

# 読文過程の行動特徴量による習熟度の予測

児玉 翠<sup>\*1</sup>, 秦 崇洋<sup>\*1</sup>, 渡邊 淳司<sup>\*2</sup>, 丸谷 和史<sup>\*2</sup>, 社家 一平<sup>\*1</sup>

\*1 NTT 未来ねっと研究所, \*2 NTT コミュニケーション科学基礎研究所,

## Performance Prediction using Behavioral Data during Reading Texts for Adaptive Learning

Midori Kodama<sup>\*1</sup>, Hata Takahiro<sup>\*1</sup>, Junji Watanabe<sup>\*2</sup>, Kazushi Maruya<sup>\*2</sup>, Ippei Shake<sup>\*1</sup>

\*1 NTT Network Innovation Laboratories, \*2 NTT Communication Science Laboratories,

In the e-learning system, estimating students' performances is an important challenge to provide personalized learning materials (drill) based on their skill level. Although previous studies estimated their performances from their own answers for questions in the drill, in order to provide personalized questions from the beginning of the drill, we have to estimate their performances not only from their answers but also other data about their behavior. In this study, we collected the students' behavioral data when they read 'Yu bi Yomu'[6,7] e-textbook, by which we could monitor a location where they read. Then we estimated their performances using support vector machine from behavioral feature quantities selected by lasso regression. As a result, students' performances (upper/lower) were estimated with high accuracy (81%), which demonstrates an availability of behavioral data for improving adaptive learning systems.

キーワード: アダプティブラーニング, 行動特徴量抽出, 機械学習

### 1. はじめに

アダプティブラーニングとは、個々の生徒の習熟度や学習進度に合わせて、学習内容やレベルを調整し教材を提供する手法であり、先行研究によりその効果が示されている [1].

学習者に最適な難易度の問題を出題するためには、学習者の習熟度を推定することが重要な課題である。学習者の習熟度はヒトの内的な状態であり、直接測定することはできない。これまでの研究では、ある問題を提示した際の問題回答データを利用して、習熟度を推定することが行われてきた[2,3]。習熟度が向上する過程には、学習内容に関する情報に触れ、それを繰り返す学習経験が介在するため、このような学習経験から習熟度を推定できる可能性がある。実際、今までにどんな問題を解いてきたかという OC (Opportunity Count) という指標を用いて習熟度を予測する研究も存在する[4].

しかし、学習経験には、問題を解くことだけではなく、教科書や参考書を読んで知識を獲得する経験も存在する。そこで本検討においては、知識獲得時における学習者の学習行動に着目し、知識を獲得する場面における学習者の振る舞いから学習者の習熟度を予測することを目指す。

もし知識を獲得する場面での測定データから習熟度を予測できるならば、問題回答データが十分に計測できていない問題演習の初期にも、個人の習熟度にあったアダプティブなコンテンツが提供できる。加えて、従来通りの問題回答データとは異なる情報源である知識獲得時の行動特徴量を従来の特徴量と合わせて用いる事で、さらなる習熟度推定の精度が上がる事も期待される。

学習者の学習時の様子をモニタリングすることは、これまでも行われており、例えば視線追尾によって、文章を読む過程での思考を推定する研究も存在する [5]。しかし、視線の測定には専用の装置が必要であり、

使用者への負荷が大きく学習を阻害する虞がある。今回、我々は使用者の負荷を抑えて行動特徴量を計測するために、なぞることで文字が浮き出る電子教材を用い、タブレット以外に装置を追加することなく、利用者が文を読む過程を観測した。

本検討においては、行動特徴量を用いた習熟度予測の一次評価として、習熟度高低の2値分類の可否を検証する。習熟度は細かい粒度で予測できることが望ましいが、高低の2値であっても演習で出題する問題の難易度調節や、再度講義の受講を促す等の学習指導に利用可能である。

