

対話行為を用いた制御可能なニューラル対話モデルの検討

An Investigation of Controllable Neural Conversation Model with Dialogue Acts

河野 誠也 *1

Seiya Kawano

吉野 幸一郎 *1*2

Koichiro Yoshino

中村 哲 *1

Satoshi Nakamura

*1 奈良先端科学技術大学院大学

Nara Institute of Science and Technology

*2 科学技術振興機構

Japan Science and Technology Agency

Dialogue act is known as an essential component of the dialogue system, which captures the user's intention and produces the appropriate response. In this paper, we propose a controllable response generation model given dialogue acts. Recent neural conversation models are based on the end-to-end approach that learns a mapping between dialogue histories and response utterances. However, it was difficult to control the contents of the response generated by the model. Several models tackled the problem of generating responses under the specified dialogue acts as a condition; however, these models still have problems on conditioned generations. In this paper, we introduced an extended framework of the generative adversarial network that optimizes both conditioned generator and discriminator which explicitly classifies dialogue act classes. Experimental results showed that our conditional response generation model improved both the response quality and controllability of neural conversation generation.

1. はじめに

対話行為とは、話者が発話において持つ何かしらの「意図」あるいは発話における「機能」であり、その意図や機能の種類として対話行為タグが定義される [Boyer 10, 河野 18]. 対話行為は、対話モデルにおける基本単位の一つとして利用されており、特に近年、対話における発話間の相互作用をモデル化する上で有用であることが知られている [Yoshino 15]. しかし、近年広く用いられているニューラル対話モデル (Neural Conversation Model; NCM) では、こうした対話行為によってシステム発話を明示的に操作することが困難である。そこで本研究では、システム応答が持つ対話行為の情報を条件として用いた敵対的生成学習の枠組みを NCM に導入する。具体的には、与えられた対話行為に基づいて応答を生成する Generator と、Generator が生成した応答が指定した対話行為に基づいた適切なものであるかを判別する Discriminator を構築し、これらの 2 つのモデルを交互に敵対的に訓練する。このような学習の枠組みの導入により、NCM が任意の対話行為に基づいた適切な応答を生成することが可能かどうかの評価を行った。

2. 関連研究

対話行為の情報をを用いて発話生成を行った研究として、Wen や Zhao らの研究がある [Wen 15, Zhao 17]. Wen らは、レストラン案内対話のドメインにおいて、任意の対話行為 (e.g. Inform) とそのスロット情報 (e.g. food=Chinese, name=Seven.Days) を用いて発話生成 (e.g. “Seven Days serves Chinese.”) を行うニューラル言語モデル (Semantically Conditioned LSTM; SC-LSTM) を提案している。SC-LSTM では、LSTM 言語モデルに「読み込みゲート」と呼ばれるゲーティング機構を新たに導入することで、単語の生成時に、指定した対話行為とスロット情報のうちの情報を使うかの制御を行う。しかしながら、SC-LSTM は限られたドメインのタスク達成を対象としているため、取り扱う発話の範囲が膨大なオー

ブンドメインの非タスク志向対話システムにおいて適用しようとした場合、対話行為とそのスロット情報の両方をフレームとして保持することは現実的ではない。

また、Zhao らは NCM が過度に一般化された応答を返すという問題 [Li 16a] に対して、対話行為の情報をを用いることでより多様かつ高品質な応答生成を行う KgCVAE (Knowledge-guided Conditional Variational Autoencoder) を提案している。KgCVAE では、Encoder の出力結果から応答の対話行為を予測し、これを Decoder への条件として使用している。また、同様のアプローチとして、NCM における Decoder に感情ラベルや話者ラベルを条件として利用し、NCM における応答内容を制御しようとするモデルも提案されている [Zhou 18, Li 16b]. これらの NCM は、SCE (Softmax Cross Entropy Loss) に基づいて Decoder の各単語生成ステップにおける単語予測の最適化を行う。しかしこれらのモデルは必ずしも生成された応答が指定したラベルに基づいた適切な応答であることを保証しない。この傾向は特に、低頻度のラベルで顕著である。これは、訓練データにおけるラベルの分布の偏りに起因する [Zhou 18].

