潜在変数の投機的サンプリングに基づく多様な雑談応答生成

佐藤 翔悦 [*1]
[*1] 東京大学大学院 情報理工学系研究科
喜連川 優 [*2,3]
[*2] 東京大学 生産技術研究所
[*3] 国立情報学研究所
{shoetsu, kitsure}@tkl.iis.u-tokyo.ac.jp

1 はじめに

会話データからの学習に基づく対話モデルは人手によるパターンの作り込みを必要とせず，比較的低コストで対話システムの構築が可能であるという点から，近年の対話応答タスクにおける主流なアプローチである．しかしそうした対話モデルでは，発話のオウム返しや典型的なパターン（例: “私もそう思います”) など，“safe response” と呼ばれる応答が頻出し，多様性の低さが問題となっている [1, 2, 3].

その原因の一つに，対話データとモデルの性質の乖離がある．多くの場合，対話においてある発話に対する適切な応答は一意ではない．一方で，近年の応答生成において標準的に用いられる Encoder-Decoder モデルでは，対話タスクを発話から応答への一対一の写像とみなして解くため，対話データに対するモデルの最適化は困難なものとなる．結果，学習データ中で頻出する単語やパターンを単純に模倣することが目的関数を最適化するための有効な手段となり，safe response が頻出する．

この問題に対し，条件付き変分オートエンコーダ (Conditioned Variational Autoencoder, CVAE) に基づく対話モデルは有望な手法の一つである [4, 5]．対話タスクにおける CVAE は，発話・応答から推定した事前・事後分布よりサンプルした潜在変数を用いて応答の生成を行う．そのため，潜在変数の違いによって同じ発話に対し複数の応答が成立する，対話データの多対多性を考慮した手法である．

しかしながら，こうしたモデルを用いたとしても期待通りの効果が発揮されず，応答の多様性について依然大きな改善が得られない場合がある [6]．本研究ではその一因が学習時に用いる潜在変数のサンプリングにあると考え，投機的サンプリングという極めてシンプルな提案手法によりその解決を図る．具体的には学習時に複数の潜在変数をサンプルし，それぞれについて損失を計算し，最も損失が小さい潜在変数を最適化に用いる．

実験では，Twitter から収集した対話データを用いた応答生成により提案手法を評価する．生成された応答に対して自動評価および人手評価を行い，提案手法の有効性を確認した．

2 提案手法

2.1 事前知識: T-CVAE

提案手法の説明の前に，本研究におけるベースラインとして採用した，Transformer-based CVAE (T-CVAE) [7] の概要を述べる．T-CVAE は Variational Hierarchical Recurrent Encoder-Decoder (VHRED) [8] などのモデルと同様，CVAE に基づくモデルである．従来のモデルがエンコーダおよびデコーダとして再帰型ニューラルネットワークを用いていたのに対し，T-CVAE は Transformer [9] を採用している．

対話タスクにおいてある発話 x および応答 y が与えられた際，T-CVAE は発話から事前分布 $p(z \mid x) \sim N(\mu, \sigma^2 \mathrm{I})$ を，発話と応答から事後分布 $q(z \mid x, y) \sim N(\mu, \sigma^2 \mathrm{I})$ を推定する．その後，分布からサンプルした潜在変数 z を用いて生成を行う．具体的な分布の推定としては発話・応答をエンコードしたベクトルによって，以下のように分布の平均 μ, 偏差 σ を計算する．

$$h_p = \mathrm{Attn}_p\left(c, E^L_{\mathrm{out}}(x), E^L_{\mathrm{out}}(x)\right) \tag{1}$$

$$\begin{bmatrix} \mu_p \\ \log\left(\sigma_p^2\right) \end{bmatrix} = \mathrm{MLP}_p\left(h_p\right) \tag{2}$$

$$h_q = \mathrm{Attn}_q\left(c, E^L_{\mathrm{out}}(x;y), E^L_{\mathrm{out}}(x;y)\right) \tag{3}$$

$$\begin{bmatrix} \mu_q \\ \log\left(\sigma_q^2\right) \end{bmatrix} = \mathrm{MLP}_q\left(h_q\right) \tag{4}$$