## 2. 方法

学習中の行動からの習熟度推定の可否、および行動特徴量と習熟度の関連性を調べるため、学習中の振る舞いを測定するデータ取得実験[6]を実施し、学習後のテスト成績との関連を分析した。

### 2.1 教材表示方法

本実験では、画面をなぞるユーザの指位置を計測し、その位置に対応する文字が出現・消滅する動的な文章表示方式‘Yu bi Yomu’[7]を用いて、タブレット端末(Apple社製 iPad)上で教材を提示した。この方式の教材を用いれば、ユーザは教材を見るために画面をなぞるため、ユーザの指の位置からユーザの読文過程を詳細に知ることができる。

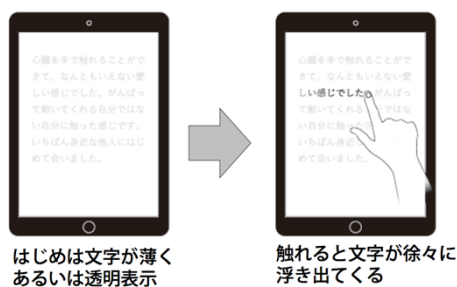


図1 教材表示装置

### 2.2 データ取得実験

データの取得実験として、入社2年目の社員47名を対象とした講義形式の社内研修において、Yu bi Yomuを導入し、研修中の読文行動の測定をおこなった。実験プロトコルを図2に示す。研修は全体で約65分あり、研修内容は通信網敷設における安全研修であった。なお、講義内容は全11章により構成されてい

た。それぞれの章において、教室のスクリーンに投影された講義スライドを用いた講師による説明の後、講師の指示のもとタブレット教材を用いた振り返りの時間を設けた。振り返り時間には講師がタブレット教材の文章を読み上げ、学習者へ講師の掛け声に合わせてタブレット教材の該当箇所をなぞるように指示した。各章の講義と振り返りが終了したのちに、全章の内容を総復習する時間を20分ほど設けた。その際、タブレット教材のうち、空欄をなぞることで重要語句が表示されるモードもしくは、空欄部分があらかじめ表示されているモードを用いて復習するように指示した。総復習時間の後に、本講義内容についての記憶テストを実施した。テストの内容は、教材中の重要な単語部分を記述式で答えさせるものであり、各問1点配点で71点満点であった。記憶テスト終了後、復習時間における学習方法や、タブレット教材の使用感についてのアンケートを実施した。

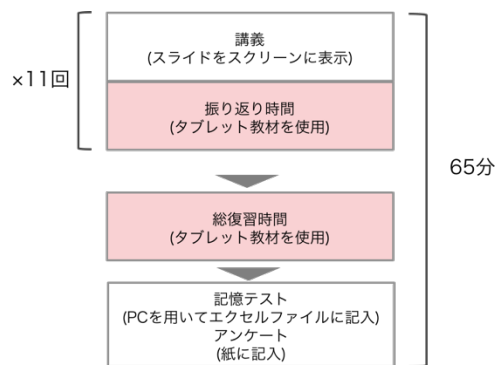


図2 研修プロトコル

### 2.3 分析方法

#### 2.3.1 取得されたデータ

本実験においては、タブレット教材使用時の行動データ取得した。すなわち、データを取得したのは各章の振り返り時間と総復習時間であって、かつタブレット教材を学習者が使用した場合に限り、スクリーンに投影されたスライドを見ている様子や、学習者がタブレット教材を使わずに復習をしている様子は計測しなかった。取得された行動データの各レコードには、時間、X座標位置、Y座標位置、X軸速度、Y軸速度、合成速度、スクロール位置が含まれている。なお、X軸はタブレットの左上を原点とした時の右方向のピクセル数、Y軸は下方向のピクセル数を表している。なお、各計測値は、サンプリングレート30Hzでされた。

