

# 人のためらい：ロボットの行動が「おいしい」とき 評価は遅れて与えられる

## The hesitation of humans: When the action of a robot is “almost” correct, they give a delayed evaluation

田中 一晶<sup>1\*</sup> 尾関 基行<sup>1</sup> 岡 夏樹<sup>1</sup>

TANAKA Kazuaki<sup>1</sup>, OZEKI Motoyuki<sup>1</sup>, OKA Natsuki<sup>1</sup>

<sup>1</sup> 京都工芸繊維大学 大学院工芸科学研究科

<sup>1</sup> Graduate School of Science and Technology, Kyoto Institute of Technology

**Abstract:** It is known that an example of near-miss, an almost correct example, is useful for learning from examples (Winston, 1975). In our previous work, we found that humans give a hesitant (delayed) evaluation when a robot's action for their command is almost correct. It is possible that the delay in the evaluation is useful information for the word meaning acquisition of “almost”, and identifying the situation of almost. In this work, we performed an experiment in which a robot indicates one or two objects according to a human's instruction, and then he/she evaluates the action using three kinds of words: “yes”, “no”, and “almost”. The result of the experiment showed that the delay of the evaluation of “almost” is longer than that of “yes” and “no”, and that the delayed evaluation is given at the situation in which the action of the robot is almost correct regardless of the kinds of evaluations.

### 1 はじめに

近い将来、日常生活の場に人の仕事を手伝うロボットが登場することが予想される。ロボットの使用者や使用環境は家庭により異なり、時々刻々と変化する。したがって、ロボットは予め決められた指示通りに動作するだけでなく、人から与えられる肯定・否定の評価を理解し、その評価に基づいて自身の行動を修正することが重要である。

人とのインタラクションを通してロボットが学習する場合、言葉による明示的な評価だけでなく、人の自然な振る舞いに含まれる非言語情報が評価として利用できることが示されている [1, 2]。[1] は、音声のピッチの急激な上昇を警告 (否定的な評価) として利用している。また、[2] は、ロボットが行動を開始してから一定時間人の発話が無いことを肯定的な評価として利用している。人の自然な振る舞いに含まれる非言語情報も評価として利用することで、ロボットはより効率的に学習することが可能になる。

また、[3] では、積み木の構造を意味ネットワークを用いて学習する上で、ニアミス (おいしい状況) から効率的に学習できることが示されている。したがって、

肯定・否定の評価だけでなく「おいしい」という評価もロボットの学習に役立つと考えられる。我々は、人-ロボット教示インタラクションを観察する中で、ロボットの行動がおいしいとき人が評価を与える際にためらうことを発見した。我々の先行研究 [4] では、人がロボットに「お手」(座って右手を出す) を教える実験において、ロボットが右手ではなく左手を出した際に、人がためらって (遅れて) 否定の評価を与える場面が見られた。また、[5, 6] では、人がためらっていることを知る上で有効に働く特徴として、発話の中断や遅延 [5]、基本周波数の変化が小さいことや発話時間が長いこと [6] が示されている。このように、人のためらいは時間的な特徴と密接に関係しており、「おいしい」という評価がためらって与えられる場合には、評価の遅延時間を利用して「おいしい」という意味の言葉を獲得することや、おいしい状況 (または行動) とはどのような状況であるかを特定することができる可能性がある。

本研究では、人が言葉で指示したものをロボットが指差すタスクを設定し、人に3種類の評価「よし」「ちがう」「おいしい」を与えてもらう実験を実施した。その結果から、以下の2点を明らかにする。

1. 「おいしい」という評価は「よし」「ちがう」より遅れて与えられるか。
2. ロボットがおいしい行動を実行した状況では、評価

\*連絡先：京都工芸繊維大学 大学院工芸科学研究科  
〒606-8585 京都市左京区松ヶ崎橋上町  
E-mail: d8821007@edu.kit.ac.jp

