

情報の発信と共有を促進するマルチモーダル 秘書エージェントシステム

Multimodal Interface Agent

中山 康子
Y. Nakayama金澤 博史
H. Kanazawa竹林 洋一
Y. Takebayashi

知識やノウハウをもっている本人に代わって問合せに答える，マルチモーダル 秘書エージェントシステムを開発した。仕事をしていて困ったときなどに言語や音声で質問をすると，その意味内容を理解して，知識情報やノウハウを蓄積したデータベースから適切な回答を探し出して提示する。秘書エージェントシステムは，知識情報をあとで検索・活用しやすくするための知識情報の構造化技術と，従来のGUI(Graphical User Interface)に代わる言語や音声を用いたマルチモーダル対話技術を組み合わせて実現した。システムの実用化に向けて，実際にコンテンツを充実させて当社の組織内で実践し，有効性を確認している。

We have developed a multimodal interface agent that supports information and knowledge sharing and inter-group communication. When a user enters a query in natural language, the agent system retrieves the relevant information from its knowledge databases and presents the appropriate answer. The system is provided with information/knowledge structuring techniques and a user interface with multimodal information using speech and natural language.

A trial system has been installed which is currently being evaluated and improved through practical use.

1 まえがき

ネットワークインフラの高度化とコンピュータのマルチメディア化の進展は，個人や組織の情報収集・発信の能力を向上させる一方で，情報洪水を招き，情報収集やコミュニケーションにおける人間の負担を増大させている。従来のマウスとメニューによる直接操作のGUIでは，膨大な情報の中から適切なものを選択するのが困難で，人にやさしい新たなヒューマンインタフェースが求められている。

一方，仕事を進めるうえで必要な知識や情報は膨大であり，特に非定型業務は個人のノウハウに依存している部分が多い。しかし，個人が所有しているノウハウは暗黙知のままどまり，貴重な知識が十分に利用されていないという問題がある。従来のグループウェアツールは情報共有の枠組を提供したが，システムが有効に機能するためには枠組のよしあしではなく，コンテンツ(情報内容)の充実が重要であることがわかってきた。知識情報共有のねらいは組織や個人の暗黙知の形式知化であり，有益な知識情報の蓄積・検索・活用方法を見極めるには，実際に組織の知的財産を地道に集めて形式知化を実践することが重要である。

以上のような問題意識に基づき，個人の知識やノウハウを容易に獲得して，知識データベースに蓄積し，言語，音声，ジェスチャによる自然な対話できめ細かく検索できる秘書エージェントシステム(図1)を開発した。現在，社内の1,000人規模の組織で，事務手続きを中心に知識データベースの内容を充実させて実証実験を行い，システムの有効性を検証している⁽¹⁾。

