新しい教育放送サービスのための情報検索システム

Agent Search System for a New Interactive Educational Broadcast Service

正会員 住吉 英樹[†], 正会員 山田 一郎[†], 正会員 村崎 康博^{††},
正会員 金 淵 培[†], 正会員 八木 伸行[†], 正会員 柴田 正啓[†]
Hideki Sumiyoshi[†], Ichiro Yamada[†], Yasuhiro Murasaki^{††},
Yeun Bae Kim[†], Nobuyuki Yagi[†] and Masahiro Shibata[†]

Abstract We propose an educational broadcast service that supports not only passive learning by simply watching broadcast programs but also active learning in which students search for and collect programs for their own purposes, such as making presentations, and exchanging opinions. To achieve this new broadcast educational service, we have constructed an education support system that enables users to search multimedia databases using agent technology. The system assumes its users to be elementary and junior high school students and supports theme-based "investigative learning." This paper describes the features of this system and presents a system overview.

キーワード: 教育, 放送, インタネット, 検索, エージェント

1. ま え が き

2002年5月末の日本における自宅でのインタネット利用可能者数は、推定で6,044万人(ネットレイティングス社)とされ、DSLやCATVを使ったブロードバンドネットワークの接続者も急増している。いまやインタネットは、双方向性を持ったメディアとして、大きな存在になったといえる。さらに、ブロードバンド化されたネットワークでは、映像の配信も可能となり、放送との関わりも深くなっている。一方、放送局もインタネットを利用して番組の紹介や番組を補完する情報の提供など、さまざまなサービスを行っているが、より魅力的なサービスの開発が求められている。そこで我々は、新しい放送サービスとして放送メディアおよび放送局のメリットを最大限に活かすことのできる、インタネットへの情報提供サービスを検討した。

検討にあたっては、放送メディアの大容量の情報を多数へ同時に送ることのできる同報性と、インタネットなど通信メディアの個別、双方向性を利用するものとし、放送と通信の機能が相互に補完するものを目指した。提案する放

送サービスは、放送番組により共通の話題を持つコミュニティを作り、インタネットを使用してそのコミュニティに向けたきめ細かな情報提供を行う。ネット上のコミュニティで発表された意見や質問、著作物を放送に使用するなど、情報を流通させる双方向性も特徴としている。

今回、この新しい放送サービスを、インタネットの普及を受けて大きく変わりつつある学校教育のフィールドに適應させた、教育放送サービスとして開発するとともに、サービスを実現するために必要となる情報提供の仕組みとして、エージェント検索システムを試作した。

開発した教育放送サービスでは、全国の学校に向けた教育番組の放送により、共通の話題を持つコミュニティを形成し、コミュニティに参加する学習者の個別の興味に答えられる情報をインタネットにより提供する。さらに、学習結果や意見の発表、交換の場の提供まで含めて教育サービスと位置付ける。

エージェント検索システムは、このサービスにおいて不可欠な、放送局の持つ情報を学習項目と関連付けて学習者に提供する、マルチメディア辞典や図鑑のような機能を実現するために、放送局にある既存のデータベースを統合して利用する枠組みを持つ。

本論文では、開発した教育放送サービスの概要と、エージェント検索システムで実現した、放送局内の情報を利用するための枠組みと、検索システムの機能について述べる。

第2章では、新しい教育サービスを提案し、その目的とサービスの形態について述べ、第3章では、サービスを実現

2002年6月28日受付, 2002年10月15日最終受付, 2002年11月8日採録

[†] NHK放送技術研究所

(〒157-8510 世田谷区砧1-10-11, TEL 03-5494-3137)

^{††} NHK技術局

(〒150-8001 渋谷区神南2-2-1, TEL 03-5455-5493)

[†] NHK Science and Technical Research Laboratories

(1-10-11, Kinuta Setagaya-ku, Tokyo, 157-8510, Japan)

^{††} NHK Engineering Administration Department

(2-2-1, Jin-nan, Shibuya-ku, Tokyo, 150-8001, Japan)

するための核となる技術として開発した、エージェント検索システムの自然言語対話機能、統合情報提示機能の実現手法について説明する。最後に、まとめと今後の課題について述べる。

