

Title	電子教材評価のためのAHP階層図設計法に関する研究
Author(s)	藤林, 由紀
Citation	
Issue Date	2003-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1706
Rights	
Description	Supervisor:落水 浩一郎, 情報科学研究科, 修士

修 士 論 文

電子教材評価のための
A H P 階層図の設計法に関する研究

北陸先端科学技術大学院大学
情報科学研究科情報システム学専攻

藤 林 由 紀

2003 年 3 月

修 士 論 文

電子教材評価のための
A H P 階層図の設計法に関する研究

指導教官 落水 浩一郎 教授

審査委員主査 落水 浩一郎 教授

審査委員 篠田 陽一 教授

審査委員 片山 卓也 教授

北陸先端科学技術大学院大学
情報科学研究科情報システム学専攻

110112 藤林 由紀

提出年月: 2003 年 2 月

概要

本論文では、電子教材評価のためのAHP（階層化意思決定法）階層図の設計法を提案する。猪俣により提案された「学習者の満足度に着目したAHPを利用した電子教材評価法」をもとに、電子教材改良個所がより明確に抽出できるように改良したAHP階層図の設計法を提案する。電子教材には繰り返し学習や、個人の要求に応じた学習が可能であるという利点があるが、目の疲労や教授者とのインタラクションが取りづらいという欠点もある。そこで評価対象の電子教材と、教室講義などの他の講義形態間を同じ評価基準を用い評価することが可能となる、学習者の主観的判断をもとにした評価法を提案する。AHP階層図の設計は、学習者の主観的判断をもとにした評価が行え、かつ電子教材の改良個所がより明確に抽出できるよう、二階層に分けそれぞれの階層において主因子分析法により抽出した評価基準を用いて行った。提案する評価法を用いて本学で開講されている講義の一部で、実際に開発したWebベース電子教材を利用し実証実験を実施した。提案する手法により評価を行い、電子教材の改良点が抽出できることを確認した。

目次

第1章	はじめに	1
第2章	電子教材評価プロセス	4
2.1	主な電子教材評価法	4
2.2	電子教材評価法のアプローチ	5
2.2.1	反応による評価	6
2.2.2	評価パラメータの抽出	7
2.2.3	評価パラメータの重要度測定：一対比較法	8
2.2.4	評価パラメータの重要度測定：AHP	8
2.3	電子教材評価プロセス	9
2.3.1	ステップ1：評価基準の抽出・AHP階層図の設計	11
2.3.2	ステップ2：評価パラメータの重要度の定量的測定	13
第3章	提案するAHP階層図設計法	15
3.1	既存のAHP階層図設計法の問題点	15
3.1.1	評価基準間の相関関係について	15
3.1.2	教材の改良点の抽出について	19
3.2	提案するAHP階層図設計法	20
第4章	電子教材の開発	24
4.1	Webベース電子教材の開発	24
4.1.1	自然言語処理論	24
4.1.2	極低温生成の科学と技術	25
4.2	データ収集システム	28
第5章	本評価法の有効性の検証	30
5.1	実証実験．1	30
5.1.1	概要	30
5.1.2	AHP階層図の設計	31
5.1.3	重要な評価基準の抽出	38
5.1.4	実証実験1：考察	40
5.2	実証実験．2	42

5.2.1	概要	42
5.2.2	重要な評価基準の抽出	43
5.3	考察	46
5.3.1	クラスター分析とは	46
5.3.2	クラスター分析結果	47
5.3.3	まとめ	49
第6章 おわりに		50
謝辞		53
参考文献		54
付録		57
	一対比較法	57
	内部従属法	59

目 次

2.1	評価プロセスの全体像	10
3.1	猪俣により提案されたAHP階層図	16
3.2	AHP階層図(相関関係なし)	18
3.3	AHP階層図(相関関係あり)	18
3.4	AHPテンプレート	22
4.1	撮影環境	25
4.2	自然言語処理論 INDEX 画面	26
4.3	Webベース電子教材 画面例	26
4.4	インタラクション画面例	27
4.5	Webベース電子教材 画面例2	27
4.6	質問紙調査 画面例	29
4.7	AHPアンケート 画面例	29
5.1	スクリープロット	33
5.2	因子分析 結果画面	34
5.3	設計したAHP階層図	37
5.4	レベル2 評価基準ごとのウエイト割合	38
5.5	レベル4 評価基準ごとのウエイト割合	39
5.6	レベル2 評価基準ごとのウエイト割合	44
5.7	レベル4 評価基準ごとのウエイト割合	44
5.8	実証実験1 クラスタ分析結果	48
5.9	実証実験2 クラスタ分析結果	48
5.10	クラスタ分析 結果比較	49

表 目 次

2.1	特性表	14
3.1	因子負荷量と他因子との差との関係	15
3.2	正しい一対比較行列	18
3.3	一対比較行列	18
5.1	評価基準ごとの得点 F 検定 $> f(1,43,0.05)$	40
5.2	実証実験 1 特性表	40
5.3	猪俣による特性表	41
5.4	評価基準ごとの得点 F 検定 $> f(1,32,0.05)$	45
5.5	実証実験 2 特性表	45
1	重要性の尺度とその定義	57

第1章 はじめに

ネットワーク技術の発展や情報メディアの技術革新、それに伴うインターネットの普及により、企業や学校など様々な場所で多種多様な教育システムが開発され、実用化されてきた。日本政府が平成13年に出した「e-Japan 戦略」においても「IT革命は産業革命に匹敵する歴史的な大変革を社会にもたらす」と言われており、高等教育においては遠隔学習システムを用いた授業の単位が認められるなど、環境は整備されてきている。これらの教育システムは、初等・中等教育においては、平成10年以降の学習指導要領変更の影響を受け「情報化への対応」、「総合的学習の時間の創設」を盛り込み、児童自身の考える力を育てるためのものへと変わりつつある。また企業内教育や大学の講義などの高等教育では、インターネットを利用したコンテンツ配信システムが主流になりつつある[1]。遠隔学習システムには、地理的制約や時間的制約を受けない、繰り返し学習が可能である等の利点があるが、同時に教授者や学習者同士のインタラクションが即座に取れないなどの問題点も存在する。これらの学習システムの評価は、受講者の学習成績に与える影響や、質問紙調査法の結果を統計解析することにより行われることが多い。しかしながら、学習システムの得失を総合的に評価する評価法は確立されていなかった。

これらのことから猪俣は、AHP (Analytic Hierarchy Process : 階層化意思決定法) による学習者の満足度に着目した電子教材の評価法を提案した[2]。AHPとは、1971年にアメリカのT.L.Saaty教授により提唱された、人間の勘や直感、フィーリングといった主観的判断を取り入れつつ合理的な決定を促す意思決定の手法である。まず意思決定の構造として解決したい「問題」と、その問題に対する最終的な選択の対象となるいくつかの「解決法(代替案)」が存在する。代替案の中から一つに絞り込むために、両者間に「評価基準」が存在し、これらは階層構造で表される。そして最終目標からみて評価基準の重要度を求め、次に各評価基準からみて代替案の重要度を評価し、最後に最終目標からみた代替案の評価に換算する。AHPではこれらの重要度測定を一対比較により行うので、これまでモデル化したり定量化が難しかった問題に対しても扱うことができる。一対比較とは、ある基準をもとに2つの要素を比較する方法である。意思決定者に「要素iは要素jに比べてどのくらい重要ですか」や「どのくらい好ましいですか」と問い、「大変重要である」や「同じくらい重要である」、「重要ではない」などの形式で答える。2つの要素間の比のみを考えれば良いので、全体を見渡す必要がなく意思決定者の負担を軽減することができる。しかしながら、一対比較により得られた数値はあくまでも2つの要素の価値の比較であるから、全体として首尾一貫した整合性をもっているかどうかわからない。AHP

Pでは、このような回答の矛盾を検知する機能として整合性の指標を提供しており、整合性が悪い場合は一対比較の見直しやデータの排除などの対策が取れるようになっている。AHPを利用することにより「評価基準がどれだけあるのか」、「評価基準がそれぞれどの程度重要なのか」、「各評価基準に対して、どの程度候補案に優劣があるのか」が明確になるという利点がある [3][4][5][6]。AHPそのものは、最終的な意思決定に上記の重要度を利用するが、猪俣は各層の評価基準の重みが評価基準の重要度を表現した情報であると考えた。

また猪俣は、Kirkpatrickによって提唱された教育評価法(Kirkpatrick's levels of training criteria)の4つのレベル、「1. 反応」、「2. 学習」、「3. 行動」、「4. 業績」のうち、第1レベルの「反応」による電子教材評価法の体系化を目指した。教育評価法とは、1959年に経営学者 Kirkpatrickによって提案された企業内教育の評価法である [7]。「反応」「学習」「行動」と段階的にステップを踏み、最終的に実施するのが難しいと言われている「業績」評価にまで近づけるという意図が含まれている。レベル1の「反応」では、アンケートにより受講者の満足度を評価し、評価結果は教育プログラムの改善に使用され、プログラムの価値を示す材料ともなる。これらを利用し、「反応」を満足度という尺度から定量的に測ることで、学習者の教育システムに対する満足度の低い部分を電子教材の改良箇所として抽出することを考えた。学習者の「反応」から判断される評価パラメータを、統計解析の手法の一つである主因子分析を利用して抽出し、AHPを利用した評価パラメータの重み付けと電子教材と比較対象となる講義形態の得点間の有意差から、評価パラメータの重要度を定量的に測定し、重要なパラメータの選定を行う。この重要な評価パラメータは、評価対象の電子教材と他の講義形態との「差」を特徴付けるものであり、電子教材を改良する手がかりを与えるものと考えられる。評価プロセスとしてはまず、質問紙調査で得たデータから主因子分析により評価基準を抽出し、それをもとにAHP階層図の設計を行う。一対比較により各評価基準の重み付けを行い、その結果を分散分析を用いて有意性の検定を実施し、電子教材を評価する重要なパラメータを選出する手法である。しかしAHP階層図の設計法、評価手段の自動化の二点において改良すべき点があった。

本研究ではこれらの問題点を解決することにより、代替案の選択において誤った判断を下す危険性を減少させ、また評価コストを下げるのが可能となる評価法の考案を目的とした。具体的には、Kirkpatrickの教育評価法の第1レベル「反応」を、満足度に影響を与える学習者の主観的判断とし、それらをもとに評価が行え、かつ教材を改良するという点を重視した新しいAHP階層図の設計法を提案する。またデータ収集・分析などの作業において、自動化を取り入れた評価法を提案する。この評価法では、学習者の満足度そのものを評価基準として扱うのではなく、満足度に影響を与える学習者が講義に対して求めるもの、例えば講義に対する集中力や、電子教材を利用する上で常に問題となる身体的疲労度などを評価基準に取り入れ、さらに階層を分けて機能性、操作性に関する要因を評価基準とする。これらにより、講義内容や受講者のモチベーションに左右されない評価が行え

ると考える。同時に、教材を教授者に関係する部分と学習環境に関係する部分とに分けることで、教材の改良を教授者が行うのか、システム設計者が行うのかを明確にした。

以下、本論文の構成について述べる。本論文は6つの章から構成される。

第2章では、電子教材の評価法開発にあたり従来の評価法について考察し、猪俣により提案された評価法の立場を明確にする。また、その評価プロセスについて詳細に述べる。今までの教育効果評価法について、どのような手法が用いられてきたのかなどをもとに、望ましい評価法について検討していく。

第3章では、AHP階層図の設計法について、猪俣により提案された手法の改良を通して新たなAHP階層図設計法を考案する。大まかな評価プロセスは猪俣の手法通りだが、AHP階層図の評価基準の選定において、電子教材の改良箇所がより具体的に抽出出来るよう階層図の設計において改良を行った。

第4章では、本学バーチャルユニバーシティプロジェクトにて我々が開発したWebベース電子教材、および本評価プロセスを円滑に進めるために実装したプログラムについて詳細を述べる。

第5章では、実証実験の結果から、提案する評価法の有効性について検証する。猪俣の実験結果と、今回提案する評価法を用いた実証実験結果とを比較することで、評価法の改良が行われたかどうかを検証する。また被験者の意識の違いが分析結果にどのような影響を与えるのかを考察し、それらを教材の改良に活かしていく手法について検討する。

最後に、第6章において、本研究をまとめ、今後の課題、展望について述べる。

第2章 電子教材評価プロセス

2.1 主な電子教材評価法

現在、主に用いられている遠隔教育システムの評価法として

- 学習成績での評価
- 学習時間による評価
- 学習コストによる評価
- 質問紙調査法による評価

などが挙げられる。これらは、評価対象である遠隔学習システムに求められる効果によって使い分けられる。以下、各評価法について説明する。

- 学習成績での評価

遠隔教育システムと従来の教室講義、他の教育システムとの比較や、受講者を成績上位群と成績下位群に分けて事前知識別の教育効果比較などの視点で評価を行う。これらは、Kirkpatrickの教育評価法のレベル2「学習」に該当する。具体的な評価法として、診断的評価法、形成的評価法、総括的評価法の3つが挙げられる[8]。診断的評価は、学習活動前に学習内容に対する知識、能力、技能などをどの程度備えているのかを把握するために実施される。形成的評価は、学習途上に学生の学習のモニタリングと調整・改善のために実施される。総括的評価は、学期、単元の終わりに教育効果の有効性検定等を目的として実施される。診断的評価として事前テスト(プレテスト)、総括的評価として事後テスト(ポストテスト)を実施し、その成績を比較する手法が一般的である。

- 学習時間による評価

遠隔教育システムを利用した単独学習では、受講者の好きな時間の受講が可能であったり、繰り返し学習が可能になるため、受講者ごとに学習時間がそれぞれ異なる。講義内容に対して予備知識を持つ受講者の学習時間は短縮されると考えられ、講義内容を初めて学ぶ受講者の場合は、学習時間は長くなると考えられる。Kirkpatrickの教育評価法において、学習コストとして考えるとレベル4「業績」に該当し、短時間で集合学習などの他の学習方法と同一の成果が得られるという点ではレベル2「学習」に該当する。

- 学習コストによる評価

主に企業内教育において、学習コストによる評価が用いられている。受講者個人のスキルアップではなく、企業としての作業効率を上げることを目的とした企業内教育では、その経済的効果測定によって教育効果の評価を行うことになる。Kirkpatrickの教育評価法においては、レベル4「業績」に該当する。教育の経済的効果は、その教育に要した費用とその教育を行ったことによって回収できる経済効果によって評価することができる。企業内教育において必要なコストは、システム開発や整備にかかる費用と教育を行う際の運営費用とを合わせたものになる。これに対する経済的効果は、受講者のスキルアップによって作業効率上がることによる人件費削減などにより算出される。

- 質問紙調査法による評価

質問紙調査法とは、対象となる人々に対して一定の様式と内容を持つ質問紙を作成し、これに対する回答を解析することで問題解決に役立つ情報を引き出す手法である。分析者の目的を明確にし、目的に応じた質問項目の作成が必要となる。質問紙調査の目的は、基礎的な統計資料を得るため、新たな問題発見のため、問題の原因や構造を解明するためなどに分かれ、さらに問題の原因や構造を解明する手法は探索的調査と確認的調査とに分けられる。探索的調査とは、問題の原因や構造に関連すると考えられる活動や意識の諸側面を網羅的に調査し、収集したデータを様々な観点から解析することによって問題の真の原因や構造を探っていく手法である。もう一方の確認的調査とは、問題の因果関係や構造を事前に予想し、予想した因果関係や構造（これを仮説という）に関連する活動や意識の特定の側面を選択的に調査し収集したデータ解析から仮説の可否を検証する手法である [10]。

2.2 電子教材評価法のアプローチ

上述したような既存の評価法を利用した場合、評価者の目的に応じて学習システムの機能的な利点や欠点を把握することや、システムが成績・コストに与える影響を測ることは可能である。しかし、個々の受講者の主観的判断をもとにした評価を行うことは困難である。例えば、遠隔教育システムを利用した講義と、従来型の教室における集合学習との違いが存在し、双方において共通の比較要素が存在しない場合が多いため、実施する評価によっては、評価対象に何らかの影響を与えてしまうような結果が生じることが考えられる。このような状況を踏まえ、猪俣は急激な変化を続ける学習環境に適切に応じた評価を可能とする、電子教材評価法を体系化することを目指した。

2.2.1 反応による評価

猪俣の電子教材評価法は、電子教材を利用した受講者が持つ「反応」の定量的測定を可能にすることに主眼をおいている。Kirkpatrickは教育を評価する際に4つのレベル(評価要素)に分割して評価を行う方法として教育評価法を提案した。4つのレベルとは

