

説得対話システムにおける話題誘導に基づく対話制御

平岡 拓也¹ Graham Neubig¹ Sakriani Sakti¹ 戸田 智基¹ 中村 哲¹

概要: 本研究では、システムが特定の目標を持って対話を進め、最終的にシステム側が意図した行動をユーザに取らせることを目的とする説得対話システムの構築に取り組む。本報告では、説得対話システムの対話制御部に対して、対話状態の確率的なモデル化とそれに基づくシステム応答決定処理を適用する。ベースラインモデルとして、嗜好によって意思決定を行うユーザモデルを仮定し、対話モデルを作成する。そして、話題誘導、説得に向けた対話戦略、興味推定といった説得対話に特化した手法を導入する。小規模な説得対話タスクにおいてシステムを構築し、その性能を実験的に評価する。実験結果から、各手法を導入するほど説得成功率及び満足度が高くなる傾向が得られることを示す。

キーワード: 音声対話システム, 説得対話, 対話制御, 話題誘導, 興味推定

Dialog management based on guiding users to new topics in persuasive dialogue systems

Abstract: In this research, we build a persuasive dialogue system that interacts with the user based on a specific goal and leads the user to take the actions that the system intends. In this presentation, we discuss the probabilistic dialog modeling aspect of this dialog system. As a baseline system, we develop a dialogue model assuming the user makes decisions based on preference. Then we introduce a model that can guide the user from topic to topic and discuss a method for estimating the user's interest in this model. We build and evaluate the system in a small task, with experimental results indicating that the proposed method are effective in this content.

Keywords: Spoken dialogue system, Persuasive dialogue, Dialogue management, Guiding users to new topics, Estimating user's interest

1. はじめに

人と機械の円滑なコミュニケーションを実現するために、対話システムに関する研究が盛んに行われている。一般的な対話システムは、言語理解部、対話制御部、言語生成部で構成される。言語理解部では、ユーザが入力したテキストから、概念を抽出し、システムが取り扱うことの出来る形式で出力する。対話制御部は、言語理解部の出力結果を基に、システムが返す応答の概念を決定する。言語生成部では、対話制御部から出力された概念に基づき、テキストを生成する。なお、テキストではなく音声を用いて対話を行う音声対話システムでは、さらに、言語理解部の前段に音声認識部と言語生成部の後段に音声合成部が必要と

なる。ここで、対話システムの応答を決定する対話制御部は、対話システムの核となる部分であり、対話の成否に大きく影響を与える。また、優れた対話制御部を構築することで、仮に音声認識部や言語理解部に誤りがあったとしても、その誤りを対話を通して修復することも可能である。

対話制御手法として、これまでに様々な枠組みが提案されている。有限状態オートマトン (Finite State Automaton: FSA) に基づく対話制御法では、ユーザからの入力に対するシステムの応答規則を、FSA の形式で明示的に記述する。対話状態を考慮した対話制御が可能となる一方で、人手により応答規則を決定する必要がある。これに対し、マルコフ決定過程 (Markov Decision Process: MDP) に基づく対話制御法 [1] では、対話状態に対するシステムの応答規則を強化学習を用いて最適化することが可能である。一方で、対話状態を決定的に求める FSA や MDP に基づく対話

¹ 奈良先端科学技術大学院大学
Nara Institute of Science and Technology

制御法では、音声認識部や言語理解部の誤りにより生じる不確定性に対応することができない。この問題に対して、対話状態の不確定性を考慮するために、ベイジアンネットワーク (Bayesian Network: BN) を用いて、対話状態を信念状態として確率的にモデル化し、信念状態に基づいた応答規則を人手で記述する枠組みも提案されている [2], [3]。さらに、これらの手法を組み合わせたものとして、部分観測マルコフ決定過程 (Partially Observable MDP: POMDP) に基づく対話制御法が提案されている [4]。POMDP に基づく対話制御法では、対話状態を信念状態として確率的にモデル化することで不確定性を考慮しつつ、強化学習によるシステム応答規則の最適化を行うことが可能である。上記の手法とは別に、対話データに基づいて、より簡易に対話制御部を構築する手法として、n-gram による対話履歴表現と重み付き有限状態トランスデューサー (Weighted Finite State Transducer: WFST) を用いた対話制御法も提案されており [5]、POMDP に基づく対話制御法との関係についても議論されている [6]。

