

難易度調整可能な多枝選択式読解問題自動生成手法と Direct Preference Optimization による難易度調整精度の向上

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、深層学習を用いて任意の難易度で読解問題を生成できる技術が提案されている。しかし、従来手法は、読解対象文中から答えを抜き出して回答する形式の問題のみを対象としており、教育現場で広く使われている多枝選択式問題を生成できない。加えて、従来手法では訓練データ中の問題の単語列に対する尤度を最大化するように深層学習モデルを訓練しており、難易度調整精度を直接最適化していないため、難易度調整精度に改善の余地が残る。そこで本研究では、自己回帰型の深層学習モデルを用いた難易度調整可能な多枝選択式問題生成手法を開発するとともに、Direct Preference Optimization により難易度調整精度を最大化する訓練手法を提案する。

1 はじめに

読解対象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 [1, 2, 3, 4, 5]。学習支援や教育評価の文脈で問題生成を利用することを想定すると、各学習者の能力に合った適切な難易度の問題を出題することが望ましいと考えられ、そのためには難易度を調整可能な問題生成手法が必要となる [6, 7]。このような背景から、近年では難易度調整機能を有する問題生成技術がいくつか提案されている [8, 9, 10, 11, 12, 13]。

しかし、従来手法には以下の課題が残る。

- 図 1 左で示されるような、読解対象文中から答え単語列を抜き出して回答する形式の問題のみを対象としており、教育現場で広く使われている多枝選択式問題を直接には生成できない。
- 訓練データ中の問題の単語列に対する尤度を最大化するように問題生成モデルを訓練しており、難易度調整の精度を目的関数として最適化していないため、難易度調整精度に改善の余地

が残る。

これらの課題を解決するために、本研究では自己回帰型の深層学習モデルを用いた難易度調整可能な多枝選択式問題生成手法を開発するとともに、DPO (Direct Preference Optimization) [14] に基づいて難易度調整精度を最大化する手法を提案する。具体的にはまず、自己回帰型生成モデルである Llama 2 [15] を用いて、図 1 右のように、任意の読解対象文と難易度から多枝選択式問題の問題文と選択枝を生成できる手法を提案する。加えて、LLM の訓練方法として近年広く利用されている DPO に基づいて、難易度調整の精度を目的関数として最適化するようにモデルを訓練する手法を提案する。

2 項目反応理論を用いた難易度調整可能な問題生成

本研究は、Tomikawa et al. [12] が提案した項目反応理論 (Item Response Theory: IRT) [16] を用いた難易度調整可能な抽出型読解問題自動生成手法を基礎とする。そこで本章では、この手法について説明する。

Tomikawa et al. では、図 1 左のように、読解対象文 c と所望の難易度 b から、問題文 q とそれに対応する答え a を生成することを目標としている。なお、答え a は読解対象文 c 中に存在する ($a \in c$) 形式が想定されている。このタスクを解く深層学習モデルを構築するためには、これらの 4 つの要素で構成されたデータセット $\{c_i, q_i, a_i, b_i | i \in 1, \dots, I\}$ が必要となる。ここで、 c_i, q_i, a_i, b_i は i 番目のデータの読解対象文、難易度、問題文、答えをそれぞれ表し、 I はデータ数を表す。しかし、一般的なデータセットには問題の難易度が含まれていない。そこで Tomikawa et al. はデータセット中の各問題に対して IRT モデルの一つであるラッシュモデル [17, 18] に基づいて難易度を推定し、難易度を含んだデータセットを構築する手法を提案している。ラッシュモデルに基づいて難易度付きの問題データセットを構



図 1: 答え抜き出し形式の問題生成 (既存研究でのタスク) と多枝選択式問題生成 (本研究でのタスク)

構築する手順は次のとおりである。まずデータセット中の各問題 q_i を性能の異なる多数の QA 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 システムを仮想的な学習者とみなして代用している。次に、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 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している。

このように構築した難易度付きデータセットを用いて、読解対象文から所望の難易度に沿った答えを抽出する BERT モデル [19] と読解対象文と答え、および所望の難易度から問題文を生成する T5 モデル [20] を構築することで、難易度調整可能な答えと問題文の生成を行う。

しかし、この手法では多枝選択式問題を生成できない。また、3.1 で詳述するように難易度を所与とした問題単語列の尤度を最大化するように問題生成モデルが訓練されており、難易度調整の精度を最適化していないため、難易度調整精度に改善の余地が残る。

