

魅力的な対話応答生成のための複数教師による知識蒸留

Ritvik Choudhary 河原大輔
早稲田大学理工学術院
{ritvik@fuji.,dkw@}waseda.jp

概要

オープンドメイン対話応答生成は、未だに人間の精度には及ばないタスクの一つである。近年の生成モデルの発展にもかかわらず、モデルは知識を活用することができず、一般的で無味乾燥な応答が多くみられる。先行研究では、感情や知識ベースなどの外部文脈をモデルに付加することでこの問題に対処している。本論文では、対話応答生成にモデルのパラメータ内知識を利用した新たなアプローチを提案する。魅力的な対話応答を生成するための2つの能力を定義し、それぞれを持つ複数の教師モデルを構築する。教師モデルから関連する知識を蒸留することによって、より魅力的な対話応答生成が可能な生徒モデルを学習する。学習された生徒モデルの自動及び人手評価を行った結果、提案手法の有効性を確認した。

1 はじめに

人工知能が日常生活に普及するにつれ、単純な受け答えを超えた魅力的な対話ができる対話システムの実現が期待されている。オープンドメイン対話応答生成タスクは、そのような対話を生み出す可能性を秘めており、これまでにも多くの研究がなされている [1, 2]。特に最近では、数千億パラメータに及ぶ超大規模言語モデル (LLM) の登場により、大きな進展がある [3]。しかし、LLM はサイズが大きく、アクセスも制限されているため、さらなる学習やスケラビリティに問題がある [4]。

そのため、本論文ではより広く使われている中規模モデル [5] に焦点を当てるが、これらのモデルは生成された応答において魅力的でないという問題がある [6]。根本的な原因の一つは、学習中にモデルが関連する知識を持たないことにあり、その結果、無味乾燥な応答が多くなってしまふことになる。

この問題に対処するため、先行研究ではペルソナ

情報の追加 [7]、外部知識ベースの利用 [8]、また情報検索の応用 [9] が検討されている。これらの先行研究の多くは、生成モデルに知識を補うことで応答の品質を高めることに焦点を当てている。一方、モデルのパラメータ内知識を直接活用する手法の一つとして知識蒸留 [10] があり、対話生成にも適用されている [11] が多くない。本研究は、上記の対話応答生成の問題を解決するために、知識蒸留に基づく新たなアプローチを試みる。人間の学び方に倣い、魅力的な対話応答生成というタスクに必要な能力を複数の教師モデルに分割し、各教師のパラメータに保持されている関連知識を蒸留しながら、生徒モデルに応答生成を学習させることを目指す。

魅力的な対話応答を生成するには、以下の2つの能力が必要であると考えられる。(1) 会話能力 (conversational ability, 対話応答の発話に対する関連性)、(2) 情報提供能力 (informativeness, または知識性)。まず、それぞれの能力に対応する2つの教師モデルを学習する。次に、2つの教師モデルから知識蒸留を行い、応答生成のタスクにおいて生徒モデルを学習する。このアプローチにより、生徒モデルは、より魅力的な対話応答生成に必要な複数の能力を効果的に学習することができる。オープンドメインの対話データセットを用いた実験の結果、様々なベースラインを上回り、提案手法の有効性が示された。

2 関連研究

知識に基づく応答生成モデル 自然言語生成において、知識に基づく seq2seq モデルの分野では様々な研究が行われている。近年の研究ではペルソナや共感情報を考慮する手法 [12] や、より適切で魅力的な応答を生成するために、知識に基づくオープンドメイン対話生成 [8] を発展させた手法もある。また、外部の知識ベースを照会し、応答を生成するように情報検索とテキスト生成を統合したアプローチもある [13]。外部知識ベースを取り込むことを意図した