### 2.3.2 前処理

記憶テストの後に行われたアンケートにおいて、タブレットのなぞり表示機能を使用しなかったと回答した被験者 8 名は、学習行動が測定されていないと判断し、分析の対象より除外した。アンケートの内容と、アンケート結果を表 1 に示す。また、機械のトラブルにより測定データが保存されなかった被験者 1 名および、指示に反する誤操作があった被験者 1 名は、分析対象より除外した。以後の分析においては、残りの 37 名を分析対象とする。

表 1 学習方法に関するアンケート

問.テスト前にどのように暗記しましたか	人数	点数平均	点数標準偏差
1.一度表示を消して指でなぞりながら暗記した	23	42.3	9.4
2.全体を表示して指なぞりを使わず暗記した	8	39.3	6.8
3.全体の表示と、指でなぞることと両方を使って暗記した	15	41.8	8.7

### 2.3.3 特徴量抽出

本実験により得られるデータより、測定できる学習行動は、1) 情報(学習内容)へのアクセスの有無、2) 情報へのアクセスの継続時間、3) 情報へのアクセスの繰り返し、4) アクセスの順序であると仮定し、それらを説明する行動特徴量を抽出した(表 2)。なお、本教材のように学習者が閲覧している部分を行や座標レベルの解像度で得られる場合に算出可能な特徴量を行単位の列に、それ以外の特徴量をページ単位の列に示す。

表 2 抽出した行動特徴量

	ページ単位	行単位
情報へのアクセス有無	・閲覧ページ数	・なぞり行数 -全ページ合計 -各ページの行数
情報アクセスの継続時間	・閲覧時間 -全ページ合計 -各ページ合計  ・閲覧時間(行数補正) -全ページ平均、分散 -各ページの値	・行なぞり時間 -全ページ合計 -各ページ合計 ・閲覧スピード(横方向) -全ページの平均,標準偏差 -各ページの平均,標準偏差 ・閲覧スピード(縦方向) -全ページの平均,標準偏差 -各ページの平均,標準偏差
アクセスの繰り返し	・ページ繰り返し回数 ・ページ再閲覧時間	・行繰り返し回数
アクセスの順序		・順方向への連続したなぞり回数 ・逆方向への連続したなぞり回数

### 2.3.4 統計分析, 変数選択

各特徴量が、記憶テストの点数の高い群(42 点以上, 17 名,上位群)と低い群(41 点以下, 20 名,下位群)において有意な差があるか、Kolmogorov-Smirnov 検定[8]を用いて検定した。なお、Shapiro-Wilk の正規性検定により約半数の特徴量の正規性が棄却された( $p < 0.05$ )ため、ノンパラメトリックな手法である Kolmogorov-

Smirnov 検定を用いた。

記憶テスト点数への寄与の大きい行動特徴量を調べるため、各行動特徴量を説明変数、記憶テストの点数を目的変数として、多重回帰分析を実施した。多重回帰においては、変数の数が膨大な場合においても効率的に変数選択ができる L1 型正則化法を用いた[9]。なお、多重回帰分析の前処理として、各特徴量の正規化(z スコア化)を実施した。

### 2.3.5 分類学習

上位群,下位群の 2 群を行動特徴量により分類する。分類器の学習には、サポートベクタマシン(Support Vector Machine : SVM)を用いた。カーネル関数の種類、SVM における損失とマージンの大きさの間のトレードオフを調整するパラメータである C, 決定境界の複雑さを調節する gamma, 多項式カーネルの次数を調節する degree をそれぞれ,表 3 の範囲で探索し、グリッドサーチ法により決定した。

表 3 グリッドサーチ法のパラメータの範囲

カーネル関数の種類	ハイパーパラメータ		
	C	gamma	degree
線形カーネル	[1,10,100,1000]	-	-
RBFカーネル	[1,10,100,1000]	[0.001, 0.0001]	-
多項式カーネル	[1,10,100,1000]	[0.001, 0.0001]	[2, 3, 4]
シグモイドカーネル	[1,10,100,1000]	[0.001, 0.0001]	-

