

拡散言語モデルを用いた知識参照の頑健化

木下貴翔¹ 三輪誠^{1,2}¹ 豊田工業大学 ² 産業技術総合研究所人工知能研究センター
{sd22039,makoto-miwa}@toyota-ti.ac.jp

概要

検索拡張生成 (RAG) において、検索ノイズへの頑健性や生成時に参照すべき外部知識の選択は喫緊の課題である。本研究では、検索ノイズへの頑健性の向上を目的に、拡散言語モデルの反復的な生成プロセスに着目し、逆拡散過程の中間状態に応じて外部知識の重要度を制御する「知識選択器」を提案する。HotpotQA を用いた実験の結果、動的な選択は確認できなかったものの、提案手法はベースラインを上回る回答精度を達成し、ノイズ文書を含む環境下での性能低下を抑えられることがわかった。

1 はじめに

外部データベースから関連情報を取得して生成に利用する検索拡張生成 (Retrieval-Augmented Generation; RAG) [1] は、モデルの知識不足を補い、ハルシネーションを抑制するための標準的な技術として広く利用されている。しかし、実環境における検索結果には、回答に不可欠な情報のほかに、無関係なノイズ (Distractor) が多分に含まれる。このため、信頼性のある回答をするためには、与えられた大量の外部知識の中から、真に必要な情報のみを正確に識別し、統合する能力が求められる。

一方、言語生成モデルとして、非自己回帰的な生成による高速な生成速度やノイズからの復元を通じた大域的な文脈制御能力から、拡散言語モデル [2, 3] が新たな生成パラダイムとして注目されている。しかし、拡散言語モデルに RAG の枠組みを適用する研究は十分になされておらず、外部知識をどのように効果的に参照させるかという方法論は十分に確立されていない。拡散モデルには、無秩序なノイズ状態から徐々に意味のある構造を形成していく「反復的な生成プロセス」があるため、生成の進行に伴って、モデルが求める情報の粒度や種類は動的に変化すると考えられる。しかし、既存の拡散言語モデルにおける外部知識参照の試みは、従来の言語モデル

と同様、外部情報を単に条件付け入力として静的に与える方法に留まっている。

そこで本研究では、拡散言語モデルに特化したノイズに頑健な RAG の実現に向けて、拡散言語モデルの生成プロセスと同期して外部知識を動的に取捨選択する「知識選択器」を提案する。提案手法は、逆拡散過程の中間表現を利用して、その時点での生成に必要な外部知識の重要度を動的に評価する。生成ステップごとの状態に応じて外部知識への注目度を制御することで、生成段階ごとに最適な知識を適応的に選択することが期待される。これにより、拡散モデルが本来持つ反復的な精緻化能力を知識参照にも拡張し、無関係な文書が混入する状況下においても、頑健な回答生成を実現することを目指す。

本研究の主な貢献は以下の通りである。

- **拡散言語モデルの生成特性を活かした動的知識参照手法の提案** 従来の静的な条件付けとは異なる、拡散モデル特有の逆拡散過程における反復的な生成プロセスに着目した、生成ステップごとの中間状態に応じて外部知識の重要度を動的に制御する「知識選択器」を提案した。
- **ノイズを含む検索結果に対する頑健性の向上** 知識選択器の導入により、回答に無関係なノイズとなる文書 (Distractor) が混入する環境下において、ベースラインの拡散言語モデルと比較して高い回答精度を達成し、提案手法のノイズに対する頑健性を示した。

2 関連研究

2.1 拡散言語モデル

拡散言語モデルは、生成空間の定義により連続拡散モデル [2, 4] と離散拡散モデル [3, 5] に大別される。従来は単語埋め込みのような連続空間上でガウシアンノイズを付加する連続拡散モデルが主流であったが、連続空間から離散的なトークン空間に写