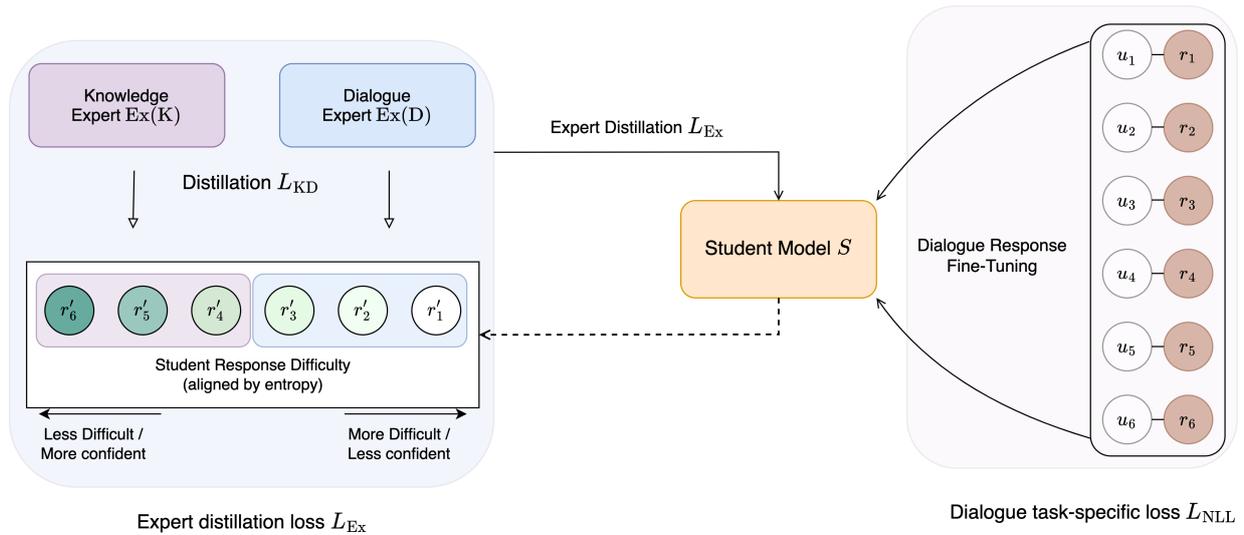


図1 提案手法の概要

既存の研究 [8] とは異なり、本研究はモデルのパラメータ内に蓄積された知識を利用することに焦点を置く。この点に関する先行研究には、言語モデルの内部知識の更新に注目しているもの [14] はあるが、提案手法は対話応答生成のタスクに内部知識を利用することを目的としている。

知識蒸留と対話応答生成 知識蒸留 (Knowledge Distillation, KD) [10] は、当初コンピュータビジョンの分野で提案されたが [15]、最近では様々な言語処理タスクで利用されている [16]。しかし、対話応答生成における知識蒸留の利用は比較的少ない。近年では、知識蒸留を用いたデータ拡張 [11] の他、対話生成のためにカリキュラム学習と KD を組み合わせるアプローチ [17] が提案されている。本研究は、このアプローチ [17] と同様の分野ではあるが、より魅力的で情報量の多い応答を生成するために、複数の専門教師の学習と蒸留に焦点を当てている点異なる。

3 提案手法

本節では、オープンドメインの対話応答生成において、複数の教師モデルから効果的に知識蒸留を行い、生徒モデルの生成結果の魅力度を向上させるための手法について述べる。まず、対話応答生成タスクを定式化し、次に提案手法を述べる。

3.1 対話応答生成

対話応答生成のタスクは、入力発話 \mathbf{u} が与えられたときの応答 $\mathbf{r} = (r_1, r_2, \dots, r_m)$ を予測する。入力発話 \mathbf{u} が与えられたときに、モデルが確率

$p(r_i | r_{<i}, \mathbf{u}; \theta)$ に従って応答 \mathbf{r} を生成できるように、モデルのパラメータ θ を学習する。モデルパラメータは次の損失関数、すなわち負対数尤度 (NLL) を最小化することで学習できる。

$$\mathcal{L}_{\text{NLL}}(\theta) = - \sum_i \log p(r_i | r_{<i}, \mathbf{u}) \quad (1)$$

3.2 複数教師による知識蒸留を用いた応答生成

本研究では、魅力的な応答生成に必要な2つの能力に着目する。会話能力と情報提供能力である。まず、オープンドメインと知識関連の対話データに対して、それぞれ $\text{Ex}(\text{D})$ と $\text{Ex}(\text{K})$ という2つのモデルを、3.1節の NLL loss に従って学習する。これらのモデルは知識蒸留 (KD) を通して、生徒モデルの教師として機能する。図1に提案手法の概要を示す。