このようにハイパーパラメータを決定した SVM 分類器について、記憶テストの高群/低群を行動特徴量から分類する際の性能を Leave one out 交差検証により確かめた。なお、SVM の入力には、2.3.4 項の多重回帰(L1 型正則化法)の結果、偏回帰係数の絶対値が 0 より大きい変数を選定した。

## 3. 結果

### 3.1 ページの時間遷移の可視化

研修中のユーザのページ遷移の様子を図 3a, 3b に示す。上位群も下位群も講義の進行に従ってページを遷移させており、上位群と下位群との間で大きな違いは見られなかった。

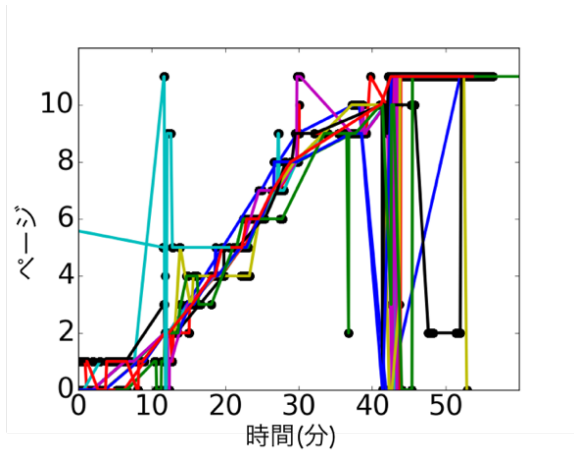


図 3a 上位群のページの時間遷移  
(各色は各被験者を示す)

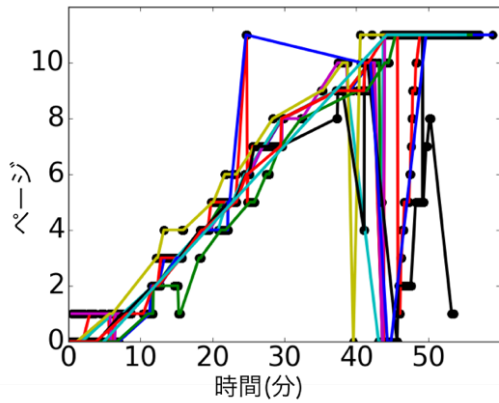


図 3b 下位群のページの時間推移

### 3.2 なぞりの様子の可視化

ユーザが教材をなぞる様子について、代表的な2名分の教材1ページ目のデータを図4a, 4bに示す。プロットは30Hzで計測される指位置を示す。プロットのカラースケールは時間を示し、青から緑に色が変化するにつれて遅い時間に計測された点を意味する。該当ページについてのデータのうち、連続して記録されたレコードは、灰色の直線で点同士を結んで表示している。図4aのような上位群の被験者は、全体を網羅してなぞる、長い時間をかけてゆっくりなぞる、同じ箇所を複数回なぞるといった行動が見受けられ、一方、図4bのような点数の下位群の被験者のなかには、時間をかけずに早くなぞる、一部しかなぞらない、といった行動が観測された。

