

エージェントを用いたCAIの高機能化

山崎健弘† 千種康民†

東京工科大学 情報工学科

本研究では、エージェントという概念を用いてCAIをより高機能化できないかを考え、対話型日本語インタフェースを目的にシステムの開発を行った。

このシステムは、先生というエージェントを作ることにより教員の手間を省くためのシステムである。複数の生徒用コンピュータと一台の教員用コンピュータをLANで接続し、生徒からの質問のうち、知識ベースにあるような簡単な質問に対してはエージェントが答えてくれる。また、知識ベースにない質問は教員のコンピュータに送信され、直接教員が回答することになる。このシステムの導入によって、教員の負担を減らせると共に生徒の学習効率を高めることができる。生徒から見ればあたかも一人の教員を独占して教育を受けているかのように見えるのである。このように見ると、このシステムの特徴は情報フィルタとして機能するところにある。また、知識ベースにない答えに教員が回答すると同時に知識ベースに追加する。

曖昧な日本語をどのように認識させ、知識ベースから解答を引き出すか処理するかが今後の研究課題である。

Intelligent CAI with a Teacher's Agent

Takehiro YAMAZAKI and Yasutami CHIGUSA

Dept. of Information Engineering Tokyo Engineering University

E-mail chigusa@cc.teu.ac.jp

We thought more optimization for Computer Assisted Instruction(CAI) in use of a concept, 'agent', and developed one intelligent CAI for Japanese text base.

In this system, making the agent as a teacher will decrease teacher's labor. More than 2 computers are connected with LAN in our system, where there are derived one teacher's computer and are one more students' computers. The answers for easy questions are from the existing knowledge base. On the other hand, if they are not existing there, the questions are transmitted to the teacher's computer, and then the answers are replied to the student's computer. At the same time these answers are appended to the knowledge base. This intelligent CAI system decreases the teacher's labor, and enables to raise students' performance for studying. Observing from students, as if they had their own teachers. In this viewpoint, the peculiarity of this system is to have a function as an information filter, and to learn new knowledge. In the learning-function, the teacher answers for only the questions which the intelligent CAI can not answer for. At the same time, the new knowledge is added to the knowledge base automatically.

In our reserch, there are the problems how to recognize fuzzy Japanese questions and how to bring out adequate answers in the knowledge base.

1 はじめに

近年、エージェント指向という概念が注目されている [1]。この概念が定着するようになるにはより人工知能（これ以後 AI と呼ぶ）の実現レベルが高くならなければならないだろう。本研究では、対象を学校教育用の CAI に絞りエージェントという概念を用いてより効率的な使い方ができないかを検討してみた。

現状の CAI (Computer Assisted Instruction) において、ディスプレイに表示される説明さえも分からない場合に、教員は一人しかいないので質問者が多い場合には全てに対応できないという問題点がある [6]。そこで、CAI を高機能化するためにユーザインタフェースとエージェントを用いた簡単な推論を組み合わせた対話システムを開発した。

2 本知的 CAI システムの特徴

質問の多いことからあらかじめ教員が CAI システムに学習させ、CAI システムが答えられるような簡単な質問は CAI システムが答え、そうでないものは直接教員が答える。これにより、

- 擬似的な質問返答者である先生（エージェント）を作ることができる。
- 生徒はあたかも教員と直接対話しているかのように見える。
- エージェントが答えられない質問のみ直接教員が答えるので、教員は楽ができる。

