

対話履歴とユーザモデルを利用した教示発話プランニング

中野 有紀子[†] 加藤 恒 昭[†]

本論文では、ユーザとの対話の履歴とユーザモデルを参照することにより、説明の内容とその提示方法を柔軟に変化させるヒューリスティクスを提案し、さらに、それを利用したプランニングによる教示生成システムの実装について報告する。そのために、まず人間同士の教示対話の分析を行い、練習、復習、失敗時再説明の3つの対話状況で、教示者の発話内容とターン方略がどのように異なるのかを明らかにした。その結果、練習では、教示者は相手の理解を逐次確認しながら行為の指示やそれについての付加的な説明を行うが、復習では、付加説明は避け、一度に多くの指示を行うことにより、対話を短縮し、簡略化する、また、失敗時再説明では、前とは異なった内容の説明を比較的短い発話に区切って学習者に与えていることが判明した。次に、これらの結果に基づき、対話履歴とユーザモデルの情報を利用した発話内容とターン決定のためのヒューリスティクスを提案し、さらにこれらを利用したプランニングに基づく教示生成システムを実現した。教示発話のプランニングは発話内容と対話方略のプランニングに分けられ、発話内容のプランニングでは、付加情報を述べるか否かを決定する際に、また対話方略のプランニングでは、ターンの単位を決定する際にそれぞれヒューリスティクスが適用される。これらのヒューリスティクスを用いることにより、対話のプランニング途中に、状況に応じて、教示の内容とターンを柔軟に変更することができるため、より自然で多様な教示の生成が可能になる。

Instruction Generation Using Dialogue History and User Model

YUKIKO I. NAKANO[†] and TSUNEAKI KATO[†]

This paper proposes some heuristics for changing the instruction content and the way of presenting it by referring to user model and dialogue history, and reports a plan-based instruction generation system that incorporates these heuristics. First, by analyzing an instruction dialogue corpus, it is examined how the content of instructors' utterance and the turn strategy vary with three dialogue situations, namely, practice, review, and re-explanation. As the result of the corpus analysis, it is revealed that, in practice, instructors direct the hearer to do some actions and give additional information about the actions. They explain the procedures step by step with confirming the hearer's understanding. In review, in order to simplify and shorten the dialogue, additional information is omitted and several actions are directed at a time. In re-explanation, instruction that has different content from the previous one is provided with confirming the hearers' understanding. Based on these empirical results, some heuristics which utilize information in dialogue history and user model are introduced in order to select the appropriate utterance content and determine a turn-taking unit. Then, a plan-based instruction generation system is realized using these heuristics. Instruction utterance planning consists of content planning and dialogue strategy planning. In content planning, the heuristics are applied for deciding whether additional information should be provided or not. In dialogue strategy planning, the heuristics are applied for deciding a turn-taking unit. These heuristics enable changing the instruction content and the turn strategy flexibly according to the dialogue situation during planning the discourse.

1. はじめに

自然言語を用いて機械等の操作手順の教示、説明を行うシステムでは、ユーザとのやりとりの履歴やユーザの理解についてのモデルを参照しながら、説明の内

容とその提示方法を柔軟に変化させることが必要である。たとえば、ある操作をすでに十分理解しているユーザに復習として再度説明する場合には、はじめて説明する場合よりも簡略な説明にとどめる方が効率的である。また、一度目の説明が十分に理解できなかったユーザに再度説明する場合には、前とまったく同じ説明を繰り返しても効果的ではない。

説明の内容決定を柔軟に行う方法として、ユーザモ

[†] NTT 情報通信研究所
NTT Information and Communication Systems Labs.

デルに記述されたユーザの知識レベルに応じて説明内容を変化させる方法^{15),22)}, 説明の丁寧さに関するパラメータを設け, その設定によって, 補足説明等を含めた丁寧な説明と簡単な説明のどちらの方略を選択するかを決定する方法^{1),17),19)}等が提案されている。しかし, これらの方法では, 参照するユーザモデルやパラメータはあらかじめ固定されており, 教示対話の途中でシステムが自動的に内容決定に関する方略を変更することは考えられていない^{*}。そのため, すでに十分理解しているユーザに復習として再度説明する場合にも, 必要以上に丁寧な説明がされてしまう恐れがある。

また, 説明内容の提示方法については, 従来のテキスト生成の研究では, ユーザからの問合せに対する答えとして, 比較的短いテキストを一度に生成する方法がとられている^{15),17),22)}。しかし, 教示対話ではシステムが主導となって一連の長い操作手順を説明するので, この場合には, 一度にすべての指示を行うのではなく, 対話の中でユーザの理解を適宜確認しながら教示を進めるのが望ましい。Cawsey^{2),3)}はこのような教示対話の特徴を考慮し, システムの各発話の終了ごとにユーザからのあいづち (acknowledgement) や明確化質問を受け付ける方法を提案している。この方法では, 短い発話のやりとりによって教示を進行させることができるため, ユーザの理解を確認しながらの対話的な説明が可能になる。ただしこの方法では, 対話の状況に応じて, どの程度頻繁にユーザからの発話を受け付けるのかを変化させること, すなわちターン²³⁾を適切に調整することが重要である。たとえば, はじめは各指示ごとに確認を求めめる方が確実であるが, 復習の説明では確認なしにいくつかの操作をまとめて指示の方が効率的である。しかし, Cawsey^{2),3)}は, このような状況に応じたターンの調整は行っていない。

そこで本論文では, 人間同士の教示対話の分析をもとに, 対話の状況に応じた教示の発話内容とターンを決定するためのヒューリスティクスを提案し, さらにこれを利用した教示生成システムを実現する。

本論文の構成は以下のとおりである。まず次章で人間同士の教示対話を分析し, 対話の状況に応じて, (1) 教示者は説明の発話内容をどのように変化させるのか, また, (2) 対話を適切に制御するために, ターンの方略をどのように変化させているのかを明らかにする。3章では, この分析結果を基に, 発話内容とターンを決定するためのヒューリスティクスを提案し, さらに

これらを利用した教示発話プランニングメカニズムを提案する。教示発話プランニングは発話内容プランニングと対話方略プランニングに分けられる。発話内容決定のヒューリスティクスは, 発話内容プランニングにおいて, 補足的な説明を述べるか否か, すなわち説明の丁寧さを決定するために用いられる。また, ターン決定のヒューリスティクスは, 対話方略プランニングで用いられ, どの時点で発話を終了してユーザに確認を求めめるのかを決定する。さらに4章では, 本メカニズムを用いた教示生成システムによる対話例を示す。