対話制御部の発展に伴い、対話システムで取り扱う対話のタスクも広がりを見せている。従来の対話システムでよく取り扱われる対話のタスクの一つとして、ユーザが明確な目標を持ち、その目標に沿った対話を行うものがある。例として、航空情報システム [7] が挙げられ、システムは対話を通してユーザの目標を正確に推定する必要がある。また、別のタスクとして、ユーザ側の目標は明確ではなく、システムとユーザが相談することで目標を決定する対話を行うものもある。この種の意味決定を支援する対話システムの例として、観光案内システム [8] が挙げられる。この他にも、明確な目標が存在しない対話もある。例として、カウンセラーを模倣した Eliza [9] に代表されるようなチャットボットや、一問一答形式で雑談を行う質問応答対話システム [10] などが挙げられる。人の対話は実に多様であり様々なタスクがあるため、今後、対話システムの応用分野を広げていく上で、システムが取り扱える対話のタスクを多様化していくことが望まれる。

本研究では、対話システムの新たな応用例として、説得対話に着目する。説得対話では、システムが予め定めた目標に対して、ユーザを誘導していき、最終的にシステムが意図した通りの意思決定をユーザにとらせることを目的とする。説得対話システムは、ユーザ側の目標が明確ではなく、ユーザの嗜好を推測し、それを手がかりに対話を進めるといった点において、意思決定支援を行う対話システムと類似している。一方で、意思決定支援用の対話システムでは、システムがユーザに対してどの意思決定をとらせるかについては対話中に決定するのに対し、説得対話システムでは、予めシステム側が意図した通りの意思決定をユーザにとらせるという点で大きく異なる。代表的な説得対話の例としては、特定の集団へと勧誘する際の対話や商品を販

売する際の対話などが挙げられる。

本報告では、説得対話システムのための対話制御部の構築に取り組む。取り扱う説得対話として、ある程度ユーザの嗜好に合致する範囲内において、システムの目標を達成する対話に着目する。対話状態の POMDP による確率的なモデル化とそれに基づくシステム応答決定処理 [4] を用いて、対話制御部を構築する。また、説得対話に特化した手法として、話題誘導、ユーザに提示する情報を制御する対話戦略、対話中に動的に変化する興味推定を提案し、対話制御部に導入する。小規模な説得対話タスクとして、研究室勧誘を対象とするシステムを構築し、その性能を実験的に評価する。実験結果から、提案法により、タスク成功率、ユーザ満足度、自然性の向上がもたらされることを示す。

以下、2 節で説得対話システムの概要について述べる。3 節では提案する説得対話制御モデルについて述べる。4 節で実験的評価について述べ、5 節で結びとする。

2. 説得対話システム

本研究で扱う説得対話タスクとして、所属研究室を選定する学生に対して、特定の研究室へと勧誘する説得対話を対象とする。本研究で構築したシステムは大きく分けて 3 つの機能をユーザに提供する。

機能 1 各研究室が取り組んでいる研究分野についての情報をユーザに提示する。

機能 2 各研究分野間の関係についての情報をユーザに提示する。

機能 3 特定の研究室への勧誘を行う。

システムが A 研究室への勧誘を目的とした説得対話の一例を表 1 に示す。表 1 において機能 1 に該当するものとして、S2 「対話は A 研究室で取り組まれています」と S4-1 「音声翻訳は A 研究室で取り組まれています」が挙げられる。また、機能 2 に該当するものとして、S3-1 「音声翻訳はコミュニケーション支援につながります」が挙げられる。そして、機能 3 に該当するものとして、S4-2 「A 研究室に入ってはいかがですか」が挙げられる。

システムは、図 1 に示すように、各研究室が取り組んでいる研究分野と研究分野間の関係についての知識を保持しており、ユーザからの応答に応じ、これらの情報を提示しながら、対話を進める。例えば、図中の「実現手段」と「研究対象」は、ある研究分野を実現する上での一手段になるという関係と、各研究室が研究の対象としているという関係を、各々表わしている。システムは、各関係を説明するテキストを同様に知識として保持しており、ユーザへの情報提示や勧誘の際に利用する。

