

ダイナミックアセスメントにおける段階反応モデルに基づいた適応型テスト

○宮澤 芳光*, 住 政二郎**, 植野 真臣***

* 長岡技術科学大学, ** 関西学院大学, *** 電気通信大学

1 はじめに

近年、ヴィゴツキーの発達の最近接領域に基づくダイナミック・アセスメントが注目を集めている。発達の最近接領域とは、現下の発達水準と支援を受けながら到達可能な発達水準とのへだたりを指している。ダイナミック・アセスメントでは、このへだたりを評価するため、評価者が受検者に対して介入を加えて、その介入に対する反応に応じて受検者を評価する [1]。具体的には、自力では解決できない課題に取り組んでいる受検者に対してヒントを与え、その反応から受検者の能力を測定する。ここでの介入は、特に足場掛けとして知られている。

植野ら [2] は、段階反応モデルを用いて適応的にヒントを提示する足場掛けシステムを開発している。このシステムは、適切な足場掛けをするため、ヒントごとの正答率を予測できるように定式化を行っている。また、課題の正答率が 0.5 に近いヒントを提示することが最も学習効果が高かったことを報告している [3]。しかし、植野らの手法では、項目の出題順序が固定されているため、受検者の能力に応じた項目選択による能力推定過程について分析されていない。

本研究では、ダイナミック・アセスメントにおける受検者への支援を項目のヒントとして、項目に誤答したときに段階的に提示し、ここで得られた解答データから受検者の能力を推定し、情報量が高い項目を出題する手法を提案する。本研究では、従来の適応型テストとの能力精度について分析する。

2 項目反応理論

本研究では、段階的なヒントを組み込んだ適応型テストを実現するため、ヒントへの反応データを項目反応理論の段階反応モデル [4] に当てはめる。まず、項目 i に対して $K-1$ 個のヒント k , ($k = 1, \dots, K-1$) を考える。ヒントがない状態を $k = 0$ とすると、課題に対して学習者の反応が誤りのとき、 $k = 1$ のヒントを提示し、更に誤りであった場合、 $k = 2$ のヒントを提示する。これを繰り返す、学習者が正答するまで ($k = K-1$) ヒントを提示する。このとき、受検者 j ($j = 1, \dots, M$) の項目 i ($i = 1, \dots, N$) に対して各ヒント提示の状態での学習者の反応データを以下のように得ることができる。

$$X = x_{ijk}, (j = 1, \dots, M, i = 1, \dots, N, k = 0, \dots, K) \quad (1)$$

ただし、

$$x_{ijk} = \begin{cases} 1 & \text{ヒント } k \text{ を提示したとき（以前を含む）に受検者 } j \text{ が項目 } i \text{ に正答} \\ 0 & \text{これ以外} \end{cases} \quad (2)$$

x_{ijK} は $K-1$ 個のヒントを全て提示しても正答できなかった場合のデータを示す。このようなデータ X が得られたとき、段階反応モデルに適用することを考える。受検者 j ($j = 1, \dots, M$) の項目 i ($i = 1, \dots, N$) に対して k ($k = 1, \dots, K-1$) 番目までのヒントを提示した後に初めて正答する確率は以下のモデルで示せる。

$$p_{ijk} = p_{ij(k-1)}^* - p_{ijk}^* \quad (3)$$

$$\begin{cases} p_{ijk}^* &= \frac{1}{(1+\exp(-1.7a_i(\theta_j - b_{ik})))}, k = 0, \dots, K-1 \\ p_{ij(-1)}^* &= 1 \\ p_{ijK}^* &= 0 \end{cases} \quad (4)$$

式 (4) では、 a_i が項目 i の識別力パラメータであり、 b_{ik} が項目 i に対してカテゴリ k の難易度パラメータであり、 θ_j が受検者 j の能力パラメータである。ただし、 $b_{i0} > b_{i1} > \dots > b_{ik} > \dots > b_{i(K-1)}$ と制約される。