2. 新しい教育放送サービス

小中学校をはじめ、多くの学校がインターネットに接続され、検索サイトを利用した調べ学習やホームページの作成、メールの交換による情報の発信や意見の交換などの学習が始まっている。政府による教育施設への高速ネットワークやコンピュータの整備計画¹⁾の発表を受け、これまで教育番組として教室で視聴するといった形でしか用いられなかった映像や音声も、コンピュータを利用した授業の中で扱われるようになってきた²⁾。この大きく変わりつつある教育環境に対応するために、インターネットなどの通信メディアを利用した、新しい放送サービスを提案する。

図1は提案する教育放送サービスの枠組みを示す。このサービスは、放送メディアの特質である同報性と、インターネットの個別、双方向性を利用する。この枠組みで、放送番組は、共通の話題を持つ学習コミュニティを作る役割を担い、ネットワークは個々の学習者を満足させる、より細かな学習情報を提供する役割を果たす。さらに、放送局も含めた学習コミュニティ内での意見交換や意見発表の場の提供も含めて教育放送サービスとする。現在のところサービスの対象には、小中学校の生徒が行う調べ学習を考えている。

調べ学習は、自分たちで具体的な課題を見つけ、調べ、調べた情報をまとめて発表するという過程そのものを学習と位置付けるものである。対象に関する知識の習得はもちろんのこと、調べる作業や情報をまとめる際に使用する各種の道具の使い方、そのまとめ方、調べ方そのものも学習の対象となる。このような学習を支援するためには、高品質な情報を効果的に提供する学習システムが不可欠となる。田村は³⁾、学習システムの開発における重要ポイントを、「面白いこと(興味)」こそが学習のインセンティブになり、駆動力になる。「興味」を呼ぶ重要な要因の一つに実社会の文脈での位置付けがある」としている。

図1 教育サービス概念図
Interactive educational broadcast service.

図2 複合情報による知識の習得
Knowledge learning from complex information.

我々は、放送局内で使用しているニュース記事や番組情報のデータベースに蓄積された情報を学習項目と連携させ、学習者に興味を湧かせる情報として提供することにした。放送局に大量に蓄積されている質の揃った情報からは、新たな知識の発見が可能になると考えている。例えば、季節や地域によるニュース記事の分類により、日本の地理的な特徴や気候に関する知識を得たり、公害のニュースや映像を見ることで、環境問題をより現実的な問題として捕えることができるであろう。また、都会の子供にとっては、なじみの薄い農作物の実例を映像クリップで見て、それが自分の生活にどのように関わっているのかということをニュース記事や番組から知る。このように、放送局にある情報を使うことで、知識を言葉だけのものから深め、また広げる学習が可能になると考えている。

図2のように、学習内容と実社会との関連性を持たせた情報を学習者に提供し、学習者の興味を高めると同時に、より確かな知識として習得できるように、学習項目と放送局内のデータベースの情報を関連付けて面白く提供するための検索システムを開発した。

3. エージェント検索システム

すでに小中学校では、インターネット上の検索サイトを利用した調べ学習が行われている。しかし、一般の検索サイトで検索によって得られる情報は大量で、それを取捨選別できるスキルをもった者には極めて有効であるが、そのスキルを持たないものには、どれが必要な情報なのか判断することも困難な作業となる。また、小中学校の生徒にとって、検索結果を絞り込むための条件設定は容易ではない。検索サイトを利用した学習では、検索の対象となっている情報が、必ずしも学習向けの情報とは限らないこと、学習者がそのページを理解できなければならないこと、などの問題点も指摘⁴⁾されている。

人手でサイトの内容が教育に利用可能かを判断し、そのサイトリストを使用した教育用ロボット型検索エンジンの研究⁵⁾も行われ、実際に利用されている。しかし、典型的な日本語ロボット型検索サービスを使っても、適切なページが得られる割合は、1検索語で1%以下、2検索語で数%という調査⁶⁾もあり、調べ学習の中で信頼でき、学習に有

図3 エージェント検索システム画面
An example of agent search system.