$\mathrm{Attn}_*(\cdot)$, $\mathrm{MLP}_*(\cdot)$ はそれぞれ注意機構 [10]，多層パーセプトロンによる演算を示す．c は注意機構のクエリとなるベクトルであり，タスクを通して学習される．$E^L_{\mathrm{out}}(\cdot)$ はエンコーダの最終層である．

訓練時は，以下の式で示される Evidence Lower Bound (ELBO) を最大化することによって，モデルの最適化を行う．D_{KL} は分布間の Kullback–Leibler divergence (KL divergence) を意味する．

$$\begin{aligned}\log p(y \mid x) &= \log \int_z p(y \mid x, z) p(z \mid x) dz \\ &\geq \mathbb{E}_{q(z|x,y)}[\log p(y \mid x, z)] \\ &\quad - D_{KL}\left(q(z \mid x, y) \| p(z \mid x)\right) \end{aligned} \tag{5}$$

この最適化により，発話・応答の両方から推定した事後分布と発話のみから推定した事前分布は近づく．その結果，テスト時においても，発話から適切な事前分布を推定した上で応答生成が可能となる．

2.2 潜在変数の投機的サンプリング

1 節で述べたように，CVAE に基づく対話モデルを用いたとしても，多様性の高い応答が得られない場合がある．そうした状況においては，モデルは注意機構などによって計算されるベクトルのみに頼り，多様な応答を生み出すために導入された潜在変数 z を無視して生成を行ってしまう．結果，潜在変数の応答への影響を考慮せずに事前・事後分布間の KL divergence について過度な最適化が行われることから，この問題は KL vanishing と呼称される．

我々は KL vanishing の一因は，分布からサンプリングされた潜在変数と与えられた訓練例の不一致にあると考えた．学習過程の初期においては推定された分布の質は低く，与えられた応答を生成するにあたり，サンプルされた潜在変数が必ずしも適切ではない場合がある [11]．

加えて，我々は十分訓練が進んだ後もこの問題は発生すると考える．事前分布と事後分布は KL-divergence の最適化によって互いに近くなる．そのため訓練例として与えられた応答とは異なる応答に対応した，ある時点のモデルにとって訓練時にサンプルすべきではない潜在変数を選んでしまう可能性が存在する．いずれの場合も，その時点のモデルにおける潜在変数空間にそぐわない最適化が行われ，学習を困難にする．

そこで本研究ではこの問題を解決するため，潜在変数の投機的サンプリング (Speculative sampling) を提案する．提案手法は学習時にのみ潜在変数を K 個サンプルし，そのそれぞれについて損失を計算した後，最も小さな損失に繋がる潜在変数のみを最適化の対象とする．具体的には，以下の式における $L(\hat{z})$ を最小化する．

$$\begin{aligned} L(z) &= -\mathbb{E}_{q(z|x,y)}[\log p(y \mid x, z)] \\ &\quad + D_{KL}\left(q(z \mid x, y) \| p(z \mid x)\right) \\ \hat{z} &= \operatorname*{argmin}_{z_k} L(z_k) \end{aligned} \tag{6}$$

3 実験設定

3.1 モデル

本研究では，T-CVAE [7] を基本となる対話モデルとする．各手法はすべて fairseq (v0.8.0) [12] 上に自ら実装を行った．主要なハイパーパラメータは Transformer-base [9] に倣い，最大 250,000 ステップ訓練を行った．学習の安定のため，事前学習した CBoW [13] ベクトルによってモデルの埋め込み層の初期化を行った．

