

K_012

電子教材のキーワード関係の可視化による評価手法に関する研究

Research for Assessment Method of Digital Learning Materials Using Visualization of Keyword Relationship

池辺 正典†
Masanori Ikebe田中 成典‡
Shigenori Tanaka古田 均‡
Hitoshi Furuta中村 健二†
Kenji Nakamura吉村 智史†
Satoshi Yoshimura

1. はじめに

近年、コンピュータやネットワークなどの情報技術を活用し、新しい学びの形態を実現する eラーニング[1]が注目されている。eラーニングの代表的なものには、WBT (Web Based Training) がある。WBT とは、学習者がインターネットを利用し、サーバに蓄積されている多様な電子教材の学習を自分のペースで進めていくものである。WBT の利点は、企業や学校などの多数の学習者に対し、低コストで一貫した教育を実施することができる点である。そのため、高等教育[2]や企業内教育[3]を中心に、WBT の導入が急速に広がっている。また、WBT の普及に伴い、多種多様な電子教材が作成されている。

WBT の学習において、最も重要なものは電子教材の品質である。2006年3月現在、eラーニングの普及促進を目的とした特定非営利活動法人日本 eラーニングコンソシアム[4]には、電子教材が 1622 件登録されているが、この中から品質の高い電子教材を見つけることは困難である。

既研究として、AHP (Analytic Hierarchy Process) を用いることにより、WBT 教材、VOD (Video on Demand) 教材と教室講義の比較を行い、各々のパラメータを評価する研究[5]や、学習者と指導者の作成するコンセプトマップの比較を行い、電子教材の評価手法を提案する研究[6]がある。しかし、既研究[5][6]においては、手作業であるため時間がかかってしまい、全ての電子教材を評価することが困難である。また、AHP やコンセプトマップは、人間の主観的な価値判断を取り入れているため、客観的に電子教材を評価することができない。そこで、本研究では、人間の主観的判断に依存しない電子教材の品質を評価する手法を提案する。

2. 研究の概要

本研究では、電子教材の重要なキーワードの関連を可視化したものを評価することにより、品質の高い電子教材を抽出することを目的とする。本システムは、図1に示すように、電子教材を入力とし、電子教材解析機能とグラフ出力機能から構成されている。電子教材解析機能では、電子教材を評価するためのキーワードを抽出するために、電子教材の解析を行う。グラフ出力機能では、キーワードの関連を可視化するために、電子教材解析機能によって抽出したキーワードの関連付けを行う。

本論文の構成は、まず、3章と4章において、本システムの各機能について順に解説する。次に、5章で本提案手法の精度を検証するために、実証実験を行う。最後に、6章で得られた成果と今後の課題についての考察を述べる。

3. 電子教材解析機能

電子教材解析機能では、入力した電子教材のキーワード解析を行う。本機能は、電子教材分割、電子教材内キーワード抽出と目次内キーワード抽出の3つの処理から構成されている。

3.1 電子教材分割

電子教材分割では、目次構成により該当する章、節や項の単位で、電子教材を複数のテキストファイルに分割する。ここで、電子教材の単一のテキストファイルを l_0 とし、分割された電子教材の集合を $LO=(l_{01}, l_{02}, \dots, l_{0m})$ と定義する。また、電子教材分割では、キーワード抽出において、キーワードを適切に分割して抽出するために、電子教材内の記号を半角から全角へ変換するなどの前処理を行う。

3.2 電子教材内キーワード抽出

電子教材内キーワード抽出では、電子教材に含まれているキーワードの重要度を評価する。まず、分割後の電子教材 l_{0m} に対して形態素解析を行う。形態素解析とは、文を構成する最小の意味単位である形態素に分解することである。形態素解析の結果から名詞または未知語をキーワードとして抽出する。単語の抽出時に、名詞または未知語が連続している場合は、複合名詞としてキーワードを抽出する。ここでは、形態素解析器である茶釜[7]を用いて、電子教材の形態素解析を行う。茶釜では、入力文の単語列に対して、品詞の出現確率の対数をリスクと定義し、実現可能な単語列、品詞列のうちリスクの和が最小となる組み合わせを出力結果とする。

