情報処理学会論文誌 Vol.53 No.12 2787–2801 (Dec. 2012)

聞き役対話の分析および分析に基づいた対話制御部の構築

目黒 豊美[1,a)] 東中 竜一郎[2] 堂坂 浩二[1,†1] 南 泰浩[1]

受付日 2012年1月23日, 採録日 2012年9月10日

概要：我々は，システムがユーザの話を聞くことによってユーザが「話したい」という欲求を満たす聞き役対話システムの構築を目的としている．本研究では，まず，聞き役対話の特徴を分析するために，聞き役と話し役に分かれて行う対話（聞き役対話）と役割を持たない対話（雑談）を収集し，対話行為の割合を比較，また，対話の流れを HMM を用いて分析した．その結果，聞き役は自己開示を減らし，質問と相槌を増やすことで，話し役に自己開示や，情報提供をさせていることが確認できた．また，聞き役は自己開示をしてから質問をすることで，話し役が話しやすい対話の流れを作っていることが分かった．これらの得られた知見をもとに，対話制御システムをルールベースで構築し，その聞き役対話制御における有効性の検証を Wizard of Oz 方式により行ったところ，提案のルールによる制御方法が人間による制御時とほぼ同等の満足度を示し，分析に基づく提案ルールの有効性を示すことができた．

キーワード：聞き役対話，対話システム，対話制御

Creating a Dialogue Control Module for Listening Agents Based on the Analysis of Listening-oriented Dialogue

Toyomi Meguro[1,a)] Ryuichiro Higashinaka[2] Kohji Dohsaka[1,†1] Yasuhiro Minami[1]

Received: January 23, 2012, Accepted: September 10, 2012

Abstract: Our aim is to build listening agents that can attentively listen to the user and satisfy his/her desire to speak and have him/herself heard. This paper describes the analysis of such listening-oriented dialogues (LoDs) to reveal the characteristics of LoDs and to build rules for LoD control based on the analysis. We collected both listening-oriented dialogues and casual conversation, and analyzed them by comparing the frequency of dialogue acts, as well as the dialogue flows using Hidden Markov Models (HMMs). The analysis revealed that LoDs and casual conversation have characteristically different dialogue flows and that it is important for listener's self-disclosure before asking questions and to utter more questions and acknowledgment than in casual conversation. We created dialogue control rules based on our analysis of the dialogues, and built our rule-based system and compared it with three other systems by a Wizard of Oz (WOZ) experiment. As a result, we found that our rule-based system achieved as much user satisfaction as human listeners.

Keywords: listening-oriented dialogue, dialogue system, dialogue control

1. はじめに

人は，誰かに，自分の主張，趣味・嗜好，経験などを話すことで，肯定感を得たり，考えを整理したり，感情をコントロールしたり，新しい気付きを得たりしている．また，傾聴ボランティアの広まりや，聞き役としてユーザの話を聞くビジネスが近年開始されるなど[*1]，聞き役のニーズが高まっている．しかし，このような対話を人が行うことはコストが高く，ユーザが話したいと感じたときに常に他の人に話を聞いてもらうことは難しい．

1 NTT コミュニケーション科学基礎研究所
NTT Communication Science Laboratories, NTT Corporation, Souraku, Kyoto 619–0237, Japan
2 NTT メディアインテリジェンス研究所
NTT Media Intelligence Laboratories, NTT Corporation, Yokosuka, Kanagawa 239–0847, Japan
†1 現在，秋田県立大学
Presently with Akita Prefectural University
a) meguro.toyomi@lab.ntt.co.jp

*1 シニア向けコミュニケーションサービスの実現に向けた取り組み：http://www.ntt-west.co.jp/news/1203/120327a.html（2012 年 5 月 29 日検索）など，聞き役のニーズは高まっている．

© 2012 Information Processing Society of Japan

	発話文	対話行為
L:	テーマは旅行ということですが，夏休みには旅行されましたかー？	質問
	僕は河口湖よりさらに西の，西湖というところに行ってきましたよー	自己開示
S:	早速ですが，私は旅行というものが大好きなんです．	自己開示
L:	お，そうなんですか！！	共感
	どんな理由がおありなんですか？？	質問
S:	今年の夏は，地元に帰ったぐらいですかね．	自己開示
	仕事で忙しくて，ちなみに河口湖は今週末行く予定です．	自己開示
	旅行好きな理由，多分刺激なんだと思います．	自己開示
L:	刺激ですか〜，いつもと違った場所に足を運ぶと心持が変わってきますものね．	共感
	今週末の，その河口湖もご旅行ですよね？	質問
	お車ですか，それとも電車で？？	質問

図 1 聞き役対話の抜粋．1 文ずつに対話行為（3.2 節参照）が付与されている．L は聞き役，S は話し役

Fig. 1 Excerpt of a typical listening-oriented dialogue. Dialogue acts corresponding to utterances are shown in parentheses (See Section 3.2 for their meanings).

そこで我々は，ユーザの話に積極的に耳を傾け「話したいという欲求」や「聞いてもらいたいという欲求」を満足させることで，ユーザの心理状態を良好な状態に導く「聞き役対話システム」の構築を目指している．本稿では，対話において一方が聞き役となり，もう一方の話し役の話を聞くような対話を「聞き役対話」，人の聞き役となる対話システムを「聞き役対話システム」と呼ぶ．**図 1** は典型的な聞き役対話の例である．図では聞き役（L，Listener の意）が話し役（S，Speaker の意）に繰り返し質問をすることで話し役の自己開示を促している．このような対話を実現する聞き役対話システムは，たとえば「ユーザの話を聞き，理解したことを伝える」，「ユーザに共感を表したり，賛同や承認をしたりする」，「話しやすいような話題へ誘導する」，「話しやすいように適切な質問をする」といったことを行うことにより，ユーザの話したいという欲求を満足させる．システム化には，以下のような利点があげられる．

- いつでもユーザが好きなときに対話を行うことができる．
- 相手との人間関係に考慮する必要がない．
- 人的コストが低い．

対話システムは大きく分けて，ユーザの発話を理解する「発話理解部」，発話理解部で得られた内容をもとに次のシステム発話を決定する「対話制御部」，そして制御部が決定した内容をもとに自然なシステム発話を生成する「発話生成部」の 3 つから構成されることが多い [1], [2]．発話理解部は，自然発話を対話行為 [3]（発話内容を「質問」「自己開示」といった抽象的な表現で表したもの）に変換し，対話制御部に渡す．対話制御部は，発話制御部が出力したユーザの対話行為をもとに，次の最適なシステムの対話行為を決定する．発話生成部は，対話制御部が決定したシステムの対話行為をもとに，システムの発話（自然言語）を生成する．対話制御部は対話行為を決定し対話の流れを決める最も重要な部の 1 つであり，本研究では聞き役対話システムの対話制御部に着目する．

本稿では，聞き役対話システムの対話制御部構築を大きく 2 つの課題に分けて行う．第 1 の課題は，聞き役対話とはどういうものかを把握し，客観的指標に基づいて聞き役対話の特徴を分析することである．この分析を通し，聞き役対話の特徴をとらえた対話制御部構築の指針を得る．第 2 の課題は，分析に基づいた聞き役対話システムの対話制御部をこの指針に基づき構築することである．

これらの課題を順に解決することで，聞き役対話システムの実現を目指す．まず，第 1 の課題に対しては，対話全体の流れを把握するために，対話行為レベルでの分析を行う．この対話行為の流れを，対話参加者が役割を持たない対話（本稿では「雑談」と呼ぶ）と比較してみることで，聞き役対話特有の流れを明らかにする．対話に制約を設けない雑談と比較することで，通常の対話の流れと聞き役対話の特徴的な流れの違いが明確になることが期待される．具体的には，対話行為を状態の出力とし，話者遷移を考慮した HMM を用いた分析を行う．

第 2 のステップとして分析に基づいたシステムを実現する．分析で得られた知見からシステムの次の対話行為を決定するルールを作成し，対話制御部を構築する．ルールを用いた対話制御の構築方法は聞き役対話のような明確なタスクを持たない対話システムの実現方法として最も一般的な方法である [4], [5], [6]．なお，HMM をそのまま制御に用いるということも考えられるが，ここでは，HMM を人手で解釈して得られたルールを用いる．これは，聞き役対話の研究が初段階であり，HMM のような統計のみに基づくシステムが満足度の高い対話行為の出力を行えない可能性があると考えたためである．

なお，本研究では，音声による対話ではなく，チャットのような，テキスト対話を対象とする．これは，言いよどみ，割込みなど，音声特有の問題を，切り離して考えたいからである．将来的には，本研究で得られた知見を音声対

© 2012 Information Processing Society of Japan

話システムへ応用する予定である．

以降，2 章では先行研究を紹介する．3 章では聞き役対話と雑談の比較分析の結果を報告する．4 章では 3 章の分析の結果をもとに構築したルールを用いた対話制御システムについて述べ，WOZ を用いた評価実験を行った結果を報告する．最後に 5 章でまとめを述べる．