3 提案手法

以上の問題を解決するために、本研究では、Tomikawa et al. の枠組みをベースに、自己回帰型の深層学習モデルを用いた難易度調整可能な多枝選択式問題生成手法を開発するとともに、DPO を用いて難易度調整精度を最大化する手法を開発する。

3.1 教師あり学習に基づく訓練

本研究では、図 1 右のように、読解対象文 c と難易度 b を入力すると、難易度に応じた問題文 q と正答選択枝 a 、および 3 つの誤答選択枝 d_1, d_2, d_3 を生成する手法を提案する。具体的には、表 1 の入力 x の形式で読解対象文 c と難易度 b を与えると、表 1

の出力 y の形式で問題文 q 、正答選択枝 a 、3 つの誤答選択枝 d_1, d_2, d_3 を生成するように、Llama 2 [15] を学習することで実現する。なお、表中の $\{ \}$ 内は実際のデータで置き換えられることに注意する。また、一般的なデータセットには難易度 b は含まれていないため、Tomikawa et al. と同様の手順で難易度を付与したデータセットを構築する。

難易度調整可能な多枝選択式問題生成を実現するために、提案手法では以下の負の対数尤度を損失関数として最小化するようにモデルの訓練を行う。

$$\mathcal{L}_{sft} = -\mathbb{E}_{(x,y) \sim \mathcal{D}} [\log \pi(y|x, \phi_{sft})] \quad (1)$$

ここで \mathcal{D} は $\{c_i, q_i, a_i, d_{i1}, d_{i2}, d_{i3}, b_i | i \in 1, \dots, I\}$ からなるデータセットであり、 π は Llama 2、 ϕ_{sft} は訓練対象のパラメータ、 $\pi(y|x, \phi_{sft})$ は x で条件付けされた出力文字列 y の尤度を表す。この訓練を行うことによって、与えられた読解対象文と難易度に沿った問題文と選択枝を y の形式で生成できるようになる。

3.2 難易度調整精度を目的関数とした追加訓練

上記の教師あり学習に基づく訓練を行うことで、指定した難易度を反映した多枝選択式問題生成が可能となる。しかし、この訓練では難易度調整精度を直接的に最大化していない。そこで本研究では、上記の教師あり学習で訓練された問題生成モデルを、難易度調整精度を最大化するように追加訓練する手法を提案する。この訓練のために、本研究では近年の LLM の追加学習手法として広く利用されている [21, 22], DPO [14] を利用する。

DPO を用いた提案手法における損失関数は次のとおりである。

$$\mathcal{L}_{dpo} = -\mathbb{E}_{(x, y_w, y_l) \sim \mathcal{D}_{dpo}} \left[\log \sigma \left(\beta \log \frac{\pi(y_w|x, \phi_{dpo})}{\pi(y_w|x, \hat{\phi}_{sft})} - \beta \log \frac{\pi(y_l|x, \phi_{dpo})}{\pi(y_l|x, \hat{\phi}_{sft})} \right) \right] \quad (2)$$

表 1: 提案モデルの入力に対応するプロンプトの構造と出力の構成

入力 (x): 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. The output format is “ $\langle c \rangle \{ \text{Option 1 (Correct Option)} \} \langle q \rangle \{ \text{Question} \} \langle d1 \rangle \{ \text{Option 2 (Distractor Option)} \} \langle d2 \rangle \{ \text{Option 3 (Distractor Option)} \} \langle d3 \rangle \{ \text{Option 4 (Distractor Option)} \}”$
Context: $\{ \text{Context} \} \langle l \rangle$
出力 (y): $\langle c \rangle \{ \mathbf{a} \} \langle q \rangle \{ \mathbf{q} \} \langle d1 \rangle \{ \mathbf{d}_1 \} \langle d2 \rangle \{ \mathbf{d}_2 \} \langle d3 \rangle \{ \mathbf{d}_3 \}$

ここで、 σ はシグモイド関数、 $\hat{\phi}_{sft}$ は 1 段階目の教師あり学習で推定されたモデルパラメータ、 ϕ_{dpo} は学習されるモデルパラメータを表し、 $\hat{\phi}_{sft}$ を初期パラメータとして訓練が進められる。 y_w は y_l よりも難易度調整精度が高い問題を表し、 \mathcal{D}_{dpo} は DPO のための訓練データセット $\mathcal{D}_{dpo} = \{x_i, y_{wi}, y_{li} | i \in 1, \dots, I\}$ である。