次に、抽出したキーワードに対して、接頭語や接尾語の除去などの前処理を行うことで、キーワードとして不適切なものを排除する。

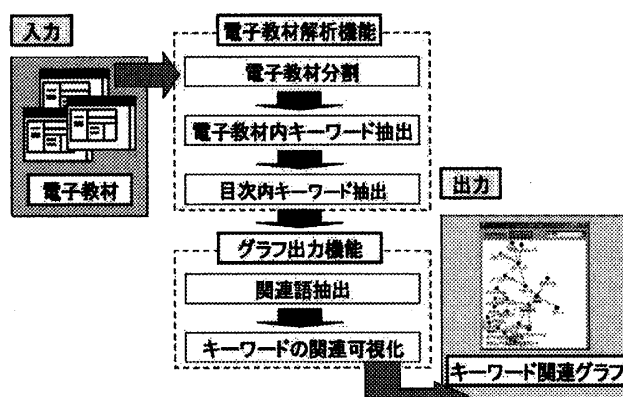


図1 本システムの概要

† 関西大学大学院 総合情報学研究科

‡ 関西大学 総合情報学部

最後に、電子教材全体の各キーワードに重み付けを行い、重要なキーワードのみを抽出するために、TF/IDF (Term Frequency Inverse Document Frequency) 法を適用する。電子教材の集合における各キーワードの重要度を算出し、ある閾値以上のキーワードのみを抽出する。次項より、本研究で用いた重み付け手法である TF/IDF を説明する。

3.2.1 TF

$tf(t_i)$ は、式(1)で定義されるように、任意の電子教材 lo_x に含まれているキーワードの集合 $T=(t_1, t_2, \dots, t_n)$ の中に、キーワード t_i がどれくらい出現するかを基準として重みの付与を行う。任意の電子教材 lo_x の中に出現する t_i の総数を $S(t_i, lo_x)$ と定義する。しかし、近年は長文を含む電子教材が増加しており、長文が含まれている電子教材は、キーワードの重みが不正に大きくなる可能性がある。そこで、 $S(t_i, lo_x)$ を電子教材中の全キーワードの出現総数で割った値を採用することで、キーワード数に依存しない重みを付与することができる。

$$tf(t_i) = \frac{S(t_i, lo_x)}{\sum_{j=1}^n S(t_j, lo_x)} \quad (1)$$

3.2.2 IDF

$idf(t_i)$ は、式(2)で定義されるように、キーワード t_i が任意の電子教材の集合 LO の中にどれくらい出現するかを基準として重みの付与を行う。 N は電子教材の総数を表し、 $df(t_i)$ は、電子教材の集合 LO の中で、キーワード t_i が出現する電子教材の数を表す。

$$idf(t_i) = \log \frac{N}{df(t_i)} \quad (2)$$

式(2)が示すように、 $idf(t_i)$ では、あらゆる電子教材に出現する一般的なキーワードの重みを小さくすることができるため、少数の電子教材のみに現れる重要なキーワードの重みを大きくすることができる。

3.2.3 TF/IDF

$tf(t_i)$ と $idf(t_i)$ の2つの尺度を組み合わせることで、電子教材に含まれているキーワード t_i の重み $w(t_i)$ を算出する。 $w(t_i)$ は、式(3)で定義されるように、 $tf(t_i)$ と $idf(t_i)$ の積で表される。

$$w(t_i) = tf(t_i) \times idf(t_i) \quad (3)$$

$w(t_i)$ の閾値は、電子教材における重みの最大となるキーワードの10%とする。閾値以下のキーワードは削除する。そして、 $w(t_i)$ が閾値以上のキーワード t_i を電子教材内キーワードの集合 $W=(t_1, t_2, \dots, t_p)$ と定義する。

3.3 目次内キーワード抽出

目次内キーワード抽出では、目次に含まれているキーワードを抽出する。電子教材の目次に対して、形態素解析を行う。分割された目次に含まれている章、節と項などの目次構成を表すキーワードを用いて、各々に対応するキーワードを抽出する。また、複数のキーワードがある場合は、キーワード抽出機能により算出した、キーワードの重みが大きいものを抽出する。抽出された目次内キーワードの集合を $C=(c_1, c_2, \dots, c_q)$ と定義する。

4. グラフ出力機能

グラフ出力機能では、電子教材解析機能で抽出したキーワードの関連性を評価し、グラフの可視化を行う。本機能は、関連語抽出とキーワードの関連可視化の2つの処理から構成されている。

4.1 関連語抽出

関連語抽出では、目次内キーワードと電子教材内キーワードとの関連性の評価を行う。キーワード間の関連性評価では、検索エンジン Google による検索結果件数を用いる。目次内キーワードの集合 C と教材内キーワードの集合 W の関連性の評価は、AND 検索結果件数を OR 検索結果件数で割った値を関連度とする。