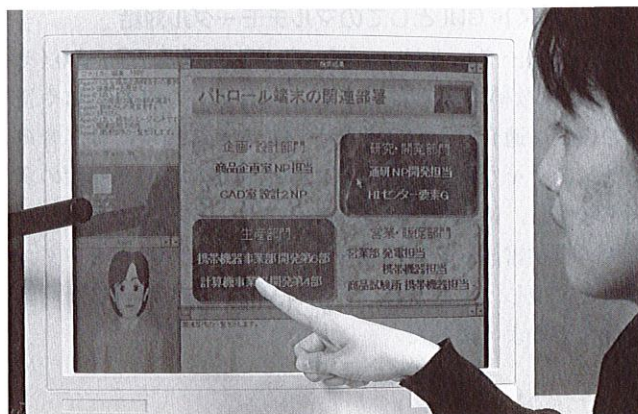


図1. マルチモーダル 秘書エージェント 言語や音声，ジェスチャを用いて質問をすると，知識やノウハウをもっている本人に代わって問合せに答える。

Multimodal interface agent

この論文では，組織における知識情報共有を促進する秘書エージェントシステム構築へのアプローチと，現在組織内で実践しているシステムの概要について述べる。

2 アプローチ

2.1 オフィス知識ベースとノウハウベースの連携

ストック情報の充実のためには，組織や個人の知識の形式知化を実際に行うことが必須(す)である。従来のエキスパートシステムは，ドメインやアプリケーションを限定し

て、問題解決に参照される知識を分析して形式化した。オフィスの知識を分析して、このような知識ベースを構築するには膨大な手間がかかる。そこで、形式知化を知識とノウハウの連携で行うことにした。知識とは、意味のある抽出された情報どうしを相互に関連づけたもので、用語の意味やものごとの機能・構造・関係などを表すまとまりをもった概念構造である。コンピュータが問題解決に利用できる記述形式で、専門家が業務を分析して記述する。一方、ノウハウは、人間が理解できる形式で記述された形式知で、組織のメンバから獲得する。

個人がストックしている情報を積極的に公開する個人情報公開システム PIP(Personal Information Provider)を先に試作した⁽²⁾。ここで開発した情報の浅い構造化を、形式化されたオフィス知識ベースと連携させ、組織レベルの本格的な知識情報共有に拡大した。

知識を体系的に記述した知識ベースと、体系化されていないが人間が理解可能なノウハウベースを構築して連携させることにより、非均質な個人のノウハウを入力しやすくし、活用できるようにする。さらに、蓄積されたノウハウから形式化できる知識を抽出して知識ベース化し、アドバイスなどの知的支援に発展させていく。

2.2 ポスト GUI としてのマルチモーダル対話

マルチモーダルとは、視覚、聴覚、触覚などの感覚(センサ)が複数種類あるということを表す。人間の認知や情動の性質に適合したマルチモーダル対話の実現には、音声やジェスチャなどの各種メディアを利用して、ユーザの意図を理解し、適切な応答を提示することが必要である。そこで、自然言語、音声、ジェスチャによるマルチモーダル対話を実現し、ユーザにとっては負担の少ない自然な対話が行なえるようにした。

3 システム構成

システムは、図2に示すように、知識データベース、対話インタフェース、マルチメディア情報構造化処理から構成され、対話インタフェースやマルチメディア情報構造化処理は、音声認識・合成、文字認識、文書理解などの各種メディア変換処理を提供する HI(ヒューマンインタフェース)ウェア⁽³⁾を利用して構築した。

3.1 知識データベース

知識情報源は、オフィス知識ベース、ノウハウをストックするノウハウベース、既存のマルチデータベースで、これらのデータはオフィス知識ベースにより“のり付け”され、相互に関連づけられて利用される。また、オフィス知識ベースは、各種のマルチメディア情報を意味レベルでのり付けすることによる各種メディア処理の統合、メディア変換のあいまい性解消、認識理解の向上の基盤となる。

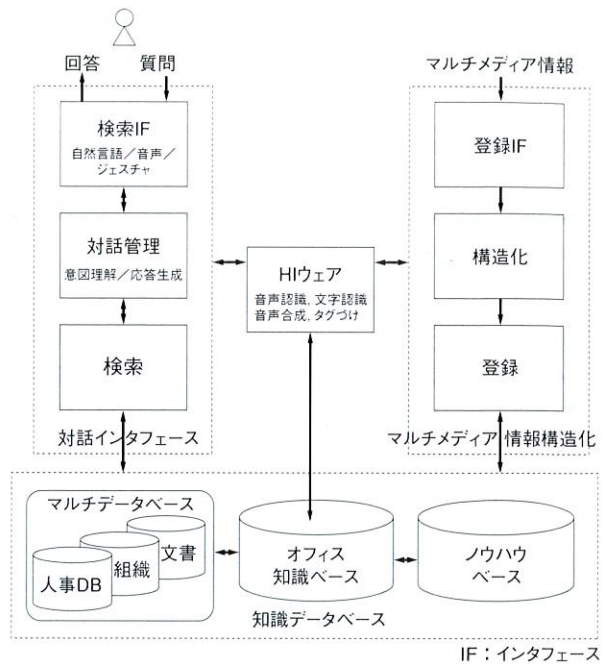


図2. システム構成 知識情報を蓄積する知識データベース、質問の意図を理解して適切な情報を検索する対話インタフェース、知識情報を自動的にデータベースに登録する情報構造化から成る。

