誤りを経験して修正する：誤りデータの正例扱いによる対照学習

薛 強 [1] 滝口 哲也 [1] 有木 康雄 [1]
[1] 神戸大学システム情報学研究科
xueqiang@stu.kobe-u.ac.jp {takigu,ariki}@kobe-u.ac.jp

概要

近年，大規模言語モデルの性能向上に伴い，対照学習（contrastive learning）を活用した誤生成抑制の手法が注目を集めている．本研究では，人間の「失敗から学ぶ」学習プロセスを模倣し，対照学習において誤りをあえて一度強化してから修正を行う新たな学習手法を提案する．具体的には，通常の正例データと負例データに加え，負例データの誤りトークンをあえて「正例扱い」するデータを作り，一時的に誤りを増幅させてから再度誤りを強く抑制する学習ステップを導入する．これによりモデルが誤りを深く認識し，修正効果の向上を狙う．OpenDialKG を用いた対話生成タスクの実験では，提案手法が従来の対照学習よりも高い性能を示すことを確認した．

1 はじめに

深層学習技術の急速な発展と大規模データの活用により，言語モデルは従来の手法では困難だった多種多様な自然言語処理タスクにおいて顕著な成果を上げている [1, 2]．一方で，モデル出力が完全に正しいとは限らず，しばしば文脈不整合や知識エラーなどの誤生成を起こすことが問題視されている．誤生成を低減するために，モデル構造の改良や大規模事後学習といった手法が検討されてきたが，近年は対照学習（contrastive learning）の枠組みが自然言語処理の分野でも注目を集めつつある [3]．対照学習では，正例と負例を対比させることで，正例の生成確率を高めると同時に，負例の生成確率を抑制する学習を行う．

しかし，人間の学習プロセスに着目すると，「誤りを一度経験してから修正する」ほうが記憶の定着や再発防止に有効である場合が多い．たとえば誤った解法を実際に試してから，その誤りを理解することで，類似の誤りが繰り返しにくくなる．本研究では，このような「失敗をあえて経験し，その後で修正する」という学習プロセスがモデルの精度向上にも寄与すると考え，負例を一時的に増幅させてから抑制する新しい対照学習手法を提案する．言い換えれば，負例を「いったん高いところに持ち上げてから落とす」ことで，誤りの修正効果をより強くすることを狙っている．

本論文の構成は以下のとおりである．まず関連研究として，対照学習と失敗学習に関する先行研究を概説する（§2）．次に，提案手法の詳細を示し，正例データ・負例データ・誤りデータの三種類の損失を組み合わせる学習方法を説明する（§3）．その後，OpenDialKG データセットを用いた対話生成タスクでの実験を行い，提案手法が既存の対照学習より優れた性能を示すことを検証する（§4）．最後に，本研究を総括し，今後の展望を述べる（§5）．

2 関連研究

2.1 対照学習

Contrastive Learning (CL) [4, 5] は，類似したサンプル同士を表現空間で近づけ，異なるサンプル間は遠ざけるという発想に基づく学習手法である．自然言語処理の分野では，文埋め込み学習 [6] や情報検索タスク [7]，文章生成における不適切出力の抑制 [8, 9] など，幅広い応用が報告されている．

特に文生成の分野では，応答の繰り返しやモラル違反，退屈な応答，文脈不整合といった誤りを抑制する目的で Unlikelihood Training [8] や Negative Training [9] が提案されており，対照学習の考え方を活用した研究も多い [10, 11]．また，要約生成における幻覚（hallucination）の抑制を目的とした対照学習手法も近年報告されている [12]．

本研究は，負例を単に抑制するだけでなく，誤り

出力を一度強化してから学習するという点が既存手法と大きく異なる．こうしたアプローチにより，モデルが誤りをより深く認識し，最終的な精度向上につながる可能性を検証する点に新規性がある．

