

# 関連度計算方式を用いたロボットの台詞に対応する動作の生成

小野彰利<sup>†1</sup> 土屋誠司<sup>†2</sup> 渡部広一<sup>†2</sup>

**概要:** 本研究では、ロボットと人間との自然な対話の実現を目指し、ロボットが発話する台詞に動作を付与して実行させる動作生成システムを構築する。提案手法では、台詞とその台詞を発話する際の動作を登録した動作知識ベースを利用する。ロボットが発話する際の動作には、ロボットが発話する台詞と関連の強い動作知識ベース内の台詞に対応する動作を採用する。関連の強さは関連度計算方式または文間関連度計算方式で算出する。

**キーワード:** ロボット, 動作生成

## Generation of Motion for Robot Speech Using Calculation Method of Degree of Association

AKITOSHI ONO<sup>†1</sup> SEIJI TSUCHIYA<sup>†2</sup>  
HIROKAZU WATABE<sup>†2</sup>

**Abstract:** In order to realize natural communication between a robot and a human, this paper proposes a motion generation system that creates motions corresponding to words spoken by robot. This system uses the motion knowledge base which stores words and motions when speaking those words. Motion corresponding to word in the motion knowledge base strongly related to word spoken by robot is adopted as motion when robot speaks. The strength of the relation is calculated by degree of association or degree of association between sentences.

**Keywords:** Robot, Generation of Motion

### 1. はじめに

現在使用されるロボットは産業用のものが主で与えられる仕事をこなすだけである。しかし、今後は産業分野のみでなく医療や介護現場といった人間により身近な環境で多種多様な役割を担うと考えられ、人間にとって親しみやすいことが求められる。そこで、ロボットには人間が持つ常識的判断のための知識を持たせた上で、人間と円滑に対話する機能が必要となる。

本研究では、ロボットと人間との自然な対話の実現を目指し、ロボットの台詞に対応する動作を生成するシステムを作成することで、対話の中で人間がロボットに親しみを持つことを可能にする。人間の発話文より感情を取得して動作を生成するシステム[1]は既に提案されている。このシステムでは人間が発話する台詞より話者感情を取得し、ロボットがその感情に反応して動作を返す。しかし、出力する動作は感情ごとに設定されており、台詞の意図、意味を考慮した動作の使い分けができないという問題がある。例えば、「大会で勝った」という入力ではロボットが「祝福」の感情を抱いて「万歳」の動作をするが、「じゃんけんに勝った」という入力においても「祝福」の感情を抱いて「万

歳」をする。人間の場合は同じ「祝福」の感情であっても、「大会」より連想される感情の程度が「強」で「万歳」の動作、「じゃんけん」より連想される程度が「弱」で「小さくガッツポーズする」というように、会話中の話題で動作が変化するのが自然である。話題とは、会話を展開する上で主題となる事柄である。本研究では話題を表す話題語を用いることで、会話中の話題を考慮したロボットの動作生成を実現する。話題語については4.4項で詳細を述べる。

さらに、人間同士の会話における台詞は感情を含むものばかりではない。ロボットの台詞を自動で生成して人間との自然な会話を展開するシステム[2] (以降、会話システム) においても感情を含むロボットの台詞は限られる。そこで、本研究では会話システムが出力するロボットの台詞に対して、感情の有無に関わらず、対応する動作を自動で生成するシステム (以降、動作生成システム) を作成する。

### 2. 研究概要

本研究では、ロボットの台詞に対応する動作を生成する動作生成システムを構築する。動作生成システムは会話システムと同時に実行し、話題語を会話システム内で取得する。また、会話システムが出力する台詞をロボットの台詞

<sup>†1</sup> 同志社大学大学院理工学研究科  
Graduate School of Engineering, Doshisha University

<sup>†2</sup> 同志社大学理工学部  
Faculty of Science and Engineering, Doshisha University

として採用する．例えば，会話システムに「レストランで料理を食べたよ」と入力（人間の台詞）すると，話題語として「料理」を取得する．そして，会話システムが出力する「いいですね」が動作生成システムの入力文（ロボットの台詞）となり，対応する動作として「両手を胸に当てる」を取得する．

動作生成システムは，動作生成のために台詞とその台詞を発話する際の動作を登録した動作知識ベースを利用する．ロボットの台詞を発話する際の動作は，動作知識ベース内の台詞とロボットの台詞との関連が強い台詞に対応する動作を採用する．例えば，「それは面白いですね」という台詞を発話する際の動作として「うなづく」が登録されている．ロボットの台詞と「それは面白いですね」との関連が強ければ，「うなづく」がロボットの台詞に対応する動作となる．関連の強さを調べる手法としては，4.2節で述べる関連度計算方式と4.6節の文間関連度計算方式を用いる．また，動作知識ベースについては5.1節で詳しく述べる