2つのキーワードが共に出現するウェブページの検索結果件数が増加することに伴い、関連度が大きくなる。ここで、関連度がある閾値以上のものを電子教材の関連用語の集合を $K=(k_1, k_2, \dots, k_r)$ と定義する。関連語抽出における閾値は、重みの最大となる関連度の5%とする。閾値以下のキーワードは削除する。

4.2 キーワードの関連可視化

キーワードの関連可視化では、キーワードの階層構造の可視化を行う。目次情報における章、節、項と教材内に含まれているキーワードの順に、キーワードを結びつけ、階層構造を可視化する。本システムにより可視化した階層構造をキーワード関連グラフとする。キーワード関連グラフの出力結果を図2に示す。

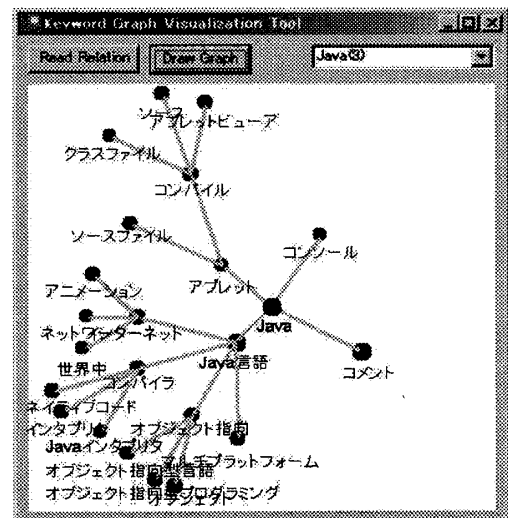


図2 キーワード関連グラフの出力結果

本システムでは、図2に示すように、章キーワードである Java を中心に配置し、節キーワード、項キーワードと電子教材に含まれているキーワードの順に、キーワードを結びつけている。本システムの出力結果には、3階層のキーワードと関連用語22個が結びついている。ここで、キーワード間の関連の強いものは近くに配置されるため、Java と Java 言語は関連が強いことから近くに、Java とコメントは関連が弱いことから遠くに出力されている。また、重要度の大きいキーワードは大きく表示されるため、Java や Java 言語などは重要であることから大きく、アプ

レットやコンソールなどあまり重要ではないことから小さく出力されている。

5. 評価実験

5.1 キーワード間の関係情報の品質評価

キーワード間の関係情報の品質評価では、本システムから出力された電子教材評価が適切であるかを確認するために、電子教材内の関連のあるキーワード同士を結びつけることができたかを検証する。

評価実験では、Javaを題材とする電子教材2点とVisual Basic, EclipseとDelphiを題材とする電子教材3点の合計5点の電子教材を本システムにより評価する。本実験に採用した電子教材の詳細を表1に示す。電子教材の詳細は、1章当りに25ページ含まれていることがわかる。

表1 電子教材の詳細

	章の数	ページ数	1章当りのページ数
Java教材A	17	170	10.0
Java教材B	12	333	27.8
Visual Basic教材	15	383	25.5
Eclipse教材	12	449	37.4
Delphi教材	10	315	31.5
平均	13.2	330.0	25.0

品質評価実験では、キーワード抽出の完全性と正確性を評価するために、 F 値によって評価を行う。本実験では、 F 値を算出し、電子教材のキーワードの抽出精度を検証する。 F 値とは、適合率 P と再現率 R から式(4)により算出する。

$$F = \frac{(\beta^2 + 1) \times P \times R}{\beta^2 \times P + R} \quad (4)$$

ただし、 β は、適合率と再現率の相対パラメータである。ここで、再現率とは、情報抽出の完全性を評価するための尺度であり、電子教材内におけるキーワードの抽出漏れの少なさを示す。再現率 R は、キーワード関連グラフにおいて、関連があると目視により評価したキーワードの総数を電子教材内の重要なキーワードの集合 W において、キーワードであると目視により評価したキーワードの総数から割ることで求めることができる。適合率 P とは、情報抽出の正確性を示す尺度であり、電子教材からどの程度のキーワードが抽出されたかを示す。適合率 P は、キーワード関連グラフにおいて、関連があると目視により評価したキーワードの総数を電子教材の関連用語の集合 K の総数から割ることで求めることができる。本実験においては、適合率と再現率の重みを対等とするために、 $\beta=1$ として精度検証を行う。