2. 教示対話の分析

本章では実際の教示対話の分析から, 対話の状況によって教示者の説明の発話内容とターンの方略がどのように変化するのかを明らかにする。

2.1 対話データ

教示者が学習者に留守番電話機の初期設定の方法を教えるという状況設定で対話をしてもらった。対話は, (a) 学習者が教示者の電話機の様子をビデオモニタを通して見ることができるマルチモーダル状況と, (b) 音声のみを使った音声状況の2種類の状況で収録した¹¹⁾。分析した対話データは音声状況30対話, マルチモーダル状況26対話, 計56対話である。

また各対話は, 電話機の組立てから始まる一連の操作についての平均約26分の対話であるが, その一部である留守番メッセージの設定についての部分のみを分析した。留守番メッセージ設定の対話は, 最初に教示者の指示に従って学習者が設定の手順を一通り練習し, それに成功したら, 復習のためにその一連の操作を再度行い, それにも成功したらこの課題については終了というシナリオで進められる。もし練習や復習で失敗したら, もう一度やり直すこととした。

2.2 対話状況の特徴

教示対話の分析では, 対話の状況の特徴付ける要因として以下に示す課題遂行回数, 学習者の理解度の2つの要因を設定する。

課題遂行回数: その操作課題にはじめて取り組んだのか, それとも2回以上か。

学習者の理解度: 学習者が説明された操作をどの程度理解しているか。ここでは, 課題遂行の結果(成功/失敗)を理解度の指標に利用する^{☆☆}。

この課題遂行回数と学習者の理解度の2つの要因の組合せによって, 以下の3種類の対話状況を表現する。

^{*} Paris²²⁾は, テキスト生成の途中で, 熟達者用と初心者用の説明方略を切替えるメカニズムは提案しているが, そこで方略決定に用いられるユーザモデルは固定的なものである。

^{☆☆} 課題に成功した場合には十分に理解していると見なし, 失敗した場合には理解が不十分であったと見なす。理解しているのに遂行には失敗してしまうという可能性はここでは考えない。

練習 (Practice): 学習者にとってその説明を聞くのは初めてであり, 教示者も学習者がその課題について熟知していることを期待していない状況. すなわち, 課題遂行回数が0, 学習者の理解度「低」の状況.

復習 (Review): 教示者は学習者にすでにその課題について一度説明しており, 前の試行において学習者が課題遂行に成功している状況. すなわち, 課題遂行回数が1以上, 学習者の理解度「高」の状況.

失敗時再説明 (Re-explanation): 少なくとも前回の試行において学習者が課題遂行に失敗している状況. すなわち, 課題遂行回数が1以上, 学習者の理解度「低」の状況.

各対話状況のデータ数は, 練習が56回, 復習が44回, 失敗時再説明が20回であった. 以下の節では, これらの対話データを分析し, 教示者の発話内容やターンの方略がどのような要因によって決定されているのかを明らかにする.

2.3 発話内容の分析

教示者の発話内容を分析する枠組には, RST (Rhetorical Structure Theory)¹⁴⁾を用いる. RSTは, テキスト構造を核 (nucleus) と衛星 (satellites) と呼ばれる部分テキスト間の修辞関係によって表現する理論である. 核は話題の中心的部分であり, 衛星は核をサポートする付加的な情報を与える部分である. また, 核は話し手の意図として聞き手に行わせたい行為を表現し, 衛星は聞き手にその行為を動機付けるための情報を与えるものとも解釈されている²⁰⁾. たとえば, 「応答1 ボタンを押してください」と「すると応答1 ランプが点滅します」という2つの発話があった場合, 前者が核で後者が衛星である. 後者は前者の行為を遂行した結果を述べており, 前者をサポートする役割を持つ.

発話内容の分析では, この枠組みを用いて, 各行為の指示を必須の核とし, それについての付加的な情報となる発話内容を衛星とし, これらの組合せによって教示対話が構成されているととらえる. 対話データ中には, Mann¹⁴⁾や Hovy⁸⁾の分類の中で, 「結果 (Result) (行為を行った結果の状態やその意味)」, 「精緻化 (Elaboration) (行為や対象物の詳細や特徴)」, 「条件 (Condition) (行為を遂行するために要求される条件)」と呼ばれる関係によって結ばれた衛星が観察された. 以上に基づき, 発話内容の分析として, 衛星の数が対話の状況によってどのように変化するのかを調べた.

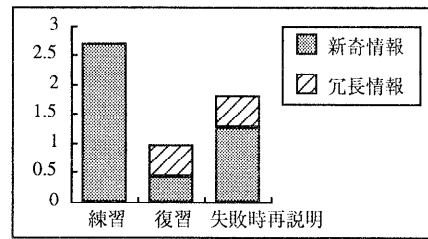


図1 付加情報の平均出現数

Fig. 1 Amount of additional information.

2.4 発話内容の分析結果

まず, 練習, 復習, 失敗時再説明の3つの対話状況によって, つまり, 課題遂行回数と学習者の理解度 (以下, 理解度) の2要因によって, 衛星 (付加的な情報) の出現数がどのように変化するかを調べた. ここでは1つの修辞関係を1つの付加情報として数えた. 各状況ごとの衛星の平均出現数を図1 (棒グラフ全体) に示す. 衛星は練習において最も多く現れ (2.70), 次が失敗時再説明 (1.80), そして復習 (0.97) の順であった*. これら3つの平均値の差は統計的に有意であった ($F(2, 117)=19.344, p<0.0001$)**.

したがって, 付加情報の量は課題遂行回数と学習者の理解度に依存しているといえる. また, シュフェ法による多重比較***を行った結果, 練習と復習の差についてのみ有意であり ($p<0.001$), 失敗時再説明と練習, および復習との差は有意ではなかった. したがって, 失敗時再説明が練習と復習のどちらにより近いのかについては, 明確な結果は得られなかった.

次に, 復習と失敗時再説明中の各衛星について, その情報がこれまでに述べられたことのない新奇な情報 (以下, 新奇情報 (Novel Information)) であるか, これまでの対話の中ですでに述べられたことのある冗長な情報 (以下, 冗長情報 (Redundant Information)) であるのかを調べた (図1). その結果, 復習では冗長情報の方が新奇情報よりも若干多かった (平均出現数

* 核の数はすべての状況で1回の教示あたり平均約6つ出現する.

