

科学研究費助成事業 研究成果報告書

平成 29 年 6 月 12 日現在

機関番号：32601

研究種目：若手研究(B)

研究期間：2015～2016

課題番号：15K16269

研究課題名(和文) 学習に関連したWebコンテンツの特定と学習履歴情報に基づく学習者特性の解明

研究課題名(英文) Identification of learning related contents and elucidation of learner's characteristics based on learning-log

研究代表者

豊田 哲也 (Toyota, Tetsuya)

青山学院大学・理工学部・助教

研究者番号：30650618

交付決定額(研究期間全体)：(直接経費) 2,200,000円

研究成果の概要(和文)：eラーニングシステム等で得られる学習関連のログデータを基に、学習者個人の学習特性を特定した学習プロフィールを開発する仕組みについて研究を行った。本研究で取り組んだ主な課題は、学習レベルに応じた学習に関連するキーワードの特定と学習コンテンツの特定、学習プロフィールの構築の3つである。日本語学習環境において実験を実施し、学習プロフィール構築に向けたアプローチを確立することが可能となった。

研究成果の概要(英文)：We studied the mechanism to develop a learning profile specifying the learning specifying the learning characteristics of individual learner's based on the learning-log data obtained by e-learning system. The main tasks of this research are identifying keywords related to learning according to the learning level, identifying learning contents, and constructing learning profiles. We have been able to establish an approach toward building a learning profile in the evaluation experiments for a learning Japanese environment.

研究分野：教育工学

キーワード：教育工学 Webの利用 教育・学習支援システム 日本語教育

1. 研究開始当初の背景

近年、大学等において e ラーニングシステムの利用が進み、その利用履歴等から個人の学習特性を明らかにする研究が進んでいる。しかし、Web 上の学習サービス・コンテンツの増加によって学習環境が多様化しており、e ラーニングの利用範囲内で個人の学習履歴や学習特性を特定することは適切ではない。そのため、e ラーニングの学習コンテンツだけでなく、その他の Web 上の学習コンテンツの利用状況を把握する必要がある。

今後も学習者にとって有益な Web 上の学習コンテンツが増加していく一方で、e ラーニングにおいても同等のコンテンツを用意することは困難であり、Web 上の学習コンテンツを利用する比重が大きくなるのが容易に想像できる。そのため、学習者の学習プロフィールの構築には、Web 上の学習コンテンツがどの程度利用されているのかを、e ラーニングの利用状況と併せて総合的に把握することが必要不可欠である。申請者は、現在の学習プロフィール構築に関する研究において、科目・単元等の理解に必要な認知的技能を測定する認知診断テストを e ラーニングシステムに実装しており、Web 上の大規模データに対するテキストマイニングやクラスタリング手法の研究に従事していた経験を含めて、有益な研究成果が得られると考え、本研究の着想に至った。

2. 研究の目的

本研究では、学習者が利用した Web 上にある様々なコンテンツの中から学習に関係するコンテンツを抽出することで、e ラーニングシステムの利用履歴と Web 上の学習コンテンツの利用状況を基に個人の学習プロフィールを構築することを目的としている。そのため解決すべき課題を以下の3つに集約し、それぞれの課題解決によって研究目的を達成可能であると考えた。以下に研究課題を列挙する。

(1) <学習関連キーワードの抽出>

これまでに Wikipedia を情報源とする学習関連キーワードの抽出手法を確立しているが、学習レベルに応じた学習関連キーワードの抽出が困難であった。これまでの方法を拡張することでこれらの解決を図る。学習者の能力やレベルに応じて学習関連キーワードに対して重み付けを行うために認知診断テストを利用し、学習関連キーワードの重要度設定を行う。

(2) <学習コンテンツの特定>

Web コンテンツの利用履歴と前述した学習関連キーワードを照合することで、学習に関連するコンテンツを特定する。この際に、Web コンテンツの種類や構造的特徴を元に機械学習によって学習コンテンツを特定する。

