

論文の内容の要旨

論文題目 並列監視評価構造を備えたヒューマノイド
の自律行動統合システムの構成法

氏名 吉海智晃

人間に密着して活動することのできるヒューマノイドを実現するためには、合目標的行動を遂行する上で利用する周囲の環境や自身の身体に関するモデルと現実の状態との間に生じた「ずれ」に対して柔軟に対応できるような機構が内部の情報処理システムの中に埋め込まれていることが必要である。たとえハードウェアとしての環境適応能力（感覚器の多様さ、運動生成能力における解空間の広さ）が高くても、このような、行動生成システムとしての適応性の高さが実現されなければ、ハードウェアの可能性を十分に引き出すことはできない。身体・環境モデルとのずれに起因する問題は、如何なるロボットを用いても避けることのできない問題ではあるが、特に、複雑、冗長な身体構成を持つロボットにおいては、その行動生成を考える上で無視できないものとなっている。

本研究では、そのような、目的とする行動の遂行において生じる身体・環境モデルと実身体・実環境との間のずれを如何に補償し、ロボットが行っている行動に反映させるかという問題に取り組む。変化に対する適応性を如何に行動システムに取り込むかといった観点からのシステムアーキテクチャの構成法は様々に考えられるが、本研究では、そういった実際の身体や環境に合わせた適応能力は、目的とするタスクとは関係なく、ロボットが動いている限り常に作用し続けることが重要であると考え、与えられた合目標的行動の文脈からは独立して、自律的に監視・評価を行う構造を用いて行動統合を行うシステムアーキテクチャの構成法を明らかにする。また、実際のヒューマノイドを用いた行動統合実験を通して、その実現可能性を示す。

そのような、人間の与える行動の文脈を実行するシステムのバックグラウンドで並列に自分自身の行動の入出力系を監視し、適時行動修正を行っていく「並列監視評価構造を有する自律行動統合システム」を構築することで、合目標的行動を記述するシステムの肥大化・複雑化を回避できる、既存の行動プログラミング資産を最小限の修正で再利用可能である、合目標的行動の記述部分とは関係なく適応能力のインクリメンタルな拡張が可能であるといった利点がある。

本論文は全7章から構成される。以下に各章の内容に沿って論文の概要を述べる。

第1章「序論」では、研究の背景・目的、そして論文全体の構成に関して述べる。

第2章「並列監視評価構造を備えた自律行動統合システムの構成法」では、提案するシステムアーキテクチャの構成要素について論じ、人間の自律的行動修正機能を整理することによって、ここで実現されるべき行動修正機能の検討を行った。まず、人間の与える行動の文脈からは独立して、自律的に行動修正を行うシステムを構成するためには、常時、感覚入力および動作出力の中から行動修正の必要性を判断することのできる状態量を抽出し、監視する機構が必要である。そして、抽出された状態量に基づいて具体的な行動修正機能を発揮し、時と場合に応じて、行動修正機能の活性を調節できる機構が必要になる。ま

た、予め人間が用意した行動修正機能の範囲では対応できないものに関して、実際の経験に基づいて修正動作を学習していくことのできる機構も欠かせない存在である。本研究では、それら3つの機構をそれぞれ、並列注意機構、反射調節機構、学習反応機構と名づけ、それらが人間の与える行動の文脈を実現する行動システム（本研究では、随意行動システムと名づける）の出力に統合される形のシステムアーキテクチャを提案する。このような自律行動統合システムにおいて実現されるべき行動修正機能としては様々なものが考えられるが、本研究では、主に運動・身体の状態に応じてそれらの調節を見通し良く扱っていくために人間の持つ反射行動を整理し、ヒューマノイドの行動修正システムにとって必要な行動修正機能を分類した。ここでは、そのような行動修正機能として、特に環境や身体の変化への即応性、順応性を保ちながら行っている運動を維持促進していくことを重視し、「運動の促進」、「運動の維持」、「安定した感覚認識の維持」、「突然の環境変化への対応」、「危険からの身体への防御」、「馴染み反応の学習」という6つの行動修正機能に関して議論を行い、実装していくことにした。従来、環境変化に対する適応性付与のアプローチとして、このような即応性を有する反射機能をロボットの行動制御に導入した形の並列駆動アーキテクチャの研究は多く行われてきた。しかしながら、本研究とそれらの多くの研究とが根本的に異なる点として、適応性付加システムを人間が与える行動記述のシステムと分離して構成し、最終的にそれらの出力を統合する方法をとることで、既存のシステムとの統合が容易である点や行っている運動の目的や身体の状態を常時監視、評価してやることで、現状に即した細かな行動修正の調整が可能である点、更に、こういった並列反応アーキテクチャの中に学習の枠組みを導入した点などが挙げられる。