** $p<0.0001$ は, 危険率, あるいは有意水準と呼ばれる値である. たとえば, 有意水準 $p<0.0001$ とは, 練習, 復習, 失敗時再説明それぞれの平均値を μ_1, μ_2, μ_3 , とした場合に, 仮説 $H_0: \mu_1 = \mu_2 = \mu_3$ が真のときそれを誤って棄却することが, 10,000回の実験のうち1回以下であるということを示す. 一般には有意水準が $p<0.05$ である場合に, 統計的に有意差があるといえる. また, $F(2, 117)$ は自由度 (2, 117) のF分布を用いて検定が行われたこと, 19.344は分散分析によって得られた値 (F比) を表す. なお, 分散分析の詳細については, Hoel⁷⁾ や Cohen⁴⁾等の参考文献を参照のこと.

*** 実験計画法において, 分散分析による主効果 (対話状況の違い全体による効果) や交互作用の検定とは別に, 処理間の差異 (各々の対話状況間の差異) を吟味するための方法.

はそれぞれ 0.55, 0.41)。しかし、この差は統計的に有意ではなかった。一方、失敗時再説明では、新奇情報の方が冗長情報よりも多く（平均出現数はそれぞれ 1.25, 0.55）、その差は統計的に有意であった ($t(38) = -2.252, p < 0.05$)。したがって、失敗時再説明では対話の履歴も教示の内容に影響を与えているといえる。

以上の結果から次のように考察できる。[考察 1] まず、はじめての説明で学習者の理解度も低い練習では、付加情報を積極的に利用した教示がなされるが、[考察 2] 復習では付加情報は省略されることが多く、教示は簡略化される。[考察 3] また、失敗時再説明では、付加情報の量については明確な結果は得られなかったが、付加情報の種類については、すでに述べられた冗長情報よりも新奇情報のほうが明らかに選好されていた。これは、失敗時再説明では、対話の履歴も付加情報の決定に影響を与えていることを示す結果である。

2.5 ターン方略の分析

本節では、教示者がどの程度の頻度で学習者に発話権を譲渡して確認を求めるといかなる要因によって決定されているのかを調べる。まず、ターンの方略を以下の 2 種類に分類した。

ターン維持方略：発話をそのまま継続し、次の発話内容を続けて述べる方略。複文中の節、あるいは並列句として表現され、その後に関連する節や句が続く。

ターン譲渡方略：発話をそこで終了し、相手（学習者）にターンをとる機会を与える方略。「～です」「～てください」等の助動詞句や終助詞が用いられる。また、文末イントネーションをとめない、その後にはポーズが置かれる⁹⁾。

たとえば、「応答 1 を押すと、応答 1 ボタンが点滅します。」という発話では、「応答 1 を押す」という行為の指示にはターン維持方略が、「応答 1 ボタンが点滅する」という付加情報にはターン譲渡方略がとられていると分析することができる。

ターン方略の分析では、発話内容が新奇情報か冗長情報であるかを新たに要因として導入し、これを「発話の冗長性」の要因と呼ぶ。したがって、ここでは、

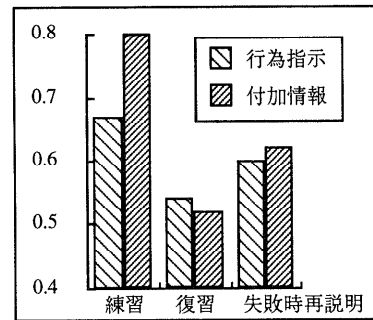


図 2 ターン譲渡方略の割合

Fig. 2 Percentage of turn release form.

課題遂行回数、理解度、発話の冗長性の 3 つの要因によってターン方略がどのように変化するかを調べる。

2.6 ターン方略の分析結果

教示中の各行為の指示（部分的な修辭関係では核にあたる）と付加情報の提示（修辭関係上の衛星にあたる）が、対話においてターン維持方略とターン譲渡方略のどちら用いて表現されているのかを調べた。図 2 に練習、復習、失敗時再説明の 3 つの状況において、行為指示と付加情報それぞれに対してターン譲渡方略が用いられた割合を示す。行為指示では、ターン譲渡方略の割合は練習で最も高く（67%）、復習で最も低かった（54%）。失敗時再説明ではその割合は 62% であった（いうまでもなく、残りはターン維持方略となる）。これらの対話状況によるターン譲渡方略の割合の差は統計的に有意であった ($F(2,117) = 6.297, p < 0.003$)。さらに、シェフェ法による多重比較を行った結果、練習と復習との差についてのみ有意であった ($p < 0.03$)。したがって、失敗時再説明が練習と復習のどちらにより近いのかについては、明確な結果は得られなかった。また、修辭関係上の衛星となる付加情報については¹⁰⁾、ターン譲渡方略は練習で最も選好され（80%）、続いて、失敗時再説明（62%）、復習（52%）の順にその割合は小さくなる（図 2）。以上のことから、行為の指示、付加情報のどちらにおいても、そのターン方略は課題遂行回数と理解度の 2 つの要因によって影響されているといえる。

さらに、これら 2 つの要因に加え、発話の冗長性の要因（その発話内容が冗長情報であるか新奇情報であるか）によっても選好されるターン方略が影響され

^{*} 本論文では、以下の理由により文末であるか否かをターン方略の分類基準として用いた。第 1 に、文の後のほうが、そうでないもの（たとえば、連用中止のようなもの）の後よりも、相手からのあいづちや明確化質問がなされることが多く⁹⁾、文末は発話権の交代（ターン）がおこる可能性の高い場所（TRP）²³⁾と見なすことができる。第 2 に、「～てください」「～か」等の文末表現をとまなう依頼や質問に対して聞き手が何も反応を返さないことは非常に少なく、したがって、このような文末表現は相手からの反応を強く要求する表現であるといえる。

¹⁰⁾ 「条件 (Condition)」の付加情報が述べられる場合には、ほぼつねに複文中の主節の前に置かれた従属節として表現されるため、対話の状況にかかわらず、つねにターン維持方略がとられるという結果になった。そこで、Condition 以外の付加情報について分析を行った。

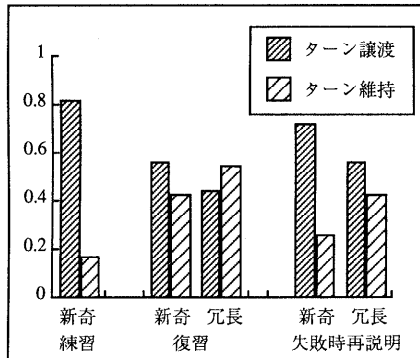


図3 新奇/冗長情報のターン方略

Fig. 3 Turn strategy for novel/redundant information.