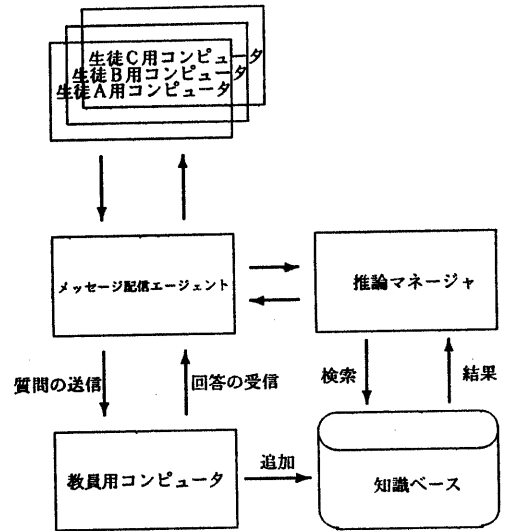


図 1: 知的 CAI システムの構成

- 教員からの知識学習により、エージェントは賢くなり、他の生徒から同じ質問があった場合には自動的に返答できるようになる。

3 知的 CAI システムの構成

CAI の教育分野を中学校程度の歴史に限定し、コンピュータへの入力、コンピュータからの出力は全てテキストベースとして応答をリアルタイムで行うようにした。

3.1 問い合わせの流れ

システムは、推論マネージャとメッセージ配信エージェントにより、まとまりのあるシステムにした。図 1 のような構成にした。

3.2 最長一致法によるトークン切り出し

最長一致法とは、文字列をマッチングさせるときに最も長い文字列にマッチングさせる方法で、仮名漢字変換 (FEP) 等に用いられている。本来、日本語解析では意味解析を行わなければならない正しいトークン切り出しは行え

いつ聖徳太子は摂政に成ったの？

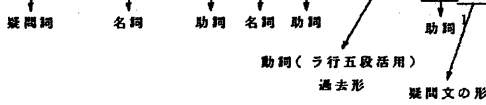


- ①の段階では“い”は辞書にないのでもう1文字読む
 - ②の段階で辞書に“いつ”という語が見つかり疑問詞の属性を返す
 - ③の段階では“聖”は辞書にない
 - ④の段階で“聖徳太子”が辞書に見つかり名詞の属性を返す
- 以後同様にして解析していく

図 2: 最長一致法

入力文

いつ聖徳太子は摂政に成ったの？



1: 疑問文の語尾という属性をもたせてある

図 3: 品詞分解

ないが、簡単のため最長一致法を用いることにした。最長一致法についての過程を図 2 に示す。

このように、1文字ずつ順次に読んでいき、曖昧さがなければ、最長一致法によって図 3 のように品詞分解される。

3.3 形態素解析と形態素生成

形態素解析によってトークンの切り出しを行うが、日本語の中には活用するものがいくつもある。それら一つ一つをトークンとして登録していたらさらに膨大な量のトークンになってしまう。そこで、活用の種類に応じて語句を活用させて解析を行うようにした。動詞の活用表を表 1 に示す。

3.4 bison による構文解析

構文解析は、bison を用いて行った。bison は LALR(1) 文法に従って解析を行うツールである。

bison や yacc はもともとコンパイラやインタプリタの自動生成のために作られたツール

活用形	第一段 (未然形)	第二段 (未然形)	第三段 (連用形)	第四段 (連用形)	第五段 (終止形)	第六段 (連体形)	第七段 (仮定形)	第八段 (命令形)
五段活用 (四段形止む)	ア	オ	イ	いん	ウ	ウ	エ	エ
上一段活用		イ	イ	イ・る	イ・る	イ・れ	イ・ろ	イ・ろ
下一段活用	エ		エ	エ・る	エ・る	エ・れ	エ・ろ	エ・ろ
カ行変格活用	こ		き	くる	くる	くれ	こい	こい
サ行変格活用	しせ	し	し	する	する	すれ	せよ	せよ

表 1: 動詞の活用

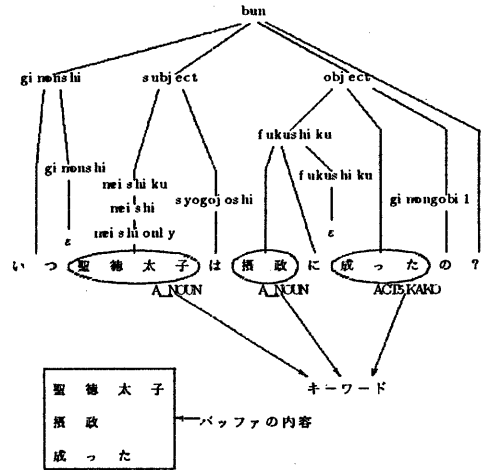


図 4: 構文解析木とキーワード抽出

である。従って、自然言語のように曖昧な表記は許されない。そういった制限上、構文はほぼ一通りに定まった文法となる。

図 4を見ると、文全体は、<疑問詞><主語><目的語><述語><疑問符>のようになっていいる。”いつ”や”どこ”などの疑問詞は疑問詞という属性に当てはめる。同様に、名詞は名詞という属性に、形容詞(終止形)は形容詞という属性に当てはめている。このようにすることで、数多くの単語を品詞という部類に分けることができる。そうすると一つ一つのトークンを用意しなくてもよくなり、簡潔化できる。解析の結果からキーワードが抽出される。

