項目反応理論を用いた難易度調整可能な多肢選択式読解問題自動生成

富川雄斗 [1]　宇都雅輝 [1]
[1] 電気通信大学大学院
{tomikawa,uto}@ai.lab.uec.ac.jp

概要

近年，教育場面において難易度調整可能な読解問題自動生成が注目されている．我々は，項目反応理論を用いて学習者の能力にあった難易度の読解問題を生成する技術を開発してきた．しかし，この手法は答えが読解対象文中に存在する抽出型の問題形式を対象としており，教育現場で広く使われている多肢選択式の問題形式には対応できない．そこで，本研究では難易度調整可能な多肢選択式問題自動生成手法を開発する．また，提案手法によって生成した問題を項目反応理論に基づいて分析し，生成の性能を評価する．

1　はじめに

読解文から問題を自動で生成する読解問題自動生成が，教育分野において注目されている．近年では，深層学習を用いることによって，高品質な読解問題自動生成が実現されている [1, 2, 3, 4, 5, 6]．

学習支援や教育評価の文脈で問題生成を利用する場合，任意の難易度で問題を生成できることが望ましい [7]．そのため，近年では，難易度の調整機能を有する問題生成技術が幾つか提案されている [8, 9, 10]．最先端手法の一つとして，Uto *et al.* [8] は項目反応理論 (Item Response Theory: IRT) [11] と事前学習済み深層学習モデルを用いて，学習者の読解力に適した難易度で問題と答えの組を生成できる手法を開発している．しかし，この手法は答えが読解対象文中に存在する抽出型の問題形式を対象としており，教育現場で広く使われている多肢選択式の問題には直接には適用できない．そこで，本研究では難易度調整可能な多肢選択式問題自動生成手法を開発する．提案手法は，BERT [12] による答え区間の抽出と T5 [13] による Answer-aware な問題生成の 2 段階で実現されていた従来の難易度調整可能な抽出型問題生成手法を，LLaMA 2 [14] による自己回帰型言語モデルを用いた手法に拡張し，読解文と任意の難易度から多肢選択式問題の問題文と選択肢群を生成できるようにする．また，提案手法によって生成した問題を項目反応理論に基づいて分析し，生成された問題の特性を評価する．本分析では，難易度調整の性能を評価するためにラッシュモデル [15] を，生成された選択肢の性質を調査するために名義反応モデル [16] を利用する．

2　項目反応理論

項目反応理論は学習者の能力と問題の難易度などの特性を定式化するテスト理論の一つであり，様々なハイステークス試験で活用されている [17, 18, 19]．ここでは，本研究で利用するラッシュモデルと名義反応モデルを紹介する．ラッシュモデルは，学習者 j が問題 i に正答する確率 P_{ij} を次式で定義する．

$$P_{ij} = \frac{1}{1 + \exp(-(\theta_j - b_i))} \tag{1}$$

ここで，θ_j は受験者 j の能力値を，b_i は問題 i の難易度を表し，正誤反応データの集合から推定される．

名義反応モデルは多値型項目反応理論モデルの一つであり，多肢選択式問題の分析では選択肢の傾向を分析する目的で利用されることが多い．名義反応モデルは，学習者 j が問題 i において K 個の選択肢から k を選択する確率 P_{ijk} を次式で定義する．

$$P_{ijk} = \frac{\exp(\alpha_{ik}\theta_j + \zeta_{ik})}{\sum_{k'}^{K} \exp(\alpha_{ik'}\theta_j + \zeta_{ik'})} \tag{2}$$

ここで，α_{ik} は問題 i における選択肢 k の識別力パラメータ，ζ_{ik} は問題 i における選択肢 k の位置パラメータを表す．名義反応モデルのパラメータは，各学習者が各問題においてどの選択肢に反応したかを表す選択反応データの集合から推定される．

3 IRT を用いた難易度調整可能な抽出型読解問題自動生成

本研究は Uto *et al.* が提案した，項目反応理論を用いた難易度調整可能な抽出型読解問題自動生成のアプローチを基礎とする．そこで本章では，この手法について説明する．

