

音声対話におけるエージェント発話行動の適応的調整

User-adaptive coordination of agent communicative behavior in spoken dialogue

堂坂浩二^{*1,*2} 金本淳志^{*2} 東中竜一郎^{*1} 南泰浩^{*1} 前田英作^{*1,*2}
 Kohji Dohsaka Atsushi Kanemoto Ryuichiro Higashinaka Yasuhiro Minami Eisaku Maeda

^{*1} 日本電信電話株式会社
NTT Corporation

^{*2} 大阪大学大学院情報科学研究科 連携講座
Graduate School of Information Science and Technology, Osaka University

In this paper, aiming at achieving smooth spoken interaction between human users and dialogue agents, we present an experimental study that evaluates a method for user-adaptive coordination of agent communicative behavior. The method adapts the pause length of agent utterances and the duration of agent gaze to minimize the discomfort that individual users perceived during the interaction. The experimental results showed that there was a significant tendency that the method made the agent's pause length and gaze duration converge during the interaction, and that the method significantly improved the perceived smoothness of the agent communicative behavior.

1. はじめに

我々は、人間・対話エージェント間の自然なインタラクションを通して、エージェントが人間に適切に働きかけることにより、発想・学習の支援、コミュニケーションの活発化等を実現することを狙いとして研究を進めている。これまで、対話を通してユーザの自発的な思考を喚起するという観点から、思考喚起型対話に着目し、その一例としてクイズ型対話を提案してきた [Higashinaka 07a, Higashinaka 07b, Dohsaka 09]。この対話では、システムが人物等の名前当てクイズを出題し、ヒント文を順に提示する。ヒント文は、ユーザが正解に容易に辿りつかないように、難しい順に自動的に並び替えられる [Higashinaka 07b]。結果として、ユーザが自ら正解を見つけようとする思考が促進され、ユーザの対話参加意欲が向上する [Higashinaka 07a]。また、クイズ型多人数対話においてエージェントの共感的発言がコミュニケーションを活発化させることが分かっている [Dohsaka 09]。

しかし、そうした思考喚起やコミュニケーション活発化も、ユーザ・エージェント間の円滑なコミュニケーションが実現できなければ、十分にその効果を発揮することは難しい。そこで、本研究では、ユーザ・エージェント間の円滑な音声対話の実現を狙いとして、エージェント発話行動を個々のユーザに対して適応的に調整することにより、ユーザが対話中で覚える違和感を減少させることに焦点をあてる。

ここで、発話行動とはエージェントが対話中に行う発話や、それに付随する顔向き・ジェスチャーなどの行動のことを言う。円滑なコミュニケーションを実現するためには、エージェント発話行動を個々のユーザに適応的に調整することが重要と考えられる [光永 06]。発話行動には音声的要素と非音声的要素があるが、本研究では、音声的要素として発話タイミング、非音声的要素として顔向きによる注視を取り上げる。

音声的要素としては、発話タイミング、発話表現・内容など様々なものが考えられるが、発話タイミングの取り方が適切でなければ、発話表現・内容がいかにも適切であっても、コミュニケーションが成立しない場合があることから、発話タイミングを適切にすることが第一の解決課題であると考えた。ここでは、特に、エージェント発話開始前のポーズ(発話の間)に着目する。発話タ

イミングに関する従来研究としては、音声対話における発話タイミングの影響に関する知覚実験がある [伊藤 08]。この実験では合成音声による機械同士の対話を人間が第三者として観察することで、発話タイミングの自然性や重要性について調べている。これに対して、本研究では物理エージェントとユーザとの間の音声対話に注目する。また、円滑な音声対話の実現を目的として、明滅光源を用いたエージェント内部状態表出手法が提案されている [船越 08]。この手法は、発話準備中・聞き取り失敗などエージェント内部状態を明滅光源により表出し、それに合わせてユーザの対話の進め方を適応させる。これに対し、本研究ではエージェント発話行動をユーザに対し適応的に調整していく。

非音声的要素としては、顔向き、視線、ジェスチャーなどがあるが、中でも顔向きによる注視は、簡易な物理エージェントで実現できる機能であり、人間同士のコミュニケーションにおいては、相手を見つめる時間の割合がその人物の印象に影響を与えることが知られている [Argyle 76]。本研究でも、エージェントがユーザを注視する時間長に着目する。