2. 先行研究

聞き役対話システムの先行研究として，まず，非言語的要素に着目した研究があげられる．Maatman らの研究 [7] では，エージェントが頷いたり，首をふったりするジェスチャを用いながらユーザの発話を聞くことで，ユーザは「聞いてもらっている」と感じるという結果を報告している．加えて，パラ言語コミュニケーションに着目したものとして，Noddy [8] がある．Noddy はタスク指向型対話システムで，音声認識を行い，タスク達成に重要な単語が発話されたときに適切なタイミングで相槌をいれることで，ユーザに「システムが聞いている」という感覚を与えるものである．

しかし，実際に対話を行う聞き役対話システムを実現するためには，非言語的要素だけではなく，言語的要素にも着目する必要がある．近年の言語的視点からの研究として，下岡らは聞き役の応答生成に着目した研究を行っている [4]．彼らは音声認識器 Julius から算出される，ユーザ発話内の格要素の単語認識信頼度 [9] を用い，高い信頼度のときには「繰返し/問い返し発話（例：お花を貰ったんですか)」または「共感（例：それは嬉しいですね)」を行い，低い信頼度のときには，「相槌（例：そうですか)」を行う．「繰返し/問い返し発話」では，そのままユーザの発話を繰り返す，または，ユーザ発話中の必須格の有無を調べ，必須格がない場合には問い返しを行う．「共感」ではユーザの感情を推定し，ユーザの感情と同じ感情を表出する．横山らの研究 [5], [6] は，対話システムとのインタラクションを長く継続させるために，傾聴モードと話題提示モードを切り替える．「話題提示モード」ではロボットが “話し手” となり話題を模索し，「傾聴モード」ではユーザの話を聞く “聞き手” となる．傾聴モードでは，「相槌」か「質問」を行う．発話文は，音声認識結果から重要な単語を抽出しテンプレートを埋めることで作成している．これらの先行研究では，カウンセリングもしくは傾聴の行動として重要と考えられている対話行為（すなわち，共感，相槌，繰返し/問い返し発話など）の中から 1 つを選択して対話制御を行っている．

我々も言語的要素に着目し，「言語的コミュニケーション」を行うことで聞いてもらっているという感覚をユーザに与えることを目的としている．この点では下岡らや，横山らの研究と同じであるが，下岡らや横山らが，カウンセラがなすべきだと思われるいくつかの対話行為に絞って聞き役を実現しているのに対し，我々は，日常会話においてユーザが話しやすい対話システムの実現を目指している．そのために，聞き役対話を収集し，分析した結果を用いて対話制御を行っている点が異なる．具体的にいえば，図 1 の聞き役の 1 発話目に含まれる自己開示のように，ユーザから明示的に自己開示を要求された場合でなくても，システムが自発的に自己開示を行うような行為が聞き役対話では重要であるが（3 章を参照），このような知見は分析によって初めて分かることである．

なお，我々は対話システムとして本手法とパラ言語，非言語の要素と組み合わせることは今後の重要な課題だと考えており，本稿で提案する対話制御部は，前述の Maatman らの研究などと相補的なものと考えている．

聞き役に類した対話を行う対話システムとして ELIZA [10] がある．ELIZA は Rogerian psychotherapist を模した対話システムである．ELIZA は，単純な聞き返しなど受動的な行動のみ行うが，我々の目指すシステムは，日常会話における聞き役であり，日常会話に現れる対話行為を含めて，積極的な質問や自己開示などの多様な制御を行う．なお，本研究が目的とする聞き役システムにも，ELIZA のように，カウンセラの典型的な行動が含まれる可能性はある．

また，日常会話を目的とした雑談システムとして，A.L.I.C.E. [11] や Rosette [*2] などの Loebner prize [12] が目指すシステムがあげられる．これらのシステムは，わざと打ち間違いを混入させるなど「人らしさ」を目指したシステムで，必ずしも対話の満足度を向上させることを第 1 の目標としているわけではない．我々のシステムは聞き役対話におけるユーザの満足度を上げることを目的としており，A.L.I.C.E. などとは対話の目的が異なる．ただし，雑談的な機能において実装が重複する可能性がある．

その他，関連研究として，タスク指向対話を HMM で分析した研究や [13] や雑談を HMM で分析した研究 [14] があるが，我々は，聞き役対話を対象としており，また，対話のモデル化だけでなく，対話制御部の構築を目的にしている点が異なる．

3. 聞き役対話と雑談の比較

聞き役対話の特徴を明らかにするために，聞き役対話と雑談を収集し，それらを比較することで対話の流れを分析する．まず，人間同士の聞き役対話と雑談を収集する．次に，分析を行うために収集したデータに対話行為を付与する（対話行為の詳しい説明は 3.2 節を参照）．

聞き役対話と雑談の間で，2 つの比較分析を行う．まず，聞き役対話と雑談の 2 つの対話タイプ間で，対話行為の割合を比較する．聞き役対話において割合が大きい対話行為

*2 http://labs.telltalegames.com/rosette/

© 2012 Information Processing Society of Japan

や，雑談と比較して多く出現する対話行為は，聞き役対話にとって重要であり，また逆に，割合が小さい対話行為や雑談と比較して少ない対話行為は，聞き役対話において重要ではないと考えられる．この分析により，聞き役の行動の特徴を明らかにするとともに，システム構築において，着目すべき対話行為を明らかにできる．たとえば，聞き役対話において聞き役は，「話し役のことを積極的に尋ねる」，「自分の情報を抑えめにする」などの特徴を含んでいる可能性がある．

しかし，対話行為の割合の比較だけでは対話にとって重要な，対話の流れの分析ができない．そこで，これら 2 タイプの対話データを隠れマルコフモデル（HMM）でモデル化し，それぞれのタイプの対話の流れを比較分析する．聞き役対話は複雑な構造を持つ可能性があり，かつ，対話の構造が既知ではないため，状態数が不明なデータ系列の構造を学習するのに一般的な手法である HMM を採用する．なお，聞き役と話し役の対話行為を分離して分析するために，それぞれの役割に対応した状態を持つ HMM を用いる．

HMM を用いた分析では，まず，HMM が「聞く過程」を適切にモデル化できているか調べるために，聞き役対話と雑談から作った HMM を用いて，それぞれのタイプの対話が適切に識別できるか確かめる．その後，作成した HMM の状態遷移を分析して，対話の流れを詳細に調べる．

3.1 対話データ収集

データ収集の実験参加者は計 16 人，8 人を聞き役とし，残り 8 人を話し役とした．男女比は半々とし，年齢は，21〜29 歳であった．それぞれの実験参加者は最初に雑談を 2 対話，その後聞き役対話を 2 対話，計 4 対話行った．なお，対話のバリエーションを確保するため，実験参加者は 4 人の異なる話者と 1 回ずつ対話をし，すべての聞き役と話し役のペアは 1 度しか対話しないようにした．雑談を収集する際には，その後に割り振られる役割について実験参加者には知らせておらず，雑談収集後にそれぞれの役割をランダムに振り，聞き役対話を収集した．すべての実験参加者について，同じ順番で対話を収集しているが，これは，聞き役対話収集後に雑談を収集すると，実験参加者が直前に行った役割を引きずる可能性があると考えたためである．たとえば，「聞き役（話し役）が聞き役（話し役）を続けてしまう」や，「聞き役（話し役）が役割がなくなったことを気にしすぎて，積極的に話しすぎて（話さなくなり）不自然になる」などの状況が起き，自然な（役割が規定されていない）状態での雑談が収録できないという懸念である．そこで，先に雑談を収集し，その後，聞き役対話を収集する必要があると考えた．また，収集前，収集中を通して，実験参加者は対話相手とは顔を合わせることはないように，聞き役と話し役は別々の部屋で収集を行った．収集には，Microsoft Live Messenger™ を使用した．

表 1 分析用データの内容
Table 1 Statistics of collected dialogues for analysis.