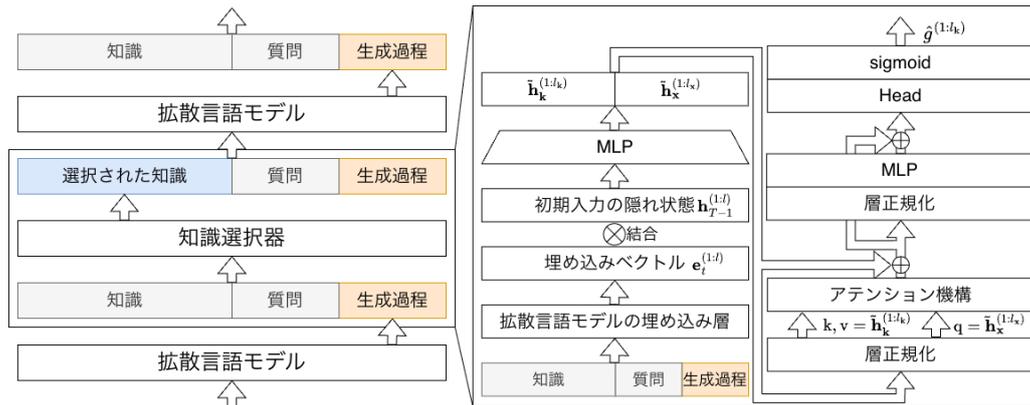


図1 知識選択器を用いた生成の構造。拡散言語モデルによる生成の前に、外部知識，質問，生成過程を知識選択器に入力して，生成に必要な外部知識だけを選択する。

像する際の丸め誤差などが問題となっていた。これに対し、離散拡散モデルは、トークン空間上で直接遷移を定義するため、テキストの離散性を自然に扱うことが可能である。特に、吸収状態（マスクトークン）への遷移を用いる手法は、事前学習済みのBERT [6] からの学習や、自己回帰型の言語モデルからの adaptation [7] が容易な点から、高い性能を実現している。本研究では、RAG において正確な知識の引用や統合が求められる点からも、テキストの離散的な操作に適した離散拡散モデルを採用する。

2.2 言語モデルにおける知識参照

大規模言語モデルの知識不足やハルシネーションを解消するために、外部知識を利用する検索拡張生成 (Retrieval-Augmented Generation; RAG) [1] が広く用いられている。一般的な RAG では、入力された質問に基づいて検索器を用いて文書集合を取得し、生成モデルにこれらの文書と質問を入力して回答を生成する。しかし、検索器は完全ではないため、取得された文書集合には回答に必要な文書 (Gold 文書) だけでなく、無関係なノイズ文書 (Distractor) が含まれることが多い。こうしたノイズ文書は言語モデルの推論を妨げ、回答性能を大きく低下させる。

この課題に対し、自己回帰型言語モデルにおいては、外部のネットワークで回答に必要な知識だけを選択させる手法 [8, 9] が提案されている。しかし、これらの既存手法は、外部知識を生成開始前に静的なコンテキストとして与えるに留まっている。

この解決のために、より動的なアプローチとして、検索と生成を交互に繰り返すことで、文脈の深まりに応じて必要な知識を適応的に参照し、回答精度を高めようとする試み [10, 11] も存在する。

しかし、これらの手法は依然として自己回帰的な生成を前提としており、拡散言語モデル特有の、ノイズ除去を通じた「反復的な生成プロセス」そのものと同期させた知識参照手法については、十分な検討がなされていない。

3 提案手法

本研究は「不完全な状態から徐々に完全な状態へと遷移する」という拡散モデルの生成プロセスに着目し、逆拡散過程の途中の生成状態に応じて、与えられた外部知識から必要なものだけを明示的に選択する手法を提案する。このために、外部知識の必要性を判断する知識選択器を既存の拡散言語モデルに追加する。この手法により、ノイズに頑健な、生成過程に応じた適応的な知識選択・参照を期待する。

