

副詞の意味に基づくロボットの動作生成に関する一考察

濱園 侑美[†] 小林 一郎[†] 麻生 英樹[‡] 長井 隆行[§] 中村 友昭[§] 持橋 大地[¶]
[†]お茶の水女子大学 [‡]産業技術総合研究所 [§]電気通信大学 [¶]統計数理研究所

1 はじめに

日本は超高齢化社会に突入しており、人手不足をロボットを利用することで解決をはかる場面が増えると考えられる。家庭内でロボットを用いる場合、ロボットが居住者と協調して暮らせる条件として、言葉や身振りを用いることで居住者の経験をロボットに伝え、ロボットはそれを学習することが必要になると考える。

このことを踏まえ、人の言葉による指示からロボットが動作を行なうことを目標に、言葉と動作の対応関係を学習することによって、初めて行なう動作であっても言葉の意味から推測し、動作を行なえるようにすることを目的としている。本研究では先行研究 [1] に対して、多様なロボット動作と曖昧な表現との対応関係をより正しく学習出来る枠組みを検証する。

2 提案手法概要

動作に関する言葉と、動作を変化させる曖昧表現の意味と動作表現の対応関係が既知であるとする。動作と対応関係が分からない未知の言葉が与えられた際に、他の言葉との意味的な関係から対応する動作を推定する手法を提案する。

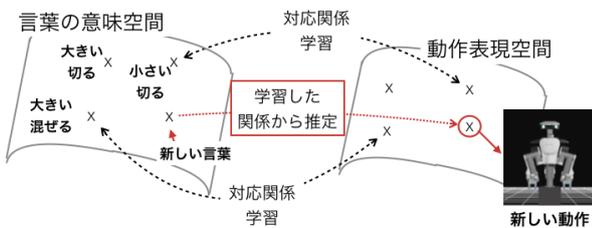


図 1: 提案手法の概要

図 1 に提案手法の概要を示す。言葉は、動作に関する言葉と、動作を変化させる曖昧表現の言葉のそれぞれにラベルをつけ、それらを組み合わせて表現する。動作を表現空間へ配置する方法は、先行研究 [1] で提案した時系列対応 AAM を用いる。また、言葉と動作の対応関係の学習にはニューラルネットワークを用いる。

3 ロボット動作

3.1 ヒューマノイドロボットの概要

(株)川田工業社製ヒューマノイドロボット HIRONXC を用いる。HIRONXC は全 24 の関節を持ち、それぞれ

A Study on Motion Generation of a Robot based on the Meaning of Adverb

Yumi Hamazono[†], Ichiro Kobayashi[†], Hideki Asoh[‡]
 Takayuki Nagai[§], Tomoaki Nakamura[§], Daichi Mochihashi[¶]

の関節角と時間 t を指定することで、 t 秒かけて関節を指定された角度へと動かすことが可能である。

3.2 動作構成

ロボットの調理動作を関節軸の基本動作から構成するために、時系列対応 AAM [1] を用いる。これは Cheng ら [2] による、動作と動作に関連している意味属性を符号化した Activity-Attribute Matrix (以下 AAM) を参考にして作成したものである。表 1 に本研究で使用する時系列対応 AAM の具体例を示す。

表 1: 時系列対応 AAM の概要

	s_1				s_2				s_n				
速く切る	0	0	-5	0.8	0	3.5	5	0.8	0	3.5	5	0.8	
速く混ぜる	0	3.5	0	0.8	3.5	-3.5	0	0.8	-3.5	0	0	0.8	
ゆっくり切る	0	0	-5	0.3	0	3.5	5	0.3	0	3.5	5	0.3	
細かく切る	0	0	-5	0.5	0	1	5	0.5	0	1	5	0.5	
ザクザク混ぜる	0	6	0	0.8	6	-6	0	0.8	-6	0	0	0.8	
	p_{x_1}	p_{y_1}	p_{z_1}	t_1	p_{x_2}	p_{y_2}	p_{z_2}	t_2	...	p_{x_n}	p_{y_n}	p_{z_n}	t_n

	x	y	z
RSY	0.1	1.8	0.7
RSP	-2.3	0.7	0.1
REP	2.1	-0.5	-2.7
RWY	0.0	0.0	0.1
RWP	0.2	0.2	2.7
RWR	0.0	-1.8	0.0

時系列対応 AAM は Activity として調理動作を例に捉え、それに伴い Attribute として右手の指先を前後に動かす (x)、左右に動かす (y)、上下に動かす (z) を基本ベクトルを設定した。また、ロボットを実際に動かすにはそれぞれの関節角を指定する必要があるため、Attribute の x, y, z それぞれに対して、係数となるベクトルとの内積をとることにより、ロボットの動作生成を可能にする。さらに Activity を生成する過程において、それぞれの Attribute の度合い、時系列性、速度等が重要となるため、 x, y, z の変動割合に速度の要素 t を組み合わせた $[p_x, p_y, p_z, t]$ を時系列に n 個並べた $[[p_{x_1}, p_{y_1}, p_{z_1}, t_1], [p_{x_2}, p_{y_2}, p_{z_2}, t_2], \dots, [p_{x_n}, p_{y_n}, p_{z_n}, t_n]]$ を与え、動作生成を可能にする。

4 言葉と動作の関係学習

言葉と動作の関係を学習するために、ニューラルネットワーク (NN) を用いる。ネットワークは図 2 に示した 3 種類を検証した。なお、Net2, Net3 の中間層での活性化関数はソフトプラス関数とする。

