

e-learning における学習者の Low-Level Interaction 特徴に基づく 行き詰まりの推定システム

An Automatic Estimation System of Learners' Impasses
Based on Features in Low-Level Interactions in an E-learning Environment

堀口 祐樹*¹ 小島 一晃*² 松居 辰則*²
Yuki HORIGUCHI Kazuaki KOJIMA Tatsunori MATSUI

*¹早稲田大学 大学院人間科学研究科
Graduate School of Human Sciences, WASEDA Univ.

*²早稲田大学 人間科学学術院
Faculty of Human Sciences, WASEDA Univ.

Mentoring is one of the effective tutoring tactics for learning support. However, mentoring requires much cost of a human tutor. Therefore, we propose an Intelligent Mentoring System (IMS). Toward the implementation of IMS, the purpose of the current study is to develop an automatic estimation model for real-time detection of impasses of learners, a mental state occurring in learning activity. One of the important features of the proposed model is to estimate the impasses through analysis of Low-Level Interaction (LLI) resources, which are fine grain temporal data of interactions with learners. Analysing LLI resources is expected to estimate mental states without dependence on learning domain. We implement the model as a system for automatic detection of learners' impasses, which uses three LLI resources including mouse movements, face movements to be close or away, and face-tilting movements and that comprises three Multi Resolution Analyze units and four Neural Network units.

1. はじめに

e-learning へのメンタリングの導入が広がっている。ここでは指導者(メンター)と学習者とがメールや掲示板などを介して密にコミュニケーションをとり、学習内容に関する補足的説明や学習活動を維持するための激励などの支援がなされている。これにより、ドロップアウト率の低下などの成果が報告されている [松田 2005]。しかし、これら運用には人的なコストや時間的なコストの増加という課題がある。

これらの課題に対して、学習者支援を自動化する取り組みが ITS(Intelligent Tutoring System) 研究センターになされてきた。ITS の多くは、学習者モデルにおいて、教材に関する知識等を表現し、学習者が持つ知識や教材に対する理解の程度を診断することが多く、またこれらの構成には、帰納推論などの技術が用いられる事が多い [溝口 1995]。これに加え、さらに学習者の行き詰まりや意欲、集中の度合いなどの心理状態を学習者モデルに取り込むことが出来れば、より柔軟で広範な学習支援の自動化が可能となるだろう。

そこで、本研究では広義の学習支援システム (e-learning システムを含む) において学習者の知識・理解状態、心理状態の両側面を推定する機能を実装し、適切な自動メンタリングを実現するために必要なモデルと技術基盤を開発する。このような学習支援機能を備えたシステムを、知的メンタリングシステム (IMS: Intelligent Mentoring System) と呼ぶことにする。ここで提唱する IMS の特徴のひとつは、学習者モデルの診断機能において学習者の心理状態を考慮する点である。この心理状態は刻々と変化するものであるため、常にモニタリングし即時的に診断とフィードバックができるようにする必要がある。そこで、IMS では既存の ITS 研究による知識・理解状態の診断技術や支援方法 (教授戦略) の決定モデル等に加え、心理状態のリアルタイムな推定結果やそれに基づく支援方法の決定モデルとあわせて統括的な学習者支援を行う (図 1)。このような IMS の実現に向けては、学習者の心理状態の推定機構と、学

習者状態の診断結果に基づく支援方法の決定機構の 2 つの技術的基盤が新たに必要となる。

1.1 本研究の目的

本研究の目的は、IMS 実現のための技術的基盤のひとつである、学習者の心理状態をリアルタイムに推定するシステムを構築することである。ここでは一般的なコンピュータ利用環境への導入を容易にするため、“特殊な装置や操作を必要としない”手法の確立を試みる。学習者の心理状態を推定する指標としては、瞳孔面積などの生理的指標が有用であるが、特殊な装置が必要となってしまう。そこで本研究では、学習者の行動的特徴である Low-Level Interaction リソースに注目し、これに基づく心理状態の推定を試みる。

