

ヒューマノイドロボットにおける身体操作を表現するための 二項動詞の獲得

長谷川大 ジェブカ・ラファウ 荒木健治

北海道大学大学院情報科学研究科

{hasegawadai, kabura, araki}@media.eng.hokudai.ac.jp

1 はじめに

近年の自然言語処理では、言語を言語が意味する物理世界や概念から切り離された単なる記号または記号列と見做し、記号を統計的に処理することによって様々なタスクを実現している。意味処理においても、三つ組や各フレーム、シソーラス、類義語辞典など代表的な手法やツールは記号を記号によって定義する循環的定義が行われ、記号が実世界と結びつくことはない。

しかし我々は、深い意味処理を実現するためには、実世界に結合された状態での記号に基づくアプローチが必要であると考えている。そしてそのためには、記号をどのように実世界と結び付けるべきかという記号接地問題 [1] に取り組まなくてはならない。ここで考慮すべきことは身体的重要性である。Brooks [2] は知覚する主体における身体的重要性を主張しているが、これは言語の記号接地においても同様であると言える。

このような理由から、我々は、ロボットが自身の知覚を通して言語を獲得するアルゴリズムを実現することが自然言語処理技術の発展に寄与すると考えている。

ロボットにおける言語獲得は既に多くの研究が行われている [3-8]。特に Roy [9] や岩橋 [10] の画像と音声から自動的に物の名前を学習するモデルなど名詞に関する研究は成功を収めている。しかし、動詞の獲得に関する研究はまだ十分とは言えない。

谷ら [11] はセンテンスとアーム付移動ロボットのモータパターンのペアをリカレントニューラルネットワークにより学習し、シンボル操作可能な動詞の獲得を実現している。杉浦ら [12] はセンテンスとアームロボットの軌道のペアを隠れマルコフモデルを用いて学習し、物体操作を表現するための動詞の獲得を行っている。しかし、「近づける」のような動詞は動作主体と対象物の距離を縮めることを意図しており、軌道やセンサパターンには依存しないため適切な表現モデルとは言い難い。

そこで本稿では、6つの素性による意味表現モデルを使用し、ヒューマノイドロボットにおける身体操作を表現するための二項動詞の獲得手法を提案する。

2 提案手法

本手法では、文を形態素に分割する能力と「右手」「左手」「頭」の3つの名詞は既に獲得されていることを前提としている。さらに、ロボットの胸を

原点とし、頭上方向を z 軸、正面を x 軸とする座標系を基準座標系とする。

2.1 システム概要

まずシステムの構成を図1に示す。ヒューマノイドロボットには KHR2-HV¹ を使用している。

学習フェーズではロボットはユーザから「右手を頭に置く」などの二項動詞を含む文を入力される。つぎにユーザから適切な動作を直接教示により入力される。教示された動作はモータ角度のパターンとして知覚される。モータパターンは動作認知モジュールによって適切な表現系に変換され、文と動作表現のペアとして用例データベースに格納される。ロボットが実際に経験した実例である用例は汎化モジュールで動詞毎に汎用性のある文脈に依存しない表現に抽象化され、ルールとしてルールデータベースに格納される。

テストフェーズではロボットはユーザから未知の文を入力される。未知の文が、ルールデータベース内に存在する動詞を含んでいる場合には、ルールを用いて行動生成を行う。

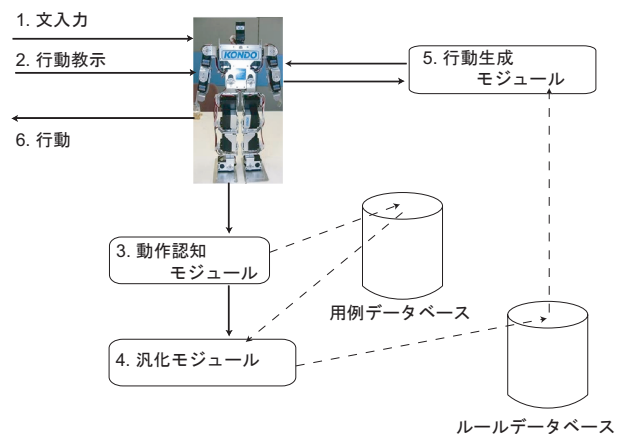


図 1: システム構成

2.2 二項動詞のための表現モデル

本手法では身体操作を表現するための動詞の表現モデルとして、以下の6つの素性を提案する。

1. 動作主体

¹Kondo Kagaku Co. Ltd,
<http://www.kondo-robot.com/>

2. 動作客体
3. 動作主点と対象点との初期距離と最終距離の变化量
4. 動作主点と対象点との最終距離
5. 動作主点と対象点の上下関係
6. 対象点を原点とし、z 軸をロボット頭上方向に固定し、動作主点の初期位置方向に x 軸を向けた座標系における動作主点の軌道