この研究では，質問応答や問題生成のベンチマークデータセットとして広く利用される SQuAD データセット [20] を用いている．SQuAD は，読解対象文 $\boldsymbol{c}_i$ とそれに関連する問題文 $\boldsymbol{q}_i$，およびその問題文に対応する答え $\boldsymbol{a}_i$ で構成され，答え $\boldsymbol{a}_i$ は読解対象文 $\boldsymbol{c}_i$ 中に存在する ($\boldsymbol{a}_i \in \boldsymbol{c}_i$)．したがって，データセットは $\{\boldsymbol{c}_i, \boldsymbol{q}_i, \boldsymbol{a}_i | i \in 1 \ldots I\}$ と表すことができる．ここで I はデータ数を表す．

しかし，SQuAD には問題の難易度が含まれていない．そこで Uto *et al.* は SQuAD のデータ中の問題に対して，IRT に基づく難易度を推定する方法を提案している．具体的には，データセット中の各問題 $\boldsymbol{q}_i$ を性能の異なる多数の QA (Question Answering) システムに出題し，正誤反応データを取得する．なお，本来は，人間の正誤反応データを取得すべきであるが，これには膨大なコストがかかるため，ここでは QA システムで代用している．そして，このデータから，ラッシュモデルによって各問題の難易度値 b_i を推定し，元々のデータセットに b_i を加えることで，難易度を含んだデータセットを構築する．このように構築した難易度付きの SQuAD データセットを用いて，Uto *et al.* では次の二つのモデルを訓練することで，答えと問題の生成を行う．

1. **難易度調整可能な答え抽出モデル:** 読解対象文から所望の難易度に沿った答えを抽出するモデル．具体的には，難易度 b_i と読解対象文 $\boldsymbol{c}_i$ を結合した文字列を入力し，読解対象文における答えの開始位置と終了位置を出力するように訓練された BERT モデルとして設計．
2. **難易度調整可能な問題生成モデル:** 読解対象文と答え，および所望の難易度から，問題を生成するモデル．具体的には，難易度 b_i と読解対象文 $\boldsymbol{c}_i$，答え $\boldsymbol{a}_i$ を結合した文字列を入力し，問題文を生成するように訓練された T5 モデルとして設計．

この手法は答えが読解対象文中に存在する抽出型の問題形式を対象としており，教育現場で広く使われている多肢選択式の問題には対応していない．

表 1: 問題生成時の入出力

入力
Create a question and 4 options with a difficulty level of {b} based on the Context.
Option 1 is the correct answer and Option 2, 3 and 4 are the distractor options.
Difficulty level -3.0 is the easiest and 3.0 is the most difficult.
Context: <tag> {context} </tag>
出力
Option 1 (Correct Option): <tag> {correct_option} </tag>
Question: <tag> {question} </tag>
Option 2 (Distractor Option): <tag> {distractor_options_1} </tag>
Option 3 (Distractor Option): <tag> {distractor_options_2} </tag>
Option 4 (Distractor Option): <tag> {distractor_options_3} </tag>

4 提案手法

本研究では，難易度調整可能な多肢選択式問題自動生成手法を提案する．多肢選択式問題自動生成の研究では，モデルの学習や性能評価に RACE データセット [21] が広く用いられている [22, 23, 24, 25, 26]．RACE は，読解対象文 $\boldsymbol{c}_i$ とそれに関連する問題文 $\boldsymbol{q}_i$，およびその問題文に対応する正答選択肢 $\boldsymbol{a}_i$ と 3 つの誤答選択肢 $\boldsymbol{d}_{i1}, \boldsymbol{d}_{i2}, \boldsymbol{d}_{i3}$ で構成され，$\{\boldsymbol{c}_i, \boldsymbol{q}_i, \boldsymbol{a}_i, \boldsymbol{d}_{i1}, \boldsymbol{d}_{i2}, \boldsymbol{d}_{i3} | i \in 1 \ldots I\}$ として表すことができる．しかし，SQuAD データセットと同様に，RACE データセットにおいても問題の難易度は含まれていない．そこで，Uto *et al.* と同様の手順で難易度を含んだデータセットを構築する．次に，難易度と読解対象文を所与として，難易度に沿った正答選択肢と問題，3 つの誤答選択肢の生成を LLaMA 2 によって実現する．具体的な提案手法の構築手順は次の通りである．