るのかを調べた。ここでは、十分なデータ数が得られた「結果」の付加情報についてのみ分析した。図3に「結果」の新奇情報と冗長情報について、ターン譲渡方略、およびターン維持方略の割合を示す。練習をはじめの説明なのですべて新奇情報であるが、それらのうち82%にターン譲渡方略が用いられていた。これは、先述の付加情報全体の分析結果と一致する。また、失敗時再説明の新奇情報では、ターン譲渡方略が73%で、ターン維持方略は残りの27%であり、明らかにターン譲渡方略が選好されていた。復習にはターン方略についての明らかな傾向は見られなかった。このことから、特に失敗時再説明では、発話の冗長性の要因もターン方略に影響を与えていることが示唆された。このように、新奇情報は対話の状況によって、その提示方法（ターン方略）が大きく変化するのに対し、冗長情報は復習と失敗時再説明とでターン方略は大きくは異ならず、平均でほぼ半数にターン維持方略が用いられていた。Griceの会話における量の格率（The maxim of quantity）⁵⁾に違反するとされる情報として冗長な発話は、実際の会話の中では相互信念の確立のために重要な機能を持つとする主張²⁵⁾があるが、ここでの分析結果は、このような情報として冗長な発話は会話の中で確かに出現するが、その提示方法が新奇な情報とは異なっていることを示唆している。

以上の結果から次のように考察できる。[考察4] まず練習では、すべて新奇情報である行為指示や付加情報の発話内容が短いターンに区切って提示される。[考察5] 一方、復習では、複数の発話内容がまとめて伝えられる傾向が強い。[考察6] また、失敗時再説明では、練習と比べてターン譲渡方略の割合はやや減るものの、新奇な付加情報についてはターン譲渡方略が明らかに選好される。以上の分析結果より、教示者によるターン方略は、課題遂行回数、学習者の理解度、発

話の冗長性の3要因によって規定されているといえる。

2.7 教示対話の分析のまとめ

すべての分析結果をまとめると、次のことが考察できる。(1) 練習では、教示者は、行為の指示やそれについての付加的な説明を短い発話に区切って学習者に与え、相手の理解を確認しながら対話を進めていく。(2) それに対して復習では、冗長な付加説明は避け、一度に多くの内容を伝えることによって、対話全体を短縮し、簡略化している。(3) また、失敗時の再説明では、これまでとは異なった内容の説明を比較的短い発話に区切って学習者に与えることにより、学習者の誤解を修正しようとする。

3. 教示発話のプランニング

本章では、教示生成システムについて説明する。まず、本システムの基本的な処理の流れについて述べ、次に、前章の分析に基づき、課題遂行回数、学習者の理解度、発話の冗長性の3つの要因を利用した、発話内容とターン方略を決定するヒューリスティクスを提案する。

3.1 プランニングメカニズム

教示生成システムの構成図を図4に示す。尚、本論文で取り扱う範囲を太枠で示した。ユーザに説明すべき課題（例：「留守番電話のメッセージの録音の仕方を教える」）がゴールとしてシステムに入力されると、まず発話内容プランニングでは、発話内容プランオペレータを用いてプランを展開することにより、教示の発話内容を決定する。プラン展開には、Mooreら¹⁸⁾と同様の、階層的な詳細化に基づくプランニングアルゴリズムを用い、また、Cawsey³⁾と同様の、対話の進行にともなってプランを深さ優先で漸次的に展開する方法をとる。このCawseyの方法は、ユーザの反応に応じて対話の途中でプランを変更することが可能であるため、一連の操作を教える長い教示対話に適している。

また、発話内容決定プランオペレータはMoore^{17),18)}の方法を用いて記述した。各オペレータはヘッダ(header)、そのプランを実行した効果(effect)、実行のための制約条件(constraint)、必須のサブゴールである核(nucleus)、さらに核に対して何らかの修飾関係を持ち、核をサポートする付加的な情報を加えるための衛星(satellite)からなる。発話内容のプランオペレータの例を図5に示す。このオペレータは、(Command S H (Done H ?act))という行為の指示が核であり、(Persuade S H (Done H ?act))と(Achieve S (BMB S H (Competent H (Done H

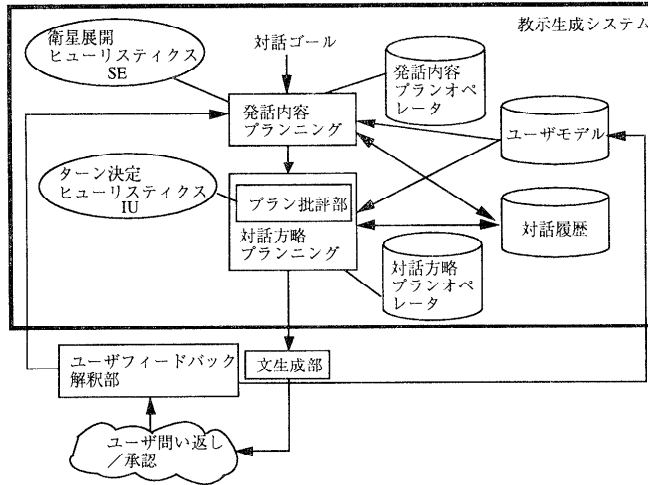


図 4 教示生成システム構成図

Fig. 4 The architecture of the instruction generation system.

HEADER	(Instruct S H ?act)
EFFECT	((BMB S H (Goal H (Done H ?act)))) (BMB S H (Know H (How-to-do ?act))))
CONSTRAINTS	((Goal S (Done H ?act)) (Step ?act ?g-goal))
NUCLEUS	(Command S H (Done H ?act))
SATELLITES	((Persuade S H (Done H ?act)) *optional*) ((Achieve S (BMB S H (Competent H (Done H ?act)))) *optional*)

図 5 発話内容プランオペレータ

Fig. 5 A content plan operator.

?act)))*が衛星である。衛星には必須なもの任意 (optional) なものがある。図 5 のオペレータの衛星はどちらも任意である。

発話内容プランニングでのプラン展開の結果、プリミティブな命題にまで展開されたものが、発話内容として順次対話方略プランニングへと出力される。ここではまず、プラン批評部で、各発話内容について、そこで発話を終了するのか、続けて発話をするのかのターンの決定を行う。これについては 3.2.3 項で詳述する。次に、ターンとしてまとめられた発話内容をどのような表現形式を用いて伝達するのかを、発話内容プランニングと同様のプランニングメカニズムによって決定する (3.2.4 項)。ここではプラン展開に対話方略プランオペレータを用いる。オペレータの例を図 6 に示す。対話方略プランオペレータはヘッダ、制約、サブゴールからなる。

こうして決定された教示発話は、文生成部に送られ

<Operator 1>	
HEADER	(Instruct-exchange ?contents)
CONSTRAINTS	((Content-type ?contents DECLARE))
SUBGOALS	((S-inform-move ?contents) (U-reply-move ?contents))
<Operator 2>	
HEADER	(Instruct-exchange ?contents)
CONSTRAINTS	((Content-type ?contents COMMAND))
SUBGOALS	((S-request-move ?contents) (U-answer-move ?contents))

図 6 対話方略プランオペレータ

Fig. 6 Dialogue control strategy plan operators.