- Level 1 反応 (Reaction)
- Level 2 学習 (Learning)
- Level 3 行動 (Behavior)
- Level 4 業績 (Results)

から構成される。このレベル分けには「反応」「学習」「行動」と段階的にステップを踏み、最終的に実施するのが難しいといわれる業績にまで近づける手段でもあるという意図が含まれている。以下、各レベルについて述べる。

- Level 1 反応 (Reaction)
アンケートにより学習者の満足度を測定する評価要素である。Kirkpatrickは反応を「教育を受ける者が、特定のプログラムに対して持つ好意の程度」と定義した。すなわち、受講者の教育プログラムに対する満足度を評価し、評価の低い要素を改善するということである。またこの評価は、プログラムの価値を示す証拠にもなる。Kirkpatrickは受講者の反応のみならず、教育を観察する立場の人の意見も取り入れ、比較することによってプログラムの効果が最も適正に示されると提案している。
- Level 2 学習 (Learning)
Kirkpatrickは、好意的反応は必ずしも学習されたことを保証しないと認めている。したがって、実際に行われた学習の量を測定すること、つまり、教育カリキュラムの学習目標に対する理解度を測定することが重要になる。このステップには、「受講者がどのような原則、事実、技法を理解し吸収したか」を測定して認識することが含まれているため、目標に対する重要な「結果管理」となる。評価方法は、理解度テストから実技演習までそれぞれの学習方法に応じて、知識、技能、態度の分類に適した評価方法が採用される。
- Level 3 行動 (Behavior)
学習の結果として得られる、受講者の行動変容を評価する評価要素である。学習前・学習後の系統的な評価、また受講者の行動を良く知る人など出来るだけ多面的な人により評価を行わねばならず、行動変容を評価するというのは、大変難しいことではある。しかし、学習したことが実際に職場等で生かされたか否かを調査する評価方法は、簡便で効果的である。

- Level 4 業績 (Results)

この評価要素では、成果に重点が置かれる。学習を通して身に着けた知識・技術等により、どの程度業績が上がったかを評価する。成果の例には次のような事項が挙げられる。

- － 経費の削減
- － クレームの減少
- － 作業時間の短縮
- － 生産の質および量
- － モラルの向上

遠隔教育システムでは、システムの優劣が受講者に与える影響は大きいと考えられ、「分かりやすい」や「良い」、「使いやすい」といった形で満足度を測り、「反応」に基づいて評価を行うのは有益であると言える。「反応」における評価要素の一例として、以下のものが考えられる。

- 講義の受講しやすさ
- 教授者の話し方、聞き取りやすさ
- 受講者に配慮された教授法
- 受講者の興味への応え方
- 受講目的の充足度、満足度
- 学習環境やシステムの充足度、満足度

2.2.2 評価パラメータの抽出

質問紙調査によって、前述したような「反応」を測るための評価基準を得る。なぜなら受講者は、質問紙調査における電子教材の機能面や内容面に関する判断基準から、電子教材の主観的・直感的な評価を行うことが出来る。しかし、評価者が無作為に多数の判断基準を呈示して比較させることは、受講者の主観的判断をもとにするために評価結果を曖昧にさせてしまうことも考えられる。これらのことから、質問項目の設計は、十分な検討が必要である。

次に、質問紙調査の結果から複数の受講者が持つ共通の判断基準を推定するため、統計処理の手法である主因子分析を用いる。主因子分析を利用し、受講者の回答データから相関の強いもの同士をグルーピングし、データ中に潜む要因を探る手法である。質問紙調査による評価のみでは、質的な局面に注目したデータが量的に得られたデータによって影響

を受けてしまうこともある。例えば、全体の評価としては重要性の低い要素に対して量的に多いデータが確保された場合、誤って重要度としては低いものが出力されてしまうことがあるからである。

2.2.3 評価パラメータの重要度測定：一対比較法

学習者の満足度を測る評価基準の重要度を定量的に測定し、重要な評価パラメータの選定を一対比較法を用いて行う。一対比較法とは嗜好型官能評価法の一つで、自分の提示した尺度（基準）に従って対象物を比較することが可能である。複数個ある要素から2つずつ取り出し比較を行い、各項目ごとに得点付けし、得られた数値をもとに分析を行なう。最終的には要素を一本の数直線状に配置させ、対象物間の関係性と順位を見ることができる手法である。被験者は全項目の選好度の順を決定する必要が無いという利点がある。一般的に、明確な尺度を持たない要素間の比率を被験者が厳密に答えるのは難しいが、一対比較法では、被験者は具体的な値でなく言葉で表現すればよい。一対比較値を獲得するために、「非常に重要である」「重要である」「同じくらい重要である」といった言葉によるファジィな表現を用いることで主観的な判断の分析が可能になり、被験者の負担を軽減できる。

2.2.4 評価パラメータの重要度測定：AHP

一対比較法により、判断基準の重要度を定量的に測定することを体系的に実施する手法として、AHP（Analytic Hierarchy Process：階層化意思決定法）を利用する。AHPは、複数の候補の中から1つを選ぶという意思決定問題に対し人間なら誰もが持つ経験や勘といった主観的な価値判断を、重要な要素として取り入れ合理的な決定を下すための手法である。AHPは1970年代に米ピッツバーグ大学のT.L.Satty教授によって開発された手法であり、問題の分析において主観的判断とシステムアプローチをうまく混ぜ合わせた問題解決型意思決定手法の1つである。実際に適用事例として、「ペルー人質事件」「首都機能移転」等の国家規模の政策から、「コンセプト評価」「企業イメージ評価」「デザイン評価」「人事評価」等の企業経営上の意思決定まで幅広く利用されている[9]。

AHPではまず、解決したい問題と最終的に選択するいくつかの代替案、代替案の中からひとつを選択するための評価基準があり、それらを三段階の階層構造に配置する。そして最終目標から評価項目の重要度を求め、次に各評価基準からみて各代替案の重要度を評価し、最後にこれらを総合目的からみた代替案の評価に換算することで最終的な決定を下すことが可能となる。しかし、一対比較法を利用することから得られた一対比較値はあくまでも2つの項目の価値の比較であり、全体として首尾一貫した整合性を持っているかどうかはわからない。そこでAHPでは、回答の矛盾を検知する機能として整合性の指標（C.I値：Consistency Index）を提供する。

以下に、他の意思決定の手法とくらべAHPを利用することの利点についてまとめる。

- 評価基準がたくさんあり、互いに共通の尺度がないような問題の解決ができる
- これまでの定量的分析では取り扱えなかった、不明瞭な要因がからむ問題の解決ができる
- 首尾一貫性のないデータを扱え、かつ、首尾一貫性の度合いが測れるのでデータの修正が可能である
- システムアプローチと主観的判断を組み合わせることで、人間の勘や経験を生かした意思決定が行える
- グループでの意見を取りまとめた意思決定が行える
- 問題を階層構造で表現することで、ある限られた条件での比較・考察を重ねていけば最終決定が下せる
- 各評価基準に対して、どの程度候補案に優劣があるのかが定量的に把握できる

2.3 電子教材評価プロセス

電子教材を改良することを目的として、質問紙調査法をもとにした主因子分析法、およびAHPを組み合わせた評価プロセスが猪俣により提案された。この評価法は、大きく以下の2つのステップからなる

ステップ1：評価基準の抽出・AHP階層図の設計

ステップ2：評価パラメータの重要度の定量的測定

この評価プロセスはそれぞれの役割を持つ4つのアクター、被験者、実験実施者、AHP階層図設計者、分析者により構成される。被験者は実際に電子教材を用いて講義を受講し、質問紙調査、AHPアンケートに答え評価を行う。実験実施者は電子教材の提供、AHPアンケートの実施を行う。AHP階層図設計者は、質問紙調査項目の設計、AHP階層図の設計を行う。分析者は質問紙調査により得たデータを分析し、因子を抽出、評価基準を選定する。さらに、AHP階層図設計者によって設計されたAHP階層図をもとにAHP評価による分析を行う。評価プロセスの全体像を図2.1に示す。図2.1はUML activity diagramであり、各アクティビティのフローが示されている[11]。楕円はアクティビティを表し、破線矢印はそのアクティビティのフローを表す。また長方形はオブジェクトを表し、直線矢印はあるアクティビティから別のアクティビティへのオブジェクトのパスを表す。

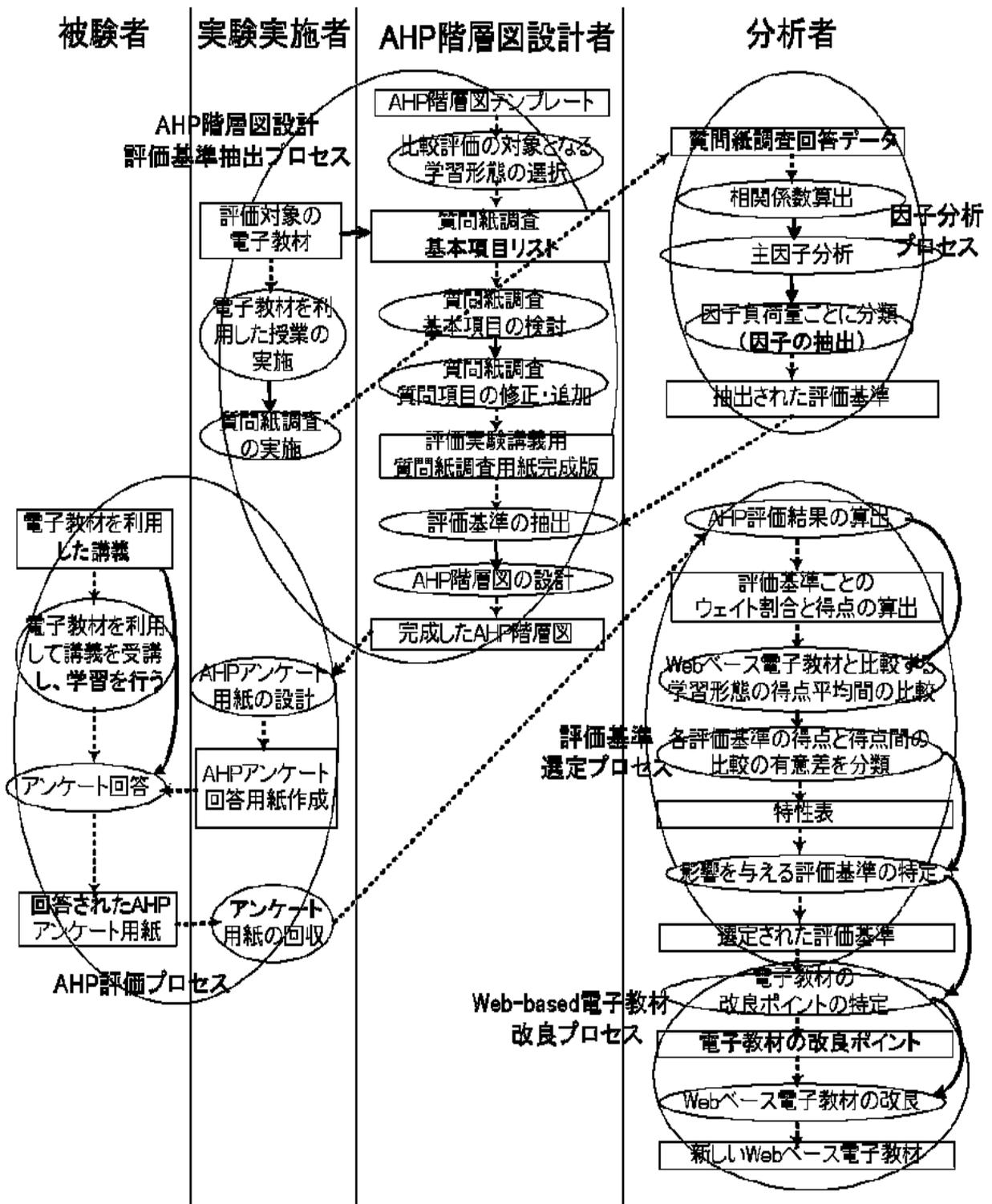


図 2.1: 評価プロセスの全体像

2.3.1 ステップ1：評価基準の抽出・AHP階層図の設計

本ステップでは、学習者の「反応」から判断される評価パラメータを質問紙調査を主因子分析を用いて分析して抽出し、AHP階層図の評価基準とすることによりAHP階層図を設計する。ここで用いる統計的多変量解析手法の一つである因子分析 (Factor Analysis) とは、ある1つの現象を一定の意図のもとに多種類の計測項目で観測した統計資料に基づき、現象の本質を構成する普遍的な因子に関して探索や実証を行う手法である。そして探索した因子には分析者自身が解釈し意味付けを行う [12]。

因子分析には探索的因子分析と確認的因子分析とがあり、利用する方向によって使い分けられる。ある概念に従い、それを測定しうるだろうと考えられる項目を集め、データを収集し、そこにどのような因子が見られるかを探索する時には、探索的因子分析を行う。また既存の尺度を用いてデータを収集し、そのデータが先行研究での因子構造と合致するか否かを検討する場合には確認的因子分析を行う。明確な仮設がなく、因子を求める必要がある本手法では、探索的因子分析を行うことになる [13][14][15][16]。

因子分析法のモデルとは、実際に観測されるデータをいくつかの変数の一次結合によって表すものである。質問項目の得点と因子負荷量、因子得点との関係を式で表すと以下のようになる。

(質問項目の得点) = (質問項目の因子負荷量) * (因子得点) + (独自部分)
例えば「菓子商品の人気」を因子の結合として表すと、次のように表すことができる。

商品の人気得点 = ブランド得点 + 味得点 + 香り得点 + 外装得点 + 価格得点
以下に因子分析モデルを代数的に示す。

$$Z_{ij} = a_{j1}f_{i1} + a_{j2}f_{i2} + \dots + a_{jm}f_{im} + d_j u_{ij} \quad (2.1)$$

ここで、 Z_{ij} とは被験者 i の項目 j に対する回答の標準得点である。 $a_{j1}, a_{j2} \dots a_{jm}$ は「因子負荷量」と呼ばれ、 $f_{i1}, f_{i2}, \dots f_{im}$ は「因子得点」と呼ばれるものである。また、 d_j は独自因子と呼ばれるもので、因子負荷量ではかりきれない部分を表す。

因子分析の手順は、質問紙調査の回答から質問項目ごとの相関係数を算出し、相関行列を作成する。次に相関行列を固有値分解し、因子負荷を求め、因子スコアを推定する、という手順で行われる。分析の結果から、質問紙調査における質問項目がどれだけ因子分析によって抽出された因子に左右されているのか考察できる。

次に、質問項目間の関係を相関係数で示す。相関係数とは2つの変数 X と Y の間に直線的な関係があるかどうかを見るためのものである。- 1 ~ 1 の間の値を取り、相関係数が + 1 に近い場合2変数には「正の相関がある」と言え、強い関係があると考えられる。0 に近ければ2変数には関係はなく、また - 1 に近ければ「負の相関がある」と言え、2変数の関係は反対の性質を持つと考えられる。2変数 X, Y の変動関係の指標である X と Y の共分散 S_{XY} は、2変数 X, Y の偏差の積の平均値である。ここで N を標本数、2変数 X, Y の平均値を X_M, Y_M とすると

$$S_{XY} = \frac{1}{N} \sum_i (X_i - X_M)(Y_i - Y_M)$$

である。相関係数の算出には、一般的に相関係数と呼ばれるピアソンの積率相関係数の式 (2.2) を利用する。

$$\text{相関係数 } r_{xy} = \frac{X \text{ と } Y \text{ の共分散}}{[X \text{ の標準偏差}] * [Y \text{ の標準偏差}]} \quad (2.2)$$
$$(-1 \leq r_{xy} \leq 1)$$

この式から算出された相関係数を質問項目ごとに並べて表したものが相関行列となる。この相関行列 R は正方行列であるので、正方行列 R は式 (2.3) のように固有値分解が可能である。

$$R = \sum_{i=1}^n \lambda_i x_i x_i \quad (2.3)$$

この時 $x_i x_i$ のように固有ベクトルを調節してやれば、固有ベクトルがすなわち因子負荷ベクトルになる。

さらに因子抽出過程において、因子数の決定には判断基準がいくつか存在するが本評価プロセスにおいては、一般的に用いられる以下の2つの基準により判断する。

1. Kaiser-Guttman 基準

固有値は、その因子で全体のどのくらいを意味するかを表しており、これが 1.0 を超えていることは、すなわち質問紙調査一項目以上の影響力がその因子にあることを示す。