効な情報を探し出すには、指導者(教師)の事前準備や支援が重要となる。さらに、前述のように、実社会との関連性や面白さが学習システムの重要なポイントとされているが、このような機能まで一般の検索サイトやWebサイトに要求するのは困難である。

一方で子供向けとされる検索サイトでは、多くのゲームやページが数多く用意されているものの、検索機能や検索結果の提示手法は、一般向けのものと大きくは変わるものではない。我々は、これらの点を考慮し、図3に示すようなWeb Browserから利用できるエージェント検索システムを開発した。

エージェント検索システムは、画面上に現れるキャラクターエージェントに対して学習者が質問すると、放送局のデータベースに蓄えられた動画や情報を使用してエージェントが答えるという、マルチメディア事典のようなシステムである。なお、表示には一般に普及しているWeb Browser(Internet Explorer 5.5~)を用いることで、特殊なアプリケーションのインストールを不要にしている。

エージェント検索システムの機能的な特徴は次の3点である。

(1) 操作の簡易化

自然言語質問文からのキーワード抽出により、キーワードの意識的な入力やAnd, Or制御の指定などの複雑な操作性を隠蔽する。

(2) 学習と関連付け可能な信頼できる情報の提供

放送局に蓄積された、番組制作情報やニュース記事など信頼できる多様な情報を、学習に利用できるように組合せ、情報間の関連性の視点もあわせて提供する。

(3) 面白く、わかりやすい情報提示

検索結果を文字情報の羅列だけでなく、画面に表示されたキャラクターの動作や合成音声など、簡易なシナリオにより定義された、動きのある情報として提示する。

エージェントという言葉に関しては、様々な解釈が存在するが、ここでは、「利用者の要求を理解し、その要求に合った処理の選択と実行および、処理失敗時の代替処理の

提案などを自動的に行うもの」という意味で使っている。以下にシステムの実現方法について詳しく述べる。

3.1 検索インタフェースとしての自然言語処理

エージェント検索システムの利用対象を小学校の生徒と想定したため、操作はできるだけ簡易にする必要がある。

本システムでは、検索に不慣れな学習者でもキーワードや検索式を意識せず利用できるように、自然言語による質問文での検索を可能にした。質問文章から、検索キーワードと処理内容を抽出し、検索処理を行い、検索結果は、質問への回答として文章を合成音声で読み上げる機能を持たせた。このような対話形式での検索を可能にするために、本システムでは、FACTS(FIPA実証プロジェクト)のエージェントテレビ⁷⁾において、ユーザプロフィール(名前や嗜好など)の獲得とシステムの制御を行うために作られた、対話処理プログラム(Natural Language Interface:NLI)を使用した。

NLIは、入力文の形態素解析(語の標準化を含む)により得られた語の出現パターンを、事前に登録した語の出現パターンと比較し、両パターンが一致した場合に、事前に定義した応答および動作を実行するものである。基本的には、ELIZA⁸⁾のパターンマッチング手法を拡張したものといえる⁹⁾。

今回のように、学習を目的として児童、生徒が検索に利用する場合、入力となる質問文が、非文法的な文章や、語の欠落、断片的な語であることが予想される。このような用途には、注目する語の出現パターンのみで質問の目的を判断する、テンプレートマッチによる対話手法は適している。また、音声認識による質問の入力を行う際に起こりうる、一部の語の誤認識や欠落にも対応することができる。

3.1.1 NLIの基本機能

入力文を形態素レベルに分割した単語と、それらの出現パターンを抽象的に表現するメタキャラクタを含んだ文字列でテンプレートを定義しておく。利用者から入力された質問文とテンプレートが一致するかどうか比較し、一致した場合に記述されている動作を行う。

このテンプレートを複数集めたものをテンプレート辞書ファイルと呼び、対話処理の知識情報に相当する。以下に対話処理を行うためにNLIが用いるテンプレート辞書、単語辞書、同義語辞書の各ファイルについて簡単に説明する。

(1) テンプレート辞書ファイル

基本テンプレートは、次のようにテンプレートパターンと返答文の2行で構成される。

テンプレートパターン
>返答文

(a) テンプレートパターン

入力となる質問文で想定される語の出現パターンを表1のメタキャラクタを使用して、図4のような構造で定義したものである。ここで「*」は、いわゆる欲張りマッチで、文字

表1 テンプレートを構成するメタキャラクタ
Template elements.

メタ キャラクタ	動作
*	0～無限個の単語に対応 (最長マッチ)
%	0～無限個の単語に対応 (最短マッチ)
+	1個の語に対応
!	一致してはいけない単語
[]	[]内で指定された単語もしくは存在しない
[]	[]内で指定された単語 (複数の場合は で区切る)
()	()内の単語はどのような順序で出現してもよい

図4 テンプレートパターンの構成
Template pattern.