		聞き役対話	雑談
対話数		16	16
発話数		850	720
平均単語数	聞き役	20.60	17.92
	話し役	26.46	21.44

聞き役に対しては，「対話相手からテーマに関する好き嫌いを聞き出してください．相手が話しやすいように心がけてください」，話し役に対しては，「話し役として，対話相手にテーマに関する好き嫌いを伝えてください」と指示を与えた．

対話開始前に，実験参加者は実験者が用意した複数のテーマの中から好きなテーマを選び，それぞれについて，30 分間チャットを行った．テーマは，Yahoo!カテゴリ名から，話しやすいと思われるものを選択した．テーマは 12 個で，食べ物，旅行，映画，音楽，芸能人，スポーツ，健康，家事および育児，パソコンおよびインターネット，動物，ファッション，ゲームである．**表 1** は収集した対話数，発話数，聞き役と話し役の 1 発話あたりの平均単語数を示している．

我々が収集したい対話は，「聞いてもらえている」と感じる対話であるが，聞き役対話の用途を考えると，尋問調になったりして，満足度を下げるような対話になってはならない．そこで，それぞれの対話の後に，実験参加者はアンケートにより，対話の満足度や対話相手に聞いてもらえていると思ったかを，10 段階のリッカート尺度に基づき評価してもらった．聞き役対話と雑談に付与されたアンケート項目の満足度（「今の対話に対する満足度はどの程度ですか？」）を比較したところ，平均値には有意差がないことが確認できた．また，「聞いてもらえていると思ったか？」についての評価値は，聞き役対話が平均 7.5，雑談 5.6 であり，聞き役対話の方が有意に高かった（Welch's pairwise t-test; $p = 0.016$）．なお，このアンケート項目については，聞き役対話に対して値が高くなるように誘導している可能性が残るため，今後，より適正な調査により検証する予定である．

3.2 対話行為の付与

聞き役対話の特徴を分析するために収集したデータに 6 種類の対話行為を付与した．対話行為の定義を**表 2** に示す．これらの対話行為は，一般的な対話を扱うために作られた DAMSL タグセット [15] と，聞き役対話と近い関係にあると思われるセラピー対話用のタグセット [16] を参考にした．今回は大まかな流れを分析するため，収集したデータ内の発話をモデリングするのに特に重要であると思

© 2012 Information Processing Society of Japan

表 2 対話行為の定義
Table 2 Definition of dialogue act.

対話行為	定義
自己開示	嗜好や感情を開示している発話
情報提供	客観的な情報を伝えている発話
相槌	対話相手の発話を促す発話
質問	相手の返答を期待した発話
共感	共感した発話や，相手への励ましなど
挨拶	対話の最初と最後に見られる社会的な発話

われる最小限の対話行為を用いた．たとえば聞き役は相手のことを聞く対話であり，話し役は積極的に自分のことを話していると考えられるため，自己開示を含めた．なお，Altman らの social penetration theory [17] では，対話参加者間の親密さが増えるほど自己開示が増えるとされており，相手のことを聞く，聞き役対話において自己開示は重要である可能性がある．共感は，対話参加者同士の親近感を高めるのに有用であるということが分かっており [18], [19]，聞き役が話し役との親近感を高めるために活用している可能性があるため，加えた．その他，日常的な対話において重要と思われる情報提供，相槌，質問，挨拶を含めた．

著者ではないアノテータ 2 人にそれぞれの発話に 7 つのタグ（6 つの対話行為と「その他」）を付けてもらった．アノテータ間のタグの一致率を Cohen's κ を用いて測ったところ，0.75 と高い一致率であった．結果として，1,177 個のタグを聞き役対話に，1,312 個のタグを雑談に付与した．タグの数と発話数が一致しないのは，テキスト対話であるという性質上 1 発話が長く，1 文ごとにタグを付けたためである．ここで 1 文とは，基本的に句点，もしくは 1 発話終了時を切れ目としたものである．しかし，1 文の途中で発話が終了していたり，句点が適切に挿入されていなかったりする場合がある．そのため，片方のアノテータが人手で適切な文区切りと思われるを決定し，その後 2 人のアノテータが独立してタグ付けを行っている．

3.3 聞き役対話と雑談の比較

3.3.1 対話行為の頻度の比較

聞き役対話と雑談の対話行為の頻度（対話中の割合）を比較した．**表 3** はそれぞれのタイプの対話における対話行為の割合を示している．表中の「その他」は顔文字や打ち間違いなど 6 つの対話行為に該当しない文が含まれている．すべての実験参加者は，雑談対話に参加後，聞き役対話に聞き役，もしくは，話し役として参加しているが，表 3 において，「聞き役–雑談」となっているカラムは，後半の聞き役対話において聞き役を行った実験参加者が，雑談時に行った対話行為の割合を表している．同様に，「話し役–雑談」のカラムは，聞き役対話において話し役を行った実験参加者が，雑談時に行った対話行為の割合を表している．

表 4 は対話行為の割合が雑談時に比べ，聞き役対話時に増加，もしくは，減少した聞き役の人数を示している．話し役にはほとんど増減が見られなかったが，雑談と比較して，聞き役対話では自己開示と情報提供の割合が減少した．反対に，相槌と質問の割合は増加した．このことから聞き役は自分の情報を少し抑え目にし，その代わり，話し役が自己開示や，情報提供をするように，対話行為を生成していることが確認できた．

表 3 対話行為の割合
Table 3 Rates of dialogue act tags.

	聞き役		話し役	
	雑談	聞き役対話	雑談	聞き役対話
自己開示	57.3%	44.5%	53.3%	66.6%
情報提供	5.2%	1.4%	5.6%	6.5%
相槌	6.9%	12.3%	6.6%	8.0%
質問	14.0%	25.8%	21.3%	4.1%
共感	3.3%	3.7%	3.2%	2.6%
挨拶	9.6%	9.8%	7.2%	10.9%
その他	3.7%	2.5%	2.9%	1.3%

表 4 聞き役対話におけるタグの割合が増加または減少した聞き役人数
Table 4 Number of listeners whose tags increased/decreased.

	増加	減少
自己開示	0	8
情報提供	0	8
相槌	8	0
質問	8	0
共感	5	3
挨拶	4	4

なお，表 3 において，雑談において聞き役のグループは質問が 14.0%，同じく話し役のグループは質問が 21.3%と差がある．ただし，聞き役と話し役の質問の割合について，平均値に差があるか検定したところ，これらの間に有意な差は認められなかった．直接対話内容を精査したが，特に値が突出している実験参加者がいるわけではなかったことから，この差は，揺れの範囲内だと考えられる．

3.3.2 HMM を用いた対話の流れの分析

聞き役対話と雑談の流れを分析するために，対話行為の系列を HMM でモデル化した．6 つの対話行為に「その他」を加えた 7 つの対話行為を用い，対話行為を聞き役と話し役それぞれを別の観測値として考え，計 14 個の観測値を用いた．モデル化には HMM のライブラリである Jahmm [*3] を利用し，独自構成の HMM を構築した．HMM の半分の状態からは，聞き役の対話行為しか出力しないようにし，残りの半分からは話し役の対話行為しか出力しないようにした．具体的には，半分の状態では，聞き役の対話行為の出力確率の初期値をゼロとし，同様に残りの半分の状態では，話し役の出力確率の初期値をゼロとした．こ

*3 http://code.google.com/p/jahmm/

© 2012 Information Processing Society of Japan

表 5　聞き役対話と雑談から学習した HMM を用いた識別実験の confusion matrix
Table 5　Confusion matrix for the recognition experiment.

	識別結果	
	聞き役対話	雑談
聞き役対話	28	4
雑談	4	28

うして，学習を行うと，初期値ゼロの出力は学習が行われず，ゼロのままとなり，結果的に，その状態からは，聞き役話し役どちらかの対話行為しか出力されないようになる．それぞれの状態から他の状態はすべて遷移可能である．このような HMM を，Speaker HMM と呼ぶ．HMM の学習には EM アルゴリズムを用いた．最適な HMM を見つけるために，状態数を各役割につき 1〜10 個（つまり，総数 2〜20 個）と変えてみて，それぞれの状態数で 100 個の HMM，計 1,000 個の HMM を学習した．この 1,000 個の HMM から Minimum Description Length（MDL）基準を用いて最適な HMM を選んだ．このような過程で HMM をモデリングするのは，EM アルゴリズムを用いた学習では初期確率に依存するので，できる限り網羅的に最適な HMM を求めるためである．

3.3.3　対話タイプの識別

学習した HMM が聞き役対話と雑談を識別できるかを検証するために，実験を行った．この実験では，特定の話者が参加した聞き役対話 2 対話と，雑談 2 対話をテストデータとしてヘルドアウトし，残りの 28 対話を学習データにして識別を行うことを，聞き役，話し役すべての話者に対して行った．未知の話者が含まれた対話が聞き役対話か，雑談であるかを識別するという観点において，オープンなテストといえる．

識別率は，平均精度は 87.5%となり，これは，明らかにチャンスレベルの 50%を超えている．これは，それぞれのタイプの対話の流れが HMM によって適切に学習できていることを示している．識別実験から得られた confusion matrix は**表 5** のようになった．識別したサンプルは全部で 64（話者 16 人 × 4）個である．なお，次節からの HMM の分析には，各対話タイプ 16 対話ずつすべてを用いて学習した HMM を用いる．

3.3.4　Speaker HMM の分析

図 2 と**図 3** はそれぞれ聞き役対話と雑談から学習した Speaker HMM を表している．図中での，L は聞き役，S は話し役を表す．また，S1 は聞き役対話収集時に聞き役を行った実験参加者（L），S2 は話し役を行った実験参加者（S）に対応している．それぞれの状態の出力確率と遷移確率を比較してみると，よく似ている構造であることが見て取れる．

たとえば，両方とも状態①と②は挨拶をする状態として

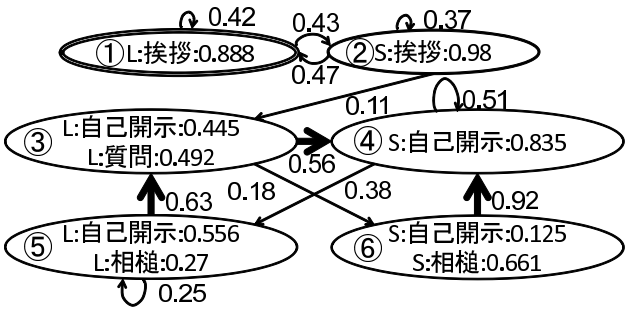

図 2　聞き役対話から作成した Speaker HMM．楕円は状態を表し，楕円内には，話者，観測値，観測確率をコロン区切りで示している．状態間の遷移確率は矢印横に示している．観測確率，遷移確率は，それぞれ 0.1 以上のもののみを示している．L は聞き役，S は話し役である
Fig. 2　Speaker HMM for listening-oriented dialogue.

