

ロボットの随伴性行動を生成するためのアーキテクチャの提案

滝本 佑介^{1*} 長谷川 孔明¹ 今井 倫太¹
Yusuke Takimoto¹ Komei Hasegawa¹ Michita Imai¹

¹ 慶應義塾大学

¹ Keio University

Abstract: 人間と自然なインタラクションを行うロボットには、外界からの刺激に対して即座に回答する随伴性が求められる。このようなロボットは人間の動作を模倣することで作成されてきた。しかし従来はモデルが複雑で目的ごとに特化しているため、行動の追加や修正が困難であるという問題点がある。本研究では随伴性を持つロボットの設計難易度を下げることが目的とした、よりシンプルで汎用性の高い行動生成アーキテクチャを提案する。

1 はじめに

ロボットが人間と自然なコミュニケーションを行うために、ロボットは人間が違和感を感じない振る舞いを行う必要がある。そのためには事前に設計された固定の動きをするのではなく、人間の動作などの外界からの刺激に対して即座に反応し、適切な行動をすることが求められる。このような行動は随伴性行動と呼ばれ、ロボットが随伴性行動を行うことで人間がロボットに関わろうとすることが知られている [1][2]。本研究では随伴性行動の生成に取り組む (図 1)。

違和感を感じない随伴性行動として、人間の動作を模倣する研究が行われている。Schulz らは目と頭に着目し、人間の生理的な動作を模倣する Humotion フレームワークを提案した [3]。また随伴性行動を生成するモデルとして、Sumioka らは振る舞いの中から適切なものを選択するモデルを提案した [4]。

しかし Humotion フレームワークでは、行動生成モデルが生起する振る舞いを内包する形になっているため複雑で、振る舞いの変更が困難であるという問題点がある。一方で Sumioka らのシステムでは、複雑な随伴性行動を表現するだけの汎用性がない。

そこで本研究では、随伴性を持つロボットの設計難易度を下げることが目的として従来の行動生成モデルを整理し、独立に設計された振る舞いを統合することで随伴性行動を生成するアーキテクチャを提案する。また、提案する行動生成アーキテクチャを使用した場合における振る舞いの設計方針も合わせて提案する。



図 1: 随伴性行動を行うロボットの例

2 関連研究

随伴性をもつロボットの研究は複数行われている。Fischer らは、人間と即座にアイコンタクトを行ったり、人間が指し示した物体を注視したりするロボットを開発し随伴性行動の有効性を示した [2]。

Schulz らは前庭動眼反射や滑動性眼球運動、衝動性眼球運動、両眼離反運動といった人間の目と頭における動作や、まばたきの種類の整理を行い、これらを内包するフレームワークを提案した [3]。

また Sumioka らは、事前に設計された振る舞いや、ロボットが行った行動に対するユーザの反応から学習した随伴性行動を用い、その中から適切な振る舞いを選択することで行動を生成するモデルを提案した [4]。

2.1 行動の生起と衝突

随伴性行動は主に外界からの刺激によって生起する。例として {注視, 頷き, 呼吸} の振る舞いをロボットが行うことを考える。それぞれの振る舞いは体の部位ご

*連絡先: 慶應義塾大学理工学部情報工学科
〒 223-0061 神奈川県横浜市港北区日吉 3-14-1
E-mail: takimoto@ailab.ics.keio.ac.jp