3.1 知識選択器の構造

事前学習済みの拡散言語モデルに、Attention, MLP, LayerNorm 等で構成されるゲート機構を追加する (図 1)。具体的には、生成ステップ t における中間生成物 (ノイズ状態), 質問, 外部知識の埋め込みを入力とする。これらを結合し, MLP と Self-Attention (外部知識を Key/Value, その他を Query とする) を通して処理した後, sigmoid 関数により各知識トークンに対するゲート値 $\hat{g} \in (0, 1)$ を算出する。この \hat{g} を外部知識の埋め込みに乗じることで情報を明示的にフィルタリングし, 当該ステップの逆拡散サンプリングに用いる。

3.2 知識選択器の学習

言語モデルの生成において、いつどのような知識が必要になるかは自明ではなく、選択する対象をラ

表 1 既存手法と提案手法の比較

手法	#Gold/#Dist.	正解率(↑)	ROUGE-L(↑)
Dream-7B	2/0	0.6239	0.8171
	2/4	0.5821	0.7784
	2/8	0.5403	0.7571
Dream-7B +	2/0	0.7579	0.8620
	2/4	0.7080	0.8193
	提案手法	2/8	0.7034

ベルとして与えるような教師あり学習だけで知識選択器を学習することは難しい。そこで、本研究では、ラベルを用いた教師あり学習に加え、拡散言語モデルの目的関数 $\mathcal{L}_{\text{diffusion}}$ を導入する。この導入により、ラベルからの学習信号による粗い選択だけでなく、生成過程に応じた、より細かい知識選択が可能になることを期待する。知識選択器の目的関数には交差エントロピーを用いる。交差エントロピーは外部知識の必要か否かを表すラベル $g^{(1:k)} \in \{0, 1\}^{1 \times k}$ と、3.1 節で述べた知識選択器の出力 $\hat{g}^{(1:k)} \in (0, 1)^{1 \times k}$ に対して式 (1) で計算する。

$$\mathcal{L}_{\text{CE}} = -\frac{1}{k} \sum_{i=1}^k (g^{(i)} \log \hat{g}^{(i)} + (1-g^{(i)}) \log(1-\hat{g}^{(i)})) \quad (1)$$

ただし、 $g \in \{0, 1\}$ は教師ラベルであり、1 は必要な外部知識、0 は不要な外部知識に対応する。

最終的な目的関数は $\mathcal{L} = \mathcal{L}_{\text{diffusion}} + \lambda \mathcal{L}_{\text{CE}}$ とし、LoRA アダプタを用いて、モデル全体を最適化する。

4 実験と解析

4.1 HotpotQA による評価

4.1.1 実験設定

提案手法の評価には、回答に必要な Gold 文書 2 つと不要なノイズとなる Distractor 文書 8 つが付随した質問応答データセット HotpotQA [12] を用いた。提案手法の Distractor 文書に対する頑健性を示すために、Distractor 文書の数をも 0/4/8 個の場合で比較し、ROUGE-L [13] と正解率で評価した。

70 億パラメタの拡散言語モデルである Dream-7B [14] を用いて、HotpotQA で学習したものをベースラインとした。提案手法では、Dream-7B に、知識選択器を導入し、HotpotQA の質問応答と正解・不正解文書で学習する。いずれも同じエポック数だけ学習した。

4.1.2 実験結果および考察

表 1 より、ベースラインは Distractor 文書の数が増加するにつれて性能が著しく低下しているのに対し、提案手法は高い性能を維持していることがわかる。特に、Distractor 文書が 4 から 8 に倍増しても、提案手法の正解率の低下は 0.05 ポイント程度に留まっており、提案する知識選択器がノイズの量に依存せず、頑健に機能していることを示唆している。

また、Distractor 文書が 0 の設定においても、提案手法はベースラインと比較して高い正解率を示している。これは、拡散モデルの目的関数も含めて知識選択器を学習したことで、Gold 文書の中でも、特に回答に重要な部分の重みを大きく、それ以外を小さくしたからだと考えられる。