3. 説得対話制御モデル

ユーザを説得する上で、ユーザが研究室を選択する際に重視する決定要因 (研究分野) の推定する必要がある。ま

た、現在の話題から勧誘目標とする研究室に関する話題へと対話を誘導する機能や、説得を行うために有利となるような情報提示機能、対話中において動的に変化するユーザの興味に応じた話題の展開機能が、より自然な説得対話を実現する上で有効であると予想される。そこで、これらの機能を備えた対話制御部を提案する。

3.1 ベースラインモデル

ベースラインモデルでは、意思決定支援に基づく観光案内システム [8] と同様に、嗜好に基づく説得を行う。ここで、嗜好とは、ユーザが特定の代替案を選択する際に重視する決定要因であり、どの決定要因がユーザの嗜好に合致するかは対話を通して不変とする。なお、代替案とは、ユーザの意思決定の候補を表わし、本研究では研究室を表わす。また、決定要因とは、代替案を選択する際の評価基準を表わし、本研究では研究室が取り組んでいる研究分野を表わす。

本研究における対話モデルにおいても、ユーザは自らの嗜好に合致した研究分野を扱っている研究室を選択すると仮定する。システムが勧誘目標とする研究室に関連する研究分野（以後、説得目標）に対するユーザの嗜好を利用して、最終的な勧誘を行うことでユーザを説得する。ユーザ

表 1 想定する説得対話の一例。S はシステム応答発話を表し、U はユーザ応答発話を表す。システムは A 研究室への勧誘を行う。

発話文
S1: どんな研究分野に興味がありますか
U1: 対話やコミュニケーションに興味があります
S2: 対話は A 研究室で取り組まれています
U2: でもコミュニケーションはしてないんですね
S3-1: コミュニケーションに興味があるんですね
S3-2: 音声翻訳はコミュニケーション支援につながります
S3-3: 音声翻訳に興味がありますか
U3: はい、あります
S4-1: 音声翻訳は A 研究室で取り組まれています
S4-2: A 研究室に入っているはいかがですか
U4: 入ります

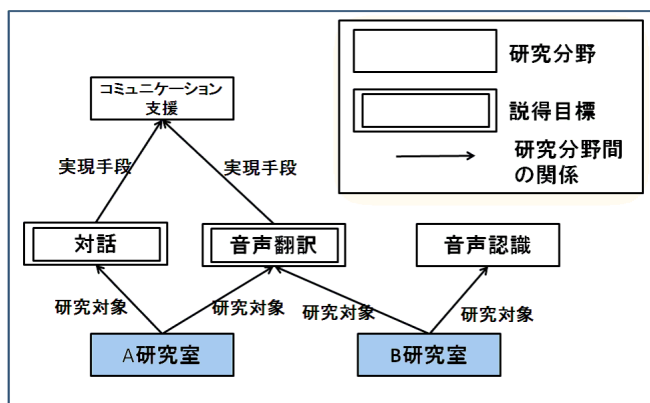


図 1 システムが保持する知識の一例

の嗜好に合う研究分野はシステムにとって未知であるため、対話を通して推定する必要がある。そこで、対話状態を次式にてモデル化する。

$$P(e_k|A, O) = \frac{P(O|A, e_k) \sum_{e'_k} P(e_k|e'_k)P(e'_k)}{P(O|A)} \quad (1)$$

ここで、 e_k は研究分野 k に対するユーザの嗜好を表し、 A はシステムの直前の発話行為、 O はユーザの発話行為を表す。また e'_k はユーザとシステムの直前のターンにおける研究分野 k に対するユーザの嗜好を表わす。対話状態に対するベイジアンネットワークによる記述を図 2 に示す。対話開始時点から現時点までの対話履歴は、マルコフ過程に基づき、対話状態（信念空間）として表現され、現時点における対話状態に基づき、システムの発話行為を決定する。具体的には下記に挙げる対話戦略に基づいて発話行為を決定する。