第3章「並列監視評価構造における並列注意機構」では、2章で検討した6つの行動修正機能において、その必要性を判断するために必要な状態量の抽出・監視機構としての「並列注意機構」の構成法について明らかにし、各状態量の抽出・監視機能をシミュレーション及び実際のヒューマノイドを利用した実験において検証する。まず、運動を促進する際に、その運動中にどの部位に最も力を出そうとしているのかを抽出・監視できれば、そのような身体部位の動きを促進可能な反射行動を選択的に活性化させれば良い。ここでは、アクチュエータ変化から計算した四肢の加速度の大きさによって「最も力を出すべき部位」を抽出する並列計算機構を構築した。次に、運動の維持が何によって成されるかと考えた場合、一つの方法として、それぞれの運動中において保持しようとしている姿勢を抽出・監視する手法が考えられる。この時、ある運動中に姿勢間の遷移がある場合には、その運動において保持しようとしている姿勢は必ずしも現在の姿勢と同じとは限らない。そこで、本研究では、SVMを利用して立位・座位・臥位の3つの基本姿勢のいずれに現在の姿勢が相当するのかを判別した上で重心速度の変化を同時に監視することにより、どの基本姿勢を目標として遷移しようとしているかまで抽出・監視する並列計算機構を構築した。また、感覚認識を安定なものとして維持するように行動修正を行うには、そもそも、現在行われている行動において、どの感覚情報が着目されているのかわからなければならない。そのために、本研究では、感覚情報と動作情報の変化の同時性に着目することによって、その因果関係を抽出する機構を構築し、それによって、感覚に基づいた行動がなされていた場合の、その感覚認識能力の維持のための行動修正の必要性を判断する。ここまで述べてきた3つの状態量に関して抽出・監視する機構は基本的には、ロボットが行っている行動の性質に対する注意機構であったといえる。このような行動修正機能により、自身の行っている行動を如何に維持促進するか、ということは状況に合わせて頑健な運動を行うために重要な機能であるが、その一方で、現在行っている行動とは無関係の、不意の刺激が生じた際に、その刺激に対して、適切に反応するように行動修正することも同等の重要性を持つといえる。そのような不意の感覚刺激は環境や状況の変化の可能性を示唆しており、それらを認識し、反応する機構がなければ、状況に応じて柔軟に行動戦略を変化させることもできない。本研究ではそのための状態量として、「予測を裏切る感覚変化」を導入する。具体的な抽出・監視機構としては、「感覚と動きの間の因果関係」を求める注意機構を拡張し、そこで着目されていなかった感覚刺激の変化に対して、どれだけ刺激の変化が予測を裏切ったのかをその変化の頻度の情報量という形で着目すべき値として計算し、監視を行う注意機構を構築した。また、継続して作業を行っていくためには身体を労わりながら行動を行っていくことが大変重要である。そのために、身体のつらさを評価する状態量として「身体負荷の評価」を考える。そのような身体負荷として、本研究では、ロボットのパフォーマンスにもっとも影響を与える、各アクチュエータの負荷を評価することにする。ここでは、ある種類のセンサ出力の変化のみに影響されすぎないように、電流値や温度など負荷を示唆する複数の感覚情報ごとに与えられた負荷評価関数の線形和の平均を取る形で抽出・監視する機構を構築した。

第4章「人間の自律調整機能に学んだ反射調節機構の構築」では、予め人間が埋め込んだ行動修正の最小単位である反射行動モジュールの出力の調節を行う機構を整理し、その構成法を明らかにする。2章で整理したように、6つの行動修正機能のそれぞれごとに、そこで求められる反射行動の性質は異なる。

そこで、それぞれの性質に合わせた出力の調節・統合を行うことにする。まず、行っている運動を促進するための反射群においては、現在の運動で最も強く力を出そうとしている部位の変化を促進するのが良いと考えられるため、3章で論じた「力を出すべき部位の抽出」として求められる状態量を主に参照し、反射行動群を選択的に活性化する形で実装している。現在の運動を維持していくための反射群においては、保とうとしている姿勢に着目して、それを助ける反射群を活性化させれば良いと考えられるため、「運動維持のための姿勢の条件」を参照して、目標とする基本姿勢を維持するのに有効な反射群の出力を活性化し、統合する機構とする。また、認識を補助する反射群の出力の調整には、現在の行動において着目している感覚情報を求めることが必要なので、「感覚と動きの間の因果関係」を参照し、そこで着目している感覚によって認識補助を行う反射群の活性を調節する。一方、突然の環境変化に対応する場合には、そのような変化を認識するために、「予測を裏切る感覚変化」を参照するのが良いが、そのような不意の変化に対して、現在の行動を中断して反応すべきかどうかをその都度判断できなくてはならない。そこで、本研究では、「感覚と動きの間の因果関係」において着目している感覚情報に関しても「予測を裏切る感覚変化」と同様に変化の頻度の情報量を計算することによって、「着目している感覚情報の変化が少なく、尚且つ、生じた新奇刺激が反応すべきほどその価値が大きき場合に、現在の行動を中断してその方向を向く定位反射を活性化する」形の統合を行う機構として実装した。そして、身体を防御するための反射群に関しては、身体を壊さないための危機回避は最優先事項であると考えているので、その出力の調整を行うような機構は設けず、身体が危険な状態と判断されたならば、必ず作用するような形にしている。そのような反射モジュールの内、アクチュエータ状態における負荷を評価して危機回避を行うようなものに関しては「身体負荷の評価」を参照することにする。