System configuration

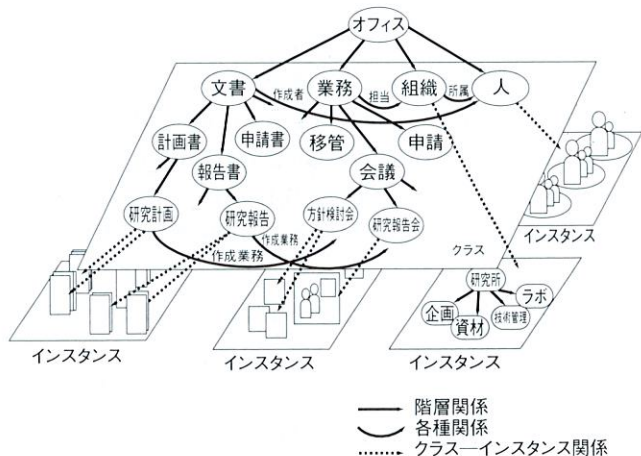


図3. オフィス知識ベースの構成 人・組織、文書、業務などの概念体系を表すクラス階層と、オフィスの実体や相互関係を表すインスタンス階層から成る。

Office knowledge base

汎(はん)用の知識ベースとしては、CYC、WordNet、EDR辞書など⁽⁴⁾が知られているが、当社では、知識情報共有に焦点を絞った知識ベースを構築中である。知識ベースは、図3に示すように、概念体系や推論の基本処理をもつクラス階層と、オフィスの実体や相互関連を記述するインスタンス階層をOODB(Object-Oriented DataBase)上に実装した。クラス階層は、人・組織、文書、業務などのオフィスの事物の相互関連や機能、業務の手順、概念・語彙(い)体

系を記述したものである。一方、ノウハウベースには記述方式や質を限定せずに、組織内で共有すると有益であると思われるものを個人が提供して蓄積した。ノウハウデータは、自然言語対話で検索できるように、内容を表す対象、行為、形態、状況などの13種類の属性や、ノウハウどうしの関連を記述する構造情報をもつ。これらの構造情報は、ノウハウデータ登録に際してユーザが記述したタイトルと本文から抽出する。ユーザからの自然言語の問合せ文を解析した結果、この構造情報を参照して、マッチングを行う。

3.2 対話インタフェース

対話インタフェースは、音声、言語、ジェスチャによるマルチモーダル入力を可能としている。ユーザからの意図を正しく理解するためには、複数の入力メディアからのあいまいな情報を統合して適切な解釈を行うことが必要となる。そこで、仮説推論に基づく入力統合方式を新たに提案し、実システムに組み込んだ⁽⁵⁾。

音声やジェスチャなどから得られた大量の認識候補の組合せの中から、高速かつ最適な解釈を導くために、仮説推論を用いて、過去に行われた入力統合結果を参照し、現在候補の中から矛盾のない候補の組合せを絞り込み、解釈処理を行う。これにより、入力メディアの冗長性と自然性、およびあいまいな入力に対する頑健性を同時に実現することができるようになった。

また、システムからの出力は、合成音声、応答文テキスト、表情をもったエージェントのコンピュータグラフィック(CG)画像、応答結果画像によるマルチモーダル出力とした。応答結果画像に対してはジェスチャによる対象物の指定を行えるようにした。さらに、自然性を高めるために、合成音声とCG画像の同期をとり、合成音の再生時は口を動かし、それ以外では適度なタイミングでまばたきや首振りを行うようにした。入力解釈に失敗した場合は、再入力の回数に応じて表情を徐々に変化させたり、合成音の韻律を変化させるようにし、システムの状態の円滑な伝達を行えるようにした。

3.3 マルチメディア情報の構造化

ユーザが登録したノウハウや報告文書などの、テキスト、イメージ、映像情報などの多種多様なマルチメディア情報は、検索・活用しやすいように構造化して蓄積している。入力されたマルチメディア情報は、HIウェアの各機能を用い、ドキュメントリーダ、音声認識、英日翻訳などにより日本語テキストに変換し、これに対してキーワード抽出、タグ付けを行うことで自動的に構造化する。図4に示すように、ユーザが入力したタイトルと本文から、内容を表す重要なキーワードを抽出して構造化を行い、ノウハウベースに登録する⁽⁶⁾。キーワード抽出は、オフィス知識ベースの概念体系を参照している。