1. 精度の異なる 400 個の QA システムを構築する．具体的には，まず RACE の検証データを用いて，huggingface[1)] で公開されている 4 つの深層学習モデル (bert-base-uncased, roberta-base, deberta-v3-large, albert-base-v1) を，入力に対応する選択肢を選ぶモデルとして Liu *et al.* [27] の方法で訓練する．次に，各モデルの **[CLS]** に対応する出力層の後に，Dropout 層を挿入する．この Dropout 率を 0.00 から 0.99 まで 0.01 刻みで変化させることで，精度の異なる 400 個の QA システムを構築する．
2. Uto *et al.* と同様の方法で RACE の訓練データに対して問題難易度 b_i を推定し，その難易度値を RACE データセットに追加することで難易度付きの RACE データセットを構築する．

1) https://huggingface.co/

3. Llama-2-7b[2)]を用いて，読解対象文と難易度を入力すると，正答選択肢，問題，3 つの誤答選択肢を出力するように訓練を行う．ここで，モデルの入力と出力の形式は表 1 の通りである．表 1 中の{b}は手順 2 で推定した問題の難易度，{context}は読解対象文，{correct_option}は正答選択肢，{question}は問題，{distractor_options_1}, {distractor_options_2}, {distractor_options_3}は 3 つの誤答選択肢を表し，それぞれデータセット中の各データで置き換えるものとする．

5 評価実験

ここでは提案手法の性能を評価する実験を行った．

5.1 実験手順

上記の手法で訓練した提案モデルを用いて以下の実験を行った．まず，RACE テストデータ中の 300 個の読解対象文に対し，それぞれ-3.0 から 3.0 まで 0.1 刻みで難易度を指定して，提案モデルによって計 18300 問の問題を生成した．実際に生成した問題例を付録 A に示す．次に，生成した問題を，4 章の手順 1 で作成した 400 個の QA システムに解答させ，選択反応データを収集した．以上の手順で収集した選択反応データを用いて，正答率，ラッシュモデル，名義反応モデルによる評価を行った．なお，以降の分析のために，提案手法の構築手順 2. で問題難易度を推定する際に，各 QA システムの能力推定値も合わせて求めた．

5.2 正答率による評価

生成した問題が指定した難易度を反映できているか確認するために，5.1 節で収集した選択反応データを正答，誤答に 2 値化し，正答率による評価を行った．各難易度の問題群に対する 400 個の QA システムの平均正答率を図 1 に示す．横軸は指定した難易度，縦軸は各難易度に対応する問題群の平均正答率を表す．この結果より，指定した難易度が高くなるほど正答率が下がることが分かる．また，指定した難易度と平均正答率の相関係数が-0.97 であったことから，正しく難易度情報を反映できていることが分かる．

2） https://huggingface.co/meta-llama/Llama-2-7b

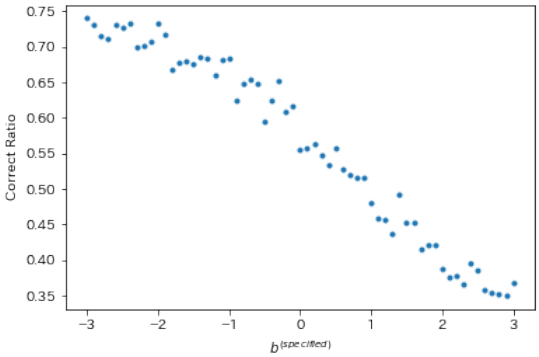

図 1: 各難易度の問題群に対する平均正答率

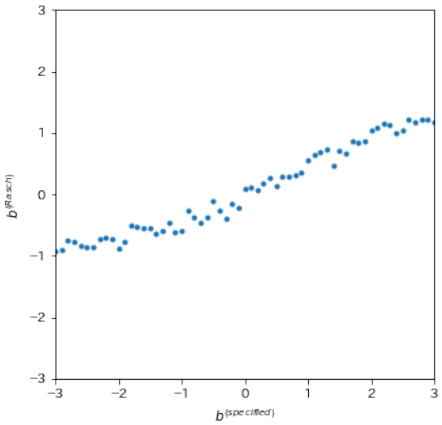

図 2: 指定した難易度と難易度推定値

5.3 ラッシュモデルによる評価

指定した IRT 尺度に応じた難易度の問題を生成できているか確認するために，ラッシュモデルによる評価を行った．具体的には，5.2 節と同様に選択反応データを正答，誤答に 2 値化し，事前に求めておいた各 QA システムの能力値を所与としてラッシュモデルを適用することで，各問題の難易度を推定した．結果を図 2 に示す．横軸は指定した難易度，縦軸は各難易度に対応する問題群に対する難易度推定値の平均値を表す．指定した難易度と難易度推定値の平均値の相関係数は 0.98 と高かったが，平均絶対誤差は 0.90 であり，十分に小さいとは言えない．この結果は，問題の識別力が関係している可能性があるが，これについての詳細な分析は今後の課題としたい．

