

## マルチメディア端末上における音声対話支援

安地 亮一 城塚 音也 桑田 喜隆 小泉 宣夫

NTTデータ通信株式会社

情報科学研究所

**要約：**金融相談端末上の相談業務を、音声対話を理解することによって知的に支援する実験システム、サイバーコンサルタントを構築した。知識の少ない相談員でも本システムが提供する知識を元に、スムーズな相談対話を実現でき、サービス品質を維持できることを目的としている。本システムは、WWW インタフェースを用いた情報提供の仕組みを持ち、分散環境にも適応できる構造を持っている。

**キーワード：**金融相談、対話支援、音声認識、WWW インタフェース

## Spoken dialog support on Multimedia terminal

Ryoichi ANCHI, Otoya SHIROTSUKA, Yoshitaka KUWATA  
and Nobuo KOIZUMI

Laboratory for Information Technology

NTT Data Corporation

e-mail:{anchi, sirotuka, kuwata, nkoizumi}@lit.rd.nttdata.co.jp

**Abstract:** An experimental system, cyber-consultant, was built which supports financial consultation performed on a multimedia consultation terminal by understanding spoken dialogue. The system enables operators to consult with clients smoothly based on information provided by the system and conserves quality of the service high. This system has a application controllable WWW browser interface, which is used to provide a necessary information for consultation, and it has architecture applicable to network distributed environment.

**Keywords:** financial consultation, dialog support, speech recognition, WWW interface

## 1 はじめに

今日、自動電話応答システムや、WWWを用いたカスタマサポートなど、コストダウンをねらってユーザサポートが自動化される傾向が強い。その中で、金融機関のカウンタで行われる相談業務は、自動化がもっとも難しい業務の一つである。これは、顧客が金融商品に対する予備知識が少なく、顧客の金銭が絡むので確実性が要求されていることに起因する。

このように、未だ対面相談業務が重要である金融機関でも、合理化のためにかぎられた資源で質の高いサービスが提供できる枠組みを考えていく必要がある。筆者らが構築したサイバーコンサルタント [1]は、音声認識、知識処理、エージェントアーキテクチャを用いて、顧客支援中のコンサルタントを支援するシステムである。

これを用いることで、金融商品について知識の比較的乏しい相談員でも、音声認識や知識処理技術によって自動的に表示される支援情報を用いて品質の高い相談業務が行えることをねらいとしている。

本稿では、サイバーコンサルタントのシステム構成について説明する。特に Microsoft社の ActiveX 技術やWWWを用いて支援情報を表示する仕組みについて詳しく述べ、本システムの情報支援の有効性について評価実験結果を用いて議論する。

## 2 コミュニケーション支援の研究動向

リアルタイムの人間同士の対話をコンピュータに理解させ、対話支援を行わせる試みとして、西本ら[2]は、キーボード対話で、人間同士の対話をコンピュータが理解し、お互いの対話中に現れるキーワード群に基づいて話題の空間や個人の視点を可視化し、言及されていない話題空間に属する、見逃されている関連情報を提示することにより対話の促進を

図っている。

また、人間同士の音声対話をモニターし、何らかの決められたアプリケーション処理を行うシステムとしては、伊藤らが家のレイアウトの決定をタスクに人間同士の対話をコンピュータが理解し、対話内容に反映してレイアウトの変更をおこなうマルチモーダルシステム[3]を、有馬らが会議室の予約の受け付け係と利用者という人間同士の音声対話から情報抽出を行い、対話の要約をリアルタイムに生成するシステム MIMI[4]を構築している。

さらに、Nagaoらはこれらの対話促進機能と、音声対話理解の機能、表情合成機能を併せ持つ Talkman[5]を構築した。このシステムは、対面した人間同士の対話を聞き、自律的に音声対話に参加し情報支援を行うものである。

積極的に人間同士の対話に口をはさむ Talkman に対し、筆者らの提案するサイバーコンサルタントは、人間同士の対話を最も優先すべきものと考え、視覚的に相談内容に応じてタイムリーに情報支援の仕組みを提供している。さらに、アプリケーション処理機能を併せ持つことで、対話者にとって必要な情報処理を対話に連動して行うことができる。今回構築した相談業務支援システムでは、小規模ながらこれを実現している。

## 3 サイバーコンサルタント

筆者らが構築したサイバーコンサルタントシステムは、年金相談業務中のコンサルタントに対して情報支援を行う仕組みである。顧客とコンサルタントが、TV 会議上で行っている年金相談の音声対話を自動的に認識理解し、それに関する支援情報をコンサルタントに提示することで、相談の品質を制御する仕組みである。