F 値とは、式(4)が示すように、再現率 R と適合率 P の調和平均である。 F 値は、再現率 R と適合率 P の両方の値が大きい場合に値が大きくなる。再現率 R 、適合率 P と F 値は0から1の範囲にあり、値が大きいほど評価が高い。電子教材の再現率 R 、適合率 P と F 値を表2に示す。

表2 電子教材の再現率 R 、適合率 P と F 値

	再現率 R	適合率 P	F 値
Java教材A	0.79	0.85	0.82
Java教材B	0.61	0.84	0.68
Visual Basic教材	0.79	0.78	0.81
Eclipse教材	0.68	0.83	0.75
Delphi教材	0.67	0.81	0.73
平均	0.71	0.82	0.76

表2に示すように、電子教材における再現率 R は0.71となり、比較的に高い精度を得ることができた。各電子教材の再現率に注目すると、値にばらつきがあった。この原因として、目視により重要であると判断したキーワードにおける本システムでのキーワード抽出精度が、電子教材によって異なったためである。本システムにおけるキーワード抽出の精査は、閾値のみを用いているため、電子教材によって、閾値以下のキーワードを多数削除してしまったと考えられる。また、電子教材における適合率 P は0.82を超え、高い精度を得ることができた。これは関連語抽出において、一般的なキーワードを排除し、関連語がうまく抽出できたためである。さらに、電子教材の F 値は0.76となり、良い結果を得ることができた。これは、再現率 R と適合率 P において、比較的に高い値を示すことができたためであると考えられる。このことから、本提案手法では、電子教材の中から関連用語を非常に高い精度で抽出できていることがわかる。最後に、本実験に採用した電子教材がほぼ同様の結果を示していることから、電子教材の詳細と品質評価の実験結果に関連がないことがわかった。このことから、電子教材の詳細によって、実験結果が左右されることがないといえる。

5.2 電子教材の比較実験

電子教材の比較実験では、複数の電子教材から品質の高い電子教材の判別を行う。キーワードが幅広くかつ詳細に配置されている電子教材は、品質の高い電子教材であるといえる。そのため、本実験では、キーワードの幅広さを評価するために、2つの電子教材の目次内キーワードの集合 C と結びつく関連用語の集合 K の数を評価する。また、キーワードの詳細さを評価するために、目次内キーワードの階層の深さを評価する。本実験の対象として、Javaを題材とした電子教材2点による比較実験を行う。

比較実験では、まず、2つの電子教材に共通の目次内キーワードの集合 C を抽出する。次に、IT専門の辞書[8]を用いて、2つの電子教材に共通して含まれているキーワードの抽出を行う。そして、辞書にあるキーワードと辞書にないキーワードの関連用語数と階層の深さを抽出する。重みが大きい辞書掲載キーワード上位5件における関連用語数と階層の深さと目次内キーワードの関連用語数と階層の深さの平均を表3と表4に示す。

表3 辞書掲載キーワードの関連用語数と階層の深さの上位5件

辞書掲載キーワード	Java教材A		Java教材B	
	関連用語数	階層数	関連用語数	階層数
コンパイル	5	1	1	1
GUI	45	3	56	3
キーボード	5	2	4	2
while文	2	2	0	1
演算子	9	3	5	2

表3に示すように、Java教材AとJava教材Bでは、関連用語数と階層数にばらつきがあることを確認した。これは、2つの電子教材において、重点をおいて説明しているキーワードの違いを表している。また、2つの電子教材の各辞書掲載キーワードにおいて、関連用語数に同様の傾向がみられた。これは、Javaを題材とした2つの電子教材において、説明する必要のあるキーワードが共通していることを表している。

表4 目次内キーワードの
関連用語数と階層の深さの平均

	Java教材A		Java教材B	
	関連用語数	階層数	関連用語数	階層数
辞書に掲載されている キーワードの平均	9.80	2.15	10.65	1.95
辞書に掲載されていない キーワードの平均	4.85	1.70	6.90	1.50
平均	7.33	1.93	8.78	1.73

表4に示すように、Java教材AとJava教材Bでは、関連用語数の平均に差が確認できた。Java教材Bは、Java教材Aよりも多くのキーワードが結びついていることがわかる。しかし、2つの電子教材の階層の深さに大きな差は見られなかった。これは、2つの電子教材のキーワードの階層の深さが類似していることを表している。