4.2 知識選択器の解析

4.2.1 実験設定

本解析では、知識選択器が、外部知識のラベルに対してどの程度正しく選択できているのか、生成過程を通して知識参照する範囲がどのように変化していくのかを調査した。

まず、知識選択器が、外部知識のラベルを正しく選択できているか、を定量的に評価するために、式 (2) に示す集合を定義した。

$$\begin{aligned} \Lambda_{\text{gold}} &= \{i \in \mathbb{N} | \mathbf{k}^{(i)} \in (\text{Gold Document Tokens})\} \\ \Lambda_{\text{positive}} &= \{i \in \mathbb{N} | \mathbf{k}^{(i)} \in (\text{Selected Tokens})\} \end{aligned} \quad (2)$$

ここで、(Gold Document Tokens) は回答に必要な文書の列であり、これに含まれるようなトークンのインデックスの集合が Λ_{gold} である。また、(Selected Tokens) は知識選択器が選択したトークンの列である。ただし、知識選択器は 0 から 1 のスカラー値 \hat{g} で選択の度合いを表現するので、0.5 を閾値として、 \hat{g} がこれ以上なら選択したものとし、これらの集合を用いて、式 (3) に示す「選択スコア」を定義した。

$$\begin{aligned} S_{\text{recall}} &= \frac{1}{\#\Lambda_{\text{gold}}} \sum_{i \in \Lambda_{\text{gold}}} \hat{g}^{(i)} \\ S_{\text{precision}} &= \frac{1}{\#\Lambda_{\text{positive}}} \sum_{i \in \Lambda_{\text{positive}}} g^{(i)} \\ S_{\text{F1}} &= \frac{2 \cdot S_{\text{recall}} \cdot S_{\text{precision}}}{S_{\text{recall}} + S_{\text{precision}}} \end{aligned} \quad (3)$$

必要な文書を網羅的に選択できるほど、 S_{recall} は大きくなる。また、 $S_{\text{precision}}$ は、知識選択器が、誤ら

表2 選択スコアの比較

手法	$S_{\text{recall}} (\uparrow)$	$S_{\text{precision}} (\uparrow)$	$S_{\text{F1}} (\uparrow)$
Dream-7B	0.0038	0.0012	0.0000
+ 提案手法	0.4833	0.0791	0.1749

ずに選択できているほど、高い値となる。そして、 S_{F1} は2つのスコアの総合したスコアとなる。

知識選択器を用いないベースライン手法では、 \hat{g} の代わりにアテンションスコアを利用した。本解析では、回答部のトークンの最後の Transformer ブロックのアテンションスコアを平均した値を用いた。具体的には、回答トークンのインデックス集合 $\Lambda_{\text{answer}} = \{i \in \mathbb{N} | \mathbf{x}^{(i)} \in (\text{Answer Tokens})\}$ に対して、式(4)のように定義した。

$$A^{(i)} = \max \left\{ \frac{1}{\#\Lambda_{\text{answer}}} \sum_{j \in \Lambda_{\text{answer}}} \alpha_{h,j,i} | h \in \{1, 2, \dots, \text{num_heads}\} \right\} \quad (4)$$

ただし、 $\alpha_{h,j,i}$ は h 番目のアテンションヘッ드의、 j 番目のトークンから i 番目のトークンへのアテンションスコアである。

さらに、生成過程を通して、選択する文書の範囲がどのように変化していくのかを定量的に解析するために、式(5)に示す選択エントロピーを定義する。

$$H = -\frac{1}{k} \sum_{i=1}^k (\hat{g}^{(i)} \log \hat{g}^{(i)} + (1 - \hat{g}^{(i)}) \log(1 - \hat{g}^{(i)})) \quad (5)$$

これは、より狭い範囲の文書を選択しているほど値が小さく、広い範囲を選択しているほど値が大きくなる。そして、選択エントロピーが生成過程を通してどのように変化しているかを定量的に測るために、各生成ステップにおける選択エントロピーの平均値を計算する。

いずれの解析においても Distractor を4つ用いた。

4.2.2 実験結果および考察

表2に選択スコア(式(3))を計算した結果を示す。表2において、提案手法はいずれのスコアで見ても、ベースラインより高い値を示し、これは Gold 文書を正しく特定できたことを示唆する。一方で、 S_{recall} に対して、 $S_{\text{precision}}$ の値が低いことから、表層的に類似した文書が含まれる場面に関しては誤った選択をしてしまっていると考えられる。