図5 単語辞書ファイル
Terminology dictionary file.

列の最後まで読込んだ後、「*」の後ろに指定されたパターンが見つかるまで、1文字ずつ前に戻る処理を行う。「%」は、その逆で1文字ずつ増やしていく。このテンプレート内で使用される語も形態素解析され、問合せ文の形態素解析後のパターンと比較が行われる。複数のパターンがマッチした場合、パターン長の長いほうが優先される。

(b) 返答文

パターンが一致したときに返される文字列で、「>」を行頭に置いて、テンプレートパターンの次の行に記述する。検索システムではシステムが答える説明文章となる。

(2) 単語辞書ファイル

テンプレート内で多くの語を並列的に記述するのを防ぐため、語群を図5のように外部の単語辞書ファイルに記述し、それをファイル名で代表することができる。

(3) 同義語辞書ファイル

同義語処理は最初に行われ、同義語辞書ファイルにコロンで区切られた語のリストの2番目以降と同じものがあれば、リスト先頭の語に置き換えられる。

次の例では、「あゆ」「鮎」「香魚」のすべてが「アユ」に置き換えられて処理される。

アユ:あゆ:鮎:香魚

これらの要素を使用し、次に示すようなテンプレートを記述することができる。

タガメ * |どんな| *
>タガメはカメムシの仲間で、体調が6cmもある大型の水生昆虫です。

ここで、「たがめってどんな虫ですか?」という質問が入力されると、形態素解析、語の標準化が行われ、次のパターンが得られる。

タガメ | っ て | どんな | 虫 | です | か | ?

このパターンとテンプレートパターンがマッチし、システムからの応答として、「タガメはカメムシの仲間で…」が返される。

3.1.2 NLIの拡張

今回のエージェント検索システムでは、NLIの基本機能を拡張し、多数の語に関する質問、応答を効率的に記述する単語辞書ファイルの機能拡張と、検索結果を統合提示するための制御機能を実装した。前述のように、テンプレート辞書ファイルには、入力が予想される語の出現パターンを記述しておく必要がある。このため、さまざまな質問に対する応答をしようとすると、入力に予想される語の数(種類)と予想される出現パターンの積だけテンプレートが必要となり、記述量が増大する問題がある。

今回の検索システムでは、質問に含まれている用語の解説が重要な機能で、語ごとに異なる返答や動作をさせようとする、記述量が大きくなり、実用的ではない。そこで図6のように、テンプレートをパターンマッチ用のテンプレートファイルと用語辞書データベースファイルに分割して記述できるようにした。これにより、用語とその説明文を用語辞書データベースとして外部に置いて、テンプレート辞書ファイルから参照しながらパターンマッチングを行

図6 テンプレート辞書ファイルの拡張
Expansion of the template dictionary file.

うことができる。

用語群とその説明をファイル名mushiという用語辞書データベースファイルとして、用語、属性、説明を順に「:」で区切る次のような形式で登録しておく。なお、属性は省略可能である。

```
タガメ::タガメはカメムシの仲間で体長が…  
トンボ::トンボは幼虫のヤゴの…  
トンボ:WHERE:日本には約200種類が…
```

この用語辞書データベースは、テンプレート内では、\$x@を先頭につけたファイル名で指定し、ファイルに内包されるすべての用語に置き換えてマッチさせることができる。また、応答文中では、用語辞書データベースの説明部分を\$x@mushi.explainという記述で使うことができる。

```
$x@mushi *どんな*  
>$x@mushi.explain
```

さらに、「いつ」、「どこ」という情報を属性(WHEN, WHERE)をつけて用語辞書の中に記述しておけば、次のような利用も可能である。

```
$x@mushi *いつ*  
>$x@mushi.when  
  
$x@mushi *どこ*  
>$x@mushi.where
```

これにより、少数のテンプレートで多数の語に関する質問に回答できる。また、用語辞書データベースファイルは、形式が単純になるため、用語の追加や変更が容易である。ニュース記事などから用語辞書データを自動生成する実験も進めている¹⁰⁾。さらに、検索結果を提示する際の提示方法や画面上に現れるキャラクターの制御などを応答文の行末に記述する拡張も行った。詳細は3.3節で述べる。