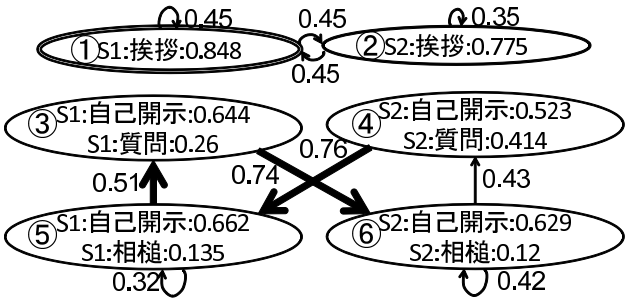

図 3　雑談から作成した Speaker HMM．S1 と S2 は雑談を行った 2 人の対話参加者を表し，聞き役対話においてそれぞれ聞き役，話し役を行った実験参加者である
Fig. 3　Speaker HMM for casual conversation.

それぞれ他の状態から独立している．このように 6 つの各状態の確率値をもとに，対応していると思われるものに図中で同じ ID（①〜⑥）を目視で割り振ったところ，2 つの Speaker HMM の比較から雑談と聞き役対話の流れの違いについて以下のことが分かった：

- 図中の③と⑤の状態を比較することで，聞き役対話中で聞き役は，自己開示を減らし，質問と相槌を増やしていることが分かる．つまり，聞き役のターンにおいて，最初に相槌を行い，最後には質問を行って，話し役のターンへ戻すという流れが重要であるということが分かる．
- 聞き役は質問の前に，自己開示や相槌を行っていることが分かる．これは，質問をし続けて尋問調になることを避け，相手の自己開示を促し，話し役が話しやすいように努めているからだと考えられる．表 3 では，雑談と比較し聞き役対話では自己開示の割合が減っているため，重要な箇所に絞って自己開示を行っていると考えられる．つまり，相手に質問するだけではなく，適切な箇所で自己開示することによって相手の自己開示を引き出している．これらは Altman らの social penetration theory の知見 [17] とも合致する．

© 2012 Information Processing Society of Japan

```
function generate_action(user_action) returns system_action
  num ← random(1..3)
  system_action ← {}
  if user_action = 挨拶 then return 挨拶 × num
  else if user_action = 情報提供 then return (共感, 情報提供 × (num-1))   … (3), (5)
  else if user_action = 自己開示 then return (user_action×num, 質問 (sub:(random.sub))…(1)
  else if user_action = 相槌 then return 自己開示 (sub:事実) × num   …(2)
  else if user_action = 質問 then
    if user_action.sub = 情報提供要求 then return 情報提供 × num   … (3)
    else if user_action.sub = preference then return 自己開示 (sub:評価+) × num
    else return 自己開示 (sub: user_action.sub) × num
  else if user_action = 共感 or 非共感 then return (繰り返し, user_action × (num-1))   … (4)
  else
    if user_action = 繰り返し or 言い換え or 感謝 or フィラー then system_action ⇐ 共感  … (5)
    else if user_action = 提案 or 承認 then system_action ⇐ 感謝
    else if user_action = 謝罪 or その他 then system_action ⇐ 繰り返し   … (6)
    else if user_action = 確認 then system_action ⇐ 言い換え   … (6)
    else if user_action = 感嘆 then system_action ⇐ 感嘆   … (4)
    system_action ⇐ 自己開示 (sub:事実) × (num-1)   … (2)
    return system_action
end
```

図 4 関数として書かれた対話制御用ルール．それぞれの対話行為は表 6 と対応する．“←” は代入，“⇐” はリストへの追加，“× num” は num 個の対話行為の繰返し，“random.sub” はサブカテゴリの対話行為をランダムに選択することを示す．ルール (1)〜(6) については 4.2 節〜4.3 節で後述

Fig. 4 Our dialogue control rules written as a function.

- 聞き役対話中では聞き役の相槌が増えている．今回のようなテキスト対話では音声対話と違って，意識的に相槌を文字として打ち込む必要があるので，聞き役対話には明示的な相槌が重要であることが分かる．

4. HMM に基づいたルールを用いた対話制御

3 章では，役割を持たない対話の代表として雑談を取り上げ，役割を持つ対話である聞き役対話を比較することで，聞き役対話の特徴を明確にすることができた．本章では 3 章の分析結果から作成したルールを用いて対話制御部を実装する．ルールに基づく制御は，聞き役対話のようなタスク指向ではない対話システムにおいて最も一般的な方法であり [4], [5], [6]，対話システムを作るうえでの第 1 歩としては最も基本的な方法である．加えて，近年さかんになっている統計ベースの手法 [20], [21] と比べてルールベースの制御方法には以下のような利点がある．

- ルールによる制御は統計モデルによる制御と比較して，完全に文脈と合わない制御を行う可能性が低く，対話が破綻する可能性が低い．
- 細かな対話の文脈に応じて確実な制御を行うことができ，メンテナンスが容易である．

欠点としては，すべての対話状態に応じたルールを書き尽くすことは不可能であるという点がある．そのため単調になったり，複雑な文脈に対するルールを細かく記述するコストが膨大になったりし，文脈によっては自然でない制御を行うときがある．これは，Meguro らが先行研究で検討中の POMDP [21] などの統計モデルを適用することで克服できる可能性がある．しかし，対話の満足度を扱うような POMDP の制御は手法が確立されているとはいえず，本稿での検討には加えなかった．これは，今後の検討課題である．

4.1 ルール作成

図 4 は 3 章での分析で明らかにされた，聞き役対話の特徴から作成した対話制御を行うためのルールである．1 つの対話行為は 1 文に対応しており，ユーザは 1 ターン内に複数文を発話することができる．ルールはユーザ発話内の最後の文の対話行為に対して適用され，次のシステムの対話行為（列）を生成する．

実際上，自然な対話を行うには，表 2 の 6 種類の対話行為では少ない．たとえば，表 3 によれば自己開示は全体の 50.9%（聞き役話し役平均）を占めている．このまま対話制御を行った場合，システムは 2 度に 1 度は自己開示することになるが，どのような内容の自己開示をすればいいかは具体的に決定する必要がある．そのため，自己開示にはその内容を具体化したサブカテゴリ（事実，経験，習慣，評価，欲求，予定）を導入し，質問にもそれぞれに対応した同様のサブカテゴリを追加した．また，そのほかに聞き役

© 2012 Information Processing Society of Japan

表 6 対話行為の定義．テンプレートは評価実験で WOZ がシステムの対話行為から自然文を作成するために使用したものである．すべてのテンプレートは対話のテーマである「食事」に合わせて作った．自己開示（sub: 評価）は，ポジティブ，ネガティブ，ニュートラルの 3 つのサブカテゴリを持っており，異なる対話行為とする．「繰返し」「言い換え」「その他」以外の対話行為には複数のテンプレートがある．テンプレートは穴埋め形式になっており “時” には，昼，朝，夕が入る．同様に，“日” には明日，今日など，その他穴埋め箇所には対話の内容に合わせて WOZ が適切な単語を挿入する

Table 6 Definitions of dialogue acts with sample generation templates.

