

# 脳型人工知能におけるプランニング

## Planning in Brain-Inspired Artificial Intelligence

荒川直哉<sup>1</sup>

ARAKAWA, Naoya<sup>1</sup>

<sup>1</sup>NPO 法人 全脳アーキテクチャ・イニシアティブ

<sup>1</sup>(NPO) The Whole Brain Architecture Initiative

**Abstract:** This article surveys engineering and neuroscientific models of planning as a cognitive function, to present them as references for realizing the planning function in brain-inspired AI. It also proposes themes for the research and development of brain-inspired AI from the viewpoint of tasks and architecture.

### はじめに

この小論は、認知機能の一つとしてのプランニングの工学的および神経科学的モデルについてサーベイを行う。その目的は、脳型 AI のプランニング機能実現の参考とするために、工学および神経科学における既存のプランニングのモデルを提示することである。また、サーベイの結果見えてきたことから、脳型 AI の研究開発が次に行うことができることを課題とアーキテクチャの観点から提案する。今後は、この小論をもとに具体的な研究依頼 (request for research) を記述していくことが考えられる。

プランニングは、未知の事態に対して新たな学習なしに (その場で) 問題を解決するという点で汎用知能にとって重要な認知機能である。汎用知能をめぐる議論の中では、プランニングは**流動的知能**の典型的な機能とされる。

工学においてプランニングは人工知能の黎明期から研究されてきており、記号的に定式化された課題については解法が存在しているが、定式化されていない実世界での問題解決は今のところ困難である。動物については、ヒトの他に例えばカレドニアカラスが複雑なプランニング課題を解くことが知られているが、他の動物たちも野生で生存していくために何らかのプランニングを行っているはずである。哺乳類の脳については、前頭前野がプランニングに関与していることが知られている。

以下では、まずプランニングについて概説を行い、工学的なモデルを紹介した後、神経科学的モデルについてのサーベイ結果を紹介する。最後に、プランニングを行う脳のモデルを作成するための方策を考える。具体的には、課題、評価のためのベースラ

インシステム、および全脳アーキテクチャの3つについて提案を行う。

### プランニングとは？

プランニングは一般的に、目標と現在の状況の表現が与えられた時に、制約を満たすような行動列 (の表現) を生成する課題である。プランニングでは、実行時に試行錯誤的に問題を解決するのではなく、実行に移す前に「頭の中で」行動列の表現を生成する必要がある。

プランニングでは目標への道筋を見つけるための探索が必要になるが、そこでは道筋を途中目標の列に分解していく**手段目的分析**が行われると考えられる。なお、実世界でのプランニングでは外部の状況が変化するので、それに対応してプランを変更していく必要がある。

### プランニングの要素

プランニングには次のような要素が必要だと考えられる。プランニングをこのように機能分解することは、後の節で神経科学的モデルを考える際にも必要となる。

#### 表象 (internal representation)

プランニングに関わる表象としては以下のようなものが考えられる。

- 現在の状況の表象
- 目標の表象
- 適用できるアクション (方策) の表象
- アクションの結果 (状況) の表象

これらの表象は現在の状況の表象の一部を除き、その場で知覚可能なものではなく、想像、想起または予測される[1] (p.200)。また、アクション (また

は無作為)の結果、変化する状況の表象を得る(想像する)ことは(心的な)シミュレーションである。

アクション(方策)は必ずしも明示的に表象される必要はない。例えば、アクションの結果の表象が目標として選択されると、アクションが分散的にコーディングされている方策により自動実行されるということも考えられる。

## 評価

アクションの結果(状況)の評価は、手段目的分析で得られる部分経路を評価し、どの道筋を優先して探索すべきか、最終的にどのプランを実行すべきかを決めるのに用いられる。

## 探索

プランニングでは目標への道筋を見つけるための探索が必要になる。探索にはさらに次のような認知機能が関与すると考えられる。

- バックトラック  
バックトラックとは、探索において、ある経路がうまくいかないまたはうまくいきそうにならないときに未探索の経路を探索する(あるいは試さなかった行動オプションを試す)ことを指す。バックトラックを行うためには探索済みの経路(あるいは試行済みの行動オプション)を記憶しておく必要がある。
- 探索の戦略(ヒューリスティクス)  
探索は、どの方策を選択するかで分岐するネットワーク(または木構造)の上で行われる。探索の順番として広さ優先探索や深さ優先探索が用いられ、見込みのない経路を捨てる「枝刈り」が行われたりする。
- 目的関数の最大化  
プランニングにおいては、単に目標を達成する経路を見つけることが求められることもあれば、報酬などの値を最大化する経路を採用することが求められることもある。