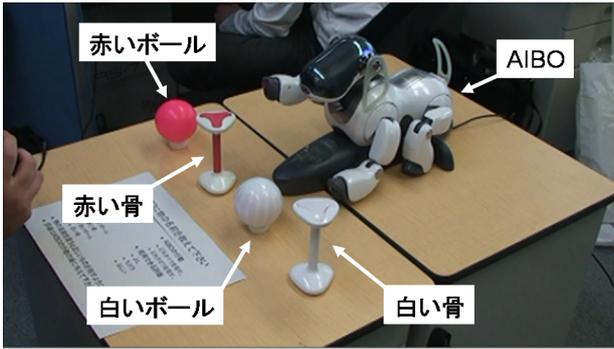


図 1: 実験の一場面

## AIBOに物の名前を教えてください。

- 物の名前
    - 赤
    - 白
    - ボール
    - 骨
    - 赤いボール
    - 赤い骨
    - 白いボール
    - 白い骨
  - AIBOの行動
    - どれか1つを指す。
    - どれか2つを指す。
  - 使用できる評価
    - よし
    - ちがう
    - おいしい
- 物の名前を言うと正しいものを指すように教えてください。
- 評価はAIBOが物を指し終えてから与えて下さい。

図 2: 実験参加者に提示した紙面

本研究では、色(赤/白)や形(ボール/骨)も物の名前の一つとして扱う。例えば、赤いボールは「赤」や「ボール」といった名前でも指示される。

の種類に関わらず遅れて評価が与えられるか。

## 2 実験

本研究では、人が言葉で指示したものを AIBO ERS-7 が指差すタスクを採用した。実験の一場面を図 1 に示す。AIBO が指差す物は図中に記している。本節では、実験の手順 (2.1 節) やプログラムの仕様 (2.2 節) など、実験設定について説明する。

### 2.1 実験の手順

まず、実験参加者には図 2 を A4 の紙 1 枚で提示し、口頭で以下の説明を行った。

- AIBO に物の名前を教えてください。物の名前を言うと正しいものを指すように教えてください。
- まず、紙に書かれている 8 種類の物の名前から 1 つ選び AIBO に指示して下さい。

- AIBO はどれか 1 つまたは 2 つ選んで指差します。
- AIBO が正しいものを指差したときには「よし」、誤ったものを指差したときには「ちがう」、おいしいと思ったときには「おいしい」と言って評価して下さい。
- AIBO は物を指し終えた後に顔を上げます。指し終えてから評価を与えて下さい。
- 実験は 30 分行います。

実験中に記録する情報は以下の 4 つである。

- 人が指示した物の名前
- AIBO の行動 (何を指差したか)
- 人の評価
- AIBO が物を指し終えてから人が評価を与えるまでの時間

また、実験後にどのような状況でおいしいと思ったか、実験参加者に口頭でたずねた。

実験参加者 5 名の属性を表 1 に示す。

表 1: 実験参加者の属性

実験参加者	属性	年齢
A	社会人 (男性)	20 代
B	社会人 (女性)	50 代
C	社会人 (男性)	20 代
D	社会人 (男性)	20 代
E	社会人 (男性)	20 代

### 2.2 プログラムの仕様

図 3 に実験で使用したプログラムのフローチャートを示す。プログラムは Visual C++ を用いて実装した。また、AIBO の開発環境である AIBO Remote Framework を使用し、無線 LAN により AIBO を遠隔操作した。

表 2 に人の指示 (物の名前) と AIBO の行動 (AIBO が指差す物) を示す。指示と隣り合う行動が、その指示に対する正しい行動である。また、AIBO が 2 つの物を指す場合には 1 つずつ指すが、その順番はランダムに決定する。

人の指示 (物の名前) や評価の音声認識は、Julius[7] を用いて行った。Julius には認識対象とする言葉 (物の名前: 8 種類、評価: 3 種類) と、「あー」「えー」といったノイズとなる言葉を予め登録しておいた。

人の指示を認識すると、AIBO が正しい物を指差す確率  $p_t$  に従って、正しい物 ( $p_t$ ) が誤った物 ( $1 - p_t$ ) どちらを指すか決定する。 $p_t$  の初期値は  $0.1(1/10)$  とし、人が評価を与える度に、以下の式に従って更新する。 $\alpha$  は  $p_t$  の上昇率、 $P$  は  $p_t$  の最大値を示しており、それぞれ  $0.02$ 、 $0.7$  とした。

