

# K-50 移植性の高い音声対話システムにおける対話戦略設計ツールの開発

## Development of Dialogue Strategy Design Tool for Highly Portable Spoken Dialogue Systems

小暮 悟<sup>†</sup>  
Satoru Kogure

伊藤 敏彦<sup>‡</sup>  
Toshihiko Itoh

中川 聖一<sup>†</sup>  
Seiichi Nakagawa

### 1. はじめに

近年、音声対話システムの研究が盛んに行なわれている。最近では、システムの音声認識誤りに対するロバストネスやユーザビリティに関する研究に加えて、システムのドメイン・タスク適用能力の向上を焦点とした研究も多い。我々もこれまでに、データベース検索に焦点を絞った、移植性の高いデータベース用音声対話システムを開発してきた。今回、その対話管理部の動作、すなわち対話戦略を GUI で編集できるシステムを開発した。

### 2. 移植性の高い音声対話システム

我々が開発している移植性の高い音声対話システム [1, 2, 3] についてまず述べる。本システムは、あるドメイン・タスクの対話例と読み込み可能なデータベースを用意する事で比較的容易にドメイン・タスクの適用が出来る。これまで、富士山観光案内、文献検索、ホテル検索等のドメイン・タスクを実際に適用した。ここでは、ホテル検索を適用した時のシステムの対話例を図 1 に示す。

```

システム：ホテル検索システムです。ご用件をどうぞ
ユーザー：品川区にあるホテルを教えてください
システム：2件見つかりました
===検索結果===
1番. ホテルワトソン,
    品川区上大崎2-26-5, S 9000円,
    D 15500円, T 15500, 17000, 20000円
2番. 東興ホテル(東京都),
    品川区西五反田2-6-8,
    S 9400円, D 15000円, T 16600, 18000円

システム：検索結果は以上です

ユーザー：2番のホテルの詳細をお願いします
システム：2番の詳細を表示します
2番の詳細情報：
    ホテル名：東興ホテル(東京都)
    住所：品川区西五反田2-6-8
    アクセス：JR山手線五反田駅から徒歩1分,
    都営浅草線五反田駅から徒歩2分,
    東急池上線五反田駅から徒歩3分
シングル料金：9400円
ツイン料金：16600, 18000円
ダブル料金：15000円
    
```

図 1: システムの対話例 (ホテル検索)

#### 2.1 対話管理部 [2]

以下に、本システムにおける 9 つの対話状態を示す。

- 初期状態：対話の開始時に遷移する。

<sup>†</sup>豊橋技術科学大学情報工工学系  
<sup>‡</sup>静岡大学情報学部

- 検索条件入力状態：検索条件を入力する。入力ユーザ主導である。
- システム主導用検索条件入力状態：検索条件を入力する。入力する検索キーの種類を逐次システムが監視しプロンプトを出力する (入力はシステム主導)。
- 検索実行状態：検索条件入力状態で入力された検索キーを使って実際にデータベース検索を行なう。
- 検索結果表示状態：得られた検索結果の表示を行なう。
- 説明状態：対話システムの使い方に関する質問に答える。
- 確認状態：システムの状態に対してユーザに「はい」「いいえ」等で質問をしてその返答ごとに決められた状態に遷移する。
- 最終状態：対話の終了時に遷移する。

これら 9 状態の間の遷移がどのような順番および条件で発生するかを決めることによって、対話戦略が決定される。対話戦略は対話スクリプトで記述される。

#### 2.2 対話スクリプト

本システムでは対話戦略を決定するのに対話スクリプトと呼ばれる XML 形式のファイルを用意する。XML ベースの対話スクリプトの文法を図 2 に示す。文法中大文字で始まる要素は XML の tag の名前、小文字で始まる物は、attribute (属性値) を示す。例えば、Dialogue が対話スクリプトを示す XML のタグで attribute として、title(システムの名前)、kind(システムの種類)を持ち、内部に、InitialState (初期状態) と任意個数の State(対話状態、これは上記の対話状態から初期状態と最終状態を抜いた 7 状態)、FinalState(最終状態) から構成されることが示されている。対話戦略を決定する一番の要素は ToState と呼ばれる tag である。これは、ある対話状態において、条件 condition が満たされた時に、Script で記載されている命令を実行してから nextstate で示されている対話状態に遷移するという tag である。ToState の XML での記述例を図 3 に示す。この例では、検索条件入力状態において、検索条件消去? というフラグがたっている場合 (『検索条件を消去して下さい』といった発話がされた場合) に、検索スロット消去という命令を実行し、検索条件を消去しましたという引数を持ってシステム出力が実行された後に、検索条件入力に遷移する (この場合は自己遷移)。

#### 2.3 対話戦略構築 GUI ツール

前述の対話スクリプトのうち初期状態における変数定義スクリプトおよび対話状態内の対話状態遷移の情報を自由に編集できるツールを構築した。また、対話の動的な対話状態変化を対話遷移図を使ってリアルタイムに表

```

Dialogue      := title kind InitialState
               State* FinalState
InitialState  := nextState Script
Script        := (Let|Run)*
Let           := name value
Run          := command argument
State        := name kind InitialScript?
               ToState+
InitialState  := Script
ToState      := nextState condition Script
FinalState   := Script
    
```