### 3. 研究環境

本研究で使用するロボットは図1に示すRobovie-R Ver.3[3]（以降，Robovie）である．Robovieは人間の動作を実現するために片腕7自由度（肩2自由度，肘2自由度，手首1自由度，指2自由度）の両腕，首3自由度，車輪2自由度，両目4自由度（片目2自由度）を備えている．本研究では，車輪以外の計21の自由度で動作を生成する．



図1 Robovie-R Ver.3

## 4. 関連技術

本章では使用する技術について述べる．

### 4.1 概念ベース

概念ベース[4]は，電子化された国語辞書等から自動的に構築された知識ベースである．ある概念  $A$  は  $m$  個の属性  $a_i$  と重み  $w_i (>0)$  の対により式(1)のように定義される．

$$\text{概念 } A = \{ (a_1, w_1), (a_2, w_2), \dots, (a_m, w_m) \} \quad (1)$$

また，概念は約9万語格納されており，概念に定義されている意味特徴を一次属性，一次属性の意味特徴を概念に対する二次属性というように任意の次元で表現する．

### 4.2 関連度計算方式

関連度計算方式[5]とは，概念ベースに定義される2つの概念間の関連の強さを定量的に表す手法である．関連度は0.0から1.0の間の実数値で表され，概念間の関連が強いほど大きな値を示す．例えば，概念「自動車」に対して，「車」，「自転車」，「馬」の関連の強さを表1に示すように数値化でき，「自動車」と関連がより強いのは「車」であると判断できる．

表1 関連度計算の例

基準概念	対象概念	関連度
自動車	車	0.92
	自転車	0.34
	馬	0.034

次に，関連度計算方式を用いるために必要な一致度および，重み比率付き関連度計算方式について述べる．

ある概念  $A, B$  において，その属性を  $a_i, b_j$ ，対応する重みを  $u_i, v_j$  とし，それぞれ属性が  $L$  個， $M$  個 ( $L \leq M$ ) とすると，概念  $A, B$  はそれぞれ式(2)式(3)となる．

$$A = \{ (a_1, u_1), (a_2, u_2), \dots, (a_L, u_L) \} \quad (2)$$

$$B = \{ (b_1, v_1), (b_2, v_2), \dots, (b_M, v_M) \} \quad (3)$$

このとき，概念  $A$  と概念  $B$  の属性一致度  $DoM(A, B)$  を以下のように定義する．

$$DoM(A, B) = \sum_{a_i=b_j} \min(u_i, v_j) \quad (4)$$

$$\min(u_i, v_j) = \begin{cases} u_i & (u_i \leq v_j) \\ v_j & (u_i > v_j) \end{cases} \quad (5)$$

ここで， $a_i=b_j$  は属性同士が一致した場合を示している．つまり，一致度とは概念  $A$  と概念  $B$  それぞれの属性を比較して一致した場合に，小さい方の重みを選択して足し合わせていった合計値ということになる．これは，小さい方の重みが互いの属性の重みの共通部分となっているため，概念  $A$  と概念  $B$  のどちらにも有効な重みであると言えることができるからである．

上述の概念  $A, B$  において，まず属性数の少ない方の概念  $A$  を基準とし，その属性の並びを固定する．その上で概念  $B$  の属性を概念  $A$  の各属性との一致度が最大になるように並び替える．並び替えた概念  $B$  の属性と重みを  $(b_{x_i}, v_{x_i})$  として次のように定義する．

$$B = \{ (b_{x_1}, v_{x_1}), (b_{x_2}, v_{x_2}), \dots, (b_{x_M}, v_{x_M}) \} \quad (6)$$

これらの概念についての重み比率付き関連度  $DoA(A, B)$  を次の式で定義する．

$$DoA(A, B) = \sum_i DoM(a_i, b_{X_i}) \times \frac{(u_i + v_{X_i})}{2} \times \frac{\min(u_i, v_{X_i})}{\max(u_i, v_{X_i})} \quad (7)$$

### 4.3 意味理解システム

意味理解システム[6]は、入力文を 6W1H (Who, What, When, Where, Why, Whom, How) +用言でフレーム分けすることで、内容を整理するシステムである。本研究では、Who フレームをさらに Who+フレーム(誰と)に分割した 7W1H+用言でフレーム分けする。例えば、「私と母は昨日、本屋で妹に絵本を買った」という入力は表 2 のようにフレーム分けされる。(使用しないフレームは省略)

表 2 フレーム分けの例

Who	Who+	What	When	Where	Whom	用言
私	母	絵本	昨日	本屋	妹	買う

### 4.4 会話システム

会話システムは人間の発言に応じた応答文を生成して会話を展開するシステムである。このシステムには、人間との自然な会話を実現するために多様な応答手法が導入されている。例として、話題に沿った応答を返す話題掘り下げ応答と人間の感情に反応した応答を返す感情判断応答においては、人間が「車を買いました」と入力すると、喜びの話者感情と「車」という話題語を取得し、「いいですね! どんな車ですか?」と応答する。

