

エネルギー科学研究科
エネルギー社会・環境科学専攻修士論文
題目： 音声対話インタフェースを用いた
CAI システムに関する研究

指導教官： 吉川 榮和 教授

氏名： 山本 専

提出年月日： 平成 12 年 2 月 7 日 (月)

論文要旨

題目： 音声対話インタフェースを用いた CAI システムに関する研究

吉川榮和研究室， 山本 専

要旨：

化石燃料の枯渇、地球温暖化などのエネルギー問題は年々深刻になってきている。このような問題は全ての人々が知っておくべきで、エネルギー問題について教育する必要がある。この要求を満たす方法の一つにコンピュータによる教育支援 (CAI:Computer Assisted Instruction) がある。

本研究では、一般の人々を対象にエネルギー教育を行うための CAI システムとして音声対話インタフェースを用いた CAI システムを実現することを目的に研究を進めた。

本論文ではまず、研究の背景として一般の人々に対するエネルギー問題についての教育の必要性に触れた後、それを実現するために CAI システムを構築することを述べる。そして、従来の CAI システムの研究について展望する。その問題点として初心者に馴染みにくい入力インタフェースの問題と一人での学習のみを実現したの多いことを述べ、その解決のために音声対話インタフェースを用いること、同期分散型 CAI システムとしての機能を実現することを述べる。この音声対話インタフェースの要素技術である音声対話技術を概説し、本研究の目的を述べる。

次いで、本研究で構築する音声対話インタフェースを用いた CAI システムに必要な機能について検討し、学習の方法を講義と質問に対する回答から構成すること、教材を一元管理してクライアントサーバ方式でシステムを実現することを述べる。

構築する CAI システムは一般的な音声対話システムの構成に基づき、音声認識部、音声合成部、文解析部、対話管理部、タスク処理部から構成される。講義、質問回答のためのデータは教材データベースに格納されている。また、質問文解析処理は音声認識された発話テキスト文字列に対するパターンマッチングとキーワード抽出によって行う。抽出したキーワード群をデータベースから検索することで音声による質問を理解し、それに回答する機能を実現する。

さらに、この構築した CAI システムについて動作確認した結果、意図した通りに動作することを確認した。また、このシステムを CAI システムとして評価するために評価実験を行った。その結果からシステムの有用性がおおむね認められた。同時に、問題点も明らかになり、自由な質問文への対応が不十分であること、さらに高度な対話の実現について検討する必要があることが分かった。

最後に結論と今後の課題を述べ、本研究を総括している。

目次

第 1 章 序論	1
第 2 章 研究の背景と目的	3
2.1 研究の背景と目標	3
2.2 CAI システム	4
2.2.1 各 CAI 技術の分類と特徴	5
2.2.2 各 CAI 技術の比較	9
2.2.3 従来の CAI の問題点	12
2.2.4 CAI システムに音声対話インタフェースを用いることの利点	14
2.3 音声対話システム	14
2.3.1 音声認識	15
2.3.2 文解析	15
2.3.3 対話管理	16
2.3.4 タスク処理	17
2.3.5 音声合成	17
2.4 昨年度構築した会話エージェント型 CAI システムとその問題点	17
2.5 本研究の目的と提案する CAI システム	19
第 3 章 構築する CAI システムの機能	21
3.1 システム機能の特徴の検討	21
3.2 想定する学習の進め方	23
3.3 構築する CAI システムの機能設計	25
3.3.1 音声対話機能	25
3.3.2 講義機能	27
3.3.3 質問回答機能	29
3.3.4 複数人での学習機能	29
3.3.5 学習者の発話についての考察	31

第 4 章 CAI システムの設計と構築	34
4.1 CAI システムのハードウェア構成	34
4.2 CAI システムのソフトウェア構成	34
4.2.1 教材データベースの構成	49
4.2.2 講義の進行、分岐の方法	56
4.2.3 質問文解析方法	58
4.2.4 複数学習の実現方法	60
4.2.5 構築したシステムに用いたソフトウェア	62
4.3 システムの構築と実験に用いたハードウェアの詳細	65
4.3.1 インタフェースを受け持つパーソナルコンピュータ	65
4.3.2 サーバコンピュータ	66
第 5 章 システムの動作確認と評価実験	68
5.1 構築したシステムの動作確認	68
5.1.1 音声対話機能	69
5.1.2 講義機能	69
5.1.3 質問回答機能	70
5.1.4 複数学習機能	70
5.1.5 動作確認のまとめ	70
5.2 評価実験	70
5.2.1 評価の目的	70
5.2.2 実験方法	71
5.2.3 実験結果	72
5.2.4 考察	72
5.2.5 システムの問題点	78
5.2.6 実験のまとめ	79
第 6 章 結論	80
謝 辞	82
参 考 文 献	83

目 次

2.1	フレーム型 CAI の学習コースの分岐	6
2.2	ITS の構成例	7
2.3	WWW 型 CAI システムの一例	8
2.4	一般的な音声対話システムの構成	15
3.1	複数人での学習のイメージ図	24
4.1	構築する CAI システムのハードウェア構成	35
4.2	CAI システムのソフトウェア構成	36
4.3	CAI システムの詳細なソフトウェア構成	37
4.4	CAI システムのソフトウェア構成 (音声認識部)	38
4.5	音声レベル検出 (通常の発話)	39
4.6	音声レベル検出 (雑音などの場合)	39
4.7	句点挿入処理	40
4.8	CAI システムのソフトウェア構成 (音声合成部)	41
4.9	CAI システムのソフトウェア構成 (文解析部)	41
4.10	文解析部の処理の流れ	42
4.11	CAI システムのソフトウェア構成 (対話管理部)	46
4.12	対話管理部の処理	47
4.13	CAI システムのソフトウェア構成 (タスク処理部)	48
4.14	データベースの構成	50
4.15	講義の構造	54
4.16	講義の分岐の処理	57
4.17	質問文解析処理 (意味)	59
4.18	質問文解析処理 (原因・理由)	61
4.19	複数学習者での学習環境を構築するときのシステム構成	62
5.1	学習者に提示される画面例	68

表 目 次

2.1	本研究で対象とする CAI によるエネルギー教育の特徴	3
2.2	WWW の特徴	9
2.3	各 CAI システムの比較	10
3.1	システムの特徴についての考察	23
3.2	音声対話の機能構成	26
3.3	講義の機能構成	28
3.4	質問回答の機能構成	29
3.5	複数人での学習に関する機能	30
4.1	発話の分類と解析用単語	44
4.2	質問の分類と解析用単語	45
4.3	講義テーブルの仕様	50
4.4	分岐テーブルの仕様	50
4.5	名前変換テーブルの仕様	51
4.6	意味テーブルの仕様	51
4.7	原因・理由テーブルの仕様	51
4.8	結果テーブルの仕様	51
4.9	対応・対策テーブルの仕様	51
4.10	講義テーブルの例	52
4.11	分岐テーブルの例	52
4.12	名前変換テーブルの例	52
4.13	意味テーブルの例	52
4.14	原因・理由テーブルの例	53
4.15	結果テーブルの例	53
4.16	対応・対策テーブルの例	53
4.17	システムの構築と実験に用いたパーソナルコンピュータの仕様	66
4.18	システムの構築と実験に用いたサーバコンピュータの仕様	67

5.1	評価実験の結果 (1)	73
5.2	評価実験の結果 (2)	74
5.3	評価実験の結果 (3)	75
5.4	評価実験の結果 (4)	76
5.5	実験での学習者の発話 (被験者 1)	76
5.6	実験での学習者の発話 (被験者 2)	77
5.7	実験での学習者の発話 (被験者 3)	77

第 1 章 序論

近年、コンピュータによる情報処理技術の進展はめざましい。コンピュータの処理能力は年々飛躍的に向上し、日々、新たな技術や新たなソフトウェアが生み出され、人間にはできなかったことだけではなく、今まで人間にしかできなかったことまでコンピュータで可能になっている。さらにインターネットの普及によりコンピュータを使って世界中どこへでも瞬時に情報を発信し、受信することが可能となった。このように日々進展し続けているコンピュータは、現在、様々な分野において使われている。その用途は、ワープロ、表計算、スケジュール管理、事務処理などのいわゆる電子文具としての用途から、科学技術計算などの専門分野の用途、また、個人の娯楽、コミュニケーションツールとしての用途まで多岐に渡り、幅広い。

その中で、コンピュータを教育の分野に応用したものとして CAI がある。CAI とはコンピュータによる学習支援 (Computer Assisted Instruction もしくは Computer Aided Instruction) のことで、コンピュータを用いて効果的な学習を実現しようとするものである。

一方でコンピュータによって情報化が進んだ社会においては個人のコンピュータ利用能力、いわゆるコンピュータリテラシーの格差が問題となっている。即ち、このような社会においてコンピュータを扱う能力が低い人々は不利益を被り、その能力の高い人々はコンピュータによる多大な恩恵を受けるといった大きな格差が生まれる問題である。

このような問題は CAI にも当てはまる。従来の CAI は、効率よく学習を進めるためには、キーボードやマウスの操作などのコンピュータ操作をある程度できることが必要とされるものが殆どであった。このため、コンピュータ操作のできない人々にとっては CAI が効果的な学習法になり得ないこともあった。

ここへきて、コンピュータの処理能力の向上とソフトウェア技術の進展によって、人間の音声を自動認識する音声認識技術が実現され、音声文字入力などの用途に用いられるようになった。音声認識技術を用いれば、キーボードやマウスの操作ができなくてもコンピュータを操作できるインタフェースを実現できると期待されている。最近ではパーソナルコンピュータにおいても計算処理能力の向上がめざましく、パーソナルコンピュータでこのような音声認識を実現するソフトウェアも実用化されている。さ

らに性能が向上すればほとんどのパーソナルコンピュータにこのような音声認識の機能が装備されるようになるであろうと予想される。

本研究ではこのような音声インタフェースの可能性を探るために「エネルギー問題」を教育内容として、音声によって誰もが効果的に学べる CAI システムの実現を試みる。

エネルギー問題は化石燃料の枯渇、地球温暖化などの様々なエネルギーに関する問題である。これらは現在大きな問題となっており、全ての人に知ってもらい、正しく理解されることが求められる問題で、また、それによって各個人が問題解決のための努力をすることが求められている。

本論文の構成は以下の通りである。

第 2 章ではまず、本研究の背景である CAI について述べ、その問題点について考察する。そして音声対話インタフェースの要素技術である音声対話技術を概説し、本研究の目的を述べる。

第 3 章では、本研究で構築する音声対話インタフェースを用いた CAI システムに必要な機能について検討し、機能設計について説明する。

第 4 章では、実際に構築した CAI システムを実現するための構成、処理の手法について説明し、実際のシステム構築に用いたハードウェア、ソフトウェアについて述べる。

第 5 章では構築したシステムについて動作確認した結果と、評価実験の結果からシステムの有用性について述べる。

第 6 章では結論と今後の課題を述べ、本研究を総括する。

第 2 章 研究の背景と目的

2.1 研究の背景と目標

CAI とは、Computer Assisted Instruction もしくは Computer Aided Instruction の略で、広義にはコンピュータを活用する学習や教育を意味する。CAI による学習では、学習者はコンピュータを用いて時間的空間的な制約なしに学習を行えるといった利点がある。つまり、CAI を用いることでより多くの人々が時間、場所を気にせず教育を受けることが可能になる。

従来から CAI については様々な研究がなされてきたが、これまでの CAI においてあらゆる人を対象とした場合、大きな問題となることがあった。それがキーボードやマウスなどのコンピュータのインタフェースの問題である。従来のコンピュータのインタフェースは入力にキーボードとマウスを用いてきたが、コンピュータの初心者や高齢者にとってこのキーボードやマウスは慣れるまで非常に扱いづらく、大きなストレスとなるものであった。このように CAI において、コンピュータを扱うこと自体が障壁になることは学習者を教育の機会から遠ざけることになり、必ず避けなくてはならない。この問題を解決するために、本研究では人間にとって最も自然なコミュニケーション手段のひとつである音声による対話に着目し、音声対話インタフェースを応用した CAI システムを構築することを目指す。

表 2.1: 本研究で対象とする CAI によるエネルギー教育の特徴

意義	一般の人々がエネルギー問題について知り、関心を高めることによってエネルギー問題解決のために努力することを促す。
対象	一般の人々を対象とする。
内容	エネルギー問題の現状と各問題の原因、技術開発、対策など多岐にわたる。主に浅く広い知識。整理されていないと把握しにくい。恒常的な教育内容の更新が必要。
手段	コンピュータによる個別学習。
場所	なるべく多くの場所で学習することを目指す。
時間	いつでも学習したいときに学習できる。

昨年度、本研究室では音声インタフェースを用いた CAI システムとして、擬人化エージェントに着目した音声会話エージェント型 CAI を試作した^[1]。この CAI システムの評価実験を行った結果、このシステムは会話機能や、擬人化エージェントの実現などの点で多くの問題があった^[2]。

本研究では、この基礎研究の成果を踏まえて各機能の実現方法について検討と改良を行い、「エネルギー教育」を対象に、音声対話インタフェースを用いた CAI システムを構築することを目標とする。

エネルギー教育のために CAI を用いる意義は、CAI の利点を活かして一般の全ての人にエネルギー問題について正しい知識を獲得させ、それとともにエネルギー問題について関心を持たせることにある。本研究で具体的な対象とする、CAI によるエネルギー教育の特徴を表 2.1 に挙げる。この CAI ではあらゆる人々を教育の対象としている。特に、エネルギー問題の社会、環境との関わりについて教育、啓蒙することを考えると、知識を広く浅くすることと、新しい技術開発のニュースや社会の動きの知識などデータの恒常的な更新が必要である。

以下では、本研究で取り扱う CAI と研究の背景となるいくつかの CAI 技術について述べた後、それぞれの長所、短所について検討し、また、従来の CAI の多くが持つ問題点についても考察する。次に、本研究で実現しようとする音声対話インタフェースの要素となる音声対話技術について述べる。さらに、昨年度の本研究室での研究の成果とその問題点について述べた後、本研究で目標とする CAI システムについて述べる。

2.2 CAI システム

「CAI は計算機に特有の機能を教育に生かし、学習者が計算機と対話しながら学習する方式で、教授と学習の個別化と、教授と学習の最適化を実現し、学習者の個性に応じて効果的に学力を形成していく教授学習システム」であると芦葉は定義している^[3]。CAI は、1950 年代後半、コンピュータ上に優れた教師の教授能力を実現しようとする試みとして人工知能の分野で本格的な研究開発が開始された。

CAI は当初、学習者の個人差に応じて教材を適切な順番で提示して学習を進める形のものであった。その後、学習の最適化、教材の自動生成等の研究が進み、個人適応型 CAI へと発展していく。また、より効果的に学習を進めるという観点から、静止画像や文章だけのよう単純な教材ばかりでなく、物理現象についての学習や訓練をより現実に近いものにするを目的にして、コンピュータによるシミュレーションを利用し

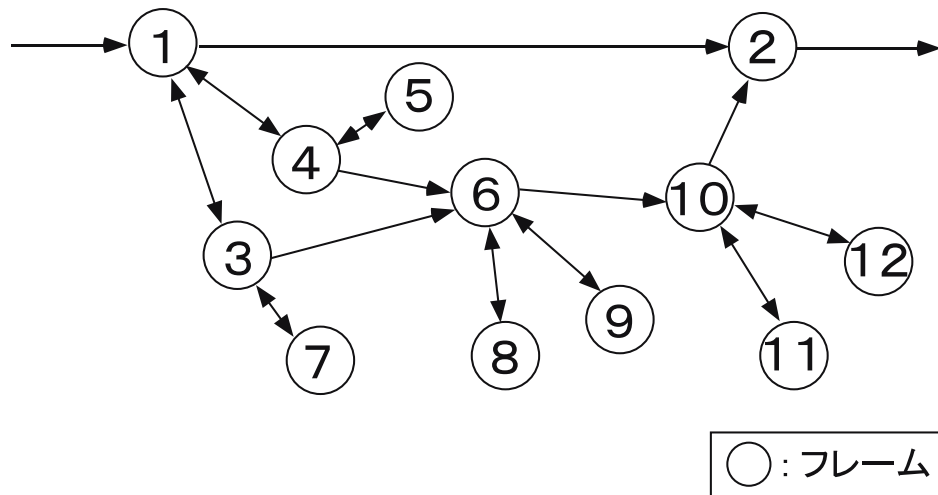
た CAI も生み出された。1970 年代になって言語処理や知識処理に関する人工知能の研究が進展すると、CAI の研究も言語や知識を扱うようになり、自然言語による人にとって自然なやりとりを目指した研究や、コンピュータ内に教育対象の体系的知識を持ち、それを用いて知的に柔軟な教授を行うことを目指した研究が盛んになった。これらは ITS(Intelligent Tutoring System) または、知的 CAI と呼ばれる。1990 年代になるとインターネット、WWW(World Wide Web)、HTML(HyperText Markup Language) の普及によって、ネットワークを利用した CAI の研究が盛んになってきた。これらの技術を用いた CAI は教材の作成が容易で、インターネットに接続できるほとんどのコンピュータで学習が行えるため、実践的な研究が数多く行われている。また、優れた学習環境をコンピュータで実現することを目的とした研究も盛んになった。ネットワークを介した協調作業 (Computer Supported Cooperative Work:CSCW) を学習するような CAI や、ネットワークを介して協調的に学習を進める分散型協調学習 (Computer Supported Collaborative Learning:CSCL) などについて活発に研究が行われている。

以上では CAI の歴史的発展を述べたが、以下、本節では菅井の考え方に従って、教授、学習理論のパラダイム変化により CAI を分類し^[4]、それぞれの特徴について述べる。

2.2.1 各 CAI 技術の分類と特徴

フレーム型 CAI・コースウェア

フレーム型 CAI(Adhoc Frame-Oriented CAI) は伝統的 CAI とも呼ばれ、フレームと呼ばれる学習の単位を順に提示して学習を行う CAI システムである。教材の製作者は、あらかじめ様々な学習者を想定して、いくつものフレームと呼ばれる小さな学習の単位に分けて教材を作成する。これをコンピュータによって提示し、学習者の「反応」(キーを押すなどの操作をいう) によって次に進む。これにより学習者は個人差に応じた自由な学習速度で学習できる。また、図 2.1 に示すようにフレーム間の関係を木構造やネットワーク構造で記述しておき、学習者の反応によって分岐を行うように学習コースを決めておけば学習者の反応に応じた対応をすることができる。これらのものはコースウェアとも呼ばれる。



- ・上級者の場合: 1→2→
- ・初心者の場合: 1→3→6→8→6→9→6→10→

図 2.1: フレーム型 CAI の学習コースの分岐

知的 CAI(Intelligent Tutoring System : ITS)

知的 CAI は ITS と呼ばれ、人工知能の研究と密接にかかわって発展してきた CAI である。清水らは、ITS とは「認知科学に依拠して、分野知識と問題解決方略、学習者モデルと教授方略をコンピュータに持たせて、学習者の理解状況に合わせた教授を行う試み」としている [5]。そのために知識表現や学習者モデル、言語処理などの人工知能の技術を応用し、学習者に対して柔軟な対応が可能になっていった。

ITS の代表的な研究に Carbonell による SCHOLAR [6], [7], [8] がある。これは、南米の地理学習用 CAI で、自然言語を用いた学習者との対話によって学習を進めることができる。SCHOLAR では、意味ネットワークを用いて南米の地理に関する知識を表現したものを持ち、これを用いて学習者に自然言語で問題を出題したり、学習者からの質問に答えることができる。

また、機器故障診断などの訓練においてシミュレーションによる訓練をコンピュータ上で行いながら、あらかじめそのタスクに関する知識と推論機構を持ったコンピュータが助言を行ったり、誤りを指摘したりすることが可能な SOPHIE^{[9], [10]} や STEAMER^{[11], [12]} などのシステムも研究されている。

このような ITS は、フレーム型 CAI と比較して、複数の戦略を用いた柔軟な教授を実現するために複雑な機構を備えており、自然言語による対話や教材を抽象化した知識構造として持つことによる柔軟な教材の提示、学習者モデルによる個々の学習者に

適した対応など、より知的な教授を可能にした。しかし、このような ITS では教材とシステムは完全には不可分で、新たな教材を開発する際には教材を知識構造に変換する作業、知識構造化した教材を取り扱う方略を検討し実装する作業が必要となる。また、学習者モデルとその取り扱いも修正しなければならず、結局システム的大幅な変更が必要となることが多く、あまり汎用には適さない。また、教材、システムともに開発には非常に大きな労力が必要となる。実際に教育システムとして実用できる程度まで知識構造化された教材を構築することは難しく、またコストもかかるため、応用の分野が限られてしまう問題もある。

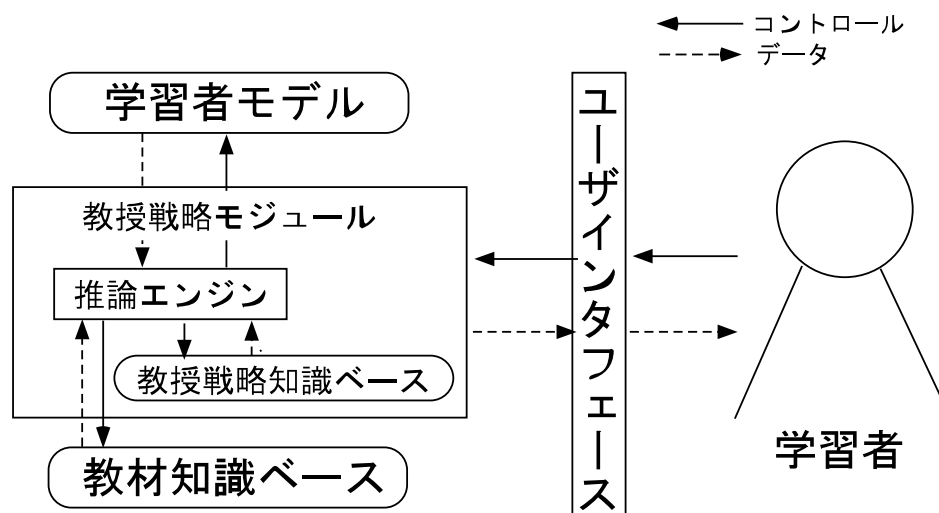


図 2.2: ITS の構成例

操作型学習環境 (Interactive Learning Environment : ILE)

ここまで挙げた 2 つの分類は「コンピュータ上に教師を作る」ことを目的とする教育システムであったが、学習者が主体的に新しい知識を獲得すること、また、既得の知識を適用・検証して理解を深めるための自立的発見学習を支援する学習環境をつくり出すことを目的とした教育システムが ILE(Interactive Learning Environment) である^[13]。ILE には、対象世界を限定して、その中でシミュレーションなどを可能にすることで学習者の設定した仮説を検証させながら発見的学習を進めていくマイクロワールド型と呼ばれるものと、対象世界を広くとって、問題解決に必要な情報と、数式、図、グラフなどの作成、操作を行える汎用的ツールを提供して自由に問題解決をさせる開放型とがある。

