

対話代行エージェントを用いた音声対話情報検索インタフェース

6C-4

岩沢 透 西村 健士

NEC C&Cメディア研究所

e-mail: iwasawa@ccm.cl.nec.co.jp

1 はじめに

音声対話による情報検索は、近年のCTIやカーナビ市場への音声認識の普及に伴いその必要性を増して来ている。しかし、音声によるコンピュータとの対話は様々な難しい問題をはらんでいる。その重要問題の一つが利用者の負担の問題である。本稿では、利用者の音声対話に要する負担を軽減する手法として、利用者に代わりコンピュータとの音声対話を代行する第3者的な対話代行エージェントを導入した3者間対話方式を提案する。そして、提案した対話方式をワインの検索に適用したシステムについて述べる。

2 音声対話による情報検索

従来の音声対話による情報検索アプリケーションとしては、レストランの検索システム[1]や電話で遠隔地の気象情報を検索するシステム[2]などがある。これらのシステムでは、自然言語を用い利用者主導でレストランや気象のデータを検索するシステムを提供している。しかし、自然言語を用いた音声対話システムは、発話語彙が定型化されていないため「システムが何を理解してくれるのかわからない(Whatの問題)」「どのような言い回しを使えるのか分からない(Howの問題)」といった問題点を持つ。本稿では、このような利用者の負担を取り除く方法として、利用者になり代わりコンピュータとの音声対話を代行する第3者の対話代行エージェントを提案する。

3 対話代行エージェントの利用

3.1 対話代行エージェントの概要と応用例

まず対話代行エージェントを用いた代行対話形式について述べる。代行対話では、対話システムに主対話エージェントと対話代行エージェントという2人の話者を配置する。主対話エージェントは、利用者と自然言語で対話しながら情報物件の検索を行なう。対話代行エージェ

ントは、利用者が前述のWhatやHowの問題で困った時に利用者になり代わり自動的に主対話エージェントとの対話を代行する。

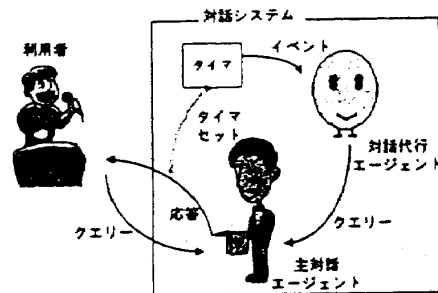


図1: 代行対話の対話形態

対話システムの動作形態を図1に示す。

利用者は、主対話エージェントと対話を行なう。主対話エージェントは、利用者からの入力を受けると、利用者に対し応答文を発話すると同時にタイマをセットする。タイマは、一定時間利用者の発話がない場合のみイベントを発する。そのイベントは対話代行エージェントに通知され、対話代行エージェントは主対話エージェントに対して発話を行なう。代行対話形式では、対話代行エージェントが必要に応じて主対話エージェントと対話するので、利用者が主対話エージェントに対して応答することなく対話を進行させることが可能となる。

対話システムは、内部に図2に示すような対話シナリオを持ち、主対話エージェントと対話代行エージェントは対話シナリオを共有している。利用者が発話すると、シナリオ中の利用者発話の項目に従い状態遷移を行なう。利用者の発話がない場合は、対話シナリオ中の利用者発話の項目が「タイムアウト」の部分に対応する状態遷移を行なう。「タイムアウト」の項目には対話代行エージェントの発話文章が付加されており、対話代行エージェントがこの発話文章を発話した後、状態遷移を行なう。

対話代行エージェントの応用例として、音声対話による「ワイン検索システム」を構築した。図3に「ワイン検索システム」の画面を示す。主対話エージェントと対話代行エージェントは画面右脇に配置されている。対話代行エージェントは主対話エージェントに話しかける時

A Spoken Dialogue Interface using Dialogue Agent.
Toru IWASAWA, Kenshi NISIMURA
C&C Media Research Laboratories, NEC Corporation

応答文: 「応答文1」		応答文: 「応答文2」	
状態ID	利用者発話: 遷移先状態ID	状態ID	利用者発話: 遷移先状態ID
1	発話1	2	発話1
	発話2		発話2
	発話3		発話3
	発話4		発話4
	タイムアウト 発話文: 「発話2」		2
2	発話1	4	発話1
	発話2		発話2
	発話3		発話3
	発話4		発話4
	タイムアウト 発話文: 「発話3」		5
3	発話1	8	発話1
	発話2		発話2
	発話3		発話3
	発話4		発話4
	タイムアウト 発話文: 「発話1」		9