話題語とは、会話する上で主題となる事柄を表す語で、会話システムにおいては会話を展開する中で 2 通りの手法で決定する。1 つ目は、関連する話題語への推移である。これはある話題語から関連する次の話題語に推移させる手法である。2 つ目は用言からの話題語の決定である。これは会話開始時の台詞に対してと 1 つ目の手法で話題語が取得できないときに用いる。以降、図 2 に示す会話文を例に各手法を説明する。

A: 昨日、レストランで料理を食べたよ。  
 B: いいですね。美味しかったですか?  
 A: うん。  
 B: レストランはどこにありますか?  
 A: 大阪だよ

図 2 話題語が取得できる会話文の例

図 2 の 1 行目は、意味理解システムによって表 3 のようにフレーム分けされ(使用しないフレームを一部省略)、用言「食べる」を取得する。

表 3 「昨日、レストランで料理を食べたよ」

Who	Who+	What	When	Where	Why	用言
		料理	昨日	レストラン		食べる

表 3 のように用言を取得できる場合、用言になる動詞とその動詞に関係する名詞を用言ごとに整理した京大格フレーム[7]を基に作成された用言格データベースを参照する。これは、動詞と同時に使用される名詞の頻度を格ごとに登録したもので、頻度が最大の格に対応する語を意味理解システムの結果より取得して話題語とする。用言「食べる」に対して京大格フレームで共に用いられる名詞と助詞を獲得すると「子供が、弁当を、先に、…」となり、「子供」は Who 格、「弁当」は What 格、「先」は When 格というように分類して数をカウントする。表 4 に用言格データベースの一部を示す。

表 4 用言格データベースの一部

Who	Who+	What	When	Where	Why	Whom	How	用言
377	408	2159	132	663	32	182	148	食べる
298	1469	7235	246	794	37	374	180	見る

表 4 より、用言「食べる」と共に用いられる名詞は What 格に分類される頻度が最大と分かる。よって、表 3 においては What フレームの「料理」が話題語となる。

図 2 において、1 行目から 3 行目では「料理」、4 行目から 5 行目では「レストラン」が話題語となる。この場合、「料理」が推移前、「レストラン」が推移先の話題語である。推移先となる話題語の候補は、会話文中の各台詞を意味理解システムでフレーム分けして取得する語のうち過去に話題語となった語以外である。話題語を推移させる際、推移前の話題語と推移先の候補とで関連度を計算し、その関連度が最大かつ閾値(0.0035)を超えるものを推移先の話題語とする。表 5 に推移前の話題語「料理」と推移先候補との関連度を示す。

表 5 推移前と推移先候補との関連度

遷移前	遷移先	関連度
料理	レストラン	0.22
	大阪	0
	昨日	0.06

表 5 より、関連度が最大かつ閾値を超える「レストラン」が推移先の話題語となる。

### 4.5 感情判断システム

感情判断システム[8]とは、単文を入力することで、発話者とその単文を発話する際に抱いている感情を出力するシステムである。扱う感情は「喜び、安心、悲しみ、罪悪感、恐れ、恥、怒り、後悔、落胆」の 9 感情と「感情なし」を加えた 10 感情である。例えば、「新車を買いました」と入力すると、発話者の感情として「喜び」を出力する。

### 4.6 文間関連度計算方式

文間関連度計算方式[9]とは、文と文との関連の強さを定量的に表すことができる計算方式である。関連の強さは 0.0 から 1.0 の実数値で表され、値が大きいほど文と文の関連が強いことを示す。本研究で用いる文間関連度は EMD[10]

を用いて算出する。EMD とは線形計画問題の 1 つであるヒッチコック型輸送問題において計算される距離尺度であり、2 つの文に出現する自立語群を離散集合とみなすことで文間関連度を算出する。例えば、文 1「東京の梅」と文 2「八王子のうめの祭り」で文間関連度を求める場合、2 文の自立語をそれぞれ需要地、供給地とする。そして、需要地と供給地の距離を自立語間の関連性、輸送量を自立語の重みと見立てることで、概念ベースにおける一致度計算によって求めることができる。このとき、自立語の重みは tf-idf[11]で重み付けする。一致度は関連性が高いと値も大きくなるため、1 から一致度を減算した値に変換する。図 3 に文 1 と文 2 における計算の様子を示す。「梅」と「祭り」間の輸送量が 1 となるのは「梅」より「うめ」に重み 2 を輸送し、残る重み 1 を「祭り」に輸送するためである。