WWW 型 CAI システム

WWW 型 CAI は、インターネット上での World Wide Web(WWW) 技術を利用した CAI の総称である。主に HTML を用いて教材を記述し、インターネットを介して配信する。学習者はインターネット上で WWW サーバへアクセスしてサーバ上にある教材を WWW ブラウザで表示して学習を進める。この種類の CAI システムは基本的にフレーム型 CAI の延長線上にあり、HTML で記述された WWW ページのひとつひとつがフレームに相当する。図 2.3 に示すように各ページには HTML の特徴であるハイパーリンク機能を活かして他のページへのリンクがいくつか付与されている。学習者はそれをマウスでクリックするなどしてリンク先のページを表示させる。こうして学習者のたどったリンク先のページが次のフレームになる。WWW 型 CAI ではこの機能によって、学習者に合わせた進行と分岐による学習者の反応への対応を実現している。

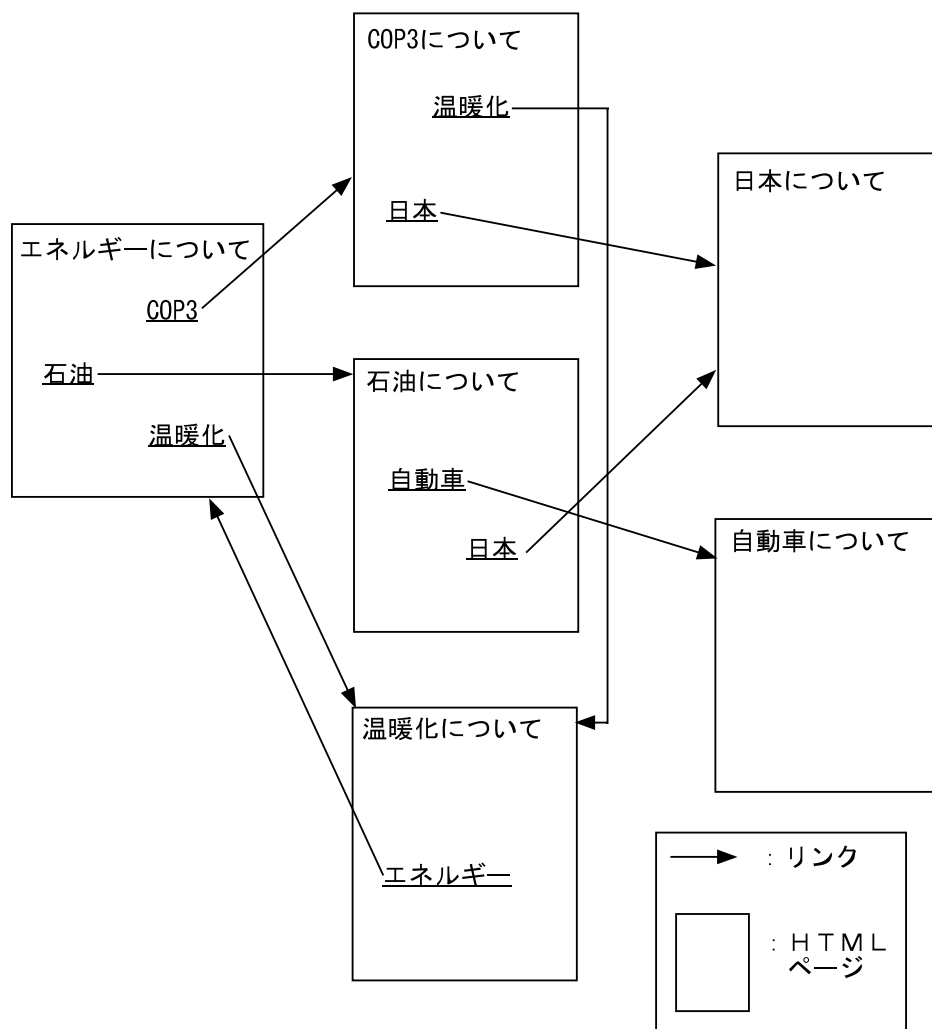


図 2.3: WWW 型 CAI システムの一例

また、WWW は表 2.2 に示すような特徴を持ち、それを CAI へ生かして効果的な学習システムが構築されている。

表 2.2: WWW の特徴

時間的、空間的制約のない非同期分散環境を実現できる。
ハードウェア・OS に依存せず、WWW ブラウザを備えるほとんどのコンピュータで実行可能である。
サーバ上の HTML ファイルを更新するだけで内容の更新が行え、非常に容易にメンテナンスできる。
現在様々なホームページ作成ツールがあり、それらを用いることにより、非常に容易に作成できる。
ハイパーリンク機能を用いて自由に他のページやコンテンツにアクセスできる。
画像、動画、音声など様々な種類のマルチメディアコンテンツを取り扱える。

WWW 型 CAI では以上のような特徴によって、非常に柔軟に効果的な教材を製作でき、また教材作成の負担も少ない。さらに教材の配信、メンテナンスも容易となっている。WWW 型 CAI は、CAI の手法としては古い伝統的 CAI に基づいているものの、現在でも教材の作成方略や、多人数での教育効果実験など数多くの実践的な研究が行われている [14],[15]。ただし、WWW を用いるため、ブラウザを使って WWW ページを閲覧するということが中心となり、受動的な学習になる傾向がある。そこで現在では学習者モデルを持ち個人に適した教材を自動的に提示したりする [16] など、知的 CAI の研究の成果を採り入れる研究が行われている。本研究室においても、インタラクションを促進するような仕組みを採り入れた WWW 型 CAI の研究が行われている [17]。

2.2.2 各 CAI 技術の比較

本研究では、広く一般の人々を対象としたエネルギー問題について知識獲得のための CAI システムを構築することを目的としている。そのためにどのような CAI の形式、機能がふさわしいかについて考察するためにここまで述べてきたフレーム型 CAI、ITS、ILE、WWW 型 CAI の 4 種類について学習形態、実行環境、インタラクショ、教材作成の 4 つの観点から比較した結果を表 2.3 に示し、以下でそれぞれについて述べる。

表 2.3: 各 CAI システムの比較

	学習形態	実行環境	インタラクション	教材作成
フレーム型 CAI	シナリオをたどる 受動的学習	動作対象となる コンピュータ	フレームを次に 進めるときのみ	コースとフレームの作成 だけで比較的容易
ITS	教師の柔軟な指導 による学習	専用コンピュータ	学習者に合わせた 柔軟で適切なやりとり	教材の抽象化による 知識構造への変換が難しい
ILE	能動的な発見学習	専用コンピュータ	学習環境との活発な インタラクション	シミュレーション環境の 構築が必要で、多大な 労力を必要とする
WWW型 CAI	WWWページの閲覧 による受動的学習	WWWブラウザの動作する すべてのコンピュータ	閲覧による受動的学習	HTMLファイルの作成、 豊富な作成支援ツール により、作成が容易

学習の形態

学習の形態とは、CAIにおいて、ユーザがどのようなスタイルで学習を行うかという観点である。

フレーム型 CAI で実現する学習は、あらかじめツリー構造に作成されたシナリオをたどる学習で、フレームの提示を受けることによる受動的学習である。

ITS(知的 CAI) は「コンピュータ上に教師を作る」ことを目指しており、学習活動における教師の知的な役割の一部をモデル化して持つことで教師の柔軟な指導機能を実現しようとするものである。従って、教師が柔軟に指導を行う形態での学習を実現できる。

ILE(操作型学習環境) は教師による指導よりも実際に試行錯誤をする中で経験から知識を獲得するような能動的な発見的学習を実現している。

WWW 型 CAI は教材の提示と練習問題などによって知識を獲得し定着させるような場合によく応用されており、WWW ページの閲覧による受動的な学習を実現する。それ以外にも、世界中に広がったインターネットにある豊富な WWW のコンテンツを用いて、その中からある課題に関して調べるといった発見的学習も可能である。

実行環境

ここでいう実行環境とは、CAIにおいて用いる計算機の制約や、広い場所、指導者を必要とするなどの学習を実行する上での環境の制約の観点である。

フレーム型 CAI はその動作対象となるための条件を満たすコンピュータ上で動作する。新たなコンピュータ上で動作させる場合は、それに対応できるようにシステムの修正が必要である。

ITS は開発の対象となる計算機環境のみで動作するものが多い。また、高度に知的な機能を実現しているものは知的処理を行うためのエキスパートシステムや処理能力の高い専用マシンなど特別なサブシステムを用いている場合も多く、その学習環境 (ITS) は限られたところでしか動作しない。

ILE についてもシミュレーションなどを行うマイクロワールド型のものなどでは高速な画像処理能力や計算能力が要求されるものがある。それらは特別な計算機環境を用いており、ITS と同様の傾向がみられる。開放型のものでは、ツール群が揃えばどのような環境でも学習できるが、効果的な学習のためには人間の指導者のいる環境が望ましい。

これに対し WWW 型 CAI は、WWW ブラウザを持つほとんどのコンピュータで実行可能であり、また、インターネットに接続していればどこでも学習可能である。このため、最も計算機環境による制約のない CAI システムであると言える。

インタラクション

ここでいうインタラクションとは、CAI を用いた学習で学習者とシステムとの活発なインタラクションがあるか、また、それによって学習意欲が向上するかどうかの観点である。

フレーム型 CAI では、学習者はフレームを進めるときにその要求を入力をするだけである。また、システムからの反応もフレームの提示だけとなるのでインタラクションは少ない。

ITS では学習者のモデルを持ち、個人の学習状況に適應した教授を実現するための仕組みを持っており、そのときの学習者の状況に応じて適切な反応を返すことで学習者とインタラクションできる。また教材を抽象化した知識構造として持つことで、固定されたテキストデータをそのまま提示するのではなく、表現を変えて提示できるなど、柔軟に取り扱える点でも他よりも優れている。ITS の中には学習者の反応をも柔軟に取り扱え、対話を行うことができるものもあり、活発なインタラクションを行える。

ILE においても、課題を通して学習者と学習環境との活発なインタラクションが行われる。それは、問題を通じた学習環境とのインタラクションであり、人間対人間のようインタラクションとは異なる。

WWW 型 CAI は表示される WWW ページを閲覧するためのものであるという WWW の性質上、教材を閲覧するだけの受動的な学習になりがちである。知的 CAI の成果を採り入れた個人に適應したシステムも研究されているものの、活発なインタラクシ

ンはあまり期待できず、学習者が退屈し、集中力の持続が難しい。

教材作成

教材作成の観点は、その CAI 教材を作成する作業の容易さについてである。

フレーム型 CAI ではフレーム単位で提示する教材を作成するだけでよいが、学習コースをデザインする必要がある、これを細かく設定するとその作業量は多くなる。それでも必要な作業は比較的少ない。

ITS では、1つの教材を制作するのにその学習対象を抽象化して知識構造に変換することが必要で、さらにその知識処理の方法も設計しなければならない。また、同様に学習者モデルも必要となる。教材とシステムが密接に結び付いている ITS では、機能が知的になればなるほど教材作成時の負担が増える。また、教材やシステムを変更することは、人工知能や ITS の専門家でなければ困難な専門性の高い作業で、同時に教材の知識もなければならぬことを考えると教材の製作コストは非常に高い。

ILE については、意図する学習方法によって教材をどのように作成するかが変化する。教育対象をシミュレーションとして構築する必要のあるマイクロワールド型においては教材として新たなシミュレーション環境を実現しなくてはならず、教材製作に必要な労力は非常に大きい。

WWW 型 CAI では、基本的に教材となるのは WWW ページとなる HTML ファイルのみであり、これの作成が教材作成作業となる。教材全体の構造や分岐はハイパーリンクを設定することで可能になる。さらにホームページ作成ツールなど様々な作成補助ツールがあり、これらを用いて様々なマルチメディアコンテンツを持った教材を非常に容易に作成することができる。

2.2.3 従来の CAI の問題点

以上では、4つの CAI の種類別に4つの観点から比較し、利点と問題点を挙げたが、本項では従来の全ての CAI に共通する2つの問題点を以下に挙げ、考察する。

1. 従来の CAI は一人だけの学習を対象にしたものがほとんどで、複数人を対象とする学習に対応できるシステムがほとんどない。そのためグループ学習の利点を活かさない。
2. 現在ほとんどの CAI で用いられているキーボード、マウスなどの入力インタフェースの操作が煩わしい。

まず、従来の CAI が複数人での学習に対応していない問題点について述べる。複数人による学習 (グループ学習) は、他の学習者とのインタラクションを通して、また他の学習者の学習行動を参考にすることなどによって学習効果を高めると考えられる。しかし、これまで CAI 研究はより効果的な個別学習を実現するためのシステムを目指すことを出発点としていた。これらのシステムは一人だけの学習を想定し、それをどのように効果的なものにするかに重点をおいていたため、教室での授業にみられるような複数人による学習の効果についてはあまり考慮されず、また実現されていなかった。近年になってようやくバーチャルクラスルーム^[18] や Sharlok^[19] などに代表されるインターネットを用いた遠隔コミュニケーションによる学習支援環境の研究が行われるようになったが、これらはあくまでも協調学習支援環境であり、教授、指導を主な目的とするものではない。本研究では、この複数人による学習を行えないという問題点を解決するために、ネットワークで結ばれた複数の学習者が同時に一つの CAI 学習環境で学習を行える「同期分散型 CAI システム」を構築することを提案する。

次にインタフェースの問題について考える。現在、ほとんどの CAI システムにおいて、入力インタフェースとしてキーボードとマウスを用いている。マイクロワールド型 ILE や体験型訓練用 CAI の中には人工現実感 (Virtual Reality:VR) を利用し、移動、操作にジョイスティックなどの特殊なデバイスを用いるものもあるが、それ以外のもものではキーボード、マウスを用いているのが現状である。キーボードは操作に熟達すれば非常に早く文章を入力できる優れたインタフェースではあるが、コンピュータ初心者にとっては、扱いづらく、また入力速度も非常に遅くなる。また、マウスは、GUI(Graphical User Interface) において非常に画期的で有用な入力デバイスであるが、初めて扱う場合、その独特の操作感ゆえにコンピュータのディスプレイ上で自分の意図する場所にマウスカーソルを移動させることが困難な人も多い。これらの入力デバイスは、ある程度コンピュータを操作したことがある学習者にとっては何ら障壁とはならないかも知れないが、コンピュータ初心者や高齢者、身体障害者などの学習者にとってまず最初の大きな障壁となる。学習対象を学ぶための余計な負担は極力少ない方がよい。

このような障壁をなくすために、初めてでも誰もが躊躇なく操作できるような新たなインタフェースが必要となる。そこで、本研究では、人間にとって自然なコミュニケーション手段をインタフェースとして用いればこれを実現できると考えた。人間の持つ自然なコミュニケーション手段の一つに音声による自然言語での対話がある。人間の知的活動の最も基本となるものが「見る、聞く、話す」であるが、このうちの「聞

く、話す」をコンピュータのユーザインタフェースとして用いることで、誰もが自然に学習を行える CAI システムを構築できると考えた。残念ながら音声対話は難聴者には向かないが、この点を除けばよりバリアフリーなインタフェースである。

2.2.4 CAI システムに音声対話インタフェースを用いることの利点

CAI システムにおいて、音声対話を用いる利点としては以下の 2 点がある。

1. 音声による対話によって学習を行うことでキーボードやマウスの操作に煩わされない。
2. 音声による対話を行うことで学習において活発なインタラクションが期待でき、興味、集中力が高まると考えられる。

以下ではこの音声対話を実現するための音声対話技術について述べる。

2.3 音声対話システム

音声による自然言語での人間との対話をコンピュータに行わせようという試みは音声言語処理分野、自然言語処理分野、人工知能分野の研究の進展と、各分野間の有機的な結び付きによって、近年一層盛んになってきた。最近の音声対話実用化に関する研究では、NTT コミュニケーション科学基礎研究所による Dug-1 がある^[20]。これは、音声による対話を用いてスケジュール管理やビデオの予約を行うシステムである。Dug-1 は逐次理解、逐次生成方式を用いることで対話特有の割り込みや短い省略文を理解可能にし、リアルタイムで自然な対話を実現している。

音声対話システムとは、ユーザとの音声による対話によって何らかの課題を遂行するシステムである。この課題はタスクと呼ばれ、情報検索、教育、機器の操作など様々なものがある。音声対話システムの構築には、音声言語処理分野の音声認識技術、音声合成技術、自然言語処理分野の形態素、構文解析技術、意味解析技術、文生成技術、人工知能分野の知識処理技術、プランニング、推論など各分野の各技術を統合して初めて実現できる。しかし、各要素技術を用いて各々独立した処理系を構築し、それを直列に接続して処理するだけでは効果的な処理はできない。全体を 1 つの複雑系として取り扱い、大きな枠組みでとらえることが必要である。

堂下らは音声対話を 1 つの大きな課題としてその実現のための枠組や要素技術についての詳細な研究を行っている^[21]。これは対話のメカニズムを解明して対話過程のモ

デルを構築し、それを計算機上にシステムとして実現することを目指したもので、その中で一般的な音声対話システムの構成に関する考察がなされている [22]。それによると、図 2.4に示すように一般的な音声対話システムは音声認識部、音声合成部、文解析部、対話管理部、タスク処理部からなる。実際の音声対話での処理の流れとしては、音声認識部で音声を文字列に変換し、文解析部で文の意味、発話の意図などを解析する。次いで対話管理部で解析した発話に対する反応が決定される。そしてタスクに関係する処理がタスク処理部で行われ、対話管理部で生成された反応としての発話が音声合成部で合成音声として出力される。以下では、音声対話システムを実現するために必要な音声認識、音声合成、文解析、対話管理、タスク処理の 5 つの要素技術について述べる。

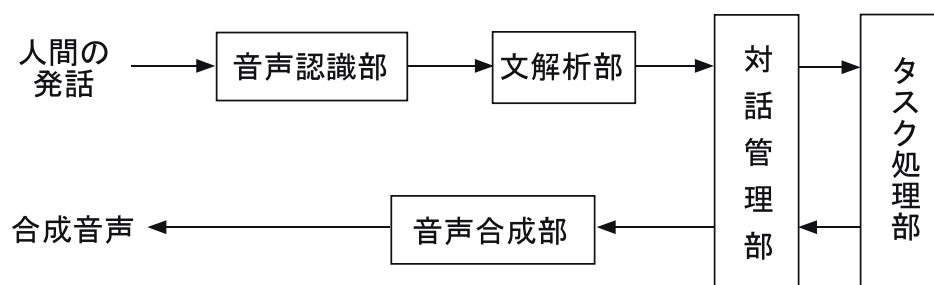


図 2.4: 一般的な音声対話システムの構成

2.3.1 音声認識

音声認識はコンピュータに接続されたマイクロフォンから入力された音声を解析して、音情報から発話した内容を文字列として取り出す処理である。現在この分野では、大量の音声言語データから音声言語モデルをつくり、それを利用して隠れマルコフモデル (Hidden Markov Model : HMM) を評価関数としたパターンマッチングによる認識を行う方式が主流となっている [23]。現在、認識率も向上して音声文字入力などにも応用されるようになったが、音声対話への応用については認識における音声対話特有の問題 (短い句だけの発話を多く含むと認識率が低くなる等) や、認識処理に時間がかかり、リアルタイム性を実現しにくいなどいくつかの問題も残っている。

2.3.2 文解析

文の解析は形態素解析、構文解析、意味解析、発話意図解析などの処理を行うことで実現される。システムによっては全てを行う場合もあるが、この中のいくつかの処

理を行ったり、いくつかの処理をまとめた1つの処理を行ったりするものもある。

形態素解析、構文解析

形態素解析は音声認識の結果として得られた文字列を単語ごとに区切り、各単語について品詞を解析する処理である。構文解析はその結果を用いて、主語、述語、修飾語等の単語間の関係を解析することで係り受けなどの文の構造を得るための処理である。しかし1つの文で何通りかの構文に解釈できる文もあり、文脈、意味と密接にかかわっている。

意味解析

構文解析した結果などから文の意味を解析する処理である。文中で使われる単語には多義性があり、構文解析して得られる単語間の関係だけでは文の意味は決まらない。そのため、意味解析によって文脈情報や常識的知識などを用いて各単語の意味を1つに決め、文の意味を抽出する。

2.3.3 対話管理

対話管理は文解析で得られた結果に対する反応を生成する処理であり、対話を実現する核となるものである。対話管理は既に述べた文解析、次項で述べるタスク処理と密接にかかわっており、その内容は発話意図の解析、プランニング、文生成など多岐にわたり、システムを性格づける。これらの処理においては、発話された文を文解析して得られた結果をより抽象化した知識構造にして扱う。

発話意図の解析

意味解析を行った結果、解釈された文の意味に基づいてそれがどのような意図を伝えるために発話されたかについて解析するのが発話意図解析である。同じ文を発話していても状況によってその意図するところが異なることは多く、この処理は現在の対話の状況、文脈、タスクに大きく依存する。

プランニング

プランニング処理はタスクの進行状況を認識し、その状況における発話の解析結果を元に対話システムの次の発話、行動をプランニングするものである。この処理によっ

てシステムの発話意図が決定される。タスク処理と密接に関係する。

文生成

文生成は認識した発話に対する反応として、コンピュータが発話するための文を生成する処理である。この処理によって抽象化されていたコンピュータの発話意図が実際の自然言語の文として生成される。

2.3.4 タスク処理

対話システムは対話を行いながら設定されたタスクを遂行するシステムであり、その実際のタスクを行うのがタスク処理である。対話管理部で生成された反応の1つとしてタスクに関する処理があり、それを行うことで目的を達成する。例えば、音声対話型の情報検索システムやCAIなどユーザと対話を行うことで知識を得るようなタスクを遂行するシステムでは、タスクの知識を格納するデータベースとその処理部からなり、データを引き出して対話管理部にわたすことでタスクを遂行する。このデータが対話管理部の文生成処理で自然言語の文に変換され、音声合成部で合成音声として発話される。また、音声対話型機械制御システムなどのようなユーザとの音声対話によって何らかの機械の制御を行ったりするものでは機械の制御信号を送ることなどがタスク処理として行われる。

2.3.5 音声合成

生成された文はコンピュータ上の文字列であるが、それを音声合成処理で、音声に変換して出力する。これが発話になる。そのとき、より自然な発音やイントネーションで流暢な発話を実現するために文を解析してイントネーションや発音を最適に制御する。