図 2: 対話シナリオ

は横を向き、それ以外はこちらを向いて利用者の発話を待つことで、利用者に発話のタイミングを教える。

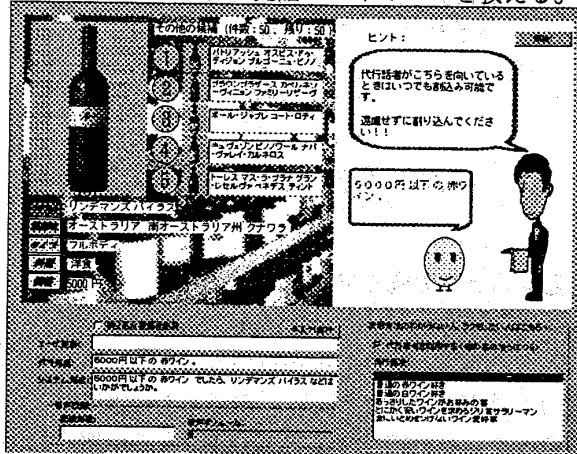


図 3: 「音声対話ワイン検索システム」画面

このシステムを用いた対話例を示す。[主]は主対話エージェントの発話を、[代]は対話代行エージェントの発話を、[利]は利用者の発話を示す。

- [主]: 「いらっしゃいませ、どのようなワインをお探でしょうか。」
- [利]: 「……」 (無発話)
- [代]: 「5000円以下の赤ワイン」
- [主]: 「5000円以下の赤ワインでしたら〇〇はいかがでしょう」
- [利]: 「……」
- [代]: 「フランスワイン」
- [主]: 「フランスワインでしたら△△はいかがでしょう」
- [利]: 「イタリアワイン」
- [主]: 「イタリアワインでしたら××はいかがでしょう」
- [利]: 「それ下さい」
- [主]: 「どうもありがとうございます。他に何かお探ししますか。」

主対話エージェントの「いらっしゃいませ」という発話に対し、利用者の無反応状態が一定時間続くと、対話代行エージェントが利用者の代わりに「5000円以下の赤ワイン」と発話する。ここで提示されたワインに対し、まだ利用者が無反応であった場合、今度は「フランスワイン」と条件を指定して検索を続ける。そして主対

話エージェントが提示したフランスワインに対し、利用者が(「イタリアワイン」と)発話すると利用者側に対話の主導権が移りその後の対話が進行する。

3.2 対話代行エージェントの効果

対話代行エージェントの利点として、まず、利用者がいちいち発話することなく対話を継続させることができるという点が挙げられる。これは、カーナビなどの画面を注視したり対話に集中できないような状況があるシステムの対話インターフェースとして有効であると考えられる。さらに、対話代行エージェントは、利用者に対して対話の語彙や方法を教示する効果も併せ持つ。音声認識の認識語彙数は有限であり、通常、人間の全ての発話を理解できるような音声対話システムを構築するのは不可能である。対話代行エージェントは、その限られた認識語彙の中でシステムが理解可能な言葉や情報検索の手段を利用者に教示する効果がある。先に示した対話例を取ると、対話代行エージェントは「5000円以下の赤ワイン」と発話することで、ワインを検索するための語彙(の一例)を利用者に提示し、さらに「フランスワイン」と発話することで、国名をキーにした検索が可能であることを利用者に教示している。その結果として、対話例に示したような対話代行エージェントの「フランスワイン」という発話の後に利用者が「イタリアワイン」と言い直すといった効果が期待できる。このような教示効果は、特に対話システム導入の初期段階でシステムが何を理解してくれるのか分からない状況において有効であると考えられる。

4 おわりに

本稿では、音声対話による情報検索を行なうインターフェースとして、対話代行エージェントを利用した対話制御手法を提案した。対話代行エージェントは、利用者の音声対話システムへの負担を軽減する効果があると考えられる。今後の課題としては、対話システムの有効性や対話代行エージェントの具体的な利用方法について検討していく予定である。

参考文献

- [1]. S. Senef and J. Polifroni, "A New Restaurant Guide Conversational System: Issues in Rapid Prototyping for Specialized Domains," Proc. International Conference on Spoken Language Processing, pp. 665-668, Philadelphia, PA, October 1996
- [2]. Jupiter's Homepage
<http://sls-www.lcs.mit.edu/jupiter/index.html>, 1998