

# バス運行情報案内システムにおける適応的な対話管理を行う ユーザモデルの評価

駒谷 和範      上野 晋一      河原 達也      奥乃 博

京都大学大学院 情報学研究科 知能情報学専攻

komatani@kuis.kyoto-u.ac.jp

## 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応答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実現する。

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、バーサイン、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。我々は京都市バス運行情報案内システム [3] において、システムに対する習熟度、ド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提案し、音声や言語、対話レベルなどの種々の特徴から決定木学習を行うことでユーザモデルを学習した [4]。これらをシステムに実装し、より大規模な評価実験を行った結果を報告する。

## 2 京都市バス運行情報案内システム と導入したユーザモデル

京都市バス運行情報案内システム [3] は、指定した停留所のいくつ手前の停留所までバスが接近しているかを知らせるシステムである<sup>1</sup>。本システムは携帯電話からも利用でき、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、バスの乗車場所・降車場所・系統番号を音声で入力し、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、項目に不足があった場合には、システムからプロンプト

<sup>1</sup>(075) 326-3116

---

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: どこで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

を行い入力を促す。対話例を図 1 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。

- システムに対する習熟度 (【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数、バーサイン (システム発話中のユーザの割り込み) の有無、沈黙の有無などを特徴として用いる。
- ドメインに関する知識レベル (【知識レベル】)  
 ユーザの当該ドメインに対する知識レベルにより提示する情報を変える。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。
- 性急度 (【性急度】)  
 音声によるコミュニケーションでは、他の手段 (ブラウジング等) に比べて情報提供への切迫性が大きい場合が多い。そこで、性急度というユーザモデルを導入する。判別にはバーサインの有無、沈黙の有無、発話継続時間を用いる。

電話により得られた対話データを用いて、これらのユーザモデルを判別するため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。電話 (コール) 回数は 215, そこに含まれるユーザの合

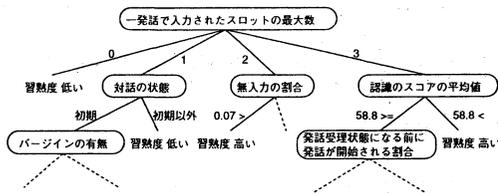


図 2: 習熟度判別の決定木の例

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」「ドメインに関する知識レベル」「性急度」の 3 点に関して人手で正解を与え、決定木学習を行った。得られた決定木のうち、習熟度に関するものの一部を図 2 に示す [4]。

### 3 ユーザモデルを用いた応答生成

#### 3.1 音声対話システムにおける協調的応答

音声対話システムにおける協調的な応答には以下が挙げられる。

- 補完的な応答 (Completion answers)  
ユーザの要求以上に情報を含めて応答を行う。しかしどのような情報をどのような時に応答に含めるかの検討が必要である。
- 矯正的な応答 (Corrective answers)  
検索に失敗した場合や、ユーザの発話に誤りが含まれている場合にその理由を伝える。
- 提案的な応答 (Suggestive answers)  
検索結果の内容がユーザにとって否定的なものであっても、関連のある情報を提供する。
- 条件を明示する応答 (Conditional answers)  
検索結果にユーザが指定した条件以外の制約がある場合に、それを提示する。
- 内包的な応答 (Intensional answers)  
検索結果が多い場合、内包的な表現としてユーザに提示する。

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。

#### 3.2 ユーザモデルを用いた応答生成戦略

習熟度、知識レベル、性急度の 3 つのユーザモデルを用いて、ユーザに適応した応答を生成する戦略について述べる。基本的に対話は混合主導対話 (mixed initiative dialogue) で行い、ユーザに自由な発話を許しながら、システムは必要に応じて自発的に質問・確認を行う。ユーザモデルを用いることで、システムのプロンプト生成を次の 2 点で変化させる。

- 対話のフロー
- プロンプト内に含める情報量

対話のフローは、習熟度および性急度により変化させる。習熟度が高い場合には、対話をユーザ主導で行い、基本的にユーザに自由に発話させる。習熟度が低いことが検知された場合には、システムが主導権を取り、システム側からその時に必要な項目を順番に尋ねていく戦略に変更する。また性急度が低い場合には、入力された内容を確認してから検索を行い結果を出力するが、性急度が高い場合には、必要な項目が入力されればそのまま検索する。

次にプロンプト内に含める情報について述べる。[1] で定義されている協調的応答では、ユーザへ情報を提示する際にさまざまな情報を付加することが検討されている。システムのプロンプトに含まれる情報は次の 2 種類に大別できる。