5.4 名義反応モデルによる評価

次に，生成する際に指定した難易度によって各選択肢を選ぶ確率が異なるかを確認するために，選択

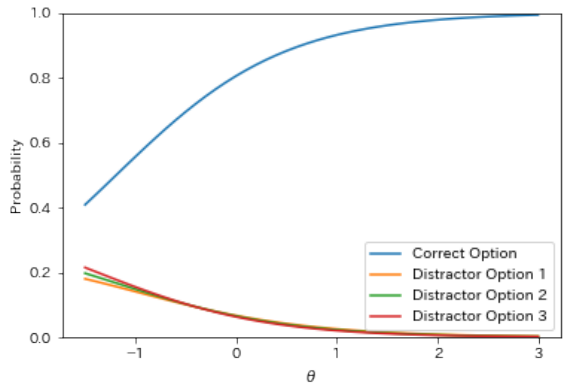

図 3: 難易度を $b^{(\text{specified})} = -3.0$ と指定して生成した問題に対する名義反応モデルの項目特性曲線

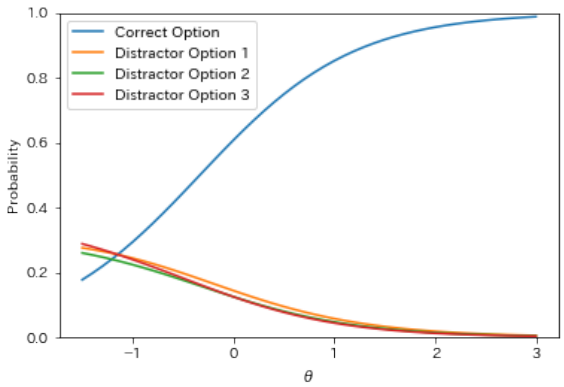

図 4: 難易度を $b^{(\text{specified})} = 0.0$ と指定して生成した問題に対する名義反応モデルの項目特性曲線

反応データを名義反応モデルによって分析した．具体的には，収集した選択反応データから，ラッシュモデルによって推定された各 QA システムの能力値を所与として名義反応モデルを適用することで，各問題の各選択肢に対応するパラメータ α_{ik}，ζ_{ik} を推定した．なお，ラッシュモデルにおける能力値と名義反応モデルにおける能力値の尺度は一致しないが，ここでは簡単のため使用している．次に，難易度を $b^{(\text{specified})} = -3.0$ と指定して生成した選択肢に対応するパラメータ α_{ik}，ζ_{ik} の平均を，4 つの選択肢ごとに求めた．このパラメータを用いて作成した項目特性曲線を図 3 に示す．横軸は学習者の能力値，縦軸は選択肢 k を選ぶ確率を表す．なお，推定した QA システムの能力値の最小値は-1.2，最大値は 2.6 であったため横軸の範囲はこれを含む範囲として-1.5 から 3.0 としている．$b^{(\text{specified})} = 0.0$，3.0 として生成した選択肢に対応するパラメータに対しても同様に作成した項目特性曲線をそれぞれ図 4，5 に示す．能力値が高くなるほど，正答選択肢の選択確率が高くなり，誤答選択肢の選択確率が低下する傾向が読み取れる．また，難易度が増加するほど，正答選択肢の選択確率が総じて低くなり，誤答選択肢の選択確率が総じて向上する傾向が読み取れる．このことから，正答選択肢や誤答選択肢は適切に生成されていることがわかる．

以上の結果より，指定した難易度が各選択肢に反映されていることを確認できた．また，個別の問題ごとに項目特性曲線を分析することで，個々の選択肢に関する詳細な情報が得られる．この分析例を付録 A に示す．