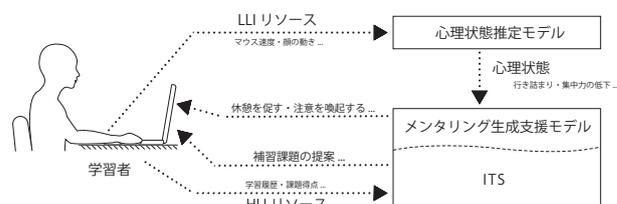


図 1: IMS の概念図

1.2 Low-Level Interaction により期待される成果

インタラクションは様々な粒度で切り取る事ができる。Ryuらは GUI におけるインタラクションを、ユーザとシステムとの循環的なやり取りであると、その最小単位を Low-Level Interaction と定義している [Ryu 2004]。本研究ではこの定義を参考に、マウスの移動速度の変化や、キーボードの打鍵時間間隔、姿勢の変化など、出来る限り細かい粒度でサンプリングした行動的特徴を Low-Level Interaction (LLI) リソースと定義する。反対に、生成された文字列や、その作業に要した時間などのリソースは、サンプリング粒度の粗い High-Level Interaction (HLI) リソースと定義する。図 2 は HLI と LLI の概念を示している。

LLI リソースを利用したアプリケーションの一例として個人

認証がある。認証の指標としては、パスワードなどの文字列やサインなどの画像 (HLI) ではなく、それらが生成される過程に得られるキーボードの打鍵時間間隔 [佐村 2006] やマウスの移動軌跡 [泉 2004] などの時間的粒度の細かな信号 (LLI) が利用されている。これらの LLI リソースは、認証シーケンスだけでなく、日常的なユーザーの作業シーケンスからも連続的にモニタリングが可能である。これにより、これらの認証システムでは作業ドメインに依存しない認証方式も提案している。

さらに LLI の利用範囲は個人認証にとどまらず、個人内での心身的な変化の抽出に対する示唆もなされている [佐村 2006]。IMS においても LLI リソースを用いることで、1) 連続的に、2) 学習ドメインに依存しない心理状態の推定が期待される。

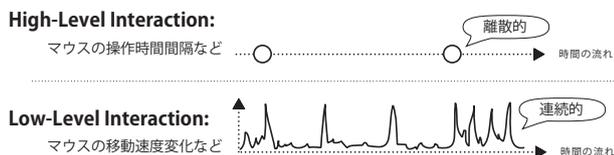


図 2: HLI と LLI の比較

2. 関連研究

学習者の行動的特徴に基づく心理状態の推定を目的とした取り組みの報告が過去にもなされている。中村ら [中村 2002] による研究では、学習者へヒントを掲示する等の支援を行う機能を複数用意し、これらの支援機能が呼び出される時間間隔やスクロール操作などの時間間隔 (HLI) に基づいて、学習者の行き詰まりの推測を行うシステムの提案がなされている。また、中村ら [中村 2007] によって、ステレオカメラを用いて取得した視線と顔の傾きの分散 (LLI) と、マウスクリックなどの端末操作の時間間隔 (HLI) を組み合わせて主観的難易度を推測するシステムが提案された。これらの HLI に基づく研究は、リソースの時間粒度が大きいため実時間推定には限界がある。また [中村 2002] では HLI リソースの取得にあたり、コンテンツ内に特殊な機能を埋め込むという制約がある。[中村 2007] では LLI リソースの取得に特殊な装置を必要としている。そのため、一般的な e-learning 利用環境における導入には更なる工夫が必要である。

そこで著者らは先行研究において、LLI リソースのひとつであるマウスの移動速度に注目し、学習者の異常行動のリアルタイム検知にある程度成功した [堀口 2008]。しかし、マウスをあまり動かさない学習者に対しては有効な推定が行えないという課題が残っていた。そこで、複数の LLI リソースに基づいて、より広範な学習者に適応可能な心理状態のリアルタイム推定を試みる手法を提案した [堀口 2009]。本報告では [堀口 2009] により得られた LLI リソースと心理状態との関連をより詳細に見直し、学習者の行き詰まりの自動推定を行うモデルを提案する。

3. LLI リソースと心理状態の関連

LLI リソースに基づく心理状態の推定モデルを構築するためには、特定の心理状態における学習者の LLI リソースの特徴を探る必要がある。そこで学習者の LLI リソースと心理状態を取得し、両者の間の関連性を抽出するための予備実験を実施した。