## 作業記憶

プランニング(経路探索)に必須であろう認知機能としては作業記憶を挙げることができる[2]。例えばバックトラックを行う際には、探索済みの経路を(短期的または中期的に)記憶しておく必要がある。

作業記憶は[3]により用語が導入されて以降、さまざまな定義付けが行われてきた。狭義の作業記憶は、**delayed memory** 課題で調べられるような直近の過去の状況の記憶であるが、ここではより一般的に心的な操作の対象となるような記憶を指すことにする。この点で、作業記憶は、直近の過去の記憶

(**retrospective memory**)とプランニングにより将来実現されるべき物事の表象である**展望記憶**(**prospective memory**)に分けて議論される。例えば[1](p.7)は過去に関する記憶を作業記憶、未来に関する記憶を **preparatory set** と呼んで区別している。

## 注意と抑制

プランニングは時間を要する課題なので、一定の時間課題に集中する(注意を持続する)か、別の物事に注意が移っても課題の途中経過に戻ることができなければならない。別の物事に注意を移さないということはすなわち抑制ということになる。ただし、逆に課題間のスイッチングが行われなければ固執になる(**cognitive flexibility**)。こうしたことから、注意と抑制は課題の執行制御の中心的な機能だといえる。

## プランニング課題

ここでは以下のような課題を典型的なプランニング課題とみなす:

課題解決のために複数のアクションを必要とし、事前に目標、用いることができるアクションおよびアクションの結果についての情報が与えられていて、実行開始後には試行錯誤なしに解決のためのアクション系列を実行するような課題

以下にいくつかの関連するであろう課題を紹介する。

### Tower of London (Tower of Hanoi)

Tower of London/Hanoi は、ヒトのプランニング課題としてよく用いられる [1](p.201)。(空間的な手がかりを除いた)プランニング機能そのものをテストするには適しているとされる [2] が、探索を含まない解法があるという点で本来のプランニング課題ではない。なお、実生活で類似の課題としては片付けタスクを考えることができる。

## 経路探索

出発点から目標に到達する経路またはその中で最短のものを見つける課題である。エージェントが実環境で試行錯誤して解くのであればプランニングといえる。試行錯誤の禁止は、時間的な制約や同一パスの複数回通行禁止 (cf. 一筆書き制約) などにより施行することができる(例えば [4] や **Porteus Maze test** を参照されたい)。さらに特定経路の通行に対して事前に別の行動をとることを前提条件として設定することにより、順序の制約をかけ

ることもできる。

## 道具の使用

カレドニアカラスは道具を使用したプランニング的な行動を行うことが知られている [5]<sup>1</sup>。具体的には木の枝を使ったり、木の枝を取るために別の道具（石など）を使ったりする道具使用行動を順番に組み立て、最終的にエサを獲得する。この際、道具を体の拡張として使ったり、学習した物理法則を用いたりする。こうした行動は興味深いですが、プランニング能力だけで実行できるわけではない。簡単な道具を用いた経路探索としては、棒などで橋を架けていくことで目的地にたどり着く課題を考えることもできる。<sup>2</sup>

## 関連する課題

N-back 課題や Ordered object task (self-ordered task) では、直近の過去に現れた複数の対象を記憶している必要がある。こうした作業記憶のことをモニタリングと呼ぶが、複数の対象の記憶という点でプランニングで必要な展望記憶と共通点がある。

前頭前野研究では、Wisconsin Card Sorting 課題などの課題スイッチング課題が、認知的柔軟性や固執 (perseveration) の性質を調べるものとしてしばしば取り上げられる。課題スイッチングがプランニング機能とどのように関係するのかは一見して明らかではないが、**作業記憶**、**注意**、**抑制**といった下位機能には共通するものがあると考えられる。なお、今得られる報酬より後に得られるより大きな報酬を優先する比較的長時間の行動計画 (プランニング) に関する delayed gratification 課題 (例えばマッシュマロ実験) でも**抑制**が重要視される。

## 工学的モデル

### 記号的モデル

情報処理が (ニューラルネットのように) アナログ・分散的なものではなく、記号の操作によって行われるモデルをここでは記号的モデルという。プランニングの伝統的な記号的モデルとして一般問題解決器 (General Problem Solver・GPS) を挙げることができる。GPS に与えられる問題は、初期状態と目標状態、オペレータ (手段) と状態間の差異とい

う形に定式化される。GPS は初期状態から目標状態への差異を埋めていくようなオペレータの列を探索して求める。あるオペレータを適用すれば目標状態に到達できるような状態を、その目標状態の下位目標 (サブ目標) という。GPS は目標状態から下位目標を求めることを (再帰的に) 繰り返し、初期状態から目標状態へのオペレータ列を発見する (**手段目標分析**)。GPS の具体例としては、1970 年代に開発された STRIPS を挙げることができる。また、局所経路のコストが事前に知られている経路探索 (プランニング) で最適解を求める場合、ダイクストラ法などの動的プログラミング的アルゴリズムが存在する。