第5章「環境や人間への馴染みを学習する行動発達機構」では、行動中に生じる、環境や人間から受ける運動の拘束に対して、その馴染み方を学習するための機構の構成法に関して明らかにし、実際にヒューマノイドを用いた行動実験を通じて、そのような馴染み学習機構が身体負荷の低減に有効であることを示す。自転車のペダルをまわす際や、狭い空間を通り抜ける際のように身体を通じて行うことのできる運動に何らかの拘束がある場合には、馴染む動作は、必ずしも内力/外力が発生する方向と同一ではなく、どの方向に身体を動かせば、より負荷を低減できるのかに関して実際の経験に基づいて学習していける機構が必要になる。ここで求められている学習法は経験に基づいて強化が行われていくものであり、それが常にオンラインで連続的に行われていく必要がある。そのような馴染み反応動作の学習に必要な機能を満たす学習法として、本研究では強化学習手法R学習を利用して四肢及び頭部の各部位ごとに並列学習環境を用意する構成を提案する。また、馴染みを行っていく上で、環境或いは人間から受ける内力/外力の方向によって馴染み方を学習していけるようにするため、3章で論じた「身体負荷の評価」を常に参照し、アクチュエータ負荷から推定した、エンドエフェクタでの内力/外力方向に基づいて強化学習のための状態空間を構成した。また、行動結果の評価のための報酬関数に関しても、「身体負荷の評価」で監視している、その肢を駆動するアクチュエータの総負荷の変化に基づいたものとした。随意行動に対する提案した学習機構を用いた馴染み修正機能を確認するために実ヒューマノイドを用い、サドルに腰を固定した状態で片脚での自転車漕ぎ実験を行い、拘束のある運動を行いながら、より各筋肉に対する負荷の少ない漕ぎ動作を探ることが可能になることを確認した。

第6章「並列注意機構による行動修正機能のシステム統合と行動実験による評価」では、身体構成の異なる複数のヒューマノイドにおける行動統合実験を通じて、3章、4章、5章で個別に論じてきた行動修正機能が、統合された一つの行動修正システムとして機能するかどうかを検討する。この章では、まず、そのためのテストベッドとして用いるヒューマノイドKenta、Kenji、HRP2のハードウェア構成及びソフトウェア構成に関してまとめた。それぞれの身体を駆動するための、センサ、アクチュエータといったデバイス及び、そこにアクセスする部分のソフトウェア構成は異なるものの、本研究で提案する、随意行動システムのバックグラウンドでロボットの入出力系を監視し行動修正を行う自律行動統合システムは、統合する随意行動システム、扱うロボットに関わらず、ほとんど同じ実装によって実現される。この章においては、統合された行動修正システムの上で複数の機能の異なる行動修正を行った例として、複雑な身体を有するヒューマノイドにおける自転車漕ぎ実験及び段差のある経路において等身大ヒューマノイドの移動を誘導する実験をおこなった。

まず、超冗長多自由度な腿駆動ヒューマノイドKenjiを用いた自転車実験では、複雑な身体を有するヒューマノイドにおいては、提案する自律的な行動修正システムが常に作用していることが、行動を維持していく上で必要であることを示すことをその目的とした。この実験においては、5章において行った自転車実験とは異なり、自転車のサドルに軽く腰掛けさせた状態で両脚でペダルを漕がせるようにした。その際に統合した行動修正機能は、バランスを維持するための姿勢反射、視覚系を安定化させるための前庭