$$p_{t+1} \leftarrow p_t + \alpha(P - p_t) \quad (1)$$

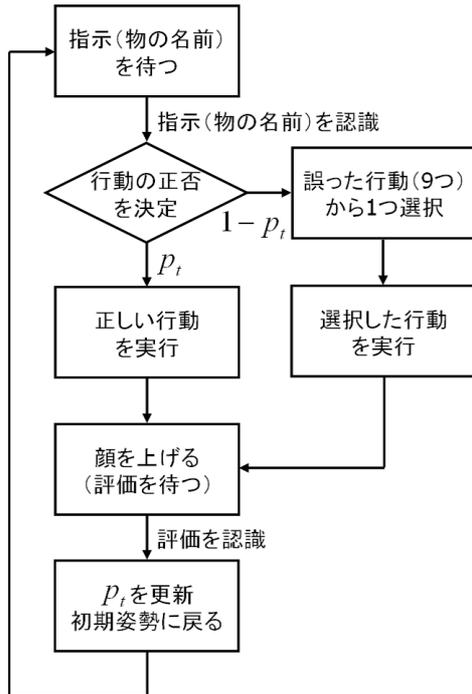


図 3: 実験で使ったプログラムのフローチャート  
 図中の  $p_t$  は AIBO が正しい物を指差す確率を示している。また、 $p_t$  の更新方法は 2.2 節を参照せよ。

誤った物を指す場合、表 2 に示した AIBO の行動の内、正しい行動を省いた残りの 9 つの行動からランダムに 1 つ選択する。

### 3 実験結果

本節では、実験結果を示し、評価の種類と評価の遅延時間の関係、状況の種類と評価の遅延時間の関係についてそれぞれ分析する。

#### 3.1 評価の種類と評価の遅延時間の関係

代表的な例として、図 4 に実験参加者 D の各評価(「よし」「ちがう」「おいしい」)ごとの遅延時間(ロボットが物を指し終えてから評価が与えられるまでの時間)の度数分布表を示す。度数分布表から分かるように、「よし」「ちがう」と比べて、「おいしい」という評価は、より長い遅延時間の階級に多く分布していることがわかる。

そこで、各評価ごとに遅延時間に違いがあるか調べるため、実験参加者全員の実験結果を用い、各評価ごとの遅延時間の平均値について、評価の種類を要因とした 1 要因分散分析を行った。その結果、主効果が認められたため ( $F(2, 8) = 12.67, p < .01$ )、多重比較を行った。図 5 に各評価ごとの平均値と多重比較の結果

表 2: 人の指示と AIBO の行動

人の指示	AIBO の行動
「赤」	赤いボール、赤い骨を指す
「白」	白いボール、白い骨を指す
「ボール」	赤いボール、白いボールを指す
「骨」	赤い骨、白い骨を指す
「赤いボール」	赤いボールを指す
「白いボール」	白いボールを指す
「赤い骨」	赤い骨を指す
「白い骨」	白い骨を指す
-	赤いボール、白い骨を指す
-	白いボール、赤い骨を指す

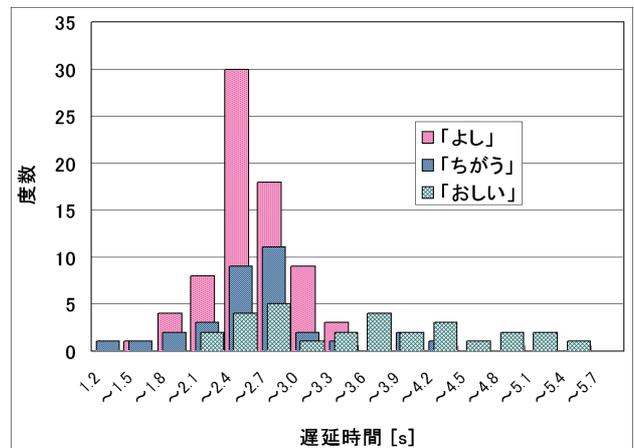


図 4: 実験参加者 D の各評価ごとの遅延時間の度数分布表

を示す。「よし」と「おいしい」、「ちがう」と「おいしい」との間にそれぞれ有意差が認められた ( $p < .01$ )。つまり、「おいしい」という評価は「よし」「ちがう」より遅れて与えられることが分かった。

#### 3.2 状況の種類と評価の遅延時間の関係

実験後、どのような状況でおいしいと思ったか、実験参加者に口頭でたずねた。共通して挙げられた状況を以下に示す。

[色や形を指示したとき]

- 1 つだけ正しい物を指した。  
例: 「赤」 赤いボールだけ指す
- 2 つ指した内、1 つだけ正しかった。  
例: 「ボール」 赤いボールと白い骨を指す

[特定の物を指示したとき]