記号的モデルには、関連する知識を事前に記号化して与えておかなければならない点、大きな問題空間ではしらみつぶしの検索に時間がかかりすぎる点、不確かな現象 (蓋然性) をうまく扱えない点などの実際上の問題がある。また、こうした点は脳におけるプランニングの実際とも一致しない。なお、ニューラルネットを用いて環境を記号化し、記号的モデルでプランニング課題を解いたハイブリッドシステムの例として [6] がある。

## 学習モデル

上記の記号的モデルでは知識を事前に記号化して与えておかなければならなかった。学習モデルでは、知識を自ら学習する。多くの場合、学習モデルの知識表現はアナログ・分散的なものであり、蓋然的な知識も表現することができる。

## 強化学習

強化学習は、一連の行動を通じて報酬が最も多く得られるような方策を学習する機械学習の手法である。哺乳類の脳では、大脳基底核 (または前頭葉・基底核・視床ループ) が強化学習を行っていると考えられる。さらに、脳は (特に前頭前野において) 次の状態と報酬を予測することができる (モデルを持っている) ので、モデルベース強化学習を行っていると考えられる。

なお、強化学習における「プランニング」は「モデルを入力とし、モデル化された環境との相互作用のためのポリシーを作り出すか改善するような計算過程」を指す ([7] 8 章)。

## ゲート付き RNN

RNN の一種である LSTM (Long Short-Term Memory) や GRU (Gated Recurrent Unit) では、内部状態を保持するか終了するか「ゲーティング」制御が行われる。プランニング・経路探索課題では、

<sup>1</sup> より学術的な論考についてはこちらの諸論文を参照されたい:

<https://unidirectory.auckland.ac.nz/profile/alexander-taylor>

<sup>2</sup> 経路探索を抽象化したバックトラックの例題については <https://www.geeksforgoeks.org/backtracking-algorithms/> を参照されたい。

現在実行中のサブ目標や検討済みのサブ課題を記憶しておく必要があるため、内部状態の保持機能（作業記憶）は必須であると考えられる。ゲート付き RNN がそのままの形で何らかのプランニング課題を学習しうるのであるのか、あるいは何らかのしくみを外付けすることによりプランニング課題を遂行できるようになるかということは、実証されるべきことである。

## メモリーネットワーク：RNN の訓練によるアルゴリズム学習

機械学習の分野では、RNN (LSTM) を適切に（教師あり学習または強化学習により）訓練することでアルゴリズム的な動作を習得することが知られている（一般的にこうした仕組みはメモリーネットワークと呼ばれる cf. [8]）。プランニングもまた（上で見たように）一定のアルゴリズムで動作するので、メモリーネットワークで学習させることが考えられる。実際、Differentiable neural computers (DNC) という手法では、経路プランニングが学習される。

## 注意モデルと MERLIN

近年の自然言語処理では、RNN ではなく「注意モデル」を用いたシステムが好成績を挙げている [9][10]。ここで「注意モデル」と呼ばれるのは、要するに文脈に応じた連想メモリのアクセスである。

ゲート付き RNN と注意モデル的な連想記憶さらに強化学習などを組み合わせて、迷路課題を含むさまざまな課題でよい成績を上げているシステムとして MERLIN [11] を挙げるができる。MERLIN は個別の時点での内部状態を記憶しておくことができるため、過去に 1 度だけ遭遇した事例を知識として利用することができる（これは海馬にヒントを得たエピソード記憶的な機能であるとされる）<sup>3</sup>。また「明示的にプログラムしていないのに MERLIN は階層的な目標志向行動を示した」とされており、プランニング的な機能を持つことが示唆されている。

## RGoal Architecture

RGoal Architecture [13] は、複数の課題空間をランドマークという状態で代表することにより課題の切り替えを行い、ランドマーク間の最適移動経路の探索（プランニング）を行うことで、その場での問題解決を図る「思考モード」を持つ強化学習シス

<sup>3</sup> その後同様なアイデアは例えば [12] などでも追求されている。

テムを提案している。ランドマークは、実際にはサブ目標であることを考えると上記の目標表象に対応する。思考モードでは強化学習アルゴリズムを用いて経路探索を行う。

## プランニングにおける前頭前野の働き

前頭前野がプランニングに関与していることはしばしば指摘されるため、ここでは主にその関係を調査することとした。前頭前野の働きの概要については主に [14][1] および [15] を参照した（日本語で読める文献としては [16] を挙げる）。[14] は、前頭前野の働きは「単一の事象から文脈と需要に対して適切な目標を生成する」ことであるという提案を行っている（8 章）。[1] は、前頭前野の基本的な働きを「新しい形の組織だった目標志向行動の表象と実行」と述べており（p.1）、プランニングを含む実行機能における時間的統合の重要性を指摘している（p.7, p. 201）。[15] は、統合的なプランニングが前頭前野の中心的な機能であるという観点から、外側前頭前野の実行制御機能についてレビューを行っている。