一方で、NCM に敵対的生成ネットワーク (Generative Adversarial Net; GAN) を導入することで、より多様かつ高品質な応答生成モデルの構築が検討されている [Li 17a]. 具体的には、応答発話を生成する Generator と、与えられた応答が学習データか Generator のどちらに由来するものかを識別する Discriminator を構築する。これらの 2 つのモデルを交互に敵対的に訓練することで、より自然な発話を生成しようとする試みである。このような、学習の枠組みが従来の NCM と比較して高品質かつ多様な応答を生成したことが報告されている。本研究はこうした GAN に基づく学習の枠組みを、指定した対話行為らしい応答を生成することを目的として導入する。

3. 提案手法

本研究では、前節で述べた NCM における GAN に基づく発話生成の枠組みを拡張し、応答の対話行為を条件として用いる NCM (条件付き NCM) に応用する。具体的には、与えられた対話行為に基づいて応答を生成する Generator (G) と、G が

連絡先: 河野誠也, 奈良先端科学技術大学院大学, 〒 630-0101, 奈良県生駒市高山町 8916 番地の 5, kawano.seiya.kjo@is.naist.jp

生成した応答が指定した対話行為に基づいた適切なものであるかを識別する Discriminator (D) を構築し、これらの 2 つのモデルを交互に敵対的に訓練する。本手法は、G による応答生成と D による応答発話の評価の枠組みに、応答の対話行為の情報を明示的に導入している点で従来手法 [Li 17a] と異なる。本研究では、このような応答の対話行為の情報を利用した GAN の枠組みの導入により、NCM が任意の対話行為に基づいた適切な応答を生成することが可能かどうかの評価を行う。

3.1 対話行為による条件付き NCM

対話行為による条件付き応答生成では、対話履歴 $M = \{M_{i-1}, M_{i-2}, \dots, M_{i-n}\}$ と応答の対話行為 d_i を用いて、応答発話 $R_i = \{w_1, w_2, \dots, w_l\}$ を生成する。ここで、 n は対話長、 l は発話長である。本研究では、最も素朴な NCM の拡張として図 1 のように、階層的な Encoder-decoder モデルに基づく NCM [Tian 17] に対して、発話のデコード時に明示的に対話行為の情報を与えるモデルを構築した。

通常の Encoder-Decoder モデルでは、Decoder における各単語出力ステップ t において、直前の隠れ状態 h_{t-1} と現在の単語 w_t のベクトル表現 x_t が入力として使用される。一方で、本モデルでは、直前の隠れ状態 h_{t-1} に加えて、対話行為 d_i のベクトル表現 e と x_t の連結ベクトル $v_t = e \oplus x_t$ を入力として使用することで、指定した対話行為に基づいた応答を NCM が生成することを期待する*1。

RNN には、LSTM や GRU と比較して高速で訓練が可能であることが知られている SRU (Simple Recurrent Unit) [Lei 17] を用いた。評価実験では、Encoder 及び Decoder で使用する SRU の層数は 2、中間層の次元は 1024、単語埋め込みベクトルの次元は 256、対話行為埋め込みベクトルの次元は 100 とした。また、Encoder における最大対話長 n は 5 に設定した。モデルの訓練は SCE により行い、開発セットにおける Perplexity が最も下がったモデルを評価に使用する。

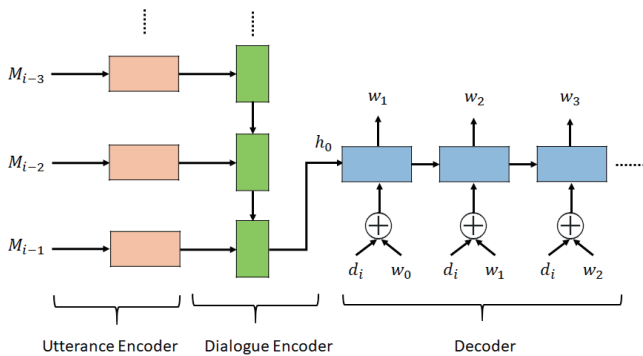


図 1: 対話行為による条件付き NCM の概要

3.2 対話行為による条件付き敵対的生成ネットワーク

本研究では、敵対的生成学習 (GAN) の手法として SeqGAN (Sequence Generative Adversarial Net) [Yu 17, Li 17a] の枠組みを採用し、対話行為による条件付き応答生成へ応用する。発話のような離散系列データの生成に GAN を適用する場合、Generator の出力は単語の離散値の系列を前提とするため、Discriminator から Generator への勾配の更新を行うことが難しい。この問題に対して、SeqGAN では、強化学習によるアプローチを Generator の更新に用いることで、離散的