表3に、生成過程の初期($t=1$)、中期($t=0.5$)、終期($t=0$)における選択エントロピーの推移を示す。

結果より、ベースラインでは生成が進行するにつ

表3 選択エントロピーの比較

手法	$H(t=1)$	$H(t=0.5)$	$H(t=0)$
Dream-7B	0.0141	0.0101	0.0106
+ 提案手法	0.6916	0.6916	0.6916

れてエントロピーが減少しており、ノイズが除去され文脈が具体的になるにつれて、アテンションが特定のトークンへと局所化することがわかった。これは、モデルが参照すべき知識が生成過程で確定していくことを裏付ける。

一方、提案手法の選択エントロピーは、生成過程を通じてほぼ一定の値を推移しており、動的な知識選択は確認できなかった。この要因として、次の2点が考えられる。1点目は知識選択器が入力となる「質問文」の情報を支配的に利用している可能性である。拡散言語モデルでは、回答部分はノイズから徐々に生成されるが、条件付けとなる質問文は全ステップを通じて固定されている。提案手法の知識選択器が、ノイズの多い生成途中の回答状態よりも、確実な情報源である質問文を強く手がかりとして Gold 文書を特定している場合、ゲート値は初期段階($t=1$)から同じ値で維持されることになる。これの動的な挙動を許容している中で生成初期から安定して必要な知識を選択できているという点は提案手法の頑健性を裏付けている。2点目は、シーケンス長と生成過程の長さの比率に起因する構造的な要因である。また、本実験ではシーケンス長に対して生成過程で変化する部分が極端に短かったため、中間表現の変化が知識選択器のゲートの開閉に影響するほどの変化ではなかった可能性がある。

5 おわりに

本研究では、知識参照におけるノイズと既存の拡散言語モデルの知識参照手法が十分に拡散言語モデルの性質を利用できていないことによる問題を指摘し、ノイズに頑健で適応的な知識参照の有効性の検証を目的として挙げた。そして、生成ステップごとの中間状態に応じて外部知識の重要度を明示的に動的に制御する「知識選択器」を提案した。HotpotQA を用いた実験において、ベースラインの拡散言語モデルと比較して提案手法のノイズに対する頑健性を示した。一方で、知識選択は静的な挙動を示した。今後は、長文生成など中間状態の変化が大きいと考えられるタスクにおける動的挙動の検証を行う。

謝辞

この成果の一部は、NEDO（国立研究開発法人新エネルギー・産業技術総合開発機構）の委託業務（JPNP25006）の結果得られたものです。また、産総研及び AIST Solutions が提供する ABCI 3.0 を「ABCI 3.0 開発加速利用」の支援を受けて利用しました。本研究に関して有益なご議論とご助言をいただきました産業技術総合研究所の浅田真生氏に深く感謝いたします。