「右手を頭に置く」を表現した場合の例を図 2 に示す。

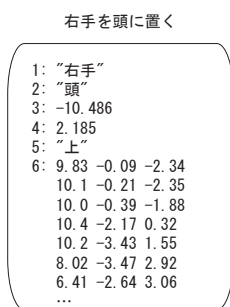


図 2: 表現モデルの例

2.3 動作認知モジュール

動作認知モジュールには入力文と教示されたモータ角度パターンが入力される。「右手を頭に置く」を例にとると、このモジュールでは、動作主体が「右手」であること、動作の客体が「頭」であることを理解したうえで、その他の素性について値を計算し、それらを入力文とペアにし用例データベースに追加する(図 3)。動作主体と動作客体の認知では、教示されたモータパターンから基準座標系での「右手」と「左手」の軌道を計算する。この軌道から、動作量の大きいものを動作主体とし、他方を動作客体としている。

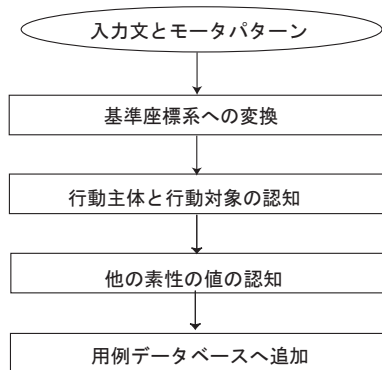


図 3: 動作の認知

2.4 汎化モジュール

汎化モジュールでは、ロボットが経験した二項動詞と動作表現のペアから汎用性の高い動作表現を生成しルールとしてルールデータベースに追加する(図 4)。重要素性の決定処理では、各用例間で素性の値の類似性が高ければ、その動詞を表現する素性として重要であるという仮定をし、類似判定を行って重要素性を決定している。これは、様々な文脈で経験した用例にも関わらず値が類似している素性は、その動詞を表現する文脈に依存しない一般的な素性であるという考えに基づいている。さらに、特徴量の平均化では、重要素性および軌道について、値を平均化している。

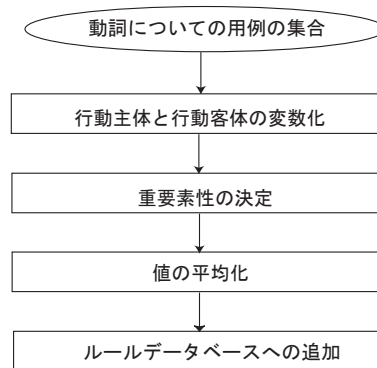


図 4: 用例の汎化

2.5 行動生成モジュール

テストフェーズでは入力文は行動生成モジュールに入力される。行動生成モジュールでは、データベース内にルールが存在する動詞ならば行動を生成する。行動を生成する場合、まず動作主体と動作客体を決定し、ルールから動作の最終位置を決定する。その後、最終位置に合わせて軌道を生成する。そして最後に、軌道をモータ角度パターンに変換し行動を生成する。

3 実験

上記のアルゴリズムを実装し、「置く」「近づける」「離す」の3つの二項動詞の獲得実験を行った。まずトレーニングセンテンスを「右手を頭に置く」「左手を頭に置く」「右手を頭に近づける」「左手を頭に近づける」「左手を右手から離す」「左手を頭から離す」の6文とし、テストセンテンスを「左手を右手に置く」「右手を左手に近づける」「右手を頭から離す」の未知の3文とする。トレーニングセンテンスを用いて各2回ずつ教示を行った場合のテストセンテンスの行動出力結果を表 3 に示す。