て日本語表層表現に変換され、音声で出力される。システムの教示発話に対してユーザが承認した場合には、次の発話のプランニングが行われる。

発話内容プランニング、対話方略プランニングを通して、サブゴールは 1 つのアジェンダで管理される。これは基本的に Cawsey³⁾ と同様である。アジェンダから次に展開するサブゴールを取り出し、それを達成

* (BMB S H P) は (BEL S (MB S H P)) の略記形であり、ある命題 P が S と H の間で相互信念になっていることを S が信じている、という意味である。

するためのプランオペレータを見つけ、さらにそのプランオペレータを展開した結果のサブゴールをアジェンダの先頭に置く。こうすることにより、プランが深さ優先で展開される。サブゴールのアジェンダによる管理は、プランの柔軟な変更を可能にするため、本論文で扱うような、対話状況が変化する教示の生成に適した方法である。

3.2 ヒューリスティクスの提案

本論文では、以上に述べた従来の文生成のためのプランニングメカニズムを基本とするが、より自然で適切な教示発話を生成するために、新たに(1)発話内容プランニングにおいて、付加的な情報を述べるか否かを決定するための衛星展開ヒューリスティクス、および(2)対話方略プランニング中のプラン批評部において、発話内容を適切なターンにまとめるためのターン決定ヒューリスティクスを提案し、これらを用いた制御をプランニングメカニズム上に実現する。これらのヒューリスティクスは2章で報告した教示対話の分析結果に基づいており、課題遂行回数、学習者の理解度、発話の冗長性の3つの要因から、発話内容やターン方略についての判断を行うものである。衛星展開ヒューリスティクスについては3.2.2項で、ターン決定ヒューリスティクスについては3.2.3項で詳述する。

3.2.1 対話履歴とユーザモデル

まず、ヒューリスティクスで参照される対話履歴とユーザモデルを定義する。これは、課題遂行回数、学習者の理解度、発話の冗長性の3要因についての情報を得るためのものである。

対話履歴には、現在展開途中のプラン木と、前回までのプラン展開結果を登録する。この対話履歴中のプランから、これまでに遂行されたゴールを知ることができる。したがって、課題遂行回数についての情報を得ることができる。また、プランの葉を調べることにより、システムから出力された発話を知ることができ、それによって各発話の冗長性についての情報を得ることができる。

一方、ユーザモデルには、ユーザの知識、信念、およびゴールを列記する。このユーザモデルは発話のやりとりのたびに更新される。たとえば、システムの発話をユーザが承諾した場合には、その時点でその発話の効果がユーザモデルに付け加えられる。ユーザが課題に失敗したことが明らかになった場合には、説明した行為に関する信念がすべて取り消される。したがって、学習者の理解度はユーザモデルを参照することによって判断することができる。

3.2.2 衛星展開ヒューリスティクス

3.1節で述べたように、発話内容プランニングでは、発話内容プランオペレータを用いてプラン展開することにより、教示の発話内容を決定する。Moore^{17),18)}は任意な衛星についての展開の方針として、まったくそれらを展開しない方法と、積極的に展開する方法^{*}の2つを設けている。しかし、これらの方法のどちらを選ぶかは、あらかじめ説明開始前に固定されており、教示の過程で動的に変更することは考えていない。比較的短い説明文の生成を対象とする場合には、この方法でも十分である。しかし、ある程度長い対話を費やして行われる教示を対象とする場合には、対話の状況によって、付加情報を積極的に述べる方略をとるのか、あるいは教示の簡略化のためにできるだけ付加情報を省略する方略をとるのかを柔軟に変更する必要がある。そこで本論文では、任意である衛星の展開方略(任意である付加情報を述べるか否か)をプランニング途中で随時決定する方法をとる。そのために、対話履歴とユーザモデルを利用した以下に示す3つのヒューリスティクスを提案する。これらはプランオペレータ選択の際に適用され、衛星の数や内容において様々であるプランオペレータの候補の中から、状況に最も適したものを選択する。SE1は2.4節の[考察1]、SE2は[考察2]、SE3は[考察3]からそれぞれ導出されたものである。

< 衛星展開ヒューリスティクス >

SE1 候補となる各プランオペレータにおいて、その核の展開履歴が対話履歴中に登録されておらず、かつ、プランオペレータを実行したことによる効果がユーザモデルに登録されていない場合には、候補の中から、任意な衛星を持つオペレータを选好せよ。

SE2 候補となる各プランオペレータにおいて、その核の展開履歴が対話履歴中に登録されているか否かにかかわらず、プランオペレータの実行結果の効果がユーザモデルにすでに登録されている(あるいはユーザの既有知識として仮定できる)場合には、任意な衛星を持たないオペレータを选好せよ。

SE3 候補となる各プランオペレータにおいて、その核の展開履歴が対話履歴中にすでに登録されているが、プランオペレータの実行結果の効果がユーザモデルに登録されていない場合には、誤修正

* 積極的に展開する場合は、ユーザモデルと対話履歴を参照しながら、ユーザの知識と重ならない衛星についてのみ展開される。

を行うために必要な衛星をとまなうオペレータを選好せよ。

はじめの説明である練習では **SE1** が適用され、任意な衛星を有するプランオペレータが選好される。復習では **SE2** が適用され、任意な衛星を持たないオペレータが選好される。また失敗時再説明では **SE3** が適用され、誤解修正に必要な衛星が展開される。ただし、本論文で扱う範囲では、誤解の原因を特定することができない。そこで、「脇道にそれた質問 (digressive question)」への応答⁶⁾、および対話失敗からの復帰^{17),19)}のために提案されたヒューリスティクス—まだ述べられたことのない付加情報についてのプランを展開せよ—をここでは適用することにする。このヒューリスティクスは 2.4 節の分析結果 (失敗時再説明では、冗長な付加情報よりも新奇な付加情報の方が選好される) によってその妥当性が裏付けられている。したがって、**SE3** では、これまでに展開したことのない衛星を持つプランオペレータが優先的に展開され、その結果、今までに述べたことのない付加情報が優先的に述べられることになる。