2.4 昨年度構築した会話エージェント型CAIシステムとその問題点

本研究室にはCAIの研究を行っている研究グループがいくつかある。筆者らの対話型CAI研究グループは昨年度、協調学習者としての教師をコンピュータ上に実現し学習効果を高めることを目的に、音声インタフェースを採り入れた会話エージェント型

CAI システムを試作した^[1]。そして、音声を用いたインタフェースの有効性と音声会話エージェントを用いた CAI の有効性を確認した^[2]。このシステムは学習者と音声による会話でコミュニケーション可能な教師エージェントをコンピュータ上に実現することと会話によって学習を進めることを目指していた。

この CAI システムは以下のような機能を持っている。

1. 学習者に対して音声で講義を行う。

講義用教材は段階、章、節、単元の階層構造を持ったものとして制作され、合成音声による説明とグラフ、写真などの資料の提示を行うことで講義を行う。講義は一定の順序で進行する。

2. 学習者に対して問題を出す。

1 つの単元が終るとそれに関する練習問題を合成音声で出題し、学習者がそれに対して解答すると正解、不正解に応じた反応の文を読み上げる。

3. 学習者からの質問に答える。

学習者はいつでも音声で質問を行うことができる。教師はそれに対して回答することができ、学習中の疑問点を解消できる。

この CAI システムは原子力発電の必要性についての教育を行うもので、一人で学習するためのものである。システムのハードウェア構成は、ネットワークで結ばれたワークステーションとパーソナルコンピュータからなる。講義と質問回答の制御にワークステーション上で Gensym 社製エキスパートシステム構築ツール G2 を動作させており、パーソナルコンピュータ上の音声認識エンジンと音声合成エンジンを用いて音声認識、音声合成を行う。このシステムではデータベースとして講義、問題、質問回答、反応の 4 つのデータベースを持ち、講義など実際のデータをパーソナルコンピュータ上に、制御データをワークステーション上に持っており、それを用いて学習を進行する。

学習者はパーソナルコンピュータの前に座り、システムによって実現された会話型教師エージェントと音声による会話を行いながら学習を進める。また、パーソナルコンピュータのディスプレイ上には講義に関連するグラフや写真が表示される。また、「すみません」などの呼びかけに対して講義を止める機能を持っていた。

この研究では、教師エージェントがあたかも人間の教師のように学習者と会話し、学習を進めることを目指していた。しかし、実際に構築したシステムを評価するために被験者実験を行った結果から、以下のような問題点が明らかになった。

- 被験者による主観評価では、「会話の成立」「協調的な情報の交換」の観点での評価が非常に低かったことなどから、円滑な会話は実現できていない。
- 質問回答については被験者による自由な質問に対して全く答えられなかったことから質問回答の機能を果たしていない。
- 練習問題で学習者の回答認識率が 46.4%と音声認識率の 72.5%に比べても非常に低く、半分以上を誤って判断しているため、練習問題の機能としての正解判定ができていない。

以上のような問題の他に、本研究において対象とするエネルギー問題について一般の人々が知識を得るための CAI として考えると以下の点で問題があった。

- 高価なワークステーションとエキスパートシステム構築ツール G2 を用いており、コストが非常に高くなってシステム導入の際の大きな障壁となる。
- ワークステーションとパーソナルコンピュータの両方に異なった形で教材のデータベースがあるため、教材の作成、配信が困難である。

本研究の目的を達成することを、昨年度のシステムにおいて考えたとき、これらのような多くの問題が障害となる。システム構成についても問題があることから、新たなシステムを構築する必要性がある。

本研究では昨年度の研究成果と問題点への対応を踏まえて、エネルギー教育用の新たな CAI システムへの発展を行う。

2.5 本研究の目的と提案する CAI システム

本研究では、エネルギー問題を広く一般市民に啓発し、エネルギーと社会、環境の関係に関する知識を身につけてもらうための有効なインターネット社会のコミュニケーション手段として以下のような特徴を持つ CAI システムを提案し、開発することを目的とする。

- 入出力に音声対話インタフェースを用いる。
- 複数の学習者が同時に学習を行える。

本研究で構築する CAI システムにこれらの 2 点の特徴を持たせることによって 2.2.3 項で挙げた従来の CAI システムの問題点を解決することを目指す。

また、音声対話については昨年度の問題点を踏まえて、新たな手法を用いることでより優れた音声対話機能を目指す。

本章で述べた背景をもとに第 3 章ではシステムの持つべき特徴について詳細に検討し、その具体的な機能、それらの実現方法について述べる。

第 3 章 構築する CAI システムの機能

本章では、本研究の目的である音声対話インタフェースを用いた CAI システムの構築にあたり、まず、第 2 章で述べた背景に基づいてシステムに付与すべき機能の特徴について検討する。次いで、実現する機能を詳細に検討する。

3.1 システム機能の特徴の検討

2.5 節で研究の目的として述べたように、従来の CAI システムの問題点を解決するために、本研究で構築しようとする CAI システムは、「複数人で学習可能」、「音声対話」の 2 つの大きな特徴を持つ。本節では、エネルギー問題を広く一般の人々が学ぶための CAI システムが備えるべきその他の特徴について 2.2.2 項で述べた 4 つの観点から考察する。

本研究の CAI システムで対象とするエネルギー教育は表 2.1 に挙げたような特徴を持つ。このような対象を広く設定し、浅く広い知識の伝達を図るような教育では誰もがいつでもどこでも使えること、広い範囲の教育内容を効果的に学習できること、関心や集中力を高める仕組みを持つことなどが重要になる。また、エネルギー問題の特徴としてデータや情報の内容の変化が比較的速いことが挙げられ、教材を持続的に頻繁に更新することが必要となる。以下でエネルギー教育のための CAI システムに必要な特徴について検討する。

まず、誰でも使えるものにするために、音声対話インタフェースを持つ CAI システムにする。

また、どこでも使えるシステムにするためには、実行環境の制限の少ないものが良い。WWW 型 CAI はこの点で非常に優れているが、ページの閲覧という WWW の枠組の制限にとらわれ音声による対話の実現が困難になる。この点に注意して、現在最も普及しているパーソナルコンピュータでシステムを構成し、入手しやすい安価なシステム構成を実現する。

さらに、広い範囲の教育内容の学習には体系的に整理された教材を用いることが効果的であるが、そこで扱い切れなかった内容についても網羅する必要がある。従って、体系的な教材と網羅的な教材の 2 つの学習を組み合わせることを目指す。これを音声

で講義をし、学習者の質問に答えるという学習形態をとることで実現する。また、体系化された講義用教材に関しては、フレーム型の CAI で用いられるような教材を体系化し、整理してフレーム単位で作成されたものがこの目的に合致するため、このような教材を用いる。

頻繁な教材の更新が必要である場合、教材の作成、更新が容易なシステムが良い。この観点からも WWW 型 CAI は優れており、本研究で構築する CAI システムではこれと同じクライアントサーバ方式による教材の配信方法をとる。クライアントサーバ方式は、中心的処理を実現する 1 台のサーバコンピュータをコンピュータネットワーク上に置き、クライアントとなるコンピュータが適宜このサーバコンピュータに接続してその処理を行わせ、結果を受け取る。この場合の送受信するデータは教材データとなる。このようにしてサーバによって教材を一元管理すれば容易に教材を更新できる。また、教材の作成を容易にするために文章を中心に教材を構成し、その補助として HTML によるマルチメディアコンテンツを用いる。このようにして教材を扱いやすくし、作成、更新を容易にする。

学習において活発なインタラクションは興味や集中力の持続と密接に関係する。学習者にとって学習が常に新鮮な刺激であれば学習への興味、集中力は高いまま持続すると考えられる。逆に単調になって新鮮さがないと興味、集中力も低下する可能性が高い。学習においてどれだけ他者とインタラクションしているかは、どれだけ他者に働きかけて新たな刺激を受けているかということに直接つながる。従って、学習の中で活発なインタラクションが行われることによって学習者の興味、集中力を高めることができると考えられる。本研究では音声対話によるインタラクションでこれを実現することを目指す。

以上の考察を表 3.1 にまとめた。この考察から、本 CAI システムは以下のような構成上の特徴を持つことが望まれる。

- ネットワークを用いたクライアントサーバ方式によってシステムを実現する。
- サーバによって教材データを一元管理する。
- パーソナルコンピュータを用いて安価なシステム構成を実現する。

このシステム構成については第 4 章で詳しく説明する。また、この考察から以下のような機能上の特徴を持つことが望まれる。

- 音声による対話で学習を実現する。

- 講義形式で学習を行う。
- 学習者からの学習内容に関する質問を受け付け、それに適切な回答をする。
- 複数の学習者が同時に学習を行うことが可能である。

以下では、これらの機能上の特徴をもとに本 CAI システムの機能を詳細に検討する。

表 3.1: システムの特徴についての考察

教育の特徴からの要求	CAI 比較の観点	要求される条件	解決策
対象: 一般の人々	- - - -	誰でも扱える インタフェース	優れた音声対話 インタフェース
場所:なるべく 多くの場所	実行環境	入手しやすい安価な システム構成	パーソナルコンピュータ で構成されるシステム
教育内容: 浅く広く	実現する 学習	体系化された教材と 網羅的な教材	講義と質問回答による学習
教育内容: 頻繁な更新	教材作成	教材の更新が容易	サーバ - クライアント方式 サーバでの教材の一元管理
		教材の作成が容易	テキスト文字列に よって構成された教材
関心、集中力 を高める	インタ ラクション	学習への関心、集中 力を高める仕組み	音声対話による インタラクション

3.2 想定する学習の進め方

まず、構築した CAI システムの機能の設計について述べる前に、このシステムで想定する学習の進め方について述べる。本研究の目的である、複数の学習者が同時に学習可能な音声対話を用いた CAI システムを構築することによって、どのような学習を実現しようとしているかについて以下で具体的に述べる。

図 3.1 に本研究で構築する CAI システムによって実現しようとする学習のイメージを示す。図 3.1 に示すように、学習者の知識獲得を図るために、このシステムではコンピュータ上の仮想教師が学習者に対して講義形式で教授を行う形態をとる。一人の仮想教師が行う学習には同時に複数の学習者が参加でき、学習者たちはこの仮想教師と音声で対話しながら学習を進める。また、この教師は学習者に対して、図や写真など

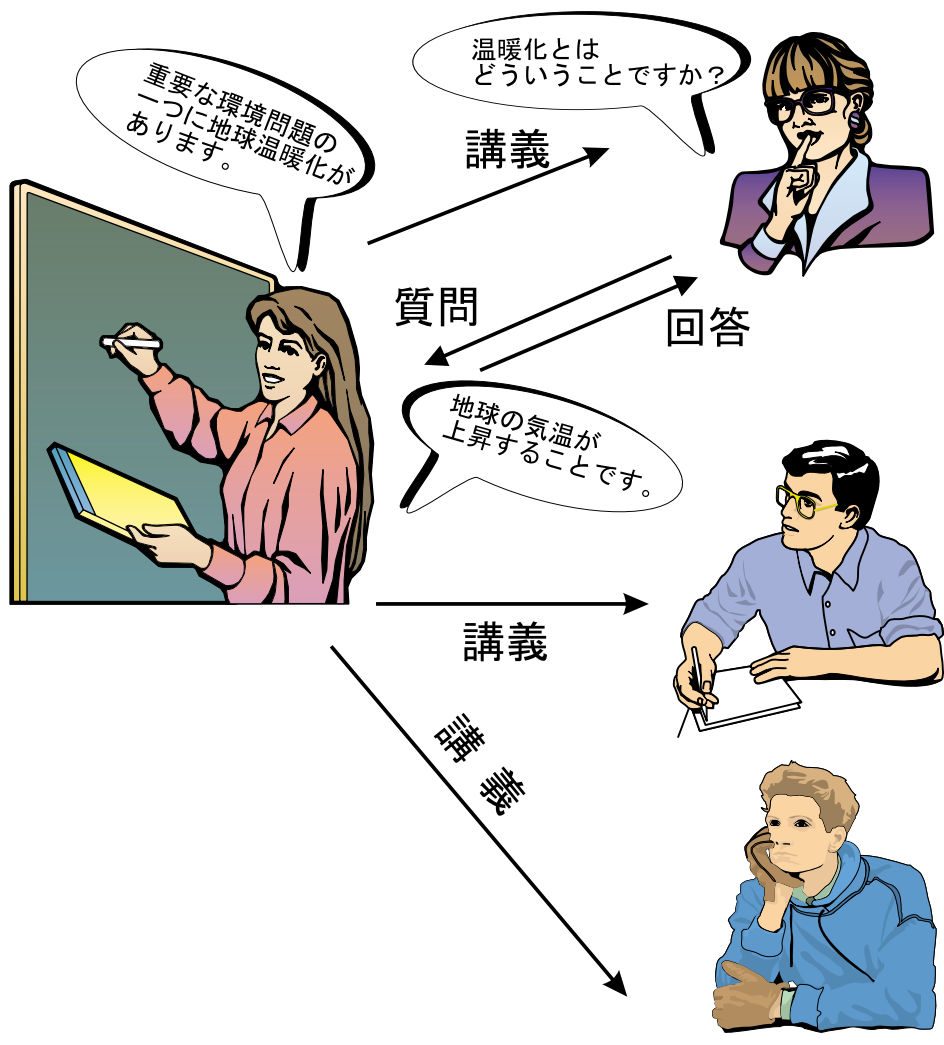


図 3.1: 複数人での学習のイメージ図

の目に見える資料を提示しながら音声による講義を行い、確認問題を出題する。また、教育内容に関する生徒の音声による質問に対して音声で答える。学習は講義を順に読み、聞き進めることで行う。学習者はその間にいつでも質問をすることができ、教師はそれに答え、また、講義を続ける。

3.3 構築する CAI システムの機能設計

本研究で構築する CAI システムにどのような機能が必要かを検討する。以下に挙げる 3.1 節において述べた機能上の 4 つの特徴をもとに機能を設計する。

- 音声による対話で学習を実現する。
- 講義形式で学習を行う。
- 学習者からの学習内容に関する質問を受け付け、それに適切な回答をする。
- 複数の学習者が同時に学習を行うことも可能である。

まず、CAI としての基本的な機能を備える必要があるために、はじめに一人で学習するために必要な機能について検討し、次いで、それを踏まえて複数人で学習するための機能について検討することにする。

以下では、音声対話、講義、質問回答、複数人での学習機能のそれぞれについて必要な機能を検討する。

3.3.1 音声対話機能

以下で、この CAI システムを実現するにあたって、音声対話を行う観点から必要な機能を表 3.2 に示し、それぞれについて以下に説明する。

音声認識機能・音声合成機能

音声対話を実現するための最も基本的な機能として必要なものは、相手の音声を「聞く」機能と、発話したい内容を音声にして「話す」機能である。つまり、音声認識機能と音声合成機能が必要となる。

実際には音声認識機能は学習者の発話をマイクを通してテキスト文字列に変換する機能、音声合成機能は発話用に用意された文字列を合成音声に変換してスピーカから出力する機能になる。

発話状態の認識機能・呼びかけ応答機能

実際の人間同士の音声対話においては、双方が同時に発話することはまれであるし、通常は避けられるべきことである。なぜなら、相手の発話を聞くことと自分が発話することの2つを同時に行うことは混乱をきたす原因となるからである。

従って、本システムのようなリアルタイムでの音声対話を実現するシステムにおいては、相手(学習者)が発話していることを検知し、その間は、発話を開始しない機能が必要となる。また、自分の発話中に相手が発話する「割り込み」があった場合、それに対して「聞く」姿勢を見せるために発話を停止する必要がある。これが発話状態の認識機能である。同様の理由で自分の発話中に「すみません」などと呼びかけられた場合にも発話を停止して「何でしょうか?」と聞き返す、呼びかけ応答機能も必要となる。

聞き返し機能

実際の人間同士の対話において、相手の発話が聞き取れず、何を言っているのかわからない場合がある。大抵の場合、そのようなときには「もう一度言って下さい」と聞き返す。音声対話システムにおいても音声認識における誤認識などで相手の発話を正しく理解できなかった場合、もう一度発話を促すことで正しい認識結果を得られるならば、発話理解の精度向上が期待できる。このような聞き返し機能を持つことによって、音声の誤認識などがあっても対話上の工夫で補うことができる。

表 3.2: 音声対話の機能構成

機能	説明
音声認識	学習者の発話音声をコンピュータで扱いやすいテキスト文字列に変換する。
音声合成	テキスト文字列の発話文を合成音声に変換して出力する。
発話状態認識	学習者が発話中かそうでないかを認識して、発話中なら教師は発話を止める。
呼びかけへの対応	「ちょっと」、「すみません」等の呼びかけに反応して、「はい、何でしょうか」と聞き返す。
聞き返し	学習者の発話の意味が分からなければ、「何を言っているのかわかりません、もう一度言ってください」と聞き返す。

3.3.2 講義機能

「講義によって学習を行う」という特徴を実現するために必要な機能を表 3.3に示し、それぞれを以下に述べる。

音声による講義機能・講義資料提示機能

まず、講義による学習を実現するにあたって必要となるのが、実際に講義を行う行為自体を実現する機能である。これが講義文を発話する講義文読み上げ機能と講義資料提示機能である。実際の教育においても講義形式をとるものでは、教師が音声で講義を行い、また同時に視覚的資料を提示することもよく行われ、効果的である。

講義文読み上げ機能は、予め講義として用意された文章を合成音声で発話することによって音声による講義を実現する機能である。また、講義資料提示機能は学習者の理解を助けるために、図や写真などの視覚的な補助資料を学習者に提示する機能である。これら2つの機能は、講義を実現するための基本機能である。

講義進行・分岐機能

よい講義を行うにはよく整理された教材が必要である。このシステムにおいてはフレーム型 CAI の教材構成を参考にして、講義読み上げ用テキストを内容のまとめりごとに講義フレームとして作成する。講義用教材はいくつもの講義フレームからなる木構造で構成され、講義ではその進行、分岐を制御する必要がある。そこでシステムに講義進行、分岐の機能を付加する。

講義進行の機能は、1つの講義フレームが終わった後、順番に次の講義フレームに進める機能である。教材は木構造を持つので講義の分岐が可能になる。この分岐を制御する機能が分岐機能である。分岐には2種類あり、いくつかの講義フレームの中からどれをはじめに講義するかを決める順序選択と、いくつかの講義フレームの中から1つだけ先に進む講義フレームを選ぶ講義分岐がある。順序選択機能では、まず、講義の中で学習者に対して問いかけが行われる。学習者はこの問いかけに対して音声で返答する。この結果から、次に進む講義が決定される。ここで選ばれなかった講義は後で行われる。講義分岐機能では、同様に講義の中で問いかけが行われ、それに対する答えによって次に進む講義フレームが決定される。この場合、選ばれなかった講義が読み上げられることはない。なお、この講義分岐によって、確認問題の正誤判定ができ、講義の枠組みの中で確認問題を実現することができる。

進行の確認

講義はフレームを順番に進めていくことで進行するが、1つの講義フレームが終わった後、強制的に次の講義へ進んでしまうと自分のペースで学習できるというCAIの利点の1つが失われてしまう。そのために1つの講義フレームが終わるごとに「次に進んでいいですか？」と確認を行い、学習者が了解した場合に次の講義へ進む。

講義再読・文の言い直し

学習者が講義中に聞き逃したり、理解できない場合には、もう一度講義を聞く要求が生じる。この要求を聞き入れるのが講義再読、文の言い直しの機能である。学習者がシステムの発話を聞き逃したとき、「もう一度言って下さい」と発話すると、現在発話している文をもう一度話す。また、学習者が講義をよく理解できない場合、「もう一度はじめからお願いします」、「戻ってください」と発話すると、教師はその講義フレームの先頭から再度講義を始める。

表 3.3: 講義の機能構成

機能	説明
講義文読み上げ	合成音声で講義の文を読み上げる。
講義資料提示	HTMLで作成された講義の補助資料を講義文の読み上げの間提示する。
講義進行	木構造で構成された講義を順番に進行する。
分岐（順序選択）	学習者に問いかけ、その答えによって、次の講義の順序を選択する。
分岐（講義分岐）	学習者に問いかけ、その答えによって、次に進む講義のコースを決定する。確認問題などの正誤判定に用いる。
進行確認	一つの講義が終わると学習者に「進んでいいですか」と問いかけ、了解の返事があると先へ進む。
講義再読	学習者の「戻ってください」、「もう一度お願いします」など講義の再要求があればもう一度その講義のはじめから発話する。

3.3.3 質問回答機能

講義だけでは言及しきれない内容についての教材として、また、講義中に疑問が生じたときに、学習者がいつでもその疑問点を解決できるように、学習者が質問すればそれに答えるような質問回答を行えることが必要である。この質問回答を実現するのに必要な機能を表 3.4にまとめ、以下にそれぞれについて説明する。

質問回答機能

質問回答を実現するための核となる機能が質問回答機能である。学習者が質問をすると、それに対して答えられる範囲内で適切な回答をする。

答えられないときの応答

学習者が質問をした際に、教師がそれに答えられない場合がある。質問されたことに関する知識がない場合、質問された単語が分からない場合、質問の意味を理解できない場合などであるが、それぞれの場合についてなぜ答えられないかを、答えられないときの応答として返す。

表 3.4: 質問回答の機能構成

機能	説明
質問回答	学習者が分からないことについて質問すると、教師の知識の範囲で適切な答えを返す。
答えられないときの応答	質問されたことが知識になく、答えられないときは「分かりません」と答える。

3.3.4 複数人での学習機能

複数人での学習は、前述の一人での学習の機能にいくつかの機能を付加して拡張することで実現する。一人の教師による学習を複数の学習者が同時に受けるような学習を実現することによって、学習者は他の学習者の学習の様子、質問などを聞くことができ、学習の相乗効果を得られる。このような「同期分散型学習」を実現するための機能を付加する。この機能によって、何人かの生徒が一人の教師の講義を受け、また質問できるような学習を実現できる。そのために実現すべき機能を表 3.5に示し、以下にそれぞれについて説明する。

複数学習者接続機能

同期分散型学習を実現するには、まず複数の学習者が同じ学習環境を共有する必要がある。ここでいう学習環境とはコンピュータ上に実現された仮想の学習環境で、一人の仮想の教師が学習を進行するものである。これには複数の学習者がコンピュータを接続して同じ学習環境を共有する複数学習者接続機能を実現する必要がある。

複数発話処理機能

このような同期分散型学習を音声対話によって実現するためには複数の学習者の発話を同時に並列に解析する必要がある。このような複数発話処理機能を付加することで、同時に複数の発話に対応する。実際には、最も早く発話した学習者に対して応答することになる。

