

強化学習を用いた駄洒落の自動生成

Generative Language Model for Japanese Puns using Reinforcement Learning

塚見 大吾郎†

全 へい東‡

Daigoro Tsukami

Heitoh Zen

1. はじめに

Transformer を始めとする生成 AI 技術の発展に伴い、テキストや画像の生成などの多様なタスクへの応用が試みられている。特に人間の評価を用いて言語モデルを強化学習で微調整する手法 (RLHF) は OpenAI の ChatGPT や InstructGPT のモデルを訓練するために用いる中心技術として注目されている。先行研究や現在の生成 AI で幅広いタスクに対応できている一方で、駄洒落などの音韻の一致を含む文の生成を行うことはできていない。本稿ではこの RLHF の手法を応用し、駄洒落の生成を対象とする。

1.1 TRL (Transformer Reinforcement Learning)

TRL は RLHF の学習方法のひとつである。TRL では言語モデルを学習する教師データを強化学習によって生成する。言語モデルを二つ使用し、一つは参照用モデル (Ref. LM) として固定し、もう一方のモデルを学習用のモデル (Active. LM) としてパラメータの更新を行う。TRL では目的関数を最大化する PPO とよぶ最適化アルゴリズムを用いて勾配の更新を行う。PPO では方策の目的関数に使用される Advantage の推定量として GAE が使用される。生成した文の日本語らしさを維持するために Kullback-Leibler Divergence (以下 KL Div.) を即時報酬とし、タスクに応じた報酬を加算して Advantage を推定する。

2. 関連研究

荒木らはユーザー投稿型の 9 つの駄洒落サイトから収集した約 70,000 件のデータの駄洒落データベースを構築している^[1]。また駄洒落を生成するタスクの例として、与えた「題」からクエリを生成し、駄洒落データベース内を検索するシステムも考案している^[2]。Yu らは LSTM を用いて駄洒落を機械学習で自動生成する手法を提案している^[3]。畠山らは初めて Transformer モデルで駄洒落の自動生成を試み、日本語の流暢さと音の反復を考慮した学習手法を提案している^[4]。著者らは GPT-2 モデルに音韻の一致する異なる単語を二つ入力し、生成された文の中から駄洒落を識別する駄洒落生成システムを検討してきた。

一方、駄洒落を認識するタスクでは、谷津らが語彙特性や子音列の音韻類似性を基にした SVM を用いた分類手法を提案している^[5]。

3. 提案手法

本研究では GPT-2 を用いた言語モデルを使い、「種」と呼ぶ単語を入力として駄洒落を生成する。図 1 に本研究で用いた TRL の学習時の構造を示した。Query に「種」を入力すると Active. LM から Response が生成される。ここで Ref. LM と Active. LM には、3.2 で述べる追加学習済みの言語モデル (以下 FT-Pun LM) を用いる。

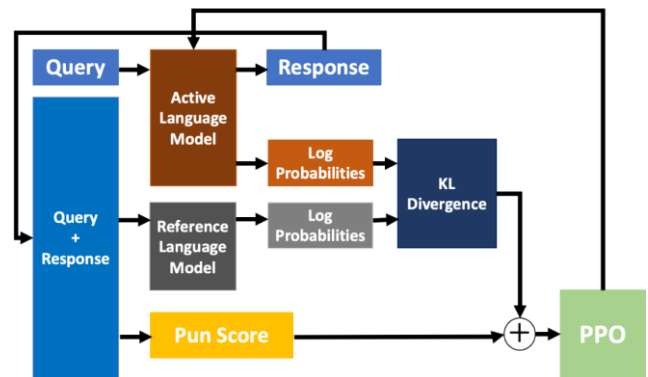


図 1 TRL による学習時の構造

Ref. LM と Active. LM に「種」を入力すると駄洒落と同等の長さを持つ文章が生成される。Ref. LM と Active. LM の Log-Probabilities から算出する KL Div.により、生成される文の「日本語らしさ」を算出する。一方、Pun Score では、ルールベースの報酬 (3.3. 参照) により「駄洒落らしさ」を算出する。これら二つの値を加算し、強化学習の報酬にすることで日本語らしさを保ちつつ、かつ駄洒落らしい文章が生成できるよう学習する。

実験では Advantage の推定に次のふたつの学習方法を試みた。

- ① Advantage の推定に KL Div. のみを用いて学習を行ったモデル (TRL-Pun LM)
- ② Advantage の推定に KL Div. に加え、3.4. で述べるトークン音韻報酬を用いて学習を行ったモデル (Proposed-Pun LM)

3.1 定義

この研究では「種」と同一の音韻の異義語を含む文を駄洒落と定義する。ただし長音や促音は無視する。

例) 箸が走った アルミ缶の上にあるミカン

3.2 追加学習

TRL によって学習をする前に、Rinna 社が公開している GPT-2 に荒木らの「北大駄洒落コーパス」^[1]に含まれる併置型駄洒落のみ約 60000 個を対象とし、「種」と駄洒落の「種」部を[MASK]したものを対応づけてファインチューニングを行う。

3.3 報酬設計

「種」と生成文の音韻の一致度に応じて報酬を決定する。計算手順は下のとおりである。

【計算手順】

- ① 「種」と生成文から長音・促音を除去し、ひらがなに変換する
- ② 生成文の先頭から「種」の文字列と比較し、以下の式に従って Score を計算する。

$$Point = \begin{cases} 1 & (\text{完全一致}) \\ 0.3 & (\text{母音一致}) \\ 0.1 & (\text{子音一致}) \end{cases}$$

$$Score = \max_{j=1,2,..,Len(\text{生成文})-Len(\text{種})+1} \sum_{i=1}^{Len(\text{種})} Point_{j,i}$$