また、一からやり直す再説明だけでなく、対話途中で学習者からの質問等によって不理解が明らかになった場合に、一部の指示についての誤解修正を行う際にも **SE3** が適用される²¹⁾。さらにこのとき、学習者の誤解がいくつか前に指示された内容に関するものであることもある。この場合には、誤解を修正した後、先ほどの話題に戻るために、**SE2** が適用されて、これまでの説明が簡潔にまとめて再度述べられる²¹⁾。この誤解修正からの復帰時の方略は、本論文で分析対象とした教示対話の中にも実際に見られた。また、これは back-on-track repetition と呼ばれ、割り込みからの復帰策 (recovery device) として、注意の焦点を戻すために有効な手段であるという報告もある¹⁶⁾。

3.2.3 ターン決定ヒューリスティクス

Cawsey³⁾ は対話方略のプランニングによって、教示を対話的に進めていく方法を提案しているが、その中では対話の状況に応じてターンを調整すること、すなわち、一度に伝達する情報量を変化させることについては考えられていない。しかし、教示においては、1つ1つユーザの確認をとりながら指示を確実に言う方が適切な場合も、複数の指示をまとめて行う効率的な方法が適切な場合もあり、対話の状況に応じて、ターンの調整を柔軟に行うことが必要である。そこで本論文では、以下に述べるターン決定のためのメカニズムとヒューリスティクスを提案する。

対話方略プランニングでは、まずターンを決定する

ためにプラン批評部が起動される。プラン批評部では、発話内容プランニングから出力された個々の発話内容に、以下に示すターン決定ヒューリスティクスが適用され、その発話内容を後続する発話内容と同一ターン中で発話すべきか否かの判断がなされる。これらのヒューリスティクスでは、衛星展開ヒューリスティクスと同様、対話履歴とユーザモデルが参照される。また、これらは 2.6 節の教示対話の分析結果に基づいており、**IU1** は 2.6 節の [考察 4] [考察 6]、**IU2** は [考察 5] から導出されたものである。

< ターン決定ヒューリスティクス >

IU1 対話履歴においてかつてその発話内容が述べられた履歴がなく、また、ユーザモデルにおいてもその発話による効果が登録されていない場合には、ユーザに発話権を譲渡する方略を選好せよ。

IU2 対話履歴においてすでにその発話内容が述べられた履歴があり、また、ユーザモデルにおいてもその発話による効果が登録されている場合には、発話をそのまま継続し、発話権を維持する方略を選好せよ。

ここで、**IU2** が選択された場合には、その発話内容は一時的に「発話内容リスト」に保持される。一方、**IU1** が選択された場合には、その発話内容とすでに発話内容リストに蓄積されている発話内容 (の列) が 1 つのターンにまとめられ、さらにそのまとまりがアジェンダの先頭に加えられて、インタラクションタイプの決定 (3.2.4 項) のためのプランニングへと進む。

練習、および失敗時再説明[☆]では **IU1** が適用され、個々の内容がターン譲渡方略を用いて生成される。一方、復習では **IU2** が適用され、各発話内容はターン維持方略がとられる。ただし復習においては、指示する行為の特徴を考慮し、いくつかの行為 (たとえば、メッセージを話すといった時間のかかる行為) の指示には **IU1** が適用される。これは、すべての内容を一度に述べてしまうという現実的には不適切な教示の生成を避けるためである。また、**IU2** は 3.2.2 項で述べた誤解修正からの復帰時にも適用され、**SE2** と組み合わせられて、談話の焦点を元に戻すための簡潔な教示が生成される。

これら対話履歴とユーザモデルの情報を利用した

☆ 2.6 節の [考察 6] で述べたように、失敗時再説明では新奇な付加情報に対してはターン譲渡方略が選好されていた。しかし、すでに述べられた発話内容に対するターン方略については、教示対話の分析からは明らかな結果は得られなかった。そこで、本論文では [考察 6] を重視し、失敗時再説明におけるすべての発話内容について、極力ターン譲渡方略を用いて丁寧な説明を行うこととした。

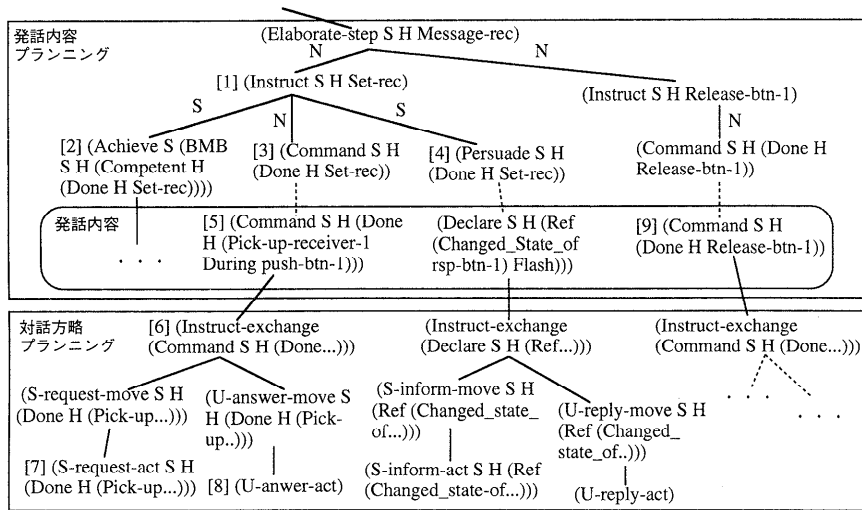


図7 教示対話プランの一部(練習)

Fig. 7 A partial instruction dialogue plan (practice).

ルールの他に、各発話内容間の修辞関係の緊密性や知識ベースにおける行為間の関連の強さは、ターン維持を愛好する要因となると考えられるが、これらの要因の考慮については今後の課題である。

3.2.4 インタラクションのタイプの決定

プラン批評部によって、一度に伝えるべき発話内容をまとめた後、このまとまりに対して対話方略プランオペレータを適用しながらプラン展開を進める(図4)。ここでは各々の発話内容のまとまりを対話の中でどのような形式で伝達するのかが決定され、その結果がシステムの出力となる。対話方略プランオペレータの例を図6に示す[☆]。たとえば、<Operator1>はシステムとユーザとの1つのやりとり(Exchange)が、システムによる叙述(S-inform-move)と、それに対するユーザの承認(U-reply-move)の2つの部分からなることを示す。発話内容(の列の最後)、すなわち、ターン譲渡を相手に知らせるきっかけとなる発話内容が、統語的、意味的に完結した叙述である場合には<Operator1>が選択され、これが依頼である場合には<Operator2>が選択される。これによって、システムの叙述(inform)に対してはユーザの応答(reply)を、依頼(request)に対しては返答(answer)を求めるプランが作られる。

