

バス運行情報案内システムにおけるユーザモデルを用いた適応的応答の生成

Generation of Cooperative Responses using User Model in Spoken Dialogue System

駒谷 和範

上野 晋一

河原 達也

奥乃 博

Kazunori Komatani

Shinichi Ueno

Tatsuya Kawahara

Hiroshi Okuno

1 はじめに

現状の音声対話システムの問題点の一つに、どのようなユーザに対しても画一的に応答が行われることが挙げられる。音声対話システムにおける協調的な対話の研究も行われている [1] が、生成される発話がどのようなユーザに対しても有用であるわけではない。そこで本研究では、音声対話システムにユーザモデルを導入し、ユーザに適応的な応答を生成する。

ユーザモデルの研究は主として自然言語対話システムにおいてユーザの知識に重点をおいて行われてきた [2]。これに対して音声対話システムでは、バリエーション、応答までの時間など、テキストには含まれないがユーザの特徴を反映する情報が多く存在する。本研究ではこれらの特徴を考慮に入れながら、3種類のユーザモデル(システムに対する習熟度、タスクドメインに関する知識レベル、性急度)を導入し、適応的な対話制御・応答生成の実現を目指す。現在開発中の京都市バス運行情報案内システム [3] による対話データを用いて、ユーザモデルの判別に有効な特徴と規則について決定木学習を行った。

2 ユーザモデルの分類と対話戦略

京都市バス運行情報案内システム [3] は、指定した停留所の何駅手前までバスが接近しているかを知らせるシステムである¹。本システムは電話で利用でき、分刻みで変わるバスの情報をどこからでも手軽に知ることができる。ユーザは、バスの系統番号・乗車場所・降車場所を音声で入力し、バスの接近情報を得る。地名は、バス停の名前そのもの以外にも、そのバス停が最寄りとなる名所・施設名でも入力可能である。ユーザは任意の項目を入力できる。項目に不足があった場合には、システムからプロンプトを行い入力を促す。対話例を図1に示す。

このシステムに対して、ユーザモデルを導入する。処理の流れを図2に示す。ユーザモデルとして、具体的には以下の3つの分類を用いる。

1. システムに対する習熟度

音声対話システムに対するユーザの慣れを表す。判別には、一回の発話で入力されたスロットの

S: 乗車駅、降車駅、系統番号をおっしゃってください
 U: 四条河原町
 S: 四条河原町からバスに乗るのですか
 U: はい
 S: どの停留所でバスを降りますか
 U: 嵐山です
 S: 四条河原町から嵐山までよろしいですか
 U: はい
 S: 11 系統の嵐山、山越中町行きのバスは、
 2 駅手前の三条京阪前を出発しています

図 1: バス運行情報案内システムの対話例

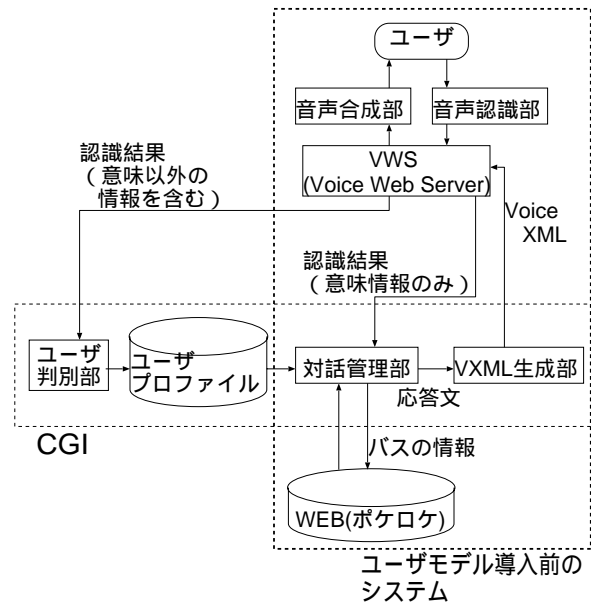


図 2: ユーザモデルのバス案内システムへの導入

数、バリエーションの有無、沈黙の有無などを特徴として用いる。習熟度の低いユーザに対しては、単語発話を誘導したり、質問時に可能な選択肢を提示するなどの対話戦略をとる。

2. タスクドメインに関する知識レベル

知識レベルの低いユーザに対しては、前提となる知識を明示したりする。バスの乗車場所・降車場所が正式なバス停名で指定されたかなどを用いる。

3. 性急度

性急度の高いユーザに対しては提示する情報を最低限に絞る。判別にはバリエーションの有無、沈黙の有無、発話継続時間を用いる。

⁰ 京都大学 大学院 情報学研究科 知能情報学専攻

¹ (075) 326-3116

表 1: 対話データにおけるユーザの分類内訳

	習熟度		知識レベル		性急度	
	発話数	割合 (%)	発話数	割合 (%)	発話数	割合 (%)
低い	743	49.8	275	18.4	421	28.2
不明	253	17.0	808	54.2	932	62.5
高い	496	33.2	409	27.4	139	9.3
合計	1492	100.0	1492	100.0	1492	100.0