2.2 失敗を活かす学習

誤りや失敗を学習に活かす手法は，強化学習における Experience Replay や Prioritized Experience Replay [13]，学習カリキュラムで難度の高いサンプルを意図的に提示する方法 [14]，および Hindsight Experience Replay [15] など，多様な形で検討されている．これらの研究では，単にエラーを排除するのではなく，エラーの分析や再利用によって学習効率や汎化性能が向上すると報告されている．本研究では，言語モデルの対照学習において，誤りトークンの生成確率をあえて高めるフェーズを設けることで，誤りの修正効果を高める点が特徴的である．

3 提案手法

提案手法では，(1) 正例データ（Posi データ），(2) 負例データ（Nega データ），(3) 誤りデータ（“nega-as-posi”データ）を同時に学習する．以下では，それぞれに対応した損失関数と，総合的な最適化戦略を述べる．

3.1 正例データ学習

まず，入力 x と正解応答 y が与えられた場合，式 (1) の損失関数で学習を行う．これは一般的な言語モデル学習と同様に，正解トークンを高確率で生成できるようにパラメータを更新するものである．

$$\mathcal{L}_{\text{Posi}} = -\sum_{t=1}^{|y|} \log p_\theta(y_t \mid y_{<t}, x), \tag{1}$$

ここで p_θ はモデルパラメータ θ に基づく条件付き確率を示す．

3.2 負例データ学習

負例データ y^- は，元の正例応答に対して文やエンティティをランダムに置換することで誤りを導入したものである（詳細は §4.2 参照）．y^- 内には本来の正例トークン（置換されなかった部分）と誤りトークン（置換や挿入によって生じた部分）が混在している．そこで，それぞれに対して異なる目的で損失を定義する．

具体的には，負例トークン集合を

$$T_{\text{nega}} = \{\, t \mid y_t^- \text{ の内誤りトークン} \,\},$$

正例トークン集合を

$$T_{\text{posi}} = \{\, t \mid y_t^- \text{ の内元の正例トークン} \,\}$$

とし，式 (2) の損失関数を用いて学習を行う．すなわち，誤りトークンの生成確率を抑える（$1 - p_\theta(\cdot)$ の項）一方で，元の正例トークンは引き続き正しく生成できるようにする．

$$\begin{aligned}\mathcal{L}_{\text{Nega}} = &-\sum_{t \in T_{\text{nega}}} \log\bigl(1 - p_\theta(y_t^- \mid y_{<t}^-, x)\bigr) \\ &-\sum_{t \in T_{\text{posi}}} \log p_\theta(y_t^- \mid y_{<t}^-, x).\end{aligned} \tag{2}$$

3.3 誤りデータ学習

上記の負例応答と同一の y^- を，“nega-as-posi”として一時的に「誤りトークンを正しく生成する」よう扱うステップを導入する．ここでは誤りトークンの集合を

$$T_{\text{err}} = \{\, t \mid y_t^* \text{ が誤りトークン} \,\}$$

とし，それ以外のトークンは損失に含めない．式 (3) に示すように，誤りトークンのみを強化する学習を行う．y_t^* は y_t^- と同じであるが，「全体を誤りトークンとして正しく生成」するため，誤りデータを y_t^* として表現している．

$$\mathcal{L}_{\text{error}} = -\sum_{t \in T_{\text{err}}} \log p_\theta(y_t^* \mid y_{<t}^*, x). \tag{3}$$

その後の再学習（§3.2 での $\mathcal{L}_{\text{Nega}}$）で，これら誤りトークンを強く抑制することで，「高いところから落とす」効果を狙う．これにより，モデルが誤りを深く認識し，結果として最終的に誤生成を低減する．

3.4 最適化

最終的な総合損失 $\mathcal{J}(\theta)$ は，以下のように定義する．

$$\mathcal{J}(\theta) = \alpha_1 \mathcal{L}_{\text{Posi}} + \alpha_2 \mathcal{L}_{\text{Nega}} + \alpha_3 \mathcal{L}_{\text{error}}. \tag{4}$$