3.1 予備実験の手続き

予備実験では、学習者に与える学習課題として、多肢選択式の英文読解問題への解答を設定した。この課題は [堀口 2009] で開発した実験ツールにより提示され、後述する LLI リソースを取得した。さらに学習者には、問題解答の間と後に心理状態の報告を求めた。

LLI リソースは、マウスと顔の動きの信号をそれぞれサンプリングレート 30Hz と 5-7Hz で取得した。マウスの LLI リソースとしては、カーソル座標とクリックの状態 $\{x, y, stat\}$ を、顔の LLI リソースはコンピュータに内蔵された web カメラから画像解析により顔の位置と傾き $\{x, y, z, \theta\}$ を取得した。画像解析にはオープンソースのコンピュータビジョンライブラリである OpenCV*1 を用いた。ステレオカメラを用いていないため、奥行き成分 z の取得には、左右の目の間の距離 w を利用した。そして、学習者の心理状態の情報は、解答中の発話プロトコルと事後のヒアリングを組み合わせて取得した。

この実験は大きく 3 つのステップで構成された。まず学習者は実験ツールを用いて提示される問題に解答した (Step1)。このとき学習者は、考えた事を可能な限り声に出すように教示された。Step1 での学習者の行動は、Web カメラとは別のカメラによって記録された。解答終了後、Step1 の映像を学習者に提示し、その時に考えていたことの報告を求めた (Step2)。問題解答中に考えていたことを忘れてしまわないように、このヒアリングは各問題の解答の終了直後に行った。Step2 でヒアリングした内容を基に、解答時の心理状態の遷移をまとめたシートを、学習者の監修のもと、実験者が作成した (Step3)。以上の 3 つのステップを合わせて 1 試行と呼ぶ。学習者は一連の手続きを繰り返し練習し、十分に慣れた上で本課題に取り組んだ。学習者の時系列の心理状態情報は、Step3 で作成したシートから抽出した。

3.2 心理状態の分類

予備実験では 1 名の学習者を対象に、4 回の練習試行の後に 3 回の本試行を実施し、データを獲得した。学習者から報告された心理状態の情報からは「スラスラわかる」「答えの見当がついた」「単語の意味は分からないけど想像がつく」など、解答行動が順調に進展していると考えられる心理状態や、「単語がわからない」「答えがどっちか迷う」などの、解答行動に何らかの問題が発生していると考えられる心理状態が確認された。後者の心理状態にあると考えられる箇所のうち「何らかのヒントが欲しい」といった助言を求める旨の報告を含むものが、3 回の本試行から 14 箇所抽出された。これらの箇所では「行き詰まり」の状態に陥っていると捉えられる。これに基づいて、本研究では「行き詰まりを」自力での問題解決が困難であると感じており、何らかの助言を求めている状態」と定義し、後述するモデルの推定対象とした。

3.3 LLI リソースの特徴

3.1 節の手続きによって得られたデータから、マウスの移動速度・顔の移動速度を算出し、各位置情報の情報などと併せてそれぞれの挙動の特徴を探る分析を行った。その結果、マウスの軌跡や顔の動かし方に関するいくつかの特徴的なパターンが抽出された。

マウスの軌跡からは、文章や選択肢をトレースする動きが観察された。また、行き詰まりが発生した時間帯においてマウスがトレースしていた場所と、被験者が行き詰まりの原因として報告した箇所は、3 回の試行で観測された 14 箇所全てにおいて一致していることが確認された。

*1 OpenCV Wiki, <http://opencv.willowgarage.com/wiki/>

また、図3に示すマウスの移動速度に注目すると、速い移動と遅い移動が周期的に繰り返されていることが分かる。これは、文章を読み進める際、単語や節などの一定のブロックごとにマウスカーソルをトレースために発生するパターンであると考えられる。このパターン中において移動速度が最大になる点の間隔（ピーク間隔）に注目すると、行き詰まりの発生した時間帯の周辺では、ピーク間隔が狭まる（マウス移動速度の変化周期が早まる）傾向が見受けられる。これは、未知の単語などの行き詰まりの原因となるブロックに出会った際、周辺のブロックや選択肢などを参照し、ヒントを探索している為に発生する現象であると考えられる。同様に、行き詰まりの発生した時間帯においては、顔の前後方向（ z 軸）の動きや傾き（ θ ）にも変化が確認された。