以降、プランニングに必要な認知機能（上記）について関連する脳部位を挙げる。なお、基本的な行為選択と実行決定は、前頭葉・基底核・視床ループが担っているというのが定説であり、少なくともプランニングの結果の実行にはこのループが関わっていると考えられる。

## 表象

「プランニングに必要な要素」の項で述べた表象は、前頭前野のどの部分と関係しているだろうか。

- 現在の状況の表象  
現在の状況の表象は知覚系の情報を処理する脳部位で扱われると考えられる。視覚に関しては、動きや場所の情報を担う背側路（where path）は前頭眼野（FEF）を含む BA8<sup>4</sup> や背外側前頭前野に情報を送り、モノの識別に関わる腹側路（what path）は、腹外側前頭前野（BA45 および BA47）に情報を送る [15]。また、眼窩前頭前野も腹側路からの視覚情報を受け取る。
- 目標の表象  
ここで、目標の表象とは「必ずしも現前していないが実現されるべき状況の知覚イメージ」であるとしておく。[15]（6 章）は、目標の表象の生成が前頭前野（特に外側前頭前野）の機能

<sup>4</sup> 以降 BA は Brodmann Area の略

であるとする。また、[17] は、背側前頭前野および運動前野が目標表象を保持すると主張している。

- 適用できるアクション（方策）の表象  
[15]（3章）は、内側前頭前野が現在の需要との関係で結果との連想に基づく行為の評価と選択に寄与していると述べている。内側前頭前野は海馬との入出力があり、海馬で行われるアクションと結果の系列のシミュレーションの結果を受け取っている可能性がある。

なお、前頭葉の中では運動野に近いほうが直接的な運動に関わり、そこから前方へ行くほど準備的なものとなる [1] (p.320)。

## 評価

強化学習では報酬に関して状況や方策が評価される。[15] では、刺激の評価（眼窩前頭前野）や行為の評価（内側前頭前野）が行われることを想定している。それぞれ状況の評価および方策の評価を行っていると考えてよいだろう。なお、内側前頭前野は部位によって評価の機能が異なり、さまざまな実験結果が報告されている。

## 探索

経路探索（迷路課題など）に関する神経科学的知見は多くない。[4] では、課題の最終的な目標の表象に背外側前頭前野が関与していることを見出している。探索はプランニングそのものともいえるので、この項で述べている他の認知機能が総合的に動員されていると考えられる。

## 作業記憶

作業記憶は、直近の過去の記憶（retrospective memory）と、プランニングにより将来実現されるべき物事の表象である展望記憶（prospective memory）に分けて議論される（上記）。

いずれの場合でも、脳内で（短期の）記憶がどのように実現されているかということが問題になる。作業記憶（短期記憶）を実現する再帰的（recurrent）な神経回路としては、同一皮質領域内の局所的なもの [18] から複数の皮質領域をまたぐ大域的なものまでを考慮することができる [1] (pp.270-283)。なお、前頭前野に関しては記憶の保持そのものを行っているというより、注意の制御（下記）などにより、どの記憶をどれだけ保持すべきかといった制御をおこなっているのではないかという議論がある [19][1] (p.394)。

展望記憶（prospective memory）については、[14] [15][2] などの文献が背外側前頭前野の役割を指摘

している。

## 注意と抑制

[1] によれば、注意には集中と（課題に関係ない情報の）排除（抑制）の側面があり（p.142）、前者は主に背外側前頭前野、後者は主に腹内側前頭前野によって担われるとしている（p.254）。同書はまた、前者を行為の準備を行う preparatory set（展望記憶・上記）と同一視している（p.143）。

さらに同書は、前部帯状回（内側前頭前野）が注意を必要とする課題（例えば Stroop 課題）や行為の準備の際に活動することを指摘している（p.319 [20][21][22]）。また p.321 では、前部帯状回および腹外側前頭前野が抑制信号を視床に送っていることを指摘している。