例

正月休みに飛行機で札幌に帰省する場合

帰りは、仙台まで飛行機で行って、仙台から新幹線を使う手がお勧め。札幌—仙台間は意外と簡単に予約がとれます。

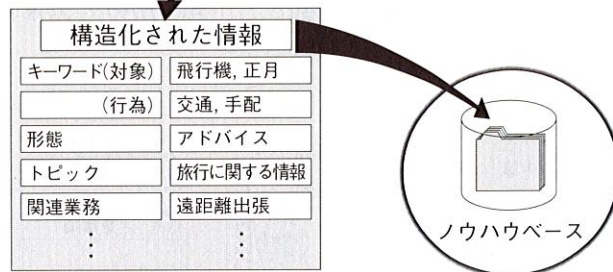


図4. ノウハウの自動獲得 ユーザが入力したタイトルと本文から、内容を表すキーワードを抽出し、データを構造化してデータベースに登録する。

Automatic structuring of know-how data

また、紙の文書のデータベース化・構造化をHIウェアの文書理解機能により進めている。

4 知識データベースの内容検索

システムは自然言語による質問文を解析し、オフィス知識ベースまたはノウハウベースから適切な回答を検索して提示する。人・組織の役割や関連文書など、体系的に整理された知識を用いて回答できる質問に対しては、知識ベースを参照して回答し、こつや秘訣、注意事項、事例、例外処理、失敗談など、体系的に記述されていないものはノウハウベースを検索して回答する。

図5では、「設備を購入する際の作成書類は？」という問合せに対して、自然言語解析の結果、知識ベースに記述されている“設備購入”オブジェクトと関係属性“作成書類”のペアがマッチし、作成する書類として“設備計画書”、“見積書”、“手配依頼書”を回答する。一方、「設備計画書を作成する際の注意事項は？」という問合せに対しては、知識ベースでは“設備計画書”オブジェクトはマッチするが、“注意事項”に該当する属性や関係が存在しないのでノウハウベースの構造情報とマッチングを行い、対象が“設備計画書”、行為が“作成”、形態が“注意事項”であるノウハウデータを検索して提示する。

5 社内実践システム(Advice/Help on Demand)

現在、社内の約1,000人規模のオフィスで日常業務に使いながら、実証実験を行っている。図6はシステムの動作画面例である。クライアントインタフェースはWWW

対話例

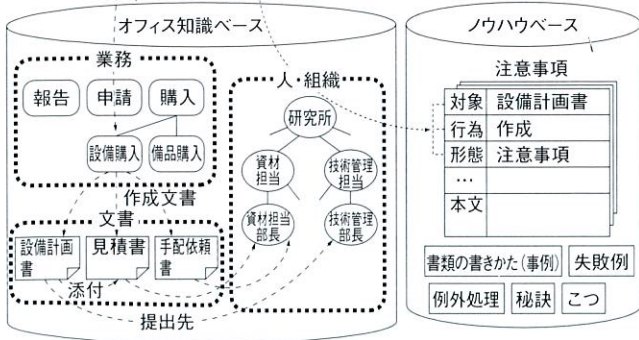
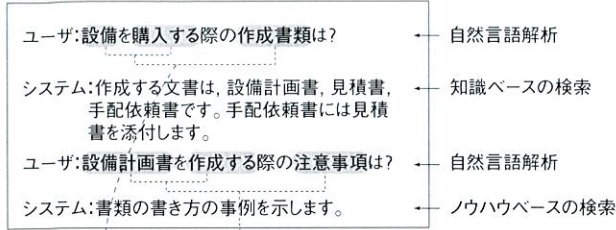