2. スクリー基準

固有値の大きい順にプロットした図 (スクリープロット) から、固有値の値が大きく落ち込む部分で決める。こうして得られた因子を因子負荷量をもとに分析者が解析、意味づけを行い命名する。

最後のステップとして、因子の特定作業を行う。全観測データ中に潜む因子を特定するために、各々の因子に寄与された因子負荷量ごとに高い負荷を持つ変数をグループ化する作業を行う必要がある。ここで因子負荷量とは、解析した結果算出された相関の強さのことをいい、変量と因子の間の相関係数に等しいから、この負荷量が多いほど相関が強いことを意味している。抽出された因子を因子負荷量から解釈しやすくするため、因子軸の回転という手法を用いる。因子軸の回転には大きく分けると直交回転と斜交回転があり、直交回転は因子間は無相関であることを仮定しており、また斜交回転は因子間に相関を仮定したものである。本評価プロセスでは直交回転の代表的なものである Varimax 回転を行う。Varimax 基準として、因子負荷の平方の分散を用いる。Varimax 法では、全ての因子について同時にこの分散を最大とする解を求めるのが特徴である。分散を最大にするということは、特徴をより際立たせることに等しい。

以上のステップで抽出した因子に、分類された項目の傾向ごとに命名する。

次に主因子分析により抽出し命名した評価基準をもとに、AHP 階層図を設計する。階層図を設計するにあたり、まず AHP 評価の問題と代替案を決定しなければならない。本

評価プロセスでは問題を「効果的な学習をするために、あなたはどちらの学習形態を選択しますか」と定義し、代替案として「Webベース電子教材」と「教室講義」とした。階層図を設計するにあたり、注意すべき点を以下に挙げる。

- 同一レベルに位置する評価基準は、互いに独立であるようにしなければならない
- 一対比較の対象となる評価基準は7個まで、多くても9個以下にする

2.3.2 ステップ2：評価パラメータの重要度の定量的測定

ステップ1において設計されたAHP階層図に基づきAHPアンケートを作成し、受講者に回答してもらう。この回答データから、評価パラメータの重み付けと電子教材と比較となる講義形態の得点間の有意差から、評価パラメータの重要度を定量的に測定し、重要な評価基準の選定を行う。本評価プロセスの特徴は、AHPにより選定される代替案の決定ではなく、AHPにおける分析過程により得られる評価基準ごとの得点結果をもとに重要なパラメータの判断を行うことにある。

AHPアンケートの実施における注意点を以下に挙げる。

- 一対比較値が確信できないときは、その値に関する感度分析を行う
- 総合的重要度は選好度を表しており、この値の順に好ましい代替案となるが、この値の差の取り扱いには注意が必要である。場合によっては、重要度の低い要素を除いて再びAHPアンケートを実施することも必要である
- グループでの意思決定には、一対比較値としては幾何平均値を用いる

評価基準に配分された得点は、AHP階層図に示された評価基準の関係をもとに上層の得点が下層の得点に加味されて算出される。さらに、複数の被験者ごとにその評価基準の得点が得られるため、全体としての得点平均間の比較の有意差を判断するため分散分析F検定を実施する。分散分析F検定は、得点平均間の差の有意差を判定することができ、評価基準ごとの差を特徴付ける有力な判定法である。分散分析F検定においては、それぞれ2群の分散を比として比較を行う。2群A, Bの分散をそれぞれ S_1, S_2 とし、その比は以下の式で示される。

$$F = \frac{S_1^2}{S_2^2}$$

この比をF値と言い、このF値を用いて2群の分散が等分散であるか否かを検定する。実証実験で得られる各群のデータはばらつきや誤差があると考えられ、そのまま2群の差が全体の差であるとは言えないからである。得点と分散分析F検定の結果から、評価基準を各カテゴリに分類する。このカテゴリを表に示し、この表を特性表と呼んだ。表2.1に各カテゴリごとの評価基準の意味を示す。表に示す特性表において、得点が高く統計的に有

意差があるエリアの評価基準のうち、「評価対象の電子教材」が比較対象の「他の学習形態」よりも低いものを電子教材改良の重要な評価基準であるとし、教材改良に有用な評価基準として選定される。これをもとに電子教材の改良を進めていく。なお、選定された評価基準から電子教材の改良を進めていく際には、AHPアンケートの回答と合わせて自由記述による回答からも改良点の推察を行う。

表 2.1: 特性表

	統計的に有意差あり	統計的に有意差なし
高い得点(100点/評価基準の項目数以上)	効果的な学習をする際に影響を与え、評価対象とする電子教材と比較する講義形態間の差を特徴付ける重要なパラメータとみなす	効果的な学習をする際に影響を与えるパラメータとみなす
低い得点(100点/評価基準の項目数以下)	評価対象とする電子教材と比較する講義形態間の差を特徴付けるパラメータとみなす	評価対象とする電子教材に影響を与えるような重要なパラメータであるとはみなさない

第3章 提案するAHP階層図設計法

既存の電子教材評価法では、主因子分析法による評価基準の抽出では共分散行列、評価基準の重み付けでは一対比較行列、得点間の有意差の検定には分散分析F検定と数学的理論に基づいて評価プロセスが構成されているが、AHP階層図の設計法においては体系化されていなかった。抽出された因子をレベル2の評価基準とし、レベル3の階層には因子に関する質問項目からAHP階層図設計者が選択し設計を行っており、レベル3の評価基準に関して設計者に左右される部分が大きかった。本研究では、階層図の設計を体系的に行える階層図設計法を提案する。提案する設計法では評価基準の独立性に着目し、学習者の満足度そのものを評価基準として扱うのではなく、満足度に影響を与える機能性、操作性に関する要因を評価基準とすることで電子教材の改良点をより明確にする。

3.1 既存のAHP階層図設計法の問題点

3.1.1 評価基準間の相関関係について

まず猪俣により設計されたAHP階層図を図3.1に示す。レベル2には、主因子分析で抽出し命名された4つの評価基準「教授者のプレゼンテーション」「講義中に提供されている機能」「学習空間の快適さ」「満足度」が配置され、AHP階層図が設計されている。ここで因子負荷量と他因子との差との関係を表3.1に示す。

表 3.1: 因子負荷量と他因子との差との関係

	因子間の差が大きい	因子間の差が小さい
因子負荷量の値が大きい	ほぼ他の因子から独立した因子であり、評価に大きな影響を与えている	他の因子に影響を及ぼしている因子であり、評価に大きな影響を与えている
因子負荷量の値が小さい	ほぼ他の因子とは独立した因子である	他の因子に対して影響を与えている因子である

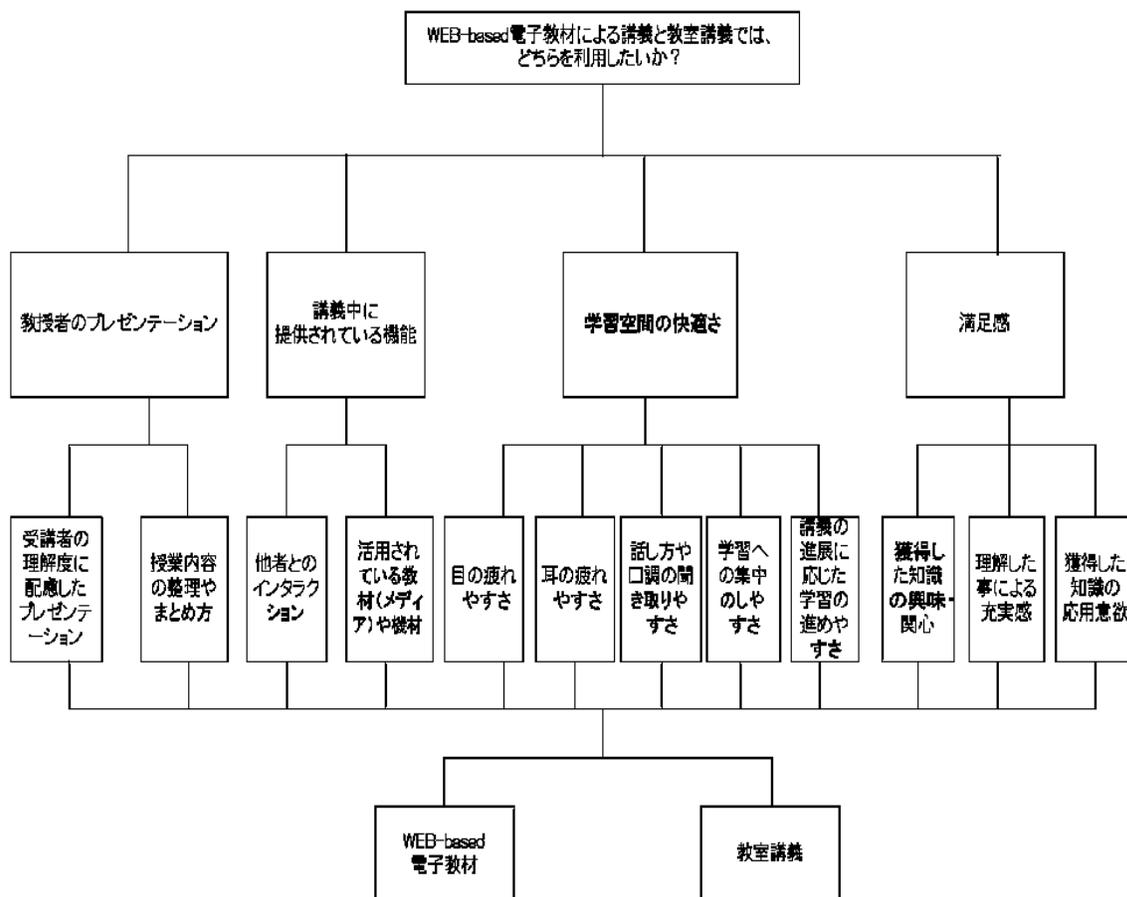


図 3.1: 猪俣により提案されたAHP階層図

レベル2に配置された評価基準「満足度」は、表3.1と因子負荷量から考察すると、他の評価基準と大きく関わっていることがわかった。そこで、実際にどの程度の関係があるのかを、各評価基準を構成している質問項目に対するデータから相関係数を求めると以下のようになった。

- 「満足度」と「講義中に提供されている機能」間 … 約0.4
- 「満足度」と「学習空間の快適さ」間 …… 約0.4
- 「満足度」と「教授者のプレゼンテーション」間 … 約0.3
- 「学習空間の快適さ」と「教授者のプレゼンテーション」間 … 約0.2
- 「学習空間の快適さ」と「講義中に提供されている機能」間 … 約0.2
- 「講義中に提供されている機能」と「教授者のプレゼンテーション」間 … 約0.2

相関の強さの表現として、一応の目安としては、相関係数の絶対値が0.2よりも小さければ「ほとんど相関がない」、0.2から0.4のときは「弱い相関がある」、0.4から0.7のときに「中程度の相関がある」、0.7を超えるときには「強い相関がある」と表現することが多いようである。しかしながらAHPを行う際に注意すべきは、同じ階層に位置する評価基準は互いに独立であるように、階層図を設計しなければならないという点である。なぜなら評価基準間に相関関係がある場合、重要度測定において正しい値を得ることができなくなるからである。実験を行った結果、もし強い相関関係が見られたなら評価基準の抽出やAHP階層図の設計をやり直す必要がある。

相関関係が強い場合、重要度測定にどのような影響がでるかを具体的な例をあげて述べる[4]。

《例》菓子商品の選択をAHPを用いて行う。最終目標は「食べたい菓子を選択する」ことであり、代替案は菓子A, B, Cとする。

- 評価基準間に相関関係がないとき
評価基準を「味」と「形」としてAHP階層図を設計し図3.2に示す。選択者にとって「味」のほうが「形」よりもa倍重要であると仮定して一対比較行列を作成すると表3.2のようになり、完全な整合性を持っている。最終的な評価基準の重み付けは、「味」 $a/(a+1)$ 、「形」 $1/(a+1)$ と求められた。
- 評価基準間に相関関係があるとき
相関係数が1の極端な例を取り上げ説明する。評価基準を「味1」「味2」「形」としAHP階層図を設計し図3.3に示す。「味1」と「味2」間の相関係数は1である。前述の例と同じく、「味」の方が「形」よりもa倍重要であると仮定し一対比較行列を作成すると表3.3のようになり、最終的な評価基準の重み付けは「味1」 $a/(2a+1)$ 、「味2」 $a/(2a+1)$ 、「形」 $1/(2a+1)$ となった。「味1」と「味2」の実体は同じものな

表 3.2: 正しい一対比較行列

	味	形
味	1	a
形	1/a	1

表 3.3: 一対比較行列

	味 1	味 2	形
味 1	1	1	a
味 2	1	1	a
形	1/a	1/a	1

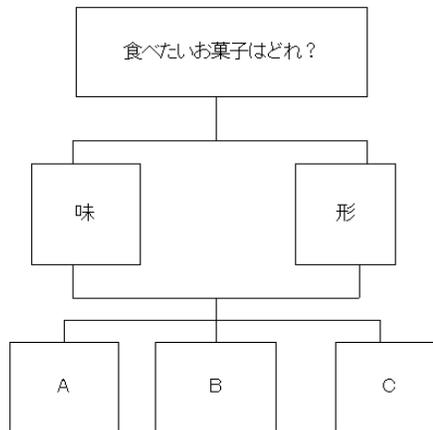


図 3.2: AHP 階層図 (相関関係なし)

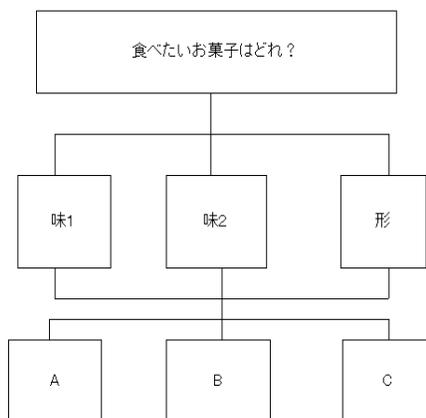


図 3.3: AHP 階層図 (相関関係あり)

ので2倍に強調されていることになり、このように従属性の強い要素を入れると、その要素は x ($1 \leq x \leq 2$) 倍に強調されてしまい、正しい重み付けが行えていない。

以上の例からもわかるように、同一レベルの評価基準は互いに独立でなければならない。独立の要素であれば、一対比較によって行列を作りその固有値と固有ベクトルを計算することにより、各要素のウェイトを決定すればそれが正当な重要度であることが保証されるからである。相関係数の絶対値がどの程度であれば、実害がない程度に独立であるのかという点に関してはまだ研究がなされていない。しかし、出来る限り各評価基準が互いに依存せず、一対比較行列の整合性が成り立つことが望ましい。

本研究では、評価基準が互いに独立になるようなAHP階層図設計法を提案する。評価基準が互いに依存しないことで、一対比較を行う際の矛盾が排除できると考える。

3.1.2 教材の改良点の抽出について

既存のAHP階層図では、レベル3にはレベル2の評価基準をそれぞれ詳細化したものとして、質問項目の中から代替案に対して適切な評価基準となるような項目を選択し設計を行っていた。つまり、明確な設計基準がなく、AHP階層図設計者に依存する部分が大きかった。以下にレベル3の評価基準を全てあげる。

- 受講者の理解度に配慮したプレゼンテーション
- 授業内容の整理やまとめ方
- 他者とのインタラクション
- 活用されている教材（メディア）や機材
- 目の疲れやすさ
- 耳の疲れやすさ
- 話し方や口調の聞き取りやすさ
- 学習への集中しやすさ
- 講義の進展に応じた学習の進めやすさ
- 獲得した知識の興味・関心
- 理解したことによる充実感
- 獲得した知識の応用意欲

レベル3に配置された評価基準は、最終的な得点が高く統計的に有意差がある評価基準のうち、「評価対象の電子教材」が比較対象の「他の学習形態」よりも低いものを電子教材改良の重要な評価基準であるとし、教材改良に有用な評価基準として選定される。猪俣による実証実験では「目の疲れやすさ」「耳の疲れやすさ」「受講者の理解度に配慮したプレゼンテーション」がWebベース電子教材「ソフトウェア設計論」の改良点として選定されている。しかし具体的に、目や耳の疲労度を軽減するにはどうすればよいのかまではわからなかった。他にも、例えば、改良すべき点として「獲得した知識の興味・関心」や「理解したことによる充実感」などが選定された場合、教材の改良を教授者とシステム設計者のどちらが行うのか、具体的にどの部分を改良すればよいのかを明確にすることができない。これはレベル3の評価基準に、質問項目をそのまま用いたことで曖昧さが残っており、教授者に関する部分と学習環境に関する部分とに分かれていなかったためだと考えられる。