- 2 つ指した内、1 つは正しかった。  
例: 「赤いボール」 赤いボールと白い骨を指す

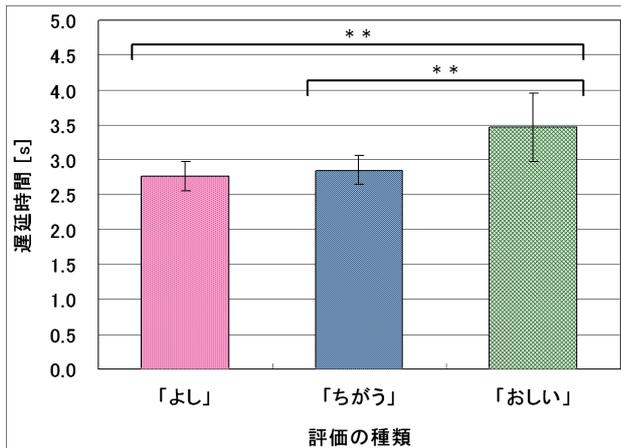


図 5: 各評価ごとの遅延時間の平均値と多重比較の結果  
図中のエラーバーは平均値±標準偏差を示している。また、\*\*は1%の有意水準を示している。

- 指した物の色または形が合っていた。  
例:「赤いボール」 白いボールを指す

上記の4つの状況をおいしい状況として定義し、おいしい状況のほか、正しい状況(指示通りの物を指した状況)、誤りの状況(それ以外の状況)に実験データを分類した。そして、各状況ごとに評価の遅延時間に違いがあるか調べるため、各状況ごとの遅延時間の平均値について、状況の種類を要因とした1要因分散分析を行った。その結果、主効果が認められたため( $F(2, 8) = 6.35, p < .05$ )、多重比較を行った。図6に各状況ごとの平均値と多重比較の結果を示す。正しい状況とおいしい状況との間に有意差が認められ( $p < .05$ )、誤りの状況とおいしい状況との間には有意傾向が認められた( $p < .1$ )。つまり、おいしい状況では、正しい状況より遅れて評価が与えられること、誤りの状況より遅れて評価が与えられる傾向があることが分かった。

## 4 考察

[3]では、積み木の構造を意味ネットワークを用いて学習する上で、ニアミスの例(おいしい状況)から効率的に学習できる方法が示されている。したがって、ロボットの学習においても、おいしいという情報は効率的に学習する上で重要であると考えられる。本節では、実験結果を考察し、評価の遅延時間を利用して「おいしい」という意味の言葉を獲得する方法(4.1節)や、評価の遅延時間を利用しておいしい状況を特定する方法(4.2節)を提案する。また、強化学習において「おいしい」という情報を利用して効率的に学習する方法(4.3節)についても提案を行う。

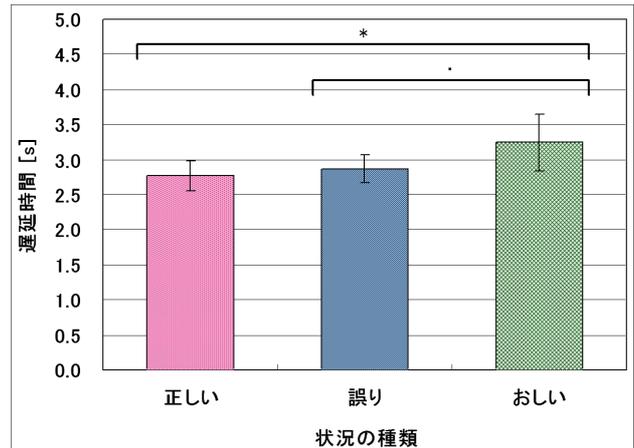


図 6: 各状況ごとの遅延時間の平均値と多重比較の結果  
図中のエラーバーは平均値±標準偏差を示している。また、\*は5%、·は10%の有意水準を示している。

### 4.1 評価の遅延時間を利用した「おいしい」という意味の言葉の獲得

3.1節に示した通り、「おいしい」という評価は「よし」「ちがう」より遅れて与えられることが分かった。したがって、評価の遅延時間は「おいしい」という意味の言葉を獲得する上で有効な情報であると考えられる。

しかし、図4に示した度数分布表の例からも分かるように、「おいしい」という評価が全て遅れて与えられるわけではない。実験の終盤では、実験参加者はどのような状況で「おいしい」という評価を与えるかある程度方向性を決めており、ためらわずに(即座に)「おいしい」という評価を与える場面も見られた。この問題は、学習が進むにつれて次第に無視できなくなることが予想されるが、実際にロボットが学習する場面においては大きな問題にはならないと考えている。本研究では、多くのデータを集めるため、実験時間中(30分)、実験参加者は常に評価を与えなければならないという設定にして実験を行ったが、我々の先行研究[8, 4]では、ロボットの学習が進むにつれて、人は評価を与えなくなっていくことがわかっている。つまり、人がためらわずに評価を与えるようになる頃には、評価の頻度が減少しているため、この問題の影響は小さくなると思われる。