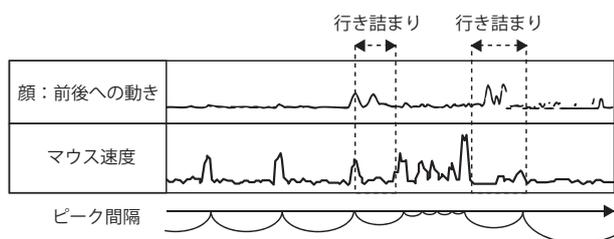


図3: マウスの移動速度・顔の前後への移動速度の変化の特徴

以上の結果から、学習者の行き詰まりにおける LLI リソースの特徴として次の2点を導いた。

特徴 1: 行き詰まりが発生するとマウスの移動速度の周期に変化が生じる（周期が短縮される）

特徴 2: 行き詰まりが発生すると顔の前後移動（ z ）と傾き（ θ ）に変化が生じる（変位量が大きくなる）

4. 推定モデルの構築

予備実験の知見に基づき、各 LLI リソースの特徴的な変化から学習者の行き詰まりを推定するモデルを構築する。このモデルでは、LLI リソースから得られた信号を入力し、行き詰まりの有無を出力する。LLI リソースの取得から推定結果の出力のプロセスは全て独立してバックグラウンドで動作させることで、リアルタイムな推定を実現する。

4.1 モデルの概要

予備実験に従い、1) マウスの移動速度、2) 顔の前後移動、3) 顔の傾きの特徴的な動きを検出するモデルを構築する。推定モデルにはこれらの3つの LLI リソース信号を入力するものとし、それぞれの信号は時間-周波数解析によって LLI リソースの特徴量を抽出する。時間-周波数解析には離散ウェーブレット変換による多重解像度解析 (MRA: Multi Resolution Analyze) を採用した。これらの特徴量から推定結果 (行き詰まりの有無) を出力する識別器には3階層フィードフォワード型のニューラルネットワーク (NN) を採用した。その概念図を図4に示す。推定モデルは、3つの MRA ユニットと4つの NN により構成される。3つの LLI 信号はそれぞれ MRA によって特徴量を抽出し、得られた特徴量から、それぞれ推定結果を NN によって出力する。この3つの NN の推定結果をさらにもうひとつの NN によって評価を行う構造とした。このように複数の LLI リソースの推定結果を総合的に評価することで、推定精度の安定性の向上が期待される。これらの推定は全て逐次的なバックグラウンド処理が可能である。MRA の為の信号の

バッファリングによる遅延は発生するものの、リアルタイムな推定を実現している。

4.2 MRA・NN ユニットの構成

周波数成分は信号の時系列信号の特徴を表す指標である。3.3節で述べた2つの特徴で表される LLI リソースの時間変化も、周波数成分として表現することで明確な特徴量の抽出が期待される。本モデルでは時間-周波数解析として MRA を採用し、3つの LLI リソース信号 {マウスの移動速度、顔の前後移動（ z ）、顔の傾き（ θ ）} の周波数成分を抽出する。なお、先述したように各 LLI リソースごとに独立した MRA ユニットが、あわせて3つ用意される。MRA の離散ウェーブレット変換に用いるウェーブレット関数 ψ とスケーリング関数 ϕ は Haar 関数とし、それぞれ10スケール分のウェーブレット成分（周波数成分）を抽出する。この周波数成分に含まれる周波数は 30Hz-0.05Hz である。

NN は入力層10つ、中間層7つ、出力層1つのユニットにより構成される。この NN は、各 LLI リソース信号の10スケール分のウェーブレット成分を入力し、行き詰まりの推定結果（ $0 \leq x \leq 1$ の値と閾値による判定結果）を出力する。これらの NN の評価結果を総合して評価する NN は、入力層が3つ、中間層が2つ、出力層が1つのユニットにより構成される。入力層には上述の NN の出力結果（ $0 \leq x \leq 1$ ）を入力し、上述の NN 同様の行き詰まりの推定結果を出力する。中間層と出力層の出力関数には、ロジスティック関数を用いる。