3.2 複数データベースの利用

実社会の情報との連携により、学習内容に対する興味を沸かせるための方法として、放送局に蓄積されたニュース

記事などを含む番組制作情報を効果的に利用する枠組みを実現した。今回のシステムでは、以下の四つのデータベースを使用している。

(1) 用語辞書データベース

教育番組「おこめ」のWeb Siteで使用されている用語集データを利用した。語句の解説が中心となる用語集のデータは、前述の自然言語対話部分の中心的なデータとなるため、対話処理プログラムから直接利用する。

(2) 映像データベース

NHK内の番組制作部門が持つ映像データベースを複製して実験的に利用した。番組の制作に用いられた映像の一部とナレーション文章などがデータベースとして蓄積されており、単純なキーワード検索が可能である。

(3) ニュース記事データベース

NHKの持つニュース原稿のデータを基にして構築されたデータベースである。ニュース記事からインデックスとして利用できる話題を自動抽出¹¹⁾し、その話題ごとに分類し蓄積されている。

(4) 番組情報データベース

NHKが構築している番組アーカイブスにも利用される、メタデータが蓄積されたデータベースである。今回はデータ内容の一部と構造を複製した実験的なデータベースとして利用した。これらのデータベースは、それぞれ利用目的が異なり、複数の視点で収集された情報といえる。当然ながら、データの内容や構造、場所などが異なるため、操作の方法や提示方法が異なる。データベースごとに異なる内容や検索方法、出力形態の違い、データベースの場所の違いを隠蔽するために、図7のような統合化ユーザインタフェース(統合化UI)を中心に置き、各データベース管理する検索インタフェースプログラム(検索IF)と連携させる構成とした。各データベースを管理する検索IFは、キーワードをパラメータとして含む特定のURL(Uniform Resource Locator)で呼び出すと、検索結果をHTMLで返すCGI(Common Gateway Interface)やASP(Active Server Pages)として実現した。

図7 複数データベースを利用するソフトウェア構成
Framework for integrated use of multiple databases.

(呼び出し例)

http://WebServer1/videoSearch.asp?search=タガメ

以下に図7を用いて①～⑨までの処理手順とデータの流れを説明する。

統合化UIは、クライアントとなるWeb Browserから、ユーザの質問文章を受け付け①、対話処理プログラムに送る②、対話処理プログラムは、対話テンプレートとのマッチングを取る③、用語に対する問合せであれば、用語DBを使用して用語の解説を返答文として返す④。統合化UIは、返答文に制御文が含まれていれば、返答文と制御文を切り分け、その制御タグに対応したデータベースの検索制御を行う⑤。制御タグで示されるコンテンツの対象、例えば、映像検索を示す<searchVideo>タグと映像データベースのURLは一对一で対応しており、対象となるデータベースをURLの呼び出しにより検索する。各データベースに付加された検索インターフェースは、統合化UIから送られてきたURLに基づき、必要な検索処理を各データベースに対して行う⑥。検索結果⑦を受け取った検索IFは、それをWebページとして表示できるようにHTMLに組み込み、統合化UIに返す⑧。統合化UIは、そのページも含めてWeb Pageにまとめ、解説を行うキャラクターエージェントの制御や返答文の読み上げ制御と同期させ、利用者の Web Browserに表示する⑨。

3.3 検索結果の統合表示

前述のように教育システムの開発においては、学習者が楽しいと感じられることが重要なポイントとしてあげられている。

本システムでは、複数のデータベースから検索された情報を単なるリストとして提示するのではなく、画面上に配置したキャラクターエージェント (Microsoft Agent Technologyを使用) が検索情報を解説しながら提示する。また返答文は、音声合成を使用し、キャラクターエージェントがせりふとして読み上げることで、利用者の興味を引くとともに、検索結果中に含まれる難しい漢字に対応した。データベースに持っている項目にマッチする検索が行われた場合には、エージェントが「いい質問だ」というように誉め、利用者の意欲を高めたり、なにも検索できなかった場

合には、謝ったりというアクションを組み込み、親しみ易いものにした。これらの機能を実現するために、3.1で述べた返答文の行末に、データベース検索制御、情報提示制御コマンド、キャラクターエージェントの動作制御などが記述できるように拡張した。書式は次のように「<」「>」で囲んだ制御タグとして記述する。