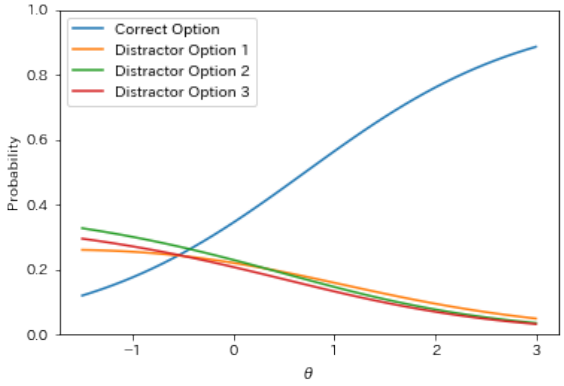

図 5: 難易度を $b^{(\text{specified})} = 3.0$ と指定して生成した問題に対する名義反応モデルの項目特性曲線

6 おわりに

本研究では，項目反応理論と LLaMA 2 を用いた難易度調整可能な多肢選択式読解問題自動生成手法を開発した．また，提案手法によって生成した問題を QA システムに解かせて収集した選択反応データを用いて，正答率，ラッシュモデル，名義反応モデルで分析した結果，指定した難易度を反映した問題生成ができていることが確認できた．

5.3 節の結果を踏まえ，今後は識別力を考慮できる 2 母数ロジスティックモデルや当て推量を考慮できる 3 母数ロジスティックモデルの活用も検討していきたい．また，提案手法ではしばしば解答不能な問題が生成される場合があるため，今後は，解答可能性を向上する機構について検討を行っていきたい．

謝辞

本研究は JSPS 科研費 23K17585, 21H00898, 19H05663 の助成を受けたものです.

参考文献

[1] Xinya Du, Junru Shao, and Claire Cardie. Learning to ask: Neural question generation for reading comprehension. In **Proc. 55th Annu. Meeting of the Association for Computational Linguistics**, pp. 1342–1352, 2017.

[2] Yao Zhao, Xiaochuan Ni, Yuanyuan Ding, and Qifa Ke. Paragraph-level neural question generation with maxout pointer and gated self-attention networks. In **Proc. 2018 Conf. Empirical Methods in Natural Language Processing**, pp. 3901–3910, 2018.

[3] Ying-Hong Chan and Yao-Chung Fan. A recurrent BERT-based model for question generation. In **Proc. 2nd Workshop on Machine Reading for Question Answering**, pp. 154–162, 2019.

[4] Jingjing Li, Yifan Gao, Lidong Bing, Irwin King, and Michael R. Lyu. Improving question generation with to the point context. In **Proc. 2019 Conf. Empirical Methods in Natural Language Processing and the 9th Int. Joint Conf. Natural Language Processing**, pp. 3216–3226, 2019.

[5] Asahi Ushio, Fernando Alva-Manchego, and Jose Camacho-Collados. Generative language models for paragraph-level question generation. In **Proc. 2022 Conf. Empirical Methods in Natural Language Processing**, pp. 670–688, 2022.

[6] Ying-Hong Chan, Ho-Lam Chung, and Yao-Chung Fan. Keyword provision question generation for facilitating educational reading comprehension preparation. In **Proc. 15th Int. Conf. on Natural Language Generation**, pp. 196–202, 2022.

[7] Maomi Ueno and Yoshimitsu Miyazawa. IRT-based adaptive hints to scaffold learning in programming. **IEEE Trans. on Learn. Technologies**, Vol. 11, No. 4, pp. 415–428, Oct.–Dec. 2018.

[8] Masaki Uto, Yuto Tomikawa, and Ayaka Suzuki. Difficulty-controllable neural question generation for reading comprehension using item response theory. In **Proc. 18th Workshop on Innovative Use of NLP for Building Educational Applications**, pp. 119–129, 2023.

[9] Yifan Gao, Lidong Bing, Wang Chen, Michael Lyu, and Irwin King. Difficulty controllable generation of reading comprehension questions. In **Proc. Twenty-Eighth Int. Joint Conf. Artificial Intelligence**, pp. 4968–4974, 2019.

[10] Yi Cheng, Siyao Li, Bang Liu, Ruihui Zhao, Sujian Li, Chenghua Lin, and Yefeng Zheng. Guiding the growth: Difficulty-controllable question generation through step-by-step rewriting. In **Proc. 59th Annu. Meeting of the Association for Computational Linguistics and the 11th Int. Joint Conf. on Natural Language Processing**, pp. 5968–5978, 2021.