Net1: 曖昧表現ラベルと動作表現ラベルからそれぞれの中間層を経た後足し合わせ、ロボット動作を出すネットワーク。中間層での活性化関数をシグモイド関数とソフトプラス関数で検証する。

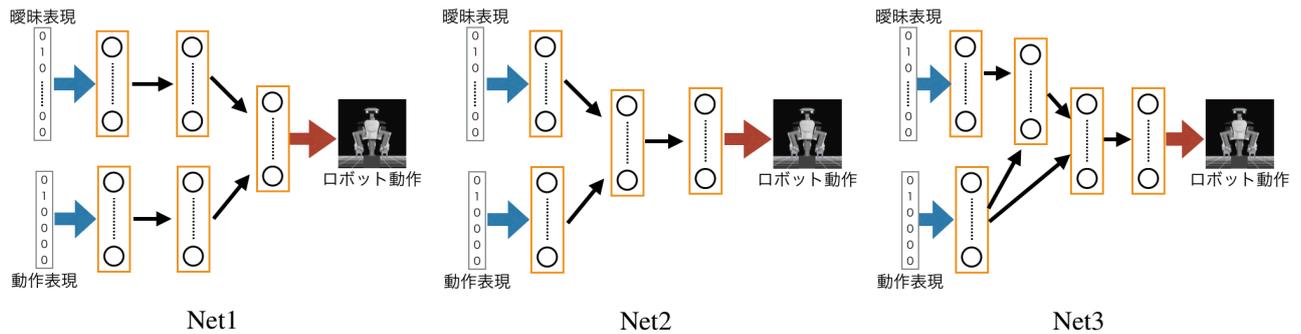


図 2: 検証する Network の概要

Net2: 曖昧表現ラベルと動作表現ラベルのそれぞれの中間層を足し合わせ、中間層を経た後ロボット動作を出すネットワーク。

Net3: 動作表現によって曖昧表現が変化すると仮定する。まず、曖昧表現ラベルと動作表現ラベルのそれぞれの中間層を足し合わせ、出力された、動作により変化した曖昧表現の中間層と動作表現の中間層を足し合わせ、中間層を経た後ロボット動作を出すネットワーク。

5 実験

NNの学習方法は誤差逆伝播法を用い、またロボット動作への出力では活性化関数にシグモイド関数を用いた。NNの出力は、動作の様々なパターンを3.2節に示した時系列対応AAMにて作成し、時刻 s_1 から s_6 の6つの動作を合わせて24次元としたベクトルとする。また、中間層のノード数は2から20を試行し、最も平均二乗誤差が小さくなった9とする。訓練データとして、6個の動作表現ラベルと、12個の曖昧表現ラベルのうちいずれか7個をそれぞれ組み合わせた全42種に対し各100個の全4200個を与えた。学習をした後、評価データは訓練データとして与えた動作表現ラベルと曖昧表現ラベルの組み合わせの42種、未知の言語表現として訓練データに用いた動作表現と曖昧表現のラベルの組み合わせ以外の30種のラベルを入力とした。

表 2: 評価結果

データセット	Network			
	Net1		Net2	Net3
	シグモイド	ソフトプラス		
全体	2.7272	2.7264	2.4751	2.4749
訓練データ	2.6698	2.6689	2.4422	2.4423
未知データ	2.8074	2.8070	2.5212	2.5206

表2は予測した動作と、評価結果として出てきた動作をそれぞれロボットに動作させ、 s_1 から s_6 での x , y , z 軸に関するそれぞれの誤差(単位: cm)と、速さの誤差の平均二乗誤差を「誤差」とする。

6 考察

Net1について、誤差の平均は活性化関数にソフトプラス関数を用いた学習で若干小さくなっており、ソフトプラス関数が有用であると言える。また、Net2, Net3がNet1よりも誤差の平均は小さくなっており、動作表現と曖昧表現の言葉を足し合わせた後、ロボット動作を出力する前に一度中間層を経たネットワーク構成の学習が有用であると言える。Net1とNet2での「ざっと」「炒める」動作を実際にロボットに動作をさせると図3のようになる。

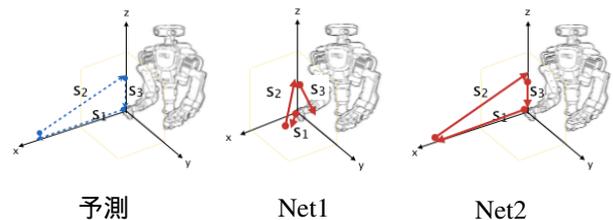


図 3: 「ざっと」「炒める」のロボット動作

また Net3 は全体の誤差と未知データの誤差の平均が最小であり、動作表現によって曖昧表現が変化するという仮定が正しいと言えるであろう。

7 おわりに

本研究ではロボット動作と曖昧表現との対応関係を学習する枠組みとして3種類のネットワークについて検証を行なった。今後の課題としては、分散意味表現を用いることにより、さらに多様な言語表現とロボット動作との関係学習を行なうことである。

参考文献

- [1] 濱園侑美, 小林一郎, 麻生英樹, 持橋 大地, Muhammad Attamimi, 中村友昭, 長井隆行, “語彙の分散意味表現とロボット動作との対応関係の学習”, 第30回人工知能学会全国大会, 2016.
- [2] Heng-Tze Cheng, Feng-Tso Sun, Martin Griss, Paul Davis, Jianguo Li, Di You, “NuActiv: Recognizing Unseen New Activities Using Semantic Attribute-Based Learning”, Mobile Systems, Applications, and Services, 2013.