このような訓練データセット \mathcal{D}_{dpo} は、 i 番目の問題データから表 1 の形式で入力 x と出力 y_w を構築し、 i 番目の読解対象文と同じ読解対象文だが異なる難易度の問題データから表 1 の形式で y_l を構築することで得られる。

4 提案手法の有効性評価実験

本章では、提案手法の有効性を評価する実験について述べる。

4.1 実験手順

多枝選択式読解問題生成の研究では、モデルの学習や性能評価に RACE データセット [23] が広く用いられているため、本実験でも RACE データセットを評価に用いる。具体的な実験手順は次の通りである

1. QA システムの構築：RACE の検証データセットを用いて性能の異なる 69 個の QA システムを構築した。具体的には、まず 10 個の事前学習済み深層学習モデルを訓練データ数を変えて訓練することで、能力の異なる 100 個の QA システムを構築した。次に、性能が極端に低い QA システムのデータはノイズとなり得るため、訓練データセット中の問題に対する正答率が 30% 以上の QA システムのみを採択した。その結果、69 個の QA システムが残った。
2. 難易度を付与した訓練データセットの構築：得られた 69 個の QA システムに訓練データセット中の全ての問題を出題し、その正誤反応データを用いてラッシュモデルで各問題の難易度 b_i を推定した。推定した難易度を訓練データセットに加えることで、難易度を付与した訓練データセットを構築した。なお、4.2 節での分析の

ために各 QA システムの能力値も推定した。

3. 難易度調整可能な多枝選択式問題生成モデルの作成：3.1 節で述べたように Llama 2 に対して、読解対象文と難易度を入力すると正答選択枝・問題文・3 つの誤答選択枝を出力するモデルを設計し、教師あり学習に基づく訓練を行うことで難易度調整可能な多枝選択式問題生成モデルを作成した。以降ではこのモデルを QG_{sft} と呼ぶ。
4. 難易度調整精度を目的関数とした追加訓練： QG_{sft} に対して 3.2 節で述べた DPO による追加訓練を行うことで、難易度調整精度を向上させた問題生成モデルを作成した。以降ではこのモデルを QG_{dpo} と呼ぶ。
5. 様々な難易度を指定した問題生成：RACE テストデータセット中の 300 個の読解対象文に対して、-3.0 から 3.0 まで 0.1 刻みで難易度を指定して、表 1 の入力 x の形式に沿った累計 18,300 個の入力データを作成し、それらを QG_{sft} 、 QG_{dpo} にそれぞれ入力することで問題を生成した。次節では、ここで生成された問題群の難易度を評価する。

4.2 生成した問題の難易度調整精度評価

ここでは提案手法の難易度調整精度を評価する。

上記の手順 5 で生成した問題群の難易度を確認するために、生成された問題群に対する 69 個の QA システム群の正誤反応データからラッシュモデルを用いて推定された難易度と、問題生成時に指定した難易度との関係を分析した。結果を図 2 に示す。図の横軸は指定した難易度、縦軸は生成した問題群に対して推定された難易度の平均を表す。黒線は原点を通る傾き 1 の直線で、黒線に近いほど指定した難易度と生成した問題群の難易度が近いことを表す。青点は QG_{sft} が生成した問題群に対する難易度、橙点は QG_{dpo} が生成した問題群に対する難易度を表す。この図より、 QG_{sft} 、 QG_{dpo} 共に指定した難易度が高くなるほど生成する問題の難易度も高くなるのがわかる。また二つの生成手法を比較すると、青

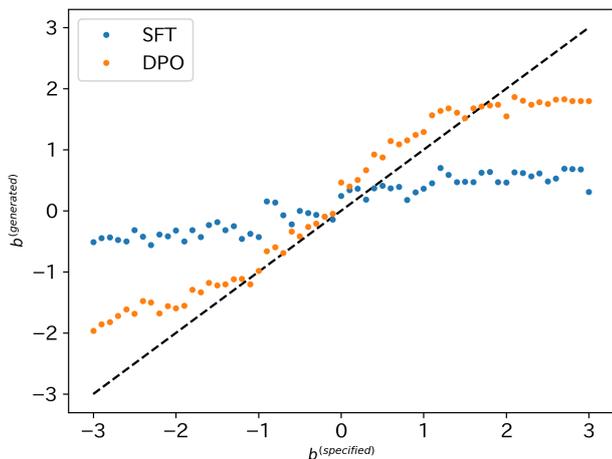


図 2: 生成する際に指定した難易度と実際に生成された問題の難易度の関係

点より橙点の方が黒線に近いことから、 QG_{dpo} の方が指定した難易度に近い難易度の問題群を生成できていることがわかる。生成された問題群に対する難易度推定値と生成時に指定した難易度との平均絶対誤差は、 QG_{sft} が1.75、 QG_{dpo} が1.17であり、この結果からも、 QG_{dpo} が指定した難易度により近い問題を生成できていることがわかる。