近年、対話エージェントの行動戦略を対話データから強化学習により獲得することを目指した研究が、タスク達成型対話 [William 07]のみならず、思考喚起型対話 [Minami 09]においても進められている。強化学習は、エージェントがユーザとの対話を通して報酬(ペナルティ)を受け取り、報酬を最大化(ペナルティを最小化)するように、最適な行動戦略を獲得するためのものであり、エージェント行動をユーザに個人適応することにも利用することができる。実際、光永等は方策勾配型強化学習に基づくロボット行動の個人適応方法を示している [光永 06]。しかし、そこでは、ロボットと人間が固定的なシナリオに基づくインタラクションを行う状況が想定されており、いつでも話者交代が可能なより柔軟なインタラクションにおいて、エージェント発話行動の個人適応が可能であるかは明らかではない。本論文では、柔軟な話者交代が可能な音声対話において、方策勾配型強化学習に基づく適応手法によって、ユーザの違和感を減少させるように、エージェント発話行動をユーザに個人適応することが可能かどうか評価した実験の結果について報告する。

2. 思考喚起型対話システム

本研究では、評価実験のために、思考喚起型対話の一例としてクイズ型対話に注目し、クイズ型音声対話システム [Minami 07] を多人数対話環境で利用できるように拡張したものをを用い



図 1: ユーザ 1 名とエージェント 2 体による対話の風景

た. 対話にはユーザと 1 体あるいは 2 体の対話エージェント(小型のロボット)が参加する. 2 体のエージェントは, クイズのヒントを順に提示する出題エージェントと, ユーザと同じ立場で対話に参加し回答を行う回答エージェントである. ユーザの発言は音声によって入力され, エージェントの発言はスピーカにより出力される. エージェントは, 顔向きにより他の参加者の顔を見るという動作を行うことができる. 図 1 にユーザとエージェント 2 体との間の対話の風景を示す.

以下はユーザと出題エージェント(出題 A)と回答エージェント(回答 A)の間の対話の例である.

- (1) 出題 A: 第 1 問, なぞなぞスタート.
最初のヒント, 戦国大名だよ. (ヒント, 自発)
- (2) ユーザ: 織田信長かな? (回答, 自発)
- (3) 出題 A: 惜しい. もう少し後の時代だよ. (評価, 応答)
- (4) 回答 A: 難しくてわからないな. (難度表出, 自発)
- (5) 出題 A: そろそろ次のヒント出すよ.
第 2 ヒント, 征夷大將軍だよ. (ヒント, 自発)
- (6) ユーザ: うーん, 誰だろう. (難度表出, 自発)
- (7) 回答 A: 難しくてわからないよね. (難度表出, 応答)
- (8) 回答 A: 徳川家光(回答, 自発)
- (9) 出題 A: 惜しい. 征夷大將軍というのはあってるよ.
(評価, 応答)
- (10) ユーザ: 徳川家康! (回答, 自発)
- (11) 出題 A: あたり! 正解は"徳川家康"だね.
すごい! よくわかったね. (評価, 応答)
- (12) 回答 A: 悔しい. 次は正解できるように頑張る!
(感想, 自発)

エージェントの発話としては, エージェントが自ら発話のターンを取得して行う自発的発話と, 他の参加者の発言に対して応答的に行う応答的発話がある. 出題エージェントの自発的発話としては, ヒント文の提示(1,5), 難度表出(「難しいかな」), 他の参加者に発言を促す発話等がある. 応答的発話としては, 他の参加者の回答を評価する発話(3, 9, 11)等がある. 回答エージェントは, 自発的発話として, 回答発話(8), 他の参加者に発言を促す発話, 自ら問題の難易度を表出する発話(4), 正解時の感想の表出(12)等を行う. 応答的発話としては, ユーザに対して行う共感的発話(7)等がある.