実験では T-CVAE および，以下の手法を T-CVAE に対して適用したモデルを比較する．

Monotonic annealing [14]: 式 (5) における第二項，KL divergence に対し変数 β によるアニーリングを行う．β は 0 から始まり，学習の進行とともに 1 へと線形に増加する．アニーリングの期間は 1 エポックとした．

Cyclical annealing [11]: KL divergence の項に対し，変数 β による周期的なアニーリングを行う．本研究では 1 周期を 1 エポックとした．β は前半の 0.5 エポックの間は 0 から 1 へと線形に増加し，後半の 0.5 エポックの間は 1 を取る．

BoW loss [5]: 潜在変数と応答に含まれる単語との関連性を強めるため，学習時に与えられた応答に含まれる単語の集合を潜在変数から推定するタスクを加え，対話タスクとの同時最適化を行う．

Speculative sampling: 2.2 節参照．潜在変数のサンプル数 K は開発データから 5 に決定した．

3.2 データセット

本稿における実験のために，我々の研究室において継続的に収集している Twitter データ上で行われたツイート・メンションの組を 1 ターンの対話

	BLEU	dist-1	dist-2	平均長
参照応答	-	4.11	30.60	12.01
T-CVAE	2.48	1.14	4.24	15.68
Mono. annealing	2.73	1.27	4.47	13.99
Cycl. annealing	2.73	1.19	4.24	14.01
BoW loss	1.97	1.56	**8.74**	10.18
Speculative sampling	**2.91**	**1.57**	6.48	**11.31**

表 1 自動評価の結果．各評価尺度ごとに，参照応答に最も近いものをハイライトした．

としてみなした上で，対話データセットの構築を行った．訓練・開発データには 2017 年および 2018 年のもの，評価データには 2019 年のものを用いた．訓練・開発・評価データのサイズはそれぞれ 18,116,756 件，191,890 件，96,276 件である．またコストの問題から，人手評価を全ての例に対して行う事は困難である．そこで実際に行われた発話・応答のみを見て，人手評価が比較的容易な 100 件の会話をテストデータから抽出し，人手評価用のデータとした．

データの前処理としては簡単なフィルタリングによってリツイートや bot によるメンションなど，対話として不適切なものを除いた後，MeCab[1)] を用いて形態素解析を行った．その後，SentencePiece[2)] を用いてサブワード分割モデルの学習および発話・応答のサブワードへの分割を行った．

3.3 評価尺度

生成された応答に対し，自動評価および人手評価を行う．自動評価の尺度としては標準的に用いられる，BLEU [15]， dist-n [1]，応答の平均長を採用する．人手評価では Adiwardana らの研究 [16] を参考に，1. 応答の明快さ・発話との関連性 (Sensibleness) と，2. 応答の独特さ・面白さ (Specificity) について，5 段階の点数付けを行うことで，提案手法が意図した効果を得られているかを定量的に確認する．

4 実験結果

4.1 自動評価

表 1 に自動評価による結果を示す．**BoW loss** を除き，追加の手法を適用したモデルはどれも純粋な **T-CVAE** と比べて若干高い BLEU スコアを達成した．しかし目標である応答の多様化が達成されたとしても，生成された応答が参照応答と類似すると

1） https://taku910.github.io/mecab
2） https://github.com/google/sentencepiece

	Sensibleness	Specificity	Average
T-CVAE	3.550	1.590	2.570
Mono. annealing	3.380	1.520	2.450
Cycl. annealing	3.425	1.530	2.478
BoW loss	2.780	**1.885**	2.333
Speculative sampling	**3.610**	1.725	**2.668**

表 2 人手評価の結果 (相関係数 $\rho = 0.602$).