KD [10] は、ある教師と生徒モデルが与えられたとき、それらの出力確率分布間の Kullback-Leibler (KL) 距離を最小にすることによって、生徒の確率分布を教師に近づける手法である。生徒 S と教師 T の予測分布 (\mathbf{r}_s と \mathbf{r}_t とする) に対して、KD は次の距離を最小化する。

$$\mathcal{L}(\text{T})_{\text{KD}} = \text{KL}(\mathbf{r}_s \parallel \mathbf{r}_t) \quad (2)$$

複数の教師モデルが存在する場合、生徒には各教師からの得る知識を最適化することが重要になる。また、知識蒸留の先行研究 [18] では、生徒と教師モデルの出力確率分布の差が大きい場合、生徒モデルの学習に悪影響を及ぼすという、容量ギャップの問題が指摘されている。

そこで、カリキュラム学習やアクティブラーニングにおける生徒中心の学習方法 [19] にならい、学習中に教師モデルの蒸留信号を効果的に調整するための指標として、生徒による応答の難易度を考慮する。先行研究 [19] に従い、本研究では生成された応答に対する生徒モデルのエントロピー H を難易度の指標とみなす。

$$H = - \sum_i p(r_i | \mathbf{r}_{<i}, \mathbf{u}) \log p(r_i | \mathbf{r}_{<i}, \mathbf{u}) \quad (3)$$

難易度を生徒モデルの自信としても考えることもできる。

上記の指標に基づき、バッチ内の応答を簡単なものから難しいものへと並び替え、各教師からの蒸留信号はその順序に従って分割する。例えば、難しい応答の場合、生徒 S は学習しやすい教師の指導を受け、もう一方の教師から、より簡単な応答について指導を受けることになる。よって、教師モデルから蒸留信号の分割を以下のように定式化する。

$$\mathcal{L}_{\text{Ex}} = \gamma \mathcal{L}(\text{Ex(D)})_{KD} + (1 - \gamma) \mathcal{L}(\text{Ex(K)})_{KD} \quad (4)$$

ここで、 γ は教師モデル間のバランスをとるためのパラメータとする。

式 (1) と (4) を組み合わせることで、生徒モデルの対話応答生成タスクで最適化すべき全体の損失関数 \mathcal{L} を以下のように定義する。

$$\mathcal{L} = \mathcal{L}_{\text{NLL}} + \lambda \mathcal{L}_{\text{Ex}} \quad (5)$$

4 実験

本研究では、多様なオープンドメインの対話データセットを用いて提案手法を評価する。様々なベースラインモデルの組み合わせと比較し、本研究のアプローチの有効性を確認する。

4.1 実験設定

本研究では様々なオープンドメインと知識的対話で fine-tuning された BART ベースの seq2seq モデルを教師と生徒モデルとして採用する。モデルの学習には Adam optimizer を使い、学習率 $3e-5$ のデフォルトパラメータで学習する。実装には Hugging Face ライブラリを使用する。

4.1.1 教師モデル

3.2 節で述べたように、会話能力を表す Ex(D) と、情報提供能力を表す Ex(K) という 2 つの教師モデ

表 1 蒸留に用いる教師モデル

モデル	Arch (BART)	学習データ	合計ターン数
Ex(D)	Large	ODD	200,000
Ex(K)	Base	WikiDialog	500,000
Ex(K _R)	Base	Reddit	500,000

ルを学習する。Ex(D) はオープンドメインの対話データセットで学習した BART-large モデルであり、Ex(K) は情報や知識が多い疑似対話データセットで学習した BART-base¹⁾ モデルである。Ex(D) の単純な日常会話に比べ、Ex(K) はより情報量の多い (多様な) データセットで学習される。そのため、知識の教師である Ex(K) は Ex(D) よりも難しいタスクを学習した教師であると想定される。すなわち、生徒にとって、Ex(K) が Ex(D) よりも学習が困難な教師である。なお、提案手法では学習時、生徒による学習サンプルの難易度に基づいて教師を選択する。したがって、学習の過程では生徒にとって難易度の高いサンプルは、学習しやすい教師 Ex(D) の信号を受け、その逆も同様である。本研究では、 $\gamma = 0.5$ に設定し、教師モデルは均等にバランスされている。