4 考察

実験の結果より、本手法により「置く」「近づける」「離す」の3つの動詞が正しく獲得されたことが確認された。さらにその際のルールを図 5, 6, 7

表 1: 実験結果

テストセンテンス	1	2	3
左手を右手に置く	成功	成功	成功
右手を左手に近づける	成功	成功	成功
右手を頭から離す	成功	成功	成功

に示すが、動詞毎に適切に素性の選択がなされていることが確認された。しかし、本手法で用いられたヒューマノイドロボットの構造上の制限により、実験可能な動作が限られているため、表現モデルの評価は十分ではない。どのような動詞が表現可能か、またどのような素性を追加するべきかを考察するために、ロボットを変更して再度実験を行うことが必要である。

@Aを@Bに置く

```

1: "@A"
2: "@B"
3: -11.447
4: 2.175
5: "above"
6: 8.43, -0.01, -6.17
   8.47, -0.03, -6.15
   8.53, -0.09, -6.16
   8.40, -0.18, -6.06
   8.32, -1.37, -3.26
   7.13, -1.80, -0.30
   5.26, -1.25, 1.561
...

```

図 5: 用例の汎化

@Aを@Bに近づける

```

1: "@A"
2: "@B"
3: -12.927
4: unimportant
5: unimportant
6: 10.3, -0.01, -4.64
   10.1, 0.19, -4.37
   9.46, 0.32, -3.84
   8.08, 0.48, -3.04
   6.79, 0.65, -2.40
   5.56, 0.83, -1.97
   5.32, 0.83, -1.94
...

```

図 6: 用例の汎化

5 まとめと今後の予定

本稿では、6つの素性からなる動詞の表現モデルを使用して、身体操作を表現するための二項動詞を獲得する手法を提案した。ヒューマノイドロボット上にシステムを実装し、実験を行った結果、「置く」「近づける」「離す」の3つの動詞を適切に獲得し、本手法の有効性を確認することができた。今後、ロボットの変更、素性の追加を行い、さらに他の動詞について実験を行っていきたいと考えている。

@Aから@Bを離す

```

1: "@A"
2: "@B"
3: 8.704
4: unimportant
5: unimportant
6: 6.47, -0.01, -0.31
   6.58, -0.01, -0.38
   7.13, -0.05, -0.60
   8.22, -0.08, -0.88
   9.17, -0.03, -1.12
   10.4, 0.052, -1.25
   11.5, 0.114, -1.34
...

```

図 7: 用例の汎化

参考文献

- [1] S. Harnad. The symbol grounding problem, *Physica D*, Volume 42, pp.335-346, 1990.
- [2] R. A. Brooks. Intelligence without representation, *Artificial Intelligence*, Volume 47, pp.139-159, 1991.
- [3] D. Roy. Semiotic schemas: A framework for grounding language in action and perception, *Artificial Intelligence*, Volume.167, pp.170-205, 2005.
- [4] L. Steels. The Emergence of Grammar in Communicating Autonomous Robotic Agents, *Proceedings of European Conference of Artificial Intelligence*, In Horn,W.,(Ed), IOS Press, p764-769, 2000.
- [5] L. Steels and Michael Spranger. The robot in the mirror, *Connection Science*, Volume 20, Issue 4, pp.337-358, 2008.
- [6] L. Steels and F. Kaplan. AIBO 's first words, *The social learning of language and meaning, Evolution of Communication*, Volume 4, Issue 1, pp.3-32, 2001.
- [7] J. M. Siskind. Grounding the Lexical Semantics of Verbs in Visual Perception using Force Dynamics and Event Logic, *Artificial Intelligence Research*, Volume 15, pp.31-90, 2001.
- [8] J. M. Siskind. Grounding language in perception, *Artificial Intelligence Review*, Volume 8, Numbers 5-6, pp.371-391, 1994.
- [9] D. Roy. Learning Words from Sights and Sounds: A Computational Model, Ph.D. thesis, MIT, 1999.
- [10] 岩橋直人. ロボットによる言語獲得言語処理の新しいパラダイムを目指して, *人工知能学会誌*, Volume 18, Number 1, pp.49-58, 2003.
- [11] Y. Sugita and J. Tani. Learning samentic combinatoriality from the interaction between linguistic and behavioral processes, *Adaptive Behavior*, Volume 13, Number 1, pp.33-52, 2005.
- [12] K. Sugiura and N. Iwahashi. Motion Recognition and Generation by Combining Reference-Point-Dependent Probabilistic Models, *Proceedings of IEEE/RSJ 2008 International Conference on Intelligent Robots and Systems (IROS 2008)*, pp.852-857, Nice, France, September, 2008.