他者の発話提示機能

このような環境では、教師対学習者の関係を複数作ることによって複数の学習者での学習を実現しているが、これだけでは他の学習者がどのような質問をしているかなどの情報が伝わらず、教師が他の学習者の質問に答えていても何について答えているかが分からない。この問題を解決するには質問をした学習者の発話が全ての学習者に提示される機能が必要である。この他者の発話提示機能は混乱することを防ぐために教師が理解した発話にのみ適用される。

表 3.5: 複数人での学習に関する機能

機能	説明
複数学習者接続機能	異なった場所にいる複数の学習者が1つの環境に接続し、一人の仮想教師によって同時に学習を行う。
複数発話処理機能	複数の学習者の発話を同時に処理し、それぞれの発話を理解する。
他者の発話提示機能	他の学習者の発話を画面上に提示し、誰のどのような発話に対して教師が答えているのかを分かるようにする。

3.3.5 学習者の発話についての考察

本研究で構築する CAI システムには以上で述べたような機能を持たせる。これを音声対話で行うために、学習者の発話についてどの発話を受け入れるかを決める必要がある。このシステムで受け入れる必要のある学習者の発話は、質問回答のための質問文、講義再読要求や、講義の一時停止要求などの要求文、また、講義において分岐の際のシステムからの問いかけに対する返答文である。この質問文、要求文、返答文を受け入れて理解できるシステムにすれば、本研究で目的とする音声対話による学習を実現できる。この3種類の分類を学習者の発話タイプ分類と呼ぶ。

以下では3つの発話タイプについてそれぞれどのようなものであるかを述べる。

質問文

学習者は講義などにおいて疑問点が生じたときにシステムに対して質問を行う。システムはそれに対して回答し、疑問点の解消を図る。このために学習者がシステムに対して発話する文が質問文である。

本研究では学習において基本的な事柄を質問するときにはまず考えられる4つの分類を質問文として想定した。それらは、「意味」、「原因・理由」、「結果」、「対応・対策」である。これらを質問タイプ分類とし、以下でそれぞれの質問タイプの説明をする。

- 意味

物事の意味や言葉の定義を聞く質問。

例:「地球温暖化とは何ですか。」

- 原因・理由

物事の原因やなぜそうなるかを聞く質問。

例:「地球温暖化の原因は何ですか。」

例:「なぜ地球温暖化は起こるのですか。」

- 結果

物事の結果についてどうすればどうなるかについて聞く質問。

例:「地球温暖化が進むとどうなりますか。」

- 対応・対策

その物事を防ぐにはどうすればよいのか、どのような対応をとればよいのかを聞

く質問。

例：「地球温暖化を防ぐにはどうすればよいですか。」

上記の質問タイプと、質問文の中で学習者が何について聞いているのかを表している「キーワード」によって学習者の質問の意図を特定できる。質問文は以下のようなものを想定する。

- 「対象」の「意味」は何か
- 「対象」が「動作、変化」した「原因・理由、結果、対応・対策」は何か

質問タイプが意味の場合、キーワードは「対象」のみである。前述の質問文の例で言えば、「地球温暖化」の「意味」を聞く質問で、キーワードは「地球温暖化」である。また、原因・理由、結果、対応・対策の3つの質問タイプではキーワードは2種類ある。それは、「対象」と「動作、変化」であり、原因・理由の質問タイプの説明で挙げた例について考えると「対象」のキーワードは「地球温暖化」であり「動作、変化」のキーワードは「起こる」である。この質問文は「地球温暖化」が「起こる」「理由」を聞いている質問であると解釈できる。以降ではこの「対象」キーワードをキーワード1、「動作、変化」キーワードをキーワード2と呼び、単にキーワードと言うときは「対象」キーワードを指すことにする。

要求文

講義中に講義を一旦停止して欲しい場合や、もう一度同じ講義をして欲しい場合がある。このような場合、学習者はその要求をシステムに伝えることで講義の一時停止や講義再読をさせることができる。このような学習者の要求をシステムに伝えるための発話が要求文である。

要求には「待って下さい」などのように講義の一時停止を要求する「一時停止要求」、
「もう一度言って下さい」など講義の再読を要求する「再読要求」、ひとつの講義フレームが終了際のシステムからの「進んでもいいですか」の問いに了解を伝える「了解」の3種類があり、これらを要求タイプとする。

返答文

分岐を行う講義では、はじめにシステムが講義の中で学習者に対して問いかけを行い、その問いかけに対する学習者の返答によって分岐先が決定される。この学習者の

返答が返答文である。例えば、「地球温暖化の原因となるものは何ですか」といったシステムの問いに対する「温室効果ガスです」などの返答が返答文である。

第 4 章 CAI システムの設計と構築

本研究では第 3 章で述べた機能を備える CAI システムを構築した。本章では構築した CAI システムについて実際のシステムの構成、機能について述べる。

4.1 CAI システムのハードウェア構成

ここでは音声対話インタフェースを持つ CAI システムを実現するためのシステム構成について述べる。3.1 節で検討した 3 点の構成上の特徴は以下のとおりである。

- ネットワークを用いたクライアントサーバ方式によってシステムを実現する。
- サーバによって教材データを一元管理する。
- パーソナルコンピュータを用いて安価なシステム構成を実現する。

以下で述べるようなシステム構成をとることでこれらの特徴を備えた CAI システムを実現する。

まず、システムのハードウェア構成を図 4.1 に示し、説明する。本研究ではクライアントサーバ方式でシステムを構成する手法をとる。このため、システムは一台のサーバコンピュータと何台かのクライアントコンピュータから構成される。サーバコンピュータとクライアントコンピュータはインターネットで接続され、ソケット通信を用いて通信することで学習と対話に関するデータ送受信を実現する。サーバコンピュータは文解析、対話管理、学習に関する処理などを行い、クライアントコンピュータは学習者とのインタフェースとして動作し、音声認識、音声合成、画像提示を行う。学習者はクライアントコンピュータに向き合って学習を進める。サーバコンピュータ上には教材データの一元管理のためにデータベース管理システムが動作しており、そこに全ての教材が格納されている。

4.2 CAI システムのソフトウェア構成

音声対話によって学習を進める音声対話インタフェースを実現するために、2.3 節において述べたような音声認識部、音声合成部、文解析部、対話管理部、タスク処理部

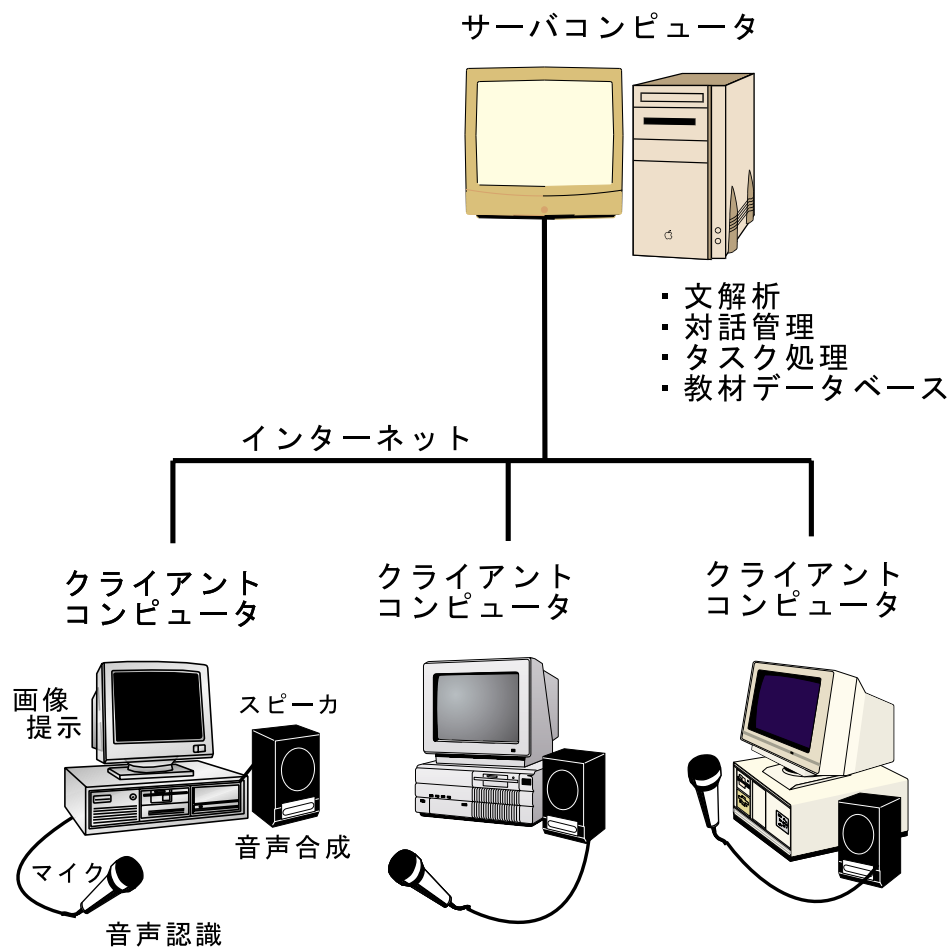


図 4.1: 構築する CAI システムのハードウェア構成

に加えて、学習者に補助資料を提示する画像提示部からなるシステム構成にする。これを図 4.2に示す。

以下でこの処理の流れの概要について説明する。まず、学習者の発話は音声認識部で自動認識され、テキスト文字列へ変換される。次に、この学習者の発話文字列は文解析部に送られ、学習者の発話意図が抽出される。さらに、この結果は対話管理部で処理されて学習者に対する反応が決められる。この反応は音声合成部にテキスト文字列を送り合成音声として発話させることや、画像提示部の WWW ブラウザに画像などの補助資料を提示させることで実現する。

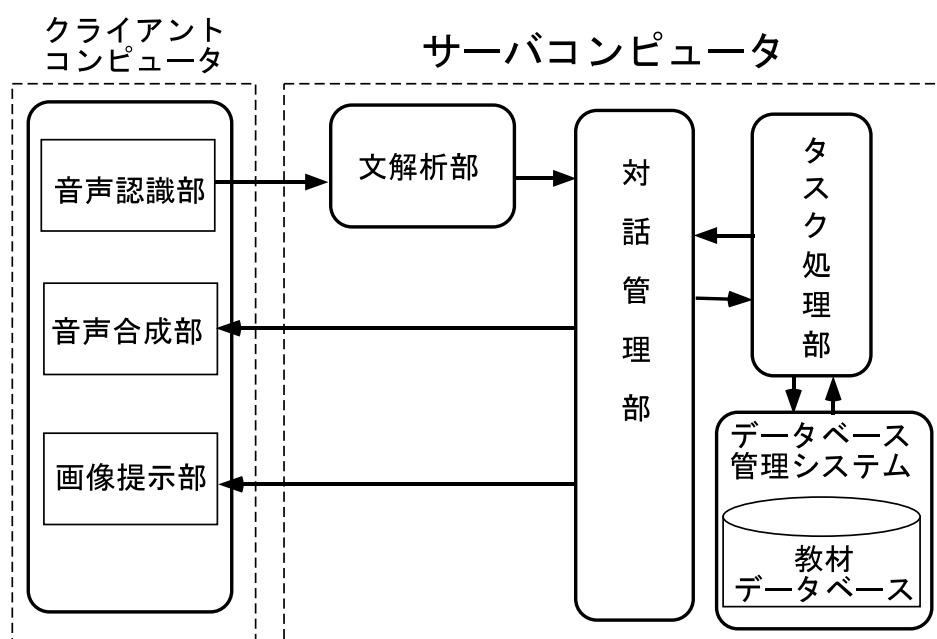


図 4.2: CAI システムのソフトウェア構成

さらに詳細なシステムの構成を図 4.3に示し、以下で各部の機能、処理について述べる。

音声認識部

音声認識部は図 4.4の太い赤線で囲まれた斜線部である。音声認識部では音声認識と発話音声レベル検出、句点挿入を行う。

音声認識処理

音声認識処理はマイクを通して入力された学習者の発話音声进行处理し、漢字仮名混じり文のテキスト文字列へ変換するものである。この処理では発話音声をリア

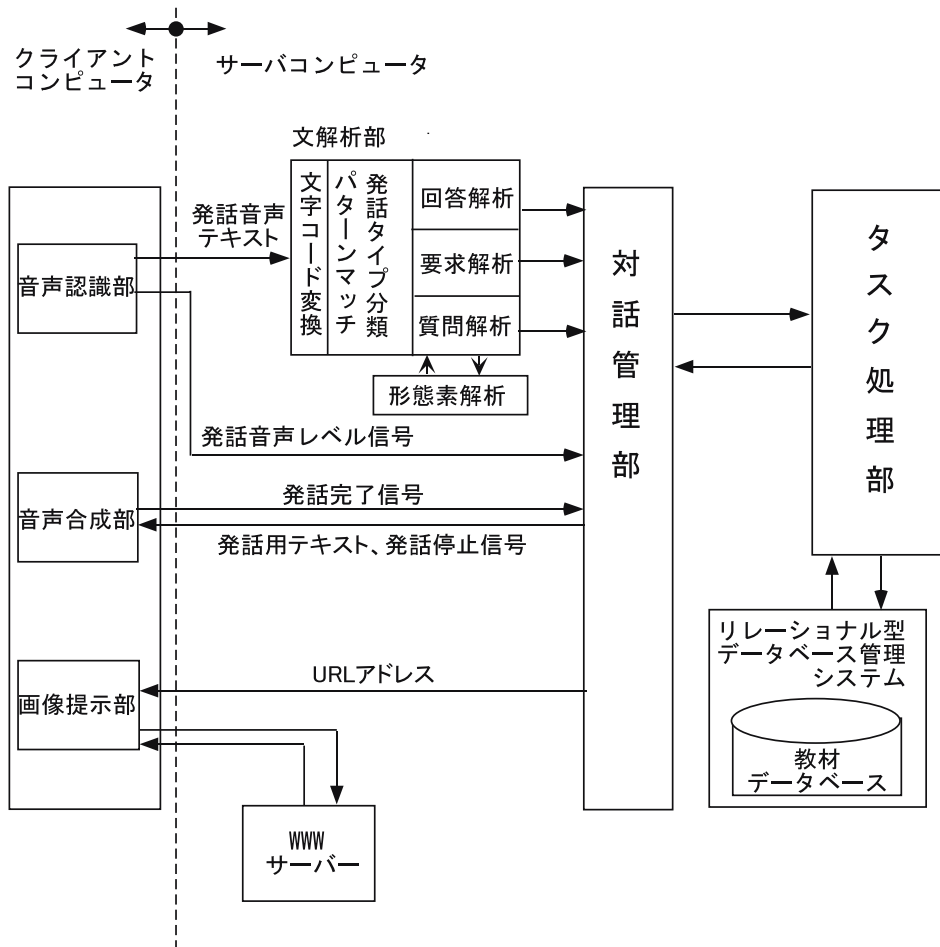


図 4.3: CAI システムの詳細なソフトウェア構成

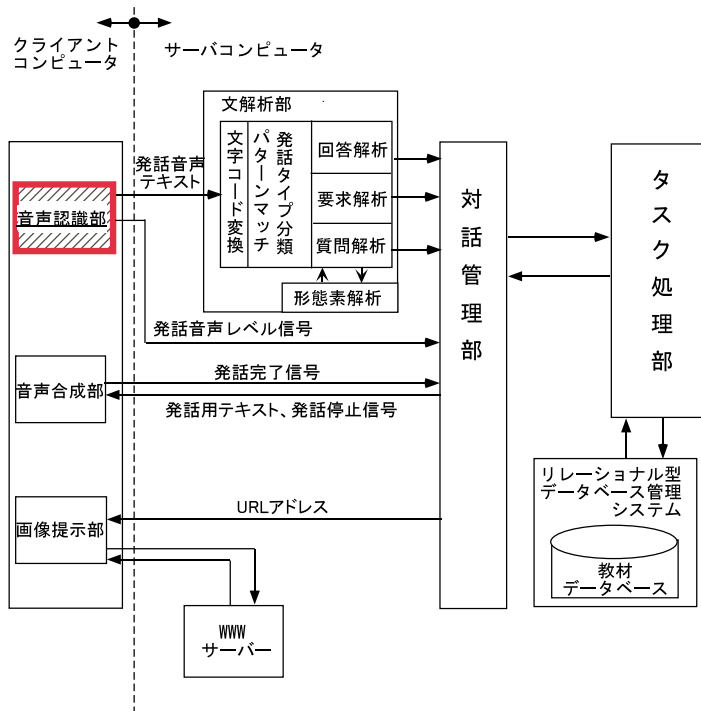


図 4.4: CAI システムのソフトウェア構成 (音声認識部)

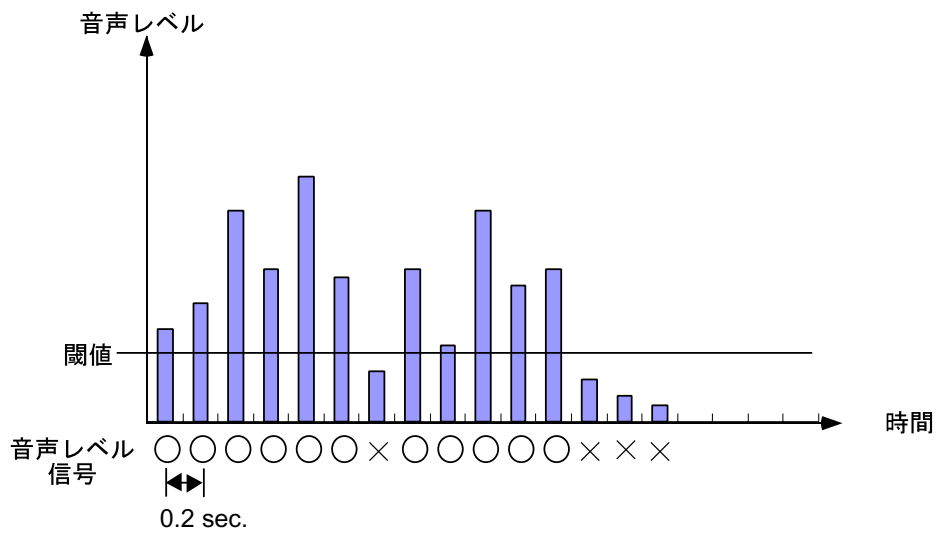
リアルタイムで自動認識する。

発話音声レベル検出

発話音声レベル検出は学習者の発話音声の音量がある一定の閾値を越えたときに発話があったことを検出し、発話音声レベル信号を対話処理部へ送信するものである。この処理もリアルタイムで行われる。この処理を例を挙げて説明する。図 4.5は「エネルギー問題とは何ですか」と発話した場合の音声レベルを表している。この場合、0.2 秒毎に音声レベルを検出し、音声レベルが閾値を越えていれば、発話音声レベル信号を送信する。また、学習者が発話していない場合に雑音が入った場合を図 4.6 に示す。この場合、雑音ではあるが音声レベルが閾値を越え、発話音声レベル信号が何度か送信される。このような雑音を発話として検出しないために対話管理部において、3 回以上連続した音声レベル信号を受け取った場合にはじめて学習者が発話中であることを検出する。

句点挿入処理

句点挿入処理は、発話音声の認識において、文の切れ目に句点を挿入する処理である。音声対話において文の切れ目を検出することは必要である。この処理は図



発話 エネルギー問題とは何ですか

図 4.5: 音声レベル検出 (通常の発話)

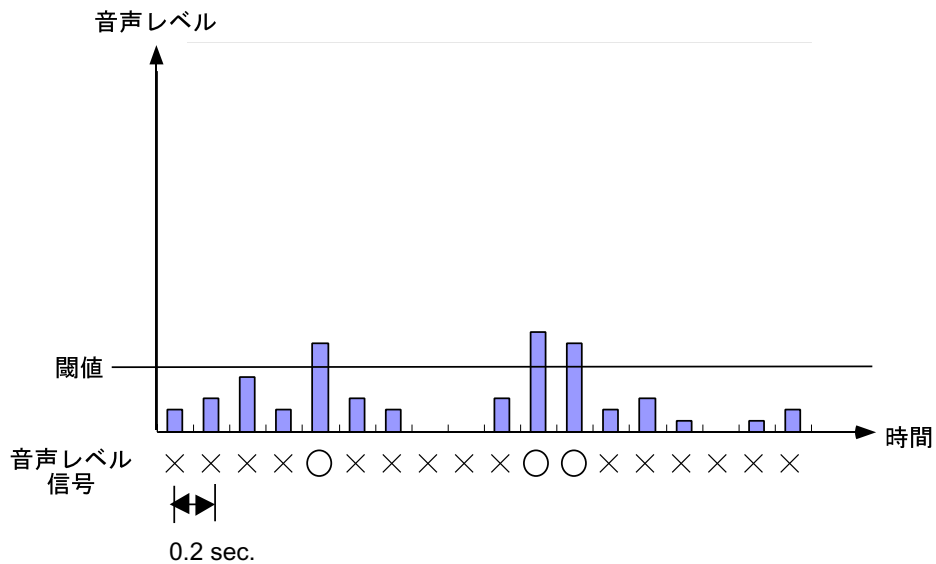


図 4.6: 音声レベル検出 (雑音などの場合)

4.7に示すように、発話音声レベルが一定時間 (0.7sec) 以上閾値を下回っている場合、発話が完了したとみなして発話認識結果に句点を挿入するものである。

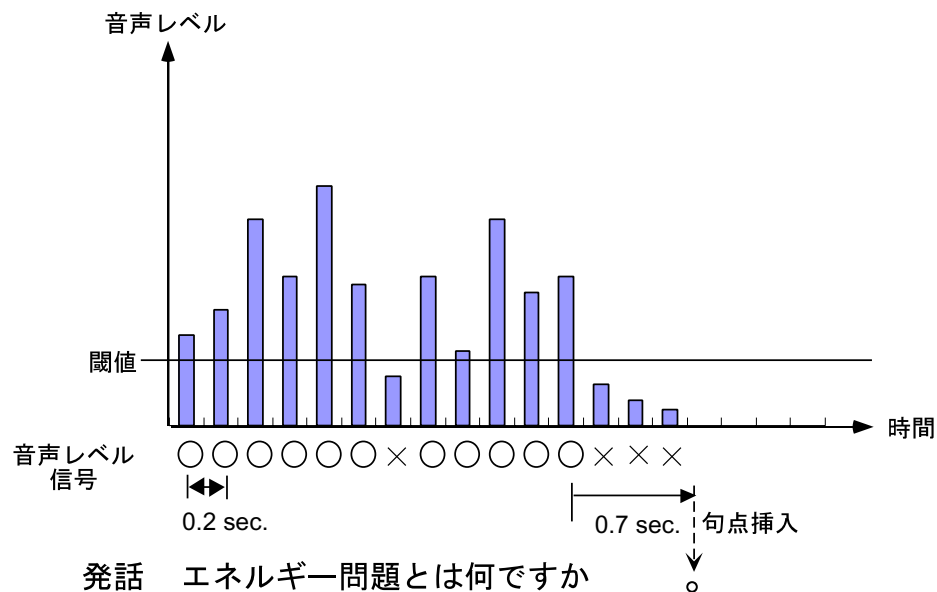


図 4.7: 句点挿入処理

音声合成部

音声合成部は図 4.8の太い赤線で囲まれた斜線部である。音声合成部では音声合成と発話完了検出を行う。

音声合成処理

音声合成処理は、対話管理部から送られてきた合成音声発話用テキスト文字列を処理して合成音声としてスピーカから出力するものである。この処理によってシステムが「話す」ことが可能になる。