これらの技術を統合して顧客との対話を支援する仕組みをサイバーコンサルタントと呼ぶ。

### 3.1 支援の種類

システムは年金相談をおこなう相談員のた

めに以下の3種類の支援を行う。

1. 対話内容のアドバイス

顧客に話すべき対話の内容や話の進め方を提示する。

2. リファレンス情報の表示

対話の内容に関連した情報を知識データベースから取得し、編集して提示する。

3. ユーザ情報の取得

対話内容から相談に必要なユーザ情報(保険の種類や職業等)を取得し、必要に応じて情報の加工(年金の受給額計算等)を行う。

現在の実験システムでは、1については、ごく基本的な知識(例:年金の受給額を計算するためには年金の加入期間と標準報酬月額を聞かなければいけない)しか使用しておらず、個人の状況や対話のコンテキストに応じた対話内容の推定と推定された発話内容に基づいたアドバイスを作成できるようにはな

っていない。2については、個人の状況(加入年金の種類等)に応じて、支援内容が編集される。現在はあらかじめ用意された知識ベースのみを使って支援内容を作成しているが、サービスの種類によっては、より大規模なデータベースの検索(インターネット上のサーチエンジン等)を利用して、リファレンス情報を収集することも有効であろう。

3.2 システム構成

サイバーコンサルタントは、次のような構成で実現されている。技術的ポイントは、

- ワードスポッティングを用いた音声認識[6]
- ルールベースに基づく知識処理
- エージェントアーキテクチャ

である。

サイバーコンサルタントは年金相談に必要な知識を数個のハイパーテキストの形式で持っている。知識ハイパーテキストは、それぞれ「国民年金」や、「繰り上げ受給」に関するトピック毎に詳細に分類され、インデック

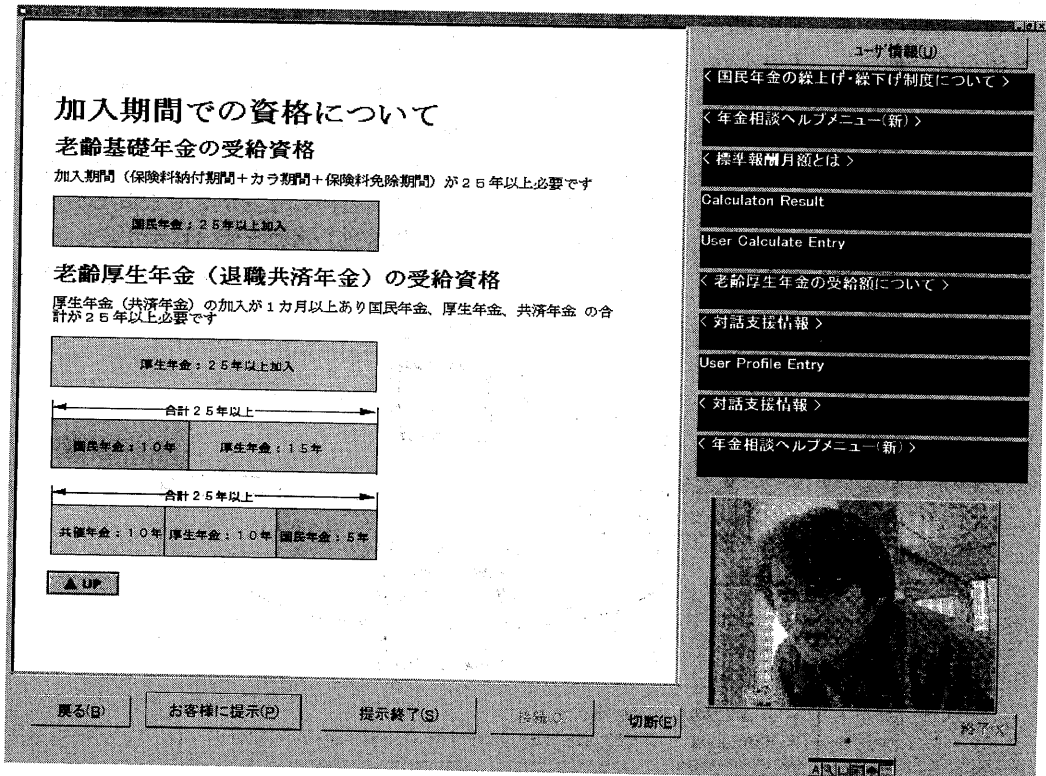


図 1 相談員端末画面

スが付けられている。以下に、サイバーコンサルタントを機能ごとに分類し、それぞれの詳細に付いて述べる。

### マルチエージェントアーキテクチャ

本システムで用いているマルチエージェントアーキテクチャは、ブラックボードモデルと呼ばれ、メッセージサーバを中心に、機能毎にエージェントが散在するアーキテクチャである。音声認識や知識処理、ユーザインタフェースを受け持つ部分は、機能毎にエージェントとして散在し、相互の通信はすべてメッセージサーバを通じて行われる。