本研究では上記の問題点を改良することを通して、AHP階層図設計法を体系化することを目指した。

3.2 提案するAHP階層図設計法

本研究では、体系的に行える階層図設計法を提案する。具体的には、学習者の主観的判断をもとにした評価が行え、かつ評価者が「誰が何を改良すべきか」という点を適切に抽出できるよう、二つの階層図を組み合わせたAHP階層図の設計法を提案する。主因子分析で得られた評価基準を、学習者の「反応」に着目し命名することで評価基準間の相関関係を弱め、学習者が講義を受講する際、何を重視するのかがウエイトとして求められる。さらに、質問紙調査の質問項目を「教授者による部分」と「学習環境による部分」にわけ、それぞれにおいて主因子分析を行い、抽出した評価基準を下半分の階層図に配置することで評価者の目的である教材の改良点を選定することが可能となる。

AHPを行う場合、より多くの意思決定の情報を採用し、それらを整理して総合的に判断することが分析結果をより一層現実的なものにする必要条件である。一般に代替案の評価は多面的であり、1つの評価基準だけで代替案の優劣を決めるのは困難である。1つの代替案が、すべての評価基準について最も優れているということは通常ありえないため、意思決定者は複数の評価基準を総合的に判断して最良の代替案を選択することになる。しかし、評価項目が多いことが必ずしも判断の客観性を保証するものではない。なぜなら、項目間の相関が高くなる可能性があるからである。AHP階層図において、同一レベルの要素には従属性の濃い要素を入れないことが望ましいが、評価基準間に相関関係が見られた場合の対処法として、以下の2つの手法がある。

- 従属性の濃い要素をまとめて1つの要素とし、他の要素と一対比較する
- 内部従属性を用いて重要度の計算を行う

要素の統合化が難しく、かつ要素間の影響力の強さを一対比較で表せる場合には内部従属法が用いられることが多いが、被験者の一対比較を行う負担が増え、評価が過度に複雑になるという問題点がある。本研究では、従属性の濃い要素を1つにまとめる手法を用いて、評価基準間の相関関係を取り除くことを考えた。猪俣のAHP階層図では、レベル2の評価基準「満足度」と他の3つの評価基準、「学習空間の快適さ」「講義中に提供されている機能」「教授者のプレゼンテーション」との間にやや強い相関関係が見られた。これは学習者の「反応」を測るための「満足度」という要素が、学習環境やシステムの機能、教授者など学習者以外の要素から影響を受けて評価されているからである。つまり学習者の「反応」は教材を評価する上で、講義を構成する様々な要因から影響を受けていることになり、猪俣によるAHP階層図において、評価基準間に相関関係が見られたのは、学習者の「反応＝満足度」を評価基準の1つと考え、反応に影響を与える他の要因と同じレベルに配置したことが原因であると考えられる。

これらを踏まえた上で、評価基準が互いに独立になるよう従属性の濃い要素をまとめて1つの要素にするならば、全てをまとめ「反応」だけで評価を行うことになる。そのためには主因子分析法により抽出した因子には、「反応」に関する項目、例えば満足度や充実感などに限定して命名する必要がある。このように、学習者の「反応」に着目し命名することで初めてKirkpatrickの教育評価法「レベル1：反応」をもとにした教材評価が可能となり、学習者が講義に何を求めているのかが明確になる。

教材に得点を付け優劣を決めるだけなら、「反応」をもとにしたシンプルな階層図による評価で可能であるが、本研究ではAHP階層図を用いて評価基準の重要度をもとにした教材改良を目的としており、学習者の「反応」に影響を与えている要因が重要となる。すなわち学習環境や教授者のプレゼンテーションが「反応」をもとに命名した各要素にどの程度影響を与えているのかを測り、その度合いから誰が教材の改良を行うのかを決定するために、レベル3には、電子教材を構成しているのは「教授者」と「システム」であり、教材改良を実施するのも教授者とシステム設計者の二者であることから、「学習環境」と「教授者のプレゼンテーション」の2つの評価基準を配置する。ここで「学習環境」は、教材の画面構成やシステムの機能などシステム設計者により設計される範囲を指し、「教授者のプレゼンテーション」は講義内容や話し方など教授者に左右される範囲を指す。さらに、最終的に改良すべき箇所を限定し抽出できるよう、質問項目を「学習環境」と「教授者のプレゼンテーション」とにわけ、主因子分析を行いレベル4に配置する評価基準を抽出し、レベル3の各評価基準の下に配置する。本研究では、このように学習者と教授者、システム設計者の各役割と目的を明確にわけた評価基準の抽出法を提案する。AHPテンプレートを図3.4に示す。

学習者の「反応」のみによるAHP階層図はシンプルであり直感的な比較になるため、定量的な評価は可能ではあるが一般的なアンケート評価結果とさほど変わらない。提案する二段階に分割したAHP階層図によって、上階層では、最終目的に対し学習者の「反応」がどの程度寄与するのか、下階層では各「反応に基づいた評価基準」のそれぞれに対

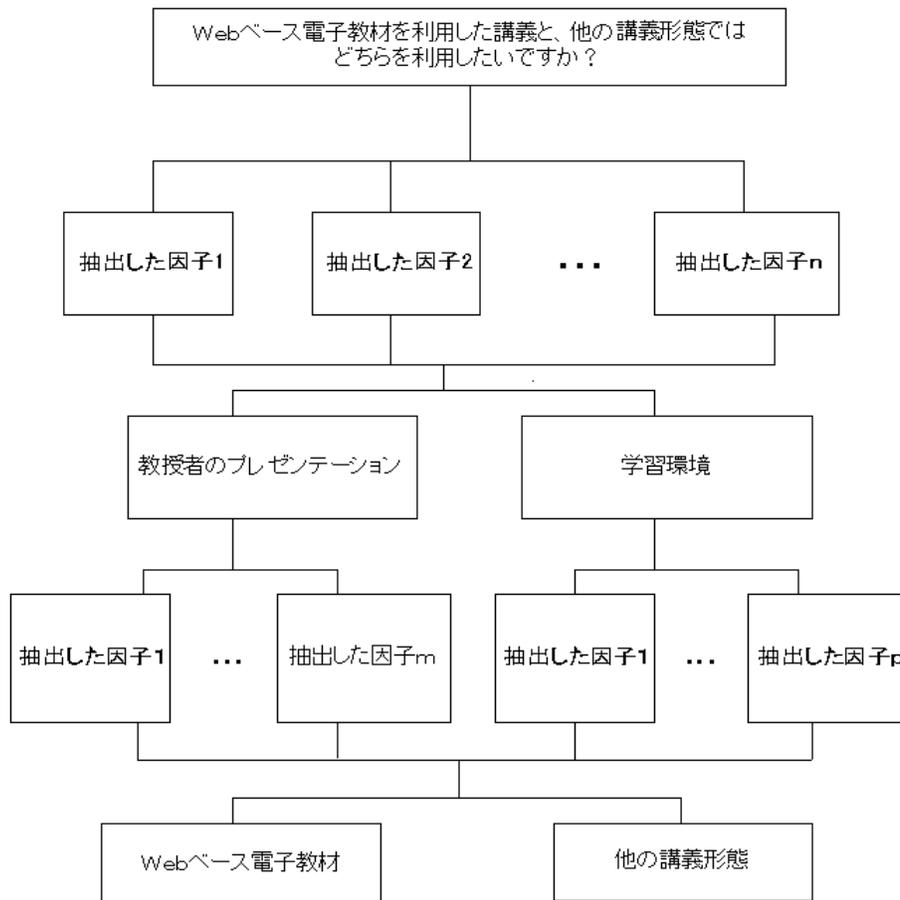


図 3.4: AHPテンプレート

して、各教材構成要素はどの程度寄与するのかを知ることが可能となる。問題を多階層構造で表すことで、直感では気付かなかった利点・弱点が浮き彫りになり、最終的に、代替案選定において弱点を理解した上で対応策を考えることが可能となる。電子教材の評価を行う際には、弱点が浮き彫りになることが教材改良につながるため重要になる。提案する階層図設計法に基づいた評価を実施することで、レベル2において学習者が講義に求めているものがわかり、レベル3, 4から学習者の主観的判断に影響を与えているものが教授者なのか、学習環境なのかウエイトとして数値化され、教授者が講義内容やプレゼンテーション等を改良するのか、システム設計者が教材の機能や画面構成等を改良すべきかを明確にできる。評価者の目的である教材の改良個所の抽出とともに、改良実施者が選定されることで教材の改良がスムーズに行える。

第4章 電子教材の開発

本章では、本学で我々が開発し実証実験の対象としたWebベース電子教材の詳細と、電子教材の評価を円滑に進めるためのデータ収集システムについて述べる。遠隔学習システムとは、計算機を利用した学習システムを指し、大きく二つの形態に分けられる。

- 同期型遠隔学習システム
- 非同期型遠隔学習システム

同期型遠隔学習システムとは、ネットワークを利用して教授者と学習者が双方同期して講義等を受講出来るシステムである。非同期型遠隔学習システムとは、あらかじめ作成された教材を学習者が自分の好きな時間・空間で個別で行う学習システムのことである。我々は、学習者の利便性、とりわけ学習に対する時間の自由度の拡大、また受講場所の自由度の拡大を行える非同期型遠隔教育システムの開発を行ってきた。

今回、我々は本学の情報科学研究科 島津 明 教授による「自然言語処理論」と、材料科学研究科の機材使用手順と実験の基礎を説明した「極低温生成の科学と技術」の2つの講義を遠隔教育システムとして開発を行った。電子教材化するにあたり、準備段階として講義の撮影を行った。図 4.1 に撮影環境を示す。映像の記録として教授者中心の映像と、教授者がプレゼンテーションをする際の手の動作などを記録するためにプレゼンテーション映像の2台のデジタルビデオカメラを利用して収録した。さらにコンテンツ化する際に必要となるタイムレコードを記録するため、デジタルビデオに記録されるタイムレコードの他に、各知識単位の開始と終了をしめすマーキング映像の挿入を行った。撮影に要した時間は、1回の講義(約60分)に平均2時間であった。

4.1 Webベース電子教材の開発

4.1.1 自然言語処理論

「自然言語処理論」は情報科学研究科 島津 明 教授による全15回の講義から構成されるWebベース電子教材である。各回はそれぞれ約60分の講義からなり、1回の講義はさらに細かい単元に分けられているため、受講者はそれぞれ必要な部分を選び受講することが可能である。Webベース電子教材はRealVideo[17]を利用して構築されており、教授者映像、プレゼンテーション映像、テキスト表示、インデックス機能、インタラ

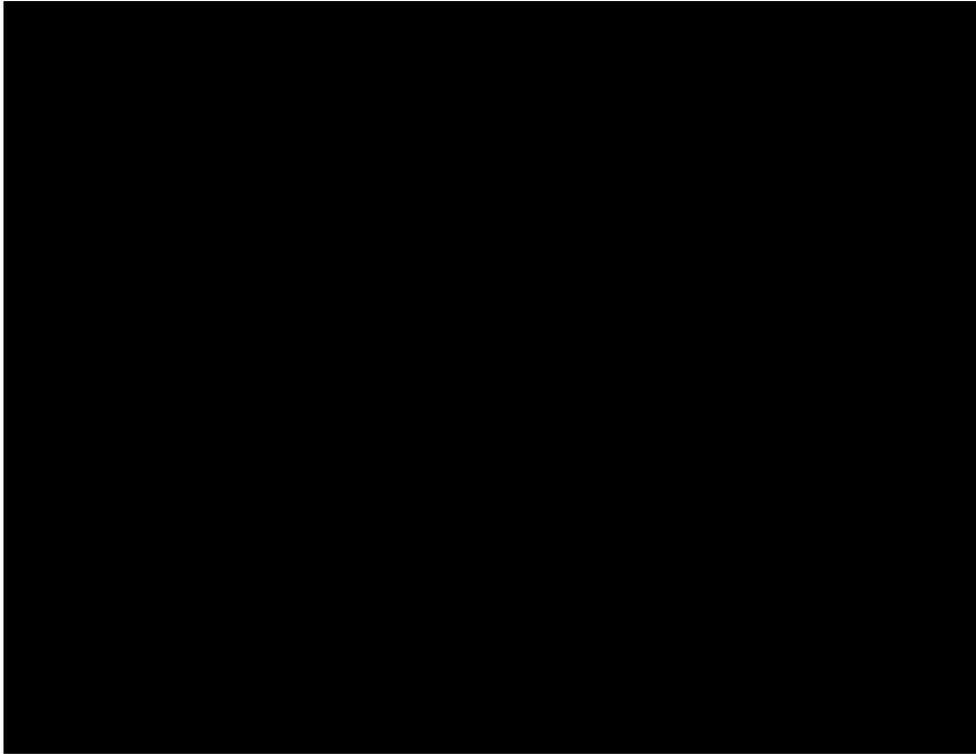


図 4.1: 撮影環境

クシヨソ機能、システム操作ツールバーから構成される。教授者映像では、教授者を中心として撮影された映像が表示される。プレゼンテーション映像には教授者の説明個所と同期したスライドによる教材資料が表示される。教授者の説明個所が詳細にわかるよう、説明に合わせFlash [18] のアニメーション機能を利用したマーキングを行っている。テキスト表示では、教授者が発話した内容がテキスト表示され受講者がノートを取りやすくなっている。テキスト表示は受講者の好みに応じて、表示/非表示の切り替えが行えるようになっている。インデックス機能には各知識単位へのリンクが提供される。インデックス画面を図 4.2、教材画面例を図 4.3 に示す。インタラクシヨソ機能では、教授者とのインタラクシヨソとしてFAQ、学習者同士のインタラクシヨソとして掲示板が提供されている。また本学における学内LANを利用して配信することを前提に、Webベース電子教材の映像、音声ともに256Kbit/sの比較的高品質なデータを利用している。

4.1.2 極低温生成の科学と技術

本節では、本学材料科学研究科の講義「極低温生成の科学と技術」のWebベース電子教材について述べる。今まで開発・利用してきた電子教材とは違い、「極低温生成の科学と技術」は教室における講義と実習室での実験説明で構成されている。第1部は堀 秀信