>タガメは…	<searchVideo タガメ>
説明文	制御文

その制御タグを解析し、制御を行うプログラムを統合化UIにJavaScript, VBScriptによる関数として実装し、Web Browserで表示している。検索、および表示に関しては、表2に示す7種類、エージェントの動作に関しては、Microsoft Agentの動作の一部を検索結果の提示用途に対応させた、表3の9種類の関数を実装している。

検索の主機能として実装した、三つの制御タグが指定されたときの表示画面を図8(a)～図8(c)に示す。返答文は、{}に囲まれる範囲に、説明文と制御タグを組合せた動作を複

表2 検索制御タグ
Control tags for searching.

動作	制御タグ
静止画表示	showImage
映像再生	playVideo
映像検索	searchVideo
番組検索	searchProg
ニュース検索	searchNews
" (キーワード+地図表示)	searchKNews
" (キーワード+地図非表示)	searchKNewsNoMap

表3 エージェント制御タグ
Control tags for character agent action.

動作	制御タグ
指定場所へ移動	moveto
挨拶	greet
拍手して喜ぶ	happy
説明読み上げ	readYougo
指差しポインティング下	pointingDown
指差しポインティング左	pointingLeft
指差しポインティング右	pointingRight
質問待ち	qWait
困惑	confused

(a) 映像検索
<searchVideo>

(b) ニュース検索
<searchNews>

(c) 番組検索
<searchProg>

図8 各検索制御タグによる検索結果 (例)
Search result of control tags.

```

%x@yougo %なに *
>{うむ、いい質問だ。$x@yougoとは、こういことだ。<happy> ... ①
$x@yougo.explain<readYougo> ... ②
$x@yougoの映像を検索してみました。興味のある…<searchVideo $x@kome 900 600> ... ③
$x@yougoに関するニュースも検索してみました。<searchKNewsNoMap $x@kome 900 600> ... ④

```

図10 シナリオの実行

Execution flow of scenario.

図9 統合情報提示

Integrated information presentation.

数記述できるように拡張し、検索結果の提示順序や方法を指定するシナリオとして機能させた。用語の説明を終わった後にデータベースの検索を行うといった、一つの質問に対する回答を複数のデータベースから取出し、提示することができる。

図11 シナリオの表示結果（映像、ニュース検索時）

Result of scenario (Video Clips and News).

図9に、最も基本的な動作として試作システムに実装した、用語解説と映像検索を順次行う際のシナリオ記述例とその表示例を示す。

ニュース記事から抽出して作成した用語辞書ファイル(yougo)に登録された語に関する質問が行われた場合には、[用語解説][映像検索][ニュース記事検索]の三つを続けて行う下記のようなシナリオを返答文として記述した。返答文は、統合化UIで1行ずつ取出し、①～④に示す各行の返答文の文字数から計算した時間だけ遅らせ、図10のように順次実行する。

この返答文が実行されると、解説文の読み上げをした後、映像検索、ニュース記事検索を行い、画面の左側に、ニュース記事の一覧が図11のように表示される。映像の検索結果に続いてニュース検索を行う場合には、地図を表示しない検索を行うタグ<searchKNewsNoMap>を指定することで、映像とニュース記事の検索結果が同時に一覧できる。また、「稲刈り」といったテレビ番組としてよく取り上げられるような語の場合には、返答文を次のようなシナリオとして設定し、番組データベースもあわせて検索する。

%稲刈り%*

>|うむ, いい質問だ. 稲刈りとは, ということだ. <happy>

実った稲を刈りとることです. 昔はかまなどを使って, 一株ずつ根元から刈り...<readYougo>

稲刈りの映像を検索してみました. 興味のあるものを...<searchVideo 稲刈り 900 600>

稲刈りに関するニュースも検索してみました. <searchKNewsNoMap 稲刈り 900 600>

番組を検索してみました. <searchProg 稲刈り 900 600>

|

番組データベースの検索結果は, 図12の例のように映像データベースの検索結果と同じ場所に表示されるので, 映像検索結果が見えなくなってしまうが, 画面左上のボタンをクリックすることで直前の, 映像, ニュース, 番組の各検索結果をいつでも表示できる. また, 各データベースの検索後に利用者から「ニュース」や「番組」という言葉を含む質問が入力されれば, 直前に入力された質問文章から抽出したキーワードを利用して該当するデータベースを検索, 提示することができる. これを実現するために, キーワードの保持機構を対話処理プログラムに持たせている.

