

注意の働きを組み込んだ階層的学習運動制御モデル A Hierarchical Learning Motor Control Model that Incorporates the Function of Attention

高屋 真絵[†], 阪口 豊[‡]
Mae Takaya, Yutaka Sakaguchi

[†]電気通信大学大学院 情報理工学研究科

The University of Electro-Communications Graduate School of Informatics and Engineering
t2232072@edu.cc.uec.ac.jp, yutaka.sakaguchi@uec.ac.jp

概要

本稿では、技能習得課程における注意の働きを検討することを目的として「意識上の働き」を担う系と「意識下の働き」を担う系から構成される階層的な強化学習モデルについて議論する。ここでは、ヒトの運動制御の実態に即した議論を行うため、感覚入力や運動出力、内部表現、情報処理の時間スケール、転移学習などの問題も陽に取り扱い、概念レベルではなくアルゴリズムレベルでのモデル構築をめざす。また、倒立振子問題を題材とした具体的なモデルの実装についても述べる。

キーワード: 注意, 意識, 身体技能, 運動学習, 階層的モデル

1. はじめに

身体運動中の感覚情報は常に脳に伝えられているが、意識上で知覚されるのはそのうちのごく一部であり、運動制御の大半は無意識のうちに実行される[1,2]。このことは、歩行中に脚や腕の動きを意識しないことから経験的に明らかである。一方で、運動中に身体や外環境の対象に注意を向けてその状態を知覚し、その知覚内容に応じて随意的に運動を修正することもできる。このような運動遂行における注意の働きに関連して、自分の身体の一部に注意を向ける「内的焦点」に比べて、外環境の対象に注意を向ける「外的焦点」の方が運動成績や学習成績が向上することが報告されている[3,4]。このような経験や知見をふまえると、ヒトの運動制御プロセスでは、注意や知覚内容に応じて処理が行われる意識上のシステムと、意識に上ることなく自動的に処理が実行される意識下のシステムが並行して機能していることが推察できる。

本稿では、身体技能の遂行・習得過程を念頭におき、運動制御における注意メカニズムについて陽に議論することを目的として、意識上のシステムと意識下のシステムから構成される二階層の強化学習モデルについて検討する。このモデルは、筆者の一人が2018年の本学会大会[5]において提案した「意識上の働き」と「意識下の働き」を組み合わせた運動制御の計算モデルの

構想を具体化したものである。以下では、モデルの全体構成について説明したのち、倒立振子課題を題材としたモデルの実装について述べる。

なお、意思決定や身体運動における意識上・下のシステムの働きを論じたものとして、Kahnemanによるシステム1・2の議論[6]やGallweyによるセルフ1・2の議論[7,8]がよく知られている。本論で議論する二つのシステムの働きはこれらのシステムと概ね対応すると考えられるが、本研究では、概念的な議論にとどまらずにモデルのアルゴリズムを具体的に記述し、数値計算によりその振舞いを例証することをめざしている。

2. モデルの構成

2.1 全体的な構成

前述したように、提案するモデルは意識上の情報処理を担うシステム（上位システム）と意識下の情報処理を担うシステム（下位システム）から構成され、二つのシステムが相互作用しながら動作する（図1）。現実の脳ではいずれのシステムも神経回路網上に実装されていると考えられるが、このモデルでは神経組織との具体的な対応関係は棚上げし仮想的な情報処理モデルとして定式化する。

上位システムは、運動課題を遂行・学習するうえで有用な情報源に注意を向ける方策を探り、下位システムに対して評価関数（報酬）を提供する。一方、下位システムは、上位システムが定めた評価関数を最大化すべく感覚情報に基づいて運動指令を決定する方策を探る。下位システムが過去に学習した感覚運動変換の一部は、上位システムが注意を向ける先を切り替えても引き続き利用できるはずであるから、下位システムの内部では過去の学習内容を新しい学習に利用する「転移学習」が働くものとする。

2.2 上位システムの仕組み

上位システムは、「課題を遂行する上での身体の望ましい状態（知覚内容）を定める」という意識上の意思決

定にあたる機能を担う。身体の状態に関して我々の意識に上る内容は、感覚信号そのものではなくそこから推論された知覚内容であり、また、運動の意図も筋への運動指令そのものではなくて「こうあるべき」という目標状態である [9]。このことを考慮して、上位システムへの入力感覚信号から計算される知覚内容（カテゴリー化された表現）とし、上位システムの出力は個体が望ましいと考える状態の知覚内容、すなわち、「主観的な目標状態」であるとする。主観的な目標状態は、身体や外環境の状態から定まる多様な知覚の中から課題成績を決定づけられると思われる特定の内容を選択したものである。言い換えれば、知覚可能な身体や外環境の様々な変数のうちどれを利用すれば（すなわち、どこに注意を向ければ）課題成績の向上に貢献するかを見出すことが上位システムの目的である。この発見（学習）には試行錯誤が必要であり、上位システムは強化学習によりこれを解決する。この強化学習を進めるための報酬は、運動課題の成績として外部から与えられるものと考えられる。