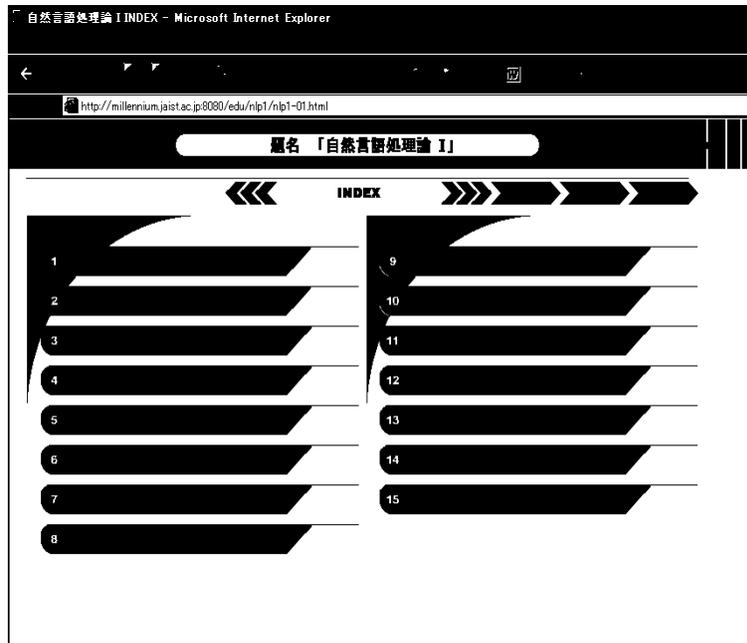


図 4.2: 自然言語処理論 INDEX 画面

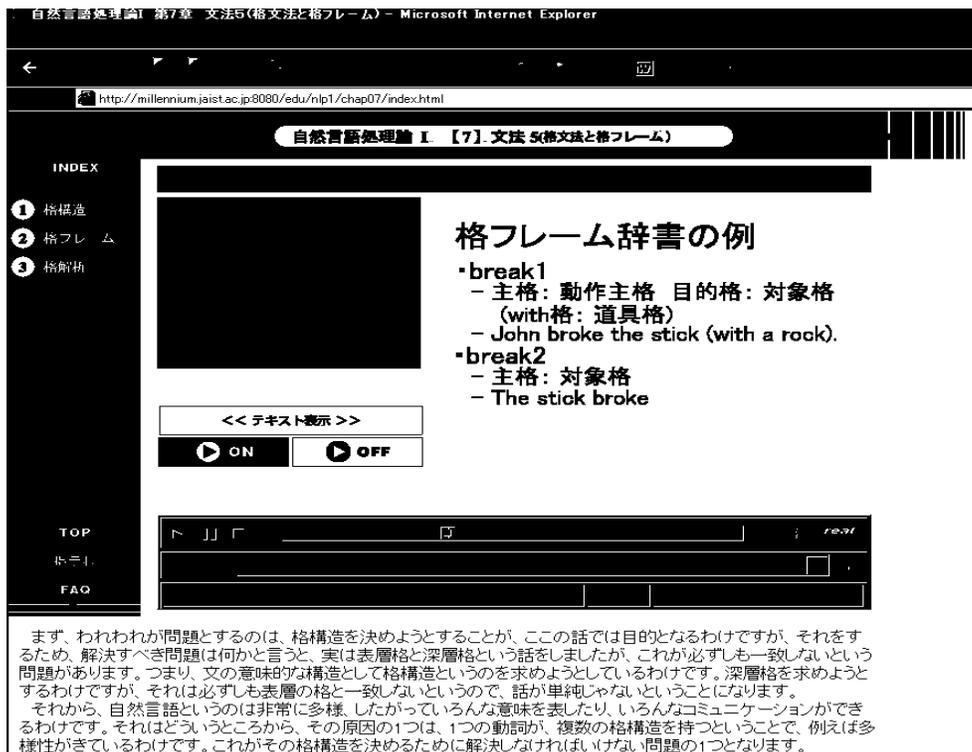


図 4.3: Webベース電子教材 画面例

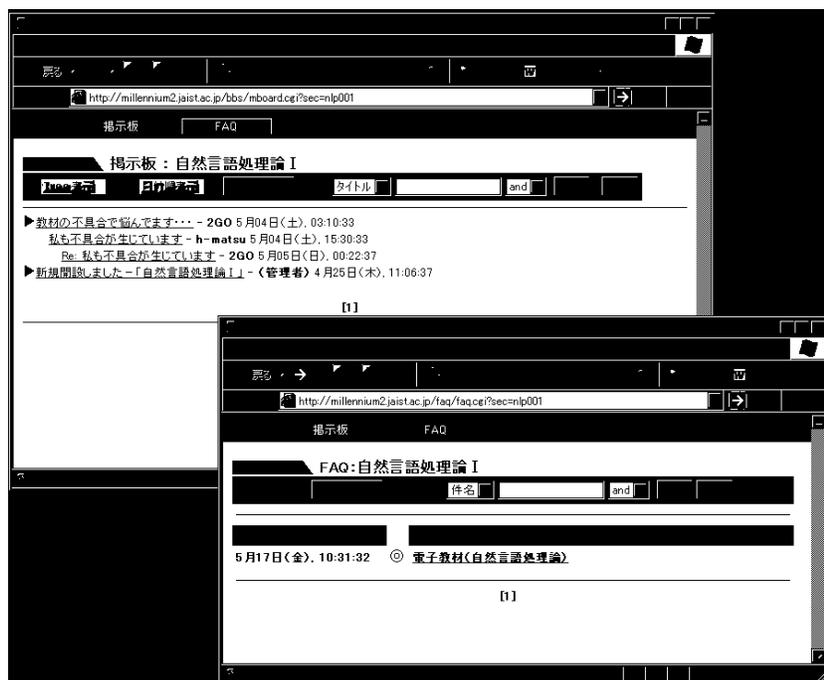


図 4.4: インタラクション画面例



図 4.5: Webベース電子教材 画面例 2

教授による「極低温生成の歴史」についての講義全4回、第2部は岩崎 秀夫 助教授による「極低温科学の世界」についての講義全4回、第3部は本学技官による「極低温生成の技術」についての講義全4回の計12回の講義から構成されている。特に第3部では教授者映像として実験映像を提供し、プレゼンテーション映像には実験原理や装置の内部についての説明資料が表示されており、今回の開発の新しい試みである。実験装置や操作に関する講義部分は教授者映像を大きく、プレゼンテーション映像をやや小さくし、画面配置を変えて設計した。実験説明をWebベース電子教材としたことで、学生は実験装置の細部まで画面上で見ることが可能になり、同時に装置の原理も学ぶことが出来る。実験装置を正確に扱うには十分な知識が必要であり、実習を行う予習としての効果が期待される。今回、この電子教材を用いた実証実験は実施しなかったが、電子教材の可能性を知るためにも今後実験を行う必要がある。

4.2 データ収集システム

教材評価プロセスにおいて、質問紙調査の実施、AHPアンケートの実施をWeb上で行えるようプログラムを実装した。

今まで質問紙調査・AHPアンケート調査ともに、実験実施者が教室に集まってもらった受講者に直接用紙を手渡し、回答データを得ていた。この方法では受講者の反応を見ながら、ほぼ受講者全員のデータ収集が可能であるが、質問紙に記入されたデータの集計プログラムへの入力が必要であり、多人数のデータを扱うにはコストが掛かりすぎている。本システムではコスト削減のため、データ収集がWeb上で実施でき、かつ、収集されたデータは集計プログラムに適用できる形になるようプログラムをJavaを用いて実装した。質問紙調査では、Webベース電子教材に対する反応に関する質問項目のほか、電子教材を受講した環境や予備知識の有無についても問えるよう設計し、また受講者が自由に意見を書き込めるスペースも用意した。それぞれ画面例を図4.6、4.7に示す。



図 4.6: 質問紙調査 画面例



図 4.7: AHP アンケート 画面例

第5章 本評価法の有効性の検証

本章では、Webベース電子教材を改良する上で、本研究において提案するAHP階層図設計法を利用した評価法の有効性について確認することを目的として行った実証実験について述べる。本AHP階層図設計法を用いることにより、評価対象の電子教材に対する改良点を適切に明示した評価基準として選定されているかを確認するために実施した実証実験について報告する。具体的には、開発した電子教材に対して本AHP階層図設計法を用いた評価を適用したことで、評価結果にどのような変化が生じたかを猪俣による実証実験の結果と比較することから考察する。また抽出した改良点において、具体的にどのような改良を行えばよいかについて検討する。

5.1 実証実験 . 1

5.1.1 概要

実証実験1において、提案するAHP階層図設計法を用いて設計したAHP階層図を利用して評価を行った。本学で開講されている東条敏教授による「人工知能特論」の講義の一部として本研究室にて開発されたWebベース電子教材を利用した授業を行った。以下、その方法の詳細について述べる。

1. 実験機器等

- ネットワーク環境
本学の学内LANを利用し、各受講者の端末までの帯域として100Mbpsまでを想定し、受講者全員による同時アクセスにも耐えられるようサーバにも余裕を持たせた構成とした。
- 被験者が利用した端末の機種構成
各被験者が所有しているPC、または各研究室に設置されているPCを端末とした。指定した機種として、CPUはPentium200MHz以上、メモリは64MB以上、OSはWindows95,98,Me,2000,NT4.0,XPとし、Mac, Solarisには対応していない。また、ブラウザはInternetExplorer5.0あるいはNetscape4以降とし、1024x768以上の解像度を推奨した。またどうしてもPCを確保できない受講者には本研究室のNotePCを貸し出すなどの対応を取った。

2. 提示資料

- 教材

本学 情報科学研究科 東条教授による「人工知能特論」を扱った。

- 利用した教材の箇所・範囲

Webベース電子教材と教室講義の比較評価を行うにあたり、授業内容にあまり差がないとされる講義部分を比較評価の対象とした。実証実験は2回に分けて行った。1回目の実験では質問紙調査を実施し、得られた回答データを分析しAHP階層図の設計を行った。2回目の実験では、設計したAHP階層図に基づいたAHPアンケートを実施した。比較評価の対象としては、Webベース電子教材の「論理推論」の第4, 5回を扱い、この回は教室講義を行わなかった。それ以外の講義は従来通り教室講義を行い、その他のWebベース電子教材の講義は受講者が復習に利用できるよう、各教室講義終了後に公開した。

- 時間

教室講義では約90分、電子教材では約50分の授業で構成した。電子教材は、受講者の好きな時間、場所（学内LAN利用限定）で受講可能である。

3. 被験者

- 人数

第1回目の実験では61名、第2回目の実験では26名であった。

第1回目の実験では講義が行われている教室まで行き、質問紙を直接学生に手渡しその場で回答してもらった。第2回目の実験では、本研究で実装したWebアンケートシステムを使用し、自由参加として回答者を募ったため人数が減少した。

- 被験者の学習履歴

被験者に対して行った個別回答によると、61人中16人が「人工知能」に関する講義を受講したことがあると答えた。これは全受講者の26%にあたる。

4. 実験期間

2002年4月16日～6月4日の講義期間に実施した。

5月14日に質問紙調査を実施し、次いで5月29日～6月4日の間にAHPアンケートを実施した。

5.1.2 AHP階層図の設計

評価基準を抽出するために質問紙調査を実施した。質問紙は基本項目と、講義に対する質問項目からなる。基本項目では受講環境や電子教材を受講した経験の有無などを問う。

講義に対する質問項目はN I M E（文部科学省メディア教育開発センター）の「SCSスペースコラボレーションシステム放送利用大学公開講座アンケート調査」[19]をもとに本学のWebベース電子教材の機能に基づいて質問項目を修正、追加したものを、全体で38問の質問として構成した。項目数は被験者の負担を考え、多過ぎないように必要最小限の項目を設定した。質問内容は、「講義に関する項目」、「システムの機能に関する項目」、「受講者の受講態度や感想に関する項目」の大きく3つに分けることができる。質問項目を以下に示す。

講義に関する質問項目

- 講師の話す速さは適当だった
- 講師の話し方は聞き取りやすかった
- 講師の話し方は単調だった
- 講義の内容はおもしろかった
- 講義の内容はよく理解できた
- 講義の内容の質は適切であった
- 講義内容の量は適切であった
- 一回の講義時間を短く感じた

システムの機能に関する質問項目

- 画像は鮮明であった
- 画面に指示された文字はよく見えた
- 画面に指示された図形（描画）はよく見えた
- 音声は明瞭であった
- 音声の音量は適切であった
- 画像と音声のずれが気になった
- 教材のインデックス画面は見やすかった
- 教材のインデックス画面は内容を探しやすかった
- 教材のインデックスを頻繁に利用した
- スライダーバーは便利だった
- スライダーバーを頻繁に利用した
- ビデオ映像画面が小さかった
- OHP画面よりビデオ映像画面に注目して視聴した
- OHP画面上のアニメーション機能はわかりやすかった

- 掲示板がある方が何かと都合良かった
- F A Qがある方が何かと都合良かった
- テキスト表示ビューワは役に立った
- テキスト表示ビューワのリンク機能は役に立った
- 自分に合った時間・空間で受講できた

受講者の受講態度・感想に関する質問項目

- 受講中ノートを取りながら視聴した
- 講義に集中して視聴した
- リラックスして視聴した
- 講義として臨場感や現実感があった
- 講義として充実感が得られた
- 教室講義との違いを大きく感じた
- 全体的に満足のいく講義だった
- 実際に役立つ講義だった
- 復習に役立った
- 教室講義がなく、電子教材だけでは内容は不十分だった
- これからもこのようなシステムを利用したい

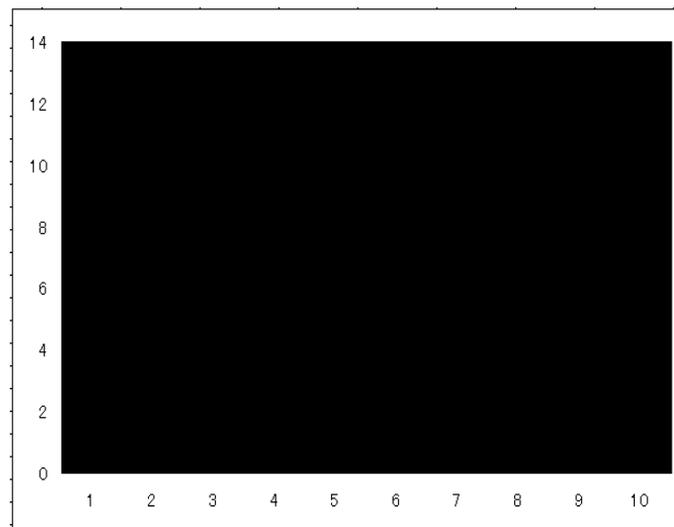


図 5.1: スクリープロット

	因子								
	1	2	3	4	5	6	7	8	9
Q10	.276	.118	.640	7.432E-02	2.767E-02	-3.806E-03	-9.152E-02	1.738E-02	2.764E-02
Q11	.213	.360	.630	-7.710E-02	.106	-1.708E-02	8.969E-02	.177	-3.513E-02
Q12	-478	4.654E-02	-2.830E-02	.149	2.262E-02	3.921E-02	.347	.635	-5.574E-02
Q13	.665	5.271E-02	.198	-.104	.145	.362	6.917E-02	-1.672E-03	-.424
Q14	.645	.139	.292	5.927E-02	.248	.235	7.536E-02	.109	-.182
Q15	.475	.208	.411	.258	6.384E-03	.342	8.375E-02	.163	-5.252E-02
Q16	.455	.266	.470	.199	7.260E-02	.409	.122	2.901E-02	.202
Q17	.270	.105	-2.504E-02	.147	4.376E-02	.858	.209	9.057E-03	2.407E-02
Q18	.110	.635	.179	.143	.392	-.106	-6.074E-02	.127	8.179E-02
Q19	.314	.696	3.770E-02	-9.854E-03	9.225E-02	7.922E-02	-.107	5.208E-02	2.861E-02
Q20	.296	.795	-7.487E-02	1.071E-02	.126	7.708E-02	-.167	-8.488E-02	2.741E-02
Q21	4.147E-02	.712	.461	-.104	4.773E-02	.111	.105	9.030E-02	2.620E-02
Q22	.150	.698	.376	9.857E-02	.153	.100	.138	-3.163E-02	-4.919E-02
Q23	-.102	-8.391E-02	-.130	.583	-6.106E-02	-.111	.210	-8.089E-02	-4.774E-02
Q24	.622	.257	.310	1.970E-02	.161	2.821E-04	1.276E-02	.162	8.826E-02
Q25	.732	.255	6.791E-02	-9.312E-02	5.478E-02	3.325E-02	6.953E-02	.149	-6.919E-03
Q26	.406	.102	2.067E-02	.268	.209	-6.294E-02	.154	.362	-.105
Q27	.335	.261	9.437E-02	.149	.528	.166	-5.163E-02	9.322E-02	.238
Q28	.176	.298	.162	-3.382E-02	.583	.255	-.235	.142	.238
Q29	-1.552E-02	-.309	-.139	-3.774E-02	.101	5.326E-02	.615	.177	-.181
Q30	.106	.145	-8.931E-03	7.507E-02	.528	-8.390E-02	.167	.119	-.179
Q31	.165	2.089E-02	.190	.111	.117	3.400E-02	-1.164E-02	.756	.159
Q32	.256	7.593E-02	.162	.764	.184	.220	-.131	.110	.137
Q33	.182	6.106E-02	.104	.735	6.449E-02	.213	-7.059E-02	.299	-9.832E-02
Q34	.189	2.891E-02	.402	.155	.251	4.052E-03	-.375	.210	-9.547E-02
Q35	.215	.118	.275	.534	.329	-3.640E-03	-.397	8.447E-02	.192
Q36	.577	-3.649E-02	.315	5.006E-02	.178	.169	.227	-4.417E-02	.216
Q37	.357	6.239E-02	7.824E-03	-3.889E-02	7.596E-02	3.575E-02	-5.911E-02	.104	.712
Q38	.595	5.123E-02	.319	4.843E-02	.155	.194	-9.999E-02	5.314E-02	.221
Q39	.205	.109	.207	4.615E-02	3.545E-02	.235	.538	4.416E-02	5.099E-02
Q40	.673	.134	.271	.184	-1.813E-02	.235	-2.063E-02	-6.965E-03	.361
Q41	.794	.163	.226	.216	.178	.157	4.901E-02	.106	2.960E-02
Q42	-.635	-4.074E-02	-3.248E-03	-2.184E-02	2.980E-02	1.174E-02	.143	8.286E-02	-4.212E-02
Q43	.711	.264	.142	.216	.322	.146	3.528E-02	9.688E-02	.157
Q44	.709	.273	.324	.238	.223	-5.648E-02	.146	-1.489E-02	.210
Q45	.372	.261	.326	.174	.360	6.280E-02	.344	-.168	.113
Q46	-.650	-.334	5.556E-02	-.124	-8.984E-02	-.147	4.390E-02	-9.747E-02	-9.444E-02
Q47	.629	.122	.170	4.325E-02	.512	6.789E-02	8.727E-02	-.203	.101