## アーキテクチャ

上で紹介した（基底核と視床を除く）脳領域とそれらの間の主要な結線を図1に示す。

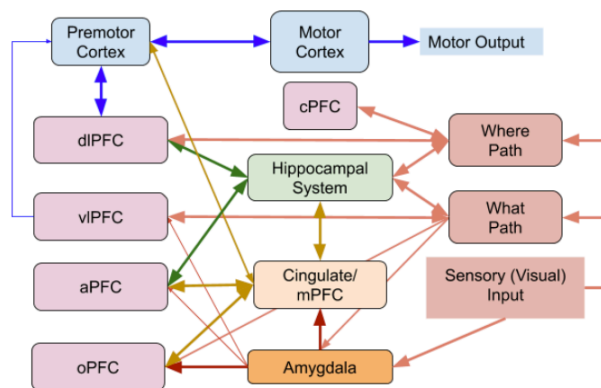


図1

結線情報は特に断らない限り [14] による。前頭葉・基底核・視床ループは図に挙げたすべての皮質に付随しており、強化学習による制御に関与していると考えられる。以下、各部位の簡単な説明を行う。

**Where Path:** 環境にある対象の場所（座標）や動きをコーディングする。

**What Path:** 環境にある対象が何であることをコーディングする。

**cPFC**（後部前頭前野:Caudal PFC): 前頭眼野(FEF)を含み、視覚的な注意に関わる。図には書いていないが、上丘などの眼球運動に関する部位に出力を持つ。

**Premotor Cortex**（運動前野）と **Motor Cortex**（運動野）は、運動の実行に関与する。

**dlPFC**（背外側前頭前野）(BA8, BA9, BA10, BA46): 主に背側路（Where Path）、海馬周辺との結合を持つ。上で述べたように、展望記憶、注意、最終的な

目標の設定などに関与するとされる。

**vIPFC** (腹外側前頭前野) (BA45, BA47, BA44) : 主に腹側路 (What Path) との結合を持つ。知覚的な目標の設定などに関与するとされる。

**oPFC** (眼窩前頭前野) : 扁桃核から情動情報、腹側路 (What Path) から知覚情報を受ける。[14] は、この領野は知覚対象の評価と選択を行うとする。

**aPFC** (前頭前野前部・anterior/polar PFC) : 主に BA10。接続については [23]、機能仮説については [24] を参照されたい。マルチタスキング (複数のタスクの切り替え) などに関与するとされる。

**Cingulate/mPFC** (帯状回・内側前頭前野) : 前頭領域として、前部帯状皮質 ACC BA24、前辺縁皮質 BA32、下辺縁皮質 BA25 を含む。帯状回はさまざまな領域に分かれるが、一般的に報酬 (評価) に関連する。

**Hippocampal System**: 海馬を中心とするいくつかの領域からなる系。Where Path の情報と What Path の情報を統合し、状況 (エピソード) の表象を作成、さらにそれらを長期記憶として固定する (前頭前野との関係については [25] を参照されたい)。

**Amygdala**: 知覚情報と情動系の情報を受け取り、状況の (情動的な) 評価に関わると考えられる [26]。海馬との間にも結線があるが、図 1 では省略した。

なお、言語に関係する vIPFC の BA44/45 は聴覚に関係する上側頭回と接続を持つ。また、dlPFC と aPFC は上側頭回のマルチモーダル領域と接続を持つ。

図 1 から前頭前野領域間の結合がそれほどないことが見て取れる。これらの領域がコミュニケーションをとる場合、共通して結合を持つ別の領域を介することになると考えられる。

図 1 に示した領域を用いたプランニングのモデルがどのようなものになるかについては、下記「提案・全脳アーキテクチャ」の項で改めて議論することにする。

## 計算論的神経科学モデル

この節では、計算論的神経科学モデルとして、前頭葉・基底核・視床ループにより作業記憶をモデル化したモデルと、ハノイの塔課題を解く前頭前野の計算モデルを紹介する。どちらも一般的なプランニング課題を解くものではない。

### PBWM

PBWM (Prefrontal cortex Basal ganglia Working Memory) モデルは、脳における実行機能のモデルで [27]、以下のように要約される。

- 内部状態を保持するか終了するかを制御機能

を持つ。

- 内部状態は前頭前野内に保持される。
- 制御は前頭葉・基底核・視床ループによる強化学習により実行される。

内部状態を保持するか終了するかを制御は、工学的モデルのところでも述べたゲート付き RNN でも行われた「ゲーティング」とも呼ばれる機能である。PBWM がそのままの形で何らかのプランニング課題を学習しうるのか、あるいはプランニング課題の学習には追加的な機構が必要なのかということは、実証されるべきことである。