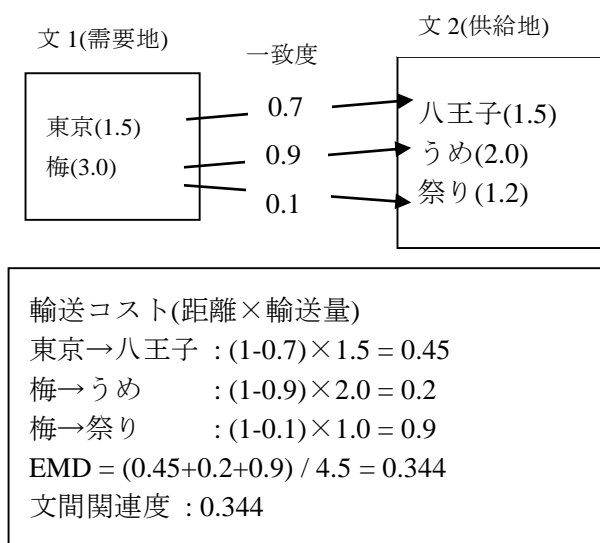


図 3 文間関連度の計算例

## 5. 提案手法

本研究では、台詞とそれを発話する際の動作を登録した動作知識ベースを作成する。動作知識ベースについては次節で述べる。提案手法では、ロボットの台詞および話題語と動作知識ベース内の台詞とで関連の強さを調べ、関連が最も強い台詞に対応する動作を、ロボットの台詞を発話する際の動作として取得する動作生成システムを作成する。関連の強い台詞は関連度計算方式と文間関連度計算方式を用いて決定する。詳しくは 5.2.4 項で述べる。

### 5.1 動作知識ベース

動作知識ベースにはロボットの台詞に対応する動作を生成するための知識を格納する。格納する知識は台詞とそれを発話する際に行う動作である。表 6 に動作知識ベースの一部を示す。動作知識ベースには動作の知識として Robovie に対象の動作を行わせるモーションファイル名を登録する。

表 6 動作知識ベースの一部

台詞	台詞を発話する際の動作 (モーションファイル名)
いいですね	首を傾けて、 両手を胸に当てる
楽しいレジャーを満喫した	右手を胸に当てる

動作知識ベース内の台詞は、ロボットの台詞が感情を含む場合の動作生成に利用するために、感情なしの台詞と感情ありの台詞に大別する。本研究における感情なしの台詞とは、4.5 節で述べた感情判断システムが「感情なし」と出力する台詞で、それ以外の台詞が感情ありの台詞である。扱う感情は感情判断システムに準拠した 9 感情である。

本節では以降、各台詞の収集方法と台詞に対応する動作の設定方法について述べる。

#### 5.1.1 感情なしの台詞

感情なしの台詞としては、中学英語の教科書[12]の中から複数人で会話する場面 (41 場面) を和訳して単文を収集した。中学英語の教科書を採用した理由は、内容が単純で台詞を収集しやすいことと、日常的な会話の場面が多いことが挙げられる。ここで収集した台詞は 440 文 (平常文:342 文, 疑問文:98 文) である。

収集した台詞の中には感情ありの台詞も存在し得るため、収集した台詞を感情判断システムに入力し、感情ありと判別される台詞 (33 文) は次項で述べる感情ありの台詞とする。よって、本研究における感情なしの台詞は、407 文である。

#### 5.1.2 感情ありの台詞

感情ありの台詞としては、前項で述べた中学英語の教科書より収集した感情ありの台詞 (33 文) に加え、感情判断システムの評価セットより収集した。この評価セットは「(主体語)+(修飾語)+名詞+動詞」で構成される単文で、「感情なし」以外の 9 感情を含む文をそれぞれ 10 文 (90 文) 収集した。よって、感情ありの台詞は 123 文である。

#### 5.1.3 動作設定

各台詞に対応する動作は実験およびアンケートで設定する。実験では中学英語の教科書より収集した台詞、アンケートでは感情判断システムの評価セットより収集した台詞に動作を設定する。

実験では中学英語の教科書を和訳した会話文を、登場人物ごとに被験者が読み上げることで教科書中の会話の様子を再現し、会話文中の単文 (440 文) に対応する動作を設定する。

アンケートでは、感情判断システムの評価セットより収集した台詞 (90 文) を被験者に提示する。被験者は提示される各台詞を発話する際に行う動作を選択記述方式、または自由記述方式で回答する。

選択記述方式では、14 種類の動作を選択肢として提示し、被験者は各台詞に対応する動作を選択する。自由記述方式

では、選択記述方式において提示する 14 種類の動作の中に適切なものがない、または選択肢の動作以外により適切な動作がある場合に自由に記述する。

アンケートの結果、選択記述方式において被験者の回答が重複した動作と、自由記述方式において回答された動作を対象の台詞を発話する際の動作とする。自由記述方式において回答された動作を必ず採用する理由は、既述の「選択肢の動作以外により適切な動作がある場合」という条件の下に回答されたもので、選択肢にある動作では表現できない台詞中の意味や意図に直接起因する動作であると考えられるためである。

なお、設定した各動作に関しては、Robovie に対象の動作を行わせるモーションファイルを作成する。

#### 5.1.4 代表的動作の設定

感情ありの台詞に動作設定を行うと、感情ごとに重複する動作が見られる。例えば、25 文ある喜びの感情を含む台詞に対応する動作の中でも、「万歳」の動作は 6 文の台詞に対して設定され、他の動作と比較して最も設定される頻度が高かった。そこで、この動作を喜びの感情を表す代表的動作として、5.2.4 項で述べる動作取得に利用する。表 7 に感情ごとに設定した代表的動作の一部を示す。