NN のトレーニングには、学習者のプロトコルから取得した行き詰まりの箇所の情報を教師信号として用いる。本推定モデルには4つの NN が存在するが、それぞれの NN は全て独立してトレーニングを行う。ただし、総合的な評価を行う NN のトレーニングには、各 LLI リソースごとの NN の評価結果を入力する必要がある。そこで NN のトレーニングは2段階に分けて行う。

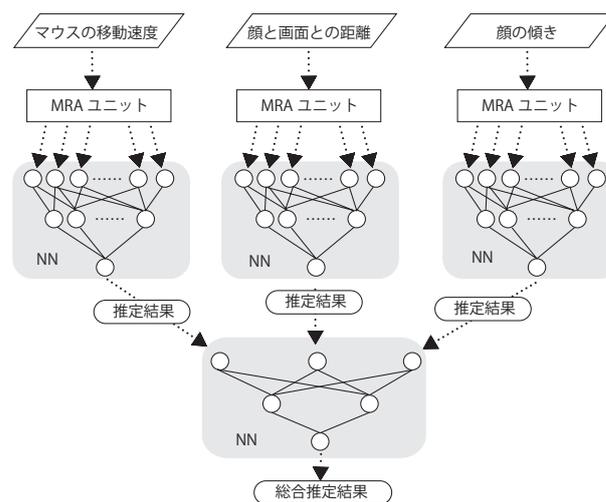


図4: 行き詰まりの推定モデルの概念図

5. モデルの実験的検証

前節のモデルの推定精度を実験により検証した。本実験では、5人の学習者から LLI リソースと心理状態を取得し、本モデルによる推定結果と学習者から報告された心理状態とを比較することで、推定結果を時間的妥当性の観点から評価した。

5.1 検証の手続き

推定モデルの検証は、1) 実験による学習者データの取得と、2) NN のトレーニング、3) 本モデルの出力結果と学習者の報告との照合による評価の3段階で行った。手続き1の学習者データの取得は予備実験と同様の実験環境において、学習者のLLIリソースと学習者の発話プロトコルとヒアリングから学習者の”行き詰まり”の時系列の情報を取得した。実験では3回の練習試行の後、10回の本試行(10問分)のデータを取得した。

取得した5名分のデータを用いて、手続き2のNNのトレーニングを行った。トレーニングには行き詰まりの発生した時間の情報を教師信号として用いた。トレーニングには10問のうち2問のみを使用するした。この2問は行き詰まりの報告を受けた問題の中から、行き詰まりの時間長が適当であると考えられる試行を任意に選択した。

手続き3の検証段階では、LLIリソースとして、i) 全てのLLIリソースを総合した推定結果、ii) マウス移動速度のみを用いた推定結果、iii) 顔の傾きのみを用いた推定結果、iv) 顔と画面との距離のみを用いた推定結果の4パターンの推定結果の出力を行った。本研究では $P(T)/P(E)$ の値(行き詰まり時の適正推定比率/行き詰まり時以外の誤推定比率)を評価指標として用いた。 $P(T)/P(E)$ の値が大いほど推定が正確であったということになり、 $P(T)/P(E) = 1$ だとランダムな推定と等価となる。

5.2 結果

図5は各学習者ごとのi~ivの推定結果に対する $P(T)$ 、 $P(E)$ の値を表している。学習者Aについては、マウス移動速度(ii)と顔の前後移動(iv)を用いた推定結果において、総合的な推定結果よりも $P(T)/P(E)$ の値が高くなった。特に、学習者Aの問題番号4においては、マウス移動速度(ii)による推定結果の $P(T)/P(E) = 4.33$ と非常に良好な結果であった。これは、学習者Aのマウス軌跡に文章をトレースする行動が含まれていたためであると考えられる。一方、学習者Aの顔の傾き(iii)に基づく推定結果については $P(T)/P(E)$ の値が0.5を下回っており、有効な推定とは言えない。学習者Bでは学習者Aとは対照的に、顔の傾き(iii)による推定結果が良好であるが、マウス移動速度(ii)と顔の前後移動(iv)の推定結果の $P(T)/P(E)$ 値が低い。