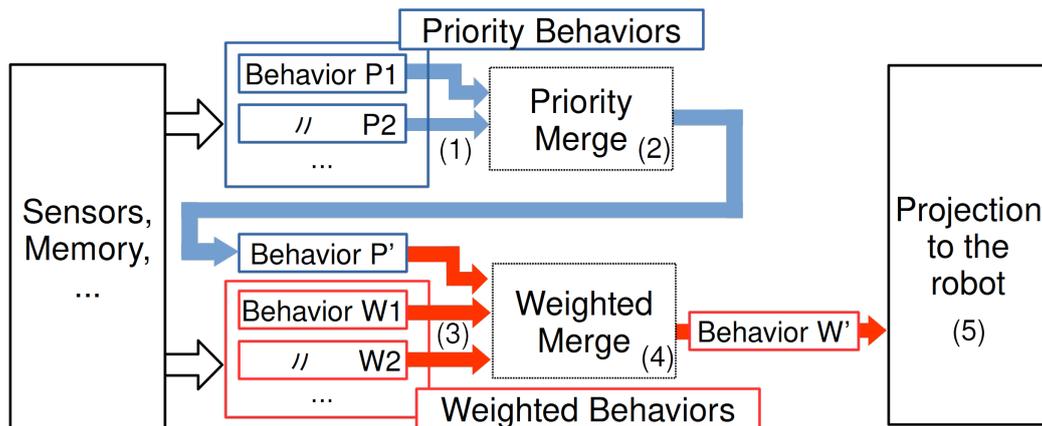


図 2: 随伴性行動生成アーキテクチャ

との動かし方の情報を持つ。この場合、人間がロボットを見た際にアイコンタクトとして人間を注視したり [2], 人間の発話が終わったタイミングで頷いたり [5] といった設計が存在する。

このように振る舞いはセンサや過去の状態を記録したメモリなどからの情報を元に、任意のタイミングで生起する。しかしロボットの同じ部位を動かす振る舞いが同時に生起した場合、動かし方が競合し適切な動作を行うことができない。この競合を解消し最終的に一つの出力に統合する方法として、選択と混合という 2 種類の統合方法が既存研究に見られる [3][4]。

2.2 選択による行動統合

ここで言う選択とは、複数個の競合する動かし方の中から 1 つのみを選択することである。これには同時に生起した振る舞いによって生じたロボットの動かし方から 1 つを選ぶだけでなく、既に出力途中の動かし方を中断し別の動きを開始することも含まれる。

例として、注視という振る舞いによって頭と目が動作している最中に、頷きという別の振る舞いが生起した場合を考える。この時、目は注視を継続する一方、頭の動かし方については注視よりも頷きを優先して選択し、頷きながら注視を行うことが望ましい。

2.3 混合による行動統合

混合とは複数個の競合する動かし方のうち、いずれかを優先するのではなく、動かし方を任意の重みにより平均化することである。

混合を行うべき振る舞いの例として、呼吸に伴う頭や腕などの周期的な振動が挙げられる。この振る舞いは常に生起し続けるため、他の振る舞いを阻害する選択ではなく、平均化を行う混合が望ましい。

2.4 既存研究の問題点

既存研究には選択と混合という 2 つの行動の統合方法がみられる。しかし選択のみを用いているシステム [4] では同一の部位で異なる振る舞いを同時に表現することができないという問題がある。また、振る舞いを独立したモジュールとして扱っていないモデル [3] には、新しい振る舞いの追加や不要な振る舞いの削除のためにモデルそのものを変更する必要があるという問題点がある。

3 随伴性行動生成アーキテクチャ

本研究では独立に設計された振る舞いを 2 種類に分類し、選択と統合に相当する 2 段階で統合を行う随伴性行動の生成アーキテクチャを提案する。これにより柔軟な随伴性行動の設計と、必要に応じた振る舞いの追加、修正、削除が可能になる。図 2 にアーキテクチャの構成図を示す。

提案するアーキテクチャは優先度付き振る舞い (Priority Behaviors) と重み付き振る舞い (Weighted Behaviors) をそれぞれ任意の個数登録することができ、センサやメモリなどから得られた情報を元に任意のタイミングで生起する (1) (3)。競合した動かし方は選択に相当する優先度付け (Priority Merge) (2) と混合に相当する重み付き平均化 (Weighted Merge) (4) の 2 段階で 1 つの振る舞いに統合される。

またロボットにより搭載するアクチュエータの数や機構が異なるため、ロボットごとに固有なデータ形式を用いることは汎用化の面から望ましくない。そこで { 頭, 目, まぶた, 腰, 右手, 左手 } といった部位ごとに抽象化されたデータを扱い、アーキテクチャの最終段階でロボットごとに固有な値に変換・出力を行う (5)。