表 7 感情を表す代表的動作

感情	感情を表す代表的動作
喜び	万歳
怒り	顔をかしげ、右手を腰に当てる
悲しみ	両手を目に当てる

また、感情なしの台詞のうち疑問文に動作を設定すると、98 文の疑問文のうち 48 文で顔をかしげ、左右どちらかの手もしくは両手を前に向ける動作が設定された。そこで、本研究では「顔をかしげ、右手を前に向ける」動作を疑問文における代表的動作とする。

## 5.2 動作生成システム

動作生成システムは台詞をテキストベースで入力すると、その台詞を発話する際に行う動作（モーションファイル名）を取得し、Robovie を動作させる。このシステムは会話システムと同時に実行することを想定して構築しており、会話システム内の処理で取得する話題語を動作取得に用いる。また、会話システムが出力する台詞が感情ありの場合は、その感情をロボットの感情とし、動作取得に利用する。図 4 に処理の流れを示す。

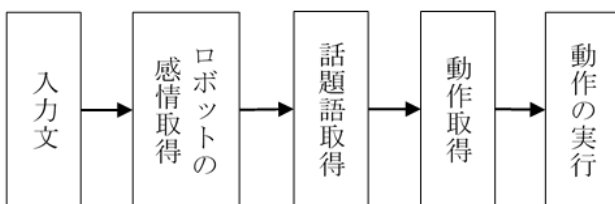


図 4 動作生成システムの流れ

### 5.2.1 入力文

本研究では会話システムの出力（ロボットの台詞）を動作生成システムの入力文として採用する。例えば、会話システムに「レストランで美味しい料理を食べたよ」と入力した際に出力される「いいですね」が動作生成システムの入力文となる。

### 5.2.2 ロボットの感情取得

ロボットの感情は前述の入力文を感情判断システムに入力することで取得する。例えば、入力文が「花を貰いました」の場合、感情判断システムは「喜び」と出力し、これを発話するロボットの感情は「喜び」となる。

### 5.2.3 話題語取得

話題語は、4.4 節で述べた通り会話システムが会話を展開する中で取得、推移する。動作生成システムは、これを順次参照して動作生成に用いる。

### 5.2.4 動作取得

動作取得は、入力文が平常文の場合（下記の A から D）と、疑問文の場合（下記の E, F）に分けて行う。

- A. 平常文 & 感情あり & 話題語あり
- B. 平常文 & 感情あり & 話題語なし
- C. 平常文 & 感情なし & 話題語あり
- D. 平常文 & 感情なし & 話題語なし
- E. 疑問文 & 話題語あり
- F. 疑問文 & 話題語なし

入力文が平常文のとき(A~D)の処理について述べる。

A のようにロボットの感情と話題語を取得できる場合は、5.1.2 項で述べた感情ありの台詞の中から、ロボットの感情と同じの感情を含む台詞を抽出し、各台詞中の自立語と話題語との関連度を求め、台詞ごとに関連度の平均値を計算する。そして、その平均値が最大の台詞に対応する動作を取得する。これで、動作知識ベースの感情ありの台詞から、話題語と関連が強い台詞を調べ、それに対応する動作を取得できる。図 5 にロボットの台詞が「お菓子を買う」、ロボットの感情「喜び」、話題語「お菓子」のときの例を示す。この場合、動作知識ベース中の「喜び」の感情を含む台詞において自立語と話題語との関連度の平均値を計算する。「甘いアイスを食べた」における関連度の平均値は 0.23 となり、同様に他の台詞においても関連度の平均を計算した結果、この値が最大の場合には「首をかしげてほほに手を当てる」が入力文に対応する動作となる。

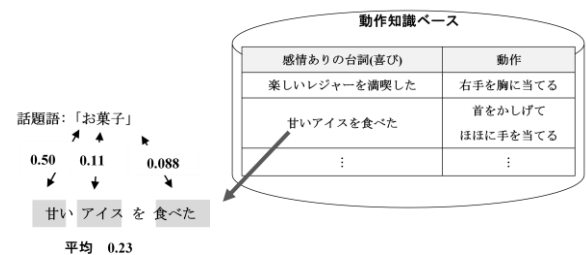


図 5 「平常文 & 感情あり & 話題語あり」の計算例

Bのようにロボットの感情あり、話題語なしの場合、ロボットの感情に対応する動作を 5.1.4 項で述べた代表的動作より取得する。これにより、ロボットの感情を動作に反映することができる。例えば、ロボットの台詞が「チームが優勝する」の場合、ロボットの感情が「喜び」となり、喜びの感情を表す代表的動作「万歳」を入力文に対応する動作とする。