このように、上位システムは運動課題を遂行するうえで有効な「注意の向け方」を試行錯誤学習により獲得しようとする。この学習により、望ましい「注意の向け方」を発見できれば課題成績は向上し、そうでないと向上しないことになる。このように、本モデルの枠組みを用いることにより、注意の向け方と課題成績との関係性が議論できる。例えば、注意の向け方を（個体が自己探索するだけでなく）外部から誘導することで課題成績を効率よく向上させることが考えられるが、これは身体技能習得における指導者の働きモデル化と捉えられる。

上位システムの処理内容はヒトの認知的情報処理に対応していることに加え、注意の向け方の戦略は頻繁に切り替えるものではないことから、上位システムの学習サイクルはゆっくりとした時間スケール（例えば、技能動作を一定回数繰り返す程度の単位）で進むものと想定している。

2.3 下位システムの仕組み

下位システムは、感覚器官から送られるすべての感覚情報を入力として受け取り、筋に対する運動指令を出力する感覚運動変換機能を担う。そして、上位システムが指定した「主観的な目標状態」に基づき、その状態を実現することを目標として動作する。これは、主観的な目標状態（知覚内容）を感覚信号空間に変換した目標値と現状の感覚信号との近さを評価関数（報酬）にするこ

とによって実現する。これに加えて、下位システムでは、エネルギー消費量や疲労などの生理的・生態学的評価量も評価関数の一部に含める（なお、このような評価量は「疲れ」「つらさ」といった情動の変数として意識上でも知覚されることから、上位システムにおいて取り扱うべき因子であるかもしれない）。

下位システムは、この評価関数（報酬）を最大にするような、感覚情報から運動指令を決定する方策を強化学習により獲得する。この感覚運動変換学習は、上位システムが選択した目標状態の知覚内容（あるいは注意の向け方）とは無関係に、個体が受け取るすべての感覚情報に対して進行する。例えば、ある運動課題を習得するにあたり上位システムが視知覚に基づく目標状態を選択したとしても、下位システムは視覚情報だけでなく体性感覚情報も使って感覚運動変換を学習するため、視覚・体性感覚の両方の感覚信号に対してそれを運動指令に変換する学習が進行する。そのため、下位システムに転移学習の仕組みが備わっていれば、上位システムが定める目標状態が切り替わったときに、注意の対象ではなかった感覚情報に関する過去の学習履歴も有効に利用されることになる。このような仕組みを仮定することにより、「意識していなかったのに身体が勝手に学習していた」という経験を説明できる。

このような機能を実現するため、下位システムでは、身体や外環境の特性を反映した内部モデルと報酬を最大化する感覚運動変換方策の保持機構を分離した構成をとる。運動指令によって生じる身体や外環境の変化の因果関係は、上位システムが定める目標状態（ひいては達成すべき運動課題の内容）とは独立であるため、因果関係に関する内部モデルは目標状態が切り替わっても有効に利用し続けることができる。このように、因果関係を学習・保持する内部モデルと、報酬を最大化する方策を学習する運動指令決定機構を分離することにより、転移学習が実現されることになる。

下位システムの担う処理は実時間での感覚運動情報処理であり、感覚器や効果器（筋）の応答特性や神経系での信号伝達と同じ時間スケールで時間発展することを想定している。しかし、運動指令の「選択」を短い周期で（例えば 10 ミリ秒ごとに）行っているとは考えづらいことから、ここでは、フィードバック制御系のような連続的に動作する感覚運動変換機構を設定してそのゲイン係数を選択する機構、あるいは、モデル予測制御系のように間欠的に運動指令を生成・更新する機構を設けることを想定している。

3. 倒立振り子課題を例題としたモデルの具体化

前章では、提案する学習モデルの全般的な構成や動作原理について概念的に説明した。本研究では、これを具体的なアルゴリズムとして記述し、その振舞いを検証するための例題として、倒立振り子を制御する問題を取り上げる。倒立振り子課題を取り上げた理由は、①強化学習分野での代表的例題としてよく知られていること、②系の構造が単純でその振舞いが比較的把握しやすいことがあげられるが、これに加えて、③従来研究でのこの問題の定式化は人間の運動制御の実態と大きく異なっており、この問題を本研究独自の方法で解決することにより本研究の特色を明確にできると考えたからである。

