

視線に基づく英語多肢選択問題解答の確信度推定

山田 健斗^{1,a)} Olivier Augereau^{1,b)} 黄瀬 浩一^{1,c)}

1. はじめに

近年はグローバル化の影響から、英語学習の必要性が非常に高まっている。英語学習では、単語や文法を暗記し、テストを通して学習の成果を確認した後、復習するという流れが一般的である。復習することによって、新たな知識を獲得し、曖昧な知識を補完できる。そのため、復習は非常に重要な学習プロセスであるといえる。

一般的に、人々は間違えた問題を復習する。しかし、正解した問題にも復習が必要な場合がある。具体的には、自身の解答が正解であると確信が持たずに解答し、偶然正解した場合である。このような問題は、正解したことに気を許してしまい、復習の機会が失われがちである。しかし、復習する際、このような問題を見落としては、十分な学習効果が得られない。すなわち、間違えた問題だけでなく、解答が正解であると確信を持てなかった問題に対しても復習を行う必要がある。さらには、復習する問題の中でも、優先順位をつけることが重要である。問題に間違える際には、その背景に知識の記憶違いが存在する。例として、確信を持って解答したにも関わらず不正解であった問題では、正解に必要な知識を勘違い、もしくは記憶違いしていることが考えられる。この勘違いや記憶違いが深刻であるほど、より多くの問題に間違える原因となりうる。そのため、このような勘違いや記憶違いに対する復習を最優先で行う必要がある。これにより、確信を持って間違えたものを最優先とする点、確信なく正解したものを復習に含める点で、通常の復習とは大きく異なり、効率の良い復習に繋がる。以上より、解答時に確信を持って解答したかどうか分かれば、見落としがなく、優先順位を考慮した復習が期待でき、学習において大変有益である。

解答に対する確信度を考慮に入れた復習の方法として、解答時に確信を持てなかった問題に印を付け、後からその箇所を復習するというものがある。しかし、この方法には、印を付けながら解答するために問題に集中できず、学習の成果を確認するというテスト本来の役割を十分に果たせな

い問題がある。これに対して、問題解答時の確信度を自動で推定し、解答者にフィードバックすることができれば、問題解答時の振る舞いを妨げることなく復習に活用できるため、有益である。そこで、本研究では、上記の問題を解決するため、確信度を自動で推定し、解答の正誤と合わせて解答者に提示するシステムの実現を目指す。

確信度を推定する方法として、視点情報の利用に着目する。これまでの研究から、視点情報には言語能力や確信の程度に関する情報が含まれる事が分かっている [1][2]。例えば、人は読んでいる文書が難しいと感じると、読む速度が遅くなったり、何度も読み返す傾向がある。視点情報はアイトラッカというデバイスによって容易に取得できる。さらに、問題解答時の妨げになりづらいという利点がある。

確信度を推定する対象として様々なものが考えられる。本研究では、確信度推定の第一歩として、多肢選択問題を取り扱うこととする。その理由は、多肢選択問題では問題文や複数の選択肢といった見るべき対象が多く存在するため、確信度に応じた視点情報が如実に現れると予想されることである。

そこで、本論文では、解答者の視点情報を用いて、多肢選択問題に対する確信度を推定する手法を提案する。提案手法では、問題解答時の視点情報から特徴量を抽出し、SVMを用いて学習した後、識別する。提案手法の有効性を検証するため実験を行い、推定精度を算出する。提案手法による推定精度は 90.1%であった。本論文では、実験によって得られた結果について考察する。なお、本研究は、大阪府立大学工学研究科倫理委員会の承認を得ている事を付記しておく。

2. 関連研究

視点情報が文書の読み方や理解度と関係があることがこれまでの研究から分かっている。本節では、読書時の視点情報や問題解答時の確信度に関する研究について述べる。

大社らの研究 [3] に、視点情報を利用してユーザが難しいと感じた文書内の箇所を推定するものがある。難しいと感じた箇所を推定するために、読み返しの回数や読む速さなどの特徴量を用いているが、精度は十分でない。