したがって、提案手法によって難易度を調整した多枝選択式の問題生成が可能であるとともに、DPOを用いた追加訓練を行うことによって難易度調整精度が向上することが確認できた。

4.3 フューショット学習との比較

LLMは、少数の入出力例を入力プロンプトに含めたフューショット学習によって様々なタスクを高精度に解くことができる [24] ため、プロンプトで難易度を指定することで難易度調整を実現できる可能性がある。そこで本研究では、GPT-4o miniを用いて難易度調整を目的としたフューショット学習による問題生成を行い、提案手法と難易度調整精度の比較を行った。具体的な実験手順は次の通りである。

1. 難易度を付与した RACE の訓練データセット中から難易度が -2.0, 0.0, 2.0 の問題をランダムに一つずつ選び、それぞれのデータに対応する入出力 x, y を作成した。これらを 1 ショット分の例示プロンプトとみなし、同様の手順で 3 ショット, 10 ショット, 30 ショット分の例示プロンプトを作成した。
2. RACE のテストデータセットからランダムに

表 2: 提案手法とフューショット学習の難易度調整の比較

指定した 難易度	QG_{sft}	QG_{dpo}	フューショット学習		
			3ショット	10ショット	30ショット
-2.0	-0.32	-1.59	-1.46	-1.30	-0.91
0.0	0.24	0.47	-1.67	-1.10	-0.75
2.0	0.47	1.55	-1.19	-1.34	-0.79

100 個の読解対象文を選択し、各読解対象文に対して難易度を -2.0, 0.0, 2.0 とそれぞれ指定して表 1 の入力 x の形式で累計 300 個の指示プロンプトを作成した。

3. 手順 1 の各例示プロンプトと手順 2 の指示プロンプトをそれぞれ結合して GPT-4o mini に入力することで、各ショットに対応した問題生成を行った。

これらの生成した問題群を 4.1 節の手順 1 で作成した 69 個の QA システムに解答させ、正誤反応データからラッシュモデルに基づいて難易度を推定した。表 2 に、この手続きで生成された問題群に対する難易度推定値の平均を、生成時に指定した難易度別に示す。表には、4.1 節の実験で QG_{sft} , QG_{dpo} にそれぞれの難易度を指定して生成された問題群に対する難易度推定値の平均も示す。

表より、 QG_{sft} や QG_{dpo} では指定した難易度が高くなるほど生成した問題の難易度も高くなるのに対し、フューショット学習では指定した難易度と生成した問題の難易度にそのような関係を確認できない。このことは、IRT の尺度に基づく難易度調整機能を獲得させるためにはフューショット学習では不十分であり、提案手法のようなモデル学習の枠組みが必要であることが示された。

5 おわりに

本研究では、自己回帰型の深層学習モデルである Llama 2 を用いた難易度調整可能な多枝選択式問題生成手法を開発するとともに、DPO により難易度調整精度を最大化する訓練手法を提案した。RACE データセットを用いた評価実験により、提案手法では難易度を調整した多枝選択式の問題生成が可能であるとともに、DPO を用いた追加訓練を行うことによって難易度調整精度が向上することを確認した。さらに、IRT の尺度に基づく難易度調整機能を獲得させるためにはフューショット学習では不十分であり、提案手法のようなモデル学習の枠組みが必要であることも示された。