Cのようにロボットの感情なし、話題語ありの場合、5.1.1 項で述べた感情なしの台詞（平常文）において、自立語と話題語との関連度を計算して、Aの場合と同様にその平均値が最大となる台詞に対応する動作を取得する。これで、動作知識ベースの感情なしの台詞から、話題語と関連が強い台詞を調べ、その台詞に対応する動作を取得できる。図 6 にロボットの台詞が「スポーツを習う」、話題語が「スポーツ」のときの計算例を示す。全平常文において関連度の平均を計算した結果、「僕はサッカーをします」における関連度の平均 0.0082 が最大であれば、「両腕を振る」が入力文に対応する動作となる。

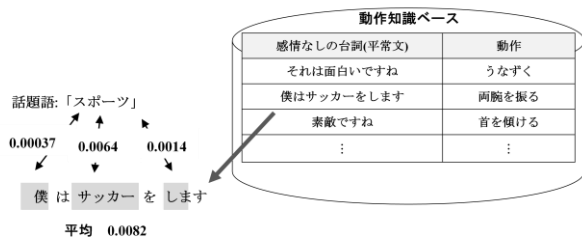


図 6 「平常文 & 感情なし & 話題語あり」の計算例

Dのようにロボットの感情と話題語の両方を取ることができない場合は、5.1.1 項の感情なしの台詞（平常文）と入力文との文間関連度が最大の台詞に対応する動作を取得する。これにより、動作知識ベースより入力文と関連が強い台詞に対応する動作を取得できる。図 7 にロボットの台詞が「ジョギングが日課です」のときの計算例を示す。全平常文において入力文との文間関連度を計算した結果、「僕はサッカーをします」における文間関連度が最大であれば、「両腕を振る」が入力文に対応する動作となる。

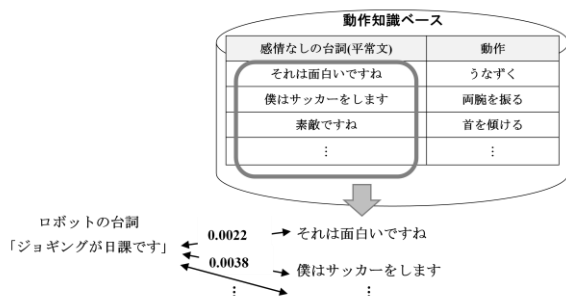


図 7 「平常文 & 感情なし & 話題語なし」の計算例

次に、入力文が疑問文のときの E, F について述べる。  
Eのようにロボットの台詞が疑問文で話題語ありの場合、5.1.1 項で述べた感情なしの台詞（疑問文）において、疑問

文中の自立語と話題語との関連度を計算して、Aの場合と同様にその平均値が最大となる台詞に対応する動作を取得する。図 8 にロボットの台詞が「水族館にはよく行きますか?」、話題語が「水族館」のときの計算例を示す。全疑問文において関連度の平均値を計算した結果、「あれは遊園地ですか?」における関連度の平均 0.06 が最大であれば、「首をかしげる」が入力文に対応する動作となる。

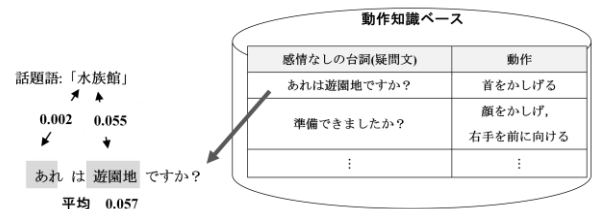


図 8 「疑問文 & 話題語あり」の計算例

Fのようにロボットの台詞が疑問文で話題語なしの場合は、5.1.4 項で述べた疑問文における代表的動作「顔をかしげ、右手を前に向ける」を入力文に対応する動作とする。

## 6. 評価

動作生成システムは会話システムの出力を入力文として採用するため、会話システムの精度に大きく左右されると予想できる。そこで、動作生成システム単体での評価（6.1 節）と、会話システムの出力を動作生成システムの入力文とする全体での評価（6.2 節）を行う。本章では以降、各評価の方法と評価結果について述べる。

### 6.1 動作生成システム単体の評価

動作生成システム単体の評価では、入力する 1 文に対して生成される動作を評価する。しかし、動作生成システムは話題語の取得を会話システムに依存しており、本評価においては、各入力文に対して話題語を別途用意する必要がある。そこで、感情ありと感情なしの両方を含む台詞（250 文）に 4.4 節で述べた用言からの話題決定手法を適用し、入力文ごとに話題語を設定した評価セットを作成した。表 8 に評価セットの一部を示す。

表 8 動作生成システム単体の評価セットの一部

入力文	話題語
綺麗な絵を見ました	絵
道で転びました	なし
約束を破ってしまった	約束

被験者は、表 8 に示すような入力文ごとに動作するロボットを見た上で、その動作が入力文に対して「○：常識的」、「△：非常識とは言えない」、「×：非常識」かの 3 段階で評価する。精度は「○：常識的」の割合とする。

### 6.2 会話システム+動作生成システムの評価

会話システムの出力を動作生成システムの入力文とする場合の評価では、会話システムと人間との会話文を用いる。会話文は 15 セットあり、各セットそれぞれ人間とロボットの台詞が 1 から 5 往復するものである。図 9 に会話文

の例を示す。この例では、1行目から3行目で話題語なし、4行目から5行目で「友達」、6行目で「学校」が話題語となる。

人間	: 学校で勉強した.
ロボット	: 誰としたのですか?
人間	: 友達とです.
ロボット	: 友達の誰さんですか?
人間	: 吉川さんとです.
ロボット	: 学校はどこにありますか?