問題解答時の視点の振る舞いに関する研究に、Tsai らの

¹ 大阪府立大学大学院工学研究科

a) kento@m.cs.osakafu-u.ac.jp

b) augereau.o@gmail.com

c) kise@cs.osakafu-u.ac.jp

研究 [4] や, Ishimaru らの研究 [5] がある. Tsai らは, 四択問題における問題と選択肢それぞれに対する時間のかけ方や, 正答者と誤答者との視点の動きに違いがあることを示した. Tsai らの研究により, 問題で問われているものに関係が深い箇所ほど長時間注視され, 誤答者は正答者と比べて問題文に読み戻す傾向が強く見られることが分かった. Ishimaru らは, 物理に関するテキストを読み, 問題を解く際の視点情報を, 学習者の物理に対する習熟度で分けて調査した. この研究により, 学習者の理解の深さと, テキストを読む際の視点情報や時間のかけ方には相関があることが明らかになった.

問題解答時の確信度に関する研究として, 小島らの研究がある [2]. 小島らは雑学や一般常識に関する四択問題を取り扱っており, 解答者がどの程度確信をもって解答したのかを確信度と定義している. 小島らは, 解答者の視点情報から, 確信度を推定する手法の考案を目的に, 解答時の視点の動きについて調査した. これによって, 確信度に応じて視点の動きに一定のパターンがあることを明らかにした. しかし, 確信度を推定するには至っていない.

3. 提案手法

3.1 提案手法の流れ

提案手法の流れについて述べる. まず, ディスプレイ上に英語の問題を表示し, 解答する際の視点情報をアイトラッカで取得する. 次に, 得られた視点情報に対して, fixation-saccade 判定を行う. fixation-saccade 判定後の視点情報を利用し, 特徴量を抽出する. 抽出した特徴量から, 確信度の推定に有効なものを選択し, SVM を用いて確信度を推定する. ここで, 確信度としては最も基本的なもの, すなわち, 確信の有無を用いる. これらの手法の内, 特徴量の選択と確信度の推定は, 被験者ごとに行う. その理由は, 問題解答時の視点情報は, 被験者の癖や問題に対する戦略などに左右されると予想されることである. 以下に詳細を述べる.

3.2 視点情報の取得

アイトラッカを用いてユーザの眼球運動を測定し, 視点情報を取得する. アイトラッカとは, 赤外線を照射することでユーザの瞳孔の位置を捉え, 眼球運動を測定する装置である. 本手法では, Tobii EyeX というディスプレイ装着型のものを用いる. これにより, ディスプレイ上に表示された問題を解答している際の視点情報を取得する.

3.3 fixation-saccade 判定

眼球運動は fixation, saccade, blink の 3 種類に分類される. fixation は視点が一定時間留まることを表し, saccade は fixation 間の素早い視点の移動を表す. blink は, 瞬きのことである. 人は fixation と saccade を繰り返すことで

表 1 特徴量

No.	feature
1	fixations count on the choices
2	fixations ratio on the choices
3	fixations count on the sentence
4	fixations ratio on the sentence
5	total fixation duration on choices [ms]
6	average fixation duration on choices [ms]
7	max fixation duration on choices [ms]
8	min fixation duration on choices [ms]
9	total fixation duration on sentence [ms]
10	average fixation duration on sentence [ms]
11	max fixation duration on sentence [ms]
12	min fixation duration on sentence [ms]
13	variance of fixation in x axis [px]
14	variance of fixation in y axis [px]
15	total saccade length [px]
16	average saccade length [px]
17	whole saccade count
18	saccade count between the choices
19	saccade count on the sentence
20	saccade count between the choice and sentence
21	total saccade duration [ms]
22	average saccade duration [ms]
23	max saccade duration [ms]
24	min saccade duration [ms]
25	total saccade velocity [px/ms]
26	average saccade velocity [px/ms]
27	max saccade velocity [px/ms]
28	min saccade velocity [px/ms]
29	reading-time [s]
30	correctness

物を見たり文章を読んだりしている. 本手法では Buscher らの手法 [6] を用いて fixation-saccade 判定を行う.