### Stewart&Eliasmith のハノイの塔対応モデル

[28] は、ハノイの塔課題を解く計算論的神経科学モデルを提案している。モデルの要点は以下のようである。

- モデルは学習を行うものではなく、解法アルゴリズムは設計者によって作り込まれる。
- 作り込まれる解法は、ハノイの塔課題の (探索的・プランニング的ではない) 決定的なアルゴリズムである。
- ハノイの塔の状態としては、注意対象、移動対象、移動先が記号的に与えられる。
- モデル (アーキテクチャ) は、視覚野、運動野、その他の新皮質、基底核、視床をモジュールとして持つ。
- ニューロンモデルとして Leaky Integrate-and-Fire (LIF) spiking neuron を用いている。
- バインディング問題に対処するために Vector Symbolic Architecture という計算法を用いている。

モデルが特定の決定的なアルゴリズムの作り込みであるという点で、汎用知能への発展を考えるのは難しいかもしれない。

## 提案

ここでは、プランニングを行う哺乳類の脳のモデルを作成するための方策を考える。具体的には、課題、評価のためのベースラインシステム、および全脳アーキテクチャの 3 つについて考察を行う。

## 課題

上記「プランニング課題」の項では、Tower of London/Hanoi 課題と経路探索課題を取り上げた。Tower of London/Hanoi 課題はヒトのプランニング機能を調べるために広く用いられているが、決定的 (deterministic) な解法があるという問題がある。

経路探索課題でも、特定の「地形」を何回も学習するとうまくいった解を丸暗記してしまうおそれがある。よって、プランニングの課題としては、学習環境とテスト環境が異なるような経路探索課題が望ましい。このため、課題を生成するソフトウェアを準備する必要がある。

課題を与える際には状況をどこまで記号化して与えるかということに留意する。一般的に状況がより多く記号化されて与えられれば難易度は低くなると考えられるが、ヒトや動物にとっては生態学的な（空間的な）手がかりが減って抽象度が上がるので必ずしも解きやすくなるとは限らない。

## ベースライン

評価のためのベースラインシステムとしては工学的モデルを用いる。課題が記号化されている場合、プランニング専用の記号的モデル（STRIPS や経路探索アルゴリズム）は 100%解いてしまうので、個別に検証する必要はない。上で述べた工学的学習モデルのうち、プランニングを行えるかもしれない LSTM などのゲート付き RNN、DNC などのメモリーネットワーク、MERLIN、RGoal Architecture は、課題が記号化されている場合でも比較対象となる。また、計算論的神経科学モデルとしては PBWM モデルが比較対象となる。

課題が記号化されておらず、画像として与えられる場合は、画像処理のための機構が別途必要になる。記号的モデルとのハイブリッドシステムとして [6] を改変したものを比較対象とすることができる。上記の工学的学習モデルや PBWM についても同様に画像処理のための機構を付加したものを比較対象とすることができる。公正な比較のために、画像処理部分は複数のモデルで共通なものを用いることが望ましい。

## 全脳アーキテクチャ

この節ではプランニングを行う哺乳類の脳のモデルについて、単純なものからより複雑なものを考えてみる。脳のアーキテクチャについては図 1 を参照されたい。

プランニングまたは手段目標分析は、目標と現状の表象からはじめて、下位目標やオペレータ（行為）、行為の結果などの表象列を作り出し、現状の表象と目標をつなぐことで完成する。この際、複数のオプションを「頭の中で」試しつつ、表象列を作り出すことから、広い意味での作業記憶（展望記憶）が必要になる。強化学習的な行動選択は前頭葉・基底核・視床ループで行われるとしても、単純な強化学

習システムは作業記憶などの機能を持たないため、プランニング課題を解くことができない。

作業記憶の計算論的神経科学モデルとしては PBWM モデル（上記）がある。また、メモリーネットワークのうち特に [29] は、計算論的神経科学的側面にも留意しており、LSTM などの RNN でモデル化される前頭前野を基底核を中心とする強化学習装置が訓練して作業記憶などの機能（課題）を遂行できるようなものとするという見立てとなっている。

プランニングにおける作業記憶の役割としては、例えば検討済みの探索経路に印づけしておくことがあるが、これもどこで行うのかを考慮しておく必要がある。これをエピソード的記憶と考えるのなら、海馬周辺との接続のモデル化が必要となる。外世界とのやりとりを伴うバックトラッキングはタスク（サブタスク）の切り替えを必要とする。タスクの切り替えには、前頭前野前部（aPFC）が関与するとされるが、脳内でシミュレートするタスクのバックトラッキングにも前頭前野前部が関与する可能性がある。

プランニングの執行制御（注意と抑制）が強化学習の結果なのであれば（PBWM モデルが想定するように）前頭葉・基底核・視床ループが制御を行ったり、([29] や [30] が想定するように) 強化学習によって訓練された前頭前野が制御を行ったりすることになる。執行制御が作り込みのものだと考えるなら、そのアーキテクチャを脳の構造の中に探すことが有用であろう。