図 5.2: 因子分析 結果画面

各質問項目に対し「かなりあてはまる」「ややあてはまる」「あてはまる」「どちらともいえない」「あてはまらない」「ややあてはまらない」「かなりあてはまらない」の7件法により回答してもらった。質問紙調査で用いられるのは5件法が一般的であり、猪俣の実証実験でも5件法が使われていたがデータのバラツキが小さく、主因子分析を行う際にはバラツキが大きい方が因子の抽出が行いやすいので、今回の実証実験では7件法を用いた。

以下、第3章で述べた提案する手法に基づきAHP階層図を設計するための評価基準の抽出を行った。

主因子分析法は統計ソフトSPSSを用いて行った[20][21][22]。主因子分析を行った結果画面を図5.2に示す。因子数の決定には固有値の値が1.0を超えておりスクリープロット(図5.1)による変動量から検討して因子数を3とした。図5.1のスクリープロットの縦軸は固有値の値、横軸は因子番号を示している。スクリープロット図から固有値の値が因子番号3から4の間で大きく落ち込んでいることがわかる。抽出された3つの因子の特性を分析するため、因子負荷量ごとに分類を行った。また、各因子負荷量の二乗和を項目数で割ったものを寄与といい、寄与率はそれを百分率で表したものである。すなわち寄与率は、変数全体の分散のうち、各因子が説明できる分散の大きさを表す。今回の3つの因子では、全体の分散の半分(51%)が説明できることになる。図5.2の結果から、抽出され

た3つの因子を学習者が講義に対して何を求めているのかがわかり、学習者の「反応」に基づいた評価ができるよう考え命名した。以下それぞれ命名した因子について述べる。

第1因子

講義に対する充実感や満足度、集中して受講できたかなどの負荷が高く、受講者が一番強く講義に求めているものであると考えられたので「講義に対する集中力」と命名した。ここで「満足度」と命名しなかったのは、そう命名することで講義そのものに対する感想を測ることにつながり、受講者が重要度を決定する際に影響を及ぼす恐れがあるからである。

第2因子

画像や音声、スライダーバーに関するものなど、受講者の身体に対して物理的な影響を与えている因子として「身体的疲労度」と命名した。

第3因子

講師の話し方が単調であったか、教授者や他の学習者とのインタラクションの容易性、リラックスして受講したかなどの負荷が高く、他者との関係から受ける影響を表す因子として「講義に対する緊張感」と命名した。

因子に命名するにあたり、本研究では「集中力」と「緊張感」とを分けて考えた。「集中している状態」＝「緊張している状態」と考えられることもあるかもしれないが、「緊張」とは心や身体が引き締まること、慣れない物事などに直面して心が張りつめること、精神的あるいは感情的な負荷、どっちつかずの気持、不安あるいは興奮を言う。ここでの不安とは、現実あるいは想像からくる威嚇、何となく容易でないという感じを言う。つまり講義に集中している状態と、他者との関係や講義内容を理解できるかどうか等から感じる緊張感は別のものであると考えられる。

上記のように、学習者の「反応」を測るための因子抽出を行った。次に、学習者の「反応」に影響を与えている因子の抽出を行った。電子教材を構成しているのは「教授者」と「システム」であり、教材改良を実施するのも教授者とシステム設計者の二者であることから、レベル3には「教授者のプレゼンテーション」と「学習環境」の2つの因子を配置する。次いで、質問項目を「教授者のプレゼンテーションに関する項目」と「学習環境に関する項目」とに分け、同様にして主因子分析を行いレベル4に配置する因子の抽出を行った。

「教授者のプレゼンテーションに関する項目」では3つの因子が抽出され、その寄与率は79.19%を示しており全体の分散の約80%を表していることがわかる。

第1因子

講義内容の量、質に関する項目の負荷が高く、講義そのものを評価するものとして「内容のまとめ方」と命名した。

第2因子

講師の話し方が単調であるか、講義をおもしろいと感じるかなどの項目が含まれた。教授者のプレゼンテーション能力に左右される項目と考えられ、「プレゼンテーションの仕方」と命名した。

第3因子

教授者の話し方の早さや、口調の聞き取りやすさに関する項目の負荷が大きく「話し方・口調」と命名した。

「学習環境に関する項目」に関しても同様に主因子分析を行い3つの因子を抽出した。抽出された3つの因子の寄与率は53%を示していた。

第1因子

スライダーバーや画面の見やすさに関する項目の負荷が大きく、システムの使いやすさに関するものとして「活用されている教材・機材」と命名した。

第2因子

掲示板やFAQなど、学習者間や教授者とのインタラクションに関する項目の負荷が大きく「他者とのインタラクション容易性」と命名した。

第3因子

画像や音声に関する項目の負荷が小さく、インデックスやテキスト表示、時間・空間の自由度などの項目の負荷が大きかった。学習の進めやすさ、好みに応じて変化できることを示す因子であることから「学習方法の自由度」と命名した。

上記のように二段階にわけてレベル2、レベル4に配置する評価基準をそれぞれ抽出した。AHP評価の問題を「効果的な学習をするために、あなたはどちらの学習形態を選択しますか?」と定義し、この問題に対する代替案として「Webベース電子教材」と「教室講義」を最終的な候補として得られた因子をもとにAHP階層図の設計を行った。レベル2には学習者の「反応」をもとに主因子分析により抽出した因子、レベル3には電子教材を構成する「教授者」と「学習環境」という2つの要因、レベル4にはレベル3をもとに主因子分析により抽出した因子を配置し、設計したAHP階層図を図5.3に示す。ここで設計したAHP階層図において、本研究の目的であった評価基準間の相関関係を弱めることができたのかを調べた。以下に結果を示す。

レベル2

- 「講義に対する集中力」と「講義に対する緊張感」… 約0.1
- 「講義に対する集中力」と「身体的疲労度」…… 約0.1
- 「講義に対する緊張感」と「身体的疲労度」…… 約0.05

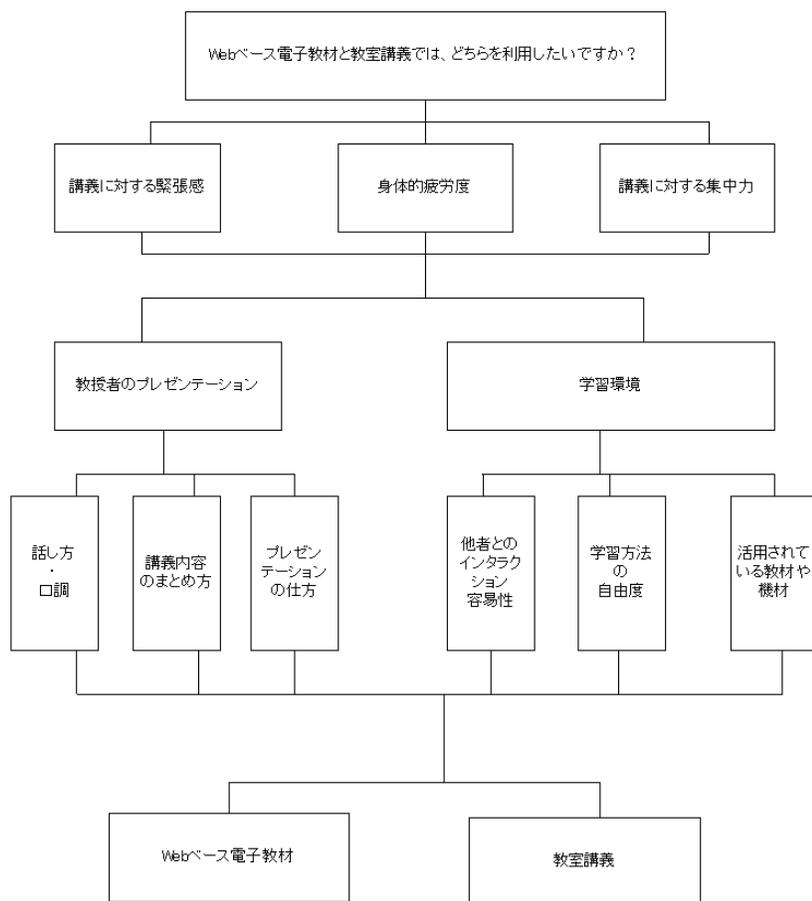


図 5.3: 設計したAHP階層図

レベル4

- 「プレゼンテーションの仕方」と「話し方・口調」… 約0.1
- 「プレゼンテーションの仕方」と「内容のまとめ方」… 約0.1
- 「話し方・口調」と「内容のまとめ方」…… 約0.2
- 「他者とのインタラクション容易性」と「活用されている教材・機材」… 約0.2
- 「他者とのインタラクション容易性」と「学習方法の自由度」… 約0.1
- 「活用されている教材・機材」と「学習方法の自由度」…… 約0.1

猪俣により設計されたAHP階層図と比べると、全体的に相関関係が弱まっているのがわかり、評価基準の重み付けにおいて、一対比較行列の矛盾が全くないときの値に一步近づいたといえる。しかしながら、相関係数がどの程度であれば、最終的な評価基準の重み付けにおいて影響が出ないのかという点については明確になっていない。今後、相関係数と重み付けの関係についてはさらに検討する必要がある。

5.1.3 重要な評価基準の抽出

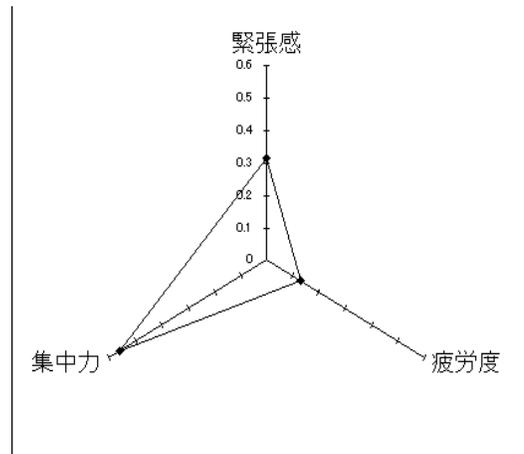


図 5.4: レベル2 評価基準ごとのウエイト割合

「人工知能特論」における第2回目の実証実験では、設計したAHP階層図をもとにしたAHPアンケートを本研究において実装したWebアンケート収集システムを利用して実施した。受講者には10日間のアンケート公開中に、一人一回参加してもらうよう呼びかけ、25名分の回答データを得た。そのうちAHP整合度(C.I値)が0.1を超えていた3名のデータを集計から除外した。はじめに、幾何平均によるグループ全体の各評

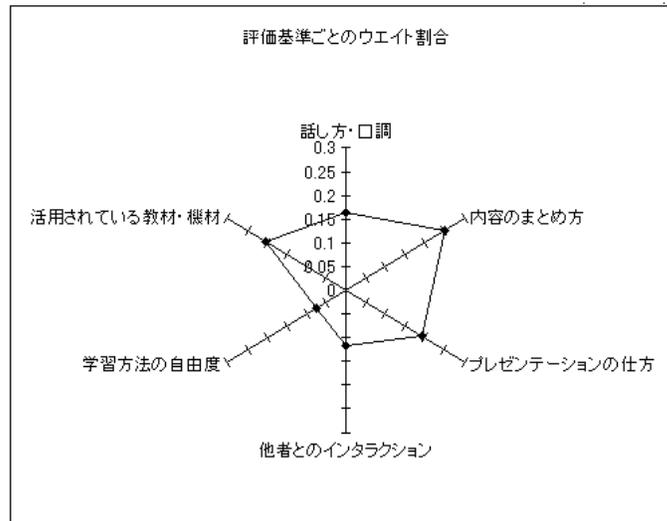


図 5.5: レベル4 評価基準ごとのウエイト割合

評価基準のウエイトを求めた。グループでの意思決定にAHPを使う場合、一対比較行列値としてグループを構成するメンバーの値の、幾何平均を用いる。一対比較行列において、対称の位置の数値が逆数を示すことが前提であり、算術平均は適さないからである。各評価基準のウエイトをレーダーチャートで示す。レベル2では学習者が講義に何を求めているのかがわかるように、評価基準を設定していたので、図5.4からは学習者は集中できる講義を強く望んでいることがわかる。これは、Webベース電子教材を利用する上で「目が疲れる」や「ヘッドホンなどを使用するため耳が疲れる」などの問題点が挙げられるが、集中し緊張して受講できる講義であれば疲労度はさほど問題にならないことを示している。また、レベル3の重み付けでは、「教授者のプレゼンテーション」(0.61) > 「学習環境」(0.39)(カッコ内はウエイト)となり、学習者が講義を評価する際、重視するのは「教授者による部分」が大半であることがわかる。図5.5から「2.内容のまとめ方」「6.活用されている教材・機材」「3.プレゼンテーションの仕方」の順でウエイトが高く、次いで「1.話し方・口調」「4.他者とのインタラクション」「5.学習方法の自由度」となった。次に全体を100点として評価基準のウエイトごとに得点を分配しWebベース電子教材と教室講義それぞれの得点を算出した。また平均得点の比較を行うため、その比較の有意差を分散分析F検定により調べた。それらの結果を表5.1に示す。得点に有意差が見られたものにだけ大小関係が示してある。また実験結果を特性表にまとめ表5.2に示す。特性表とは、得点と分散分析F検定の結果から評価基準をカテゴリに分類したものである。ここで得点の高低を決める基準値は、100点満点を評価基準数で割った数値である。得点が高く、かつ統計的に有意差のあるエリアの評価基準のうち「評価対象の電子教材」が比較対象の「他の講義形態」よりも得点の低いものを電子教材改良の重要な評価基準であるとし、これをもとに電子教材の改良を行っていく。同時に質問紙調査での自由記述による回答からも改良点を検討する。得点は高いが有意差がないエリアの評価基準は、効果的な

学習をする際に影響を与えるパラメータとする。

表 5.1: 評価基準ごとの得点 F 検定 > f (1,43,0.05)

評価基準	Web		教室講義	F 検定	得点
2 . 内容のまとめ方	9.399		15.616	2.915 < F	25.02
6 . 活用されている教材・機材	8.411	<	11.753	11.005 > F	20.16
3 . プレゼンテーションの仕方	6.14	<	13.145	8.246 > F	19.28
1 . 話し方・口調	5.4	<	11.04	12.431 > F	16.44
4 . 他者とのインタラクション容易性	4.211	<	7.497	5.979 > F	11.71
5 . 学習方法の自由度	5.466	>	1.921	23.208 > F	7.39
合計	39.027		60.972	11.43 > F	100

表 5.2: 実証実験 1 特性表

	統計的に有意差あり	統計的に有意差なし
高い得点 (> 16.67)	3 , 6	2
低い得点 (< 16.67)	1 , 4 , 5	

今回の実験では、得点が高くかつ得点間の比較において有意差がみられた評価基準は「活用されている教材・機材」と「プレゼンテーションの仕方」であった。これらは効果的な学習をする際に影響を与え、かつ電子教材と他の講義形態間の差を特徴付けるパラメータであるとみなす。得点は低い有意差が見られた評価基準は「話し方・口調」「他者とのインタラクション容易性」「学習方法の自由度」であった。得点は高かったが比較に有意差が見られなかった評価基準は「内容のまとめ方」であった。これは受講者にとって「内容のまとめ方」は重要であるが、同じ教授者である限り、電子教材と教室講義の間で差はないということを示している。

5.1.4 実証実験 1 : 考察

AHP階層図の設計法を変更したことで、結果にどのような変化があらわれたのか、猪俣の実験結果 [[2] から抜粋 : 表 5.3 に示す] と比較した。今回の実験の評価基準数が猪俣のもの比べて少ないのは、AHP階層図を学習者の「反応」に関するレベルと、教授者・システムに関する項目との二段階に分けたからである。2つの表を比較するため、同じよ

うな内容を示す猪俣の評価基準に今回の実験と同じ番号を振り当てた。以下に猪俣の用いた評価基準を示す。

1. 話し方や口調の聞き取りやすさ
2. 内容の整理やまとめ方
3. 受講者の理解度に配慮したプレゼンテーション
4. 他者とのインタラクション
5. 講義の進展に応じた学習の進めやすさ
6. 活用されている教材・機材
7. 目の疲れやすさ
8. 耳の疲れやすさ
9. 学習への集中しやすさ
10. 獲得した知識の興味・関心
11. 理解したことによる充実感
12. 獲得した知識の応用意欲