総合的な推定結果(i)において、高い評価値を示した学習者Cは、個々のLLIリソースによる推定結果においても良好な推定が行われている。特に顔と画面との距離のLLIリソースでは $P(T)/P(E) = 10.77$ と、非常に高い評価値を示している。その一方で、学習者Cは問題解答時にほとんどマウスカーソルを動かすことがなかったため、マウスからの推定には失敗している。しかし、総合推定結果(i)を見ると、非常に正確な推定に成功したといえることができる。

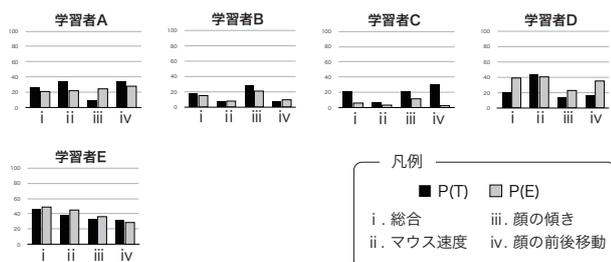


図5: 検証実験の結果

5.3 考察

複数のLLIリソースのMRAによる特徴量を用いる本手法によって、[堀口 2008]の課題であったマウスをほとんど操作しない学習者(学習者C)に対しても $P(T)/P(E)$ 値の平均3.45の推定精度を得た。また、マウスで文字をトレースする行動が確認された学習者(学習者A)に対しては、マウスLLIのみによる推定結果の $P(T)/P(E)$ 値が最大4.33を示すなど、良好な結果が確認され、IMS実現への可能性を見出した。しかし、総合推定結果(i)の $P(T)/P(E)$ 値の全体平均は1.48であり、良好な推定精度を得られているとは言えない。これには、NNのトレーニングの最適化、学習者の心理状態の取得方法等の問題があると考えられる。

6. おわりに

本研究では、IMS実現に向けた、複数のLLIリソースに基づく学習者の行き詰まりのリアルタイム推定システムを提案した。しかし、NNのトレーニングの最適化や学習者心理状態の取得方法の検討などの課題が残った。今後の展望として、推定システムの更なる精度の向上と、推定可能な心理状態の多様化、キーボード等のLLIリソースの拡大による対応コンテンツの拡充が挙げられる。

参考文献

- [松田 2005] 松田 岳士, 本名 信行, 加藤 浩: eメンタリングガイドラインの形成とその評価, 日本教育工学会論文誌, Vol.29, No.3, pp.239-250 (2005)
- [溝口 1995] 溝口 理一郎: 知的教育システム, 情報処理, Vol.36, No.2, pp.177-186 (1995)
- [Ryu 2004] Hokyong Ryu, Andrew Monk; Analysing interaction problems with cyclic interaction theory: Low-level interaction walkthrough, PsychNology Journal, Vol. 2, No. 3, pp. 304-330 (2004)
- [佐村 2006] 佐村 敏治, 西村 治彦: テキスト入力によるキーストロークダイナミクス, 情報知識学会誌, Vol.16, No.2, pp.63-68 (2006)
- [泉 2004] 泉 正夫, 長尾 若, 宮本 貴朗, 福永 邦雄: マウス操作の特徴を用いた個人識別システム, 電子情報通信学会論文誌, Vol.J87-B, No.2, pp.305-308 (2004)
- [中村 2002] 中村 喜宏, 赤松 則男, 桑原 恒夫, 玉城 幹介: 操作時間間隔の変動に着目したCAI学習の行き詰まり検知方法, 電子情報通信学会論文誌.D-1, Vol.J85-D-1, No.1, pp.79-90(2002)
- [中村 2007] 中村 和晃, 角所 考, 美濃 導彦: e-learning環境における学習者の観測に基づく主観的難易度の推定, 画像の認識・理解シンポジウム (MIRU2007) (2007)
- [堀口 2008] 堀口 祐樹, 小島 一晃, 松居 辰則: e-learningにおける学習時の潜在的な意識変化の抽出, 2008年度人工知能学会全国大会(第22回)論文集 (2008)
- [堀口 2009] 堀口 祐樹, 小島 一晃, 松居 辰則: e-learningにおける学習者のマウスの動きと顔の動き情報を用いた行き詰まりの検出手法の提案, 第56回先進的学習科学と工学研究会資料 (SIG-ALST-A901), pp.21-26(2009)