4.1.2 データセット

本研究では様々なオープンドメインの対話と知識に基づく対話データセットを用いて、教師と生徒モデルの fine-tuning を行う。

Open-domain dialogue datasets (ODD) 高品質な日常対話データセットとして DailyDialog [20]、DailyDialog++ [21]、Cornell Movie-Dialogs Corpus [22] の 3 種類をマージして教師 Ex(D) と生徒モデルの学習に用いる。各モデルの学習サンプル数を 200,000 に固定する。

知識疑似対話データセット 知識教師 Ex(K) の学習では、Wikipedia から抽出した疑似対話データセットである WikiDialog [23] を用いる。さらに、Reddit から抽出した対話データセット [13] を用いて別の教師 Ex(K_R) を学習し、比較する。

なお、知識の対話データがより複雑であるため、教師モデルの学習サンプル数を 500,000 に設定する。

1) BART-base は、Ex(K) について比較的安定した学習ができることから選択。

表 2 生成された応答の自動評価結果。ODD は 4.1.2 節の Open-Domain Dataset の集合である。

モデル	Arch	蒸留	学習	自動評価	
				データ	BLEU-4 PPL
Vanilla	Base	-	ODD	1.56	190.4
Ex(D) Distill	Base	Ex(D)	ODD	1.64	33.55
Ex(K) Distill	Base	Ex(K)	ODD	1.85	44.67
ME-Ex(K _R)	Base	Ex(K _R) + Ex(D)	ODD	1.71	30.60
提案手法	Base	Ex(K) + Ex(D)	ODD	1.88	35.37

表 3 各評価観点において提案手法の勝率 (%)

比較相手	評価観点	
	Engagement	Informativeness
Vanilla	83	67
Ex(D) Distill	79	78
Ex(K) Distill	75	84
ME-(w/ Ex(K _R))	79	84

4.2 評価

提案手法の有効性を確認するため、ODD テストセットで生成された応答について自動評価と人手評価を行う。なお、教師モデルによって学習条件が異なるため、比較対象は同規模のベースラインに限定する。

ベースライン 4つの異なるベースラインを学習し、比較する。まず、4.1.2 節で述べた ODD で直接学習したバニラ BART-base モデル (以下「バニラ」) は提案手法の同サイズの生徒モデルと直接比較する。次に、複数教師の効果を確認するため、どちらか一つの教師モデルのみで学習したベースライン (標準的な知識蒸留に相当、以下「Ex(D)-Distill」と「Ex(K)-Distill」) を比較する。最後に、Wikipedia ベースのモデルを知識教師として使用する効果を知るため、Reddit ベース (4.1.2 節) の教師モデル Ex(K_R) を学習し、Ex(K) と入れ替えた組み合わせ (「Multi Expert (w/ Ex(K_R))」) 以下、「ME-Ex(K_R)」) を提案手法と比較する。

自動評価 生成された応答に対して自動評価を行う。先行研究に従い、BLEU とパープレキシティ (PPL) [24] のスコアを算出する。

人手評価 自動評価指標は、生成された対話応答の質とあまり相関がないことが先行研究で報告されている [25]。そこで、ACUTE-Eval の手法 [2] に従い、提案手法とベースラインモデルの応答をペア

ごとに人手評価を行う。評価は、Amazon Mechanical Turk を介して 2 つの指標で行う。一つは会話能力 (Engagement) であり、その応答をした相手とさらに会話を継続したいかどうかを表す。もう一つは情報提供能力 (Informativeness) であり、その応答が知識に富んでいるかどうかを表す。評価者となるクラウドワーカーには、テストデータから無作為に選んだ 100 個の応答を上記の各指標に基づいて比較するように求める。各応答は 7 人の評価者による多数決で勝敗を決定し、提案手法の勝率を算出した。

4.3 結果

評価結果を表 2、表 3 にまとめる。BLEU では提案手法が各ベースラインを上回ることが確認された。さらに、提案手法の PPL も低い値となっている。なお、提案手法の PPL が最小値ではない原因は、生成された応答が多様であるためだと考えられる。しかし、前述したように、n-gram からなる指標は応答の品質を正確に表現していないことが示されている [25]。そのため、生成された応答の人手評価も実施した。人手評価では、提案手法によって全てのベースラインに対して双方の指標が改善した。さらに、単一教師 (通常の KD) のベースラインより提案手法が全体として勝率が高いため、会話能力と情報提供能力の双方がより魅力的な回答を生成する上で重要な役割を果たすことも示された。提案手法とベースラインによる生成応答の例を付録 A に示す。