3.4 特徴量の抽出

fixation-saccade 判定後の視点情報から, 特徴量を抽出する. 使用する特徴量を表 1 に示す. 特徴量は, 視点情報を表すものとして, fixation 関連のものが 14 個, saccade 関連のものが 14 個である. その他, 視点情報に関わりのない特徴量として, 問題の解答時間と解答の正誤が 1 個ずつ, 合計 30 個である. 一般的に, ユーザが問題を解答する際の振る舞いとして, 悩んだ選択肢に視点が集まり, そうでない選択肢には視点が集まらない傾向がある. そこで, fixation と選択肢との対応を取るために, 各選択肢を囲う矩形を図 1 のように設定する. この矩形に含まれる fixation は, その選択肢, あるいは問題文に対するものとして判定する. アイトラッカの精度の都合上, 視点が対象の文字列から多少ずれることを考慮し, 矩形を大きめに設定してある.

3.5 特徴量の選択

表 1 に示した 30 個の特徴量は, 確信度の推定を行う上で有効であるか自明ではないため, 特徴量を選択する必要

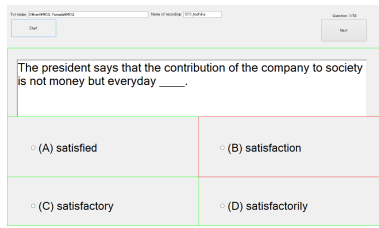


図 1 矩形の設定

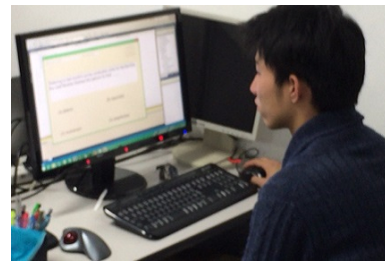


図 2 実験の様子

がある．そこで，特徴選択手法として遺伝的アルゴリズムを用いる．遺伝的アルゴリズムにおける適応度を，確信度の推定精度とする．

3.6 確信度の推定

選択した特徴量を用い，SVM によって解答者の設問ごとの確信度を推定する．確信度は「自身の解答に正解である確信を持たた」，「正解である確信を持たなかった」の 2 クラスである．

4. 実験

本節では，提案手法を用いた実験について述べ，この実験によって推定された確信度の精度について検証を行う．

4.1 実験条件・実験手順

実験の様子を図 2 に示す．被験者に図 3 の上のような問題に解答してもらい，その際の視点情報を記録した．被験者は日本人大学生および大学院生 11 名(男性：9 名，女性：2 名)である．本実験ではアイトラッカとして Tobii EyeX を用いた．使用した問題，およびアンケートのフォーマットを図 3 に示す．問題には，TOEIC の Part5 を想定した四択穴埋め問題 80 問を使用した．アンケートの内容は，「その問題に自信をもって解答できたかどうか」であり，このアンケートの結果をもとに，確信度を 2 つのクラスに分類した．

実験手順は，まず被験者に対してキャリブレーションを行った．キャリブレーションとは，アイトラッカが取得する視点情報を被験者ごとに補正することである．その後，被験者は問題に解答し，1 問解くごとにその問題に関するアンケートに回答する．この手順を 80 問解答し終えるまで繰り返した．

次に，得られた特徴量から，確信度の推定に有効なものを選択した．選択には遺伝的アルゴリズムを用いた．詳細には，始めに 200 個の親世代をランダムに生成した．各個体は，30 次元のバイナリデータで表現され，符号の 1 つ 1 つが各特徴量を表現している．これらを交叉させることによって 200 個の次世代個体を生成し，各個体の適応度を求めた．これを，300 世代繰り返した．今回は，交叉させる個体の選択方法としてルーレット方式，交叉方法として一様交叉を採用した．また，変異確率を 3% とし，適応度が