図 9 全体の評価で使用する会話文の例

被験者は上記のような会話文の人間の台詞を話していると仮定し、ロボットの台詞ごとに Robovie が実際に動作する様子を見た上で、会話文ごとに以下の指標に○△×で回答する。各指標の精度は○の割合とする。

- ① ロボットの台詞と動作が合っていたか
- ② 人間味のあるロボットであったか
- ③ 動作はスムーズであったか
- ④ 感情を正しく表現できていたか

### 6.3 評価結果

動作生成システム単体と会話システム+動作生成システムでの評価結果をそれぞれ図 10 と図 11 に示す。

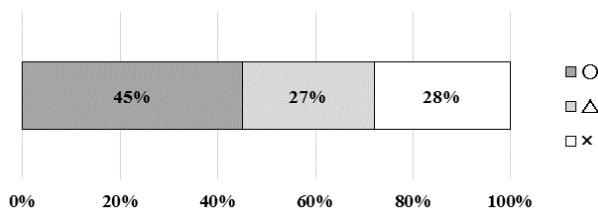


図 10 単体の評価結果

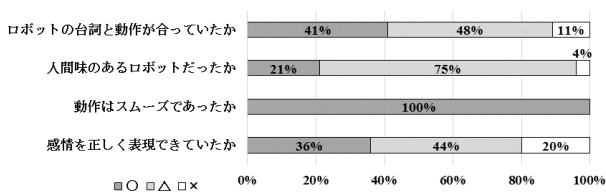


図 11 全体の評価結果

## 7. 考察

本章では 6 章の動作生成システム単体と会話システム+動作生成システムでの評価結果に対して考察を述べる。

### 7.1 動作生成システム単体における考察

図 10 より動作生成システム単体の精度は 45%となった。評価では同じ台詞において出力される動作でも被験者ごとに回答のばらつきが見られた。そこで以降、多数の被験者で重複した回答を各台詞における回答として考察を述べる。

常識的と回答された例を図 12 に示す。

入力文	関連の強い台詞
ギターを弾いた	ギターと演奏します
値段が上がった	その金額はフェアなのかもしれませんね

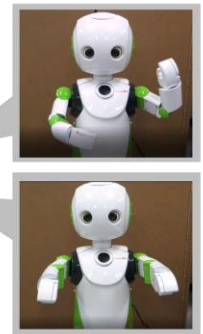


図 12 常識的と回答された例

入力文「ギターを弾いた」においては話題語が「ギター」で、5.2.4 項で述べた動作取得のうち C の処理が適応される。結果、動作知識ベース内の台詞「ギターを演奏します」において話題語「ギター」と各自立語との関連度の平均値が最大で、この台詞に対応する動作が行われた。このように、常識的な動作を取得できた入力文には、動作知識ベース内の台詞に含まれる語と同一の語、または関連の強い語が含まれる傾向が見られた。よって、動作知識ベース内の台詞と、入力文または話題語との関連の強さを考慮したロボットの動作決定手法に一定の有効性を示すことができた。

次に非常識または非常識とは言えないと回答された例を図 13 に示す。

入力文	関連の強い台詞
河豚に中(あた)った	納豆を食べます
ゲームをした	パソコンのそばにあるじゃないですか



図 13 非常識または非常識とは言えないと回答された例

入力文「河豚に中(あた)った」においては話題語が「河豚」、ロボットの感情「恐れ」で、動作取得のうち A の処理が適応される。結果、動作知識ベース内の「恐れ」の感情ありの台詞 13 文の中で、話題語と自立語との関連度の平均値が最大の台詞は「納豆を食べます」となり、対応する動作が出力された。これは話題語「河豚」と「納豆」および「食べる」という概念間の関連が強い傾向にあり、他に話題語と関連の強い語を含む台詞が動作知識ベースになかったためである。実行される動作は納豆を食べる様子を表すもので、入力文「河豚に中(あた)った」の「河豚」または「中(あた)る」のどちらの語にも対応しなかったため、被験者は非常識または非常識とは言えないと回答したと考えられる。

以上のことから、動作知識ベース内の台詞に入力文または話題語と関連の強い語が不足していたために、常識的な動作を出力できなかったと考える。そこで今後は会話コーパスなどより台詞をさらに収集して、動作知識ベース内