倒立振り子問題は **pole balancing**，すなわち「掌に棒を立ててバランスする」という身体技能の問題を抽象化・単純化した問題であるといえるが、この抽象化・単純化によりヒトの運動制御がもつさまざまな特性が捨象されている。実際、強化学習の例題としての倒立振り子課題では通常、系は棒と棒の土台となる台車から構成され、台車の位置・速度、棒の傾き・角速度といった物理的状態変数が入力情報として与えられたときに各時刻で台車に右方向に力を加えるか左方向に加えるかを決定する問題として定式化される。これに対し、ヒトの脳が受け取る入力情報は棒の視覚情報や掌にかかる触覚情報などの感覚信号であり、台車の位置・速度などの物理的な状態変数を直接受け取るわけではない。また、この課題を実行するには視線の向き（すなわち、注意を向ける先）が重要な因子であるが、通常の定式化で、知覚や注意の働きといった意識上の働きに相当する要素が取り扱われることはない。

また、**pole balancing** において手の動きはほぼ無意識に決定されている（手を右に動かす・左に動かすなどといちいち意識的に決めているわけではない）ことから、この課題の遂行においては意識下の処理過程が不可欠な役割を担っている。以上のことから、倒立振り子課題は「意識上の働きと意識下の働きの組み合わせ」という本研究の問題意識を議論するのに必要な要素を含んでいるといえる。以上の考察に基づき、ここでは、モデルが受け取る感覚情報、知覚に上る内容、運動出力などをヒトの課題実行時の状況に近い形で定式化したうえで、前章で述べた二階層モデルを適用することにより、その振舞いを調べる（図1）。

具体的には、棒や手は左右方向にしか動かない2次元平面内の問題を設定し、手は右方向に収縮力を発揮する筋と左方向に収縮力を発揮する筋の二つによって制御されていると考える。また、感覚器としては、棒の視覚像をとらえる視覚受容器、手の姿勢や動きの情報を伝える固有受容器、棒から掌にかかる力の大きさや向きを伝える触覚受容器をもつものとする。ただし、数値実験での計算コストを抑えるため、視覚受容器は深層強化学習 (DQN) [10]で使われるような解像度の高い画像ではなく粗い解像度の画像を用いるなどの単純化を行う。以上の設定の下で、ここでの目標は、受け取る感覚情報に基づいて棒をバランスさせるような筋指令を出力することである。外部から与えられる報酬（客観的な課題成績）は、棒を倒さずにバランスさせることができた継続時間（手の位置や棒の傾きが一定の範囲に収まっている時間）を用いる。

下位システムは以上で定めた感覚情報を受け取り、筋への運動指令を生成する。これに対して、上位システムは、感覚情報から計算される知覚内容を受け取り、主観的な目標状態（知覚内容）を出力する。棒の傾きを例にとると、下位システムは棒の物理的状態がもたらす視覚・触覚受容器の出力を入力情報として受け取るが、上位システムでは、「どの程度（大きくまたは小さく）」「左右どちらの方向に」棒が傾いているかといったカテゴライズされた情報（知覚内容）を入力情報として受け取る。同様に、上位システムの出力は「棒が傾いていない」「棒が止まっている」といった目標状態の知覚内容となる。

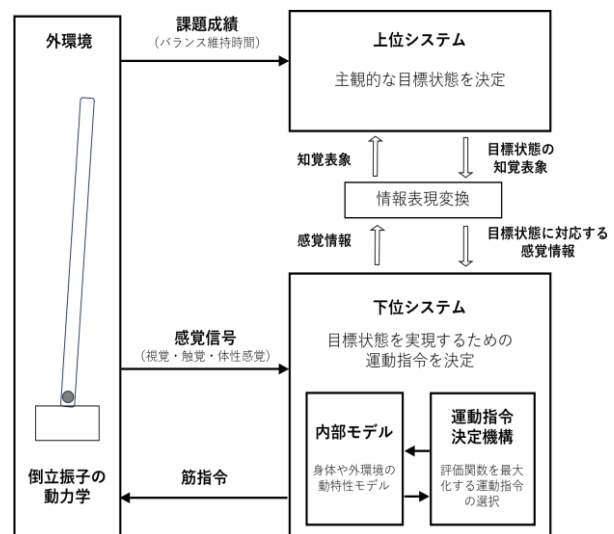


図1 提案する階層的強化学習モデル

このように、上位システムは入力・出力ともに離散的な情報を取り扱うため、ここでは強化学習アルゴリズムとして Q 学習[10]を採用することを想定している。望ましい目標状態（注意を向ける先）は時々刻々と変わるものではなく、一定回数のエピソード（課題を開始してから棒が倒れてしまうまでの一連の事象）を経験したときの成績に基づいて切り替えると考えられることから、目標状態を切り替える（意思決定を行う）時間単位は一定回数のエピソードのまとまりとし、Q 学習のための報酬にはそれらのエピソードにおける平均バランス持続時間を用いる。なお、この例題では課題内容が一定であることから、Q 学習における行動価値関数 $Q(s, a)$ における状態 s は一定値として扱い、行動 a のみの関数として定めることになる（課題内容や知覚内容に応じて注意の向け方を変える必要がある例題では、状態 s はそれらを反映することになる）。