ここで示した統合表示方法は, 実装の一例でありコンテンツの内容や, 利用対象等に応じてコンテンツの提供者がシナリオとして自由に組み立てることができる.

3.4 他のシナリオ技術との比較

インターネットコンテンツの記述言語の代表的なものとしてSMIL¹²⁾がある. SMILは, SMILプレイヤーと呼ばれる専用のソフトウェアで再生することを前提に, 動画を主とする, 各種メディアの同期関係に注目した再生手順や表示位置などを記述し, リッチなマルチメディアコンテンツとして提示するために開発された言語である. 一方, 本システムのシナリオは, データベースに蓄えられた情報をコンテンツ化して提示するために特化した記述方法として設計した. エージェントが話すせりふ(説明文章)を中心に, 検索制御タグを組合せる簡易な構文によるテンプレートを用意しておくことで, データベースに蓄積された情報を, エージェントのせりふにあわせて, 順次Web Browser上に提示することができる.

4. む す び

本論文では, 放送の特徴とインターネットの特徴を活かした, 新しい放送サービスを提案し, その適用例となる放送教育サービスとそこで使用するエージェント検索システムについて報告した. 放送局に蓄積されている各種の情報を有効に活用した, 新しい情報提供サービスの枠組みを構築できた. また, 検索結果を組合せて提供する, 簡易シナリオ機能の実装により, 従来の検索システムから一歩進んだコンテンツ提供システムが構築できた.

約400の語を持つ用語データベースと接続されたテンプレートによる対話検索では, 質問入力後, 約2秒で返答となる解説文が得られる. 映像検索, およびサムネール表示でも約2~3秒の処理時間で検索結果が得られ, Webインタフェースを持つデータベースとしては, 十分な応答速度と考える. また, 映像検索の多くは, キャラクタエージェントが用語解説を行っている期間に処理されるので, 利用者には待ち時間として感じさせない. 質問文章による検索結果が, ユーザの質問意図を反映しているかに関しては, 対話テンプレートの構成や記述量で大きく左右されること, 試作したシステムでは, 発見的な学習の支援を目的として, できるだけ多くの検索結果が得られるようにテンプレートを設計していること, などもあり, 十分な評価を行っていない. しかし, 試用時のユーザの反応からは, おおむね良好なものであると考えている. 今後の利用実験を通して評価を進める予定である.

試作したシステムは, 2001年11月に行われた「NHK 教育フェア2001」で展示した. 展示では, 将来の教育放送の利用形態をイメージした, 30分弱のデモ授業として来場者に操作を体験してもらった. アンケートは実施していないが, 授業中の利用者の観察と授業後の聞き取りを適宜行った.

参加した子供は, 想定した小学校高学年より年齢の低い子供もいたが, 用語解説をキャラクタエージェントが音声合成で読み上げることで, 漢字に関する問題もなく, 検索結果の提示を集中して見ていた. 利用した子供たちからは, 「おじいさん(エージェント)が面白い」などの声があり, エージェントの動きや応答(しゃべり)による学習への導入, 提示情報のシナリオ化が有効に機能しているものと判断した. また, ニュース記事など, 社会の実例と関連させた学習教材として, 来場した教育の専門家(教師, 教材ソフト開発者)から高く評価され, すぐにでも使いたいとの要望も受けた.

図12 シナリオの表示結果

(映像 [隠れている], ニュース, 番組, 検索時)

Result of scenario.

(Video Clips [hidden], News and TV program information)

参加者は、途中で席を離れることもなく熱心に操作しており、数量的な評価が困難な教育システムでは、貴重な実験結果が得られたと考える。NHK内の教育番組制作者からも高く評価され、現在実験中のデジタル教材の一部として組込む予定である。

今後は、教育現場におけるシステムの利用実験により学習効果の検証と改良を進めていくとともに、活動したことをまとめて発表する学習過程を支援する学習レポートオサリングシステム¹³⁾の開発とあわせ、新しい形の教育放送サービスの実現を目指す。また、複数データベースのアクセスプロトコルの汎用化を進め、より簡易な方法で高度なデータベースの統合利用を可能にし、教育だけでなく広範囲な用途に利用できるシステムとする。