は限らないことから，直接的な BLEU スコアの改善には繋がりにくい．そのため，この結果は各手法によってモデルの学習が改善されたことで，ノイジーな出力が減った事による副次的な効果であると考えられる．加えて，**T-CVAE** の応答平均長が最も長く，他のモデルではいずれも減少していることも同じ言葉の反復などのノイズの抑制を示唆している．

dist-1/2 については **BoW loss** および提案手法である **Speculative sampling** が他のモデルと比較して高い値を達成した．BLEU スコア及び後述する 4.2 節における結果も示すように，**BoW loss** がその代償として発話との関連性を損なっているのに対し，提案手法は高い BLEU スコアを保ちつつ，多様な単語を含む応答を生成出来ている．

4.2 人手評価

表 2 に人手評価による応答の Sensibleness, Specificity, およびその平均値を示す．全体としては，自動評価と概ね類似する傾向であった．**Monotonic annealing** および **Cyclical annealing** による大きな改善は確認されず，**BoW loss** と **Speculative sampling** を適用したモデルが高い Specificity を示した．また，**BoW loss** の Sensibleness は低く，出力の多様化の代償に応答の明快さが損なわれている．

一方で，提案手法である **Speculative sampling** は Sensibleness, Specificity が共に高く，潜在変数空間についての学習手法を改善することで応答生成の性能向上が果たされたと言える．

4.3 分析

訓練の結果，各モデルの分布にどのような変化が見られるかを確認するため，表 3 に開発データに対する KL divergence, 事前分布の平均 μ_p および偏差 σ_p のノルムについて，それぞれ平均を示す．

まず，**Monotonic annealing** および **Cyclic annealing** について，**T-CVAE** と比較して KL divergence の値は若干上昇した．しかし，自動・人手評価の結果と同様大きな変化は見られず，十分に訓練を行った後

	KL-divergence	$\|\mu_p\|$	$\|\sigma_p\|$
T-CVAE	0.003	0.240	21.742
Mono. annealing	0.060	0.216	15.155
Cycl. annealing	0.088	0.330	15.268
BoW loss	24.073	36.949	92.003
Speculative sampling ($K = 2$)	0.621	0.638	21.746
Speculative sampling ($K = 5$)	1.573	0.953	20.949
Speculative sampling ($K = 10$)	2.192	1.053	20.347
Speculative sampling ($K = 20$)	2.910	1.120	18.743

表 3　KL-divergence および事前分布の平均 μ_p，偏差 σ_p のノルム．

発話	急募 喉の痛みの緩和方法
T-CVAE	病院に行った方がいいですよ。
Mono. annealing	葛根湯を飲んでみては?
Cycl. annealing	お大事にしてください...!
BoW loss	胃腸炎にならなくていいと思います。
Spec. sampling	ビタミン C を摂るといいよ。

表 4　生成された応答例．

は，KL vanishing が発生している．

BoW loss の分布は大きく異なる性質を示した．KL divergence および分布の平均・偏差のノルムは，他のモデルと比べ非常に大きな値となっている．**BoW loss** における潜在変数から応答に含まれる全ての単語を推定する，というタスクはモデルに潜在変数空間を個々の応答の意味に対応させる非常に強い制約として働いており，KL vanishing は発生していない．しかしこの制約を満たす潜在変数空間を学習するのはモデルにとって非常に困難である．そのため事前分布の範囲が大きく拡大した結果，発話との関連性が薄い応答に対応した潜在変数が頻繁にサンプルされるようになり，文脈に沿わない応答を生成しがちになったと考える（表 4）．

提案手法について，4.1 節，4.2 節で評価した $K = 5$ のモデルの他に，異なる K の値で訓練した時の各統計量も示した．どの K の値を設定した場合においても KL divergence は **T-CVAE**，つまり $K = 1$ の場合と比較して高い値を示し，提案手法の狙い通り KL vanishing の抑制が行われたと考える．

興味深いことに K の値を増やすほど，KL-divergence の値も増加した．提案手法では K の増加とともに訓練時に潜在変数の候補は増えるため，与えられた応答に対しより適切な潜在変数の選択が可能になる．そのため，これは与えられた訓練例とサンプルされた潜在変数との不一致が学習の妨げとなっているという，我々の仮説を裏付ける結果になったと考える．