対話行為	定義，テンプレート
挨拶	対話の最初と最後での挨拶や，テーマの確認を行う発話 e.g.[こんにちは],[テーマは “テーマ名” です]
情報提供	客観的な情報を提示している発話や，他人からの情報提供や伝聞として発話している発話 [“食べ物” は “地名” が有名です]
自己開示	発話者が自らの情報（個人的な情報，意見，感情など）を表現している発話
sub: 事実	[“日” の “時” ご飯は “食べ物” を食べました]
sub: 経験	[「“食べ物” を食べたことがあります]
sub: 習慣	[私は “食べ物” をよく食べます]
sub: 評価 (ポジティブ)	[“食べ物” が好きなんです]
sub: 評価 (ネガティブ)	[“食べ物” が嫌いなんです]
sub: 評価 (ニュートラル)	[“食べ物” は好きでも嫌いでもないです]
sub: 欲求	[“日” の “時” ご飯は “食べ物” が食べたいです]
sub: 予定	[“日” の “時” ご飯は “食べ物” を食べる予定です]
sub: その他	上記のサブカテゴリに属さない自己開示
相槌	聞いていることを示したり，相手の発話を促すための発話 [はい] [ええ]
質問	対話相手に何かを尋ねて，答えを求めている発話
sub: 情報提供要求	[“食べ物” 教えてください]
sub: 事実	[“日” の “時” ご飯は何を食べましたか？]
sub: 経験	[“食べ物” を食べたことがありますか？]
sub: 習慣	[“食べ物” をよく食べますか？]
sub: 評価	[“食べ物” は好きですか？]
sub: 欲求	[“日” の “時” ご飯には何を食べたいですか？]
sub: 予定	[“日” の “時” ご飯は何を食べる予定ですか？]
sub: その他	上記のサブカテゴリに属さない質問
共感・同意	自分も同意していることを表している発話 [私もです]
非共感・非同意	相手の意見を否定したり逆の意見を述べたりしている発話 [そんなことないですよ]
確認	相手の話している内容を確認する発話 [本当ですか？]
提案	対話相手の行動を促す発話 [ぜひ “食べ物” を食べてみてください]
繰り返し	対話相手の発話や，その発話に含まれる内容をそのまま繰り返している発話
言い換え	対話相手の話している内容を自分の言葉に置き換え話している発話
承認	心配したり，好意を表したり，褒めたりしている発話 [すごいですね]
感謝	感謝の意を表している発話 [ありがとう]
謝罪	謝罪の意を表している発話 [すみません]
フィラー	次の文章を考えるときの発話 [えーっと],[うーん]
感嘆	感情を表している発話 [あはは],[おぉ！]
その他	上記のどれにも属さない発話

に重要とされている「共感」も細分化し，共感に近い「言い換え」「繰返し」なども新たに追加した．最終的に，**表 6** に示されるような 32 種類の対話行為セットを用いた．

ルールは基本的に我々の分析に基づくが，応答できない対話行為があると対話に支障をきたすかもしれないという懸念から，補完的に先行研究からのルール（4.3 節参照）も導入した．なお，括弧付きの数値が付与されていないルールは「一般ルール」と呼び，協調的な対話における一般的な行動（挨拶には挨拶を返すなど）を指し，どのようなシステムにおいても実装されるべき部分と考え，導入したものである．また，ルール中の if 文の条件のどれにもマッチしない対話行為については，下から 3 行目の，system action に自己開示（sub: 事実）をリストへ追加するというルールが実行され，すべての入力対話行為に応答できるようになっている．

© 2012 Information Processing Society of Japan

4.2　分析に基づくルール

分析に基づくルールであるが，3.3.4 項で，話し役の自己開示の後に聞き役がまず自己開示を行い，続いて質問するという分析結果を得た．この分析をもとに，ユーザの自己開示の後に，システムが自己開示と質問（サブカテゴリはランダムに選択）をするという図 4 のルール (1) を作成した．また上記の分析結果と，social penetration theory から，自己開示はシステムに対する親密感を上げるために有効な手段であると考え，ユーザが繰返し，言い換え，感謝などを発話したときには，自己開示を返し，次の発話を引き出すルール (2) を作成した．

加えて，聞き役対話における情報提供が起こる回数が低いという知見から，ユーザが情報提供か情報提供要求をしたときのみ，システムが情報提供を返すようにするルール (3) を作成した．

分析において，相槌が増えているという知見が得られていた．しかし，本ルールではシステムは相槌を扱わないことにしている．これは，後述する WOZ 実験においてシステムの応答までに間が空いてしまうことによるもので，このせいで相槌本来の目的である「発話と発話の間を埋める」「適切なタイミングで相手の意図を理解していることを伝える」という役割 [22], [23] を果たせず，対話を不自然にさせてしまう可能性があると考えたからである．相槌については，今後の検討課題としたい．

なお，人同士の聞き役対話中では，1 発話中に平均 2.2 個（標準偏差 1.1 個）の対話行為を発話している．そこで，ルール中においても，ほとんどのルールが 1 発話中に 1〜3 個の対話行為を発話するようになっている．後述する評価実験で行われた対話では，実際に本手法によるシステムは 1 発話あたり平均 1.8 個の対話行為を選択しており，人間が自由に聞き役を行ったときの対話行為数をおおむね再現できていることが分かる．

4.3　先行研究に基づくルール

4.2 節のルールだけではすべての対話行為に応じたルールを作成することが難しいため，ルールが作成できなかった対話行為については先行研究を用いてルールを作成した．まず，協調的な対話では相手に合わせることがしばしば起こるという先行研究を利用し [24]，ルール (4) ではユーザが共感や非共感をしたときにはシステムも同じ対話行為を行う．また，共感がユーザを対話への興味を持続させる効果があるという先行研究 [19] に基づき，ルール (5) では共感を，繰返し，言い換え，情報提供などの後に積極的に行っている．繰返しや言い換えは聞き役が理解を示すうえで有効な手段であるという聞き役の研究 [25] に基づき，ルール (6) では謝罪，提案などの後に繰返しや言い換えを行うようにした．

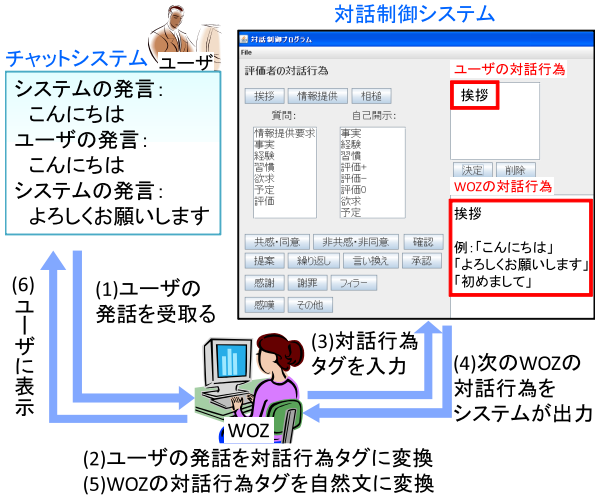

図 5　実験フロー．WOZ はチャットシステムと，対話制御システムを操作し，ユーザ役の実験参加者はチャットシステムの画面のみを見ている．ユーザはチャット画面に自然文を入力し，その入力文から WOZ は対話行為に変換する．対話行為を対話制御システムに入力すると，自動的に次の WOZ の対話行為が出力され，それに基づき次の発話をチャットシステムに入力する
Fig. 5　The flow of the Wizard of Oz experiment.

4.4　評価実験

本ルールを用いた対話制御部の評価実験を行った．対話システムの制御部としての性能を見るためには，実ユーザとの対話を通して性能を評価する必要がある．しかし，提案の制御方法では，対話行為レベルの入出力しか行えず，ユーザ入力文の理解（発話理解部）とシステム発話の生成（発話生成部）の機能は持っていない．そこで，発話理解部と生成部は人間 [Wizard of Oz（WOZ）] が行うことで，自律的な対話システムに近い状態での評価を行う．評価者のユーザからは WOZ は見えず，聞き役システムが自律的に動いていると説明する．

対話制御システムはユーザの対話行為を入力とし，出力として次のシステムの対話行為を出力する．評価実験では，それぞれの対話行為に対して，対話行為から自然文を生成するためのテンプレートを用意し，このテンプレートに基づいて，WOZ が文を作成した．比較のためのシステムは，3 種類用意した（4.4.1 項）．それぞれのシステムはまったく同じ実験インタフェース上で動作し，WOZ はどのようなシステムを，どのような順番で操作するか知らない．WOZ はチャットシステムと，対話制御システムを操作する．ユーザ役の実験参加者はチャットシステムの画面のみを見ている．実験フローは図 5 ようになっており，流れは以下のとおりである：

(1) ユーザはチャットシステムに自然言語の文で発話を行う．
(2) WOZ はユーザの発話を受け取り，対話行為の定義をもとに対話行為タグに変換する．
(3) WOZ は変換したユーザの対話行為を対話制御システ

© 2012 Information Processing Society of Japan

ムに入力する．

(4) システムの対話制御部により，次になすべき対話行為が画面上に出力される．

(5) WOZ はシステムの対話行為のテンプレートの中から最適と思うものを選択，ブランク部分の穴埋めを行い発話を生成する．

(6) WOZ がシステム発話としてチャットシステムに入力する．

対話のテーマは日常的なテーマである「食事」に設定した．

実験には評価者として 15 人（男性 8 人，女性 7 人），WOZ として 3 人（女性 2 人，男性 1 人）が参加した．評価者らが評価するシステムの順番はカウンタバランスが保たれるよう考慮されており，順番による影響は評価結果には影響しないようにした．実験参加者に筆者は含まれない．対話を行う評価者と WOZ は別の部屋に分かれ，評価者は 2 人の WOZ と対話を行った．それぞれの WOZ とは 1 システムにつき 1 回ずつ，提案システムと比較システムの 3 つで計 4 対話行った．つまり，評価者は計 8 対話を行った．それぞれの対話は 15 分間行われ，評価者は対話終了後に「いい聞き役だったと思うか？」という設問に 7 段階（1：まったくそう思わない〜7：非常にそう思う）のリッカート尺度で答えた．