この仕組みは、単純なプロトコルで通信を行うため、実装が容易で機能拡張がしやすいという利点がある。また、役割毎にプロセスが分けられており、同一マシン上で複数のエージェントを動作させることも、地理的、ネットワーク的に分散した環境でも柔軟に対応できる。

### 音声認識部[6]

音声認識は、顧客側の相談端末とオペレータ側双方で動作し、それぞれ年金相談に必要なキーワードが認識対象語彙として登録されている。ユーザが発話を中断した時、認識処理は一時的に中断され、それ以前に認識されたキーワードがメッセージサーバに送られ、新たな発話待ち状態に入る。これらの動作はすべて自動的に行われるため、オペレータと

顧客は、音声認識が作動していることをまったく意識せず自然な相談対話を続けることができる。

### 知識処理部

音声認識によって対話から抽出されたキーワードは、その時点では何の意味も持たない。そのキーワードを元に、対話支援のための情報を検索する必要がある。本稿で述べる実験システムの知識処理部は、入力キーワード列に簡単なパターンマッチングを用いることで、対応する支援情報へのポイントを得る。情報へのポイントは URL としてメッセージサーバを通じて相談員端末へと転送される。

### プレゼンテーション部

顧客用店頭端末と相談員用端末は、Windows95 を用いて実装されている。相談員側端末には、知識処理部からの情報提示コマンドに従って支援情報が表示され、提示した情報の題名が、時系列順にボタンとして表示されている。相談員は、この履歴ボタンを用いて、最近の話題をすばやく再表示することができる。

また、対話に応じて自動的に提示された支援情報が、グラフなどの視覚的に分かりやすいものならば、相談員は自分の判断で支援画面を顧客側端末にも表示させることができる。相談員用端末の画面を図 1 に示す。

プレゼンテーション部の構造詳細について

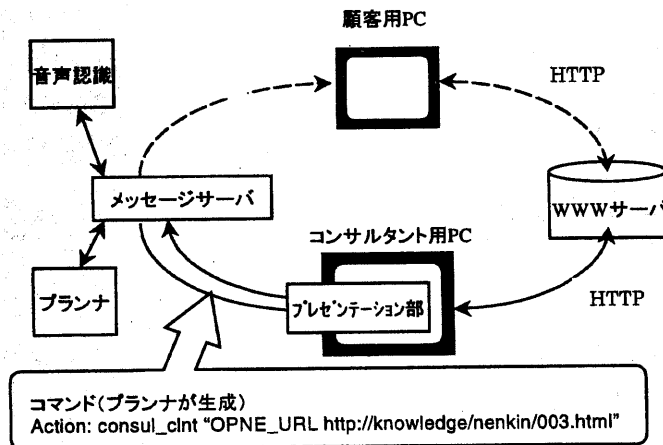


図 2 メッセージの流れ

は次章で詳しく述べる。

## 4 端末構成

本章では、コンサルタント側 PC と、TV 会議で接続された顧客側 PC 上に実装された情報支援のプレゼンテーション部について詳しく述べる。

サイバーコンサルタントが提供する支援情報は、内部に HTML 形式で保存されている。プランナが、発話されたキーワード列をもとに生成した情報へのポインタは、知識情報を指し示す URL で表現される。

この URL は、メッセージサーバを経由してプレゼンテーション部へ転送される。プレゼンテーション部は、支援情報のインデックスを URL として受け取り、その URL に対応する HTML ファイルを表示することでコンサルタントに情報を提示することができる。メッセージの流れを図 2 に示す。

支援情報へのポインタが URL で与えられるので、支援情報のデータベースの場所がローカルマシン上、ネットワークを介したマシン上のどちらであっても、一意に指し示されたアドレスで示されるので、情報の場所を意識すること無く情報を取り出すことができる。マルチエージェントアーキテクチャと共存することで、柔軟なシステム構成が可能である。

### 4.1 WWW ブラウザインタフェース

プレゼンテーション部には、メッセージサーバとの通信機能と、支援情報を表示する HTML ビューアの機能が含まれ、メッセージサーバからのコマンドに応じて、自動的に WWW から支援情報である HTML ファイルを表示することができる。また、コンサルタントは、支援情報に含まれるハイパーリンクをクリックすることで、従来通りの Web インタフェースで情報検索を行うこともできる。