5　関連研究

対話応答における応答の多様化のためのアプローチは以下の 3 つに大別される．1. 個人適応に代表される，対話における発話外の情報を用いる手法 [17, 18, 19, 20]，2. 生成時または学習時の目的関数によって応答を多様化する手法 [1, 21]，3. 応答の生成過程にランダム性を導入する，変分オートエンコーダに基づく手法 [14, 5, 22, 11, 3, 23] である．

提案手法は 3 つ目の区分に該当し，特に CVAE における KL vanishing の解決を目的とした手法である．同様の試みのうち，Bowman ら [14] や Fu ら [11] は，学習初期の信頼性の低いモデル上で推定された分布の最適化が性能低下に繋がると考え，KL divergence の重みのスケジューリングを行った．He らも同様の考えから，エンコーダの最適化を優先して行うことでその解決を試みた [24]．

一方，Zhao らの研究は，潜在変数に対して追加の損失によってより直接的な制約を与えることで KL vanishing を抑制するものである [5]．Gao らの研究も潜在変数空間に対する制約によって，応答との対応付けを行うという点で類似している [3, 23]．

提案手法は前者のスケジューリングによって学習を安定させる試みに近い．しかし我々は，潜在変数と訓練例の不一致は KL divergence の最適化の影響により常に発生し得ると考え，学習の進行度を問わず適用可能な手法を提案した．また，提案手法はモデルの構造や目的関数に変更を加えない．そのため，既存手法との併用が可能である．

6　おわりに

本研究では対話タスクにおける応答生成の多様化のため，変分オートエンコーダの学習時に生じる KL vanishing の解決のための手法を提案した．KL vanishing の一因は学習時の訓練例とサンプルされた潜在変数の不一致にあると考え，その解決を図るべく潜在変数の投機的サンプリングを提案した．生成された応答に対し自動評価および人手評価を行い，提案手法による効果を確認した．

謝辞

本研究の一部は JSPS 科研費 19J14522 の支援を受けたものである．また，本研究で用いた Twitter 対話データセットは，同研究室の豊田正史教授の支援を受けて構築した．

参考文献

[1] Jiwei Li, Michel Galley, Chris Brockett, Jianfeng Gao, and Bill Dolan. A diversity-promoting objective function for neural conversation models. In *Proceedings of NAACL 2016*, pp. 110–119, 2016.

[2] Ruqing Zhang, Jiafeng Guo, Yixing Fan, Yanyan Lan, Jun Xu, and Xueqi Cheng. Learning to control the specificity in neural response generation. In *Proceedings of ACL 2018*, pp. 1108–1117, 2018.

[3] Xiang Gao, Sungjin Lee, Yizhe Zhang, Chris Brockett, Michel Galley, Jianfeng Gao, and Bill Dolan. Jointly optimizing diversity and relevance in neural response generation. In *Proceedings of NAACL 2019*, pp. 1229–1238, 2019.

[4] Iulian Serban, Alessandro Sordoni, Ryan Lowe, Laurent Charlin, Joelle Pineau, Aaron Courville, and Yoshua Bengio. A hierarchical latent variable encoder-decoder model for generating dialogues, 2017.

[5] Tiancheng Zhao, Ran Zhao, and Maxine Eskenazi. Learning discourse-level diversity for neural dialog models using conditional variational autoencoders. In *Proceedings of ACL 2017*, pp. 654–664, 2017.

[6] Xiaoyu Shen, Hui Su, Shuzi Niu, and Vera Demberg. Improving variational encoder-decoders in dialogue generation. In *Proceedings of AAAI 2018*, pp. 5456—-5463, 2018.

[7] Tianming Wang and Xiaojun Wan. T-cvae: Transformer-based conditioned variational autoencoder for story completion. In *Proceedings of IJCAI 2019*, pp. 5233–5239, 2019.