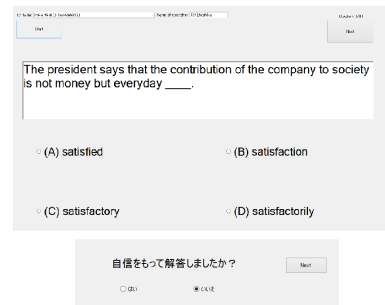


図 3 問題とアンケートのフォーマット

高い個体を次世代に残すエリート保存戦略を採用した．

その後，SVM を用いて確信度を推定し，その精度を求めた．SVM を使用する際，被験者 1 名が解答した問題の内，1 問分のデータをテストデータ，残りの 79 問分のデータを学習データとする Leave-One-Document-Out Cross-Validation により，被験者依存の推定を行った．これを被験者 11 名に対して行った．

4.2 実験結果・考察

各被験者が解答した問題 80 問に対して選択された特徴量を用いて確信度を推定した．また，比較実験として，特徴量に解答時間(表 1 の No.29)のみを用いた場合と，解答の正誤(表 1 の No.30)のみを用いた場合で確信度を推定した．推定により得られた結果を図 4 に示す．選択された特徴量によって被験者 11 名の確信度を推定した精度は，最高値は 95.0%，最低値は 85.0%，平均は 90.1%であった．一方，特徴量に解答時間のみを使用して確信度を推定した場合，推定精度の最高値は 81.3%，最低値は 62.5%，平均は 75.2%であった．また，特徴量に解答の正誤のみを使用して確信度を推定した場合，推定精度の最高値は 81.3%，最低値は 35.0%，平均は 63.5%であった．提案手法と比較手法で符号検定 [7] を行った結果を表 2 に示す．これより，被験者ごとの検証では，全てのケースについて帰無仮説が棄却できたわけではないが，全体を通してみると，比較手法の二者に対して，提案手法の優位性が示された．

30 個の特徴量から特徴選択を行った結果，全ての被験者に採用された，または採用されなかった特徴量は存在しなかった．各特徴量の採用回数を表 3 に示す．採用回数が多かった特徴量として，問題文の fixation 持続時間の合

表 2 符号検定の結果．**は有意水準 1%，*は 5%で帰無仮説 (提案手法と比較手法に差がないとする仮説) が棄却されることを示す．

比較手法	全体	A	B	C	D	E	F	G	H	I	J	K
解答時間のみ	**											
解答の正誤のみ	**	**		**	*		**	*		**	*	**

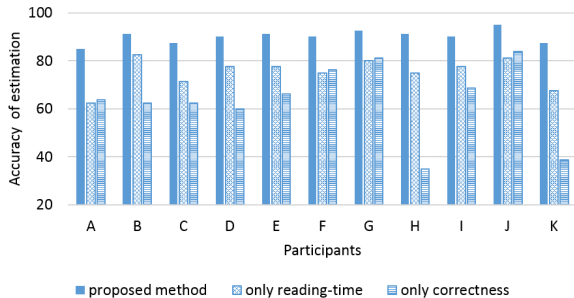


図 4 確信度の推定精度

表 3 特徴量の採用回数

採用回数	特徴量の番号
6 回	9, 12, 15, 16, 21, 23, 24
5 回	3, 6, 7, 8, 13, 17, 18
4 回	4, 14, 25, 27, 28, 29
3 回	1, 2, 5, 11, 20, 26
2 回	10, 19
1 回	22, 30