3. エージェント発話行動の適応的調整手法

個々のユーザとインタラクションしながら, その場でエージェント発話行動を適応させるためには, できるだけ短時間で発話行動を調整する必要がある. そこで, 本研究では, 適応的調整手法として, 少ないデータから短時間で局所的に最適な行動戦略を求めることが可能な方策勾配型強化学習を用いる [Kohl 04]. 図 2 にエージェント発話行動の適応的調整手法を示す. この手

- (1) $\theta \leftarrow$ 方策パラメタのベクトル(長さ n)
- (2) $\varepsilon \leftarrow$ パラメタごとの微小変動幅のベクトル(長さ n)
- (3) $\eta \leftarrow$ 更新幅
- (4) $\max \leftarrow 0$ (これまでの報酬の絶対値の最大値を保持)
- (5) while (対話続行中)
- (6) for $i = 1$ to T
- (7) for $j = 1$ to n
- (8) $r_j \leftarrow$ $\{-1, 0, +1\}$ からランダムに選択した数
- (9) $R_j^i \leftarrow \theta_j + \varepsilon_j * r_j$
(R^i は各パラメタを微小変動させた方策)
- (10) for $i = 1$ to T
- (11) 方策 R^i に基づいて 1 ヒント対話を実施し, 報酬を獲得
- (12) for $j = 1$ to n
- (13) $\text{Avg}_{+\varepsilon_j} \leftarrow \theta_j$ に $+\varepsilon_j$ を加えて生成されたすべての方策の報酬の平均値
- (14) $\text{Avg}_{0,j} \leftarrow \theta_j$ に微小変動を加えなかった場合のすべての方策の報酬の平均値
- (15) $\text{Avg}_{-\varepsilon_j} \leftarrow \theta_j$ に $-\varepsilon_j$ を加えて生成されたすべての方策の報酬の平均値
- (16) if ($\text{Avg}_{0,j} > \text{Avg}_{+\varepsilon_j}$) 且つ ($\text{Avg}_{0,j} > \text{Avg}_{-\varepsilon_j}$)
- (17) $a_j \leftarrow 0$
- (18) else
- (19) $a_j \leftarrow \text{Avg}_{+\varepsilon_j} - \text{Avg}_{-\varepsilon_j}$
- (20) $A \leftarrow \frac{A}{|A|} \times \eta$
- (21) $\forall j, a_j \leftarrow a_j \times \varepsilon_j$
- (22) $\text{maxC} \leftarrow$ このサイクルでの報酬の絶対値の最大値
- (23) if $\text{maxC} > \max$
- (24) $\max \leftarrow \text{maxC}$ (\max を更新)
- (25) else
- (26) $A \leftarrow A \times \frac{\text{maxC}}{\max}$
- (27) $\theta \leftarrow \theta + A$

図 2: エージェント発話行動の適応的調整手法

法は, 報酬の絶対値の大きさにより勾配を調整するように, Kohl 等のアルゴリズム [Kohl 04] を改良したものである.

エージェント発話行動は方策と呼ばれる行動戦略に基づいて決まり, この方策をユーザに適応させていく. 方策は方策パラメタのベクトル $\theta (= \{\theta_j\})$ として表される. 本実験では次の 4 つの方策パラメタを用いる ($n=4$).

- 自発的発話の間 (θ_1): 直前の発話から, エージェントが自発的発話を開始するまでの時間
- 応答的発話の間 (θ_2): エージェントが, 他の参加者の発話の後, それに対して応答的発話を開始するまでの時間
- 注視時間 (θ_3): エージェントが, 自らが発話を行うとき, 或いは, 相手の発話を聴いているとき, 顔向きにより相手を注視する時間の長さ
- ヒント提示間隔 (θ_4): クイズ対話においてヒントを提示する時間間隔

本実験では, 2 種類の発話の間と注視時間に加えて, クイズ型対話に特有のパラメタとしてヒント提示間隔も考慮に入れた.

エージェントの方策を適応させるために, まず, 現在の方策を微小変動させた T 通りの方策 R^i を用意する(図 2 の 7~9). R^i は, 方策パラメタ θ_j ごとに決まる微小変動幅 ε_j が与えられるとき, 各パラメタに $+\varepsilon_j, 0, -\varepsilon_j$ のいずれかをランダムに加えることにより生成される. 本実験では T は 10 とした.