表 5.3: 猪俣による特性表

	統計的に有意差あり	統計的に有意差なし
高い得点 (> 8.33)	6,7,8	2,9,11
低い得点 (< 8.33)	3	1,4,5,10,12

猪俣によるAHP階層図では、「目の疲れやすさ」「耳の疲れやすさ」は、「充実感」など学習者の満足度に関する項目との相関が強く、AHPでは相関が大きい評価基準があるとその重要度は本当の値よりも大きくなるという性質があるため、それらの得点が高くなり、逆に相関の弱い「話し方・口調」「プレゼンテーションの仕方」の得点が低くなるという問題があった。表5.2と表5.3とを比較すると、評価基準間の相関関係が弱まったことで、特性表における評価基準の移動がおこった。今回の実験では得点が低く有意差がないカテゴリに属す評価基準がなくなり、評価対象のWebベース電子教材と教室講義との差を明確にする評価基準が得られたと考えられる。さらに、統計的に有意差のある評価パラメータを多く抽出することができた。「目の疲れやすさ」「耳の疲れやすさ」は、学習者

にとって重要度が低いことがわかっているので、評価対象の電子教材の改良点は「活用されている教材・機材」「プレゼンテーションの仕方」、次いで「話し方・口調」であることがわかる。階層図を二段階に分け、学習者が判断する「学習への集中しやすさ」などを他の評価基準より上の階層に配置したことで、学習者の主観的判断をもとにした評価を、電子教材と教室講義という異なった講義形態の間で同じ評価基準を利用して行うことが可能となった。

5.2 実証実験 . 2

5.2.1 概要

実証実験 2 において、実証実験 1 で設計した A H P 階層図が異なる電子教材に対しても利用可能かどうかを検証するために、同 A H P 階層図を用いた評価を行った。本学で開講されている片山 卓也 教授による「ソフトウェア設計論」の講義の一部として、本研究室にて開発された W e b ベース電子教材を利用した授業を行った。以下、その方法の詳細について述べる。

1. 実験機器等

- ネットワーク環境
本学の学内 L A N を利用し、各受講者の端末までの帯域として 100Mbps までを想定し、受講者全員による同時アクセスにも耐えられるようサーバにも余裕を持たせた構成とした。
- 被験者が利用した端末の機種構成
各被験者が所有している P C、または各研究室に設置されている P C を端末とした。指定した機種として、C P U は Pentium200MHz 以上、メモリは 64MB 以上、O S は Windows95,98,Me,2000,NT4.0,XP とし、Mac , Solaris には対応していない。また、ブラウザは InternetExplorer5.0 あるいは Netscape4 以降とし、1024x768 以上の解像度を推奨した。またどうしても P C を確保できない受講者には本研究室の N o t e P C を貸し出すなどの対応を取った。

2. 提示資料

- 教材
本学 情報科学研究科 片山教授による「ソフトウェア設計論」を扱った。
- 利用した教材の箇所・範囲
W e b ベース電子教材と教室講義の比較評価を行うにあたり、授業内容にあま

り差がないとされる講義部分を比較評価の対象とした。比較評価の対象としては、Webベース電子教材の第3回「オブジェクトモデル2」と第5回「動的モデル1」を扱い、この回は教室講義を行わなかった。それ以外の講義は従来通り教室講義を行い、その他のWebベース電子教材の講義は受講者が予習・復習に利用できるよう、実験初日に全て公開した

- 時間

教室講義では約90分、電子教材では約80分の授業で構成した。電子教材は、受講者の好きな時間、場所（学内LAN利用限定）で受講可能である。

3. 被験者

- 人数

18名による。

- 被験者の学習履歴

被験者に対して行った個別回答によると、18人中4人が「オブジェクト指向技術」に関する講義を受講したことがあると答えた。

4. 実験期間

2002年10月3日～11月28日の講義期間中に実施。

10月31日～11月11日の間にAHPアンケートを実施した。

5.2.2 重要な評価基準の抽出

「ソフトウェア設計論」における実証実験では、設計したAHP階層図をもとにしたAHPアンケートを本研究において実装したWebアンケート収集システムを利用して実施した。受講者には12日間のアンケート公開中に、一人一回参加するよう呼びかけ、18名分の回答データを得た。そのうちAHP整合度(C.I値)が0.1を超えていた1名のデータを集計から除外した。

各評価基準のウェイトをレーダーチャートで図5.6に示す。レベル3の重み付けでは、「教授者のプレゼンテーション」(0.56) > 「学習環境」(0.44)(カッコ内はウェイト)となり、学習者が講義を評価する際重視するのは、実証実験1と同じく「教授者による部分」が大半であることがわかる。図5.7から「内容のまとめ方」「活用されている教材・機材」「プレゼンテーションの仕方」の順でウェイトが高く、次いで「他者とのインタラクション」「学習方法の自由度」「話し方・口調」となった。次に全体を100点として評価基準のウェイトごとに得点を分配しWebベース電子教材と教室講義それぞれの得点を算出した。また平均得点の比較を行うため、その比較の有意差を分散分析F検定により調べた。それらの結果を表5.4に示す。得点に有意差が見られたものにだけ大小関係が示してある。また実験結果を特性表にまとめ表5.5に示す。

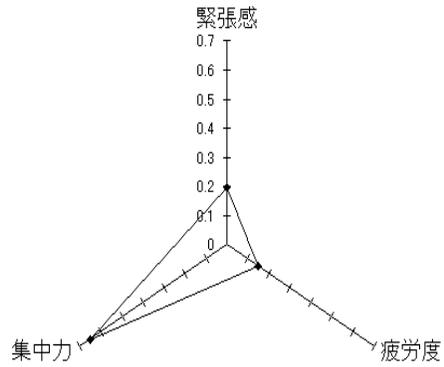


図 5.6: レベル 2 評価基準ごとのウエイト割合

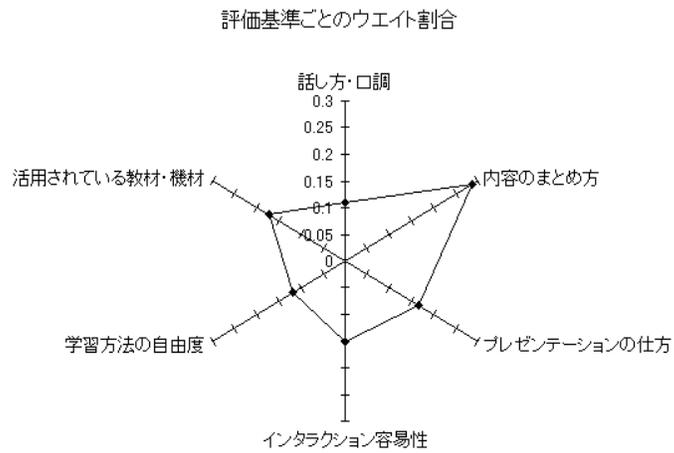


図 5.7: レベル 4 評価基準ごとのウエイト割合

表 5.4: 評価基準ごとの得点 F 検定 > f (1,32,0.05)

評価基準	Web		教室講義	F 検定	得点
2 . 内容のまとめ方	17.688		10.968	2.475 < F	28.66
6 . 活用されている教材・機材	10.395	>	6.811	4.434 > F	17.21
3 . プレゼンテーションの仕方	8.126		8.251	0.11 < F	16.38
4 . 他者とのインタラクション容易性	8.532		6.588	0.767 < F	15.12
5 . 学習方法の自由度	9.768	>	1.949	195.987 > F	11.71
1 . 話し方・口調	4.611		6.313	2.656 < F	10.92
合計	59.12		40.88	16.222 > F	100

表 5.5: 実証実験 2 特性表

	統計的に有意差あり	統計的に有意差なし
高い得点 (> 16.67)	6	2
低い得点 (< 16.67)	5	1 , 3 , 4

得点が高く、かつ得点間の比較において有意差がみられた評価基準は「6. 活用されている教材・機材」であり、効果的な学習をする際に影響を与え、電子教材と他の講義形態間の差を特徴付けるパラメータであるとみなす。得点は低い有意差が見られた評価基準は「5. 学習方法の自由度」であり、電子教材と教室講義との差を特徴付けるパラメータであるとみなす。本来の目的は、AHPを用いて教材の改良すべき点を抽出することであり、その手法は講義形態間の得点に有意差があり、かつ「評価対象の電子教材」が「他の講義形態」よりも得点が低いものを改良点として抽出することにあつた。しかし、今回の実験では、合計得点がWebベース電子教材 (59.12) > 教室講義 (40.88) となり、電子教材の方が得点が高くなり、同時に有意差ありとなった「6. 活用されている教材・機材」「5. 学習方法の自由度」の2つの評価基準はともに、教室講義よりもWebベース電子教材の得点が高くなったため、教材改良点とはならなかった。このような結果になった場合、現段階では質問紙調査での自由記述による回答から教材の改良点を検討するしか対処する方法がない。また「統計的に有意差なし」となった評価基準が、実証実験1では無かったが、今回の実験では「1. 話し方・口調」「3. プレゼンテーションの仕方」「4. 他者とのイン

タラクション容易性」の3つに増えた。次節においてこれらの原因を探り、教材の改良点を得るための次段階の手法を考察する。

5.3 考察

実証実験1「人工知能特論」の結果と、実証実験2「ソフトウェア設計論」の結果を比較し、その違いを考察する。実証実験1では最終的な得点は教室講義(60.07) > Webベース電子教材(39.03)であったが、実証実験2ではWebベース電子教材(59.12) > 教室講義(40.88)となり、レベル4の評価基準ごとのウエイトには大きな違いはなかったが、代替案の得点で大きな差が出た。その結果、実証実験2において改良点の抽出に至らなかった。本節では、得点に大きな差が出た原因を探るため、クラスター分析を行った。

5.3.1 クラスター分析とは

クラスター分析とは、異なる性質のもの同士が混ざり合っている集団(対象)の中から、互いに似たものを集めて集落(クラスター)を作り、対象を分類しようという方法を総称したものであり、数値分類法とも呼ばれる。対象のもつ性質を距離や似たものの度合いで調べ、クラスターに分類する[23]。一般的によく利用される階層的クラスター分析では、対象間の非類似度(値の小さい方が類似性が高いことを表わす)を手掛かりにして、樹形図(デンドログラム)を構成することが目的である。クラスターを融合する際に基準となる方法には、最短距離法、最長距離法、群平均法、重心法、メジアン法、ワード法、可変法など様々な手法がある。ワード法には、分類感度が高く最も明確なクラスターを作るという特徴があり、本研究ではユークリッド平方距離によるワード法でクラスター分析を行い母集団の分類を行った。

以下、ワード法によるクラスター分析について説明する。

n 個の個体について、 p 個の変数 $X_{i1}, X_{i2}, \dots, X_{ip} (i = 1, 2, \dots, n)$ があるとする。初期状態として、 n 個のクラスターがあるとする(各クラスターは1個体ずつを含むと考える)。このとき、ユークリッド平方距離は、

$$d_{ij}^2 = \sum_{k=1}^p (X_{ik} - X_{jk})^2 \quad (i, j = 1, 2, \dots, n)$$

により、定義される。

ユークリッド平方距離の最も近いクラスターを併合して、1つのクラスターとする。クラスターaとクラスターbが併合されてクラスターcが作られるとする。 d_{ab}, d_{xa}, d_{xb} を、クラスターaとクラスターbが併合される前の各クラスター間の距離としたとき、併合後のクラスターcとクラスターx($x \neq a, x \neq b$)との距離は式5.1で表される。

$$d_{xc}^2 = \frac{(n_x + n_a)}{(n_x + n_c)} d_{xa}^2 + \frac{(n_x + n_b)}{(n_x + n_c)} d_{xb}^2 - \frac{n_x}{(n_x + n_c)} d_{ab}^2 \quad (5.1)$$

2 個のクラスターが 1 個のクラスターにまとめられたので、総クラスター数が 1 個減る。クラスター数が 1 になるまで式 5.1 を繰り返す。構成されたデンドログラムは、どこでクラスターを切断するかによって分類の仕方が違ってくる。そのため最終的な解析は、分析者の知識や解釈の仕方に左右される。

5.3.2 クラスター分析結果

2 回の実証実験で得た各評価基準に対する Web ベース電子教材の得点を変数とし、被験者をクラスター分析により分類した。クラスター分析には、ユークリッド平方距離によるワード法を用いた。実証実験 1 「人工知能特論」では大きく二つのクラスター（クラスター A は 6 人、クラスター B は 16 人）に分かれた。それぞれのクラスター別に評価基準ごとの平均得点を求め、図 5.8 にグラフに表した。クラスター A は、全体的に得点が高く電子教材に対して肯定的な意見の被験者であり、またクラスター B では全体的に得点が低く、電子教材に対して否定的な意見の被験者であると言える。今回の分析により、このように同じ教材を利用し評価が分かれる原因までは知ることはできないが、受講環境や当該科目に関する基礎知識の有無が関わっていると推測される。

実証実験 2 「ソフトウェア設計論」でも、同様に二つのクラスター（クラスター C は 8 人、クラスター D は 9 人）に分かれた。それぞれのクラスター別に評価基準ごとの平均得点を求め図 5.9 に示す。実証実験 1 との大きな違いは、クラスター別の平均得点に大きな差がないことである。評価基準「4. 他者とのインタラクション容易性」においてのみ得点の高低差がはっきりと表れてはいるが、その他の評価基準ではほとんど差が出ていない。つまり「4. 他者とのインタラクション容易性」の得点によって、クラスターが分かれていることになる。全てのクラスターをまとめて図 5.10 に示す。図 5.10 から実証実験 2 のクラスター C, D とともに、実証実験 1 において電子教材に対して肯定的な意見を持つクラスター A と得点の付け方が類似していることがわかる。つまり実証実験 2 では、受講者全体が電子教材に対して肯定的な意見を示したために、最終的な代替案選択において、電子教材 > 教室講義 となり、教材の改良すべき点を抽出するに至らなかったのである。また、クラスター間で一番大きく得点差が表れているのは「4. 他者とのインタラクション容易性」であることがわかる。これは、電子教材では視聴しながら学習者同士での相談が可能であることから、学習者同士のインタラクションを重視する受講者は高い得点を付け、教授者とのインタラクションを重視する受講者は、その場で質疑応答が可能な教室講義に高い得点を付けたからではないかと推測される。

以上のような結果から、受講者の意識の違いが最終的な得点付けに大きな影響を与え、教材の改良点抽出においても影響を与えていることがわかった。クラスター間の得点差が大きかった「他者とのインタラクション容易性」に関して教材改良を行ったなら、電子教材の最終的な得点は上がることが予想され、改良の余地があるといえる。しかし、このような受講者の意識の違いを、どのように教材の評価、改良に活かしていくのかは今後検討すべき課題である。