3.1 振る舞いの構成

アーキテクチャに振る舞いを登録するためには以下の3要素を指定する必要がある。

1. 振る舞いが生起する条件
2. 部位ごとの動かし方
3. 部位ごとの優先度 p または重み w

振る舞いが生起する条件はセンサや過去の状態などの情報を用いて指定する。条件が成立した時刻から振る舞いが開始される。

部位ごとの動かし方は単一の動作だけではなく、時系列データであることも想定している。これにより頷きや腕振りといった、動作の途中で動きの方向が変化する振る舞いも容易に設計することが可能である。

優先度 p と重み w は $0 < p, w$ を満たす実数値である。重要な振る舞いほど大きな値を設定することで、他の振る舞いを抑制する。

3.2 優先度付け (Priority Merge)

優先度付けは、その時刻において生起した、もしくは既に動作中の振る舞いのうち、優先度 p が指定されたものが競合した場合に行う。統合は部位ごとに行い、それぞれ優先度 p が最も高い動かし方を選択し統合された振る舞いとする。

また2つ目の統合である重み付き平均化のため、統合された振る舞いの優先度 p を重み w とみなす。

3.3 重み付き平均化 (Weighted Merge)

重み付き平均化は優先度付けと同様に、重み w が指定された振る舞いが競合した場合に行う。競合した部位ごとに重み w により平均化を行い、最終的な統合された振る舞いを生成する。

重み付き平均化は優先度付けよりも自由度の高い統合方法であり、 $w \rightarrow 0$ で競合のない場合のみ出力され、 $w \rightarrow \infty$ で他の動かし方を完全に抑制することができる。

3.4 振る舞いの設計方針

提案したアーキテクチャに登録する振る舞いは優先度付き振る舞いと重み付き振る舞いのいずれかを選択する必要がある。この選択の目安として優先度付き振る舞いには随意運動を、重み付き振る舞いには不随意運動・反射運動を割り当てる。ここで言う随意運動とは意思のある振る舞いのことであり、不随意運動・反

射運動はそれ以外の、意思のない主に生理的な振る舞いのことである。

振る舞いの優先度付けによる統合は、動かし方のいずれか一つのみを選択するため、自由度が低い一方、単純で振る舞いの設計が容易であるという特徴がある。そこで意思のある運動である随意運動は、部位ごとに見ると動きが混ざること無いと仮定し、優先度付き振る舞いとすることで設計の難易度を下げることが可能である。

これとは対照的に、不随意運動・反射運動は随意運動ほど行動自体が複雑ではないため、優先度よりも複雑な表現が可能な重み付き振る舞いとすることで柔軟な振る舞いの設計を行うことができる。

このように随意運動と不随意運動・反射運動を分けることは全体的な設計難易度を下げることにつながるが、設計の目安に過ぎず、必ずしも意思の有無で振る舞いを分ける必要はない。例えば、頭の向きが変化した場合にも目は同じ方向を見続ける前庭動眼反射は本来反射運動であるが、注視という随意運動を継続的に行うことで実現されるため、反射運動を別に用意する必要はない。

4 実装

4.1 ロボット

ロボットは頭や目、まぶたなどが可動する robovie-mR2[6] を使用した。robvie-mR2には頭に3自由度、片目にまぶたを含めて3自由度、腰に1自由度、片腕に4自由度の可動軸が存在する。

4.2 センサ

入力機器には汎用性を考えて特別な装置は用いず、今回使用した robovie-mR2には入力機器が付属していないためマイクを内蔵する Web カメラを別途設置した(図3)。また各振る舞いの生起条件の入力として、表1に示すセンサ情報を用意した。人の顔の位置推定には OpenFace[7] を用い、物体検出には HOG 特徴量 [8] を使用した。