- ユーザから特定の研究分野や研究について質問が無く、システムが現在の話題から別の話題に切り換える際、最も嗜好に合致すると推定した話題を新たに提示する。
- 勧誘を行う際に、勧誘する研究室が取り組んでいる研究分野の中でユーザの嗜好が高いものを順に三つ取り上げる。

3.2 話題誘導

説得対話においては、現状の話題から勧誘目標及び説得目標に関する話題へと対話を誘導することで、より勧誘目標や説得目標についての情報提示を行うことが出来ると考えられる。これにより、説得時に有効的に利用できるユーザの嗜好を、より効率的に推定することができる。

説得対話システムの発話行為として、現状の話題から所望の話題への誘導を導入する。システムは直前のユーザの発話の中の話題に着目しその話題に直接関係のある研究分野から説得目標への誘導に向けて、最も嗜好が高いと推定される話題を新しい話題として提示する。具体的には下記手順に基づいて次の話題を提示する。

- (1) システムの知識において、研究分野と研究分野間の関係をそれぞれノードとエッジとし、重みが $\log(P(e_k))$ である重み付きグラフを作成する。
- (2) 開始点 (e_{k1}) と終点 (e_{kn}) をそれぞれ現在の話題と勧誘目標とし、重みの総和が最大となる系列 $e_{k1} \dots e_{kn}$ を求める。
- (3) e_{k2} を次の話題に選ぶ。

例えば、表 1 に示すシステムが A 研究室を推薦する対話例において、直前のユーザの発話「でもコミュニケーションはしてないんですね」には「コミュニケーション」が含まれている。この時、図 1 より、「コミュニケーション」は説得目標である「対話」と「音声翻訳」に直接関係がある

ので、「コミュニケーション」から「対話」か「音声翻訳」へと誘導する発話を行う。仮に、「音声翻訳」が「対話」に比べてより嗜好に合致しているとシステムが推定している場合、「音声翻訳」へと誘導する応答を行う。

3.3 説得に向けた対話戦略

ユーザに対し、説得目標とは異なる研究室が有利になる情報を提示することは、説得が失敗する可能性を高める危険性がある。そこで、システムがそのような話題を出来る限り避ける対話戦略を導入する。

システムは、説得対象とは異なる研究室に関連する情報や取り組んでいる研究分野については、可能な限り情報提示は避け、ユーザが要求した場合にのみ提示する。また、その際には、対話状態におけるユーザ嗜好状態に基づき、ユーザの研究室選びに大きく影響を与えないと推定される研究分野についての情報を優先的に提示する。例えば、システムが図1の知識を持っており、A研究室に勧誘することが目的の場合において、音声認識がユーザの嗜好に合致すると推定された時、システムから「音声認識はB研究室でとり組まれています」という発話は行わない。また、ユーザから「B研究室はどのような研究をやっていますか」と質問された場合は、システムは音声認識よりも嗜好が低いと推定される音声翻訳を取り上げて、「B研究室では音声翻訳の研究を行っています」と返答する。

3.4 対話中の興味状態推定

説得対話では、ユーザの嗜好状態の推定に有益な情報を、対話を通して引き出す必要がある。その際に、ユーザの興味に沿って話題を誘導することで、より自然な対話を実現できると考えられる。また、多くの情報を自然な対話の流れで引き出す上で、ユーザの興味がある話題で対話を進めることは有益である。ユーザ嗜好に沿って同様のことを行うことも考えられるが、仮にユーザの嗜好と合致する研究分野に関する話題のみを続けると、ユーザの興味は減少し、十分な情報を引き出せなくなる可能性がある。逆に、ユーザの嗜好とは異なる研究分野においても、それに関する話題に対して興味を示す際には、その話題を通じて多くの情報を引き出せる可能性もある。そこで、対話を通して変化していくユーザの興味状態を推定し、適切な応答を行うために、対話状態を次式にてモデル化する。

$$P(e_k, i_k | A, O) = \frac{\sum_{e'_k, i'_k} P(e_k | e'_k) P(i_k | i'_k) P(e'_k) P(i'_k)}{P(O|A) \cdot P(O|A, e_k, i_k)} \quad (2)$$