図 2: 対話スクリプトの文法定義

```

<state name="検索条件入力"
kind="検索条件入力">
:
<to nextState="検索条件入力"
condition="検索条件消去?">
<script>
<run command="検索スロット消去"
argument="" />
<run command="システム出力"
argument="検索条件を消去しました" />
</script>
</to>
</state>
    
```

図 3: ToState の対話スクリプト記述の例

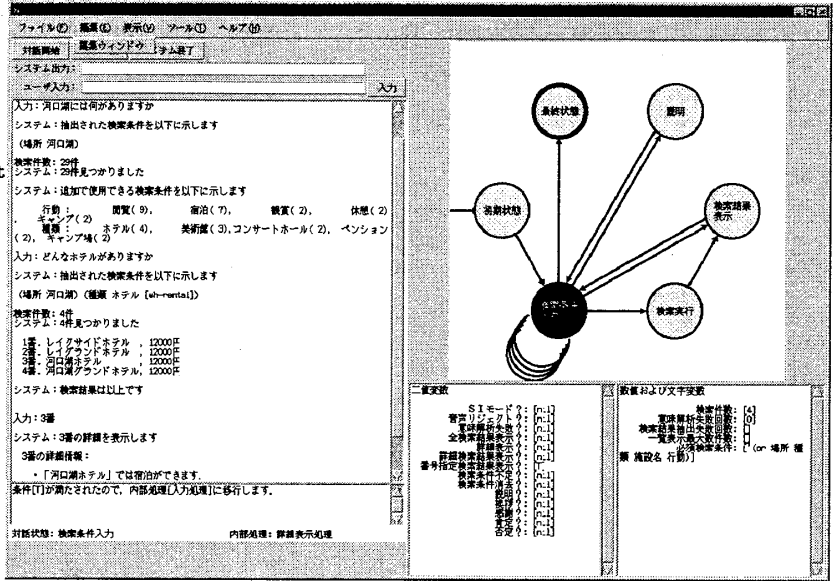


図 4: 対話システムの実行と対話状態の遷移例

示す機能を追加した。これにより、自分が意図した通りに対話状態が遷移するかどうかを確認できる。対話遷移図を伴う対話システムの実行例を図 4 に示す。この例では「河口湖にあるホテル」を検索していて、現在の対話状態が「検索条件入力」であることを示している。なお、本システムは、システムコア部を Perl で、GUI 部を Perl/Tk でそれぞれ構築している。

本システムにより XML テキストベースで編集をするよりも以下のような利点がある。

- 編集途中でいつでも対話システムを実行できる
- 対話システムを逐次的に動作させシステム構築者の意図通りにシステムが動作するかどうか確認できる

### 3. 評価実験

富士山観光案内、文献検索、ホテル検索の 3 種類のドメイン適用を行ない、こちらで用意した標準的な対話スクリプト上でシステムを使用してもらい、その後、本システムの編集機能で、いくつかの対話システムの振舞の変更課題を与え、それが達成できるかを調べた。

20 代の男性 2 名 (対話システムの知識あり) に以下の変更課題を与えた。

- システム開始時に表示されるメッセージを自分の好きなように変更する
- 標準で一覧表示できる件数を 10 件から 20 件に変更する
- 現在、意味理解あるいは検索に失敗した回数の合計が 2 回以上になった時にプロンプトを表示しているが、これを 4 回以上に変更する

予備評価実験の結果、前 2 つの課題は両者とも、最後の課題は 1 名のみが変更課題を達成した。なお、変更課

題の達成にかかる時間は数分程度である。また、1 点から 5 点の主観的評価において、「編集システムは使い易いか」という設問に対し、3.5 点 (平均) を得ている。これはまだ編集ツールがそれほど使い易いわけではないことを示している。その一方で、「編集ツールを使う場合と XML ソースを直接編集する方法ではどちらが楽そうか (1 点..XML の方が良い, 5 点.. 編集ツールの方が良い)」という設問に対しては 4.5 点 (平均) を得ている。これは対話戦略の編集ツール自体は有用であることを示している。一方、対話システムの一時停止機能、および対話状態遷移図の表示は両者とも 4 点 (平均) で「ある程度は役に立っている」という結果を得た。

### 4. まとめ

対話戦略を決定する対話スクリプトを GUI で編集できるツールを作成し、XML の記述になれていない対話システム構築者でも簡単に対話時のメッセージや対話状態遷移を修正出来ることを示した。

### 参考文献

- [1] 小暮悟, 中川聖一, “データベース検索用音声対話システムにおける移植性の高い意味理解部・検索部の構築と評価”, 情報処理学会論文誌, Vol. 43, No. 3, pp. 714-733, 2002.
- [2] 小暮悟, 伊藤敏彦, 中川聖一, “移植性の高い音声対話システムにおける対話スクリプト構築ツールの試作”, 情処研報, 2002- SLP- 40- 24 (合同: 2002- HI- 97- 24), pp. 139-144, 2002.
- [3] 梅田将満, 小暮悟, 中川聖一, “データベース検索用音声対話システムにおける意味理解部の移植性の検討”, 情報処理学会, 第 64 回全国大会 2 分冊, pp. 41-42, 2002.