

報告番号	※甲	第	号
------	----	---	---

主論文の要旨

論文題目

A Constructive Approach for Investigating the Emergence of the Primitive Functions of Mind (構成論的手法に基づく原初的な心の機能の発生メカニズムの解明)

氏名

養谷 顕一

論文内容の要旨

イメージを頭の中でつくり主観的に経験をシミュレートする能力は動物と区別できる人間の特徴であると言われている。我々人間の持つこのイメージ化の能力は、将来の自己のシミュレートを可能とするだけでなく、他者の心の状態を推測することをも可能にする。一般に、目標達成に向けた一連の行動を前もって意識的に考える能力はプランニング能力と呼ばれており、外的環境における自己の振る舞いとその結果をシミュレートする試行錯誤によって、現実の世界での生存リスクを減らすことができるという利点を持つ。また、他者の心の状態を推測できる個体は心の理論を持っていると言われ、他者を出し抜き、他者に合わせて行動するなど巧妙で適応的な社会的振る舞いが可能となる。

自己や他者といった問題は長い間、哲学で議論がされてきた。さらに近年の心理学や神経科学の発展によって、これらの問題に対して科学的にアプローチする試みが行われてきた。これらの研究の多くは、現象を分解し、部分に分けて理解する還元論的手法に基づいている。つまり、心という複雑な機能を、様々な個別の要素に分解し、それぞれの働きを実験的に調べることでそれらのメカニズムを推測していく。このような手法が膨大で価値ある事実を明らかにしてきたことは疑問の余地はないが、ふるまいの観察をもとに内部構造を正確に推測するのは困難であるという欠点も存在する。一方、近年では現象の本質を抽出して計算論的モデルを構成し、計算機での現象の再現により、進化的プロセスやメカニズムを理解する構成論的手法が盛んである。構成論的手法は様々な角度から心のモデルを作り、計算機を動かす、仮説演繹を行っていくことが可能であるだけでなく、より知的で柔軟な情報処理が可能な新しいシステム開発へ応用が可能という特徴を持つ。本研究では構成論的手法に基づく計算モデルを用いて原初的な心の機能の発生に関する知見を得ることを目的としている。一つ目の焦点はプランニング能力が進化する環境(条件)を明らかにすることである。二つ目の焦点は、脳を機能レベルでモデル化し、心の理論の脳内メカニズム基盤の獲得に関する手がかりを得ることである。

第二章ではプランニング能力の進化について論じている。ヒトはプランニングに基づく振る舞いの柔軟性を持っている。プランニングのような高次の機能を持つヒトの脳は全体重のわずか 2%程度の重量比であるにもかかわらず消費エネルギーは身体全体のおよそ 20%と、とてもコストがかかる代物である。さらに、プランニング能力と深く関わると考え

られている前頭葉は大腦皮質の約 30%もの領域を占めている。では、いかにしてヒトはプランニングのような高次な機能を獲得したのであろうか。

自然選択は生物の持つ複雑な形質の発生を説明する有力なメカニズムの一つと考えられている。ヒトの心の進化的基盤を検討する進化心理学では、人間の心理メカニズムには文化を超えた普遍性があり、それは自然淘汰の産物であるとする。ヒトの知能の進化に関しては主に生態的なものと社会的なものの二つの異なる観点から説明がされてきた。生態的要因に基づく説明では、我々の知能は環境から資源を引きだし、同時に他種にとっての資源として利用されることを防ぐために進化したのではと考える。一方、社会的要因に基づく説明では、ヒトの知能はむしろ騙しや協調など巧妙で適応的な社会的な振る舞いを可能とするために進化したのではないかと主張する。本研究では以上のような進化心理学における仮説を基に、計算機を用いた進化シミュレーションによる構成論的手法を用いてプランニング能力の進化に関する知見を得ることを狙っている。特に、プランニング能力が進化的に必要となる物理的環境、及び社会的環境を明らかにすることを目的としている。物理的環境としては問題の難易度に、社会的環境としては相互作用の仕方に着目している。