ここで $\alpha_1, \alpha_2, \alpha_3$ はそれぞれの損失の重みである．本研究では，これらを適宜調整しつつ，ミニバッチ内で正例，負例，誤りデータを同時に学習させる形で最適化を行う．

4 実験

本章では，OpenDialKG を用いた知識対話生成タスクに対して提案手法を適用し，その有効性を検証する．実験設定やベースライン，評価指標を概説し，実験結果の考察を行う．

4.1 データセット

OpenDialKG [16] は，知識グラフに基づくオープンドメイン対話を収集したデータセットである．各対話ターンに対応する推論経路がアノテーションとして付与されており，構造化された知識を活用するタスクに適している．本研究では，先行研究 [17] に倣い，データを **Test Seen** と **Test Unseen** に分割して使用した．

4.2 負例データの種類

本研究で扱う負例データ（y^-）は，元の正例応答を改変して文脈不整合や誤知識を含む応答を作ったものである．具体的には以下の三種類を用いた．なお，ランダム Span 置換およびランダム知識置換は，[18] の手法を参考にしている．

ランダム置換文　正例応答の一部を，他の対話データからランダムに抜き出した文片で置き換える．

- **正例応答例**: *“Yes! Eugene also wrote and starred in A Mighty Wind, have you heard of it?”*
- **ランダム文片例**: *“Thank you. Did Caroline Goodall happen to be in the movie White Squall?”*
- **負例応答例**（上記を組み合わせ）: *“Yes! Eugene also wrote and starred in A Mighty Wind, Thank you. Did Caroline Goodall happen to be in the movie White Squall?”*

ランダム知識置換　正例応答のエンティティ（人名や作品名など）を，別のエンティティに置き換える．

- **正例応答例**: *“Thank you. Did Caroline Goodall happen to be in the movie White Squall?”*
- **置換例**: “Caroline Goodall” → “Tom Hanks”, “White Squall” → “Jurassic Park”
- **負例応答例**（置換後）: *“Thank you. Did Tom Hanks happen to be in the movie Jurassic Park?”*

ランダム Span 置換　正例応答中の任意の重要スパン（作品名・動詞句・形容詞句など）を，他の文から抽出したフレーズに差し替える．

- **正例応答例**: *“Yes! Eugene also wrote and starred in A Mighty Wind, have you heard of it?”*
- **置換対象スパン例**: *“wrote and starred in A Mighty Wind”* → *“was nominated for a Grammy”*
- **負例応答例**（置換後）: *“Yes! Eugene also was nominated for a Grammy, have you heard of it?”*

4.3 ベースライン

ベースラインとして，事前学習済みの言語モデル GPT2 [19] を用い，通常の対照学習（正例と負例のみ）でファインチューニングを行う．具体的には，式 (4) から $\mathcal{L}_{\text{error}}$ の項を除去した

$$\mathcal{J}_{\text{baseline}}(\theta) = \alpha_1 \mathcal{L}_{\text{Posi}} + \alpha_2 \mathcal{L}_{\text{Nega}}$$

を最小化する形で学習する．$\alpha_1 = 1, \alpha_2 = 0.5$ と設定した．対話履歴は最大 3 発話までに制限し，バッチサイズは 16，学習率は 3e-5，オプティマイザには AdamW を用いる．生成時は greedy サーチを使用した．

一方，提案手法（ours）では，式 (4) において $\alpha_1 = 1, \alpha_2 = 0.5, \alpha_3 = 0.8$ と設定し，正例・負例・誤りデータを同時に学習させる．これにより，誤りトークンをあえて強化してから抑制するステップを導入する点がベースラインと異なる．

4.4 評価指標

生成応答の品質を評価するため，F1 [20]，ROUGE-L [21]，BLEU [22]，NIST [23]，Meteor [24]（以下，MT と略），Knowledge-F1（KF1） [25]，および Entity-F1（EF1）を用いた．各指標の概要は以下のとおりである．