の台詞に含まれる語彙の数と同時に動作の種類を増やし、より多様な入力文に対応する動作知識ベースを作成する必要があると考える。その中で、動作取得における計算対象の台詞が増えることで、計算の負荷が増えて処理時間が長くなり、ロボットの反応が鈍くなるという問題が発生する。対策として、人間が日常的な会話の中で頻繁に使用する語彙や、発話する際に動作で表現されやすい語彙を調査し、それらを含む台詞を優先的に採用する方法が考えられる。例えば、5.1.3 項で述べた動作設定の実験では、「食べる」という語では手を口に運ぶ動作というように、発話する際に動作で表現されやすい語が確認された。反対に国名や名前などの固有名詞に加え、「数学」や「セール」などの動作で表現するのが難しい語が存在することも確認された。この点を利用して、動作知識ベースが無作為に拡張されるのを防ぐ必要があると考える。

## 7.2 会話システム+動作生成システムにおける考察

動作生成システムを会話システムと同時に実行する際の問題点として、動作取得において会話の中で既に行った動作を考慮しないことが挙げられる。図 14 に例を示す。

人間	: タイタニックを観た。
ロボット	: タイタニックは好きですか？
人間	: うん。
ロボット	: 誰とタイタニックを観ましたか？
人間	: 一人で観ました。
ロボット	: 一人ですか
人間	: うん。
ロボット	: どこで観ましたか？




図 14 全体の評価における動作例

図 14 の例では 4 文あるロボットの台詞のうち、3 文において Robovie は同図に示す動作を行った。この動作は動作知識ベース内の疑問文に対して設定された動作の 1 つであり、この会話文中の個々の疑問文には合致する動作であったと考えられる。しかし、会話文においては同じ動作を重ねて行うことになり、台詞単位では合致する動作であっても会話文全体を通してみると不自然なものになった。そこで今後の課題として、会話の中で行った動作を履歴として保持することで、特定の動作を会話の中で連続して、または重ねて行わないようにする処理が必要である。

## 8. おわりに

本研究では、台詞とその台詞を発話する際の動作を登録した動作知識ベースを作成した。そして、動作知識ベースを用いることで、動作とあらかじめ対応付けした台詞と入力文（ロボットの台詞）、または会話文中より取得する話題語との関連の強さを動作取得に反映する動作生成システムを作成した。動作生成システム単体の評価結果より精度が 45% で一定の成果が見られ、台詞と話題語、もしくは台詞間の関連の強さより動作を決定する手法が有効であったと考える。

**謝辞** 本研究の一部は、JSPS 科研費 16K00311 の助成を受けて行ったものです。

## 参考文献

- [1] 長村里穂, 吉村枝里子, 土屋誠司, 渡部広一, “話し手の感情を考慮したロボットの感情と動作の生成”, 情報科学技術フォーラム FIT2014, J-044, pp.367-368, 2014.
- [2] 金子稜, 吉村枝里子, 土屋誠司, 渡部広一, “話題を考慮した自然な会話システムの構築”, 研究報告知能システム ICS, 2016-ICS-183(2), pp.1-8, 2016.
- [3] “Robovie-R3 | ヴイストーン株式会社”, [http://www.vstone.co.jp/products/robovie\\_r3/](http://www.vstone.co.jp/products/robovie_r3/), 13.July.2016.
- [4] 奥村紀之, 土屋誠司, 渡部広一, 河岡司, “概念間の関連度計算のための大規模概念ベースの構築”, 自然言語処理, Vol.14, No.5, pp.41-64, 2007.
- [5] 井筒大志, 渡部広一, 河岡司, “概念ベースを用いた連想機能実現のための関連度計算方式”, 情報科学技術フォーラム FIT2002, E-39, pp.159-160, 2002.
- [6] 篠原宜道, 渡部広一, 河岡司, “常識判断に基づく会話意味理解方式”, 言語処理学会第 8 回年次大会発表論文集, pp.651-654, 2002.
- [7] 河原大輔, 黒橋禎夫, “高機能計算環境を用いた Web からの大規模格フレーム構築”, 情報処理学会自然言語処理研究会資料, 2006-NL-171-12, pp.67-73, 2006.
- [8] 小野彰利, 芋野美紗子, 土屋誠司, 渡部広一, “コミュニケーションロボットのための自然言語入力に基づいた感情判断システムの構築”, 情報科学技術フォーラム FIT2015, J-007, pp.321-322, 2015.
- [9] 藤江悠五, 渡部広一, 河岡司, “概念ベースと Earth Mover’s Distance を用いた文書検索”, 自然言語処理, Vol.16, No.3, pp.325-349, 2009.
- [10] Y.Rubner, C.Tomasi, L.Guibas “The earth mover’s distance as a metric for image retrieval”, Int.J.Comput.Vision, Vol.40, pp.99-121, 2000.
- [11] G.Salton and C.Buckley. Term-weighting approaches in automatic textretrieval. Information Processing and Management, Vol.41, No.4, pp.513-523, 1998.
- [12] “NEW HORIZON 1~3”, 東京書籍, 2015.
- [13] 宮崎奇, 中村友昭, 金子正秀, “台詞に含まれる感情表現を反映したロボット動作の自動生成”, 映像情報メディア学会技術報告, メディア工学研究会, ME2015-46, vol.39, no.8, pp.41-44, 2015.