発話完了検出

発話完了検出は、システムが発話し終ったとき、発話完了信号を対話管理部に送信するものである。対話管理部はこの発話完了信号を受け取ることでシステムの発話出力が完了したことを知る。

文解析部

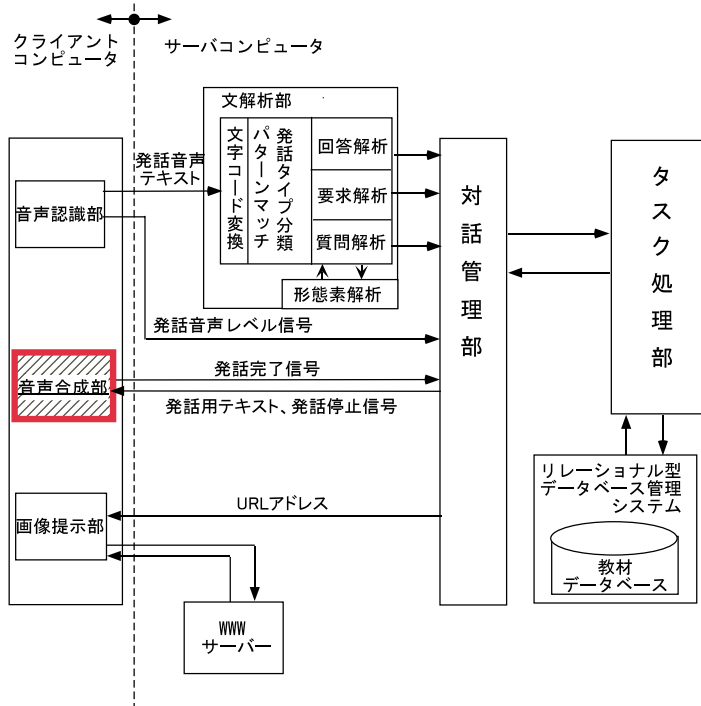


図 4.8: CAI システムのソフトウェア構成 (音声合成部)

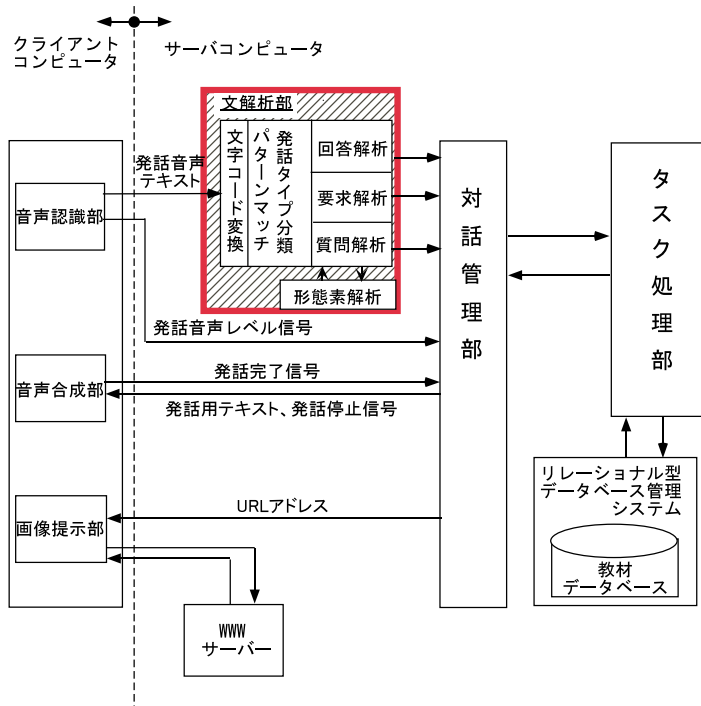


図 4.9: CAI システムのソフトウェア構成 (文解析部)

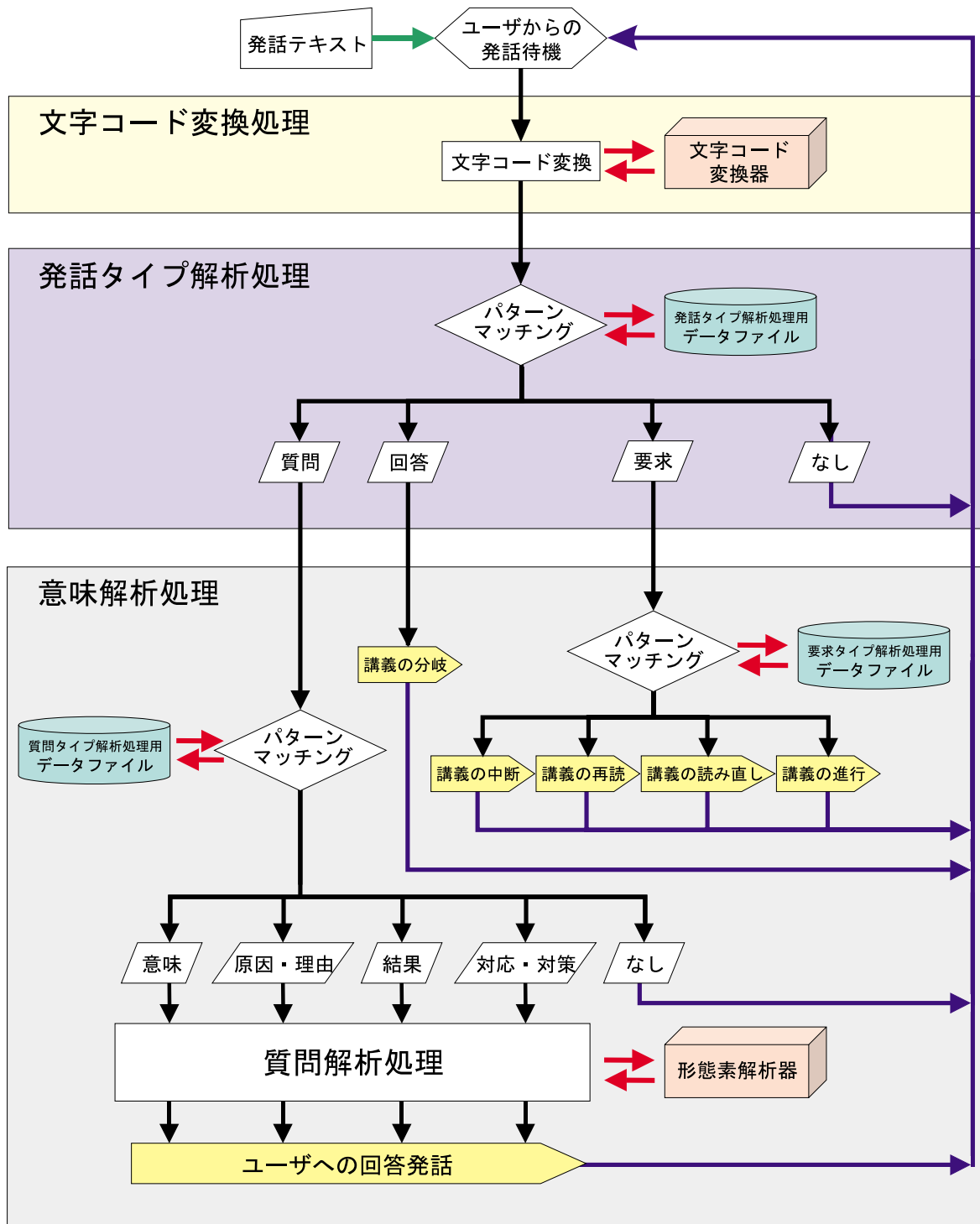


図 4.10: 文解析部の処理の流れ

文解析部は図 4.9の太い赤線で囲まれた斜線部である。文解析部では音声認識によって認識された文を解析処理することにより学習者の発話の意図を抽出する。この処理は文字コード変換、発話タイプ解析、意味解析の一連の3つの処理から構成される。また、意味解析では形態素解析を行う。これらの処理の流れを図 4.10に示し、以下で説明する。

文字コード変換処理

このシステムではサーバコンピュータとクライアントコンピュータが通信する際、データとして日本語文字列を扱う。しかし、日本語文字列には数種類の文字コード規格が存在し、これらの中で互換性がないため文字コードを統一する必要がある。このため、文字コード変換フィルタを通し、サーバ側の文字コードに変換する。これが文字コード変換処理である。

発話タイプ解析処理

この CAI システムでは 3.2節、3.3節で述べたとおり講義と質問回答によって学習を構成している。これらの機能を全て用いて学習を進行していくために 3.3.5項で考察したとおり、以下の3種類の学習者の発話を処理する。

- 質問
学習の疑問点について質問するための発話。
例:「地球温暖化とは何ですか。」
- 回答
分岐を行う講義において、講義の中で教師から質問された場合に、学習者が返す回答の発話。
例:「温室効果ガスです。」
- 要求
講義中に講義を一旦停止して欲しい場合や、もう一度同じ講義をして欲しい場合などに学習者の要求を伝える発話。
例:「もう一度お願いします。」

学習者の発話に対してこの3種類の分類を行うのが発話タイプ解析である。この処理は学習者の発話の中にそれぞれの発話タイプを特定できる特有の語を含んでいるかどうか、パターンマッチング処理を行うことで実現する。表 4.1にその一例

を示す。

表 4.1: 発話の分類と解析用単語

発話の種類	質問、要求、回答のタイプ	パターンマッチング用単語の例
質問		どうして
		なぜ
		ですか
		ますか
		何
		か。
要求	停止要求	すみません
		まってください
	講義再読要求	もう一回最初から
	講義読み直し要求	もう一度言って下さい
	了承	はい
回答	回答	です。
		。(上記以外)

意味解析処理

発話タイプを解析した結果から、更に意味解析処理を行って発話の意味を特定する。この意味解析処理には発話タイプごとに質問解析処理、要求解析処理、回答解析処理の3種類があり、それぞれの処理を行うことによって学習者の発話の意図を解析する。以下でそれぞれの処理の概要について述べる。

- 質問解析処理

質問解析処理は学習者の質問文から質問タイプとキーワードを抽出し、学習者が何について知りたいかを解析する処理である。学習者の質問として想定されるものについて検討し、意味、原因・理由、結果、対応・対策の4つの分類を質問として扱うことにした。この4つの分類を質問タイプとし、この分類によって学習者の質問に対応する。

この解析処理ではまず、質問文に対してパターンマッチングを行うことにより質問タイプを特定する。質問タイプ特定用の単語の例を表 4.2に示す。

表 4.2: 質問の分類と解析用単語

質問のタイプ	パターンマッチング用単語の例
意味	の意味は
	とは何
原因・理由	なぜ
	原因は
結果	影響は
	結果は
対応・対策	対策は
	防ぐ

この質問タイプの特定を行った後、何について聞いているのかを表す「キーワード」を抽出する。この処理は、形態素解析による単語の品詞情報を利用したキーワード抽出である。この処理については 4.2.3項の質問文解析手法で詳しく述べる。

- 要求解析処理

本システムが受け入れる学習者からの要求は講義の中断、講義の再読、講義の読み直し、了承(講義の進行)の4種類である。これらは学習者が発話した要求文に対してパターンマッチングを行うことによって、どの要求であるかを解析する。その解析に用いる単語の例は表 4.1の要求の項目に示した通りである。この要求解析で要求が解析された結果は対話管理部で処理され、講義の中断、講義の再読、講義の読み直し、講義の進行が行われる。

- 応答解析処理

分岐を行うときなど講義の中でシステムが学習者に問いかけをすることがある。このときの学習者の応答を解析するのが応答解析処理である。この処理は、形態素解析による単語の品詞情報を利用したキーワード抽出である。この応答は講義の分岐の際に用いられる。

対話管理部

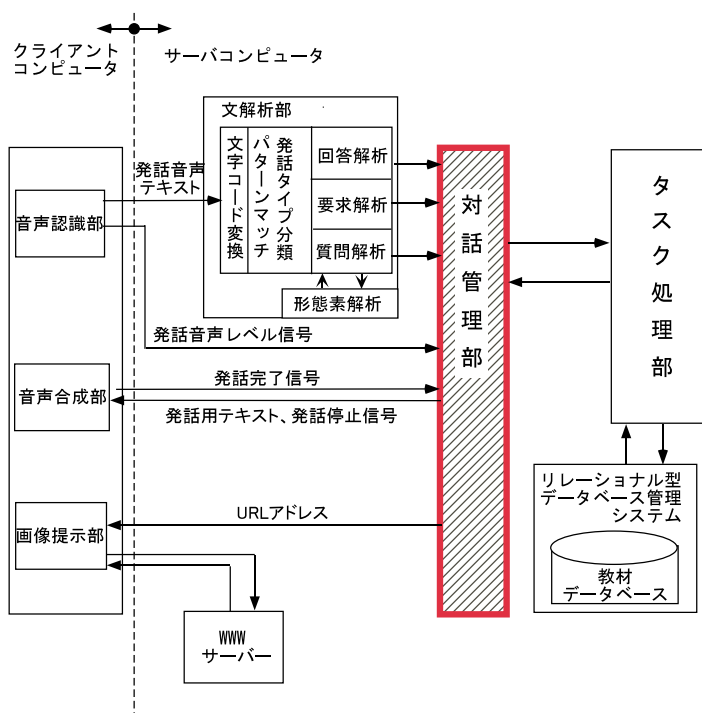


図 4.11: CAIシステムのソフトウェア構成 (対話管理部)

対話管理部は図 4.11の太い赤線で囲まれた斜線部である。対話管理部では対話に関する反応生成、同期などの処理を行う。ここでは、学習者の発話の解析結果、発話音声レベル信号、発話完了信号を受信する。これらの信号を処理してどのような反応を返すかを決定する。この反応は対話ルールと呼ばれるある一定の規則に従って決定される。

図 4.12に対話管理部の処理の流れを示し、この規則について説明する。学習者が発話した時、発話音声レベル信号と認識された学習者の発話が対話管理部へ送られる。まず、対話管理部では、現在、システムが講義中であるか、学習者へ何らかの反応をしている最中かを判断する。そして、質問に対する回答など学習者への反応中であるときには、システムは質問回答などを中断させないために、学習者の発話を無視して学習者への反応を続ける。逆に、システムが講義中である場合、学習者へ反応を返す。対話管理部で3回以上連続した音声発話レベル信号を受け取ると、学習者が発話したことを認識する。そして、現在の講義の発話を停止し、学習者の発話が終るのを待つ。

次に、文解析部から学習者の発話の解析結果が得られる。発話内容が質問である場合、この処理結果をタスク処理部に渡し、タスク処理部での教材データベース検索に

よって回答を得る。この回答文を対話管理部から音声合成部へ送り、学習者への反応として発話する。

このようにして対話管理部ではシステムが対話を行うための処理をし、対話を円滑に行うための役割を担う。

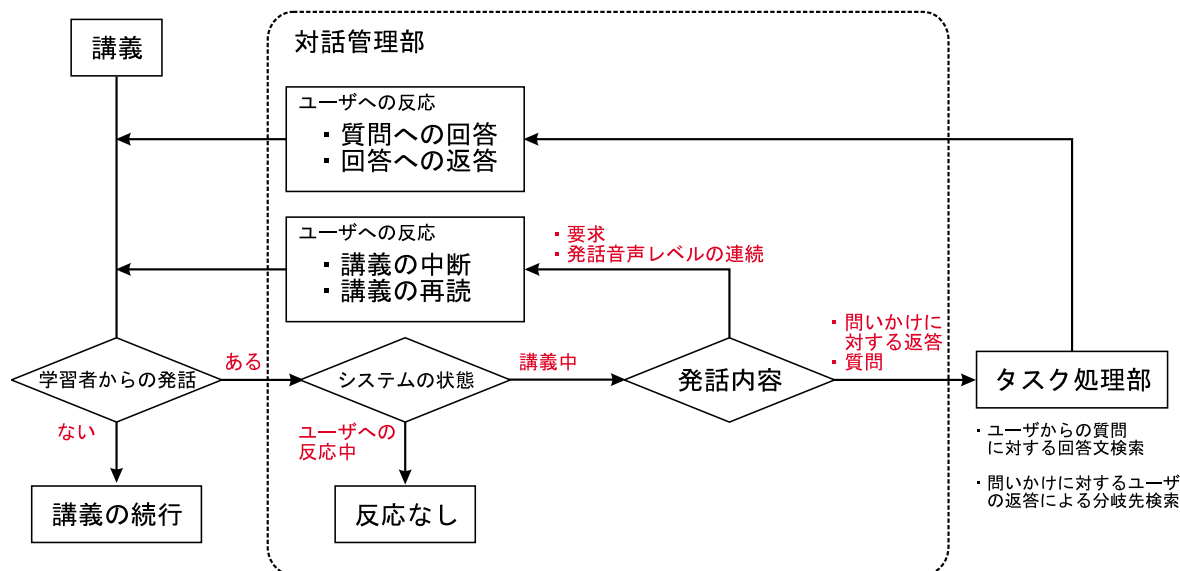


図 4.12: 対話管理部の処理

タスク処理部

タスク処理部は図 4.13の太い赤線で囲まれた斜線部である。タスク処理部では、データベース管理システムに教材データを保持し、そのデータを検索することで講義、質問回答などの学習に関連する処理を行う。

講義進行処理

講義進行処理は通常の講義の進行の際に行われる。この処理では、対話管理部から要求があった時、次の講義の発話用テキストと補助資料の HTML ファイルのある場所を示す URL アドレスを教材データベースから検索し、対話管理部へ渡す。

分岐処理

講義には、通常の講義のほかに、学習者の反応に応じたインタラクティブな学習を行うために学習者に問いかけをして、その返答によって次に進む講義を決める分岐の機能を持っている。

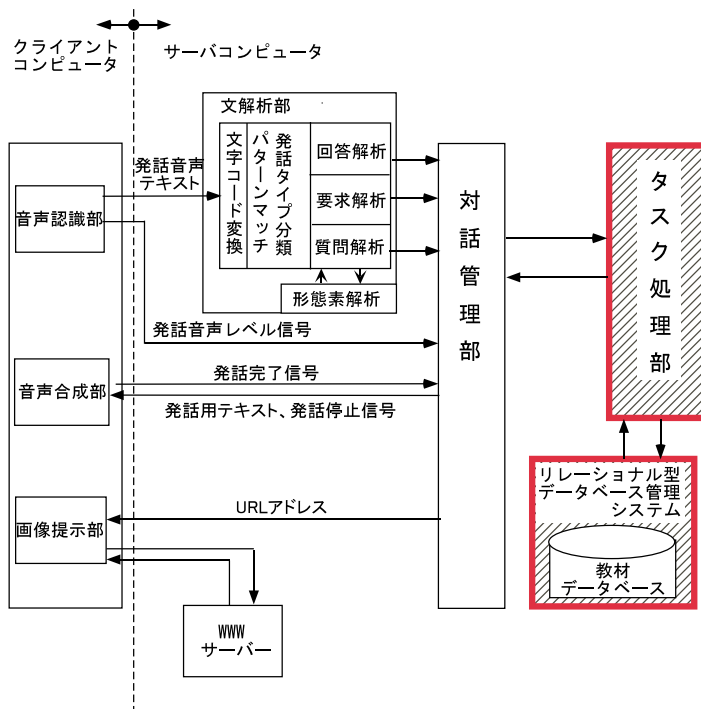


図 4.13: CAI システムのソフトウェア構成 (タスク処理部)

分岐には、次の講義の進行順だけを変更する順序選択や、いくつかの講義の中から 1 つを選択して進む講義分岐の 2 種類がある。この様な分岐を処理するのが分岐処理である。分岐を行う講義では、学習者に問いかけをし、学習者の返答によって次の講義を決める。教材データベースにはこの分岐を制御するデータが格納されている。タスク処理部においてこの分岐のために行う処理は学習者の返答から分岐を制御するデータを検索し、次の講義が何かを決定することである。この実現手法については 4.2.2 項で詳しく説明する。

質問回答検索処理

質問回答処理では、文解析部で解析された質問文の質問タイプとキーワードが対話管理部から送られ、それを教材データベースの中から検索する。この実現手法については 4.2.3 項で詳しく説明する。

また、このタスク処理部の教材データの詳細については 4.2.1 項で説明する。

以下ではこの構築したシステムでどのようにして機能を実現しているかをデータベース、講義、質問回答、複数学習者での学習のそれぞれについて説明する。

4.2.1 教材データベースの構成

これまでに述べたように本研究で構築する CAI システムでは学習で用いる全ての教材を教材データベースとして一元管理する。この教材データは 3.1 節で述べたように作成、更新が容易でなければならない。本システムでは、教材として全ての発話用の文章を用意して、これを教材データとする。このように発話用の文章をテキスト文字列の形で直接教材として用いる手法では、教材の提示 (発話) において常に決まった文章しか提示できないという問題点はあるものの、教材作成の労力が少なくすむという利点がある。この手法では、知的 CAI などのように教材を抽象化して作成する作業と比較して、文章を作成するだけの比較的簡単な作業で教材作成が行える。以下で、本システムの学習機能を実現する観点から教材データベースについて述べ、次いで講義用教材、質問回答用教材の構成についてそれぞれ述べる。

教材は全てテキスト文字列で記述されており、教材データベースに格納されている。データベースの構成を図 4.14 に示す。教材データベースはリレーショナル型データベース管理システム (Relational Database Management System: RDBMS) 上のデータベースである。リレーショナル型データベースではいくつかのテーブルからデータベースが構成される。テーブルは 2 次元の表の構造でデータを保持しており、この表にはいくつかの項目がある。このような構造のテーブルから所望のデータを検索することができる。

本研究で構築するシステムのデータベースは講義データベース、質問回答データベースと名前変換データベースの 3 つのデータベースからなる。講義データベースは講義テキストを持つ講義テーブルと、分岐情報を記述した分岐テーブルからなり、また、質問回答データベースには質問の種類に応じた意味テーブル、原因・理由テーブル、結果テーブル、対応・対策テーブルの 4 つのテーブルがある。また、名前変換データベースにはデータベースの検索の際、統一された名前に変換するための名前変換テーブルがある。