動眼反射・視機性眼球運動、右脚の馴染み反応学習である。ただし、馴染み行動修正を右脚のみとしたのは、この実験では、左右の足裏はペダルに固定されており、ペダルを持ち上げる「引き足」の力も利用可能であるため、片側の脚の馴染みのみで十分馴染みつつクランクを回すことが可能であると考えられたためである。これらの行動修正機能は身体構造による拘束によって互いに影響し合い、補完しあうように作用すると考えられる。そして、実際に随意行動である自転車漕ぎを身体の負荷や運動状態に合わせて適応的に修正していく機能が統合されたシステムの下で作用することを確認した。

次に、等身大ヒューマノイド HRP2 を用いて段差のある経路において移動を誘導する実験では、モデルベースな行動生成手法が十分通用するヒューマノイドにおいても、提案する行動修正システムを組み込むことは容易であり、導入した自律的な行動修正システムによって、比較的単純な目的志向の行動の、変化に対する硬直性を補うように全体の行動が出力され、有効に作用していることを示すことを目的とした。ここで行った実験では、途中で段差のある経路に対して、段差に対する考慮なしに与えられる歩行パターンを随意行動とし、そこに、バランスを維持する姿勢反射、段差に足先が衝突した際にそれを乗り越える動作を生成する台のせ反射、各肢のアクチュエータ負荷が高まった時に音声で通達する身体負荷通達のための反射、の3つの行動修正機能を統合している。そして、シミュレーション上および実機において、提案する行動修正システムを用いることで、何も考慮していない単純な歩行軌道を与えるだけで、そのバックグラウンドで作用する行動修正の枠組みによって、適応的に段差を乗り越えるような行動になるように修正することを確認し、HRP2 においても提案する行動修正システムは十分有効に機能することが確認できた。

第7章「結論」では本研究を総括し、今後の課題を述べるとともに、並列監視評価構造を備えた自律統合行動システムが基盤となり動かされていくヒューマノイドシステムの展望について考察した。

以上が各章の概要である。最後に、本研究の主張点とその意義、そして、それらが本研究を通じて如何に達成されたかに関してまとめる。

まず、本研究では、ヒューマノイドのシステムアーキテクチャとして、人間が与える行動の文脈である随意的行動システムと、その入出力系を並列に監視・評価することで行動修正を行う自律行動統合システムに大きく分離して捉える見方を示し、その構成法を明らかにした。このように、自律的に行動修正を行うシステムを主たる行動生成システムから切り離して構築することで、既存の行動システムの上で記述された行動プログラム資産を損なうことなく、運動状態や身体状態に応じた行動修正を付加した形で再利用が可能である、行動修正機能のインクリメンタルな拡張が可能になるといった利点がある。本研究では、3,4,5章における注意機能とそれに基づく個々の修正機能の構築、そして、それらを統合した結果として、6章の身体構成の異なる複数のヒューマノイドを用いた行動統合実験を通して、このようなシステムアーキテクチャの構成をとることで、既存の行動システムの上で記述された目的志向の単純な随意行動を、身体負荷の低減、随意行動の維持促進といった観点において適応的に行動修正するのに有効であることを確認し、その実現可能性を示した。

また、本研究では、並列監視・評価構造として注意機構、反射調節機構、反応学習機構を実装することによって、行動修正機能の並列、同時実行を実現する方法を明らかにした。このような構成をとることによって、並列監視構造としての注意機構により行動修正に必要な状態量を監視、抽出し、それを反射の分類ごと、学習機構ごとに評価して、同時に複数の反射行動或いは行動学習を適応的に調節することが可能になる。また、運動促進のための反射調節や環境変化への定位反射、馴染み動作学習といった時定数の異なる行動修正の同時実行が可能になると考えられる。本研究では、3,4,5章を通じて、これらの機構の具体的な構成法を明らかにし、それらの基本機能をそれぞれシミュレーションおよび実ヒューマノイドを用いた行動実験を通して確認した。そして、6章における2種類の異なる身体構成を有するヒューマノイドの行動実験において、これらの個々の時定数の異なる行動修正機能が統合され、同時に実行されることを確認し、その実現可能性を示した。

更に、本研究においては、環境・人間の干渉への馴染み動作学習機構の構築により身体負荷に基づいて直接馴染ませ方を強化学習する手法を示した。この、馴染み反応動作を学習する枠組みの導入によって、環境や人間からの干渉によって運動が制約される中でも実際の行動経験に基づいて身体負荷を軽減するように行動を修正し、それを再利用することが可能になると考えられる。本研究では、5章において、オンラインで連続的に学習が可能な強化学習の一手法である R 学習を利用した馴染み学習機構の構成法に関して明らかにし、拘束のある運動の中で馴染み方向を探る学習を行う行動統合実験として、5,6章においてヒューマノイド Kenji を用いた自転車漕ぎ実験を行うことで、馴染み学習機能の効果を検証し、その実現可能性を示した。