表 1: センサ情報の例

センサ情報	入力機器
人の顔の3次元位置	Web カメラ
物体の3次元位置	Web カメラ
人が注視している対象	Web カメラ
発話の開始と終了	マイク

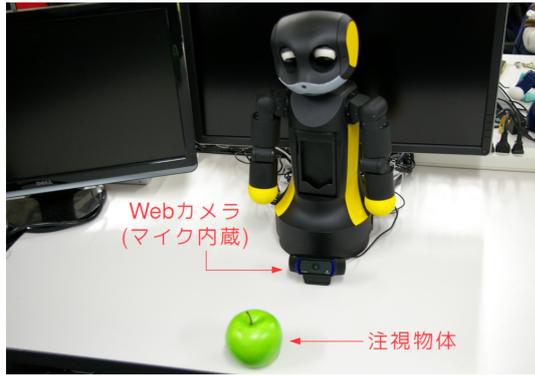


図 3: 実装されたロボットが物体を注視する様子

4.3 振る舞い

振る舞いの例とその種類を表 2 に列挙した。P は優先度付き振る舞い、W は重み付き振る舞いを意味している。注視や指差しには人の顔と物体の 3 次元位置情報を用い、頷きには発話の終了情報を使用する。図 3 に実装されたロボットが物体を注視する様子を示す。物体を注視している間も重み付き平均化によりまばたきが混合して出力されたり、人がロボットを注視すると新たに振る舞いが生起しロボットも人を注視したりする。

表 2: 振る舞いの例

振る舞い	種類
人・物体の注視	P
頷き	P
物体の指差し	P
腕振り	P
呼吸を表す周期的な動き	W
周期的なまばたき	W
眼球移動に伴うまばたき	W

5 まとめ

従来行動生成モデルを整理し、独立に設計された振る舞いを優先度付けと重み付き平均化の 2 段階で統合することで随伴性行動を生成するアーキテクチャを提案した。これにより従来よりも柔軟な振る舞いの追加、修正、削除が可能になり、随伴性行動を行うロボットの容易な設計に繋がる。

今後は、今回提案した生成アーキテクチャに登録する振る舞いの具体的設計内容について取り組む予定である。

参考文献

- [1] Movellan, J. R.: An infomax controller for real time detection of social contingency, *Proceedings. The 4th International Conference on Development and Learning, IEEE*, pp.19–24 (2005)
- [2] Fischer, K., Lohan, K., Saunders, J., Nehaniv, C., Wrede, B., Rohlfing, K.: The impact of the contingency of robot feedback on HRI, *In Collaboration Technologies and Systems (CTS), 2013 International Conference on. IEEE*, pp.210–217 (2013).
- [3] Schulz, S., Lier, F., Kipp, A., Wachsmuth, S.: Humotion A Human Inspired Gaze Control Framework for Anthropomorphic Robot Heads, *HAI '16 The Fourth International Conference on Human Agent Interaction*, pp.207–214 (2016)
- [4] Sumioka, H., Yoshikawa, Y., Asada, M.: Reproducing interaction contingency toward open-ended development of social actions: case study on joint attention. *IEEE transactions on autonomous mental development*, pp.40–50 (2010)
- [5] Watanabe, T.: Human-entrained embodied interaction and communication technology for advanced media society, *RO-MAN 2007-The 16th IEEE International Symposium on Robot and Human Interactive Communication*, pp.31–36 (2007)
- [6] Matsumura, R., Shiomi, M., Nakagawa, K., Shinozawa, K., Miyashita, T.: A Desktop-Sized Communication Robot: “robovie-mR2” *Journal of Robotics and Mechatronics*, Vol.28 No.1, pp.107–108 (2016)
- [7] Baltru, T., Robinson, P., Morency, L. P.: OpenFace: an open source facial behavior analysis toolkit, *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp.1–10 (2016)
- [8] Felzenszwalb, P. F., Girshick, R. B., McAllester, D., Ramanan, D.: Object detection with discriminatively trained part-based models. *IEEE transactions on pattern analysis and machine intelligence*, Vol.32 No.9, pp.1627–1645 (2010)