謝辞

本研究は JSPS 科研費 24H00739, 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 **Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics**, pp. 1342–1352, 2017.
- [2] Qingyu Zhou, Nan Yang, Furu Wei, Chuanqi Tan, Hangbo Bao, and Ming Zhou. Neural question generation from text: A preliminary study. In **Proceedings of the 12th Natural Language Processing and Chinese Computing**, pp. 662–671, 2018.
- [3] Ruslan Mitkov and Le An Ha. Computer-aided generation of multiple-choice tests. In **Proceedings of the HLT-NAACL 03 workshop on Building educational applications using natural language processing, North American Chapter of the Association for Computational Linguistics**, pp. 17–22, 2003.
- [4] Andreea Dutulescu, Stefan Ruseti, Denis Iorga, Mihai Dascalu, and Danielle S McNamara. Beyond the obvious multi-choice options: Introducing a toolkit for distractor generation enhanced with NLI filtering. In **Proceedings of the 25th International Conference on Artificial Intelligence in Education**, pp. 242–250, 2024.
- [5] Subhankar Maity, Aniket Deroy, and Sudeshna Sarkar. A novel multi-stage prompting approach for language agnostic MCQ generation using GPT. In **Proceedings of the 46th European Conference on Information Retrieval**, pp. 268–277, 2024.
- [6] Maomi Ueno and Yoshimitsu Miyazawa. IRT-based adaptive hints to scaffold learning in programming. **IEEE Transactions on Learning Technologies**, Vol. 11, No. 4, pp. 415–428, 2017.
- [7] Ghader Kurdi, Jared Leo, Bijan Parsia, Uli Sattler, and Salam Al-Emari. A systematic review of automatic question generation for educational purposes. **International Journal of Artificial Intelligence in Education**, Vol. 30, pp. 121–204, 2020.
- [8] Yifan Gao, Lidong Bing, Wang Chen, Michael R Lyu, and Irwin King. Difficulty controllable generation of reading comprehension questions. In **Proceedings of the 28th International Joint Conference on Artificial Intelligence**, pp. 4968–4974, 2018.
- [9] 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 **Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing**, pp. 5968–5978, 2021.
- [10] Masaki Uto, Yuto Tomikawa, and Ayaka Suzuki. Difficulty-controllable neural question generation for reading comprehension using item response theory. In **Proceedings of the 18th Workshop on Innovative Use of NLP for Building Educational Applications, Association for Computational Linguistics**, pp. 119–129, 2023.
- [11] 富川雄斗, 鈴木彩香, 宇都雅輝. 項目反応理論に基づく難易度調整可能な読解問題自動生成手法. 電子情報通信学会論文誌 D, Vol. J107-D, No. 2, pp. 53–66, 2024.
- [12] Yuto Tomikawa, Ayaka Suzuki, and Masaki Uto. Adaptive question-answer generation with difficulty control using Item Response Theory and pre-trained transformer models. **IEEE Transactions on Learning Technologies**, 2024.
- [13] Teruyoshi Goto, Yuto Tomikawa, and Masaki Uto. Enhancing diversity in difficulty-controllable question generation for reading comprehension via extended T5. **Proceedings of the 32nd International Conference on Computers in Education**, 2024.
- [14] Rafael Rafailov, Archit Sharma, Eric Mitchell, Christopher D Manning, Stefano Ermon, and Chelsea Finn. Direct preference optimization: Your language model is secretly a reward model. In **Proceedings of Advances in Neural Information Processing Systems 36**, pp. 53728–53741, 2023.
- [15] Hugo Touvron, Louis Martin, Kevin Stone, Peter Albert, Amjad Almahairi, Yasmine Babaei, Nikolay Bashlykov, Soumya Batra, Prajjwal Bhargava, Shruti Bhosale, et al. Llama 2: Open foundation and fine-tuned chat models. arXiv:2307.09288, 2023.
- [16] Frederic M Lord. **Applications of item response theory to practical testing problems**. Routledge, 2012.
- [17] Georg Rasch. **Probabilistic models for some intelligence and attainment tests**. ERIC, 1993.
- [18] WJ van der Linden. **Handbook of Item Response Theory: Volume 2: Statistical Tools**. CRC Press, 2017.
- [19] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. In **Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics**, pp. 4171–4186, 2019.
- [20] Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J Liu. Exploring the limits of transfer learning with a unified text-to-text transformer. **Journal of machine learning research**, Vol. 21, No. 140, pp. 1–67, 2020.
- [21] Albert Q Jiang and *et al.* Mixtral of experts. arXiv:2401.04088, 2024.
- [22] Abhimanyu Dubey and *et al.* The Llama 3 herd of models. arXiv:2307.09288, 2024.
- [23] Guokun Lai, Qizhe Xie, Hanxiao Liu, Yiming Yang, and Eduard Hovy. RACE: Large-scale reading comprehension dataset from examinations. In **Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing**, pp. 785–794, 2017.
- [24] Tom B Brown and *et al.* Language models are few-shot learners. arXiv:2005.14165, 2020.