参考文献

- [1] Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Kuttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, Sebastian Riedel, and Douwe Kiela. Retrieval-augmented generation for knowledge-intensive nlp tasks. In **Proceedings of the 34th International Conference on Neural Information Processing Systems**, 2020.
- [2] Xiang Lisa Li, John Thickstun, Ishaan Gulrajani, Percy Liang, and Tatsunori B. Hashimoto. Diffusion-lm improves controllable text generation. In **Proceedings of the 36th International Conference on Neural Information Processing Systems**, 2022.
- [3] Aaron Lou, Chenlin Meng, and Stefano Ermon. Discrete diffusion modeling by estimating the ratios of the data distribution. In **Proceedings of the 41st International Conference on Machine Learning**, 2024.
- [4] Shansan Gong, Mukai Li, Jiangtao Feng, Zhiyong Wu, and Lingpeng Kong. DiffuSeq: Sequence to sequence text generation with diffusion models. In **The 11th International Conference on Learning Representations**, 2023.
- [5] Subham Sekhar Sahoo, Marianne Arriola, Yair Schiff, Aaron Gokaslan, Edgar Marroquin, Justin T Chiu, Alexander Rush, and Volodymyr Kuleshov. Simple and effective masked diffusion language models. In **Proceedings of the 38th International Conference on Neural Information Processing Systems**, 2024.
- [6] 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: Human Language Technologies, Volume 1 (Long and Short Papers)**, 2019.
- [7] Shansan Gong, Shivam Agarwal, Yizhe Zhang, Jiacheng Ye, Lin Zheng, Mukai Li, Chenxin An, Peilin Zhao, Wei Bi, Jiawei Han, Hao Peng, and Lingpeng Kong. Scaling diffusion language models via adaptation from autoregressive models. In **The 13th International Conference on Learning Representations**, 2025.
- [8] Derong Xu, Xinhang Li, Ziheng Zhang, Zhenxi Lin, Zhihong Zhu, Zhi Zheng, Xian Wu, Xiangyu Zhao, Tong Xu, and Enhong Chen. Harnessing large language models for knowledge graph question answering via adaptive multi-aspect retrieval-augmentation. In **Proceedings of the AACL Conference on Artificial Intelligence**, 2025.
- [9] Tu Ao, Yanhua Yu, Yuling Wang, Yang Deng, Zirui Guo, Liang Pang, Pinghui Wang, Tat-Seng Chua, Xiao Zhang, and Zhen Cai. Lightprof: A lightweight reasoning framework for large language model on knowledge graph. In **Proceedings of the AACL Conference on Artificial Intelligence**, 2025.
- [10] Zhenrui Yue, Honglei Zhuang, Aijun Bai, Kai Hui, Rolf Jagerman, Hansi Zeng, Zhen Qin, Dong Wang, Xuanhui Wang, and Michael Bendersky. Inference scaling for long-context retrieval augmented generation. In **The 13th International Conference on Learning Representations**, 2025.
- [11] Liang Wang, Haonan Chen, Nan Yang, Xiaolong Huang, Zhicheng Dou, and Furu Wei. Chain-of-retrieval augmented generation. In **Proceedings of the 39th International Conference on Neural Information Processing Systems**, 2025.
- [12] Zhilin Yang, Peng Qi, Saizheng Zhang, Yoshua Bengio, William W. Cohen, Ruslan Salakhutdinov, and Christopher D. Manning. HotpotQA: A dataset for diverse, explainable multi-hop question answering. In **Conference on Empirical Methods in Natural Language Processing**, 2018.
- [13] Chin-Yew Lin. ROUGE: A package for automatic evaluation of summaries. In **Text Summarization Branches Out**, 2004.
- [14] Jiacheng Ye, Zhihui Xie, Lin Zheng, Jiahui Gao, Zirui Wu, Xin Jiang, Zhenguo Li, and Lingpeng Kong. Dream 7b: Diffusion large language models, 2025.

A 実験の詳細

4 節における実験では AdamW でモデルを最適化した。表 4 に 옵ティマイザの設定を示す。また、学習率は全学習ステップの内、初めの 1% のステップで表 4 に示す値まで線形に上昇させ、残りのステップで 0 まで線形に減少させた。

表 4 オプティマイザの設定

設定	値
拡散言語モデルの学習率	1e-5
知識選択器の学習率	1e-4
β_1	0.9
β_2	0.999
weight decay	1e-2

学習に用いたデータは HotpotQA の学習データの内、16,000 件のみを用い、バッチサイズは 128、最大シーケンス長は 4,096 トークンとした。また、評価用データ 7,400 件の内、1,048 件を評価に、残りをテストに用いた。

ノイズスケジュールには $\alpha_t = 1 - t$ を用いた。また、目的関数については $\lambda = 2.0$ とした。

学習に用いた LoRA の設定を表 5 に示す。

表 5 LoRA の設定

設定	値
対象パラメタ	q_proj, k_proj, v_proj, o_proj, up_proj, down_proj
ランク	8
α	16