教材データベースのテーブルの仕様を表 4.3 ~ 4.9 に示し、また、その例を表 4.10 ~ 4.16 に示し、以下でそれぞれについて説明する。

講義データベース

講義は、図 4.15 に示すように講義フレームによって構成される木構造をしている。木構造は最上位の階層に 1 つの根をもつ。この図では「エネルギー問題」が根にあた

教材データベース

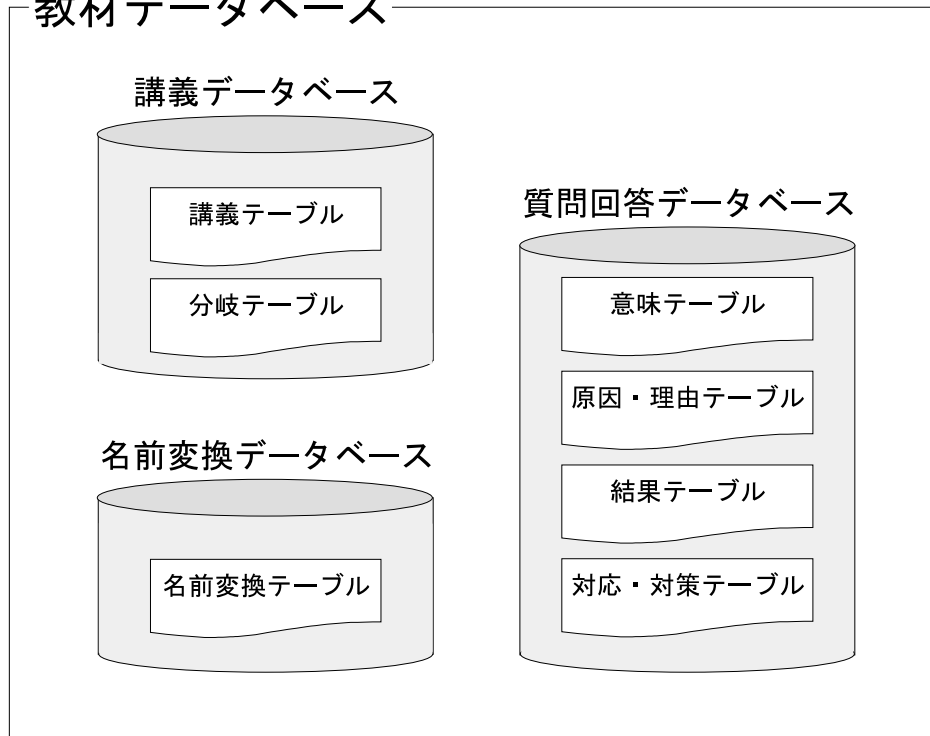


図 4.14: データベースの構成

表 4.3: 講義テーブルの仕様

タイトル	講義フレームのタイトル
内容	システムが読み上げる講義文
URL	表示画面用 URL
講義番号	講義フレーム番号
親講義番号	親の講義フレーム番号
フラグ	未読・既読
属性	講義形式・問いかけ形式・質問形式

表 4.4: 分岐テーブルの仕様

分岐元講義番号	問いかけ形式の講義フレーム番号
分岐先講義番号	進行先の講義フレーム番号
キーワード	問いかけに対するユーザの返答

表 4.5: 名前変換テーブルの仕様

入力語	ユーザから発話された単語
検索用単語	検索用に変換した単語

表 4.6: 意味テーブルの仕様

キーワード	「意味」検索用単語
定義	キーワードの定義 (ユーザへの回答)

表 4.7: 原因・理由テーブルの仕様

キーワード	「原因・理由」検索用単語
+ 要素 (増えるなど)	キーワードが + 要素となる場合の原因・理由 (ユーザへの回答)
- 要素 (減るなど)	キーワードが - 要素となる場合の原因・理由 (ユーザへの回答)
発生要素 (起こるなど)	キーワードが起こる原因・理由 (ユーザへの回答)
単独 (動詞がない場合)	キーワードの原因・理由 (ユーザへの回答)

表 4.8: 結果テーブルの仕様

キーワード	「結果」検索用単語
+ 要素 (増えるなど)	キーワードが + 要素となる場合の結果 (ユーザへの回答)
- 要素 (減るなど)	キーワードが - 要素となる場合の結果 (ユーザへの回答)
発生要素 (起こるなど)	キーワードが起こった場合の結果 (ユーザへの回答)
単独 (動詞がない場合)	キーワードの結果 (ユーザへの回答)

表 4.9: 対応・対策テーブルの仕様

キーワード	「対応・対策」検索用単語
+ 要素 (増えるなど)	キーワードが + 要素とならないための対応・対策 (ユーザへの回答)
- 要素 (減るなど)	キーワードが - 要素とならないための対応・対策 (ユーザへの回答)
発生要素 (起こるなど)	キーワードを起こさないための対応・対策 (ユーザへの回答)
単独 (動詞がない場合)	キーワードの対応・対策 (ユーザへの回答)

表 4.10: 講義テーブルの例

タイトル	内容	URL	講義番号	親講義番号	フラグ	属性
温室効果	温室効果とは、...	http://...	5	4	未読	通常講義
異常気象	温暖化が進むと、...	http://...	10	7	未読	順序選択
集中豪雨	異常気象により、...	http://...	11	10	未読	通常講義
干ばつ	温暖化による降水...	http://...	13	10	未読	通常講義
農業不振	温暖化による土地...	http://...	16	7	未読	講義分岐

表 4.11: 分岐テーブルの例

分岐元講義番号	分岐先講義番号	キーワード
4	6	二酸化炭素
10	13	水不足

表 4.12: 名前変換テーブルの例

入力語	検索用単語
地球温暖化ガス	温室効果ガス
温暖化ガス	温室効果ガス
グリーンハウスガス	温室効果ガス
温室効果ガス	温室効果ガス
化石燃料	化石資源
化石資源	化石資源

表 4.13: 意味テーブルの例

キーワード	定義
温暖化	人間の活動により、温室効果ガスが増加して温室効果が高まり、地表面の温度が上昇することです。
化石資源	石炭、石油など、地中に存在しているエネルギー源となるものです。無限に存在しているものではありません。

表 4.14: 原因・理由テーブルの例

キーワード	+ 要素 (増えるなど)	- 要素 (減るなど)	発生要素 (起こるなど)	単独 (動詞がない場合)
温暖化	温室効果ガスの増加により、温室効果が起こるためです。	温室効果ガスが減少すると温室効果が弱まるためです。	温室効果ガスの増加により、温室効果が起こるためです。	温室効果ガスの増加により、温室効果が起こるためです。
化石資源	新たな油田などが発見されたり、発掘技術が進歩するためです。	限りある資源であり、近年、大量に消費されているためです。	有機物が地中に堆積し長い時間をかけて、変化するためです。	-

表 4.15: 結果テーブルの例

キーワード	+ 要素 (増えるなど)	- 要素 (減るなど)	発生要素 (起こるなど)	単独 (動詞がない場合)
温暖化	海水面の上昇、異常気象などが起きます。	今までどおりの気候で平穏に暮らせます。	海水面の上昇、異常気象などが起きます。	海水面の上昇、異常気象などが起きます。
化石資源	増加しても一時的なもので、将来的には、枯渇に向かうでしょう。	エネルギー源として、他のものに頼る必要が出てくるでしょう。	エネルギーを作り出したり、様々な用途にも使えます。	-

表 4.16: 対応・対策テーブルの例

キーワード	+ 要素 (増えるなど)	- 要素 (減るなど)	発生要素 (起こるなど)	単独 (動詞がない場合)
温暖化	温室効果ガスの増加を止め、森林を保護することなどが重要です。	-	温室効果ガスの増加を止め、森林を保護することなどが重要です。	温室効果ガスの増加を止め、森林を保護することなどが重要です。
化石資源	-	省エネや再生利用可能なエネルギーの利用を促進することです。	-	-

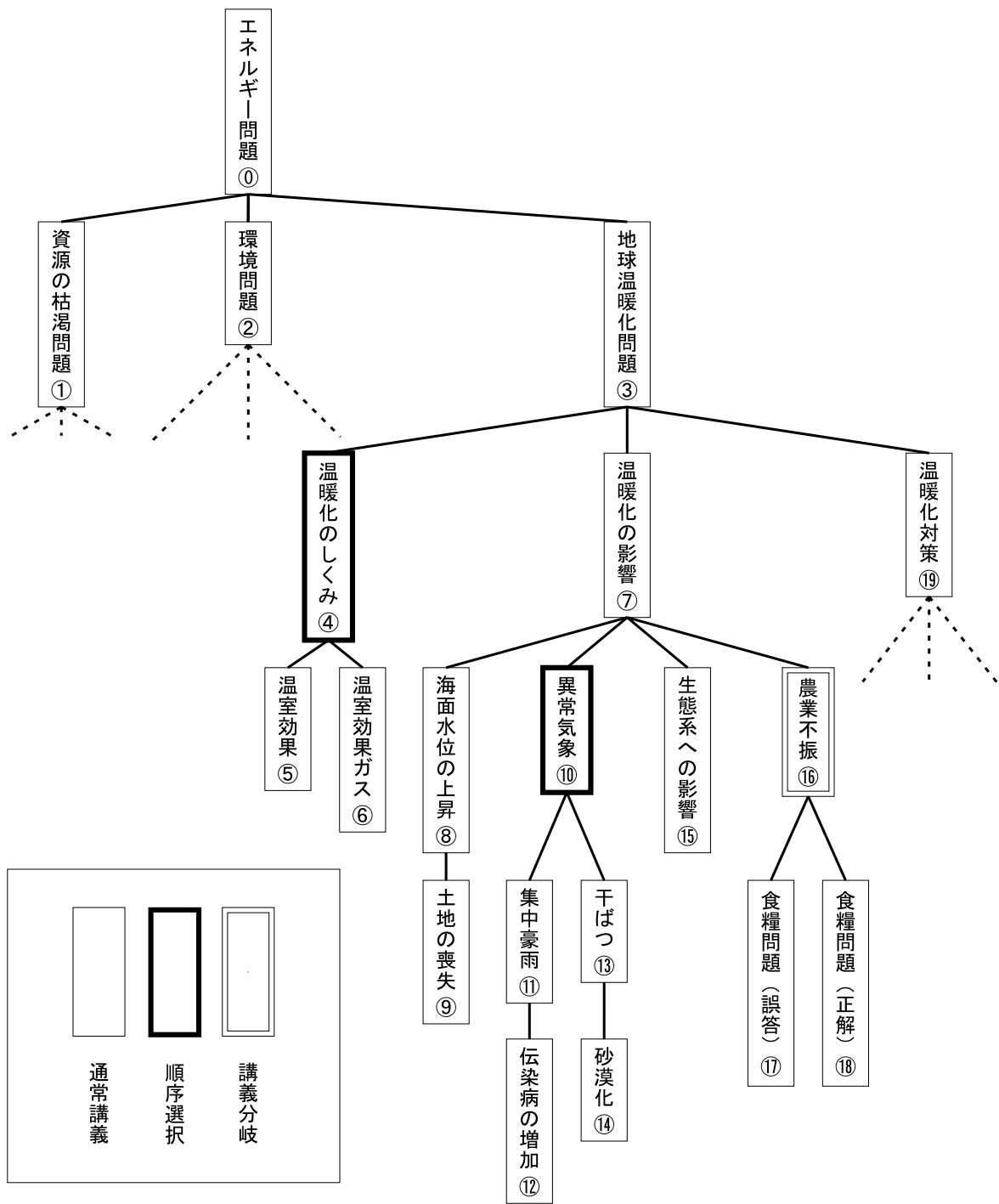


図 4.15: 講義の構造

る。また、線で直接結ばれた上位と下位の階層の講義フレームは親と子の関係にある。例を挙げると、「異常気象」が親で、「集中豪雨」「干ばつ」が子にあたる。この木構造に先順走査の順で根を始点とする通し番号を振ると、この通し番号が講義の進行順となる。図 4.15の丸で囲んだ数字が通し番号である。通常の講義進行ではこの通し番号の順に講義が進行する。一度行った講義フレームには説明済であることを示す既読のフラグが付けられる。また、分岐を行う講義を区別するために講義属性が付けられている。図 4.15の太線の講義が順序選択で二重線の講義が講義分岐である。その他の講義は通常講義である。

これらの情報を保持するために講義テーブルの仕様は表 4.3のようになっている。また、その例を表 4.10に示す。講義テーブルには項目として講義タイトル、講義の本文、HTML ファイル表示のための URL、講義番号、親の講義番号、既読・未読フラグ、講義の属性が記述してある。この中で、親の講義番号には木構造の中での親にあたる講義の番号が格納されていて、これを分岐の際に用いる。また、既読・未読フラグは講義進行の際に使用し、既読であればその講義は行わず次の講義へ進む。

分岐テーブルにはその問いかけの返事として期待される単語が全て記述してある。このテーブルはキーとしてキーワード(単語)、項目として分岐元の講義番号と分岐先の講義番号をもつ。この構造によって講義分岐の問いかけの際にどの講義の分岐において、学習者がどの単語を発話すればどの講義へ分岐するかを知ることができる。

質問回答データベースは名前変換テーブルと4つの文タイプによるテーブルからなるが、名前変換テーブルは、省略語、同義語による表現の揺れを吸収するためのテーブルである。自由な発話を解析する場合、同じものを指していても表現が違ふことが多くある。例えば「地球温暖化」を意味するのに単に「温暖化」と省略して表現したり、「化石燃料」と「化石資源」の2つの言葉で同じものを指したりする。この表現の揺れを吸収するためにデータベースの内部では、1つのものを指す場合にはある特定の1つの名前を使用する。このために質問回答で用いる同義語、類義語、省略語全てをそれぞれの単語のデータベース内での名前に変換する必要がある。名前変換テーブルにはキーに全ての単語をとり、項目としてデータベース内での名前を持たせる。これを各単語について検索し、データベース内での名前に変換することで単語表現の揺れを吸収する。

学習者の質問文は発話文解析を行って文型タイプとキーワードを抽出する。このキーワードを全て名前変換テーブルで変換し、その結果を用いて検索する。

意味テーブルはキーとしてキーワードをとり、項目には質問に対する回答文を持っ

ている。また、それ以外の原因、結果、対策テーブルはキーとしてターゲットとなるキーワードを持ち、項目にはそれぞれの動作(変化)における回答文をもつ。具体的に原因・理由テーブルに例をとってみると、各項目は、ターゲットが「増加」する原因、ターゲットが「減少」する原因、ターゲットが「起こる」原因などが記述されている。「温室効果ガスが増えるのはなぜですか」もしくは、「温室効果ガスが増える原因は何ですか」という質問文が発話された場合、発話文解析で抽出されたターゲットのキーワードを名前変換テーブルで変換すると「温室効果ガス」のデータベース内での名前は同じく「温室効果ガス」であり、これを検索に用いる。文型タイプは原因・理由であるので、原因データベースを検索する。回答文の検索の際のキーワードは「温室効果ガス」、検索する項目は「増加」する原因についての回答の項目である。この検索結果が「温室効果ガス」の「増加」の「原因」についての質問に対する回答である。

4.2.2 講義の進行、分岐の方法

講義の構造、通常進行

本システムの講義用教材はいくつもの講義フレームから構成される木構造を持ったものである。この様な構造を持った講義の進行をどのような方法で行うかについて以下で説明する。

まず、講義テキストを発話し、補助教材用 HTML ファイルを表示する。そして、既読フラグのついていない講義を通し番号の小さい順から進めることで、次に述べる講義の分岐があった場合でも問題なく進行ができる。

講義の分岐

この CAI システムはインタラクティブな学習を実現するために、学習者の応答に応じて次に行う講義を変更できる仕組みを持つ。これが順序選択と講義分岐である。この2つはほぼ同じ仕組みで実現されるため、同じ「分岐」として説明する。分岐とは、分岐を行う講義において学習者に対して問いかけを行い、それに対する学習者の回答によって次に行う講義を子の講義フレームの中から決めるものである。便宜上、この分岐を行う親の講義フレームを分岐元講義、子の講義フレームを分岐先講義と呼ぶ。通常の講義と区別するために分岐元講義にはそれを示すラベルを付与してある。この処理を図 4.16 に示し、順を追って説明する。

まず、講義を行うために教材データベースから講義のデータを読み込む。このとき

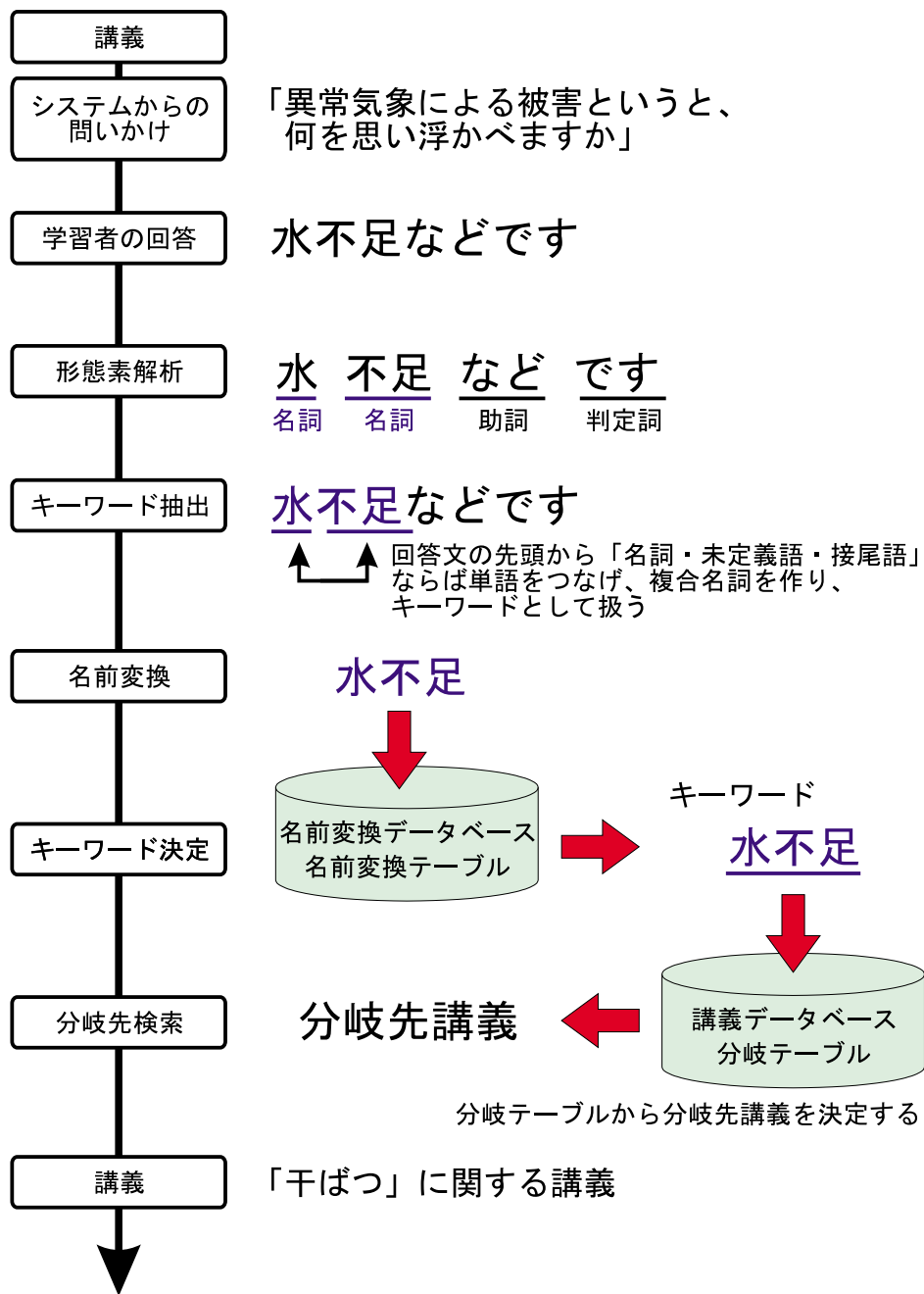


図 4.16: 講義の分岐の処理

講義に分岐を示すラベルがあった場合、その講義が分岐元講義であることをシステムが認識する。分岐元講義では講義文の最後に問いかけの文があり、学習者に対して何らかの回答を促す。この問いかけに対する学習者からの発話が行われると、その認識結果が文解析部で解析され、キーワードとして得られる。このキーワードによって分岐先講義が決定されるが、どのキーワードの場合にどの講義へと分岐するかの情報が分岐用データとしてデータベースの中にあり、これを参照することで分岐先講義が決定される。このあと、順序選択ではそのまま講義が進められ、他の分岐先の候補だった講義フレームも提示されるが、講義分岐では、他の分岐先の候補には既読フラグが付けられ、その講義が行われることはない。

4.2.3 質問文解析方法

構築した CAI システムにおいて対話的学習を実現するもう 1 つの機能が質問回答機能である。この機能を実現するために、学習者の質問文を解析し、回答文を検索する必要がある。ここではその方法と実際の処理について説明する。

図 4.10 で示し、4.2 節でも述べた通り、質問回答処理は音声認識、文字コード変換、発話タイプ解析、質問タイプ解析、質問文解析、質問回答文検索を順に行うことで実現する。ここでは、質問文解析において重要と思われる発話タイプ解析、質問タイプ解析、質問文意味解析、質問回答文検索の一連の処理について実際の解析例に基づいて説明する。図 4.17 に質問回答の処理を示す。

例えば学習者が「温暖化ガスとは何ですか」という質問文を発話した場合、音声認識処理によって文字列へ変換された質問文が入力される。まず、この文に対してパターンマッチングを行って発話タイプを分析する(図 4.10、表 4.1 参照)。この場合、質問文に「ですか」があり、質問であることが分かる。さらに質問タイプ解析のためのパターンマッチングを行って、質問タイプを解析する(表 4.2 参照)。この場合、質問文に「とは何」があるため、意味を問う質問であると解析される。この結果に対して質問文解析を行う。これは、質問文の中からキーワードを抽出することで何についての質問であるかを解析するものである。この処理は図 4.17 に示すような手順で行われる。まず、質問文を単語単位に区切り品詞を特定する形態素解析を行う。次に、キーワードを抽出する手がかりとなる語を見つける。この場合、助詞「と」と助詞「は」の連続である「とは」が手がかり語となる。次にキーワードを抽出する。意味についての質問の場合、手がかり語の直前の名詞をキーワードとして抽出する。このとき、名詞 1 つだけを取り出すのではなく、一連の名詞、接頭辞、接尾辞、未定義語を 1 つのキーワー