本手法では、受検者の能力を推定するため、テスト理論の適応型テストのアルゴリズムを応用する。適応型テストは、受検者の解答履歴から逐次能力を推定し、情報量が最大の項目を受検者に出題する Computer Based Testing である。具体的には、フィッシャー情報量が最大の項目を出題する。段階反応モデルのフィッシャー情報量は以下の通りである [5]。

$$I_i(\theta_j) = 1.7^2 a_i^2 \sum_{k=0}^K \frac{(p_{ij(k-1)}^*(\theta_j) q_{ij(k-1)}^*(\theta_j) - p_{ijk}^*(\theta_j) q_{ijk}^*(\theta_j))^2}{p_{ijk}(\theta_j)} \quad (5)$$

q_{ijk}^* は $1 - p_{ijk}^*$ である。本手法は段階反応モデルを用いた適応型テストを用いることで効率的に受検者の能力を推定する。

本手法では、テスト理論の適応型テストのアルゴリズムを応用し、段階反応モデルの適応型テストに用いた。アルゴリズムは以下の通りである。

- (1) 能力推定値の初期値は 0 に設定
- (2) アイテムバンクから情報量が最大の項目を選択
- (3) 受検者が選択された項目に解答
- (4) システムが受検者の解答の正誤を判断
- (5) 誤答の場合、最初のヒントを提示し (3) へ（受検者が正答するまでヒントを提示し続ける。項目に正答するか全てのヒントを提示して誤答したとき、この反応データを用いて受検者の能力を推定）
- (6) 受検者の推定能力値と一つ前の推定値の差が小さければテストを終了。それ以外は (2) へ上記のアルゴリズムによって受検者の能力を推定する。

3 ヒントの設計

アイテムバンクには、100 問からなる 4 択の英語の語彙・文法問題を準備した。これは研究参加校の大学が英語のプレイスメント・テストで毎年使っているものである。この 100 問に対応したヒントを設計するために、プレイスメント・テストの結果から習熟度の異なる 14 名を対象に定性的調査を行った。参加者には、プレイスメント・テストから難易度順に抜粋した 25 問を受験してもらい、解答と解答中の発話データを収集した。分析の結果、解答パターンと、文法と語彙の習熟度の低さから誤答となるケースが多くあることが分かった。この結果を踏まえ、語彙・文法に関連した 2 段階のヒントをすべての設問に対して設計した。全てのヒントが提示されたにも関わらず誤答したとき、解答と日本語訳が提示される。図 1 に、ヒントが提示されたシステムの画面例を示す。項目への誤答時、ヒントは選択肢の下部に表示される。

項目のパラメータを推定するため、37 人の大学生の反応データからベイズ推定を用いてパラメータを推定した。本手法では、このパラメータを用いて受検者の能力を逐次推定しながら、フィッシャー情報量が最大の項目を出題する。

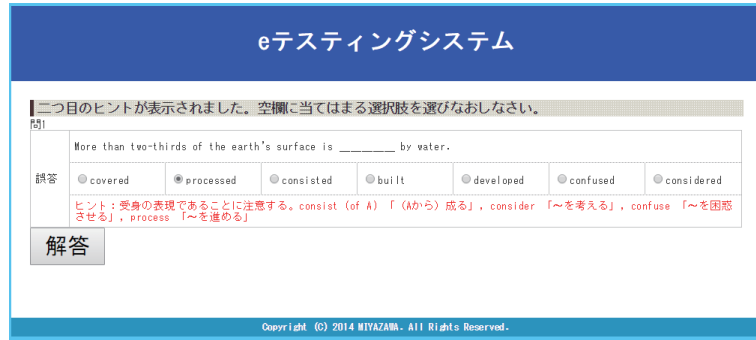


図 1: システムの画面例

表 1: シミュレーションでの能力推定過程

項目数	1	2	3	4	5	6	7	8	9
提案手法	0.664 (0.434)	0.829 (0.343)	0.905 (0.290)	0.934 (0.254)	0.950 (0.234)	0.960 (0.222)	0.963 (0.201)	0.960 (0.194)	0.965 (0.191)
適応的 ヒント	0.492 (0.337)	0.718 (0.392)	0.804 (0.341)	0.836 (0.312)	0.872 (0.285)	0.894 (0.263)	0.913 (0.243)	0.919 (0.231)	0.928 (0.215)
CAT	0.431 (0.389)	0.680 (0.437)	0.776 (0.387)	0.805 (0.357)	0.825 (0.318)	0.844 (0.300)	0.858 (0.276)	0.869 (0.260)	0.881 (0.247)
項目数	10	11	12	13	14	15			
提案手法	0.965 (0.186)	0.966 (0.181)	0.968 (0.181)	0.971 (0.178)	0.971 (0.175)	0.974 (0.170)			
適応的 ヒント	0.941 (0.204)	0.954 (0.194)	0.961 (0.182)	0.968 (0.168)	0.976 (0.158)	0.984 (0.148)			
CAT	0.889 (0.230)	0.895 (0.217)	0.898 (0.206)	0.899 (0.199)	0.905 (0.192)	0.910 (0.186)			