計, 問題文の fixation 持続時間の最小値, saccade の回数, saccade 距離の合計, saccade 持続時間の最小値, 問題文内の saccade の回数, 選択肢-問題間の saccade の回数の 7 つで, 11 名中 6 名に採用された. これら 7 つの内, saccade の回数と saccade 距離の合計以外の 5 つが頻繁に選択されたことで, 関連研究 [2][4][5] で得られた知見を検証する結果となった. しかし, 多くとも被験者全体の半分でしか採用されなかったことから, 解答時の確信度を表現する特徴量には個人差があるといえる. 最も採用回数が少なかった特徴量は, 選択肢間の saccade の回数と解答の正誤の 2 つで, どちらも 11 名中 1 名のみ採用された. 選択肢間の saccade の回数については, 選択肢を絞り切れないために確信を持っていない場合に有効であると想定していたが, ほとんど採用されることはなかった. 以上のことから, 被験者の性格や問題を解く際の戦略, 癖など, 被験者への依存が大きいと, 有効な特徴量は被験者によって異なると考えられる.

5. まとめ

本論文では, 優先度に応じた復習の重要性に着目し, 優先度を考慮に入れた復習を補助するシステムの実現のため, 学習者の問題解答時の確信度を推定する手法を提案した. 本手法では, 問題解答時の視点情報を取得した. 視点情報には個人差があるため, 被験者ごとに遺伝的アルゴリ

ズムを用いて特徴選択を行った. 確信度の推定には SVM を用いた. 被験者 11 名の平均推定精度は 90.1%であった. 実験によって有効な特徴量として, 問題文の fixation 持続時間の合計, 問題文の fixation 持続時間の最小値, saccade の回数, saccade 距離の合計, saccade 持続時間の最小値, 問題文内の saccade の回数, 選択肢-問題間の saccade の回数の 7 つが得られたが, これらの採用率は約 5 割に留まっており, 問題解答時の振る舞いは, 被験者の性格や問題を解く際の戦略などから影響を受けるため, 有効な特徴量は被験者によって異なることが分かった.

今後の課題として, 新たに視点情報以外を取得するデバイスを選定することが挙げられる.

謝辞

本研究の一部は, JST CREST (JPMJCR16E1), 日本学術振興会科学研究費補助金挑戦的萌芽研究 (15K12172), ならびに大阪府立大学キーププロジェクトの補助による.

参考文献

- [1] Keith Rayner. Eye movements in reading and information processing: 20 years of research. Psychological bulletin, Vol. 124, No. 3, p. 372, 1998.
- [2] Kazuaki Kojima, Keiichi Muramatsu, and Tatsunori Matsui. Experimental study toward estimation of a learner mental state from processes of solving multiple choice problems based on eye movements. In Proceedings of the 20th International Conference on Computers in Education, ICCE 2012, pp. 81-85, 2012.
- [3] 大社綾乃, Kai Kunze, Olivier Augereau. 学習補助のための視点情報に基づく文書アノテーション. 電子情報通信学会技術研究報告 IEICE technical report :信学技報, Vol. 115, No. 23, pp. 161-166, may 2015.
- [4] Meng-Jung Tsai, Huei-Tse Hou, Meng-Lung Lai, Wan-Yi Liu, and Fang-Ying Yang. Visual attention for solving multiple-choice science problem: An eyetracking analysis. Computers & Education, Vol. 58, No. 1, pp. 375-385, 2012.
- [5] Shoya Ishimaru, Syed Saqib Bukhari, Carina Heisel, Jochen Kuhn, and Andreas Dengel. Towards an intelligent textbook: eye gaze based attention extraction on materials for learning and instruction in physics. In Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, pp. 1041-1045, 2016.
- [6] Georg Buscher, Andreas Dengel, and Ludger van Elst. Eye movements as implicit relevance feedback. In CHI '08 extended abstracts on Human factors in computing systems, pp. 2991-2996, 2008.
- [7] Janez Demšar. 2006. Statistical comparisons of classifiers over multiple data sets. Journal Machine learning research 7, Jan (2006), 1-30