本研究にあたり、映像の提供や各種のアドバイスを頂いたNHK番組制作局菊江賢治氏に感謝する。

〔文 献〕

- 1) 文部省学習情報課：“「ミレニアム・プロジェクト」により転機を迎えた「学校教育の情報化」－「総合的な学習」中心から「教科教育」中心へ”，(2000)，http://www.manabinet.gr.jp/it_ed.pdf
- 2) NHKフルデジタル教材：“おこめ”，<http://www.nhk.or.jp/okome/>
- 3) 田村：“教育システム情報/オーバビュー”，教育システム情報学編，教育システム情報ハンドブック，実教出版，pp.3-8 (2001)
- 4) 近藤：“既存WWW上の資料の教材としての利用”，教育システム情報学編，教育システム情報ハンドブック，実教出版，pp.229-232 (2001)
- 5) 新井，小棚木，石橋，斎藤，篠原，河波：“教育用ロボット型検索エンジン”，信学技報，教育工学研究会，ET99-42 (1999)
- 6) 山本，清水：“環境教育に関連するキーワードをインタネットで検索した場合の学習情報に関する一検討”，日本教育工学誌，23，Suppl，pp.63-66 (1999)
- 7) 村崎，金，柴田，重野：“FACTS プロジェクトにおけるエージェントテレビの研究開発”，信学技報，AI2000-25，(2000)
- 8) Stuart C.Shapiroほか，大須賀節雄監訳：“人工知能大辞典”，丸善 (1991)
- 9) 住吉，山田，村崎，金，井上，八木：“エージェントを利用したマルチメディア教育支援システム”，信学技報，教育工学研究会，ET2001-35 (Sep. 2001)
- 10) 山田，柴田，金：“ニュース原稿を利用した用語集作成の検討”，言語処学第8回年次大論集，B4-4，pp.507-510 (2002)
- 11) 山田，柴田：“ニュース記事の定型性を利用した話題要約の検討”，言語処学自然言語処理研究，NL148-7，pp.45-50 (2002)

12) SMIL (Synchronized Multimedia Integration Language)，<http://www.w3c.org/TR/REC-smil/>

13) 住吉，八木，柴田：“ネットワークによる意見交換を目的とした映像レポート作成システム”，映像学メディア工学研究，ME2002-8，pp.1-6 (2002)

住吉 英樹 1980年，広島県立広島工業高校電気科卒業。同年，NHKに入局。広島放送局を経て，1984年より，放送技術研究所に勤務。現在，マルチメディアサービス業務実施グループにて，コンピュータを応用した番組制作システムの研究に従事。正会員。

山田 一郎 1991年，名古屋大学情報工学科卒業。1993年，同大学院修士課程修了。同年，NHKに入局。名古屋放送局を経て，1996年より，放送技術研究所に勤務。現在，マルチメディアサービス業務実施グループにて，自然言語処理，データマイニングに関する研究に従事。正会員。

村崎 康博 1990年，福井大学工学部電子工学科卒業。1992年，同大学院修士課程修了。同年，NHKに入局。1995年より，放送技術研究所勤務。1997年より，放送におけるエージェントの研究に従事。現在，同技術局特許部に所属。正会員。

金 淵培 1983年，サンパウロ大学工学部卒業。1984年，来日。(株)ブラビス・インターナショナルに入社し，機械翻訳システムの研究・開発を行う。1991年，NHK入局。現在，NHK放送技術研究所ヒューマンサイエンスに所属。ニュース用機械翻訳，自然言語処理，エージェントの研究などに従事。工学博士。正会員。

八木 伸行 1980年，京都大学大学院電気工学専攻修士課程修了。同年，NHKに入局。甲府放送局，放送技術研究所，技術局，編成局を経て，現在，放送技術研究所(研究企画)担当部長。画像処理，映像処理，番組制作技術，デジタル放送システムなどの研究開発に従事。工学博士。正会員。

柴田 正啓 1981年，京都大学大学院工学研究科電子工学専攻修士課程修了。同年，NHKに入局。1984年より，同放送技術研究所勤務。情報検索，映像インデキシング，映像データベース，映像制作システム，エージェント技術の放送への応用の研究，およびMPEG-7の標準化に従事。正会員。