	自発的 発話の 間 (sec)	応答的 発話の 間 (sec)	注視 時間 (sec)	ヒント 間隔 (sec)
初期値	4.96	0.53	3.04	27.7
微小変動幅	0.50	0.20	0.30	2.5

表 1: 方策パラメタの初期値と微小変動幅

次に、各方策 R^i に基づいて1ヒントあたりの対話(ヒント対話)を10通り(T通り)試行し、各ヒント対話ごとに報酬を獲得する(10~11)。1ヒント対話中に行われるエージェント発話行動は同一の方策パラメタに基づく。第4節で説明するように、報酬としては、1ヒント対話中にユーザが知覚する違和感の大きさを負の報酬(ペナルティ)として与える。10通りの方策すべてに関してヒント対話を実施した後、各ヒント対話において獲得した報酬に基づいて、方策ベクトルの勾配 $A = \{a_j\}$ を求める(12~19)。

勾配は正規化し、更新幅 η と、各成分の重み ϵ_j を掛ける(20~21)。更新幅は適応の速度を調整するための変数で、本実験では1.0とした。次に、現在の適応サイクルにおいて報酬の絶対値の最大値(maxC)を求め、これまでの適応過程において経験した報酬の絶対値の最大値(max)との比によって勾配の大きさを調整する(22~26)。

この10回のヒント対話の試行と引き続くパラメタ更新が適応の1サイクルとなる。これを繰り返すことで、個々のユーザの違和感を減少させるように方策パラメタを個人適応させていく。

4. 実験

実験では第2節で述べたクイズ型音声対話システムを用いた。実験には32名の参加者(ユーザ)が参加した(男性16名、女性16名)。うち16名がユーザ1名と出題エージェントの間の2人対話を実施し、残り16名がユーザ1名と出題エージェントと回答エージェントの間の3人対話を実施した。2人対話、3人対話いずれのグループも男女は同数であった。

クイズ型対話において問題として出題する人物は事前に決めておき、5人物を1セットとした問題セットごとに問題を出題し、各問題セットに含まれる人物のPageRankTMスコアをほぼ等しくすることにより、問題セットごとのクイズの難易度を揃えた。問題提示の順序効果を相殺するため、問題セットの提示順序はユーザごとにランダムに変更した。クイズのヒント文はWikipediaから抽出し、正解することが難しい順に並べ替えた[Higashinaka 07b]。

各ユーザは、システムの使い方に慣れるため、最初に1問題セットの練習対話を行った。練習対話終了後、評価実験のための対話を1問題セットごとに小休憩を挟みながら行った。対話は、出題人物数に関わらず、ユーザが150個のヒントを受け取るまで続けた。対話は1ユーザあたり約2時間続いた。第3節で述べた適応手法に従って、対話を通してエージェント発話行動の方策パラメタを調整していった。ヒント対話10回ごとにパラメタを更新し、対話全体で15回のパラメタ更新を行った。

方策パラメタの更新は、1ヒント対話中にユーザが知覚した違和感の大きさに基づいて行った。本実験では、ユーザがエージェント発話行動に違和感を覚えたとき、手に持ったボタンにより直接知らせるという方式をとった。ユーザにはエージェントの発話の間と注視時間に関する違和感にのみ着目するように指示した。ボタンには「3」、「2」、「1」という数字が書かれており、ユーザは違和感のレベルを3段階で評価して、該当するボタン(「3」、「2」、「1」)を押す。ユーザが、1ヒント対話中に押したボタンの合計値を1ヒント対話の時間長で正規化した値を負の報酬(ペナルティ)とした。ユーザの振る舞いや対話の状況からユーザの違和感を自動推定することは今後の課題である。

		自発的 発話の 間 (sec)	応答的 発話の 間 (sec)	注視 時間 (sec)	ヒント 間隔 (sec)
2 人 対 話	平均値	5.04	0.62	3.09	25.9
	最小値	3.90	0.39	2.40	19.5
	最大値	6.17	1.18	3.69	31.2
	標準偏差	0.72	0.21	0.36	2.73
3 人 対 話	平均値	4.86	0.62	3.15	27.4
	最小値	4.07	0.35	2.52	22.0
	最大値	5.54	0.90	3.58	32.7
	標準偏差	0.43	0.18	0.27	2.48

表 2: 方策パラメタの最終値の統計量

		自発的 発話の 間	応答的 発話の 間	注視 時間	ヒント 間隔
2 人 対 話	初期相対誤差	0.030	0.078	0.024	0.021
	終期相対誤差	0.015	0.031	0.013	0.021
	t値(df=15)	2.3	3.2	2.2	0.070
	p値	0.036	0.0066	0.041	0.95
3 人 対 話	初期相対誤差	0.023	0.12	0.026	0.022
	終期相対誤差	0.010	0.032	0.012	0.009
	t値(df=15)	2.3	4.5	2.7	2.9
	p値	0.038	< 0.001	0.016	0.011