ここで、 i_k, i'_k はそれぞれ研究分野 k に対するユーザの現時点における興味と直前のターンにおける興味を表わす。対話状態に対するベイジアンネットワークによる記述を図3に示す。その他の変数に関しては、式(1)と同様である。この対話状態を用いることで、これまでの対話を通じて推定されるユーザの嗜

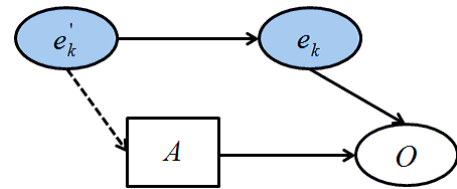


図2 ベイジアンネットワークによるベースラインモデルにおける対話状態の記述

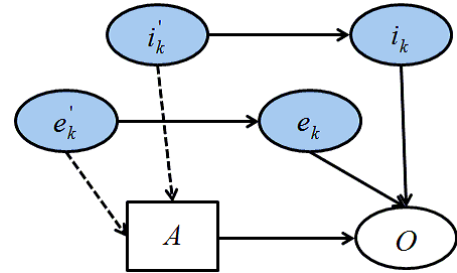


図3 ベイジアンネットワークによる興味状態推定を考慮したモデルにおける対話状態の記述。

好と興味を考慮して、システムの発話行為を決定することが可能となる。なお、嗜好は対話を通して不変なものであるのに対し、興味は対話を等して動的に変化するという点で、この2つは本質的に異なる。興味状態は、下記の対話戦略において利用される。

- ユーザから特定の研究分野や研究に対して、質問が無い場合等において、本システムが現在の話題から別の話題に切り換える際、最も興味があると推定した話題に移動する。
- また、現在の話題に関連した話題に基づいて説得目標へ誘導する際は、ユーザが最も興味を示すと考えられる話題を選択する。

4. 実験的評価

4.1 実験条件

表4に示す4つのシステムについて主観評価を行う。説得対話タスクとして、2つの研究室A, Bの中から、システム側で設定する研究室へと勧誘する説得対話を取り扱う。

対話状態をモデル化の上で必要となる変数間の遷移確率や観測確率については、出来る限り適切な対話状態のモデル化が行われるように、人手により決定する。例として、嗜好状態および興味状態の遷移確率を表2, 3に示す。対話状態からシステム応答発話を選択する方策関数についても、各システムに対して、出来る限り適切な応答が生成されるように、人手で決定する。

各システムを評価する際の実験手順を以下に示す。

- (1) 研究分野は14種類用意し、ユーザは対話を行う前に、嗜好に沿った研究分野を2つ選定するよう指示する。
- (2) 各研究室が取り組んでいる研究分野を決定する。まず、ユーザが嗜好に沿って選んだ研究分野の内、少なくとも一つを両方の研究室で取り組んでいるものとする。各研究室が残りの13種類の研究分野に取り組んでいるかについては無

表 2 嗜好の変移確率

$P(e_k e'_k)$	$e'_k = 0$	$e'_k = 1$
$e_k = 0$	1	0
$e_k = 1$	0	1

表 3 興味の変移確率

$P(i_k i'_k)$	$i'_k = 0$	$i'_k = 1$
$i_k = 0$	0.7	0.7
$i_k = 1$	0.3	0.3

作為に決定する。

- (3) 各研究室が取り組んでいる研究分野, および, システムが勧誘目標とする研究室については未知という状況の下で, 対話を開始する。
- (4) システムの勧誘に対して, ユーザが A, B どちらかの研究室に入るという意味決定を行う, もしくは, ユーザからシステムとの対話を打ち切ると要望のあった場合に, 対話を終了する。
- (5) ユーザに対して満足度と自然性について, アンケートを行う。

各ユーザが個々のシステムを評価する順序はランダムとすることで, 順序による影響を極力減らす。

音声認識誤り, 言語理解誤り, 合成音声による明瞭性低下などの影響を無くし, 対話制御部のみの評価を行うために, ユーザとシステムの間の人を仲介させて, 該当箇所を人手で代用する。被験者数は 8 人であり, 各被験者は 4 つのシステムに対してそれぞれ 2 回対話を行う。以下に示す 3 種類の評価基準により, システムの比較を行う。