5 おわりに

本論文では、モデルのパラメータに含まれる知識を学習、活用することで、より魅力的な対話応答生成のためのアプローチを提案した。具体的には、2 つの異なるドメインの対話について学習した複数の教師モデルから知識蒸留を行う手法を提案した。これにより、対象となる生徒モデルは両方の教師ドメインから学習することができ、自動評価と人手評価の双方で、より魅力的で情報量の多い対話応答生成ができることが確認された。

謝辞

本研究は LINE 株式会社との共同研究の助成を受けて行った。

参考文献

- [1] Apoorv Kulshreshtha, Daniel De Freitas Adiwardana, David Richard So, Gaurav Nemade, Jamie Hall, Noah Fiedel, Quoc V. Le, Romal Thoppilan, Thang Luong, Yifeng Lu, and Zi Yang. Towards a human-like open-domain chatbot. In *arXiv*. 2020. abs/2001.09977.
- [2] Stephen Roller, Emily Dinan, Naman Goyal, Da Ju, Mary Williamson, Yinhan Liu, Jing Xu, Myle Ott, Eric Michael Smith, Y-Lan Boureau, and Jason Weston. Recipes for building an open-domain chatbot. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, pp. 300–325, Online, April 2021. Association for Computational Linguistics.
- [3] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. Language models are few-shot learners. *Advances in neural information processing systems*, Vol. 33, pp. 1877–1901, 2020.
- [4] Conglong Li, Minjia Zhang, and Yuxiong He. The stability-efficiency dilemma: Investigating sequence length warmup for training gpt models. In *Advances in Neural Information Processing Systems*, 2022.
- [5] Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Veselin Stoyanov, and Luke Zettlemoyer. BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pp. 7871–7880, Online, July 2020. Association for Computational Linguistics.
- [6] Jiwei Li, Michel Galley, Chris Brockett, Jianfeng Gao, and Bill Dolan. A diversity-promoting objective function for neural conversation models. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pp. 110–119, San Diego, California, June 2016. Association for Computational Linguistics.
- [7] Peixiang Zhong, Chen Zhang, Hao Wang, Yong Liu, and Chunyan Miao. Towards persona-based empathetic conversational models. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 6556–6566, Online, November 2020. Association for Computational Linguistics.
- [8] Emily Dinan, Stephen Roller, Kurt Shuster, Angela Fan, Michael Auli, and Jason Weston. Wizard of wikipedia: Knowledge-powered conversational agents. In *International Conference on Learning Representations*, 2019.
- [9] Kurt Shuster, Spencer Poff, Moya Chen, Douwe Kiela, and Jason Weston. Retrieval augmentation reduces hallucination in conversation. In *Findings of the Association for Computational Linguistics: EMNLP 2021*, pp. 3784–3803, Punta Cana, Dominican Republic, November 2021. Association for Computational Linguistics.
- [10] Geoffrey Hinton, Oriol Vinyals, and Jeffrey Dean. Distilling the knowledge in a neural network. In *NIPS Deep Learning and Representation Learning Workshop*, 2015. abs/1503.02531.
- [11] Rongsheng Zhang, Yinhe Zheng, Jianzhi Shao, Xiaoxi Mao, Yadong Xi, and Minlie Huang. Dialogue distillation: Open-domain dialogue augmentation using unpaired data. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 3449–3460, Online, November 2020. Association for Computational Linguistics.
- [12] Tatsuya Ide and Daisuke Kawahara. Multi-task learning of generation and classification for emotion-aware dialogue response generation. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Student Research Workshop*, pp. 119–125, Online, June 2021. Association for Computational Linguistics.
- [13] Ritvik Choudhary and Daisuke Kawahara. Grounding in social media: An approach to building a chat-dialogue model. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Student Research Workshop*, pp. 9–15, Hybrid: Seattle, Washington + Online, July 2022. Association for Computational Linguistics.