[11] F. M. Lord. **Applications of Item Response Theory To Practical Testing Problems**. Routledge, 1980.

[12] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. In **Proc. 2019 Conf. of the North American Chapter of the Association for Computational Linguistics**, pp. 4171–4186, 2019.

[13] Colin Raffel and *et al.* Exploring the limits of transfer learning with a unified text-to-text transformer. **J. of Mach. Learn. Res.**, Vol. 21, No. 140, pp. 1–67, Jan. 2020.

[14] Hugo Touvron and *et al.* Llama 2: Open foundation and fine-tuned chat models. **arXiv**, 2023.

[15] Georg Rasch. **Probabilistic models for some intelligence and attainment tests**. The University of Chicago Press, 1981.

[16] R. Darrell Bock. Estimating item parameters and latent ability when responses are scored in two or more nominal categories. **Psychometrika**, Vol. 37, pp. 29–51, Mar. 1972.

[17] Masaki Uto and Maomi Ueno. Empirical comparison of item response theory models with rater's parameters. **Heliyon**, Vol. 4, No. 5, May 2018.

[18] Masaki Uto. A Bayesian many-facet Rasch model with Markov modeling for rater severity drift. **Behavior Res. Methods**, Vol. 55, pp. 3910–3928, Oct. 2022.

[19] Masaki Uto. A multidimensional generalized many-facet Rasch model for rubric-based performance assessment. **Behaviormetrika**, Vol. 48, pp. 425–457, 2021.

[20] Pranav Rajpurkar, Jian Zhang, Konstantin Lopyrev, and Percy Liang. SQuAD: 100,000+ questions for machine comprehension of text. In **Proc. 2016 Conf. Empirical Methods in Natural Language Processing**, pp. 2383–2392, 2016.

[21] Guokun Lai, Qizhe Xie, Hanxiao Liu, Yiming Yang, and Eduard Hovy. Race: Large-scale reading comprehension dataset from examinations. In **Proc. 2017 Conf. Empirical Methods in Natural Language Processing**, pp. 785–794, 2017.

[22] Xiaorui Zhou, Senlin Luo, and Yunfang Wu. Co-attention hierarchical network: Generating coherent long distractors for reading comprehension. In **Proc. AAAI conf. artificial intelligence**, pp. 9725–9732, 2019.

[23] Jeroen Offerijns, Suzan Verberne, and Tessa Verhoef. Better distractions: Transformer-based distractor generation and multiple choice question filtering. **arxiv**, 2020.

[24] Ho-Lam Chung, Ying-Hong Chan, and Yao-Chung Fan. A BERT-based distractor generation scheme with multi-tasking and negative answer training strategies. In **Proc. Findings of the Association for Computational Linguistics 2020**, pp. 4390–4400, 2020.

[25] Yifan Gao, Lidong Bing, Piji Li, Irwin King, and Michael R. Lyu. Generating distractors for reading comprehension questions from real examinations. In **Proc. AAAI conf. artificial intelligence**, pp. 6423–6430, 2019.

[26] Xin Jia, Wenjie Zhou, Xu Sun, and Yunfang Wu. Eqg-race: Examination-type question generation. In **Proc. AAAI conf. artificial intelligence**, pp. 13143–13151, 2021.

[27] Yinhan Liu and *et al.* RoBERTa: A robustly optimized BERT pretraining approach. **arxiv**, 2019.

A　実際に生成された問題例とその分析

難易度を-3.0，0.0，3.0 と指定して，実際に生成された問題例をそれぞれ表 2 に示す．また，これらの問題をそれぞれ 400 個の QA システムに解かせ，収集した選択反応データから推定した名義反応モデルの項目特性曲線を，それぞれ図 6，7，8 に示す．これらの図より，指定した難易度が高くなるほど正答選択肢を選ぶ確率が低くなることが確認できる．また，図 6 を見ると，誤答選択肢 3 は能力値全域において選択確率が低いため，この選択肢は迷わしとしてあまり機能していないことが分かる．次に，図 7 では，能力値の低い学習者は誤答選択肢 1 を選択する確率が高いため，誤答選択肢 1 は能力値が低い学習者に対しての迷わしとして，適切に機能していると言える．図 8 では，誤答選択肢の選択傾向が特徴的である．まず，低能力帯においては誤答選択肢 1 と 2 が同程度に高い選択確率を有しており，正答選択肢の選択確率は著しく低い．このことは，能力が低い学習者は誤答選択肢 1 と 2 に過度に反応する（惑わされる）傾向があることを意味している．また，中程度の能力帯においては，誤答選択肢 1 の選択確率が上昇し，誤答選択肢 2 の選択確率が減少していることがわかる．このことは，この能力帯の学習者にとっては誤答選択肢 2 は迷わしとして機能しなくなり，誤答選択肢 1 に対する確度が上がることを意味している．そして，高能力帯の学習者にとっては，これらの迷わしは機能せず正答選択肢を高確率で選べるようになっていることが読み取れる．