説得成功率 システムが勧誘目標とする研究室を, 最終的に被験者が選んだ割合

満足度 利用者が対話に満足したかを 5 段階で評価した結果 (1: 満足でない, 3: どちらとも言えない, 5: 満足である)

自然性 利用者が対話の自然性を 5 段階で評価した結果 (1: 自然でない, 3: どちらとも言えない, 5: 自然である)

4.2 実験結果

説得成功率, 満足度, 自然性に関する結果を図 4, 5, 6 に示す。この結果から, 話題誘導, 説得に向けた対話戦略, ユーザの興味推定といった機能を導入するほど評価が高くなる傾向がみられる。話題誘導, 説得に向けた対話戦略とユーザの興味推定は, 加算的に説得正解率と平均満足度の改善に寄与する。また, 話題誘導とユーザの興味推定は, 満足度および自然性の改善に寄与する。

5. おわりに

本報告では, システムの意図した行動をユーザに取らせるように対話を導く説得対話システムにおいて, 対話制御部の構築を行った。対話状態を確率的にモデル化する仕組みを適用し, 話題誘導, 説得に向けた対話戦略とユーザの興味推定を導入した。主観評価実験を行った結果, 話題誘導および興味推定により, システムの性能が向上することを確認した。今後は, 実際

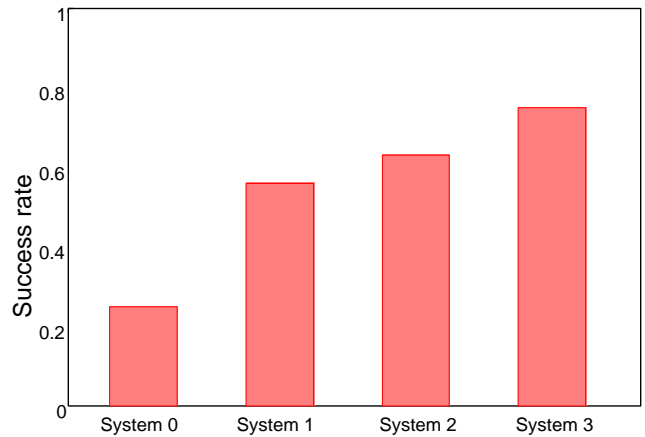


図 4 説得成功率

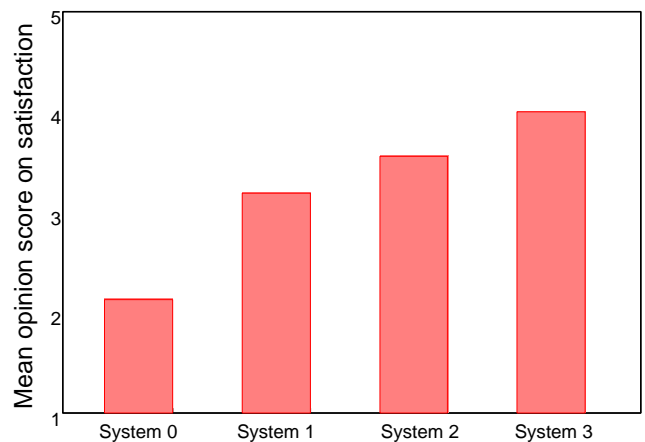


図 5 平均満足度

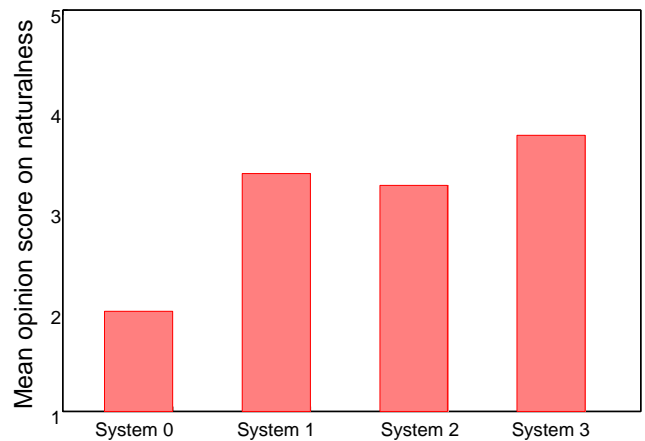


図 6 平均自然性

の対話データに基づくシステムの最適化を行う。

参考文献

- [1] Esther Levin, Roberto Pieraccini, Wieland Eckert, A Stochastic Model of Human-Machine Interaction for Learning Dialog Strategies, IEEE Trans.SAP, Vol.8, No.1, 2000.
- [2] Stephen G Pulman, Conversational games, belief revision and baysian networks, Proceedings of the Association for Computational Linguistics, the Netherlands meeting,

表 4 評価用システム

	話題誘導 (3.2 節)	対話戦略 (3.3 節)	ユーザの興味 (3.4 節)
System 0	考慮しない	考慮しない	考慮しない
System 1	考慮する	考慮しない	考慮しない
System 2	考慮する	考慮する	考慮しない
System 3	考慮する	考慮する	考慮する

表 5 ユーザの対話行為 (O)

対話行為	説明
Answer(Research Field)	研究分野 (Research Field) を含んだシステムの質問への返答
Agreement	システムの発話への同意
Disagreement	システムの発話への不同意
SetQuestionOther(Lab)	研究室 (Lab) がシステムが提示した研究以外に取り組んでいる研究の質問
SetQuestionConcrete(Lab, Research Field)	Lab または Research Field 間の関係の具体化に関する質問
CheckQuestionConcrete(Lab, Research Field)	Lab または Research Field 間の関係の確認
AddressSuggestHold	勧誘に対して、決定を保持する意思の表明
Mumble	システムが理解出来なかった発話

表 6 システムの対話行為 (A)

対話行為	説明
SetQuestionInterest(Research Field)	興味のある研究分野 (Research Field) を質問