(3) <学習プロフィールの構築>

e ラーニングの利用履歴と特定した Web 上の学習コンテンツを紐付けし、学習者個人の特徴を明らかにする。学習者の学習状況や利用特性などを考慮した上でこれらの学習プロフィールを構築する。

これらの研究課題の解決によって、e ラーニングシステム以外での学習者の学習ログを収集することが可能であると考えられる。

3. 研究の方法

本研究では、研究目的に列挙した3つの研究課題を順に解決する。まず、学習に関連するキーワード群を学習者の能力に応じて重要度を設定する。次に Web の閲覧履歴をブラウザから取得し、Web コンテンツの構造情報とテキスト情報を元にコンテンツの種類に応じたタグ付けを行い、Web コンテンツが学習に関連するコンテンツであるかどうかを機械学習によって同定する。これらの情報を元に、e ラーニングの利用履歴と Web 上で利用した学習コンテンツの利用情報を総合的に分析することで、個人の学習方法や学習に対する積極性などの学習プロフィールを構築する。本研究は日本語学習環境でのそれぞれの具体的な研究方法は以下のとおりである。

(1) 閲覧した Web ページが学習コンテンツであるか否かを判断するために、判断基準となるデータを e ラーニングシステムおよびこれに実装した認知診断テストを活用する。認知診断テストは学習者の認知的技能の習得確率を測定するテストである。本研究で対象とする日本語教育における語彙の認知診断テストはすでに構築済みであるが、それぞれの語彙項目の成績に応じて語彙の重要度を設定する。ここで、習得確率はあくまで認知的技能の習得確率であるので、語彙に対する理解度をそのまま設定するのは弊害がある。例えば漢字の意味理解に関する習得確率が最大であったとしても、個々の漢字の理解度は異なる。そこで、本研究では学習者の語彙に対する理解度を可読性スコアとして定義し、学習者が語彙に対してどの程度読むことができるかを理解度の指標とした。ここで得られた語彙とその可読性スコアを学習関連キーワードとする。

(2) 得られた学習関連キーワードが閲覧した Web コンテンツの中にどの程度含まれているかを元に学習コンテンツの同定を行う。Web コンテンツの特徴情報は、コンテンツ内のテキスト情報とコンテンツの構造情報である。テキストデータは(1)で得られた学習関連キーワードと照合することで、これらのキーワードが含まれているかどうかを判定するために用いる。この際、キーワードが含まれていればそれらのキーワードをタグ

としてコンテンツに付与する。これらのタグの数と種類に応じて、コンテンツが学習関連コンテンツであるかどうかを機械学習によって判断する。具体的には、クラスタリングによってコンテンツをいくつかのクラスに分割し、クラスタの重心付近のコンテンツを学習関連コンテンツと仮定する。この仮定の学習関連コンテンツのキーワードの可読性スコアを元に可読性を測定し、閾値以上のコンテンツを学習関連コンテンツとし、閾値以下のコンテンツは学習関連コンテンツではないものとする。この操作を、複数回クラスタ数を変化させて繰り返し、学習関連コンテンツを同定する。

同様にコンテンツの構造情報を抽出するとともにこれらをタグ付ける。構造情報のタグの種類は、HTML タグで構成し、コンテンツの種類を分類する目的で用いる。これは後述の学習プロフィール構築のために用いる。