4. 教示生成の例

本章では、教示生成システムの処理を実行例とともに

に示す。本システムは、音声、テキスト、図、ポインティング等を用いて留守番電話機の初期設定について説明するマルチモーダル対話システム¹⁰⁾の一部であり、その音声ガイダンス生成部として利用されている。教示生成システムの出力は文生成部に送られて日本語表層表現に変換され、画面上の図やポインティングと同期した発話が音声合成装置から出力される¹⁰⁾。また、本対話システムでは、ユーザの入力可能な応答をメニュー形式で提示し、そのメニューの中からユーザが応答を選択することによって対話が行われる^{☆☆}。

練習の対話状況のプラン展開例を図7に示す。発話内容プランニングでは、教示の内容をプラン展開によって決定するが、その際、衛星展開ヒューリスティクス(SE1~SE3)を用いてプランオペレータの優先順位を決定する。ここでは、はじめての説明なのでSE1により、任意な衛星を有するオペレータが愛好される。たとえば、録音の開始を教示する[1](Instruct S H Set-rec)では応答1ボタンを同定するための[2](Achieve S (BMB S H (Competent H (Done H Set-rec))))と、操作の結果生じる状態を述べるための[4](Persuade S H (Done H Set-rec))の2つ

^{☆☆} たとえば、ユーザが「えっ?」を選択した場合には、直前の指示が理解できなかったものと見なし、それに関するゴールについて再プランニングし、再度説明を生成する。すでに説明された操作についての質問を選択した場合には、その操作の再説明と元の話題へ復帰するための発話を生成する。また、システムとの対話は練習の対話状況からはじめられ、一通り説明が終了した後に、設定がうまくいったかどうかを尋ねるダイアログが提示される。その際、ユーザが「はい」を選択すると、復習の説明が、「いいえ」を選択すると失敗時再説明が続いて生成される。

[☆] 対話方略プランオペレータの記述には、Cawsey³⁾と同様、談話分析の枠組み²⁴⁾を利用した。exchange, move, actは談話分析の用語である。

の衛星が展開されている。このプランをさらに展開した結果出力される発話内容は、順次対話方略プランニングに送られる。対話方略プランニングでは、まずプラン批評部で、各発話内容に対してターン決定ヒューリスティクス (IU1, IU2) を適用し、一度に伝達する発話内容のまとまりを作る。ここではすべての発話内容に対して IU1 が適用され、ターン譲渡方略が選択される。続いて、プラン批評部で処理されたこれらの発話内容 (例: [6] (Instruct-exchange (Command S H (Done H (Pick-up-receiver-1 During Push-btn-1)))))) に対して、対話方略プランオペレータを用いたプランニングが行われ、対話のプランが展開される。たとえば、[6] (Instruct-exchange (Command S H (Done H (Pick-up-receiver-1 During Push-btn-1)))) は、最終的にはシステムからの行為指示 [7] (S-request-act S H (Done H (Pick-up-receiver-1 During Push-btn-1))), とそれに対するユーザの応答 [8] (U-answer-act) に展開され、これらがシステムの出力となる。図 8 <練習> にこの部分についてのシステムとユーザとの対話例を示す。

次に、すでに操作を理解しているユーザに対して、再度復習のために説明する場合の例を示す。ここでは、発話内容プランニングでは SE2 が適用され、任意な衛星を持たないオペレータの優先順位が高くなる。たとえば、図 7 で展開されていた 2 つの衛星 ([2], [4])

は、復習の説明では省略される。また、対話方略プランニングでは、「メッセージを話す」といった時間のかかる行為以外には、各発話内容に IU2 が適用されてターン維持方略が選択され、これらは発話内容リストに一時的に保持される。ここでは、[5] (Command S H (Done H (Pick-up-receiver-1 During Push-btn-1))) と、[9] (Command S H (Done H Release-btn-1)) の 2 つの発話内容に対してターン維持方略が適用され、さらに (Command S H (Done H Speak-message)) にターン譲渡方略が選択された時点でこれらがまとめられて 1 つの発話内容列となる。このまとまりに対して、さらに対話方略プランオペレータを用いたプランニングが行われる。その結果、図 8 <復習> に示すような対話が生成される。

また、練習での説明が十分に理解できていないユーザに対して再度説明する場合には、発話内容プランニングでは SE3 が適用される。ただし、3.2.2 項で述べたように、ここではこれまでに展開したことのない衛星を持つプランオペレータを優先的に展開することとした。たとえば、練習で「応答 1 ボタンを押しながらハンドセットをあげる」に対する付加情報として述べられた「応答 1 ランプが点滅する」という内容は今回は選ばれず、かわりに、「そうすることによって録音状態になる」ことが述べられる。また、対話方略プランニングでのターン決定には、IU1 が適用され、極力ターン譲渡方略が用いられる。ターン決定以降のプランニングは上記の練習や復習と同様に行われる。失敗時再説明の対話例を図 8 <失敗時再説明> に示す。

5. 結 論

本論文では、教示対話において練習、復習、失敗時再説明といった対話状況によって教示者の発話内容とターン方略が異なっていることに着目し、人間同士の教示対話の分析結果に基づき、対話の状況に応じた教示発話を生成するためのヒューリスティクスを提案した。さらに、それらを用いた制御を教示生成システム上に実現した。

衛星展開ヒューリスティクスにより、付加情報を積極的に述べるか否かのプラン展開方針がプランニング途中に決定される。これによって、教示進行中に説明内容を柔軟に変更することが可能になる。また、ターン決定ヒューリスティクスにより、対話の状況に応じてターンを適切に調整することが可能になる。さらに、これらのヒューリスティクスは、全体の復習や一からのやり直しの再説明等の長い談話の生成にも、また、部分的な誤解修正とそれからの復帰という短いサブグ

```

<練習>
...
S: 応答1ボタンを押しながらハンドセットをあげてください。
U: はい。
S: 応答1ボタンの横にあります赤い応答1ランプが点滅します。
U: はい。
S: 応答1ボタンを離してください。
U: はい。
S: 応答メッセージを喋ってください。
...
<復習>
...
S: 応答1ボタン押しながらハンドセットを取って、応答1ボタンを離して、応答メッセージを喋ってください。
U: はい。
...
<失敗時再説明>
...
S: 応答1ボタンを押しながらハンドセットをあげてください。
U: はい。
S: そうすると録音状態になります。
U: はい。
S: 応答1ボタンを離してください。
U: はい。
S: 応答メッセージを喋ってください。
...

