

2H-4

一般ユーザを対象とした
オンライン・データベース検索支援システム

西村健士, 日吉まゆみ, 高島洋典, 金子朝男

日本電気(株) C&C情報研究所

1. はじめに

近年のオンライン情報サービスの発展にともない、金融情報、企業情報、新聞記事情報などが公衆回線を通じて手軽に利用できるようになってきた。このように様々な情報が多量に流通する状況においては、ユーザに必要な情報だけを選択的に取り出し、それを有効に活用するための支援システムの存在が望まれる。対象を家庭内の一般ユーザに想定し、このような支援システムについて検討を行なった。

2. システムに要求される機能

通常、データベース検索は、データベースから得られたデータをなんらかの形に加工したり、得られたデータを何かの意志決定の参考にするという目的を持って行なわれる。データベースの利用支援は、ユーザが本来必要としている情報は何か、また、その情報を得るためにはデータベースからどういうデータを引き出せば良いのかまでを考えて行なうことが望ましい。

最近のオンライン・データベースの量的質的充実や、パソコン等の情報機器の普及をみると、情報検索に不慣れた家庭内の一般ユーザに対しても、オンライン・データベースを利用する環境は整ってきて

いると言って良いだろう。問題は、これらのユーザがどこにどういう情報があるのか、どうやってその情報を得たら良いのかがわかりにくいことである。このようなユーザに対してオンライン・データベースの利用を支援するシステムには、以下の機能が必要とされる。

- ①ユーザの抱える問題の分析
- ②問題解決のための計画立案
- ③データベースの検索実行
- ④検索されたデータの処理

①ユーザの持つ問題を解決するために問題を分割し、分割した各部分に対してデータベースからどういう情報を得たら良いのか分析する。②アクセスすべきデータベースの決定、検索式の生成、検索データの処理方式の決定などを行なう。③データベースに接続し、実際に検索を行い、検索結果を取り込む。④データの処理を行なう。データの形態としては自然言語のテキストも考えられるから、自然言語テキスト中からの情報抽出も行なう。

3. システムの構成

図1にここで検討したシステムの構成を示す。インタフェース部では自然言語やメニューを融合した

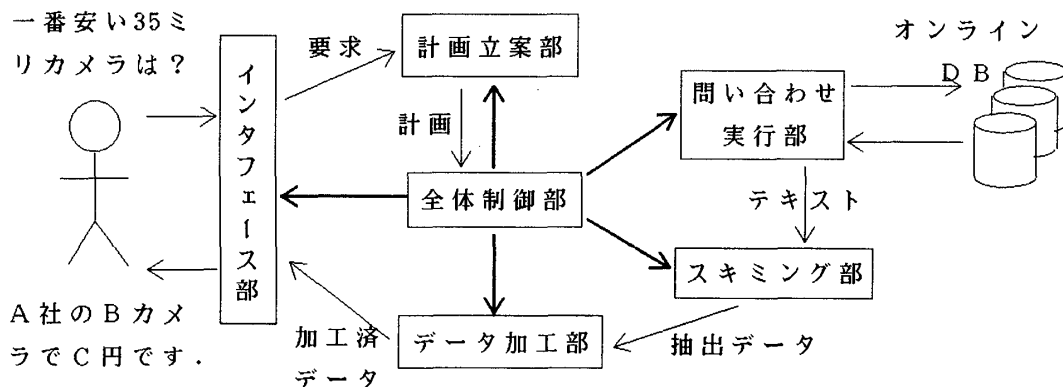


図1. システムの構成

On-line Database Information Retrieval Assistant for Novices

Kenshi NISHIMURA, Mayumi HIYOSHI, Yosuke TAKASHIMA, Asao KANEKO

C&C Information Technology Research Laboratories, NEC Corporation

対話によって、ユーザの求めている情報が何であるかを把握する。計画立案部はデータベースの検索に関する計画や検索されたデータを処理する計画を立てる。問い合わせ実行部はオンライン・データベースへの接続や検索コマンドの生成・実行、検索データの取り込みなどを行なう。検索されたデータがもしテキストであれば、スキミング部において必要な情報を抽出する^[1]。データ加工部は、検索されたデータやスキミングによって得られた抽出データをユーザの意図にあった形に加工する。全体制御部がシステム全体の制御を行なう。