プランニングが必要なタスクとして、適切な順序で行動系列を生成しなければならないタスクが考えられる。そのような状況が存在する最小モデルとして本研究では積木の世界を用いている。各エージェントの遺伝子にはプランニングリミットを設定し、リミットが長ければより深い先読みが可能となるがその一方で認知的コストがかかるというトレードオフが存在する設定としている。プランニングリミットの進化シミュレーションの結果、問題の難易度が上がるほど、より長いリミットを持った個体が進化することが分かった。また 2 体のエージェントが共同で問題を解く場面を想定した場合には、両エージェントのリミットが解に与える影響は相補的で、一方のリミットが十分大きければ精度が良くなる傾向がみられた。このため、集団中の多くをリミットが大きい個体がしめる場合、リミットが極端に小さい個体(フリーライダー)が認知的コストを抑えて相対的に高い適応度を得られ、集団に侵入できる為、独立で問題を解くよりもプランニングが進化しにくい傾向がみられた。さらに、(言語等による)情報共有のあるモデルと無いモデルとを比較した場合、特に問題の難易度が高い時、前者では上記のようなフリーライダー問題が特に深刻であるにも関わらず、情報共有によるプランニングの精度の向上というメリットによって、情報共有のないモデルと同程度にプランニング能力が進化し、結果、より高い適応度を得ることが分かった。

第三章ではヒト認知のより社会的な側面である心の理論の発生について論じている。社会場面における協調的行動決定は人間が社会を形成する上で根源的な過程と考えられる。そこには他者の意図を推測し、それに合わせて自己の役割を動的に決定するという複雑かつダイナミックな行動決定過程が存在する。他者の意図や目的といった心の状態を推測できる個体は一般に、心の理論を持っていると言われており、ヒトの協調にとって欠かせない機能の一つである。それでは、人間の協調行動を成立させている脳内神経機構はいかなるものであろうか。

近年の脳計測技術の進展によって心の理論と関連のある領野については次第に明らかにされつつあるが、個別の領野がいかに連携し社会性に関わっているのかについてはまだよく分かっていない点が多い。一方、より曖昧性のない現象理解とモデル構築が可能な構成論的アプローチが盛んであり脳科学と工学との融合が進められている。他者意図推定に基づく協調行動に関する研究についてもこれまでいくつかの試みが行われてきているが、その多くは心の理論の機能は設計者によりあらかじめ手続き的に定義されていた。本研究で

は、ア prioriに定義せずにその創発に焦点を合わせ、脳を機能レベルでモデル化し心の理論の脳内メカニズム基盤の獲得と合わせて解明することを狙っている。脳の機能モデルは、「脳は多数の機能モジュールの複合体であり、対象との相互作用に合わせてモジュール組み合わせを動的に探索する」(FPC モデル)を採用している。タスクに応じたモジュールの選択的活性化パターンの獲得に焦点を当て、2体のエージェントのコミュニケーションを必要とするタスクでその効果を検証した結果、交流する相手の内部状態(心を読む度合い)に応じた心の理論を処理する脳回路(活性化パターン)が自動的に作られることが分かった。その過程においては他者意図誘導などを初めとする多様な行動決定過程を可能とするネットワークが確認された。さらに初対面の相手でも過去に獲得した活性化パターンを迅速に呼び出して適応できることが示された。

本研究ではヒトの原初的な心の機能としてプランニング能力と心の理論に焦点を当て、構成論的手法に基づく計算機シミュレーションを行った。第二章ではプランニング能力の進化的獲得メカニズムの解明を目的として、論理的操作が必要な積木の世界をタスクとしてプランニングの進化シミュレーションを行った。実験の結果、(1)問題の難易度が上がるほどプランニングが進化すること、(2)(特に問題の難易度が高い時)言語による情報共有はプランニングの進化にとって重要であることが分かった。ヒトと他種の生態的な違いは、言語進化を語る上で重要であるが無視されてきた要因と言われているが、本研究は、「なぜヒトだけが高度な言語能力を発達させてきたのか」という問いに対する、生態的要因(ここでは問題の難易度)を考慮した一つの回答とも解釈可能と考えられる。今後はさらに両形質の共進化ダイナミクスについても検討する必要があるだろう。第三章では他者の心を読む能力「心の理論」の起源に注目し、脳のダイナミクスを含む社会場面の行動モデル構築を行い、複数エージェント間の協調タスクを対象に計算機シミュレーションを行った。結果、インタラクションする相手の内部状態に応じた、心の理論を処理する脳回路が学習を通しダイナミックに獲得されることを示すことができた。今後はさらに対人相互作用機能を組み込んだロボットと人との協調作業タスクなどを検討し、モデルの現象説明能力とより広い場面への適用可能性も評価していく必要があるだろう。提案手法により、従来の条件間の賦活部位の差を比較するだけの検討では分からない脳賦活データの背後にある脳のメカニズムに関する手がかりや、脳の情報処理の工学的応用につながることを期待できると考えている。