- **F1**：生成された応答と参照となる正解文の単語レベルにおける F1 スコアを算出し，両者の語彙的な重なりを評価する．
- **ROUGE-L（RL）**：生成応答と参照文との最長共通部分列に基づき，両者の構造的な類似度を測定する．
- **BLEU**：BLEU-2 および BLEU-4（B2, B4）を使用．n-gram 精度の幾何平均とペナルティ項に基づいて評価する．
- **MT (Meteor)**：単語単位の精度と再現率の調和平均に着目し，語形変化やシノニムなどもある程度考慮する．
- **Knowledge-F1 (KF1)**：生成応答とデータセット

表 1　各手法によって生成された応答文の評価結果.

Nega Data	Method	Test Seen						Test Unseen					
		F1	RL	B4	MT	KF1	EF1	F1	RL	B4	MT	KF1	EF1
ランダム置換文	baseline	21.29	22.16	2.40	20.49	11.33	9.63	20.95	21.56	2.15	20.07	10.15	7.56
	ours	21.65	22.46	2.55	20.02	11.98	10.79	21.17	21.94	2.20	19.59	10.60	8.12
ランダム知識置換	baseline	21.20	22.19	1.91	18.99	11.54	8.03	20.94	21.94	1.78	18.64	10.35	6.74
	ours	22.59	22.48	2.34	19.66	11.98	9.25	20.93	21.84	2.12	18.88	11.14	6.42
ランダム Span 置換	baseline	21.33	22.02	2.09	19.38	13.17	12.31	20.25	21.25	1.80	18.56	11.91	8.67
	ours	21.62	22.49	2.47	20.25	13.84	10.94	20.92	21.87	2.06	19.63	12.43	7.44

表 2　ランダム置換文負例を含むデータセットにおけるアブレーション実験の結果．Base model は提案手法を指す.

Method	F1	RL	B4	MT	KF1	EF1
Test Seen						
base model	21.65	22.46	2.55	20.02	11.98	10.79
-w/o $\mathcal{L}_{\text{error}}$	21.33	22.02	2.09	19.38	13.17	12.31
$\mathcal{L}_{\text{error}} \rightarrow \mathcal{L}_{\text{Posi}}$	21.14	22.07	2.35	19.95	13.16	12.97
$\mathcal{L}_{\text{error}} \rightarrow \mathcal{L}_{\text{Nega}}$	20.79	21.73	2.20	19.12	13.27	11.71
Test Unseen						
base model	21.17	21.94	2.20	19.59	10.60	8.12
-w/o $\mathcal{L}_{\text{error}}$	20.95	21.56	2.15	20.07	10.15	7.56
$\mathcal{L}_{\text{error}} \rightarrow \mathcal{L}_{\text{Posi}}$	20.09	21.09	1.80	18.72	12.01	8.98
$\mathcal{L}_{\text{error}} \rightarrow \mathcal{L}_{\text{Nega}}$	19.94	20.91	1.75	18.27	12.34	10.08

において注釈済みの正解知識文との一致度を F1 スコアで測定し，応答の情報性や関連性を評価する.

- **Entity-F1 (EF1)**：SpaCy[1] を用いて抽出したエンティティのみを対象に F1 スコアを算出し，知識の正確性を評価する.

4.5　結果と考察

全体的な比較（表 1）　表 1 は，OpenDialKG のテストデータ（Seen/Unseen）に対する各手法の評価結果を示したものである．まず，いずれの負例生成方法（ランダム置換文，ランダム知識置換，ランダム Span 置換）においても，提案手法（ours）はベースラインを上回る数値を示す傾向が見られた．特に，F1 や BLEU などの基本的な精度指標で向上が確認されることに加え，KF1 や EF1 においても多くの場合で数値が改善し，知識正確性の面でも有効性が示唆される.

また，Unseen データでの性能向上が一部で小さくなるケースはあるものの，Seen/Unseen いずれでも提案手法が一定の改善をもたらすことが分かる．これは，誤りを強化してから抑制する学習ステップが，モデルの汎化能力にも寄与していると考えられる.