(3) 学習プロフィールを構築するために、e ラーニング上での学習ログデータと学習関連コンテンツの閲覧履歴データを収集し、これらを総合的に分析することで学習者の学習特性を明らかにする。学習特性は、主に学習行動とその活動時間を基本とし、e ラーニングに関してはシステムとその活動内容の利用時間を、Web コンテンツの閲覧履歴については、テキスト量とコンテンツの構造情報を元に利用時間を推定して用いる。これらの情報から学習に対する積極性や継続性、学習方法の種類や活動時間の長さなどを学習者の学習特性とする。これらを個々の学習者ごとに分析し、学習プロフィールを構築する。学習に対する積極性と継続性は e ラーニングの利用履歴でも判断可能であるが、e ラーニングを利用していない期間にも学習を継続していた可能性はある。そこで、学習コンテンツの利用状況を踏まえた上で積極性と継続性を総合的に判断する。学習方法の種類は、e ラーニングの利用状況に加えて、同定した Web 上の学習コンテンツの種類を考慮する。「動画や画像などのコンテンツを利用する頻度が多いので、テキスト以外の情報を多く利用する学習を好んでいる」、「クイズ形式のコンテンツをテスト前期間中に多く利用している」といった具合に、e ラーニングと Web のそれぞれのコンテンツをどのように利用しているかを分析し、学習者がどのような学習傾向、学習特性を持っているのかを把握する。これらのデータは以後の学習者の分析・パターン分類に用いることができるよう、e ラーニング上に記録できるよう e ラーニングシステムの機能拡張を実施する。

4. 研究成果

本研究を進めるにあたって、当初予定していた計画の大きな変更点がいくつかあった。以下に成果を示すとともに、研究遂行上の変

更点について述べる。

(1) e ラーニングシステムや認知診断テストの情報のみでは学習者の語彙に関する理解度を測定するのは困難であったことから、学習者の語彙の可読性を記録・推定する仕組みを新たに構築する必要があった。そこで、Web 上のコンテンツに対して、学習者が語彙の可読性の是非を記録できる Web アドインを開発し、読むことのできる語彙、読むことの困難な語彙を記録し、これらのデータを e ラーニングシステムに組み込む仕組みを開発した。

(2) (1)に関連して、学習者の Web 閲覧履歴を収集することは、学習者のプライバシー保護の観点から困難であったことと、普段の Web の閲覧傾向として、学習のために Web を利用する時間が少ないこと、の 2 点が事前のアンケート調査で確認することができた。このことから、語彙の可読性を記録した Web コンテンツのみを収集し、これらのコンテンツの中で学習関連コンテンツの同定を行うこととした。

(3) データ収集前の事前調査では、学習者が主に利用するサイトを申告してもらったが、主に辞書サイトなどの言葉の意味を調べるために Web が利用されていることがわかり、Web アドインを改良して、Web ページ中に選択した箇所の言葉の意味を表示する機能を追加した。これにより、辞書サイトで調べる手間が省くことができ、普段の何気ない Web の閲覧を通して、簡単に意味を調べることが可能となった。また、語彙の可読性判断の機能に加えて、学習者自身がどんな言葉を調べたのかを Web アドインで記録させることを可能にしたことにより、効率的に情報収集することが可能となった。

(4) e ラーニングシステムと普段の Web 閲覧履歴を総合的に分析することで得られる学習者のプロフィール構築に関しては、前述の通り情報収集の方式等を変更したため、総合的な分析には至っていない。現時点で本研究の最終目標である上記目的を達成するために鋭意努力中である。一方で、e ラーニングシステムから得られる学習ログデータの分析を同時並行で進め、e ラーニング上での学習者の学習傾向や学習特性についてはいくつかの知見を得ることができた。1 つに、日本語学習者の学習時間帯やどのタイミングで学習を行うかについては、学習者の成績とはつながらず、学習者の母語の影響が大きいことが分かった。すなわち、日本語学習者の学習傾向は個人差よりも出身国や母語といった環境が大きく影響していることがわかった。また、実験期間後に実施したアンケートでは、e ラーニングによる学習に苦手意識を持つ学習者が一定数おり、それらの学習

者は紙媒体などでの学習を好む傾向があった。この傾向はeラーニングでの学習成績と相関関係はなかったものの、以後のeラーニング環境での授業支援の在り方について検討を残すこととなった。