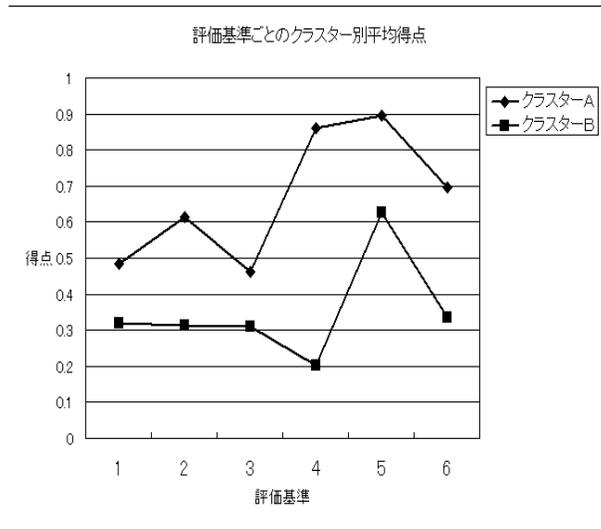


図 5.8: 実証実験 1 クラスター分析結果

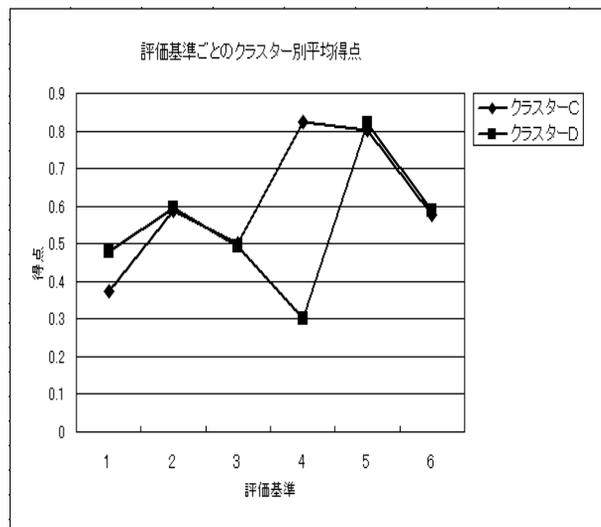


図 5.9: 実証実験 2 クラスター分析結果

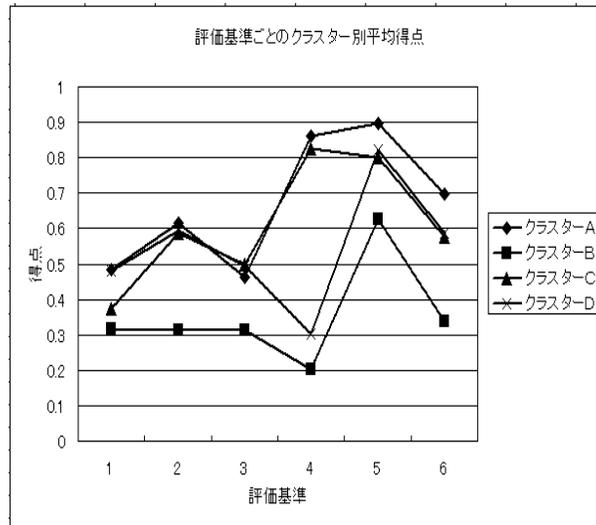


図 5.10: クラスタ分析 結果比較

5.3.3 まとめ

今回の実証実験1では、AHPを用いた評価基準の重み付けと電子教材と教室講義との得点間の有意差から「活用されている教材・機材」、「プレゼンテーションの仕方」が改良すべき個所として選定された。また実証実験2では、電子教材が教室講義よりも得点が高くなったことからAHP評価では改良すべき個所が選定されなかったが、受講者をクラスターに分けたときの得点の差異から改良の余地があるのは「他者とのインタラクション容易性」であるとわかった。以上のことから、電子教材へのフィードバックに一步近づいたといえるが、実際にこれらに関して改良を行うとき、具体的にどの機能・個所をどのように改良すればよいかという点は明確になっていない。これは今後考察すべき重要な課題であるが、一つの手法として、具体的な改良の方針については質問紙調査の自由記述文を用いることが考えられる。自由記述文には受講者が思い思いのことを自由に記述するため、非定型であいまいであり、教材改良に使える有益なデータと有益ではないデータが混在する。そこでAHP評価法の補助的な役割として、自由記述文を評価基準に基づきスクリーニングすることを考える。例えば、実証実験2の結果のように、クラスター分析により「他者とのインタラクション容易性に関して改良の余地あり」と求められた場合、その評価基準に対して低い得点を付けた受講者の意見を取り入れ改良することにより、全体的な電子教材の得点が上がることが予想される。つまり、受講者に実施した質問紙調査で得た回答から、提案したAHP評価法により改良すべき箇所として選定された評価基準や、クラスター分析で改良の余地があると求められた評価基準に対して、低い得点を付けた受講者の意見を参考に改良箇所を詳細化し、改良を行う際の方針が得られると考えられる。

第6章 おわりに

本研究において学習者の主観的判断をもとにした評価が行え、かつ教材の改良箇所が明確に選定できる電子教材評価法のAHP階層図設計法を提案した。評価プロセスは、まず電子教材受講者に対し質問紙調査を行い、回答データから主因子分析法を用いて評価基準を抽出する。次いでAHPを利用した評価パラメータの重み付けと、電子教材と他の講義形態との得点間の有意差と、評価パラメータの重要度を定量的に測定し電子教材の改良点を示す重要な評価パラメータの選定を行う手法である。同時に、受講者をクラスター分析によりグルーピングしたときに、グループを決定付ける要素となった評価基準についても改良の余地がある箇所として考慮する。現在、教材評価は質問紙調査によるものが主流であるが、電子教材の評価は評価者においても未知数であり、何を基準に評価を行えば良いのか基本的な要素さえも不明確なまま評価を進めていくことにもなりかねない。本評価法を利用することにより、複数の受講者の意見を統合し、統計的分析により抽出された評価基準により評価をすすめていくことが可能となる。受講者が持つ判断基準を把握し、他の講義形態の良い点を電子教材に取り入れることが電子教材の開発には必要であると考えられる。

本研究では、AHP階層図の設計を「学習者」の目的に対するウエイトを求める第一段階と、第一段階のウエイトに合わせて教材の改良箇所を選定する第二段階の二段階に分けたことにより、学習者の主観的判断により忠実な評価が可能となった。講義は「学習者」と「教授者」、「学習環境」から作り出されるが、講義の良し悪しを評価するのは学習者自身であり、Kirkpatrickによって提唱された教育評価法における4つのレベル「1. 反応」「2. 学習」「3. 行動」「4. 業績」のうち、第1レベルの「反応」による評価とは、学習者の受講後の感想による評価をいい、感想に影響を与える要因は含んでいない。主因子分析法により抽出した因子に学習者の主観的判断もとに命名することで、第1レベル「反応」に忠実に基づいた評価が可能となった。また教材の改良を目的として、質問項目を「教授者」と「学習環境」とに分けて主因子分析を実施し第二段階の評価基準を抽出する。範囲を限定して主因子分析を行うことで、各評価基準は独立となりAHPの重み付けを正確に実施することが可能になった。

提案するAHP階層図設計法の有効性を確認することを目的として、今回の実証実験結果を猪俣による実験結果と比較した。評価基準の抽出、改良すべき箇所の選定において、この結果から提案した階層図設計法の有効性を確認した。

本論文をまとめるにあたり、最後に残された課題と今後の展望を示す。

提案した評価法では

- 教材を「教授者」や「システム」等の構成要素に分けること
- 評価項目をレベルに分割して評価を行うこと

を基本要素としている。本研究において「レベル1：反応」をもとにした評価は可能となったので、今後「レベル2：学習」に対する評価を検討し、最終的に「レベル3：行動」、「レベル4：業績」に基づいた評価へとつなげていく必要がある。

レベル2、3に対する評価については具体的に以下のような見通しを立てた。

- 「レベル2：学習」
実際に行われた学習の量、つまり教育カリキュラムの学習目標に対する理解度を測定する。このステップには受講者がどのような原則、事実及び技法を理解し吸収したかを測定して認識することが含まれる。教えている知識や技術、技能の試験を行うことを指すのではなく、重要なことは学習目標で掲げたことが出来るか否かを確認することである。
現在、本研究室において研究・開発が進められている学習者の受講時間などをもとにしたナビゲーションシステムを利用し、学習時間による評価を取り入れる。また事前テスト・事後テストを実施することで多角的に電子教材の評価を行うことで測定できると考えられる。
- 「レベル：3行動」
学習の結果として得られる受講者の行動変容を測定する。まず受講者自身の学習による変容度とは何かを定義する必要がある。その上で、受講数ヶ月後に受講者にインタビューなどを実施し、変容度を測定する。

上記以外にも検討すべき項目を今後の課題として述べる。

- 評価基準の抽出において、質問紙調査の基本となる基本項目リストの設計法について検討していく必要がある。新たに作成される電子教材に関しても必要とされる評価基準を的確に抽出できるようにするため、基本項目を自動的に選択できる手法を検討する必要がある。
- ネットワークを利用して遠隔地において提供される学習システムにおけるインタラクティブネスの評価に対して、本評価法が適用できるか否かの検証を行う。インタラクティブネス固有の質問項目の設計が重要となる。また重要な評価パラメータの抽出までがシステムティックに自動的に行える電子教材評価システムを開発する。
- 質問紙調査で得た自由記述文を、AHP評価法により抽出された改良すべき評価基準に基づいてスクリーニングすることにより、改良箇所を詳細化し改良の方針を得ることに利用する。

新しい情報技術の開発とともに進化する電子教材に対しても、系統的な手順により評価を行い改良すべき箇所を抽出できる本評価法は、今後より良い電子教材を作り出すために有用であると思われる。

謝辞

本研究を行うにあたり、終始変わらぬご指導賜りました落水 浩一郎 教授に心から深く感謝申し上げます。

日頃から研究に協力していただき、大変有益な御助言、御指導いただきました落水研究室 助手 村越 広享博士に深く感謝申し上げます。村越 広享 博士は、本研究のご指導をしていただく中、平成14年5月22日に急病のためご逝去されました。亡くなられました故人のご冥福を心からお祈りいたします。

本研究の審査委員として、数々の御助言賜りました片山 卓也 教授、篠田 陽一 教授に深く感謝申し上げます。

本研究を進めるにあたり、御助言、御協力いただきました猪俣 敦夫 博士に深く感謝申し上げます。被験者として実験に御協力いただいた本大学院の学生諸君に感謝申し上げます。

最後に、学業において支えてくださいました落水研究室の諸兄、また全てにおいて常に暖かく応援し支えてくださいました友人、両親に厚く御礼申し上げます。

皆様、本当にありがとうございました。

参考文献

- [1] 先進学習基盤協議会 (ALIC) 編著, eラーニング白書 2001/2002年版, オーム社
- [2] 猪俣 敦夫, 遠隔学習における電子教材のAHPを利用した評価法に関する研究, 北陸先端科学技術大学院大学 情報科学研究科 博士論文,2002.
- [3] T.L.Satty, Analytic Hierarchy Process, McGrawHill, 1980.
- [4] 刀根 薫, ゲーム感覚意思決定法, 日科技連,1986.
- [5] 木下 栄蔵, 入門 AHP 決断と合意形成のテクニック, 啓学出版,1992.
- [6] 刀根 薫, 眞鍋 龍太郎 編, AHP 事例集, 日科技連,1990.
- [7] 小松 秀国 他, 教育工学事典,pp100-103, 日本教育工学会編, 共立出版,2000.
- [8] 野嶋 栄一郎 他, 教育工学事典,pp87-90, 日本教育工学会編, 共立出版,2000.
- [9] 木下 栄蔵, わかりやすい意思決定論入門, 日科技連,2000.
- [10] 辻 新六, 有馬 昌宏, アンケート調査の方法: 実践ノウハウとパソコン支援, 朝倉書店,1987.
- [11] G.Booch,J.Rumbaugh,I.Jacobson, Unified Modeling Language Semantics and Notation Guide 1.0, Rational Software Corporation,San Jose CA,1997.
- [12] 統計科学のための電子図書システムのWebページ, <http://www.sci.kagoshima-u.ac.jp/ebse>.
- [13] 浦上 昌則, 論文を読むために必要な統計知識, <http://www.nannzann-u.ac.jp/urakami/pdf/vol1.pdf>.
- [14] 小杉 考司, 因子分析法について, <http://www002.upp.so-net.ne.jp/kossun/>.
- [15] 大村 平, 評価と数量化のはなし, 日科技連,1995.
- [16] 大村 平, 統計解析のはなし, 日科技連,2000.
- [17] Real Video Inc, <http://real.com>.

- [18] Macromedia Flash, <http://macromedia.com/>.
- [19] National Institute of Multimedia Education, 放送利用の大学公開アンケート調査シート, NIME SCS,1996.
- [20] SPSS Japan, <http://www.spss.co.jp/>.
- [21] 宮脇 典彦, 和田 悟, 阪井 和男, SPSS によるデータ解析の基礎, 培風館,2001.
- [22] 石村 貞夫, SPSS による分散分析と多重比較の手順, 東京図書,2000.
- [23] 中山 功, クラスタ分析とは, <http://opinion.nucba.ac.jp/nakayama/Cluster/semi2.html>.

本研究に関する発表論文

- [1] 藤林 由紀, 落水 浩一郎, 電子教材のAHPを利用した評価法, 電子情報通信学会 教育工学研究会, 信学技法, ET2002-32-43, pp29-34, Sep.2002.

付録

一対比較法

一対比較法は嗜好型官能評価法の一つである。自分の提示した尺度（基準）に従って、対象物を比較することができ、これからどちらがどれだけよりその基準に合っているかを分析することが可能である。対象を2つずつ比較し、得られた数値をもとに分析を行なう。要素*i*は要素*j*と比較して、「どのくらい好ましいですか」や「どのくらい重要ですか」などの問いかけに答える形式で行う。最終的には対象を一本の数直線状に配置させ、対象物間の関係性と順位を見ることができる。

上述の重要度の算出法として、*n*個の要素のうち任意の要素*i*と要素*j*の組み合わせに対して、意思決定者に「要素*i*は要素*j*に比べてどのくらい重要ですか」と問い、その答えに応じて一対比較値 a_{ij} を与えて $n \times n$ 行列 $A = [a_{ij}]$ を作る。ここで $a_{ij} = 1, a_{ji} = 1/a_{ij} (i \neq j)$ と仮定する。*n* 個の要素があると $n(n-1)/2$ 回の一対比較行列 A ができる。この一対比較行列から各要素の重要度を求める。一対比較において、2つの要素を比較する際には、表 A.1 の尺度を用いる。一対比較行列から重要度を算出する方法を以下に述べる。*n* 個の評

表 1: 重要性の尺度とその定義

重要性の尺度	定義
1	同じくらい重要
3	少し重要
5	かなり重要
7	非常に重要
9	極めて重要
2,4,6,8	それぞれの中間の時に補間的に用いる。
重要でない時は逆数を用いる	

価項目 I_1, I_2, \dots, I_n があり、その本来のウエイトが w_1, w_2, \dots, w_n であるとする。そのとき、項目 I_i と I_j の重要度の一対比較値 a_{ij} は $a_{ij} = \frac{w_i}{w_j}$ という関係を満たすはずである。したがって一対比較行列 $A = [a_{ij}]$ は次のような形となる。

$$\begin{pmatrix} w_1/w_1 & w_1/w_2 & \cdots & w_1/w_n \\ w_2/w_1 & w_2/w_2 & \cdots & w_2/w_n \\ \vdots & \vdots & \ddots & \vdots \\ w_n/w_1 & w_n/w_2 & \cdots & w_n/w_n \end{pmatrix}$$

そして、A の右側からウエイトベクトルを乗じると、以下の形になる。

$$A_w = \begin{pmatrix} w_1/w_1 & w_1/w_2 & \cdots & w_1/w_n \\ w_2/w_1 & w_2/w_2 & \cdots & w_2/w_n \\ \vdots & \vdots & \ddots & \vdots \\ w_n/w_1 & w_n/w_2 & \cdots & w_n/w_n \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = n \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}$$

この行列式から、ウエイトベクトルは A の固有ベクトルであり、n は固有値であることがわかる。行列の 2 行目以下は第 1 行列の定数倍であるから、行列 A の階数は 1 となる。つまり行列 A の n 個の固有値のうち、1 つだけが非零で他はすべて零である。一般に固有値の和 = 行列の体角要素の和 = n となるので、非零の固有値は n のみであり、理想的な行列 A の最大固有値となる。現実の一対比較行列 A が上式の形をしていることは期待できないが、ほぼ近い形をしているとみれば A の最大固有値と固有ベクトルを求めればその固有ベクトルが各評価項目のウエイトとして採用できるわけである。つまり生成した一対比較行列が、理想的な一対比較行列とどの程度異なるかを知ることが必要となり、整合性の指標を用いる。

一対比較の判断の首尾一貫性の尺度として、AHP では整合度指数 (コンシステンシー指数: Consistency Index: 以下 C.I とする) を用いる。行列 A には n 個の固有値があり、その和が n となることがわかっている。したがって行列 A の最大固有値を λ_{max} とすると、 $\lambda_{max} - n$ は、 λ_{max} 以外の固有値の大きさを示す指標とみることができる。そして (n-1) 個の固有値でこの指標を有するので、固有値 1 個あたりのこの指標の平均値は

$$C.I = \frac{\lambda_{max} - n}{n - 1}$$

となる。行列 A が完全に整合性がある場合この値は 0 であり、この値が大きくなればなるほど不整合性は高いと考えられる。ただし T.L.Satty 教授は経験から C.I 値が 0.1 以下 (場合によっては 0.15) であると整合性ありとすることを提案している。

内部従属法

AHPでは本来、次のような独立条件を仮定している。

- 同一レベルにある評価基準間は互いに独立である
- 各レベルは独立である

これらの条件を満たさない場合、特に同一レベルの評価基準間において従属性がある場合には、内部従属法という手法を用いて対応する必要がある。内部従属法とは、同一レベルの評価基準間に従属性があるとき、その相互関係の数行列を用いて分析する手法である。まず評価基準間の影響力の強さを一対比較し、この一対比較行列の固有ベクトル（重み）を求め、評価基準間の従属関係を整理する。この従属数行列と、各評価基準が独立であると仮定したときの重みを掛け、真の評価基準の重み（従属性を考慮した重み） w_c は求められる。しかし内部従属法には、評価基準間の一対比較を行うことで意思決定者の負担が増え、評価が過度に複雑になるという欠点もある。