一方、下位システムは感覚情報と上位システムが出力する主観的目標状態とのずれを最小化するように筋指令を決定する。そのために、まず、上位システムが定めた主観的目標状態（知覚内容）を下位システムが取り扱う感覚信号の表現に変換する。この変換は知覚内容と感覚信号の対応関係を保持する変換機構により実現する（ここでは簡単のためこの変換機構はすでに獲得されているものとする）。下位システムは、出力した筋指令とその結果生じる感覚信号との関係性（因果関係）を保持する内部モデルを用いることにより、ある筋指令を生成したときに将来どのような感覚信号が得られるかを予測する。こうして求めた感覚信号の予測値と目標状態に対応する感覚信号との誤差を計算することによって、それぞれの筋指令に対する報酬予測値が求められるので、下位システムは報酬予測値が最大になるような筋指令を選択し、それを出力する。なお、内部モデルが不正確な学習途上ではこの方法では必ずしも最適な筋指令が求められないことから、筋指令を試行錯誤的に生成してさまざまな因果関係を経験して内部モデルの学習を進めることにより、よりよい運動指令を発見することが可能になる。以上のプロセスは、フィードフォワード制御方式の一種であるモデル予測制御の枠組みを参考にしつつ、筋指令を確率的に決定する仕組みを導入することにより実現することを考えている。

現在、本節で説明してきた方式をアルゴリズムとして書き下し、プログラムとして実装する作業を進めているところである。残念ながら本稿では数値実験の結

果を示すことはできない。学会発表の場で実験の一部を報告できるようにしたいと考えている。

4. むすび

本稿では、ヒトの運動制御における注意の働きを議論するため、意識上と働きと意識下の働きを担う二つの階層からなる運動制御システムについて述べた。

本論では、現実の脳組織との対応関係を考えずに、機能の側面から仮想的な情報処理モデルについて論じてきたが、本来は、意識に関わると考えられている脳組織（大脳など）とそうでない脳組織（小脳など）との対比も考慮したモデル化を行うことで、より実質的な脳モデルに近づけることができると考える。意識にかかわる種々の計算理論[1]も参考にしながらさらに検討を進めていきたい。

本研究は、科学研究費補助金・挑戦的研究（萌芽）19K22866 の補助を受けて行われた。この場を借りて謝意を表する。

文献

- [1] ジュリオ・トノーニ, マルチェッロ・マッスィミーニ, 花本知子 (訳) (2015) “意識はいつ生まれるのか 脳の謎に挑む統合情報理論”, 亜紀書房.
- [2] アントニオ・ダマジオ, 千葉敏生 (訳) (2022) “ダマシオ教授の 教養としての「意識」 機械が到達できない最後の人間性”, ダイアモンド社.
- [3] Wulf, G., (2007) “Attention and Motor Skill Learning” Human Kinetics.
- [4] Wulf, G., (2013) “Attentional focus and motor learning: a review of 15 years”, International Review of Sport and Exercise Psychology, Vol. 6, No. 1, pp.77-104.
- [5] 阪口豊, (2018) “「意識の働き」を組み込んだ感覚運動制御の計算モデル”, 2018 年度日本認知科学会第 35 回大会論文集, pp.130-138.
- [6] ダニエル・カーネマン, 村井章子 (訳) (2014) “ファスト&スロー あなたの意志はどのように決まるか?” 上・下巻, 早川書房.
- [7] W.T. ガルウェイ, 後藤新弥 (訳) (2000) “インナーテニス”, 日刊スポーツ新聞社.
- [8] W.T. ガルウェイ, 後藤新弥 (訳) (2000) “心で勝つ! 集中の科学 新インナーゲーム”, 日刊スポーツ新聞社.
- [9] 乾敏郎, 阪口豊, (2020) “脳の大統一理論 自由エネルギー原理とはなにか”, 岩波書店.
- [10] Sutton, R.S. and Barto, A.G. (奥村エルネスト淳ほか監訳) (2022) “強化学習 (第2版)”, 森北出版.