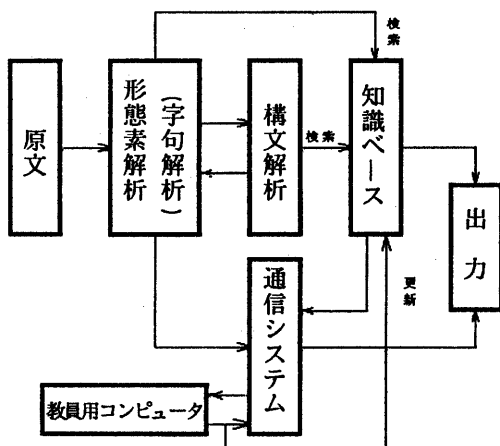


図 5: システム全体の流れ

3.5 知識ベース検索

知識ベース検索は3.4節で説明したキーワード抽出の結果を用いる。バッファに格納されたトークンを用いて完全一致探索を行い、一致数の最も多い一文を選び出力する。

3.6 知識ベースの学習

解析が失敗したり、知識ベースに該当する文が存在しなかった場合には、教員からの解答を待ち、その解答を知識ベースに追加する。学習を行うことで、以後、他の生徒から同じような質問があった場合にはキャッシングされる。すなわち教員への通信による質問は次第に減るはずである。

● 学習

図1の経路のうち、教員用コンピュータ→メッセージ配信エージェント→推論マネージャ→知識ベースというパスを経る。

システム稼働中の知識ベースの学習は、教員が生徒からの質問に対しカンマでトークンを区切って入力する。

入力の例)

1923年,に,関東大震災,が,起きた,。

3.7 通信システム

本研究の通信は、推論マネージャが知識ベースを検索し該当するデータがなかった時と、形態素解析や構文解析がうまく行かなかったときに、直接教員のコンピュータに質問内容を送信し、教員がその質問に答え、生徒に送り返すものである。

パソコンのLANのファイル共有機能を用いて、ファイル単位の送受信を行うことによってメッセージの交換を実現した。具体的には、生徒にID番号をもたせ、ID番号を含んだファイル名¹が存在するかしないかによって送信が行われているかいないかを判別する。

3.8 データ構造

データ構造という観点からは、知識がある基本的な単位に基づいて記述されていて、知識の追加や削除あるいは修正が容易にできるような表現形式が望ましい。このことを考慮すると、エディタで簡単に知識を追加、削除、修正できるような形が望ましいので、辞書は全てテキスト形式とした。また、本システムで用いる辞書には2種類あり、トークンベースと知識ベースである。トークンベースと知識ベースの違いは以下のようになっている。

トークンベース 知識ベースで使用するキーワードを取り扱う。

知識ベース 実際にCAIで使用する知識を取り扱う。

3.8.1 トークンベース

トークンベースは検索テーブルを用意してポインタでリンクをしておく。これを用いて、二分探索で検索する。このようにすると高速検索が可能である。

トークンベースの例
成, な, ラ行五段

¹MS-DOSのファイル名の限界により最高でも8文字+拡張子3文字である。

聖徳太子, しょうとくたいし, 固有名詞
摂政, せっしょう, 一般名詞

⋮
⋮

3.8.2 知識ベース

テキスト上での表現は, 知識ベース検索がしやすいうようにカンマで区切って表記する. 概ね次のような形になる.

593年, に, 聖徳太子, が, 摂政, に, 成った.,

⋮
⋮

4 システム実行例

以下における'OK!'は構文解析が成功したことを, 'searching'は知識ベースを検索中であることを示す.

4.1 質問に対する回答を持つ場合

(生徒用コンピュータ)

質問>いつ聖徳太子は摂政に成ったの?
Ok!
searching

回答>聖徳太子は593年に推古天皇の摂政に成った.

次に, 人名を聞く形として, 次のように文を入力してみる.

(生徒用コンピュータ)

質問>1016年に誰が摂政に成ったの?
Ok!
searching

回答>1016年に藤原道長が摂政に成った.