1. 対話管理情報
2. ドメイン情報

対話管理情報は、対話の進め方についてユーザに提示する情報である。すなわち、確認を行う際に「はい、またはいいえで答えてください。」のように答え方を指示したり、対話のフローをシステム主導に切り替える際に「順におうかがいします。」というように、これからの対話の流れに関する情報をユーザに提示したりする。これらは「システムに対する習熟度」に対応する内容であるため、習熟度が低い場合にはこれらの内容をシステムの応答に付加し、習熟度が高い場合には省略する。

ドメイン情報は、ユーザモデルの「ドメインに関する知識レベル」に応じて付加/削除する。つまり、ドメインに関する知識レベルの低いユーザに対しては、最寄りのバス停名の説明などを付加するが、別ルートの提案など複雑な内容は提示しない。

このように、対話の全ての場面のプロンプトに対して【習熟度】【知識レベル】により決定される発話内容

【習熟度：低】【知識レベル：低】【性急度：低】の場合

S: 利用するバス停、系統番号をどうぞ

U: 円山公園

S: 順におかがいします。  
 (【習熟度：低】により対話の流れの説明を付加)  
 乗車場所は円山公園でよろしいですか。  
 (【性急度：低】により対話のフローとして確認を行う)  
 はい、またはいいえで答えてください。  
 (【習熟度：低】により答え方の説明を付加)

U: はい

S: 降車場所をバス停名、または最寄りの有名施設を教えてください  
 (【習熟度：低】により答え方の説明を付加)

U: 東福寺です

S: 降車場所は東福寺でよろしいですか。  
 はい、またはいいえで答えてください。

U: はい

S: 検索結果は次のようになります。  
 (【習熟度：低】により対話の流れの説明を付加)  
 円山公園の最寄りのバス停は祇園になります。  
 (【知識レベル：低】によりドメイン情報を付加)  
 202 系統の九条車庫・西大路九条行きバスは、  
 2 分手前の東山三条を出発しました。  
 ...

図 3: ユーザモデル導入後の動作例

を用意し、それを【性急度】により取捨選択して出力する。つまり、性急度の低いユーザには協調的応答として、以上で述べた内容をプロンプトに追加し、ユーザへ出力する。一方、性急度の高いユーザにはこれらの情報を省略し、対話が冗長となることを防ぐ。

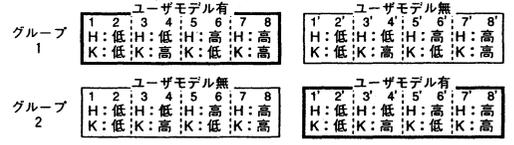
以上を実装した後の動作例を図 3 に示す。最初の発話で入力された項目数が 1 つであることなどから、決定木によりユーザの習熟度が低いと判別される。これにより、対話のフローがシステム主導へと変更される。同様に決定木により性急度が低いことが検知されているため、対話の流れの説明や答え方の説明が付加される。これらは性急度が高い場合にはいずれも省略される。

## 4 ユーザモデルを実装したシステムによる評価実験

ユーザモデルを実装したシステムを用いて、被験者による実験を行った。被験者は本システムを使用したことのない本学の学生 20 名である。実験は研究室で行い、音声入力には接話マイクを用いた。

### 4.1 実験手順

まず被験者にシステムの概要について説明し、実験条件やシナリオを書いた紙を渡した。8 つのシナリオを 2 セット用意し、ユーザモデルを導入していないシ



(ただし、H: 性急度, K: 知識レベル)

図 4: 実験の順序

ステムと導入後のシステムに対して、シナリオに従いバスの運行情報を検索してもらった。各シナリオは、例えば「あなたは京都に観光に来ています。銀閣寺を訪れたあと、円山公園に行こうと思います。このような状況を想定して実験してください。」のようなもので、実際に入力すべき具体的なバス停名やバスの系統番号などは指定していない。また、性急度を変化させる場合には「携帯電話の通話料を安くすませたいので、できるだけ急いでください。」のような条件をつけた。知識レベルの低い状況は、京都市内でも本学から遠い場所を設定することで作り出すようにした。

アンケート項目には実験前にあらかじめ目を通しておいてもらい、一方のシステムを使い終わるごとに記入してもらった。これはユーザの印象が混同されることを防いだり、ユーザの記憶の負荷を軽減するためである [5]。アンケートは、「対話がうまくいかないときに、わかりやすく誘導してくれましたか?」のような内容の 8 項目である。それぞれの項目に対して、例えば「非常にわかりにくかった」から「非常にわかりやすかった」までのような 7 段階のいずれかをチェックしてもらった。

さらに、実験中にシステムから聞いた情報(乗車するバス停名、系統番号、到着までどれくらいかかるか)を紙に記入してもらうようにした。これにより、実験条件を実際の使用状況に近づけている。