表 3: 方策パラメタ更新における初期相対誤差と終期相対誤差の平均値、及び、統計検定結果

実験で用いた方策パラメタの初期値と微小変動幅は表1の通りである。パラメタ初期値は予備実験によって定めた。微小変動幅は初期値の1/10程度を目安とした。ただし、応答的発話の間の微小変動幅は人間の知覚の特性を考慮して、200 msecとした。

実験開始前(練習対話後)と実験終了後に、エージェントの発話の間と注視時間の適切さについて、ユーザの主観評定を尋ねるアンケートを実施した。アンケートの設問は次の通りであり、それぞれ7段階のリッカート尺度を用いた。

- エージェントの発言のタイミングは適切だと感じましたか？
- エージェントがあなたを見る時間の長さは適切だと感じましたか？

5. 結果

5.1 方策パラメタ値の推移と収束性

各ユーザとの対話における方策パラメタの最終値の統計量を表2に示す。パラメタの初期値は予備実験によって定めたため、パラメタ最終値の平均値は初期値から大きく変化していないが、最大値、最小値、標準偏差をみると、パラメタ最終値は個々のユーザによってばらばらな値になっていることが分かる。

適応的調整手法が適切に働いているのであれば、方策パラメタは対話進行に従って収束の傾向を見せるはずである。そこで、パラメタの収束性を分析するために、パラメタ更新の相対誤差の推移について調べた。パラメタ更新の相対誤差は、更新 $k-1$ 回目のパラメタ値 p_{k-1} と更新 k 回目のパラメタ値 p_k が与えられるとき、 $|p_k - p_{k-1}|/p_{k-1}$ と定義される。

各ユーザとの対話で、最初3回のパラメタ更新における相対誤差(初期相対誤差)の平均値と、最後3回の更新における相対誤差(終期相対誤差)の平均値を比較した。表3は、各パラメ

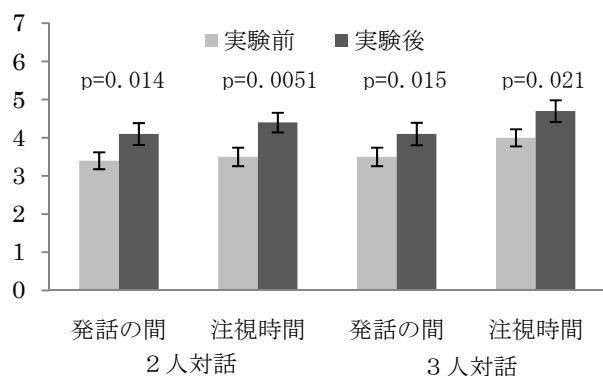


図 3: 実験前後でのエージェントの発話の間と注視時間の適切さに対するユーザ評定の平均値・標準誤差・統計検定結果

タの初期相対誤差と終期相対誤差の平均値と、2つの平均値の差の統計検定結果を示している。最初3回の更新よりも、最後3回の更新の方が、パラメタ更新の相対誤差が小さくなる傾向がみとれる。各群は正規性を仮定することができた。対応ありのウェルチのt検定により、ユーザごとの初期相対誤差と終期相対誤差の平均値の差を検定した。その結果、表3に示す通り、2人対話のヒント提示間隔を除いて、初期相対誤差と終期相対誤差の平均値の間に有意な差がみられた。このように、発話の間と注視時間に関しては、インタラクションを通して方策パラメタは収束する傾向があることが分かった。ヒント提示間隔が収束しづらい理由としては、ヒント提示間隔は、発話の間や注視時間に比べて、長い時間間隔であるため、様々な要因に影響を受けやすいことや、ヒント提示間隔の適切さはユーザの知識の量やクイズに対する関心の程度に依存しやすいことが考えられる。

5.2 ユーザ主観評定

次にアンケートによるユーザ主観評定の結果について述べる。図3は、実験開始前と終了後において、エージェントの発話の間と注視時間の適切さに関してユーザが評定した結果を示している。この図から分かる通り、2人対話、3人対話ともに、ユーザが知覚するエージェントの発話の間と注視時間の適切さは、実験前よりも実験後の方が高まっている。

アンケートによる評定値は順序尺度であるので、ノンパラメトリックな検定法としてウィルコクソンの符号付順位検定を用いて、実験開始前と終了後の評定値を比較した。各群は等分散性を仮定することができた。検定の結果、図3に示すように、2人対話と3人対話の双方において、発話の間と注視時間の適切さに関して、実験前後でユーザ評定値に有意な差があることが分かった。