表 2: ユーザモデルの判別精度 (%)

	1	2	3	4
習熟度	91.0	80.8	75.3	83.6
知識レベル	91.7	73.9	63.7	-
性急度	90.7	74.9	73.7	90.9

3 決定木によるユーザの判別実験

バス運行情報案内システムの実際の対話データを用いて、ユーザモデル判別の決定木学習を行った。学習には決定木学習アルゴリズム C5.0 を用いた。本実験で用いた対話データは、2001 年 12 月 10 日から 2002 年 5 月 10 日の間に収集されたものである。電話 (コール) 回数は 215、それに含まれるユーザの合計発話回数は 1492 である。これらの各発話に対して、システムに対する習熟度、タスクメインに関する知識レベル、性急度の 3 点に関して、「高い」「不明」「低い」の 3 段階のタグを手で付与し、正解とした。表 1 はこれらの回数の集計結果である。ユーザモデルの判別はユーザの発話ごとに毎回行った。

学習された決定木の判別精度を表 2 に示す。表 2 の 1. ~ 4. は以下の実験条件を指す。また図 3 に 3. で学習された習熟度の決定木の例を示す。

1. 全データを学習データとしたときの判別精度 (closed test)
2. 全発話をランダムに 10 個のブロックに分け、10-fold cross validation を行った結果
3. 発話ではなくコールをランダムに 10 個のブロックに分け、10-fold cross validation を行った結果
4. 対話戦略に影響を与える違い (習熟度における「低い」とそれ以外など) のみに着目した判別結果 (実験条件は 3. と同様)

4 システムの動作例

以上で述べたユーザモデルを実装した後の動作例を図 4 に示す。最初の発話で入力された項目数が 1 つであることなどから、決定木によりユーザの習熟度が低いと判別され、入力方法に関して説明が付加される。また性急度が低いことから次善案が提示される。

これらは協調的な応答 [1] と呼ばれるものだが、ユーザによっては冗長な場合がある。提案するユーザモデルを導入することにより、これらをユーザに応じて生成することが可能となる。

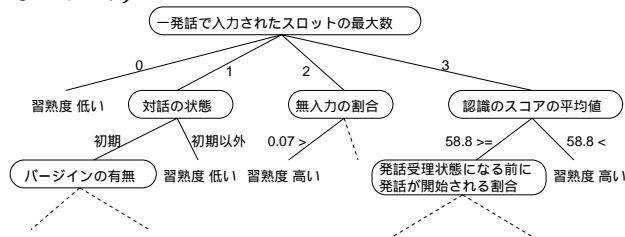


図 3: 習熟度判別の決定木の例

S: 利用するバス停、系統番号をどうぞ
 U: 円山公園からバスに乗りたいんですが
 (習熟度:低 性急度:低 の場合)
 S: どこでバスを降りますか
バス停名か最寄りの有名施設を教えてください
 (習熟度:低 説明付加)
 U: 東福寺です
 S: 円山公園から東福寺までですね
 U: はい
 (知識レベル:低 の場合)
 S: 円山公園ですと、バス停の名前は祇園になります
 (知識レベル:低 説明付加)
 202 系統の九条車庫・西大路九条行きのバスは、
 2 つ手前の東山三条を出発しました。
同様の区間を走る 207 系統の東福寺・
九条車庫行きのバスは 3 つ手前の
 四条西洞院のバス停を出発しました。
 (知識レベル:低 説明付加)
 (性急度:低 次善案の提示)

図 4: ユーザモデル導入後の動作例

5 おわりに

本研究では、ユーザに適応した対話を行うために、音声対話システムへのユーザモデルの導入を行った。京都市バス運行情報案内システムにおいて有効なユーザモデルについて検討し、システムに対する習熟度、タスクメインに関する知識レベル、性急度の 3 つのユーザモデルを導入した。対話中のユーザの発話内容から得られる特徴と音声の特徴から、ユーザモデル判別の決定木の学習を行い、その結果に応じた応答が生成されるようにシステムに実装した。

また実際の対話データを用いて決定木学習を行った。その結果、習熟度については 83.6% の精度で判別できた。

参考文献

- [1] David Sadek. Design considerations on dialogue systems: From theory to technology -the case of artimiss-. In *Proc. ESCA workshop on Interactive Dialogue in Multi-Modal Systems*, 1999.
- [2] 熊本忠彦. 自然言語対話システムにおける協調的応答の生成. *人工知能学会論文誌*, Vol. 14, No. 1, pp. 3-10, 1999.
- [3] 安達史博, 河原達也, 奥乃博, 岡本隆志, 中嶋宏. VoiceXML の動的生成に基づく自然言語音声対話システム. *情報処理学会研究報告*, SLP-40-23, HI-97-23, 2002.