4.4.1　評価用システム

評価用システムは以下の 4 つである．

(1) **ルールベースシステム**　4.1 節の我々の提案方法に基づきルールにより対話制御を行うシステムである．

(2) **HMM システム**　このシステムは，対話行為がタグ付けされた聞き役対話データから学習した HMM を用い，各対話時点においてつねに最も予測確率が高い対話行為を選択する．HMM を使用する方法として，最も予測確率が高い対話行為を選択する方法と，各対話時点において，予測確率に合わせて確率的に対話行為を選択する方法が考えられる．ここで採用した，最も予測確率が高い対話行為を選択する方法では，単調な対話になる可能性がある．しかし，ユーザの振舞いは一定ではなく，ユーザの振舞いによって対話の流れも変わることから，単調な対話のみが生成されるわけではない．また，予測確率が最大の対話行為を選択する方法と確率的に選択する方法を比較すると，確率的な手法では，話の流れを致命的に悪くしてしまう対話行為が出力される可能性が出てきてしまう．本実験では，ルールベースシステムの比較対象として，致命的に話の流れが悪くなる可能性があるものは避けるのが妥当であると考え，予測確率が最大となる対話行為を選択する方法を選んだ．このシステムは，ターンごとに次のユーザの対話行為を推定する．ユーザの対話行為の確率分布に従いランダムに選択し，ユーザの対話行為がスキップ（何もしないことを表す対話行為）だった場合には，システムの対話制御を続ける．推定結果がスキップ以外だった場合には，推定を止め，対話行為を出力する．この HMM の学習には新たに 1,260 対話（67,801 個の対話行為）を収集し，使用した．新たに対話データを収集したのは，拡張した 32 種類の対話行為を用いた統計的なモデル化に多くのサンプルが必要となったからである．収集する方法は，3.1 節と同様の方法を用いた．このデータに表 6 の対話行為をアノテートしたデータを用いて 3.3.2 項の方法で HMM を学習した．ただし，状態数は，対話行為数・学習データ数に鑑み，人手で 18 個に固定し，学習した．学習した HMM は付録 A.1 のようになった．この図によれば，質問や共感の出力確率が高い状態もあるため，つねに最も予測確率の高い対話行為を選択しても，様々な対話行為が選ばれることが確認できる．

(3) **Human システム**　WOZ は自由に対話行為数と対話行為を選択し，対話することができる．本実験においては上限値を持ったシステムといえる．

(4) **ランダムシステム**　等確率（1/32 ずつ）でランダムにシステムの対話行為を生成する．1 発話における対話行為の個数は 1〜3 の中からランダムに選ばれる．

なお，このほかに，検討段階の 3 つのシステムと同時に評価実験を行ったが，今回の評価とは関係のないシステムであるため，本稿での分析からは除外している．これらのシステムを除外しても，上記のシステムのカウンタバランスが保たれるようコントロールされている．

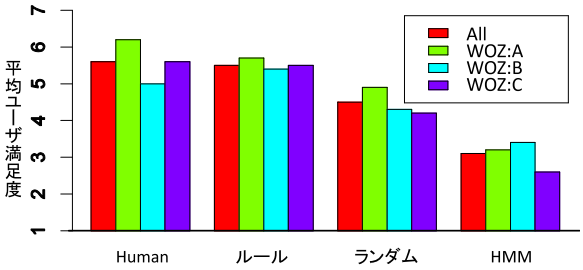

図 6　平均ユーザ満足度
Fig. 6　Averaged user satisfaction ratings.

4.5　評価結果

図 6 は「システムはいい聞き役だったと思うか？」という主観評価の結果である．提案手法のルールベースシステムは，WOZ 全員の値（図 6 内の ALL）において，Human システムとの間以外に有意な差があった（t-test: $p < 0.01$）．これは，比較用の 2 システム（ランダムシステム，HMM システム）より提案手法が優れており，人と同程度の満足度を得ることができるということを示している．つまり，予測確率が最大となる対話行為を選択する HMM をそのまま用いるより非常に高い満足度を得ているということが分

© 2012 Information Processing Society of Japan

発話文	対話行為
L: テーマは最も印象に残っている食べ物です	挨拶
S: わかりました	共感・同意
L: わかりました	繰り返し
S: 最も印象に残っている食べ物は池袋の中華料理屋で食べた蚕の串焼きですね	自己開示（sub: 事実）
L: 昨日の晩ご飯は中華料理を食べました.	自己開示（sub: 事実）
昨日の晩ご飯は池袋で食べました.	自己開示（sub: 事実）
池袋で食べることが多いですか？	質問（sub: 習慣）
S: そうですね.	共感・同意

図 7 提案方法による対話制御を用いた対話．L は聞き役，S は話し役である

Fig. 7 An excerpt of dialogue with the proposed system.

発話文	対話行為
L: 昨日の晩ご飯は焼き魚を食べましたよ	自己開示（sub: 事実）
S: 焼き魚もおいしそうですね	自己開示（sub: 評価+）
L: 私はあじの塩焼きが作れます	自己開示（sub: 事実）
S: 私もあじの塩焼き作れますよ	共感・同意
L: 一昨日の昼ご飯は焼き鮭を食べました	自己開示（sub: 事実）
S: 私は，焼き鮭はあまり好きではありません	自己開示（sub: 評価−）
L: 私は鮭フライが作れますよ	自己開示（sub: 事実）
S: 私は鮭が苦手なので鮭フライは作りません	自己開示（sub: 評価−）

図 8 HMM による対話制御を用いた対話．L は聞き役，S は話し役である

Fig. 8 An excerpt of dialogue with the HMM system.

かる．また，HMM は他のシステムに比べて最も満足度が低く，他のシステムと有意な差があった．

なお，今回の実験において WOZ:B は提案方法を使ったときの方が満足度が高かった．これに対して，Human システムを用いたときの WOZ:B の満足度は，他の WOZ に比べて劣っていた．このことは，聞き役として良い対話戦略を持っていないような聞き役には，本手法の方が有効な対話戦略として働く可能性を示しており，提案ルールが，場合によっては，人間より優れた対話戦略を持つ可能性を示している．

4.5.1 得られた対話の分析

図 7 と**図 8** は提案方法と HMM をそれぞれ対話制御として使ったときの対話例である．提案手法を用いた対話では，WOZ は自己開示の後に質問をしており，3.3.4 項の分析に基づいた生成がなされていることが分かる．このように聞き役対話の特徴的な流れの再現が，高い満足度につながったと考えられる．これに対して，HMM は学習データの中で頻出し，高い出現確率を持っている「自己開示（sub: 事実)」の対話行為を繰り返し出力している．ユーザが様々な対話行為を話しても，システムは同じ対話行為を繰り返しており単調でユーザ満足度が低くなったと考えられる．適度な自己開示はユーザの自己開示を促す効果があり，システムとユーザの社会的関係を強める効果があるが，HMM システムのような単調で過度な自己開示はユーザに「システムが積極的に聞いている」という感覚を与えなかったようである．一方，ランダムシステムは HMM より満足度が高い．ランダムシステムはすべてのタグは一様に出力しており，対話の流れなどはいっさい考慮せず生成しているが，単調さはない．そのため，単調な HMM と比較したときにより人間らしいと感じ，満足度が高くなったと思われる．

表 7 提案システムによって使用されたルールの割合

Table 7 Rates of rules used by the proposed system.

	使用割合
分析に基づくルール	82.7%
先行研究に基づくルール	5.9%
一般ルール	11.4%

表 8 それぞれのシステムの対話系列から学習した HMM 間の Kullback-Leibler 情報量

Table 8 Kullback-Leibler divergence between HMMs.