4.2 質問に対する回答を持たない場合

(生徒用コンピュータ)

質問>誰が万有引力を見つけたの?
Ok!
searching

知識ベースにありません.
お待ちください.

(教員用コンピュータ)

ID No.1の生徒からの質問>誰が万有引力を見つけたの?

回答>ニュートン, が, 万有引力, を, 見つけた.,

このとき, 知識ベースにこの回答が追加される.

(生徒用コンピュータ)

回答>ニュートンが万有引力を見つけた.

5 ネットワーク上のエージェントの将来

もし, 巨大なシソーラスが完備できAIがほとんど完全に使えるようになり, 機械翻訳の実現レベルが非常に高くなったとき, エージェント指向は強く現れるだろう. エージェント指向という考え方をを用いると次のようなことが実現できるだろう [1]. 例えば, 英語の文献を本研究のような対話形式で検索するシステムを使う場合を考える. 仮にこのシステムは, 入出力が全て英語であるエージェントで, 他に和英翻訳エージェント, 英和翻訳エージェントが存在するとする. エージェント指向になっていれば, ここで日本人ユーザが日本語でその文献をあたかももともと日本語の文献であったかのように検索することができるだろう. どのような動作をするかという点, まず, 入力した質問文が和英翻訳エージェントによって英語の質問文に翻訳される. 次に, その質

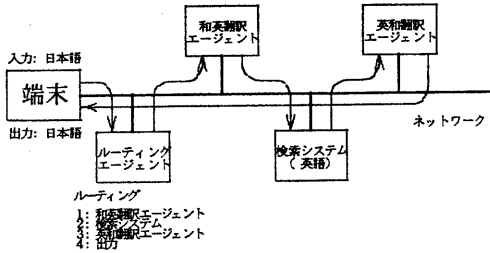


図 6: ネットワークを利用したエージェント

問文から知識ベースを検索して英文による出力をする。最後に、英和翻訳エージェントにより、和文に翻訳する。といった具合である。システムが全てエージェント指向になっていれば、文献検索システムでなくて他の英語のみでしか使えないシステムでも日本語で使えるのである。

このように見ると、エージェントはオブジェクトに AI を組み込んだようになっている。これがネットワーク上で動いていたら、いろんな経路を通ることで天文学的数字の応用が利くだろう。図示すれば図 6 のようになる。

6 おわりに

定型文でかつ曖昧性のない文章（最長一致法でうまくいったもの）では、きちんと構文解析はできている。初期の目標とする知的 CAI の実現に一步近づいているといえる。ただし、意味解析を全く行っていないので、単に文中にでてきた名詞の数の多少で知識ベースから引き出す文が決定してしまう。文中で何が重要視されていて何をユーザが聞きたいのかを文中から見つけた措置を取らなければ効率よく検索できない。

今後の課題として、

- 意味ネットワークや概念ネットワークを用いて意味解析を行うこと。
- 形態素解析、構文解析、意味解析の間で何

度でもバックトラックできるようにする。

- DCG (PROLOG での確定節文法) 等を用いて語順の融通性を持たせる。

参考文献

- [1] 所 真理雄 著：日経バイト'94 10 月号 創刊 10 周年記念 特別インタビュー 「エージェント指向が今後 20 年間話題の中心になる」 pp.116 ~ 121
- [2] ソフトバンク：C-MAGAZINE '94 9,10 月合併号 「日本語処理系学」 pp.22 ~ 52
- [3] 五月女 健治 著：bison, flex 啓学出版
- [4] 太原 育夫 著：人工知能の基礎知識 近代科学社
- [5] 疋田 輝雄・石畑 清 著：コンパイラの理論と実現 共立出版
- [6] 大須賀 節雄 編：ヒューマンインタフェース 「自然言語理解におけるヒューマンインタフェース」(石崎 俊 著) 「CAI とヒューマンインタフェース」(大槻 説平 著) オーム社
- [7] 石崎 純夫 監修/森 文彦・花岡 かほる 共著 エキスパートシステム構築技法入門 「システム AI をめざして」 オーム社
- [8] 情報処理学会：情報処理 '93 11~'94 3 月号
- [9] 長尾 真 著：NTT Communication Magazine 情報人 '94 秋号 情報大学