

読解対象文の難易度を考慮した読解問題自動生成手法

Difficulty-Controllable Reading Comprehension Question Generation Considering Levels of Reading Passages

富川 雄斗^{*1}, 宇都 雅輝^{*1}
Yuto Tomikawa^{*1}, Masaki Uto^{*1}

^{*1} 電気通信大学

^{*1}The University of Electro-Communications

Email: {tomikawa, uto}@ai.lab.uec.ac.jp

あらまし: 近年, 学習者の読解力にあった難易度の読解問題を深層学習を用いて自動生成する読解問題自動生成手法が提案されている. しかし, 既存手法では, 問題生成の元になる読解対象文の難易度を考慮していないという課題がある. そこで, 本研究では, 読解対象文の難易度も考慮した問題生成技術を開発する. 具体的には, 問題の難易度とともに読解対象文の難易度も推定できる項目反応理論モデルを開発し, それにより得られる難易度値に基づいて学習者の読解力にあった難易度の読解対象文選定と問題生成を行う方法を提案する.

キーワード: 読解問題, 問題生成, 深層学習, 言語モデル, 項目反応理論, 言語生成

1 はじめに

読解力育成のためには, 学習者に様々な読解文を与え, それに関連する多様な読解問題に取り組みせる学習支援が有効といえる. しかし, 人手で多様な読解問題を作成することは時間的・費用的コストが高いという問題がある. この問題を解決するために, 深層学習を用いた読解問題自動生成技術が近年注目を集めている.

問題生成技術を学習支援に利用する場合, 単に問題を生成するだけでなく, 学習者の読解力に応じた難易度の問題を生成することが重要である. このような生成を目指し, 近年では, 難易度を調整可能な読解問題生成技術も提案されている. 例えば, 鈴木・宇都 [1] は, 項目反応理論 (Item Response Theory: IRT) に基づいて推定された問題難易度と読解対象文を入力として, 問題と答えを生成する深層学習ベースの問題生成手法を提案している. このような既存手法は, 生成する問題と答えの難易度調整のみを目的としているが, 実際には読解対象文の難易度も問題の難易度に影響すると考えられる. 例えば, 難易度の低い読解対象文から難易度の高い問題を生成することは, 難易度の高い読解対象文から難易度の高い問題を生成する場合に比べて困難と予想できる. また, 効果的な学習のためには, 各学習者のレベルに合った難易度の読解対象文を与えることも重要といえる.

そこで本研究では, 読解対象文の難易度も考慮した問題生成技術を開発する. 具体的には, まず, 問題の難易度とともに読解対象文の難易度を推定するために, IRT モデルの一つである Rasch モデルを階層ベイズ化した新たなモデルを開発する. さらに, 提案 IRT モデルを用いて推定した読解対象文の難易度を基に学習者の能力値にあった難易度の読解対象文を選定し, その文章に基づいた適切な難易度の問題を生成する方法を提案する.

2 従来手法

本研究では, 鈴木・宇都 [1] が開発した難易度調整可能な読解問題生成技術を基礎技術として利用する. この

手法では, 次の手順で問題生成モデルを構築する.

1. 読解対象文とそれに関する問題文および答えで構成されるデータセット (例: SQuAD データセット) を用い, 各問題に対する学習者の正誤反応データを得る. なお, 既存手法では, 人間の学習者を QA (Question Answering) システムで代用している.
2. 得られた正誤反応データに, 基礎的な IRT モデルの一つである Rasch モデルを適用し, 各問題の難易度を推定する.
3. SQuAD データセット内の読解対象文と IRT で推定した問題難易度を入力とし, 問題と答えを生成するように深層学習モデルを訓練する. なお, 深層学習モデルには BERT と GPT-2 を用いており, 答えの抽出には BERT を, 問題文生成には GPT-2 を使用している.

以上で訓練されたモデルに任意の読解対象文と所望の難易度を入力することで, 難易度調整可能な問題生成を実現している.

3 提案手法

1 章で述べたように, 本研究の目標は, 読解対象文の難易度を考慮した問題生成技術の開発である. これを実現するために, まず, 問題の難易度とともに読解対象文の難易度も推定できる IRT モデルを提案する.

3.1 読解対象文の難易度推定のための IRT モデル

本研究では, 読解対象文と問題の難易度を推定できる IRT モデルとして, Rasch モデルを階層ベイズ化した次式のモデルを提案する.

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

$$\theta_j \sim \mathcal{N}(0, 1), \quad b_i \sim \mathcal{N}(b_i^{(r)}, \sigma), \quad b_i^{(r)} \sim \mathcal{N}(0, 1) \quad (2)$$

ここで, θ_j は学習者 j の能力, b_i は問題 i の難易度, $b_i^{(r)}$ は問題 i に対応する読解対象文 r の難易度を表す. また, σ はハイパーパラメータである. 提案モデルのグラ

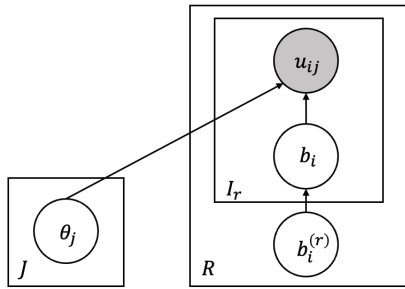


図1 読解対象文の難易度を考慮したIRTモデル. J は学習者数, R は読解対象文数, I_r は r 番目の読解対象文に対応する問題数を表す.

フィカルモデルを図1に示す.