今後の課題として、当初予定していた計画では、eラーニング上で稼働する認知診断テストとの連携により、学習者の能力に応じたWebコンテンツの特定を行う上で、認知診断テストの役割は非常に重要であったが、様々な語彙をカバーするためには大量の問題とそれに応じた認知診断テストを作る必要がある、専門家の負担が非常に大きくなることがわかった。認知診断テストを有効に応用させるためには、問題バンクの充実と自動的なテスト作成が急務であると考えられる。また、学習者からの情報収集では、想像以上に協力者を募ることができず、十分満足なデータ収集ができなかった。また、得られたデータに関しても、閲覧しているWebサイト中の学習に関連していると思われるものには偏りが多く、ほとんど同じようなページを閲覧していることがわかった。日本語学習の環境においては、学習者が選ぶ教材がWeb上の物だと限られている現実があるため、これらの状況打開のためにも、日本語学習に関するWeb教材の開発と普及が望ましいと考えられる。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計 5 件)

豊田哲也, 谷部弘子, 孫媛: eラーニングのログデータを用いた日本語学習者の学習行動予測, 東アジア日本語教育・日本文化研究学会誌, Vol. 20, No. 1, pp. 21-35 (2017).

豊田哲也, 島田めぐみ, 谷部弘子, 孫媛: 日本語学習eラーニングシステムにおける学習者の小テストの学習活動分析, 東アジア日本語教育・日本文化研究学会誌, Vol. 19, No. 1, pp. 23-38 (2016).

鈴木雅之, 豊田哲也, 山口一大, 孫媛: 認知診断モデルによる学習診断の有用性の検討 教研式標準学力検査 NRT「中学1年数学」への適用, 日本テスト学会誌, Vol. 11, No. 1, pp. 81-97 (2015).

Suzuki, M., Toyota, T., and Sun, Y.: How learners use feedback information: Effects of social

comparative information and achievement goals, Proceedings of the 37th Annual Conference of the Cognitive Science Society, pp. 2308-2313 (2015).

Shimada, M., Sun, Y., Toyota, T., and Yabe, H.: Learners' Awareness and Learning Effect of Blended Learning, Proceedings of the 6th International Conference on Computer Assisted Systems for Teaching & Learning Japanese, pp. 51-58 (2015).

[学会発表](計 5 件)

Shimada, M., Sun, Y., Yabe, H., and Toyota, T.: Development of A Japanese Grammar Cognitive Diagnostic Web Test, The 7th International Conference on Computer Assisted Systems for Teaching & Learning Japanese, Waseda University (2017/8/6). (採録決定済み)

豊田哲也, 島田めぐみ, 谷部弘子, 孫媛: 日本語語彙能力測定のためのWebアドインの開発, 日本教育工学会第32回全国大会, 大阪大学, 豊中市 (2016/9/18).

Toyota, T., Sun, Y., and Yabe, H.: Action Prediction of Japanese Language Learners Based on the e-Learning Log Data, 2016 the 20th International Conference on East Asian Research Society for Japanese Language Education and Japanese Culture, College of Hilo, Hawaii University, Hawaii, U.S.A. (2016/8/7).

豊田哲也, 島田めぐみ, 谷部弘子, 孫媛: 日本語学習に特化した学習コンテンツの特定, 日本教育国学会第31回全国大会, 電気通信大学, 府中市 (2015/9/22).

Toyota, T., Shimada, M., Yabe, H., and Sun, Y. : Analysis of Students' Performance on Online Quizzes for Developing Japanese Language e-Learning Materials, 2015 the 19th International Conference on East Asian Research Society for Japanese Language Education and Japanese Culture, Seinan Gakuin University, Fukuoka, Japan (2015/8/22).

6. 研究組織

(1) 研究代表者

豊田 哲也 (TOYOTA, Tetsuya)
青山学院大学・理工学部・助教
研究者番号：30650618

(2) 研究分担者

なし

(3) 連携研究者

なし

(4) 研究協力者

孫 媛 (SUN, Yuan)
国立情報学研究所・情報社会相関研究系・
准教授

島田 めぐみ (SHIMADA, Megumi)
日本大学・総合社会情報研究科・教授

谷部 弘子 (YABE, Hiroko)
東京学芸大学・留学生センター・教授