アブレーションスタディ（表 2）　表 2 は，ランダム置換文データを用いたアブレーションスタディの結果である．ここで，*-w/o* $\mathcal{L}_{error}$ は $\mathcal{L}_{\text{error}}$ を除去した通常の対照学習の設定，$\mathcal{L}_{error} \rightarrow \mathcal{L}_{Posi}$ は本来誤りデータ（$\mathcal{L}_{\text{error}}$）で学習する箇所を単純に正例学習（$\mathcal{L}_{\text{Posi}}$）に置き換えた設定，$\mathcal{L}_{error} \rightarrow \mathcal{L}_{Nega}$ は誤りデータを負例学習（$\mathcal{L}_{\text{Nega}}$）に置き換えた設定である.

結果を見ると，提案手法（base model）がもっとも高い指標を示し，$\mathcal{L}_{\text{error}}$ を除去すると性能が低下することが分かる．また，誤りデータを正例あるいは負例として単に扱うだけでは，性能が十分に向上しないことも示唆される．これは，「誤りトークンをあえて正例扱い」した後に再度抑制する工程が一連の流れとして重要であり，その相互作用が精度向上につながっていると考えられる.

5　まとめ

本研究では，大規模言語モデルにおける誤生成を低減するため，負例応答を一度強化してから修正を行う対照学習手法を提案した．従来の対照学習に．あえて誤りトークンを「正例扱い」するステップ（誤りデータ学習）を追加することで，人間が「失敗を経験してから学ぶ」学習プロセスをモデルにも導入するアイデアである．OpenDialKG を用いた知識対話生成タスクの実験では，提案手法がベースラインに比べて多くの指標で改善を示し，誤りトークンを深く認識して抑制する有効性を確認した．将来的な展望としては，本研究のアイデアを強化学習的な枠組みや大規模事前学習へのフィードバックループに応用することで，対話システムのみならず，幅広い自然言語生成タスクでの性能向上が期待される.

1)　https://spacy.io/api/entityrecognizer/

謝辞

本研究の一部は，JSPS 科研費 JP23K20733 の支援を受けたものである．

参考文献

[1] Daniel De Freitas, Minh-Thang Luong, David R. So, Jamie Hall, Noah Fiedel, Romal Thoppilan, Zi Yang, Apoorv Kulshreshtha, Gaurav Nemade, Yifeng Lu, and Quoc V. Le. Towards a human-like open-domain chatbot, 2020.

[2] Minlie Huang, Xiaoyan Zhu, and Jianfeng Gao. Challenges in building intelligent open-domain dialog systems. **ACM TOIS**, Vol. 38, pp. 1 – 32, 2020.

[3] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In **International conference on machine learning**, pp. 1597–1607. PMLR, 2020.

[4] Sumit Chopra, Raia Hadsell, and Yann LeCun. Learning a similarity metric discriminatively, with application to face verification. In **CVPR**, pp. 539–546 vol. 1, 2005.

[5] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey E. Hinton. A simple framework for contrastive learning of visual representations. In **ICML**, pp. 1597–1607, 2020.

[6] Tianyu Gao, Xingcheng Yao, and Danqi Chen. Simcse: Simple contrastive learning of sentence embeddings. In **EMNLP**, pp. 6894–6910, 2021.

[7] Vladimir Karpukhin, Barlas Oğuz, Sewon Min, Patrick Lewis, Ledell Yu Wu, Sergey Edunov, Danqi Chen, and Wen tau Yih. Dense passage retrieval for open-domain question answering. In **NAACL**, pp. 5835–5847, 2021.

[8] Sean Welleck, Ilia Kulikov, Stephen Roller, Emily Dinan, Kyunghyun Cho, and Jason Weston. Neural text generation with unlikelihood training. In **ICLR**, 2020.

[9] Tianxing He and James R. Glass. Negative training for neural dialogue response generation. In **ACL**, pp. 2044–2058, 2020.