SuggestEnteringLab(Lab)	研究室 (Lab) に入るように提案
CheckQuestionInterest(Lab, Research Field)	Lab または Research Field についての説明を受けたいか質問
InformRelation(Lab, Research Field)	Lab または Research Field 間の関係の説明
InformMrSystemCan'tExplain(Lab Research Field)	Lab または Research Field 間の関係の説明が出来ないことを説明

1996.

pp.789–798, 2004.

- [3] Eric Horvitz, Tim Paek, A computational architecture for conversation, Proceedings of the Seventh International Conference on User Modeling, Springer, New York, pp.201–210, 1999.
- [4] Jason D. Williams, Steve Young, Partially Observable Markov Decision Processes For Spoken Dialog Systems, Computer Speech and Language, vol.21, pp.393–422, 2007.
- [5] Chiori Hori, Kiyonori Ohtake, Teruhisa Misu, Hideki Kashioka, Satoshi Nakamura, Dialog management using weighted finite-state transducers, INTERSPEECH, pp.211–214, 2008.
- [6] Yasuhiro Minami, Ryuichiro Higashinaka, Kohji Dohsaka, Toyomi Meguro, Eisaku Maeda, Trigram dialogue control using POMDPs, Spoken Language Technology Workshop, pp.324–329, 2010.
- [7] Wayne Ward, The CMU Air Travel Information Service: Understanding Spontaneous Speech, Proceedings of the workshop on Speech and Natural Language, pp.127–129, 1990.
- [8] Teruhisa Misu, Komei Sugiura, Kiyonori Ohtake, Chiori Hori, Hideki Kashioka, Hisashi Kawai, Satoshi Nakamura, Modeling Spoken Decision Support Dialogue and Optimization of its Dialogue Strategy, ACM Transactions on Speech and Language Processing, Vol.7, pp.10:1–10:18, 2011.
- [9] J. Weizenbaum, ELIZA – A Computer Program for the Study of Natural Language Communication between Man and Machine, Communications of the Association for Computing Machinery, Vol.9, pp.36–45, 1966.
- [10] 西村竜一, 西原洋平, 鶴身玲典, 李晃伸, 猿渡洋, 鹿野清宏, 実環境研究プラットフォームとしての音声情報案内システムの運用, 電子情報通信学会論文誌, Vol.J87-DII, No.3,