プランニング（手段目標分析）においては行為の結果を予想（シミュレート）することも必要になる。脳モデルにおけるシミュレータの位置としては、前頭前野、海馬あるいは前頭前野と海馬の組み合わせといった部位が考えられる。海馬で行われる過去の状況のリプレイ（プリプレイ）も、当該課題空間内で過去の経験（事例・エピソード）に基づく近似解を示すことができる。

外側前頭前野は注意によって感覚情報に焦点を当てる。特に重要な知覚情報については、前頭前野と後部知覚野の連携による conscious access または「意識のなだれ」により注意が集中する [31]。海馬周辺のエピソード記憶は、背外側前頭前野からの注意によりアクセスされるようである（図 1）。プランニング（手段目標分析）における下位目標の呼び出しは、エピソードまたはエピソードを一般化した状況の表象を長期記憶から呼び出すことで可能になる（上記の工学モデルでは MERLIN の連想記憶や RGoal Architecture における複数の課題空間を代表するランドマークが対応する）。

経路探索課題では、どこに何があるかという情報を処理しなければならない。脳においては where 情報と what 情報は別の経路（それぞれ背側路と腹側路）で処理される。プランニング機能の比較のためには、ベースラインモデルでも where 情報と what 情報を分離した脳の視覚情報処理の仕組みを共通で利用したほうがよいだろう。

ここで、やや踏み込んだ形で、脳におけるプランニングのモデルを提案してみる。アルゴリズムは簡単なものなので、学習も容易だと思われる。脳に関して妥当なものであるかどうかは読者の判断にまかせることとなる。

- 現在の状況から全体的な目標を展望記憶にセットする。
  - 目標が知覚可能な場合：  
腹外側前頭前野に相当する部位が目標となる対象を選定する。
  - 目標が知覚されていない場合：  
現在の状況から海馬周辺の連合野に相当する部位で過去の状況が複数想起され、その中から背外側前頭前野に相当する部位が目標（モノまたはコト）を選定する。
- 現在の状況と全体的な目標から副目標を展望記憶にセットする。  
関連する脳部位については全体的な目標と同様である。  
副目標の想起は次のような尺度によって重み付けがなされる。
  - 現状や目標の属性を条件とする副目標の条件付き確率
  - 現状や目標との属性距離（現状から目標に「近づく」副目標が優先される。）
  - 期待報酬（報酬につながる副目標に重み付けされる）
- シミュレーションによるプランニング  
現状または優先度の高い副目標から別の目標に自らの行為によって到達する方策が想起できるかどうかを確認する。そうした方策が想起された副目標の優先度を方策の評価に基づいて高める。  
これを現在の状況から全体的な目標に到達する副目標と方策の列が見つかるまで繰り返す。方策の想起は、海馬周辺の連合野に相当する部位または前頭葉・基底核・視床ループに相当する部位が担当する。

## 結辞

最後にここで議論したモデル化の工学的な意味合いと神経科学的な意味合いについて述べる。工学

（人工知能）におけるプランニングは、汎用知能の工学的な実現のために重要であるが、課題が記号的に与えられていない場合、現在の技術ではヒトの水準で解決できるようにはなっていないため、脳に学ぶことには意味がある。一方、脳を模倣したアーキテクチャによりプランニング遂行が可能になれば、脳でのプランニングのモデルとして神経科学においても参考になるものと考えられる。

## 謝辞

本研究は新学術領域「脳情報動態」の助成により行われました。また、執筆にあたり、一杉裕志氏、Gideon Kowadlo 氏および山川宏氏から有用なコメントをいただきましたので、ここに謝意を表します。

## 参考文献

- [1] Fuster, J.: *The Prefrontal Cortex*, Fifth Edition, Academic Press (2015)
- [2] Owen, A.: Cognitive planning in humans: neuropsychological, neuroanatomical and neuropharmacological perspectives, *Prog Neurobiol.*, Vol.53, No.4, pp.431-50 (1997)
- [3] Baddeley, A. and Hitch, G.: Working Memory, *Psychology of Learning and Motivation*, Vol. 8, pp.47-89 (1974)
- [4] Saito, N. et al.: Representation of Immediate and Final Behavioral Goals in the Monkey Prefrontal Cortex during an Instructed Delay Period, *Cerebral Cortex*, Vol. 15, No.10 (2005)
- [5] ターナー, P.M.: 『道具を使うカラスの物語』 (2018) (Turner, P.M.: *Crow Smarts: Inside the Brain of the World's Brightest Bird*, HMH Books for Young Readers (2016))
- [6] Asai, M. et al.: Classical Planning in Deep Latent Space: Bridging the Subsymbolic-Symbolic Boundary, *AAAI-2018*, pp.6094-6101 (2018)
- [7] Sutton, S. and Barto, A.: *Reinforcement Learning -- an Introduction*, A Bradford Book (2018)
- [8] Wang, J.X. et al.: Learning to Reinforcement Learn, [arXiv:1611.05763](https://arxiv.org/abs/1611.05763) [cs.LG] (2016)
- [9] Vaswani, A. et al.: Attention is all you need. in *Proc. of NIPS 2017* (2017)
- [10] Devlin, J. et al.: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, [arXiv:1810.04805](https://arxiv.org/abs/1810.04805) [cs.CL] (2018)
- [11] Wayne, G., et al.: Unsupervised Predictive Memory in a Goal-Directed Agent, [arXiv:1803.10760](https://arxiv.org/abs/1803.10760) [cs.LG] (2018)
- [12] Hung et al.: Optimizing agent behavior over long time scales by transporting value, *Nature Communications*