[10] Shaojie Jiang, Ruqing Zhang, Svitlana Vakulenko, and M. de Rijke. A simple contrastive learning objective for alleviating neural text degeneration, 2022.

[11] Xin Li, Piji Li, Yan Wang, Xiaojiang Liu, and Wai Lam. Enhancing dialogue generation via multi-level contrastive learning, 2020.

[12] Shuyang Cao and Lu Wang. Cliff: Contrastive learning for improving faithfulness and factuality in abstractive summarization, 2021.

[13] Tom Schaul, John Quan, Ioannis Antonoglou, and David Silver. Prioritized experience replay. In **International Conference on Learning Representations (ICLR)**, 2016.

[14] Yoshua Bengio, Jérôme Louradour, Ronan Collobert, and Jason Weston. Curriculum learning. In **Proceedings of the 26th Annual International Conference on Machine Learning (ICML)**, pp. 41–48, 2009.

[15] Marcin Andrychowicz, Filip Wolski, Alex Ray, Jonas Schneider, Rachel Fong, Peter Welinder, Bob McGrew, Josh Tobin, Pieter Abbeel, and Wojciech Zaremba. Hindsight experience replay. In **Advances in Neural Information Processing Systems (NeurIPS)**, pp. 5048–5058, 2017.

[16] Seungwhan Moon, Pararth Shah, Anuj Kumar, and Rajen Subba. OpenDialKG: Explainable conversational reasoning with attention-based walks over knowledge graphs. In Anna Korhonen, David Traum, and Lluís Màrquez, editors, **Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics**, pp. 845–854, Florence, Italy, July 2019. Association for Computational Linguistics.

[17] Lang Qin, Yao Zhang, Hongru Liang, Jun Wang, and Zhenglu Yang. Well begun is half done: Generator-agnostic knowledge pre-selection for knowledge-grounded dialogue. In Houda Bouamor, Juan Pino, and Kalika Bali, editors, **Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing**, pp. 4696–4709, Singapore, December 2023. Association for Computational Linguistics.

[18] Weiwei Sun, Zhengliang Shi, Shen Gao, Pengjie Ren, Maarten de Rijke, and Zhaochun Ren. Contrastive learning reduces hallucination in conversations, 2022.

[19] Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. Language models are unsupervised multitask learners. **OpenAI blog**, Vol. 1, No. 8, p. 9, 2019.

[20] Emily Dinan, Stephen Roller, Kurt Shuster, Angela Fan, Michael Auli, and Jason Weston. Wizard of wikipedia: Knowledge-powered conversational agents, 2019.

[21] Chin-Yew Lin. Rouge: A package for automatic evaluation of summaries. In **ACL**, 2004.

[22] Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. Bleu: a method for automatic evaluation of machine translation. In Pierre Isabelle, Eugene Charniak, and Dekang Lin, editors, **Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics**, pp. 311–318, Philadelphia, Pennsylvania, USA, July 2002. Association for Computational Linguistics.

[23] George Doddington. Automatic evaluation of machine translation quality using n-gram co-occurrence statistics. In **Proceedings of the Second International Conference on Human Language Technology Research**, HLT '02, p. 138–145, San Francisco, CA, USA, 2002. Morgan Kaufmann Publishers Inc.

[24] Michael Denkowski and Alon Lavie. Meteor universal: Language specific translation evaluation for any target language. In Ondřej Bojar, Christian Buck, Christian Federmann, Barry Haddow, Philipp Koehn, Christof Monz, Matt Post, and Lucia Specia, editors, **Proceedings of the Ninth Workshop on Statistical Machine Translation**, pp. 376–380, Baltimore, Maryland, USA, June 2014. Association for Computational Linguistics.

[25] Emily Dinan, Stephen Roller, Kurt Shuster, Angela Fan, Michael Auli, and Jason Weston. Wizard of wikipedia: Knowledge-powered conversational agents. In **ICLR**, 2019.