今回のプロトタイプ・システムでは対象領域を「パーソナル・ニューメディア機器に関する新製品情報」に設定した。オンライン・データベースとしては商業新聞の新製品情報をとりあげた。このデータベースは自由語による検索が可能で、新聞記事の見出しや本文などが蓄積されている。

4. 検索対象の表現方法と処理の流れ

4. 1. 検索対象の表現方法

対象領域のオブジェクトは階層化して整理されている。例えば、製品オブジェクトの下位にカメラオブジェクトがあり、そのさらに下位にステルカメラオブジェクトがある。各オブジェクトには属性が定義されており、属性は下位のオブジェクトに継承される。例えば、“価格”という属性は製品オブジェクトに与えられており、それ以下のすべてのオブジェクトにも付属しているものとして扱われる。モジュール間で受渡しするデータはこの属性が基本となる。

4. 2. インタフェース部の働き

インタフェース部は、ユーザのほしい情報はという処理を通じて得られるのか、その処理を行なうためにはどのような情報が前提としてわかっているのか、という観点からユーザの質問を解析する。例えば、

一番安い35ミリカメラは？

という質問は、

“タイプ”が「35ミリ」であるステルカメラオブジェクトのうち、“価格”が最小であるものを知りたい

というふうに解析される。解析結果を構成するものは、①オブジェクトの種類（ステルカメラオブジェクト）、②属性（“タイプ”、“価格”）③属性にかかる条件（値が「35ミリ」）、④属性を処理対象とする処理の形態（最小値）である。

インタフェース部は、ユーザとのやりとりを通してユーザの言及する処理形態を判別し、その表現を計画立案部に渡す。処理形態のパターンは、質問文のアンケート調査の結果を分析して抽出した。

4. 3. 計画立案部の働き

計画立案部は、検索とデータ処理の計画を立てる。検索の対象となるのは記事自体ではなく、その記事中に含まれている情報（オブジェクトの属性値）である。ユーザの入力した自然言語のテキストと類似した意味を持つ新聞記事の見出しを検索するシステムが提案されているが^[2]、われわれのシステムでは、見出し文全体の意味構造を対象とするのではなく、オブジェクトの属性値を求めるために見出し文を解析し、見出し文から属性値が得られなかつたらさらに本文中から抽出するように試みる。計画立案部はこれらの計画を行ない、それを全体制御部に渡す。

4. 4. その他のモジュールの働き

全体制御部は、計画にしたがって問い合わせ実行部、スキミング部、データ加工部に指示を出す。問い合わせ実行部はオンライン・データベースのコマンドを生成し、手順に従ってキーワードによる検索を行ない、検索されたテキストを取り込む。スキミング部はオブジェクト名と属性名を渡され、対応する属性値を読みとる。データ加工部は処理の形態と属性名とスキミングで得られた属性値を渡され、データの処理（加工）を行なう。加工されたデータはインタフェース部に渡され、ユーザに提示される。

5. おわりに

一般ユーザを対象としたオンライン・データベース検索支援システムにおいて、機能構成とデータの表現方法について検討した。現在、システムを構築して評価を行なっている。今後の課題として、対象領域の拡大とそれともなうデータベースの選択方式の検討があげられる。また、各処理段階でのユーザとのインタラクションのあり方も検討していきたい。

参考文献

- [1]西村, 島津, 高島, “情報サービスにおける自然言語インタフェース”, 情処学会シンポジウム, 1987年11月
[2]杉山, 秋山, 伊吹, 川崎, 内田, “自然言語理解に基づく情報検索システムIRIS”, 情処研報 自然言語処理 58-8, 1986年