	Human	HMM	ルール	ランダム
Human システム	N/A	4.7	1.4	1.7
聞き役対話（4.4.1 項）	0.9	4.4	2.8	2.4

4.5.2 各ルールからの出力分布の分析

提案システムでは分析に基づくルールと，先行研究に基づくルールを併用して実験を行った．先行研究に基づくルールは補完的に導入したが，それらのルールが多く使用されていると，提案ルールが正当に評価されていない可能性が残ってしまう．そこで，確認のため，実際に使われたルールの内訳を精査した．その結果，**表 7** のようになった．この表から分かるとおり，システムの動作は，提案ルールからのものが大半であると確認できた．また，1 対話中に先行研究のルールからの出力は，平均 1.0 個のみ（最頻値 0）と非常にわずかであった．これらの結果から，我々の提案ルールに基づくルールが適切に評価されていることが確認でき，また，先行研究に基づくルールはごく補完的に使われていることも確認できた．

4.5.3 KL 情報量を用いた分析

表 8 は，各システムの対話行為の出現分布と人が自由に対話を行った対話（Human システムを使用して得られた対話データと，4.4.1 項の HMM システム学習に用いた聞き役対話データ）の対話行為の出現分布の近さを見た結果である．客観的指標として，Kullback-Leibler（KL）情報量を用いた．この値が低ければ，近い出力分布を持っているということを示している．

ここでの KL 情報量の算出についてであるが，まず，状態数は 6 個とし，それぞれのシステムを使用して得られた対話の対話行為列から，Speaker HMM を学習した．状態数が 6 個なのは，KL 情報量を比較する元である Human システムを使用して得られた対話データに対して，最適と

© 2012 Information Processing Society of Japan

思われる状態数を，3.3.2 項と同様に MDL 基準を用いて求めたところ，学習データが少ないことも影響し，6 個が最適となったからである．また，もう 1 つの比較元の HMM の学習データである聞き役対話は 1,260 対話と比較的多く，6 以上の状態数の HMM を学習することは可能と考えられるが，比較先である各システムの対話データが 30 対話と少ないため，比較先のデータ数に合わせて，こちらも状態数を 6 個としている．

次に，比較元の HMM の各状態と，比較先の HMM の各状態すべてのペアについて，対話行為の出力確率分布ベクトルのコサイン類似度を求め，類似度の最も高い状態のペアを対応付ける．そして，対応付けられた状態の組合せそれぞれについて，KL 情報量を以下の式を用いて計算した．

$$D(P||Q) = \sum_{i=1}^{N} P(i) \log \frac{P(i)}{Q(i)} \tag{1}$$

ここで，P は各システムを使用した対話から学習した HMM の状態の出力確率分布，Q は Human システム，もしくは，聞き役対話から学習した HMM の状態の出力確率分布，N は対話行為の数（32 個）を指す．各 HMM における 6 つの状態のそれぞれについて，対応付けられた状態との KL 情報量を求め，これらの平均をもって HMM の KL 情報量の値とした．

1 行目は Human システムとの対話から学習した HMM と，それぞれのシステムとの対話結果から学習した HMM の KL 情報量である．提案手法の方がランダム，HMM システムより Human システムとの KL 情報量が小さくなっており，提案手法は人間の対話行為の分布をうまく表現できていることを示している．

2 行目は，人間同士が自由に対話を行った際に対話データから学習した HMM と，それぞれのシステムの対話から学習した HMM の KL 情報量である．人間が自由に対話をしている状況との情報量を比較することで，本来の聞き役の対話行為の分布が再現できているかを調べることができる．その結果，提案手法の方がランダムシステムより聞き役対話データとの KL 情報量が大きく，対話行為の出力分布に関しては，人間同士の対話と比べ遠いということが分かった．ユーザ満足度（図 6）は，提案手法の方がランダムシステムより高い値を示しており，提案手法は一定の決められた対話行為の中で適切な対話行為を選択できてはいるが，今後は，より多様な対話行為を出力し，実際の人間の分布に近くなるようにしていきたい．

聞き役対話から確率を学習したにもかかわらず HMM と Human システムとの KL 情報量の値が大きい理由は，HMM が最大の出力確率の対話行為を選択していたため，一部の対話行為に偏りが生じ，結果的に，人間の分布とは遠いものになってしまったためと考えられる．

なお，人同士が自由に対話したときの聞き役対話データと Human システムとの KL 情報量は小さい．このことは，完全に自由に話したときの対話のタグの流れと，本実験でタグを自由に選びテンプレートを用いたときの対話行為の出力分布が類似していることを示している．これは，テンプレートを使うことによる自然文作成の制約は人の対話戦略を大きく変えておらず，実験内の対話は不自然な環境で行われていないことを示している．

5. おわりに

本稿では，まず，聞き役対話の特徴を分析するために，聞き役と話し役に分かれて行う対話（聞き役対話）と役割を持たない対話（雑談）を収集し，比較した．それぞれの対話タイプの対話行為の割合を比較し，対話の流れを HMM を用いて分析した．その結果，聞き役は自己開示を減らし，質問と相槌を増やすことで，話し役に自己開示や，情報提供をさせていることが確認できた．また，聞き役は自己開示をしてから質問をすることで，話し役が話しやすい対話の流れを作っていることが分かった．この特徴的な流れは，我々にとって予想ができなかった新たな発見である．続いて，得られた知見をもとに，対話制御システムをルールベースで構築し，その有効性の検証を行った．Wizard of Oz 方式により評価実験を行ったところ，提案のルールによる制御方法が人間が自由に制御を行ったときとほぼ同等の満足度を示し，提案ルールの有効性を示すことができた．

本稿の貢献は，聞き役対話の特性を対話行為レベルの分析によって明らかにし，満足度の高い聞き役対話システムの対話制御部を，分析に基づき実現したことの 2 つである．

今後の課題として，4.4.1 項で用意した HMM システムは，予測確率が最大となる対話行為を選択する方法をとっていたが，確率的に対話行為を選択する方法も考えられる．確率的に選択する方法では，多様な対話行為が出力され，提案システムの性能を上回る可能性があるため，今後検討していきたい．

また，本稿では，手作業で作成したルールによって対話制御を実現したが，Meguro らはすでに統計的モデルを用いた研究も進めている [21]．今後は，本稿で提案したルールベースによる制御と現在研究が進められている統計的手法による制御を組み合わせ，WOZ 方式で検証実験を行っていく予定である．また，完全に自動で動く対話システムを構築するために，今後は発話理解部，発話生成部の構築を目指していく予定である．

最後に，本稿では，我々は「話したい」システムと「聞いてもらいたい」システムが同一であるとしている．しかし，「システムはいい聞き役だったと思うか」という評価を上げることが「ユーザが話したいと思う」システムの実現につながるのかは明らかでない可能性がある．今後さらなる分析を行い，より話したくなり，かつ，いい聞き役と評価されるシステムに向けて，研究を進めていきたい．

© 2012 Information Processing Society of Japan

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

[1] Higashinaka, R., Miyazaki, N., Nakano, M. and Aikawa, K.: Evaluating discourse understanding in spoken dialogue systems, *ACM Trans. Speech and Language Processing* (*TSLP*), Vol.1, pp.1–20 (2004).

[2] 中野幹生：話がはずむ音声対話システム，(〈特集〉フィールドを広げる自然言語処理)，情報処理学会誌，Vol.40, No.4, pp.365–369 (1999).

[3] Stolcke, A., Ries, K., Coccaro, N., Shriberg, E., Bates, R., Jurafsky, D., Taylor, P., Martin, R., Ess-Dykema, C. and Meteer, M.: Dialogue act modeling for automatic tagging and recognition of conversational speech, *Computational Linguistics*, Vol.26, No.3, pp.339–373 (2000).

[4] 下岡和也，徳久良子，吉村貴克，星野博之，渡部生聖：音声対話ロボットのための傾聴システムの開発，人工知能学会研究会資料，SIG-SLUD 58, pp.61–66 (2010).

[5] 横山祥恵，山本大介，小林優佳，土井美和子：高齢者向け対話インタフェース—雑談継続を目的とした話題提示・傾聴の切替式対話法，情報処理学会研究報告音声言語情報処理，SLP-80, No.4, pp.1–6 (2010).

[6] Kobayashi, Y., Yamamoto, D., Koga, T., Yokoyama, S. and Doi, M.: Design targeting voice interface robot capable of active listening, *Proc. HRI*, pp.161–162 (2010).

[7] Maatman, R.M., Gratch, J. and Marsella, S.: Natural behavior of a listening agent, Lecture Notes in Computer Science, Vol.3661, pp.25–36 (2005).

[8] Hirasawa, J., Miyazaki, N., Nakano, M. and Kawabata, T.: Implementation of coordinative nodding behavior on spoken dialogue systems, *Proc. ICSLP*, pp.2347–2350 (1998).

[9] Lee, A., Shikano, K. and Kawahara, T.: Real-time word confidence scoring using local posterior probabilities on tree trellis search, *Proc. ICASSP*, Vol.1, pp.793–796 (2004).

[10] Weizenbaum, J.: ELIZA–A Computer Program For the Study of Natural Language Communication Between Man and Machine, *Comm. ACM*, Vol.9, No.1, pp.36–45 (1966).

[11] Wallace, R.S.: The Anatomy of A.L.I.C.E., *Parsing the Turing Test*, Springer (2009).

[12] Mauldin, M.L.: ChatterBots, TinyMuds, and the Turing test: Entering the Loebner Prize competition, *Proc. AAAI*, pp.16–21 (1994).