- [14] Daliang Li, Ankit Singh Rawat, Manzil Zaheer, Xin Wang, Michal Lukasik, Andreas Veit, Felix Yu, and Sanjiv Kumar. Large language models with controllable working memory. *arXiv preprint arXiv:2211.05110*, 2022. abs/2211.05110.
- [15] Yuang Liu, Wei Zhang, and Jun Wang. Adaptive multi-teacher multi-level knowledge distillation. *Neurocomputing*, Vol. 415, pp. 106–113, 2020.
- [16] Victor Sanh, Lysandre Debut, Julien Chaumond, and Thomas Wolf. DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. In *5th Workshop on Energy Efficient Machine Learning and Cognitive Computing @ NeurIPS 2019*, 2019.
- [17] Qingqing Zhu, Xiuying Chen, Pengfei Wu, JunFei Liu, and Dongyan Zhao. Combining curriculum learning and knowledge distillation for dialogue generation. In *Findings of the Association for Computational Linguistics: EMNLP 2021*, pp. 1284–1295, 2021.
- [18] Aref Jafari, Mehdi Rezagholizadeh, Pranav Sharma, and Ali Ghodsi. Annealing knowledge distillation. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, pp. 2493–2504, Online, April 2021. Association for Computational Linguistics.
- [19] Burr Settles and Mark Craven. An analysis of active learning strategies for sequence labeling tasks. In *proceedings of the 2008 conference on empirical methods in natural language processing*, pp. 1070–1079, 2008.
- [20] Yanran Li, Hui Su, Xiaoyu Shen, Wenjie Li, Ziqiang Cao, and Shuzi Niu. DailyDialog: A manually labelled multi-turn dialogue dataset. In *Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pp. 986–995, Taipei, Taiwan, November 2017. Asian Federation of Natural Language Processing.
- [21] Ananya B Sai, Akash Kumar Mohankumar, Siddhartha Arora, and Mitesh M Khapra. Improving dialog evaluation with a multi-reference adversarial dataset and large scale pretraining. *Transactions of the Association for Computational Linguistics*, Vol. 8, pp. 810–827, 2020.
- [22] Cristian Danescu-Niculescu-Mizil and Lillian Lee. Chameleons in imagined conversations: A new approach to understanding coordination of linguistic style in dialogs. In *Proceedings of the 2nd Workshop on Cognitive Modeling and Computational Linguistics*, pp. 76–87, Portland, Oregon, USA, June 2011. Association for Computational Linguistics.
- [23] Zhuyun Dai, Arun Tejasvi Chaganty, Vincent Y Zhao, Aida Amini, Qazi Mamunur Rashid, Mike Green, and Kelvin Guu. Dialog inpainting: Turning documents into dialogs. In *International Conference on Machine Learning*, pp. 4558–4586. PMLR, 2022.
- [24] Gábor Melis, Chris Dyer, and Phil Blunsom. On the state of the art of evaluation in neural language models. In *International Conference on Learning Representations*, 2018.
- [25] Chia-Wei Liu, Ryan Lowe, Iulian Serban, Mike Noseworthy, Laurent Charlin, and Joelle Pineau. How NOT to evaluate your dialogue system: An empirical study of unsupervised evaluation metrics for dialogue response generation. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pp. 2122–2132, Austin, Texas, November 2016. Association for Computational Linguistics.

表 4 各モデルにおける応答生成の例

発話	応答 (提案手法)	応答 (ME-Ex(K _R))	応答 (バニラベースライン)
Removal Point? What're you talking about!	It's Removal Point. That's where they put the boat.	I'm not talking about anything. I'm talking	I'm not talking I'm not talking
I know why you broke up with me in college. Can you forgive me?	I know, but I cannot forgive you.	I don't forgive you.	I don't know what you are talking about.
I live in Haitian District where there are lot of universities .	That's great! What kind of universities do you like in Haiti?	Congrats!	Do you have a good education?
Shes very striking, don't you think? Are these glasses a fashion where she comes from?	No, I don't think so. Its probably just fashion.	No.	I don't know. I don't think so.

A 対話応答生成結果の例

主なベースラインと提案手法による発話に対する生成応答の例を表 4 に示す。