本システムでは、プログラムからコントロール可能な HTML ビューアとして、Microsoft Internet Explorer ActiveX コントロールを用いた。以下に詳しく述べる。

## Microsoft Internet Explorer ActiveX コントロール

プレゼンテーション部は、Microsoft Windows95 上で、開発環境に VisualBasic4.0 を用いて作られている。HTML ビューアには、Microsoft Internet Explorer ActiveX コントロール[7]を用いた。この方法を用いると、開発者は作成するアプリケーション上に簡単に Web インタフェースを追加することができる。

ActiveX とは、ソフトウェアを、細かい機能ごとに分け、部品化することでソフトウェア開発を容易に、またコードの共有を可能にする技術で、Microsoft 社が推し進めているものである。その中でも MSIE コントロールは Microsoft Internet Explorer の核となるもので、無償で配布されている MSIE をインストールすると自動的にそのコンピュータで使用可能になる。開発者はその部品を用いることで、自由にアプリケーションに Web インタフェースを追加することができる。

## 5 実験と評価

Wizard Of Oz 法を用いてこのような支援形態がユーザにとって有用か評価を行った。社内総務相談を対象に、相談員がいかに効率よく情報にたどり着いて、相談対話が早く終了するかを評価した。評価指標は相談終了までの所要時間である。

被験者は相談員で、顧客役にはあらかじめ質問内容が与えてある。実験者は、被験者の情報検索画面をリモートコントロールし、自動的に情報を提示する。以下に実験を行った相談の例を示す。

### 相談例

「広島支社の最寄り駅を教えてください」  
「住宅ローンはいくらまでかりられますか？」

実験の結果、情報支援を行わない場合より情報支援を行ったことで、相談タスク達成時間は約 11%~51% 短縮されることが分かつ

た。すなわち、サイバーコンサルタントの情報支援は有効であることが立証された。

さらに、アンケートの実施によりユーザインタフェースの使用感に関して有用な情報を獲得できた。

●支援情報が自動的に得られたものなのか、自分で探し出したものなのかかわかりにくい。

●検索作業を唐突に中断させられ混乱する。

●支援された情報の、全情報中の構造的位置が不明確。

筆者らは、これらの情報を吸収し本システムのインタフェースを改良することを計画している。詳細については、別途報告することとする。

## 6 まとめ

本稿では、サイバーコンサルタントのコンセプトと構成について述べた。また、端末上の情報提示の仕組みを詳しく述べた。システムの拡張性や構築時の容易性を考えると、今回用いたような、市販の部品を積極的に用いることも、有効な手段と言えるであろう。

また、本システムの持つ情報支援の仕組みの有効性を実験によって実証した。しかし、本実験では、Wizard of Oz法を用い、提示情報正解率が100%の状況を模擬していたので、音声認識や知識処理技術の性能の問題から、現実的な実験とは言えない。今後、提示情報正解率や提示に要する時間等の性能が、本情報提供手段に及ぼす影響について詳しく調査する必要がある。

## 謝辞

本研究を進めるにあたり、多大なご助言を頂いた、管村昇情報科学研究所長に感謝いたします。

---

[1] 安地, 城塚, 桑田, "相談業務を対象とした対話支援環境の構築", 情報処理学会研究報告, 96-HI-69, pp33-38 (1996, 11)

---

[2] 西本、角、間瀬, Augmented Informative Discussion Environment "AIDE", 第2回知能情報メディアシンポジウム予稿論文集, pp259-266 (1996, 12)

[3] 伊藤, 木山, 関, 小島, 張, 岡, "同時複数話者の会話音声およびジェスチャのリアルタイム統合理解による Novel Interface System", 情報処理学会研究報告, 95-SLP-7, pp17-22, (1995.7)

[4] 有馬、亀山、河合 "自動対話要約システムの構築", 音響学会講論集, pp157-158, (1996,3)

[5] Akikazu Takeuchi and Katashi Nagao. Communicative facial displays as a new conversational modality. In *Proceedings of the ACM/IFIP Conference on Human Factors in Computing Systems (INTERCHI'93)*, pp. 187-193. ACM Press, (1993).

[6] 磯部, 森島, 小泉, "ガベージモデルを用いたキーワード抽出", 音響学会講論集, pp31-32, (1996,9)

[7] <http://www.microsoft.co.jp/intdev/sdk/docs/iexplore/default.htm>