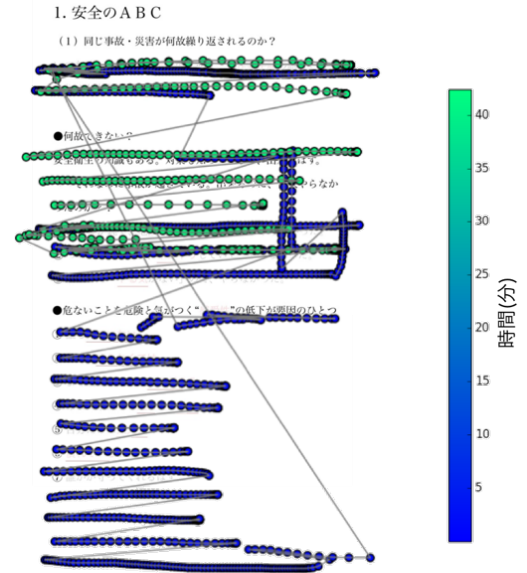


図 4a 57点(上位群)の学習者のなぞりの様子

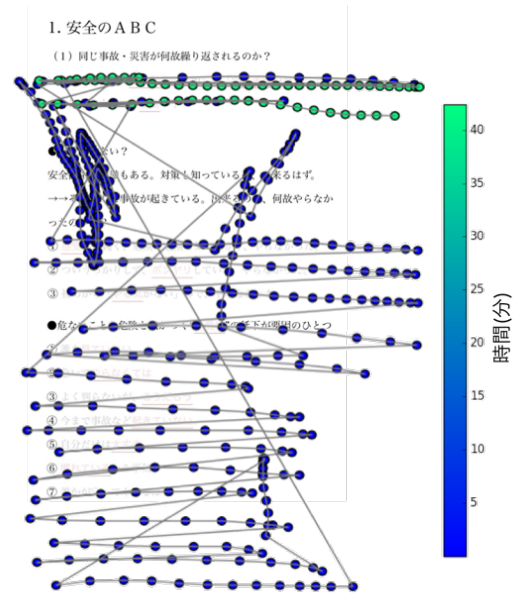


図 4b 27点(下位群)の学習者のなぞりの様子

### 3.3 なぞり行の時間変化の可視化

総復習時間中になぞった行の時間推移を図5a, 5bに示す。図5aは上位群の代表4名分のなぞり行の時間推移を示す。順方向に連続して読むことを何度も繰り返している様子が観測された。図5bは下位群の代表4名分のなぞり行時間推移を示す。一部しかなぞらない様子や、逆方向に教材をなぞる様子、使用時間が短い様子などが見られた。

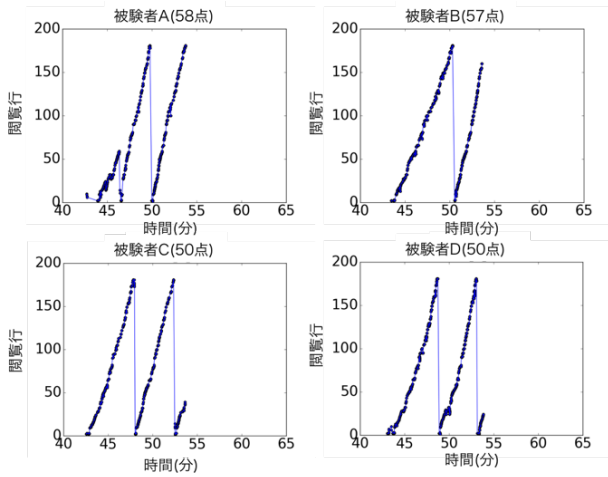
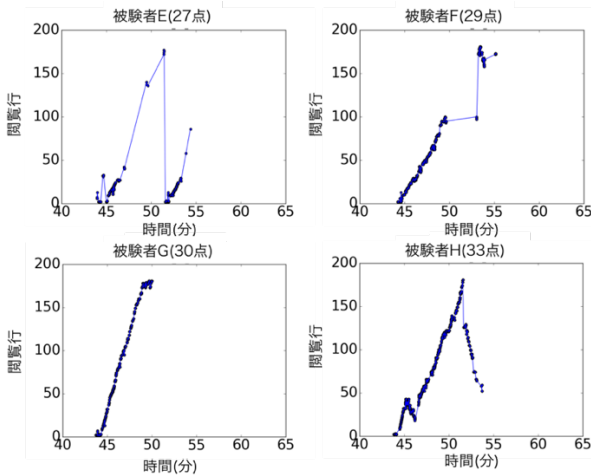