評価の遅延時間を利用して「おいしい」という意味の言葉を獲得する方法としては、サポートベクターマシンやベイジアンネットワークなどの教師あり学習手法を用いて、評価の遅延時間と肯定/否定/おいしいという評価との関係を学習する方法が考えられる。これにより、未知の言葉(「もうちょっと」「あと少し」など)で評価が与えられた場合には、その評価の遅延時間から、どの評価に属するか推測できる。もちろん、肯定

／否定の評価を遅延時間だけで分類することは難しい。肯定／否定の言葉は韻律情報を用いて分類できることが多くの研究で示されており [9, 10, 11]、また、人のためらいも韻律情報を用いて検出できることが示されているため [5, 6]、韻律情報を併用することで、より高い精度で評価の分類が行えることが期待される。

## 4.2 評価の遅延時間を利用したおしい状況の特定

3.2 節に示した通り、おしい状況では、正しい状況より遅れて評価が与えられること、誤りの状況より遅れて評価が与えられる傾向があることが分かった。この分析を行ったのは、実験の中で評価が変化した状況があり、実験参加者がためらって「よし」や「ちがう」と評価した場面も見られたからである。つまり、評価の種類に関わらず、おしい状況では、評価が遅れて与えられると言える。

おしい状況と誤りの状況の違いは有意傾向であった原因は、1 名の実験参加者 (E) が、おしい状況の定義 (実験参加者全員が共通して挙げたおしい状況) 以外でも、おしいと考えていた状況があったからである。その状況を以下に示す。

[特定の物を指示したとき]

- 2 つ指した内、どちらか (または両方) の色または形が合っていた。

例: 「赤いボール」 赤い骨と白いボールを指す

他の実験参加者は、このような状況ではためらわずに「ちがう」という評価を与えており、おしい状況とは考えていなかったと思われる。したがって、おしい状況の考え方が異なる多数の人がロボットに評価を与えるような場合には、評価の遅延時間からおしい状況を特定することは難しい。しかし、特定の一人がロボットに評価を与える場合を想定すると、評価の遅延時間はおしい状況を特定する上で十分に有効に働くと考えられる。

評価の遅延時間からおしい状況を特定する手段として、Fisher の直接確率計算法を紹介する。[12] では、PC 画面上の迷路内をランダムに動作するエージェント (迷路の正しい道筋を知っている) を人が声で誘導するタスクにおいて、状況 (各マス目での行動の正否) と指示・評価の対応付けを行う上で、状況と発話が共起する確率を Fisher の直接確率計算法を用いて計算している。状況と評価の遅延時間との共起確率を計算する場合には、遅延時間をいくつかの区間に区切る必要がある。その区間の決定方法として、度数分布表の階級を決定する際に用いられるスタージェスの公式が有効であると考えている。

スタージェスの公式とは、データ数  $n$  に対して階級

数  $k$  を決める公式である (式 (2))。階級幅  $c$  は階級数  $k$  とデータの範囲 (データの最大値 - 最小値)  $r$  より式 (3) で計算される。

$$k = 1 + \log_2 n \quad (2)$$

$$c = r/k \quad (3)$$

階級幅  $c$  を遅延時間の区間とし、Fisher の直接確率計算法を適応することで、評価の遅延時間からおしい状況を特定する (長い遅延時間 (の区間) と高い確率で共起している状況をおしい状況と判断する) ことが可能になると考えられる。

## 4.3 強化学習における「おしい」という情報の利用

人 - ロボット教示インタラクションにおいて、ロボットの望ましい状態 (および行動) を人が一つ一つ指示するのは負担が大きい。したがって、ロボットは環境から得られる報酬 (人から与えられる評価を含む) に基づいて強化学習 [13] も行えることが重要である。

人が目標とするロボットの望ましい状態は、人が「おしい」と評価した状態の近くに存在する可能性が高い。したがって、「おしい」という情報を利用するためには、連続な状態空間 (および行動空間) が扱える必要がある。