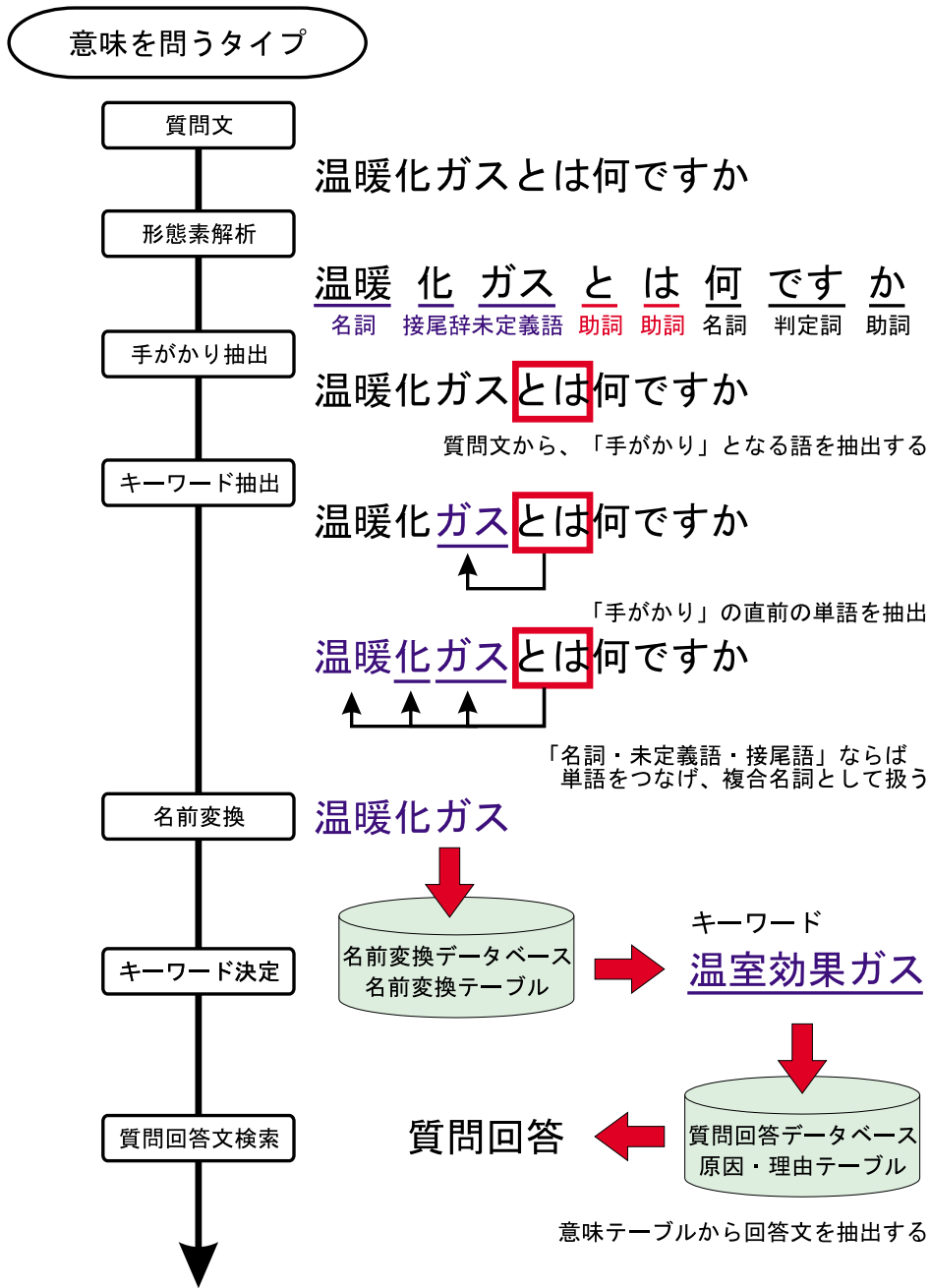


図 4.17: 質問文解析処理 (意味)

ドとして取り出す。この例では、名詞「温暖」、接尾辞「化」、名詞「ガス」がそれにあたり、キーワードとして「温暖化ガス」が抽出される。

このようにして質問文が解析され、質問タイプとキーワードが検出されると、データベースからそれに対する回答文を検索する。まず、回答文検索の前に名前変換を行う。この場合、「温暖化ガス」が「温室効果ガス」に変換される。この変換されたキーワードを用いてデータベースから回答を検索する。質問タイプが「意味」であるので、「意味」の質問回答データベースから「温室効果ガス」の回答文を検索する。この検索によって得られた回答の文を合成音声で読み上げることで、質問に答える。

もう1つ例を挙げて説明する。「なぜ地球温暖化が起きるのですか」という質問の処理を図 4.18に示す。まず、この場合、発話タイプ解析処理において「ですか」があることから、質問文であると解析される。次に質問タイプ解析処理で「なぜ」があることから、原因・理由を問う質問であると解析される。質問文解析処理で形態素解析を行い、手がかり語の抽出を行うと、助詞の「が」が手がかり語として抽出される。このような原因・理由を問う質問の場合、手がかり語の直前の「名詞」、「接頭辞」、「接尾辞」、「未定義語」の連続をキーワード1、手がかり語の直後の「動詞」、「形容詞」、「名詞」などをキーワード2とする。この例では、「地球温暖化」がキーワード1、「起きる」がキーワード2となる。この2つのキーワードを名前変換すると、それぞれ「地球温暖化」、「生起」となる。この2つのキーワードを用いて「原因・理由」の質問回答データベースから回答文を検索する。この例での実際の検索はキーワード1を検索用単語とし、キーワード2の項目の回答文を検索する。この処理によって得られた回答文を読み上げることで質問に回答する。

4.2.4 複数学習の実現方法

複数の学習者による学習は、図 4.19に示すような構成で実現する。

複数学習者で学習する場合には学習者一人に一台のクライアントコンピュータをそれぞれ用いて学習する。サーバコンピュータはそれぞれのコンピュータと一対一のソケット通信を行う。サーバの受信側には、クライアントの数だけ文解析部があり、これがそれぞれのクライアントの送信側通信路とつながっている。また、サーバの送信側は1つの対話管理部から全てのクライアントコンピュータに対する通信路が設定されている。サーバコンピュータから全てのクライアントコンピュータへ送られるデータは、同じ内容で、一人での学習の場合と同じく発話用テキストと、補助資料提示用の URL アドレスなどである。これとは別に、複数人学習の場合のみ、他の学習者の発話テキ

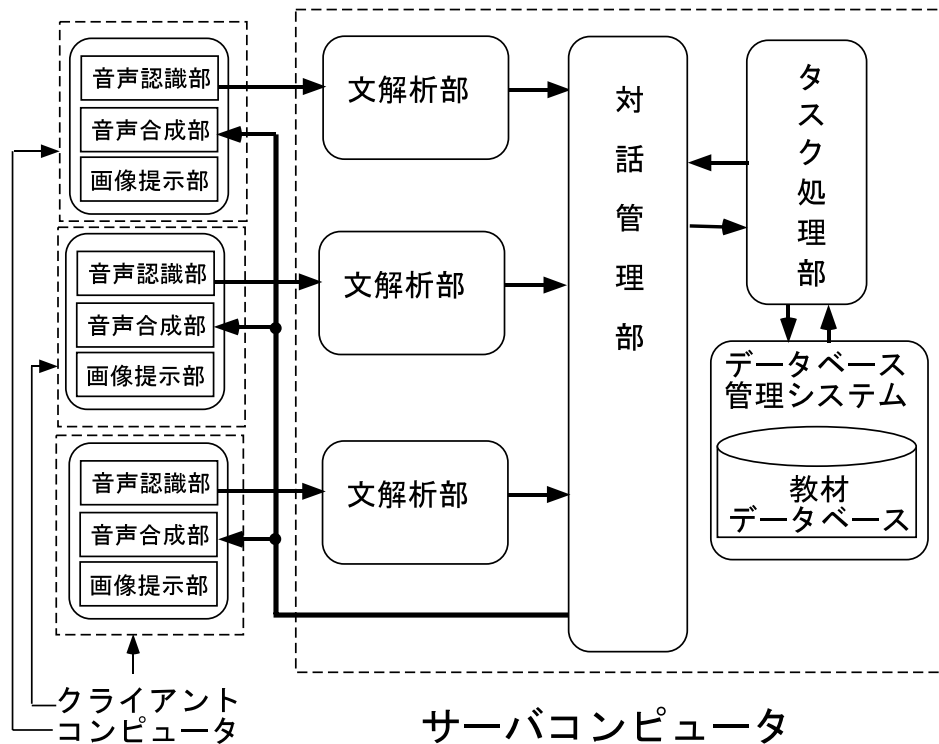


図 4.19: 複数学習者での学習環境を構築するときのシステム構成

ストも送られる。また、各クライアントコンピュータからサーバへ送られるデータも一人での学習の場合と同様に学習者の発話と各種信号である。このような構成にすることで、それぞれの学習者からの発話は同時に処理され、サーバにおいて発話の意味解析ができた学習者の発話から順に処理していく。文解析部で発話の解析ができると、それが対話管理部に送られ、その学習者の発話に対する反応が全ての学習者に発話される。また、同時に発話解析に成功した学習者の発話も全ての学習者のコンピュータディスプレイ上に表示される。このようにして同期分散型学習環境を実現する。

4.2.5 構築したシステムに用いたソフトウェア

本項では、CAIシステムを構成するソフトウェアについて述べる。本研究で構築したCAIシステムは以下に挙げるソフトウェアから構成される。

1. 音声認識エンジン IBM 社製 Via Voice Millenium (クライアントコンピュータ)
2. 音声合成エンジン 東芝社製 東芝音声システム (クライアントコンピュータ)
3. 本研究で作成したインタフェースプログラム (クライアントコンピュータ)

4. 形態解析システム JUMAN (サーバコンピュータ)
5. リレーショナル型データベースシステム PostgreSQL (サーバコンピュータ)
6. 本研究で作成したサーバプログラム (サーバコンピュータ)

これらのソフトウェアはクライアントコンピュータ上で動作するものと、サーバコンピュータ上で動作するものがある。以下でこれらのソフトウェアについて、図 4.3 のソフトウェア構成と対応させて説明する。

クライアントコンピュータ上で動作するソフトウェア

クライアントコンピュータ上では、音声認識部、音声合成部、画像提示部が動作している。この中で、音声認識部の音声認識機能を実現するために IBM 社製 ViaVoice Millenium Pro の音声認識エンジンを、また、音声合成部の音声合成機能を実現するために東芝社製 東芝音声システム Ver.4 の音声合成エンジンを用いた。

本研究で構築したインタフェースプログラムはこれらの音声認識エンジン、音声合成エンジンを制御し、それ以外の機能を実現する。以下でそれぞれのソフトウェアの特徴について説明する。

- **ViaVoice Millenium Pro**

ViaVoice は IBM 社製の音声認識ソフトウェアで不特定話者に対応した連続音声認識システムである。隠れマルコフモデル (HMM) をベースにし、統計的言語モデルを用いた音声認識を行う。ソフトウェアの処理のみで音声認識を実現しており、専用ハードウェアを必要としない。ViaVoice では約 80,000 語の単語を持つ基本辞書を備え、最適な条件下で 90%以上の認識率を実現している。さらに開発用に強力な API を備えており、容易に音声認識エンジンを利用したアプリケーションを構築できる。

- **東芝音声システム Ver.4**

東芝音声システムは音声認識、音声合成を行うエンジンとそれらを利用するアプリケーションのパッケージ群で構成されるが、本システムにおいては、音声合成エンジンのみを用いた。この音声合成エンジンはテキスト文字列で入力された文を解析し、自然な発音、イントネーションになるようにして合成音声として出力するソフトウェアである。

- 本研究で作成したインタフェースプログラム

インタフェースプログラムは音声認識システムと音声合成システムの制御を行う。また、画像提示部として WWW ブラウザの機能を備えている。このインタフェースプログラムはサーバプログラムと TCP/IP ソケットによる通信でデータの受渡しを行うことによって、学習を実現する。

このプログラムは Microsoft 社製 VisualC++ によって構築した。

サーバコンピュータ上で動作するソフトウェア

サーバコンピュータ上では文解析部、対話管理部、タスク処理部が動作している。この文解析部の形態素解析処理を実現するために形態素解析システム JUMAN を用いた。また、データベース管理システムとしてリレーショナルデータベース管理システム PostgreSQL を用いた。そして本研究で作成したサーバプログラムでこれらのシステムの制御やインタフェースプログラムとの通信、その他の機能を実現している。以下でそれぞれのソフトウェアの特徴について説明する。

- 形態素解析システム JUMAN

JUMAN は京都大学長尾研究室で開発された日本語形態素解析システムで計算機による日本語の解析の研究を目指す多くの研究者に共通に使える形態素解析ツールを提供するために開発されたものである^[24]。JUMAN を用いると、テキスト文字列の文を単語毎に区切り、単語の基本形、品詞、詳しい品詞を知ることができる。

- リレーショナルデータベース管理システム PostgreSQL

PostgreSQL は、Unix 系のオペレーティングシステム上で動作するリレーショナル型データベース管理システムである。無料で利用できるソフトウェアであり、市販のデータベースソフトのような高度なアプリケーション構築機能などは持たないものの、リレーショナル型データベース言語である SQL が使え、数十万件のデータを実用的なスピードで検索できるなど、データベースの基本機能としては十分なものを備えている。PostgreSQL はもともとはカリフォルニア大学バークレー校で開発され、その後ボランティアの手によって維持され、発展して現在に至っている。

- 本研究で作成したサーバプログラム

本研究で作成したサーバプログラムは、サーバコンピュータの受け持つ処理の中で、データベース管理と形態素解析以外の全ての処理を行う。このプログラムは複数のインタフェースプログラムと TCP/IP ソケット通信を行うことができ、これによって複数人での学習を実現する。このプログラムは C 言語で開発した。

4.3 システムの構築と実験に用いたハードウェアの詳細

4.3.1 インタフェースを受け持つパーソナルコンピュータ

システムを構築するためにパーソナルコンピュータに要求される条件

システムのクライアントコンピュータでは、音声認識エンジンと音声合成エンジンを用いて音声認識と音声合成とをリアルタイムで行う必要がある。また、クライアントコンピュータは学習者がその前に座って学習を行うためのもので、できるだけ安価で入手しやすい方が良い。音声認識エンジン、音声合成エンジンは市販のものを用いるため、プラットフォームは市販の実時間動作可能な音声認識エンジン、音声合成エンジンが動作する環境でなければならない。これらの条件を満たすためにシステムを構成するパーソナルコンピュータとしてオペレーティングシステム (Operating System:OS) に Windows98 を採用した PC/AT 互換コンピュータを用いた。

このプラットフォームはコストパフォーマンスが非常に優れており、CPU として Intel 社製 Pentium3 450MHz クラスのものを採用したシステムが 15 万円以下で入手できる。このクラスのもので十分な処理速度が得られ、動作する音声認識エンジンも 1~2 万円の安価なものが何種類も販売されている。

この他、音声認識と音声合成のための音声入出力機能とサーバコンピュータに接続するためのネットワーク機能も必要である。音声認識と音声合成を同時に行うため音声入出力機能には入力と出力を同時に扱うことが可能なものが要求される。このために入力と出力それぞれに別の処理機能を備えた Full Duplex の音声入出力サブボードを装備させる。また、ネットワークとしてはイーサネットを用いる。

システム構築、実験に用いたパーソナルコンピュータの仕様

実際のシステム構築と実験に用いたパーソナルコンピュータは EpsonDirect 社製 PC/AT 互換コンピュータである。その仕様を表 4.17 に示す。

表 4.17: システムの構築と実験に用いたパーソナルコンピュータの仕様

		備考
CPU	Intel 社製 Pentium III 450MHz	
Operatig System	Microsoft 社製 Windows98	
System Chipset	Intel 社製 440BX	
メモリ	256MB	
ネットワーク	100BASE-TX	100Mbit/sec
音声入出力	Creative 社製 Ensoniq Audio PCI64	Full Duplex

4.3.2 サーバコンピュータ

システムを構築するためにサーバコンピュータに要求される条件

サーバコンピュータは文解析、対話管理、タスク処理などを行い学習、対話の全てを司る役割を担う。従って、これには安定して動作し、しかもなるべくコストの掛からないものが望ましい。

本研究で構築する CAI システムでは、文解析の形態素解析処理に形態素解析システムを利用する。また、教材データベースを扱うためにリレーショナルデータベースシステムを利用する。このため、形態素解析システムとリレーショナルデータベースシステムが動作する環境が必要となる。これらの条件を満たすものとして OS に Linux を採用した PC/AT 互換コンピュータを用いた。

このコンピュータはパーソナルコンピュータと同じアーキテクチャのコンピュータでコストも同様に安価であるが、オペレーティングシステムとして UNIX 系 OS である Linux を採用することで、安定した動作を実現できる。また、UNIX 用の様々なアプリケーションプログラムをほとんど変更無しに移植できる利点がある。

システム構築、実験に用いたサーバコンピュータの仕様

実際のシステム構築、実験に用いたサーバコンピュータは PC/AT 互換コンピュータである。その仕様を表 4.18 に示す。

表 4.18: システムの構築と実験に用いたサーバコンピュータの仕様

		備考
CPU	Intel 社製 Pentium III 500MHz × 2	
Operatig System	Linux	
System Chipset	Intel 社製 440BX	
メモリ	256MB	
ネットワーク	100BASE-TX	100Mbit/sec

第 5 章 システムの動作確認と評価実験

5.1 構築したシステムの動作確認

本研究で実際に構築したシステムが設計した通りに動作するかを確認するために以下で述べるような教材を作成し、動作確認を行った。この教材の講義は図 4.15に示す構成となっている。教材はエネルギー問題のうち、地球温暖化を中心とした 18 個の講義フレームから構成される。また、質問回答用のデータは、意味、原因・理由、結果、対応・対策の 4 つの質問タイプに対応し、意味については 1 項目、その他の 3 タイプについてはプラス要素、マイナス要素、生起、単独の 4 項目から構成される。キーワードの単語数は 50 語で、それぞれについて質問に対する回答文を延べ 13 項目分用意した。したがって、質問回答データベースは合計 650 項目のデータから構成される。

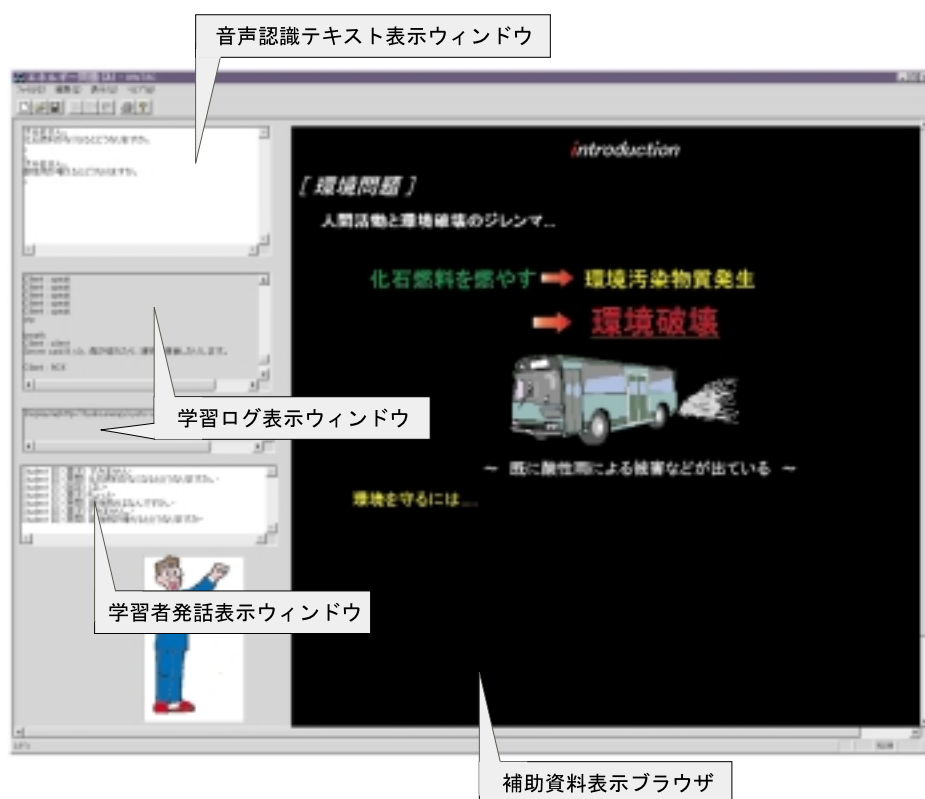


図 5.1: 学習者に提示される画面例

このような教材を作成し、教材作成が容易であることを確認した。教材作成はデー

データベースのテーブルの中の欄に文章を入力するだけの作業であった。また、何回かの動作確認の間に教材の変更を行い、教材の更新が容易であり、その結果がすぐ学習に反映されること、また、教材の変更を行っても正常に動作することを確認した。

実際に学習者に提示する画面例を図 5.1 に示す。学習者はこの画面を見ながら音声で学習を行う。このインタフェース画面には講義用の補助資料を提示する補助資料表示ブラウザがあり、学習者はこれを参考にして講義を進める。また、音声認識の結果を表示する音声認識テキスト表示ウィンドウと学習ログ表示ウィンドウがあり、現在行っている学習の全ての発話を記録している。また、学習者発話表示ウィンドウにはシステムが認識した発話が表示され、どの学習者が何を言ったのかが分かるようになっている。

動作確認は、筆者を含むシステム製作者 2 名が行い、システムの全ての機能について、それぞれの場面で意図した通りの動作を行うかについて何回か試行する形式で行った。この動作確認において、システムは学習の始めから終わりまで意図した通りの動作をし、音声対話によって学習を進行できることを確認した。

以下ではそれぞれの機能ごとに動作確認をした結果を述べる。

5.1.1 音声対話機能

講義は、はっきりした合成音声により正常に行われた。講義中に何か発話すると講義が中断され、「すみません」と発話すると「はい、何ですか」と聞き返した。また、この後、続けて質問、要求を行うことができ、システムはその質問、要求に適切に答えた。このとき発話の分類として想定していない発話を行ったところ「もう一度言って下さい」と聞き返してきた。

5.1.2 講義機能

講義は設計した順序で進行した。また、1つの講義フレームが終るごとに「次に進んでもいいですか」とシステムから確認があり、「はい」と学習者が音声で答えることで次へ進んだ。さらに、順序選択の講義では問いかけがあり、それに答えることで講義の順序が意図した順番に変更された。作成した教材の講義分岐は問題形式であった。この講義では、問題形式の講義が読み上げられ、それに対する回答を学習者が行うとその正誤に応じて次に行われる講義が分岐した。講義中に「もう一度はじめから」と発話するとその講義フレームのはじめの文から再度講義が始まった。また、「もう一度