図 5a 上位群のなぞり行時間推移



5b 下位群のなぞり行時間推移

### 3.4 統計分析

表 4 に記憶テストの得点と各特徴量得点の重回帰分析の偏回帰係数と、記憶テスト得点の上位群と下位群の二群に分け、Kolmogorov-Smirnov 検定を実施した際の統計量および  $p$  値を示す。重回帰分析の結果、順方向への連続したなぞり回数の偏回帰係数が大きく、また、Kolmogorov-Smirnov 検定の結果、X 方向なぞりスピード平均は上位群と下位群で有意な差があることが示された。

表 4 特徴量の重回帰、統計分析結果

	特徴量	偏回帰係数	KS統計量	$p$ 値	
ページ単位	情報アクセスの有無	閲覧ページ数	0.0	0.18	0.91
	情報アクセスの継続時間	総閲覧時間	0.0	0.35	0.17
		閲覧時間(行数補正)平均	0.0	0.27	0.43
	アクセスの繰り返し	閲覧時間(行数補正)標準偏差	0.0	0.27	0.43
		ページ繰り返し回数	0.024	0.25	0.56
	ページ再閲覧時間	0.71	0.17	0.92	
行単位	情報アクセスの有無	なぞり行数	0.0	0.30	0.33
	情報アクセスの継続時間	行なぞり時間	0.032	0.36	0.14
		X方向なぞりスピード平均	-0.66	0.45	0.033
		X方向なぞりスピード標準偏差	-0.13	0.27	0.43
		Y方向閲覧スピード平均	0.020	0.13	0.99
		Y方向閲覧スピード標準偏差	0.032	0.24	0.60
	アクセスの繰り返し	行繰り返し回数	0.0	0.34	0.20
アクセスの順序	順方向への連続したなぞり回数	3.3	0.40	0.081	
	逆方向への連続したなぞり回数	0.0	0.06	1.00	

なお、本発表においては主に教材全体を通じて観測される行動特徴量について記述し、ページごとに測定されるパラメータについての詳細は割愛する。

### 3.5 分類、予測

多重回帰分析にて偏回帰係数の絶対値が 0 よりも大きい 8 変数選択し、これらを入力として SVM 分類器により上位群/下位群を分類した。なお、グリッドサーチ法により、カーネルの種類およびハイパーパラメータを、RBF カーネル、 $C=1000$ ,  $\gamma=0.0001$  と決定した。分類器の評価に leave-one-out 交差検証を実施し、クラス分類の混同行列を表 5 に示す。予測と実測の合致する正解率が 0.81 での分類がなされた。また、精度と再現率の調和平均である F 尺度は上位群が 0.80, 下位群が 0.82 となった。

表 5 分類の混同行列

	予測上位	予測下位
実測上位	14	3
実測下位	4	16

## 4. 考察

統計分析の結果、X 方向のなぞりスピード平均が上位群と下位群で有意な差があることおよび、重回帰の結果より、順方向への連続したなぞり回数の偏回帰係数が大きかったことから、本実験の研修内容においてゆっくり読むことや順方向に繰り返し読むことが習熟度に関連が強いことが示された。なお、これらの特徴量は、本教材のように学習者の閲覧箇所を行や座標レベルの解像度で測定可能な場合に算出可能である。閲覧履歴の測定粒度は、従来ではページレベルの測定が行われているが[10]、さらに細かい行や座標レベルの測定により、アダプティブラーニングや学習モニタリングに効果的な特徴量を得ることができると考えられる。

X 方向なぞりスピード平均以外の特徴量は Kolmogorov-Smirnov 検定では有意にならなかったものの、回帰分析(L1 型正則化法)により選択された 8 つの特徴量を用いた RBF カーネルの学習により、8 割程度の正解率で上位群/下位群が予測された。習熟度は単一の行動特徴量の大小では記述できず、複数の特徴量による判定を組み合わせることではじめて分類可能