このモデルを用いることで, 各問題への学習者の正誤反応データの集合から, 各問題の難易度 b_i と各読解対象文の難易度 $b_i^{(r)}$ を推定できる. なお, パラメータ推定には, No-U-turn サンプラに基づくマルコフ連鎖モンテカルロ法を用いる.

3.2 学習者の能力に合った読解対象文選定と問題生成

本研究では, 提案IRTモデルを用いて推定した読解対象文の難易度を基に, 学習者の能力値に合った読解対象文を選定した上で, 適切な難易度の問題生成を行う方法を提案する. 先行研究[2]より, 効果的な学習のためには, 学習者が50%の確率で正答できる難易度の問題の出題が有効とされている. したがって, 本研究では, 学習者が50%の確率で正答できる難易度になるように, 読解対象文の選定と問題の生成を行う.

提案IRTモデルでは, $b = \theta$ のときに学習者の正答確率が0.5となる. また, 式(2)より, b は $b^{(r)}$ を平均とする正規分布に従うため, $b^{(r)} = b = \theta$ の読解対象文を利用することで適切な難易度の問題生成が期待できる. 以上より, 本研究では, $b^{(r)} = \theta$ となる読解対象文を選択し, 問題難易度として $b = \theta$ を指定して問題生成を行う. ただし, 実際の学習環境において学習者の能力値 θ が事前に測定されている機会は少ない. そこで本研究では, 適応型テストの枠組みで, θ の推定と能力推定値 $\hat{\theta}$ に応じた読解対象文選定と問題生成を行う手法を提案する.

提案手法では, まず, 2章で説明した鈴木・宇都と同様の手法で, 難易度調整可能な問題生成のための深層学習モデルを訓練する. ただし, 問題難易度の推定には提案IRTモデルを利用する. その後, 訓練された問題生成モデルを利用して, 以下の手順で適応的な読解対象文選定と問題生成を行う. 1) 学習者の能力値の初期値を $\hat{\theta} = 0.0$ とする. 2) $\hat{\theta}$ に近い難易度 $b^{(r)}$ の読解対象文を選定する. 3) 選定した難易度の読解対象文と問題の難易度 $b = \hat{\theta}$ を所与として, 訓練済み問題生成モデルで問題文と答えを生成する. 4) 生成した問題を学習者に出題し, 正誤反応データ u を取得する. 5) これまでに問題を出題した問題の難易度 b と正誤反応データ u から学習者の能力推定値 $\hat{\theta}$ を更新する. 6) 手順2~5を繰り返す. 以上の手順により, 学習者の能力値に応じた難易度の読

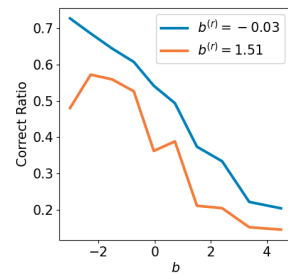


図2 QAシステムの正答率

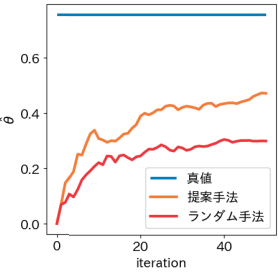


図3 適応的出題の結果

読対象文の選定と問題の生成が実現できる.

4 実データ実験

ここでは, 実データ実験を通して, 適切な難易度の読解対象文を選定することの重要性を示す. 本実験では, 難易度に差がある読解対象文から多様な難易度の問題を生成し, それらの問題に対するQAシステムの正答率に基づいて, 読解対象文の難易度差によって生成される問題の難易度に差が生じるかを評価した. 図2に実験結果を示す. 図の横軸は問題の難易度, 縦軸はQAシステムの正答率, 青線は $b^{(r)} = -0.03$, 橙線は $b^{(r)} = 1.51$ の読解対象文から生成した問題を表す. 図から, 読解対象文の難易度が高いほど指定する問題の難易度によらず正答率が低くなる傾向が確認できる. このことは, 読解対象文の難易度が生成される問題の難易度に影響を与えることを示しており, 適切な難易度の問題生成のためには読解対象文の選択が重要であることを示唆している.

次に, 適応的出題の性能を評価する実験を行った. ここでは, 読解対象文の選定方法として, 提案手法による適応的選定とランダムな難易度での選定の2種類を比較した. なお, 生成する問題の難易度はともに $b = \hat{\theta}$ とした. 各手法で推定された能力値 $\hat{\theta}$ の推移を図3に示す. 図の横軸は問題出題数, 縦軸は $\hat{\theta}$, 青線は θ の真値, 橙線は提案手法で得られた $\hat{\theta}$, 緑線はランダム手法で得られた $\hat{\theta}$ を表す. この結果より, 提案手法では, 学習者の真の能力値がより高精度に推定されたことがわかる. このことから, 提案手法は学習者の真の能力値により適した難易度の読解対象文選定と問題生成が可能といえる.

5 まとめ

本研究では, 読解対象文の難易度を考慮した読解問題自動生成手法を開発した. 今後の課題として, 人間の評価者を用いた実験などが挙げられる.

参考文献

- [1] 鈴木彩香, 宇都雅輝: Transformerモデルを用いた難易度調節機能付き読解問題自動生成手法, 日本行動計量学会大会抄録集, pp.126-129, 2022.
- [2] M. Ueno and Y. Miyazawa: IRT-based adaptive hints to scaffold learning in programming, IEEE Transactions on Learning Technologies, vol.11, no.4, pp.415-428, 2018.