[13] Shirai, K.: Modeling of spoken dialogue with and without visual information, *Proc. ICSLP*, Vol.1, pp.188–191 (1996).

[14] 磯村直樹，鳥海不二夫，石井健一郎：HMM による非タスク指向型対話システムの評価，電子情報通信学会論文誌，Vol.J92-D, No.4, pp.542–551 (2009).

[15] Jurafsky, D., Shriberg, L. and Biasca, D.: *Switchboard SWBD-DAMSL Shallow-Discourse-Function Annotation Coders Manual* (1997).

[16] Ivey, A.E. and Ivey, M.B.: *Intentional Interviewing and Counseling: Facilitating Client Development in a Multicultural Society*, Brooks/Cole Publishing Company (2002).

[17] Altman, I. and Taylor, D.: *Social Penetration: The Development of Interpersonal Relationships*, Holt Rinehart and Winston (1973).

[18] Clark, M.S. and Reis, H.T.: Interpersonal Processes in Close Relationships, *Annual Review of Psychology*, Vol.39, pp.609–672 (1988).

[19] Higashinaka, R., Dohsaka, K. and Isozaki, H.: Effects of Self-Disclosure and Empathy in Human-Computer Dialogue, *Proc. SLT*, pp.108–112 (2008).

[20] Williams, J.D. and Young, S.: Partially observable Markov decision processes for spoken dialog systems, *Computer Speech & Language*, Vol.21, pp.393–422 (2006).

[21] Meguro, T., Higashinaka, R., Minami, Y. and Dohsaka, K.: Controlling listening-oriented dialogue using partially observable Markov decision processes, *Proc. COLING*, pp.761–769 (2010).

[22] 泉子・K・メイナード：会話分析（日英語対照研究シリーズ（2）），くろしお出版 (1993).

[23] 今石幸子：聞き手の行動—あいづちの規定条件，阪大日本語研究，Vol.5, pp.95–109 (1993).

[24] Cialdini, R.B.: *Influence: Science and Practice*, Allyn & Bacon (2000).

[25] Buck, G.: *Assessing listening*, Cambridge University Press (2001).

付　　録

A.1 4.4.1 項の HMM システムで用いた Speaker HMM

楕円は状態を表し，楕円内に，話者，出力確率が大きい順から 3 つの観測値（対話行為）と，その観測確率を示す．楕円内の先頭の番号は，便宜上付与した状態の通し番号である．通し番号 0〜8 は聞き役用の状態，9〜17 は話し役用の状態である．L は聞き役，S は話し役を指す．「自己開示_評価+」は，自己開示–評価–ポジティブの対話行為，「自己開示_評価–」は，自己開示–評価–ネガティブの対話行為を指す．状態間の遷移確率（図が煩雑になるため，0.125 以上のもののみ表示）は矢印横に示す．

© 2012 Information Processing Society of Japan

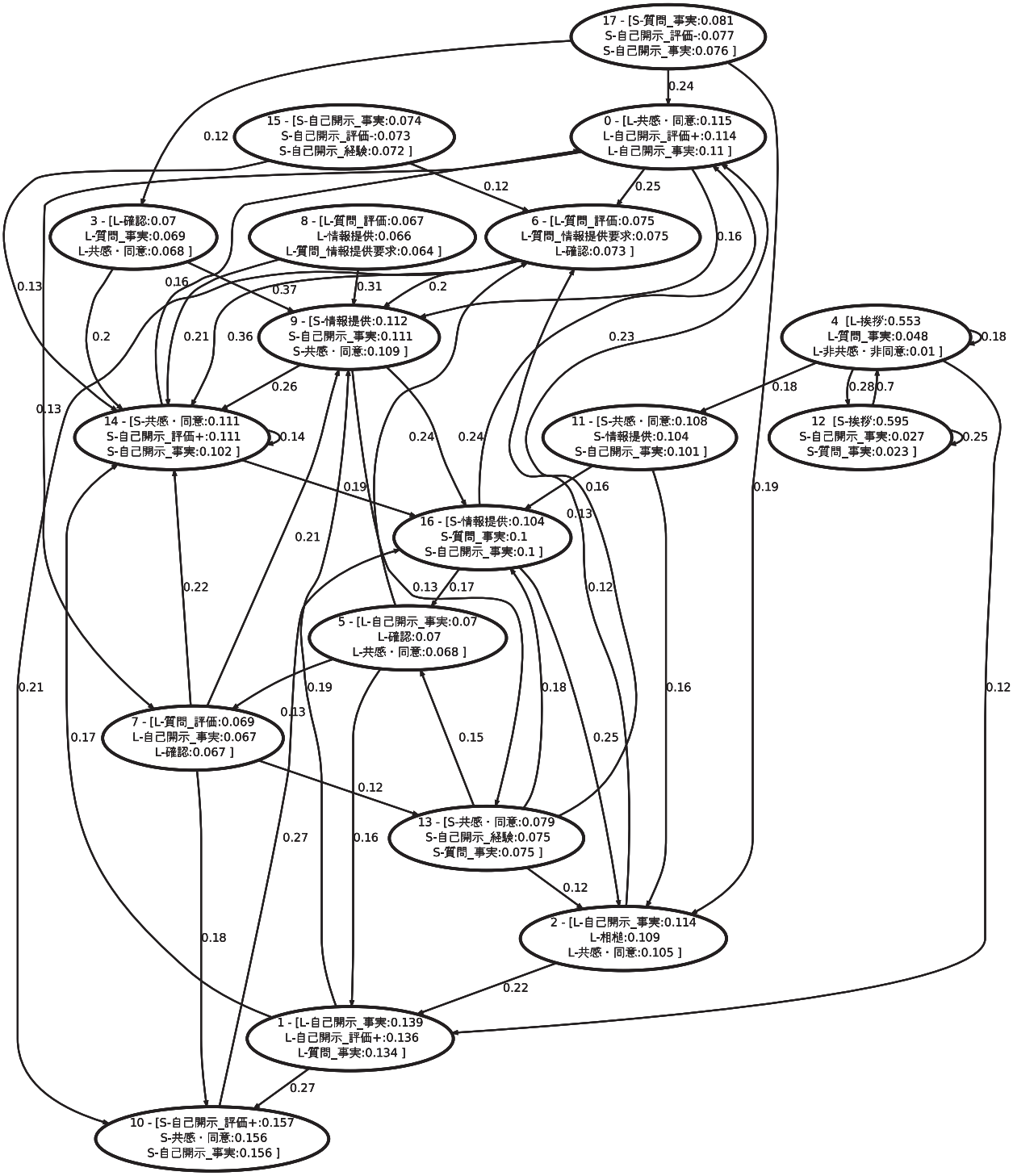

図 **A·1**　4.4.1 項の HMM システムで用いた Speaker HMM

Fig. A·1　The Speaker HMM for the HMM system in Section 4.4.1.

© 2012 Information Processing Society of Japan

目黒 豊美

2006 年東北大学工学部卒業．2008 年同大学大学院工学研究科博士課程前期修了．同年日本電信電話株式会社入社．現在，NTT コミュニケーション科学基礎研究所勤務．対話システムの研究に従事．

東中 竜一郎　（正会員）

1999 年に慶應義塾大学環境情報学部卒業，2001 年に同大学大学院政策・メディア研究科修士課程，2008 年に博士課程修了．博士（学術）．2001 年に日本電信電話株式会社入社．現在，NTT メディアインテリジェンス研究所にて勤務．音声言語メディアプロジェクトにて，質問応答システム，音声対話システムの研究に従事．言語処理学会，人工知能学会各会員．2004 年 11 月より 2006 年 3 月までシェフィールド大学客員研究員．

堂坂 浩二　（正会員）

1984 年大阪大学基礎工学部情報工学科卒業．1986 年同大学大学院博士前期課程修了，日本電信電話（株）入社．2002 年 MIT 客員研究員．2012 年より秋田県立大学システム科学技術学部教授．対話システム，ヒューマンロボットインタラクションの研究に従事．情報処理学会平成 9 年度論文賞，情報処理学会創立 45 周年記念論文優秀論文賞，人工知能学会 2010，2011 年度全国大会優秀賞受賞．言語処理学会，人工知能学会，電子情報通信学会各会員．博士（情報科学）．

南 泰浩　（正会員）

1986 年慶應義塾大学理工学部電気工学科卒業．1991 年同大学大学院博士課程修了．同年日本電信電話株式会社入社．1999～2000 年 MIT 客員研究員．現在 NTT コミュニケーション科学基礎研究所主任研究員．音声認識，対話処理の研究に従事．1993 年日本音響学会粟屋潔学術奨励賞，2003 年電子情報通信学会論文賞受賞．2005 年テレコムシステム技術賞受賞，2006 年情報処理学会創立 45 周年記念論文優秀論文賞受賞．2008 年テレコムシステム技術賞受賞．音響学会，言語処理学会，IEEE 各会員．工学博士．

© 2012 Information Processing Society of Japan