## 参 考 文 献

になるものだと考えられる。この背景には、最終的な習熟度は、実験中の学習のみならず、事前の知識にも影響を受けること、また、学習中の振る舞いは、個人の学習の戦略にも影響を受けることが想定される。本実験は、入社して間もない事前知識の少ない社員を対象としたものであるが、講義の内容は一般常識や普段の業務中の中で獲得されうる内容のものであった。今後、事前の知識の影響を排除して行動特徴量と習熟度の関係を検討するためには、事前の知識レベルの調査や統制が必要である。

また、本実験は教材文章中の単語を暗記し、記述式の記憶テストにて回答させるものであった。しかし、数学や理科のように、暗記だけではなく理論立った理解が必要な学習内容も存在し、そのような学習内容の場合は学習中の行動の性質も異なることが予想される。本技術がどのような教科、どのような学習内容に適応可能なのかは、今後検証が必要である。

## 5. おわりに

より早い時期のアダプティブラーニング機能の提供や、習熟度推定の精度向上を目指し、従来通りの問題回答データとは異なる情報源である知識獲得時の行動特徴量を用いて習熟度の予測可否を検討した。結果、知識獲得時の行動特徴量から 81%の正解率で習熟度の上位群/下位群が予測された。また、本実験の研修内容においてゆっくり読むことや順方向に繰り返し読むことが習熟度に関連が強いことが示された。今後は、異なる学習内容での適応可能性の検討および、問題回答データと組み合わせた場合の推定精度評価をおこなっていききたい。

## 謝辞

本研究を遂行するにあたり、NTTME ネットワークサービス事業本部 羽柴正治氏、高橋博之氏にご協力頂きました。

- (1) Chen, Chih-Ming, Hahn-Ming Lee, and Ya-Hui Chen: "Personalized e-learning system using item response theory.", *Computers & Education*, 44(3), pp237-255 (2005)
- (2) 月原由紀, 鈴木敬一, 廣瀬英雄: "項目反応理論による評価を加味した数学テストと e-learning システムへの実装の試み", *コンピュータ & エデュケーション*, Vol.24, pp.70-76 (2008)
- (3) Kevin H Wilson *et al.*: "Back to the basics: Bayesian extensions of IRT outperform neural networks for proficiency estimation.", *arXiv preprint arXiv:1604.02336* (2016)
- (4) Yan Wang, Korinn Ostrow, Seth Adjei, and Neil Heffernan: "The Opportunity Count Model: A Flexible Approach to Modeling Student Performance." In *Proceedings of the Third ACM Conference on Learning @ Scale (L@S '16)*, ACM, New York, pp.113-116 (2016)
- (5) 花房亮, 山岸秀一, 松本慎平, 加島智子: "機械学習処理に基づいたプログラミング読解中の視線軌道の自動分類", *人工知能学会全国大会論文集*, 3N3-2 (2015)
- (6) Kazushi Maruya, Junji Watanabe, Hiroyuki Takahashi, and Shoji Hashiba: "A learning system utilizing learners' active tracing behaviors.", In *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge (LAK '15)*, ACM, New York, pp.418-419 (2015)
- (7) Kazushi Maruya *et al.*: "Yu bi Yomu: Interactive reading of dynamic text." In *Proceedings of the 20th ACM international conference on Multimedia*. ACM, p.1499-1500 (2012)
- (8) Massey Jr, Frank J: "The Kolmogorov-Smirnov test for goodness of fit." *Journal of the American statistical Association* Vol.46, No.253, pp.68-78 (1951)
- (9) 廣瀬慧: "Lasso タイプの正則化法に基づくスパース推定法を用いた超高次元データ解析.", *数理解析研究所講究録*, Vol.1908, pp.57-77 (2014)
- (10) 三原宏一郎, 寺邊正大, 橋本和夫: "ページ閲覧時間を考慮した Web ログマイニング手法の提案.", *情報処理学会研究報告知能と複雑系 (ICS) 2007.67 (2007-ICS-148)*, pp.39-44 (2007)