代表的な強化学習 (TD-Learning、Q-Learning など) は離散的な状態・行動を対象としているが、強化学習において連続的な状態を扱う手法としてタイルコーディングが知られている [13]。タイルコーディングとは、タイルと呼ばれる格子で連続状態空間をマス目 (タイル) に区切るコーディング手法であり、連続値が与えられると、その値が属するタイルを離散状態として得ることができる。このようにコーディングしておくことで、評価が遅れて与えられた場合や「おしい」という意味の評価が与えられた場合には、周辺の状態 (タイル) においても同時に学習 (状態 / 行動価値を更新) を行い、周辺の状態に遷移するように探索を行うことで、素早く目標状態に到達することが期待できる。

## 5 まとめと今後の展望

本研究では、人が指示したものをロボットが指差す実験を実施し、「おしい」という評価は「よし」「ちがう」より遅れて与えられること、評価の種類に関わらず、おしい状況では評価が遅れて与えられることを示した。これらの結果から、評価の遅延時間を利用して「おしい」という意味の言葉を獲得する方法や、評価の遅延時間を利用しておしい状況の特定する方法を提案した。また、強化学習において「おしい」という情報

を利用する方法についても提案を行った。今後、これらの提案手法の有用性を明らかにしていきたい。

## 謝辞

本研究は科研費 (21500137) の助成を受けたものである。

## 参考文献

- [1] Komatsu, T., Utsunomiya, A., Suzuki, K., Ueda, K., Hiraki, K. and Oka, N.: Experiments Toward a Mutual Adaptive Speech Interface That Adopts the Cognitive Features Humans Use for Communication and Induces and Exploits Users' Adaptations, *International Journal of Human-Computer Interaction*, Vol. 18, No. 3, pp. 243–268 (2005).
- [2] 田中一晶, 左祥, 嵯峨野泰明, 荒木雅弘, 岡夏樹: No News 規準が有効な条件: 誘導教示の意味学習場面での実験的検討, 電子情報通信学会論文誌 (A), Vol. J92-A, No. 11, pp. 784–794 (2009).
- [3] Winston, P.: Learning structural description from examples, *Psychology of Computer Vision* (1975).
- [4] 田中一晶, 尾関基行, 荒木雅弘, 岡夏樹: ロボットへの教示場面における「間」の重要性: ロボットの行動の遅れは学習効率を向上させ教えやすい印象を与える, 人工知能学会論文誌, Vol. 25, No. 6, pp. 703–711 (2010).
- [5] Carlson, R., Gustafson, K. and Strangert, E.: Cues for Hesitation in Speech Synthesis, in *In Proceedings of Interspeech 06* (2006).
- [6] Komatsu, T. and Nagasaki, Y.: Can We Estimate The Speaker's Emotional State from Her/His Prosodic Features? -Effects of F0 Contour's Slope and Duration on Perceiving Disagreement, Hesitation, Agreement, and Attention, *Proceedings of the 18th International Congress on Acoustics*, pp. 2227–2230 (2004).
- [7] 河原達也, 李晃伸: 連続音声認識ソフトウェア Julius, 人工知能学会誌, Vol. 20, No. 1, pp. 41–49 (2005).
- [8] Tanaka, K. and Oka, N.: A robot that learns in stages utilizing scaffolds: toward an active and long-term human-robot interaction, *International Conference on Advances in Computer Entertainment Technology (ACE2008)*, Vol. 352, pp. 195–198 (2008).
- [9] Breazeal, C. and Aryananda, L.: Recognition of Affective Communicative Intent in Robot-Directed Speech, *Autonomous Robots*, Vol. 12, No. 1, pp. 83–104 (2002).
- [10] 藤江真也, 江尻康, 菊池英明, 小林哲則: 肯定的/否定的発話態度の認識とその音声対話システムへの応用, 電子情報通信学会論文誌, Vol. J88-D-II, No. 3, pp. 489–498 (2005).
- [11] 光本浩士, 濱崎敏幸, 大多和寛, 田村進一, 柳田益造: 終助詞「ね」の韻律による皮肉と賞賛の識別, 電子情報通信学会論文誌, Vol. J84-D-II, No. 5, pp. 851–853 (2001).
- [12] 岡夏樹, 増子雄哉, 林口円, 伊丹英樹, 川上茂雄: Fisher の直接法を用いたインタラクションデータからの意味学習, 知能と情報, Vol. 20, No. 4, pp. 461–472 (2008).
- [13] Sutton, R. S. and Barto, A. G.: Reinforcement Learning: An Introduction, *The MIT Press* (1998).