[8] Iulian V Serban, Alessandro Sordoni, Yoshua Bengio, Aaron Courville, and Joelle Pineau. Building end-to-end dialogue systems using generative hierarchical neural network models. In *Proceedings of AAAI 2016*, 2016.

[9] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Ł ukasz Kaiser, and Illia Polosukhin. Attention is all you need. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in NIPS 2017*, pp. 5998–6008, 2017.

[10] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. Neural machine translation by jointly learning to align and translate. In *Proceedings of ICLR 2015*, 2015.

[11] Hao Fu, Chunyuan Li, Xiaodong Liu, Jianfeng Gao, Asli Celikyilmaz, and Lawrence Carin. Cyclical annealing schedule: A simple approach to mitigating KL vanishing. In *Proceedings of NAACL 2019*, pp. 240–250, 2019.

[12] Myle Ott, Sergey Edunov, Alexei Baevski, Angela Fan, Sam Gross, Nathan Ng, David Grangier, and Michael Auli. fairseq: A fast, extensible toolkit for sequence modeling. In *Proceedings of NAACL 2019 (Demonstrations)*, pp. 48–53, 2019.

[13] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. Distributed representations of words and phrases and their compositionality. In *Proceedings of NIPS 2013*, pp. 3111–3119, 2013.

[14] Samuel R. Bowman, Luke Vilnis, Oriol Vinyals, Andrew Dai, Rafal Jozefowicz, and Samy Bengio. Generating sentences from a continuous space. In *Proceedings of CoNLL 2016*, pp. 10–21, 2016.

[15] Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. Bleu: a method for automatic evaluation of machine translation. In *Proceedings of ACL 2002*, pp. 311–318, 2002.

[16] Daniel Adiwardana, Minh-Thang Luong, David R So, Jamie Hall, Noah Fiedel, Romal Thoppilan, Zi Yang, Apoorv Kulshreshtha, Gaurav Nemade, Yifeng Lu, et al. Towards a human-like open-domain chatbot. *arXiv preprint arXiv:2001.09977*, 2020.

[17] Shoetsu Sato, Naoki Yoshinaga, Masashi Toyoda, and Masaru Kitsuregawa. Modeling situations in neural chat bots. In *Proceedings of ACL 2017, Student Research Workshop*, pp. 120–127, 2017.

[18] Saizheng Zhang, Emily Dinan, Jack Urbanek, Arthur Szlam, Douwe Kiela, and Jason Weston. Personalizing dialogue agents: I have a dog, do you have pets too? In *Proceedings of ACL 2018*, pp. 2204–2213, 2018.

[19] Pierre-Emmanuel Mazare, Samuel Humeau, Martin Raison, and Antoine Bordes. Training millions of personalized dialogue agents. In *Proceedings of EMNLP 2018*, pp. 2775–2779, 2018.

[20] Eric Chu, Prashanth Vijayaraghavan, and Deb Roy. Learning personas from dialogue with attentive memory networks. In *Proceedings of EMNLP 2018*, pp. 2638–2646, 2018.

[21] Ashutosh Baheti, Alan Ritter, Jiwei Li, and Bill Dolan. Generating more interesting responses in neural conversation models with distributional constraints. In *Proceedings of EMNLP 2018*, pp. 3970–3980, 2018.

[22] Xiaodong Gu, Kyunghyun Cho, Jung-Woo Ha, and Sunghun Kim. Dialogwae: Multimodal response generation with conditional wasserstein auto-encoder. *arXiv preprint arXiv:1805.12352*, 2018.

[23] Xiang Gao, Yizhe Zhang, Sungjin Lee, Michel Galley, Chris Brockett, Jianfeng Gao, and Bill Dolan. Structuring latent spaces for stylized response generation. In *Proceedings of EMNLP-IJCNLP 2019*, pp. 1814–1823, 2019.

[24] Junxian He, Daniel Spokoyny, Graham Neubig, and Taylor Berg-Kirkpatrick. Lagging inference networks and posterior collapse in variational autoencoders. 2019.