- ③ スライド 1 でずらして比較を行い、最終的に以下の式を用いて報酬を計算する

$$Reward = \frac{Score}{Len(\text{種})}$$

3.4 トークン音韻報酬

トークン単位で音韻の影響を学習させるためにトークン音韻報酬 (以下 TP Rew.) を考案した。トークン単位で「種」の音韻に一致する単語が含まれている場合に正の報酬を与え、その次のトークンに残りの音韻が含まれていない場合にはペナルティとして負の報酬を与える (図 2)。

この TP Rew. と KL Div. を即時の報酬として Advantage の推定を行い、最適化を行う。

	種	アルミ缶						
例文1	(TEXT)	[MASK]	の	上	に	ある	みかん	
	(SR)	0	0	0	0	0.2	0.2	
例文2	(TEXT)	[MASK]	の	上	に	ある	りんご	
	(SR)	0	0	0	0	0.2	-0.2	

「ある」→Good
「りんご」→Bad

図 2 トークン音韻報酬の計算方法

4. 結果

500 個の「種」に対して、ファインチューニングのみのモデルと 2 つの提案手法の、あわせて 3 つのモデルで生成し、駄洒落としてのスコアを比較した。結果を図 3 に示す。Proposed-Pun LM が約 15% の「種」に対してスコア 1 の文を生成しており、3 つのモデルのなかで最も優れていた。生成された駄洒落の一部を下に抜粋する。

生成した駄洒落文の例：

- ・シリアの経済は尻上がり
- ・毛布なんてもう古い!
- ・レタスが枯れたっす

5. 考察

TRL の手法を用いることにより駄洒落の生成率が大幅に向上したことから、強化学習を用いた教師データ生成が有効だったといえる。

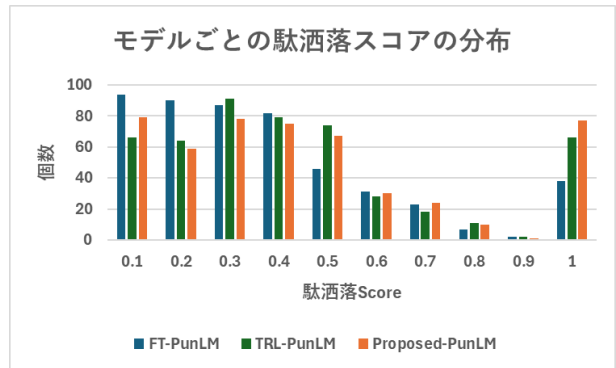


図 3 モデルごとの駄洒落スコア分布

本研究で提案したトークン音韻報酬を学習に組み込むことで更に生成率が向上している為、Advantage の推定に考案したトークン音韻報酬が有効であったといえる。

6. おわりに

強化学習を用いた言語モデル学習方法である TRL により、駄洒落を生成するモデルを提案した。トークンの音韻を強化学習の報酬に加味することにより、生成される文に含まれる駄洒落の割合が向上した。

今後は駄洒落の定義を広げ、例えば「すみま千羽鶴」のような重畳型駄洒落の生成に取り組む予定である [6]。また、他のタスクへの応用についても検討する。駄洒落は日本語固有の音韻が反復される点にある。このような特徴を有する文には、他に語呂合わせや回文などがある。提案したモデルをこれらのタスク向けに学習させ、性能を検証する予定である。

参考文献

- [1] 荒木 健治, 佐山 公一, 内田 ゆず, 谷津 元樹 “駄洒落データベースの構築及び分析”, 人工知能学会第 2 種研究会 ことば工学会研究資料, Vol. 3, No. 1, pp. 1-15 (2018).
- [2] 荒木 健治, “駄洒落データベースを用いた駄洒落生成システムの性能評価”, 人工知能学会第 2 種研究会 ことば工学会研究資料, Vol. 3, No. 8, pp. 39-48 (2018).
- [3] Zhiwei Yu, Jiwei Tan and Xiaojun Wan. “A neural approach to pun generation”, In Proc. of the 56th Annual Meeting of the Association for Computational Linguistics (ACL), pp. 1650-1660 (2018).
- [4] 畠山 和久, 徳永 健伸, “Transformer を用いた日本語併置型駄洒落の自動生成”, 言語処理学会第 27 回年次大会発表論文集, pp. 455-460 (2021)
- [5] 谷津 元樹, 荒木 健治, “子音の音韻類似性及び svm を用いた駄洒落検出手法”, 知能と情報 (日本知能情報ファジィ学会誌), Vol. 28, No. 5, pp. 833-844 (2016).
- [6] 井上 蒼一郎, 谷津 元樹, 森田 武史, “重畳型駄洒落ユーモアにおける常識的知識グラフを用いた潜在表現抽出”, 言語処理学会第 28 回年次大会発表論文集, pp. 1209-1214 (2022).

† 千葉大学大学院融合理工学府数学情報科学専攻
Chiba University, Graduate School of Science and Engineering, Div. of Mathematics and Informatics

‡ 千葉大学情報戦略機構
Chiba University, Digital Transformation Enhancement Council