以上に述べたパラメタの収束性、及び、ユーザ主観評定の分析結果は、柔軟な話者交代が可能な音声対話において、方策勾配型強化学習に基づく適応手法が、2時間程度のインタラクションを通して、エージェントの発話の間と注視時間を個々のユーザに合わせて適応的に調整することが可能であることを示している。

6. おわりに

人間と対話エージェントの間の円滑な音声対話を実現することを狙いとして、エージェント発話行動を個々のユーザに適応的に調整する手法を評価した結果について報告した。適応的調整手法として、ユーザが対話中に知覚する違和感を減少させるように、方策勾配型強化学習に基づいてエージェントの発話

の間と注視時間を段階的に調整するという方法を用いた。ユーザが違和感をボタンによって知らせるという対話環境において、手法を評価するための実験を行った。実験の結果、2時間程度のインタラクションを通して、エージェントの発話の間と注視時間が収束する有意な傾向があり、さらに、ユーザが知覚するエージェントの発話の間と注視時間の適切さが、適応の結果、有意に向上することが分かった。これらの結果は、柔軟な話者交代が可能な音声対話において、方策勾配型強化学習により、エージェント発話行動を個々のユーザに適応させることが可能であることを示している。今後の課題としては、ユーザ違和感の自動推定、対話の状態に応じたエージェント行動の適応、基本的な発話行動だけでなく、ユーザの関心レベルに応じた対話の話題の適応などがある。

謝辞: 本研究の一部は、科研費(新学術領域)「人とロボットの共生による協創社会の創成」における計画研究「ロボットのコミュニケーション戦略の生成」(21118004)の助成を受けたものである。

参考文献

- [Argyle 76] Argyle, M. and Cook, M.: Gaze and Mutual Gaze, Cambridge University Press, 1976.
- [Dohsaka 09] Dohsaka, K., Asai, R., Higashinaka, R., Minami, Y., and Maeda, E.: Effects of conversational agents on human communication in thought-evoking multi-party dialogues, Proc. SIGdial 2009, pp.217-224, 2009.
- [船越 08] 船越孝太郎, 小林一樹, 中野幹生, 山田誠二, 北村泰彦, 辻野広司: 明滅光源を用いた内部状態表出による音声コミュニケーションの円滑化, 第22回人工知能学会全国大会論文集, 1D2-9, 2008.
- [Higashinaka 07a] Higashinaka, R., Dohsaka, K., Amano, S., and Isozaki, H.: Effects of quiz-style information presentation on user understand, Proc. Interspeech 2007, pp.2725-2728, 2007.
- [Higashinaka 07b] Higashinaka, R., Dohsaka, K., and Isozaki, H.: Learning to rank definitions to generate quizzes for interactive information presentation, Proc. ACL 2007 (Poster Presentation), pp.117-120, 2007.
- [伊藤 08] 伊藤敏彦, 北岡教英, 西村良太: 音声対話における発話タイミングの影響に関する知覚実験, 情報処理学会研究報告 2008-SLP-72-18, pp.99-104, 2008.
- [Kohl 04] Kohl, N. and Stone: Policy gradient reinforcement learning for fast quadrupedal locomotion, Proc. IEEE International Conference on Robotics and Automation, Vol.3, pp.2619-2624, 2004.
- [Minami 07] Minami, Y. et al.: The World of Mushrooms: human-computer interaction prototype systems for Ambient Intelligence, In: Proc. ICMI 2007, pp. 366-373, 2007.
- [Minami 09] Minami, Y., Mori, A., Meguro, T., Higashinaka, R., Dohsaka, K., and Maeda, E.: Dialogue control algorithm for Ambient Intelligence based on partially observable Markov decision processes, Proc. IWSDS2009, 2009.
- [光永 06] 光永法明, クリスチャンズミス, 神田崇行, 石黒浩, 萩田紀博: 方策勾配型強化学習によるロボットの対人行動の個人適応, 日本ロボット学会誌, Vol. 24, No.7, pp. 820-829, 2006.
- [Williams 07] Williams, J. D. and Young, S.: Partially Observable Markov Decision Processes for Spoken Dialog Systems, Computer Speech & Language, Vol. 21, No. 2, pp. 393 - 422, 2007.