言って下さい」と発話したところシステムは発話中もしくは発話直後の文をもう一度読み直した。

5.1.3 質問回答機能

学習中にいくつかの質問をしたところ、質問回答用のデータに回答が存在するものについては正しく答えた。また、データにない質問には「すみません、分かりません」と答えた。

5.1.4 複数学習機能

2台のクライアントコンピュータを用いて1つのサーバコンピュータに接続して2人で学習を行うことができた。この動作確認で2人の学習者で学習のはじめから終わりまで進行できることを確認した。また、システムは質問などの学習者の発話を正しく理解し、それに対して設計で意図した通りの適切な反応を返した。質問を2人同時に行った場合、早く音声認識が完了した方の質問に答えた。その際、もう一方の学習者の画面に質問文が表示されることを確認した。

5.1.5 動作確認のまとめ

構築したシステムにおいて基本的な動作確認を行い、以下のことを確認した。

- 教材の作成、更新が容易であること。
- それぞれの対話機能、学習機能が意図した通りに動作すること。
- 同時に複数人の学習者によるグループ学習が行えること。

5.2 評価実験

構築した CAI システムが実際の学習支援システムとして使われた場合の学習機能、対話機能の性能を評価するために評価実験を行った。以下でその実験について述べる。

5.2.1 評価の目的

この実験の目的は、本研究で構築した CAI システムを実際の教育に用いた場合に、対話機能、講義機能、質問回答機能がどのように働くかを調べ、また、どのような問

題点があるかを明らかにすることである。さらに、学習者の関心や興味は維持されたかどうか、対話が成り立つかどうか、インタフェースとしての扱いやすさはどうかなど、総合的にどの程度教育システムとして機能するかを調べるための実験でもある。

5.2.2 実験方法

実験では、実際に被験者に一人で本システムを使って学習してもらう。このときの観察結果と学習の後に実施するアンケートの主観的評価の2つから評価を行う。

本実験は、この音声対話 CAI システムを用いた実際の学習において学習者はどのような発話をし、それに対してシステムはどのように動作するのか、また、どの程度音声対話による学習を実現できるのかについて調べるための実験である。従って、学習者には、学習を進行するために必要な最小限のことを説明するにとどめ、どのようなことを言えば答えるかなどの具体的なことは告げないこととした。学習者に告げたのは以下の2点のみである。

- 学習中にシステムから問いかけがあれば学習者はそれに答える必要がある。
- 学習中にシステムに対して自由に質問してよい。

また、上記のような実験の目的より、被験者は本システムの機能を全く知らない者とした。

本 CAI システムのこの特徴の有効性を調べるという観点から評価の対象となる項目を以下の4項目とした。

1. この CAI システムを用いた学習において、関心、興味は高まったか。
2. 音声による質問回答はうまく機能しているか。
3. 音声による対話がどの程度実現できたか。
4. 音声対話型インタフェースの意義はあるか。

この4つの観点をもとに4段階の評価と各項目の自由記述からなるアンケートを作成し、評価に用いた。なお、4段階評価は、4が最も評価が高く、1が最も低い。

実際の実験は、5.1節の動作確認で用いたシステム、教材を使用して以下の手順で行った。まず、音声認識の精度を向上させるために、予め被験者の音声の特徴を音声認識システムに登録する作業を行った。同時に、音声認識システムに元々登録されていない単

語、認識されにくい単語を登録した。その後、主観的評価の観点を被験者に理解してもらうために、事後に答えるアンケートを被験者に読んでもらった。そのとき、学習中にシステムから問いかけがあること、学習において分からないところはいつでも自由にシステムに対して質問してよいことを告げた。その後、被験者に本 CAI システムを用いて、はじめから最後まで通して学習してもらった。そして、その様子をデジタルビデオに記録した。最後にアンケートに答えてもらい主観的評価を行ってもらった。

5.2.3 実験結果

本研究室の大学生と本専攻の大学院生計 3 人に対して実験を行った。その結果のうち、4 段階の評価を 0~3 の数値にしたものと自由記述の結果について表 5.1~5.4 に示す。

また、この実験において学習者の発話について分析した結果を表 5.5~5.7 に示す。全試行回数は学習中に学習者が発話した回数であり、その試行のうち音声認識で一文字でも誤認識されたものが音声認識の失敗回数、正しく認識された回数が音声認識の成功回数である。また、成功回数のうち、システムが設計とは違う動作をした回数がシステム誤動作回数である。また、質問に関してはシステムが適切な回答を答えた回数も示す。

この結果より発話が正しく認識された割合 (全試行回数/音声認識成功回数) は、それぞれ被験者ごとに 79.4%(27/34)、78.3%(18/23)、65.5%(19/29) であった。また、学習者の自由な質問に対して答えられた数は、それぞれ被験者ごとに 15 回中 3 回 (20%)、4 回中 1 回 (25%)、9 回中 0 回 (0%) であった。システムが 1 問も答えられなかった被験者は、音声認識の精度がやや悪かった。システムの誤動作は、システムからの講義進行の確認の問いに学習者が「はい」と答えても次へ進行しないこと、何も発話していないのに「はい、何でしょう」と聞き返しをすることであった。

次項で以上で述べた結果について考察する。

5.2.4 考察

アンケートの数値評価は 0~3 点の数値で表されており、0 点台が「問題あり」、1 点台が「問題はあるが評価できる」、2 点台が「評価できる」と解釈できる。この結果より、「再度学習したいと思うか」が 0.3 点であり、大きな問題があるが、自由記述の結果をみると、被験者は既にエネルギー問題についてかなり詳しい知識を持っており、再度学習する必要性や意欲を感じなかったためであることが最大要因であると分かる。そ

表 5.1: 評価実験の結果 (1)

評価項目 (評価点の平均)	評価点	評価の理由
A1 学習に対する興味の程度 (1.3)	1	・よく取り上げられるテーマだったから。
	2	・元々興味ある内容だったから。
	1	・内容が当然のことすぎてほとんどすべてのことが既知のことだった。
A2 学習中退屈な気分の発生 (1.3)	1	・質問しても答えてもらえる事が少ないので、質問しようと思わなくなる。
	2	・音声認識自体が新鮮だったから。
	1	・上記理由により、内容に興味をもてない上、興味のあることに対する質問にも答えてもらえなかったため。
A3 再度学習することへの意欲 (0.3)	0	・反応が悪い。
	1	・質問にもっと回答してもらえたら。
	0	・上記 2 の理由により。
B1 システムの質問の認識 (1.0)	1	・1 つの形式の質問しか認識されなかった。
	1	・半分以上が認識されなかった。
	1	
B2 システムの質問への回答 (0.7)	1	・「..... とは何ですか?」という形式の質問しか回答されなかった。
	1	・半分以上が回答なし。
	0	・一度も答えてもらえなかった。
B3 システムからの質問回答の適切さ (2.0)	3	・適切だった。
	3	・認識された質問については、適切な回答を受けた。
	0	・回答してもらえなかつた。

表 5.2: 評価実験の結果 (2)

評価項目 (評価点の平均)	評価点	評価の理由
C1 システムからの 適切な応答 (0.7)	1	・途中、「はい」の認識が悪くなった。
	1	・半分位しか適切に応答されなかった。
	0	・「進んでもいいですか」「はい」 しか会話がなりたたなかった。
C2 システムからの 不適切な応答はないか (0.7)	1	
	1	・「はい」が伝わらないことがあった。
	0	・私の発話に対する適切な対応は下記の通り。 先に進むことと、「ちょっといいですか？」 といったのに対して、「はい、なんでしょう？」 とききかえしてくれるの2通りのみであった。

れ以外の理由としては、「システムが質問に答えない」ことも挙げられている。それを示すように「質問に答えたか」の項目で 0.7 点と評価が低い。しかし、昨年システムでは、自由な質問に対して 1 度も答えられなかったのに比べ本研究のシステムでは 4 つの質問について答えられている。この点では評価できるが、まだまだ回答できる確率が低く改良が必要であると考えられる。また、適切な対話ができているかという観点の評価も 0.7 点で低い。このことから本システムで実現した対話の機能だけでは、人間の学習者が求めるレベルにまで達していないといえる。

反対に、「質問に対する回答の適切さ」と「合成音声の講義の聞き取りやすさ」はそれぞれ 2.0 点、2.3 点と評価が高く、本システムの質問回答機能と音声による講義機能が評価されている。

学習者の発話の分析結果から、以下のことがいえる。音声認識の成功率がまだ低く(平均 74.4%)、それが学習の円滑な進行を妨げていると考えられる。特に質問回答に関しては音声認識の失敗が多く、音声認識の認識率が改善すればより多くの質問に答えられると考えられる。また、それによって評価実験においてもさらに高い評価を得られると期待できる。

また、実験の観察結果から、以下のことがいえる。

- 回答がデータベースにあり音声認識で正しく認識された質問には、全て答えられている。

表 5.3: 評価実験の結果 (3)

評価項目 (評価点の平均)	評価点	評価の理由
D1 音声講義の聞き取りやすさ (2.3)	3	・聞きやすかったが人間らしくはなかった。
	2	・一部聞き取りにくかったが、聞き漏れはない。
	2	・一度聞き取れない部分があったが、他はききとれた。
D2 音声講義の人間らしさ (1.0)	1	・聞きやすかったが人間らしくはなかった。
	1	・音声の強弱、高低が不自然なことが、しばしばあった。
	1	・比較的人間らしいとは思いますが、非人間的な部分が少しあるため、人間と同様には聞こえなかった。
D3 文字を読む形式と比較して 音声講義のよさ (1.7)	1	・欲しい時に欲しい情報が得られないので、ストレスになる事もある。
	2	・「文字を読む」より楽だ。ただし、少しあるため、図 (OHP?) などが必要。
	2	・聞きのがしたところをもう一度聞くことができなかつたため、これよりは自由に進んだりもどったりできる表示方式にもメリットはあると思うが、動画をみながら音声で講義してくれるのはテレビ授業みたいでわかりやすかった。
D4 音声入力の効果 (1.3)	1	・認識率次第だと思った。
	3	・うまく認識さえしてくれれば、非常に楽だ。
	0	・現在の認識率では、ストレスがたまるので良くない。ただ、学習時にキーボードを使うのもわずらわしいためほぼ 100 % 認識率が期待できるのであれば、大変よい方法になると思う。

表 5.4: 評価実験の結果 (4)

評価項目	評価
E1 改善点	・対応できる質問を増やす。講義の流れがわかりにくいので、流れを生徒に提示するようにする。
	・質問・回答プロセスの成功率を高めることを望みます。
	・発話および質問に対する認識率・回答率の向上、および安定性向上。
E2 自由意見	・左に表示される文章を読んでしまって、講義を聞かなくなってしまうので、消した方がいいかも。
	・既知の内容だったので、すんなり頭の中に入りましたが、未知のことを学ぶとすれば少々講義が速かった気がします。「もう一度説明して下さい。」 と言え、回答してくれたのでしょうか？
	・質問者は考えがまとまっていなような状態で質問する場合もある。 その時、「何をいっているのかわからない。」と言うのではなく、考えがまとまるように、アドバイスをすべきではないか。

表 5.5: 実験での学習者の発話 (被験者 1)

	全試行回数	音声認識の失敗回数	音声認識の成功回数	システムの誤動作回数	学習者からの質問への適切な回答回数
講義進行要求の「はい」	15	0	15	3	-
講義進行要求の「もう一度」	1	0	1	0	-
質問	15	6	9	0	3
分岐	3	1	2	0	-
講義中断	20	-	-	3	-

表 5.6: 実験での学習者の発話 (被験者 2)

	全試行回数	音声認識の 失敗回数	音声認識の 成功回数	システムの 誤動作回数	学習者からの 質問への適切 な回答回数
講義進行要求の 「はい」	15	0	15	3	-
講義進行要求の 「もう一度」	0	0	0	0	-
質問	5	4	1	0	1
分岐	3	1	2	0	-
講義中断	7	-	-	4	-

表 5.7: 実験での学習者の発話 (被験者 3)

	全試行回数	音声認識の 失敗回数	音声認識の 成功回数	システムの 誤動作回数	学習者からの 質問への適切 な回答回数
講義進行要求の 「はい」	13	2	11	2	-
講義進行要求の 「もう一度」	1	0	1	0	-
質問	12	7	5	0	0
分岐	3	1	2	0	-
講義中断	14	-	-	2	-

- 自由な質問では、主語や目的語などが省略された質問をされることが多く、現在の講義の文脈情報を持たなければ解釈できない質問が多い。
- 本システムの質問文解析の手法では比較的単純な質問しか解析できず、複雑な構造を持った文には対応できない。
- 本システムのような単純な規則による会話の「間」のとり方では人間の通常の会話の感覚と合わないために学習者が発話しようとする時などに戸惑うことがある。

5.2.5 システムの問題点

前述の考察より明らかになった本システムの問題点は以下の通りである。それぞれについて検討する。

- 質問回答において、文脈を考慮した解析ができない。
- 質問回答において、複雑な文には対応できない。
- 複雑な対話の実現できない。
- 会話の間のとり方が不自然である。

まず、文脈を考慮した質問文の解析ができない問題であるが、実験において学習者は質問をする際に、現在の文脈で明らかなのは省略して発話することが多かった。実際の例を挙げると、「砂漠化」の講義をしている時に「日本でも起こりますか。」のような質問があった。この質問文には主語がないが、明らかに「砂漠化」が主語であると分かる。本システムの解析手法ではこのような質問文の解析はできない。

次に、複雑な質問文に対応できない問題であるが、本システムでは、単純な質問の形式を想定してパターンマッチングやキーワード抽出を行っている。そのため、複文や重文などの複雑な構造を持った文を解析できない。

効果的な対話が行われない問題については、本システムでは、1つの発話に対して1つの決まった反応を返すような単純な対話のやりとりのみを機能として取り扱っている。実際の人間の対話ではその場面に応じて柔軟に反応を返してうまく情報を引き出そうとする。このような機能があれば、より使いやすい音声対話 CAI システムを実現できる。例を挙げると、学習者からの質問文が何について聞いているのか分からなかった場合に、「それは何について聞いているのですか」と焦点を明確にする聞き返しをす

ることによってその情報を聞き出し、学習者の要求に答えようとするなどである。本システムではこのような複雑な対話は実現できていない。

会話の「間」に関しては、音声認識のタイムラグがあり、実際の人間と同じような「間」で会話を行うような機能は実現しにくい。そのため、本研究で構築したシステムでは、この点を考慮していない。この問題は、あいづち、うなずきなどを用いたり、実際の人間の会話において「間」がどのように使われているかについて詳細に検討する必要がある。

5.2.6 実験のまとめ

本研究で構築したシステムの各機能の効果を調べるために評価実験を行った。その結果、音声による講義、質問回答の機能については一応の評価がなされたが、実際の学習においてはさらに高度な機能が要求されることが分かった。また、実験結果により今後の課題となる問題点を明らかにした。

第 6 章 結論

本研究では一般の人々を対象としたエネルギー問題を教育するための CAI システムを実現することを目指して、その設計、構築、評価を行った。この CAI システムでは一般の人々を対象としており、その中にはコンピュータ操作に慣れていない人々も含まれる。そのため、このような人にも使いやすいインタフェースとして、音声による対話を用いたインタフェースを持たせることを提案した。また、同期分散型学習環境としての機能を持たせて複数人での学習を実現し、学習効果を高めることを目指した。

このような CAI システムを実現するための機能について検討し、以下の機能、構成を持たせることとした。

- 音声による対話で学習を実現する。
- 講義形式で学習を行う。
- 学習者からの学習内容に関する質問を受け付け、それに適切な回答をする。
- 複数の学習者が同時に学習を行うことも可能である。
- ネットワークを用いたクライアントサーバ方式によってシステムを実現する。
- サーバによって教材データを一元管理する。
- パーソナルコンピュータを用いて安価なシステム構成を実現する。

そして、このような要求をみたす CAI システムを実際に構築した。

構築したシステムが意図した通りに動作するかどうかを調べるために製作者による各機能の動作確認を行い、基本的な規定の機能が正常に動作することを確認した。これより、教材の作成、更新が比較的容易なこと、入手しやすい安価なシステム構成で実際に動作することを確認した。また、CAI としての有効性を確かめ、問題点を明らかにするために製作者以外の研究室学生に参加してもらって本 CAI システムの評価実験を行った。その結果、音声による講義、質問回答の機能自体はおおむね評価された。しかし、以下のような問題点が明らかになった。

- 質問回答において、文脈を考慮した解析ができない。

- 質問回答において、複雑な文には対応できない。
- 複雑な対話の実現できない。
- 会話の間のとり方が馴染みにくい。

今後の課題としては5.2.5項の問題点での考察より、文脈を考慮した質問文解析機能、複雑な文でも対応できる質問文解析機能、高度な対話を実現する機能、会話の「間」を考慮した対話機能について改良を加え、実装することでさらに使いやすいシステムにすることが挙げられる。

謝 辞

本研究を進めるにあたり、ご多忙にも関わらず、研究の方向性から論文の書き方に至るまで、研究全般に渡って暖かく御指導下さいました吉川榮和教授に深く感謝致します。

研究を進める上で常に適切な御指導、御助言を下さいました下田宏助教授に心より感謝致します。

研究の方向性について多くの御助言を下さいました本研究室博士課程学生の WuWei さんに感謝致します。

実験などいろいろな面でこの研究を手伝って下さった修士課程学生の伊藤京子さんに感謝します。

教材作成からプログラミングまで大変な作業を引き受け、最後まで研究を支えてくれた修士課程学生の笹井寿郎君に心から感謝します。

研究を進める上で様々な御助言を下さいました博士課程学生の石井裕剛さん、小澤尚久さん、大林史明さんに感謝します。

最後になりましたが、いつも暖かく見守ってくれた谷友美秘書、藤岡美紀秘書をはじめ全ての研究室の方々に感謝します。

参考文献

- [1] 伊藤京子：会話型エージェントによる学習支援システムの研究, 京都大学工学部電気電子工学科学士論文 (1999).
- [2] 伊藤京子、山本専、Wu Wei、下田宏、吉川榮和：会話型エージェントによる学習支援システムの研究, ヒューマンインタフェース学会研究会報告集, Vol.1, No.3, pp.1-6 (1999).
- [3] 芦葉浪久：CAI コースウェア作成技法, 東京書籍, pp.42-43 (1989).
- [4] 菅井勝雄：CAI への招待 [理論編:教育工学のパラダイム変換], 同文書院, pp.95-122 (1989).
- [5] 清水康敬, 赤堀侃司, 市川伸一, 中山実, 伊藤紘二, 永岡慶三, 岡本敏雄, 吉崎静夫, 近藤勲, 永野和男, 菅井勝雄：教育工学の現状と今後の展開, 日本教育工学会雑誌/日本教育工学雑誌, Vol.22, No.4, pp.201-213 (1999).
- [6] Carbonell, J.R. : “AI in CAI”: An Artificial Intelligence Approach to Computer-Assisted Instruction , IEEE trans. Man-Machine Systems, Vol.11, No.4, December, pp.190-202 (1970).
- [7] Carbonell, J.R. : Mixed-initiative Man-Computer Instruction Dialogues, Massachusetts Institute of Technology Doctoral dissertation (1970).
- [8] Etienne Wenger 著, 岡本敏雄, 溝口理一郎 監訳：知的 CAI システム - 知識の相互伝達への認知科学的アプローチ -, オーム社, pp.31-41 (1990).
- [9] Brown, J.S.;Burton, R.R.; and Bell, A.G. : SOPHIE: a step towards a reactive learning environment, Int Jrnl Man-Machine Studies, Vol.7,pp.675-696 (1975).
- [10] Etienne Wenger 著, 岡本敏雄, 溝口理一郎 監訳：知的 CAI システム - 知識の相互伝達への認知科学的アプローチ -, オーム社, pp.53-79 (1990).

- [11] Hollan, J.D.; Hutchins, E.L.and Weitzman, L. : STEAMER: an interactive inspectable simulation-based training system, AI Magazine, Vol.5, No.2, pp.15-27 (1984).
- [12] Etienne Wenger 著, 岡本敏雄, 溝口理一郎 監訳 : 知的 CAI システム - 知識の相互伝達への認知科学的アプローチ - , オーム社, pp.83-92 (1990).
- [13] 大槻説平 : 発見的学習とその支援環境, 人工知能学会誌, Vol.8, No.4, pp.411-418 (1993).
- [14] 石村英樹, 高橋信, 吉川榮和 : ヒューマンインタフェースの大学院教育に関する非同期分散型教育支援環境の構築, 第 12 回ヒューマン・インターフェース・シンポジウム論文集, 12 , pp.7-12 (1996).
- [15] 二階堂義明, 下田宏, 高橋信, 吉川榮和 : WWW を用いた CAI システムの構築と評価実験, Human Interface News and Report , Vol.12, No.3, pp.325-332 (1997).
- [16] Nakabayashi, K. : “CALAT: an intelligent CAI system using the World Wide Web”, Syst. Comput. Jpn, Vol.28, No.9, pp.17-25 (1997).
- [17] 大林史明, 下田宏, 吉川榮和 : インターネットを用いた認知心理学に基づく知的教育環境の構築, 第 13 回ヒューマン・インターフェース・シンポジウム論文集, 13, pp.593-598 (1997).
- [18] 下村武久, 吉田敦也 : WWW におけるバーチャルクラスルームの設計, 第 13 回ヒューマン・インターフェース・シンポジウム論文集, 13, pp.609-614 (1997).
- [19] 緒方広明, 矢野米雄 : アウェアネスを指向した開放型グループ学習支援システム Sharlok の構築, 電子情報通信学会論文誌 D-II, Vol.J80-D-266, No.4, pp.874-883 (1997).
- [20] 中野 幹生, 堂坂 浩二, 宮崎 昇, 平沢 純一, 田本 真詞, 川森 雅仁, 杉山 聡, 川端 豪 : 柔軟な話者交代を行う音声対話システム DUG-1, 言語処理学会第 5 回年次大会論文集 (1999).
- [21] 堂下修司, 新美康永, 白井克彦, 田中穂積, 溝口理一郎 共編 : 音声による人間と機械との対話, オーム社, pp.3-24 (1998).

- [22] 堂下修司, 新美康永, 白井克彦, 田中穂積, 溝口理一郎 共編 : 音声による人間と機械との対話, オーム社, pp.332-334 (1998).
- [23] 堂下修司, 新美康永, 白井克彦, 田中穂積, 溝口理一郎 共編 : 音声による人間と機械との対話, オーム社, pp.29-32 (1998).
- [24] 黒橋禎夫, 長尾真 : 日本語形態素解析システム JUMAN version 3.61 (Manual), 京都大学大学院情報学研究科, <http://pine.kuee.kyoto-u.ac.jp/> (1999).