図5. 知識データベースの内容検索 自然言語の質問文を解析し、質問の意図に合った情報を知識データベースから検索して提示する。

Retrieval of knowledge database contents

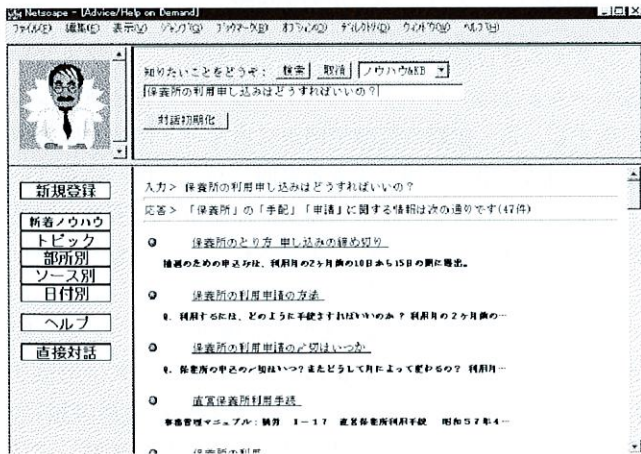


図6. 社内実践システム(Advice/Help on Demand) 1,000人規模のオフィスで日常業務に実験的に使用している。自然言語で質問を入力すると、質問の内容に合った検索結果を提示する。

"Advice/Help on Demand" trial system

(World Wide Web)上に構築し、キーボード入力による自然言語で問合せができるようにしている。現在、事務手続き、研究関連、計算機の使いかたなどに関するノウハウが約10,000件登録されている。新しいノウハウを登録する機能、参照したノウハウに評価やコメントを返す機能、未登録の質問には管理者が回答する直接対話機能をもっている。

6 あとがき

秘書エージェントシステムについて、梓組の提案だけでなく、実際に知識データベースにコンテンツを入れて実践している事例について述べた。知識をストックし、共有・活用できるようにするという課題の解決策として、オフィス知識ベースと、浅い構造をもつノウハウベースの連携、さらにHIウェアを用いたメディア変換による情報構造化が、オフィスの形式知の量と質を向上させた。また、言語、音声、ジェスチャによるマルチモーダル対話により自然な対話を実現した。今後は、知識ベースをさらに拡張し、ユーザの状況や意図を理解し、アドバイスやヘルプなどの適切な応答を行う知的支援機能を拡張する。

また、現在、システムの応用として、顧客からの問合せに対応するヘルプデスクシステムを開発中である。

文 献

- (1) 中山康子, 他: 知識情報共有システムの開発と実践, インタラクシオン'97, pp.103-110(1997)
- (2) 福井美佳, 他: コミュニケーション支援のための個人情報公開システム(PIP), 情報処理学会第52回全国大会(1W-7), pp.6-187-188(1996)
- (3) 杉山博史, 他: コモンHIサービス環境の開発, 情報処理学会第52回全国大会(2W-1), pp.6-193-194(1996)
- (4) D.Lenat, et al: CYC, WordNet, and EDR: Critiques and Response, Communications of ACM, 38-11, pp.33-48(1995)
- (5) 河野恭之, 他: 仮説推論に基づくマルチモーダル入力統合方式, インタラクシオン'97, pp.33-40(1997)
- (6) 真鍋俊彦, 他: WWWを用いたオフィスでの知識情報共有の実践, 情報処理学会54回全国大会(1S-7), pp.3-293-294(1997)



中山 康子 Yasuko Nakayama

研究開発センター 情報・通信システム研究所主任研究員。知識処理・知的インタフェースの研究開発に従事。情報処理学会, 人工知能学会会員。
Communication & Information Systems Research Labs.



金澤 博史 Hiroshi Kanazawa

研究開発センター 関西研究所研究主務。音声認識・音声対話技術の研究開発に従事。電子情報通信学会, 日本音響学会会員。
Kansai Research Lab.



竹林 洋一 Yoichi Takebayashi, D.Eng.

研究開発センター 情報・通信システム研究所部長。工博。音声情報処理・知的インタフェースの研究開発に従事。電子情報通信学会, 情報処理学会, 日本音響学会, 人工知能学会会員。
Communication & Information Systems Research Labs.