表 2: 実際に生成された問題例

読解対象文

It is true that the first Olympic Games of the modern times in 1896 were not open to women. Baron Pierre de Coubertin, who gave new life to the Games, was very much a man of his time. He didn't like women showing up for the Games. But since this time, things have changed. History shows that the number of women participating in the Olympic Games has been increasing over the years, since their first participation in the 1900 Olympic Games and especially in the last 30 years. This development is of social, political and cultural significance. The progress is the result of the cooperation of the entire Olympic Movement and of proper measures that the International Olympic Committee(IOC), International Sports Federations(IFs), and National Olympic Committees(NOCs)have taken. Especially in the last 20 years, the IOC has enlarged the women's programme at the Olympic Games, in cooperation with IFs and the Organising Committees for the Olympic Games(OCOGs). This development was further strengthened by the IOC's decision that all sports must have women's events. The IOC also started to work on increasing the number of women at leadership level in sport in 1981. Former President Juan Antonio Samaranch, who wanted to have women as IOC members made a great contribution. Many sports volunteers have worked to increase the number of women joining in the Olympic Games. This progress has been made also thanks to the strong determination of women from different countries of the world who have worked to make sure women are fully represented in all sports.

指定した難易度: -3.0

問題文: Who was strongly against women's participation in the Olympic Games?

正答選択肢: Baron Pierre de Coubertin, 誤答選択肢 1: The IOC, 誤答選択肢 2: The OCOGs, 誤答選択肢 3: National Olympic Committees

指定した難易度: 0.0

問題文: Which of the following is true?

正答選択肢: The number of women participating in the Olympic Games has been increasing over the years.

誤答選択肢 1: The first Olympic Games of the modern times in 1896 were open to women.

誤答選択肢 2: The number of women participating in the Olympic Games has been decreasing over the years.

誤答選択肢 3: The number of women participating in the Olympic Games has been stable over the years.

指定した難易度: 3.0

問題文: In which Olympic Games did women take part for the first time?

正答選択肢: 1900, 誤答選択肢 1: 1896, 誤答選択肢 2: 1980, 誤答選択肢 3: 2000

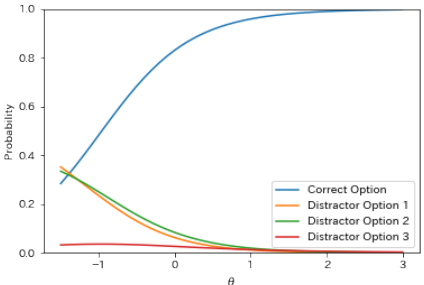

図 6: 指定した難易度が-3.0 の問題による名義反応曲線

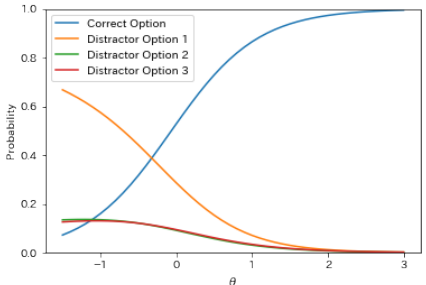

図 7: 指定した難易度が 0.0 の問題による名義反応曲線

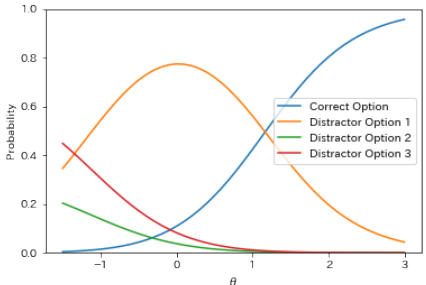

図 8: 指定した難易度が 3.0 の問題による名義反応曲線