表3と表4から、キーワードの辞書の有無に注目すると、辞書掲載されているキーワードは辞書に掲載されていないキーワードよりも、多くのキーワードが結びつき、階層の深さも大きくなるのがわかる。これは、辞書にあるキーワードが一般的に重要であることを示している。以上の結果から、Java教材Bは、Java教材Aよりも、関連用語が多く結びついている点で、品質が高いといえる。

5.3 学習結果による電子教材評価

学習結果による電子教材評価では、電子教材の比較実験の結果が正しいかを検証するために、実証実験を行う。実証実験では、ランダムに選出したJava未習得の大学生10名を対象とし、2つのグループに分け、それぞれ異なる電子教材の学習を行う。一方のグループは電子教材A、他方のグループは電子教材Bの学習を行う。学習が終わった後に、2つのグループにJava認定資格の問題集[9]のテストを50問実施した。2つのグループのテストの正解率を表5に示す。

表5 テストの正解率

	Java教材A	Java教材B
正解率	45%	56%

表5に示すように、Java教材Bを学習したグループは、Java教材Aを学習したグループの正解率を上回った。この結果から、電子教材の比較実験結果であるJava教材Bは、Java教材Aよりも品質が高いことが実証された。

以上より、本システムでは、まず、電子教材の詳細に依存することなく、電子教材から関連用語を高い精度で抽出できることを実証した。次に、目次内のキーワードの関連用語数と関連の深さを評価することで、複数の電子教材から品質の高い電子教材を判別できることを定量的に表現した。最後に、学習結果による電子教材評価において、2つのグループによる電子教材学習後のテスト結

果において、電子教材の比較実験の結果が正しいことを実証した。このことから、本提案手法の有効性を実証することができた。

6. おわりに

本研究では、本システムによる複数の電子教材の評価手法の提案を行った。本システムを用いることにより、人間の主観的な判断に依存しない教材の評価を実現した。

現状の問題点として、本システムのキーワード抽出手法では、閾値以下となるキーワードが抽出できないことやキーワードが適切に分割されないことがあった。そのため、本システムの各処理の精度を向上させる工夫が必要である。また、本研究では、電子教材のキーワードのみの評価を行ったため、電子教材に含まれている文章の構造、図表やレイアウトに関する評価や意味的な解析は行われていない。そのため、キーワード以外にも、電子教材の評価手法を考案する必要がある。さらに、本研究で使用した電子教材は、HTML形式とWORD形式であり、VOD教材やFLASH教材などの評価は行われていない。そのため、今後はインタラクティブな電子教材の評価手法を考案する必要がある。

今後の発展としては、本システムの精度の向上と電子教材の再利用と再編集を行う。具体的には、分割した電子教材に対して、メタデータの付与を行う。また、ウェブから電子教材と関係の深い学習資源の獲得を行う。そして、既存の電子教材とウェブから取得した学習資源を用いた電子教材の再利用と再編集を行う予定である。さらに、電子教材の各キーワードからSemantic Web技術で使用されているメタデータの集まりであるオントロジーの構築を目指す。各キーワードに関係情報を持たせることで、電子教材の学習者に提示する情報量を増減させることや複数のキーワードがどのような関係になっているかという情報を提供することが可能となる。

参考文献

- [1] Randy Garrison, Terry Anderson : E-Learning in the 21st Century: A Framework for Research and Practice, Routledge Falmer (2003).
- [2] Roger Schank : Designing World-Class E-Learning: How IBM, GE, Harvard Business School and Columbia University Are Succeeding at E-Learning, McGraw-Hill (2001).
- [3] 菅原良 : eラーニングの発展と企業内教育, 大学教育出版 (2005).
- [4] 日本eラーニングコンソシアム. <http://www.elc.or.jp/>
- [5] 猪俣敦夫, 村越広享, 落水浩一郎 : 電子教材の評価パラメータ抽出法, 教育工学研究会技術研究報告, 電子情報通信学会, Vol.101, No.433, pp.5-12 (2001).
- [6] 鈴木満, 猪俣敦夫, 村越広享, 東条敏, 落水浩一郎 : コンセプトマップを用いた電子教材評価手段の提案, 教育工学研究会技術研究報告, 電子情報通信学会, Vol.101, No.41, pp.41-48 (2001).
- [7] 茶釜. <http://chasen.naist.jp/hiki/ChaSen/>
- [8] IT用語辞典 e-Words. <http://e-words.jp/>
- [9] 八木裕乃, 明壁敦子, 須澤秀人 : 徹底攻略 Java アソシエイツ問題集[310-019]対応, インプレス (2006).