### 4.2 実験計画

シナリオの順序は図 4 のように設定した。被験者を 2 つのグループに分け、半数は「ユーザモデル有り → ユーザモデル無し」の順にシステムを使用し、半数は逆の順序でシステムを使用した。使用するシナリオの集合は、各グループの半数ずつで前半と後半を入れ替えた。

ユーザの主観的な指標として、各システムに対するアンケート結果を利用する。また対話時間やターン数、音声認識誤りの回数、パージインの回数などの客

表 1: 全体の実験結果

		duration	turn	error
グループ 1	UM 有	7149	513	61
(UM 有 → UM 無)	UM 無	4351	411	23
グループ 2	UM 無	6318	569	50
(UM 無 → UM 有)	UM 有	5166	392	23

duration: 合計対話時間 (秒)  
turn: 合計ターン数, error: 音声認識誤り回数

観的な指標 [6] もログとして記録しておく。

### 4.3 実験結果と考察

実験結果をグループごとに分割したものを表 1 に示す。グループ 1 は「ユーザモデル有 → ユーザモデル無」の順にシステムを使用し、グループ 2 はその逆の順序で使用した。また、ユーザごとのアンケート結果の概略を表 2 に示す。表 2 ではアンケートの 8 項目のうち、後半で使用したシステムの方が主観的な評価が高かった項目の数を 2 列目に示している。

ユーザは本システムを使用した経験がないため、実験の間にシステムに対して急速に慣れていく。このことから表 1 のように、グループ 1、グループ 2 の双方で、後半に使用したシステムでの合計対話時間が短くなっている。同様の理由で、合計ターン数や音声認識誤り回数も削減されている。表 2 に示されるユーザの主観的な評価の合計も、どちらのグループでも後半に使用したシステムの方が良くなっている。音声認識誤り回数も後半のシステムでは減少している。ユーザごとに詳しく見ると、音声認識誤りが多いユーザは主観的な評価も悪くなる傾向にある。

現在実験結果に対してより詳しい分析・考察を進めているが、今回得られた知見の一つとして、システムの操作方法を覚えたユーザは、システムの対応が変化することを好まないという点があげられる。このことは、ユーザモデルを導入した際にユーザの主観的な評価があまり向上していない理由の一つであると考えられる。使用方法を覚えて習熟度が向上したユーザに対しては動作を簡素化し、習熟度の低いユーザに対しては丁寧に応答を行うような設計が必要である。

## 5 おわりに

京都市バス運行情報案内システムに対してユーザモデルを実装し、システムを使用したことのないユーザ

表 2: アンケート結果と音声認識誤り回数

グループ 1 (UM 有 → UM 無)				グループ 2 (UM 無 → UM 有)			
user	eval	A	B	user	eval	B	A
#1	+1	3	5	#11	+3	4	0
#2	+2	17	4	#12	0	14	5
#3	+4	6	1	#13	+1	2	2
#4	+4	5	3	#14	-4	7	6
#5	+2	4	6	#15	+1	0	0
#6	-4	4	0	#16	+3	2	3
#7	-1	3	2	#17	-4	7	1
#8	+8	13	2	#18	+1	2	4
#9	0	2	0	#19	+5	10	2
#10	+6	4	0	#20	-2	2	0
計	+22	61	23	計	+4	50	23

eval: 後半に使ったシステムの方が評価が高い項目数  
A: UM 有の場合の音声認識誤り回数  
B: UM 無の場合の音声認識誤り回数

20 名による評価実験を行った。提案したユーザモデルはシステムに対する習熟度、ドメインに関する知識レベル、性急度の 3 つである。これによりユーザに適した対話管理を行う。

評価実験により、現段階の設計では、音声認識誤り回数がユーザの主観的な評価に影響を与えていることを確認した。またシステムの応答生成の設計がユーザの評価に大きな影響を与えるという傾向が得られた。

今後さらに詳しい分析・考察を行い、得られた知見に基づき実験条件やシステムの対話戦略を改善して、再度実験を行う予定である。

## 参考文献

- [1] Sadek, D.: Design Considerations on Dialogue Systems: From Theory to Technology -The Case of Artemis-, *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).
- [4] 上野晋一, 駒谷和範, 河原達也, 奥乃博: 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, *情報処理学会研究報告*, SLP-42-2 (2002).
- [5] Over, P.: TREC-7 Interactive Track Report, *Proc. of the 7th Text REtrieval Conference (TREC7)* (1999).
- [6] Möller, S.: A New Taxonomy for the Quality of Telephone Services Based on Spoken Dialogue Systems, *Proc. of 3rd SIGdial Workshop on Discourse and Dialogue* (2002).