```

図 8 システムとユーザとの対話例

Fig. 8 Example dialogues between the system and a user.

イアログの生成にも適用可能であり、多様な教示発話の生成を行うことができる。

今後の課題としては、発話内容とターン方略に影響を与える要因として、本論文で取り上げた課題遂行回数、学習者の理解度、発話の冗長性の3つの観点に加えて、他の要因も考慮することがあげられる。たとえば、発話内容の特徴による影響や^{12),13)}、人間の記憶容量の問題²⁶⁾等がある。今後、これらの観点を考慮することにより、教示生成システムのより細かな制御が可能になると考える。

謝辞 本研究の機会を与えて下さった坂間保雄知の通信処理研究部長、小原永グループリーダー、大山芳史氏、また、貴重なコメントをいただいた菊井玄一郎氏、中嶋秀治氏、長谷川隆明氏に心より感謝いたします。

参考文献

- 1) Carletta, J.: Risk-taking and Recovery in Task-Oriented Dialogue, Ph.D. Thesis, University of Edinburgh (1992).
- 2) Cawsey, A.: A Computational Model of Explanatory Discourse; Local Interactions in a Plan-Based Explanation, *Computers and Conversation*, Paul, L., Gilbert, N. and Frohlich, D. (Eds.), Academic Press (1990).
- 3) Cawsey, A.: *Explanation and Interaction: The Computer Generation of Explanatory Dialogues*, MIT Press (1993).
- 4) Cohen, P.R.: *Empirical Methods for Artificial Intelligence*, MIT Press (1995).
- 5) Grice, H.P.: Logic and Conversation, *Syntax and Semantics. Vol. 3: Speech Acts*, Cole, P. and Morgan, J.L. (Eds.), Academic Press (1975).
- 6) Haller, S.M.: Recognizing Digressive Questions During Interactive Generation, *Proc. 7th International Workshop on Natural Language Generation*, pp.181-188 (1994).
- 7) Hoel, P.G.: *Elementary Statistics, 4th edition*, John Wiley & Sons (1976). 浅井 晃, 村上正康 (訳), 初等統計学, 培風館 (1985).
- 8) Hovy, E.H.: Automated Discourse Generation Using Discourse Structure Relations, *Artificial Intelligence*, Vol.63, pp.341-385 (1993).
- 9) 泉子 K. メイナード: 会話分析, くろしお出版 (1993).
- 10) Kato, T., Nakano, Y.I., Nakajima, H. and Hasegawa, T.: Interactive Multimodal Explanations and their Temporal Coordination, *12th European Conference on Artificial Intelligence (ECAI-96)*, pp.261-265, John Wiley & Sons (1996).
- 11) 加藤恒昭, 中野有紀子: マルチモーダル対話における参照物同定要求の意味内容と形式, 人工知能学会言語・音声理解と対話処理研究会 SIG-SLUD-9501-1, pp.1-8 (1995).
- 12) Linden, K.U.: Generating Precondition Expressions in Instructional Text, *Proc. 32th Annual Meeting of the Association for Computational Linguistics*, pp.42-49 (1994).
- 13) Linden, K.U. and Martin, J.H.: Expressing Rhetorical Relations in Instructional Text: A Case Study of the Purpose Relation, *Computational Linguistics*, Vol.21, No.1, pp.29-57 (1995).
- 14) Mann, W.C. and Thompson, S.A.: Rhetorical Structure Theory: A Theory of Text Organization, Technical Report ISI/RS-87-190, USC/ISI (1987).
- 15) Mittal, V.O. and Paris, C.L.: Generating Natural Language Descriptions with Examples: Differences between Introductory and Advanced Texts, *Proc. 11th National Conference on Artificial Intelligence*, pp.271-276 (1993).
- 16) Mooney, D.J., Carberry, S. and McCoy, K.F.: Capturing high-level structure of naturally occurring extended explanations using bottom-up strategies, *Computational Intelligence*, Vol.7, pp.334-356 (1991).
- 17) Moore, J.D.: *Participating in Explanatory Dialogues: Interpreting and Responding to Questions in Context*, MIT Press (1995).
- 18) Moore, J.D. and Paris, C.L.: Planning Text for Advisory Dialogues, *Proc. 27th Annual Meeting of the Association for Computational Linguistics* (1989).
- 19) Moore, J.D. and Swartout, W.R.: A Reactive Approach to Explanation, *Proc. 11th International Joint Conference on Artificial Intelligence*, pp.1504-1510 (1989).
- 20) Moser, M. and Moore, J.D.: Toward a Synthesis of Two Accounts of Discourse Structure, *Computational Linguistics*, Vol.22, No.3, pp.409-420 (1996).
- 21) 中野有紀子, 加藤恒昭: 教示対話における誤解の修正—対話履歴を利用した修正対話の生成, 1996年人工知能学会全国大会 (1996).
- 22) Paris, C.L.: The role of the user's domain knowledge in generation, *Computational Intelligence*, Vol.7, No.2, pp.71-93 (1991).
- 23) Sacks, H., Schegloff, E.A. and Jefferson, G.: A Simplest Systematics for the Organization of Turn-taking in Conversation, *Language*, Vol.50, No.4, pp.696-735 (1974).
- 24) Stenström, A.: *An Introduction to Spoken Interaction*, Longman (1994).

- 25) Walker, M.A.: Redundancy in Collaborative Dialogue, *Proc. 14th International Conference on Computational Linguistics*, pp.345-351 (1992).
- 26) Walker, M.A. and Rambow, O.: The Role of Cognitive Modeling in Achieving Communicative Intentions, *Proc. 7th International Workshop on Natural Language Generation* (1994).
(平成9年5月7日受付)
(平成9年9月10日採録)



中野有紀子 (正会員)

1965年生。1988年東京女子大学文理学部心理学科卒業。1990年東京大学大学院教育学研究科教育心理学専攻修士課程修了。同年、日本電信電話(株)入社。現在、NTT情報通信研究所知的通信処理研究部勤務。自然言語処理、対話処理、心理言語学の研究に従事。人工知能学会、言語処理学会、ACL各会員。



加藤 恒昭 (正会員)

1959年生。1981年東京工業大学電気電子工学科卒業。1983年東京工業大学総合理工学研究科電子システム専攻修士課程修了。同年、日本電信電話公社(現NTT)横須賀電気通信研究所に入所。自然言語処理、対話処理、マルチモーダルコミュニケーションに関する研究に従事。現在、NTT情報通信研究所知的通信処理研究部主幹研究員。工学博士。電子情報通信学会、人工知能学会、言語処理学会、ACL各会員。