4 シミュレーション実験の結果と考察

開発したシステム, 100 問のアイテムバンク, 設計したヒントを活用してシミュレーションによって能力推定過程を分析する. 真の能力値 $\theta_j = 1.0$ の 1000 名の受検者の反応データを生成し, 出題項目の数は 15 項目に固定した. 表 1 は, 提案手法と適応的なヒント提示の手法, 従来の適応型テスト (CAT) の能力推定過程を示している. 表 1 の各値は, 出題項目数ごとに受検者 1000 名の推定能力値を平均した値である. 表 1 から提案手法の能力推定値が早い段階で収束していることが分かる. これは, 正誤の 2 値を扱う情報量関数が特定の能力値にのみ高いのに対し, 段階反応モデルの情報量関数が能力値に広く高いことが要因だと考えられる. また, 情報量が高い項目を選択しているため, ヒントのみ適応的な手法よりも推定精度が高い. すなわち, 従来の適応型テストに比べ, 本手法は, 少ない項目数で能力推定ができることが分かった.

5 被検者実験

ここでは, ダイナミック・アセスメントのための適応型テストの動作を検証するため, 被験者実験を実施した. アイテムバンクの項目数は 100 であり, 被験者数は 39 名である. 出題項目の数は 25 項目に固定した. これは, 本システムを毎回の授業の冒頭で使ってもらうことを念頭においてのことである.

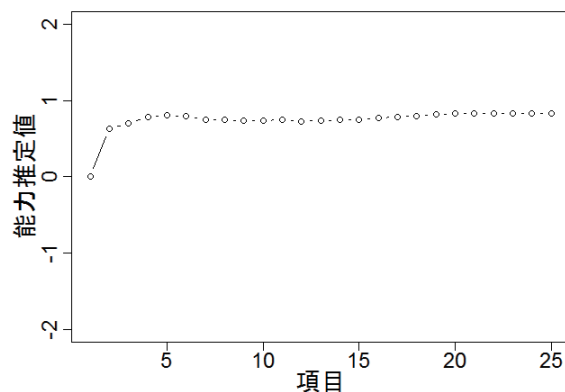


図 2: 被検者実験での能力推定過程

実験の結果，学生の能力に応じて適切な設問とヒントが提示されていることが分かった。特に習熟度の低い学生にとってヒントは有用であり，ヒントを活用して正答を選択できるようになっていた。また，ヒントを活用する過程で，誤りに気づき，学習を深めている様子も観察された。

図 2 は，出題項目数ごとに被検者全員の推定能力値を平均した値であり，推定過程を示す。横軸が項目であり，縦軸が能力推定値である。本実験では，能力推定過程を分析するため，項目数を 25 項目と固定した。この図から受検者が項目を解答するごとに能力推定値が収束していることがわかる。また，シミュレーション結果と同様に初期の項目から収束した能力推定値に近い推定値であることがわかる。

6 おわりに

本研究では，ダイナミック・アセスメントのための適応型テストを提案した。ここでは，ダイナミック・アセスメントの支援をヒントとして段階反応モデルを用いて適応型テストに組み込んだ。本研究では，実際に段階反応モデルのパラメータを推定し，このパラメータを用いて適応型テストの能力推定過程を確認した。今後，提案手法と従来の適応型テストとの能力推定過程をさらに詳細に比較し，本手法の有効性を検証していく。

参考文献

- [1] H C Haywood and C S Lidz. *Dynamic Assessment in Practice: Clinical and Educational Applications*. Cambridge University Press, 2006.
- [2] 植野真臣, 松尾淳哉. 項目反応理論を用いて適応的ヒントを提示する足場かけシステム. 電子情報通信学会論文誌, Vol. J98-D, No. 1, pp. 17–29, 2015.
- [3] Maomi Ueno and Yoshimitsu Miyasawa. Probability based scaffolding system with fading. In Cristina Conati, Neil Heffernan, Antonija Mitrovic, and M. Felisa Verdejo, editors, *Artificial Intelligence in Education*, Vol. 9112 of *Lecture Notes in Computer Science*, pp. 492–503. Springer International Publishing, 2015.
- [4] Fumiko Samejima. *Estimation of latent ability using a response pattern of graded scores*. Psychometrika monograph supplement, 2006.
- [5] 豊田秀樹. 項目反応理論 [入門編]. 朝倉書店, 2002.

E-mail: miyasawa@vos.nagaokaut.ac.jp
 会員番号: 30000171