Vol. 10 (2019)

- [13] 一杉裕志ら: RGoal Architecture:再帰的にサブゴールを設定できる階層型強化学習アーキテクチャ, 第9回人工知能学会汎用人工知能研究会 (SIG-AGI) (2018)
- [14] Passingham, R. and Wise, S.: *The Neurobiology of the Prefrontal Cortex: Anatomy, Evolution, And The Origin Of Insight*, Oxford Psychology (2012)
- [15] Tanji, J. and Hoshi, E.: Role of the Lateral Prefrontal Cortex in Executive Behavioral Control, *Physiol Rev.*, Vol.88, No.1, pp.37-57 (2008)
- [16] 虫明元: 『前頭葉のしくみ—からだ・心・社会をつなぐネットワーク』 共立出版 (2019)
- [17] Yamagata, T. et al.: Distinct information representation and processing for goal-directed behavior in the dorsolateral and ventrolateral prefrontal cortex and the dorsal premotor cortex., *J. Neurosci.*, Vol.32, No.37, pp.12934-49 (2012)
- [18] Arnsten, A.: The Neurobiology of Thought: The Groundbreaking Discoveries of Patricia Goldman-Rakic 1937–2003, *Cereb. Cortex.*; Vol.23, No.10, pp.2269–2281 (2013)
- [19] Bledowski C. et al.: What “Works” in Working Memory? Separate Systems for Selection and Updating of Critical Information, *J. Neurosci.*; Vol.29, No.43, pp.13735–13741 (2009)
- [20] Posner, M. and Petersen, S.: The Attention System of the Human Brain, *Annu. Rev. Neurosci.* Vol.3, pp.25-42 (1990)
- [21] Posner, M. et al.: Localization of Cognitive Operations in the Human Brain, *Science*, Vol.24, pp.1679-90 (1988)
- [22] Raichle, M.: Images of the Mind; Studies with Modern Imaging Techniques, *Annu. Rev. Psychol.*, Vol.45, pp.333-56 (1994)
- [23] Petrides, M. et al.: Efferent Association Pathways from the Rostral Prefrontal Cortex in the Macaque Monkey, *Journal of Neuroscience*, Vol.27, No.43, pp.11573–86 (2007)
- [24] Ramnani, N. et al.: Anterior prefrontal cortex: Insights into function from anatomy and neuroimaging, *Nature reviews Neuroscience*, Vol. 5, No.3, pp.184-94 (2004)
- [25] Eichenbaum, H.: Prefrontal–hippocampal interactions in episodic memory, *Nature Review Neurosci.*, Vol.18, No.9, pp.547-558 (2017)
- [26] ルドゥー, J.: 『エモーショナル・ブレイン—情動の脳科学』 (2003) (Ledoux, J.: *The Emotional Brain: The Mysterious Underpinnings of Emotional Life* (1998))
- [27] O'Reilly et al.: Making Working Memory Work: A Computational Model of Learning in the Prefrontal Cortex and Basal Ganglia, *Neural Comput.*, Vol.18, No.2, pp.283-328 (2006)
- [28] Stewart, T. & Eliasmith, C.: Neural Cognitive Modelling: A Biologically Constrained Spiking Neuron Model of the Tower of Hanoi Task, in *Proc. of 33rd Annual Conference of the Cognitive Science Society* (2011)
- [29] Wang, J.X. et al.: Prefrontal cortex as a meta-reinforcement learning system, *Nature Neuroscience*, Vol. 21, pp.860–868 (2018)
- [30] Yamakawa, H.: Attentional Reinforcement Learning in the Brain, *New Generation Computing* (2020)
- [31] ドゥアンヌ, S.: 意識と脳 思考はいかにコード化されるか, 高橋洋 (訳), 紀伊國屋書店 (2015) (Dehaene, S.: *Consciousness and the Brain: Deciphering How the Brain Codes Our Thoughts*, Viking. (2014))