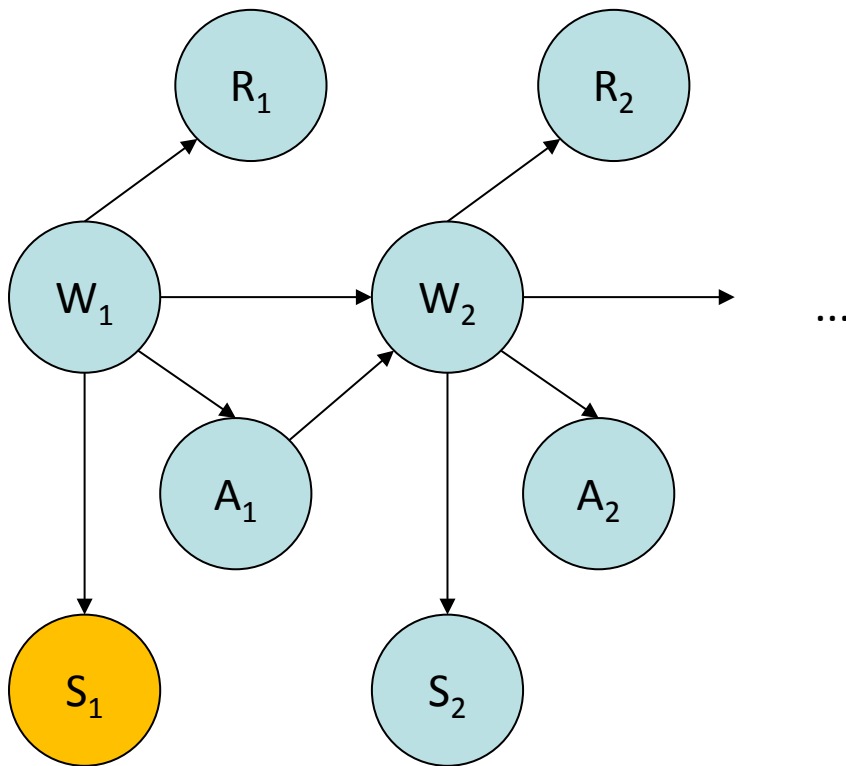
思考の目的は、開始状態と目的状態が与えられた時に、その間をつなぐ最尤経路を求めること。
(ただし目的を達成する時刻 t は小さい方がよい。)



典型例：現在の自分場所と行きたい場所の間をつなぐ移動経路と行動系列を頭の中で探索。

思考の目的の定式化： 効用（報酬）ベースの場合

思考とは、観測値 S_1 が与えられた時、累積報酬期待値を最大化する最適行動列を求めること。



ゴール指向は、報酬指向の特殊な例と考えることができるので、以後、報酬指向をもとに考察を進める。

最適行動列を求める方法

- 時刻 t までの累積報酬期待値の厳密解の計算量は $O(2^t)$ 。
- 近似解法として loopy belief revision (max-product) があるが、問題が複雑になると精度は期待できない。
- 他の近似解法の例
 - グリーディー法
 - モンテカルロ
 - タブーサーチ
 - …
- どれが適切かは状況による。(ノーフリーランチ定理)
- 人間は状況に応じて、不器用ながら適宜近似解法を選択し実行しているように思う。
 - 仮にそうだとすると、どのような機構があればそれが可能か？

幼児の思考？

- 思い付いたやり方で試す。失敗してもまた同じやり方を試す。最後には泣く。
 - グリーディー法？
- 2つのお菓子のうち1つを即座に選択する。あとでそっちも欲しいと泣く。
 - グリーディー法？
- ある程度知恵がついてくると、お菓子を選ぶとき「どちらにしようかな」を納得がいくまで繰り返す。
 - モンテカルロ？

今後の研究の方向性

- 前頭前野は、累積報酬期待値の最大化を近似計算するだけではなく、近似計算アルゴリズム自体を経験によって学習するのではないか？
- それはどのような機構で実現可能になるだろうか？
 - 幸いなことに、解剖学的構造は、その機構は決して複雑ではないことを示唆。

まとめ

- 前頭前野周辺のアーキテクチャを推定するための重要なヒントをいくつか提示した。
 - 前頭前野周辺の4つの並行した大脳皮質－基底核ループは、階層型強化学習を行っているのではないか。
 - 思考は最適意思決定の近似計算ではないか？
 - 近似計算の方法そのものも経験から学習するのではないか？
- **このような研究を始める人がぜひとも必要です！！**