*1 入力単語 w_t 及び対話行為 d は埋め込み層を用いて密ベクトル表現 x_t, e へと変換される。

な系列データに GAN を適用できることを示している。以下に GAN が持つ Generator, Discriminator の構成を示す。

3.2.1 Generator (G)

Generator では、応答生成における単語選択のプロセスをマルコフ決定過程とみなし、各ステップにおける単語生成がある政策に基づいて選択された行動によって決定されると考える。SeqGAN では、Discriminator によって推定された応答文の妥当性についてのスコア $Q(M, R, d)$ を Generator の報酬として利用し、この期待報酬を最大化するように Generator を方策勾配法により更新する [Williams 92]。本研究における対話行為による条件付き応答生成においては、その目的関数 J と勾配を次式で定義する*2。ここで、 π は応答 R の生成確率である。本研究では、前節で述べた対話行為による条件付き NCM を Generator として用いる。

$$J(\theta) = \mathbb{E}_{R \sim p(R|M,d)}(Q(M, R, d)|\theta) \quad (1)$$

$$\nabla J(\theta) \approx Q(M, R, d) \nabla \log \pi(R|M, d) \quad (2)$$

$$= Q(M, R, d) \nabla \sum_t \log \pi(w_t|M, d, w_{1:t-1}) \quad (3)$$

3.2.2 Discriminator

本研究では、以下の 2 種類の Discriminator について検討を行う。

- Implicit-Discriminator; 対話履歴 M , 応答発話 R , 対話行為 d を入力として、与えられた R が本物 (学習データ由来のもの) か偽物 (Generator 由来のもの) であるかの 2 クラス分類を行う。Generator によって生成された R が本物であると推定される確率を $Q(M, R, d)$ として利用する。これは、対話行為タグを素朴に Discriminator の入力として用い、生成結果が制御されることを期待するものである。
- Explicit-Discriminator; 対話履歴 M , 対話行為 d に対応する応答発話 R を入力として、 R が Generator 由来のものであるとき偽物、学習データ由来のものであるときその対話行為を分類するように Discriminator を訓練する。つまり、対話行為の種類数を N としたとき、 $N+1$ クラスの多クラス分類を行う。ここで、 $Q(M, R, d)$ は、Generator によって生成された R が学習データ由来の対話行為 d であると推定される確率である。生成文が各対話行為の特徴を反映しているかを Discriminator が明示的に予測することにより、各対話行為の特徴が生成文に反映されることを期待する。

Discriminator には、図 2 に示すような Utterance Encoder と Dialogue Encoder で構成される階層型 RNN を Discriminator として用いる。また、Generator の場合と同様に、RNN には SRU を Discriminator の訓練の高速化のために採用する。評価実験においては、SRU の層数は 1、中間層の次元は 1024、単語埋め込みベクトルの次元は 256、対話行為埋め込みベクトルの次元は 100、全結合層における中間層の次元は 512 に設定した。また、Discriminator の更新における損失関数には SCE を用いる。

4. 評価実験

4.1 データセット

対話行為による条件付き NCM の訓練と評価には DailyDialog Corpus [Li 17b] を用いた。本コーパスには、13,118 対話における発話に、Inform (4,6532 発話)、Questions (29,428 発話)、Directives (17,29 発話)、Commisive (9,724 発話) の

*2 本研究では簡化のために文全体で単一の報酬を使用している点に留意されたい。

表 1: 対話行為による条件付き NCM の評価

モデル	Perplexity	Average	Greedy	Extreme	Distinct-1	Distinct-2	Accuracy
条件付き NCM (greedy)	36.6614	0.7912	5.5075	0.5298	0.0336	0.0160	0.8644
Adversarial-Implicit (greedy)	39.2864	0.7882	5.6239	0.5394	0.0297	0.0069	0.8585
Adversarial-Explicit (greedy)	39.6993	0.7867	5.6661	0.5395	0.0311	0.0010	0.8843
条件付き NCM (sampling)	36.6614	0.7917	5.1785	0.4905	0.0767	0.2650	0.8176
Adversarial-Implicit (sampling)	39.2864	0.7870	5.2566	0.5035	0.0594	0.2093	0.8062
Adversarial-Explicit (sampling)	39.6993	0.7865	5.2782	0.5041	0.0583	0.1186	0.8573
条件付き NCM (beam)	36.6614	0.7847	5.5680	0.5381	0.0325	0.0006	0.8707
Adversarial-Implicit (beam)	39.2864	0.7793	5.6668	0.5411	0.0267	0.0031	0.8615
Adversarial-Explicit (beam)	39.6993	0.7796	5.7167	0.5431	0.0274	0.0017	0.8865

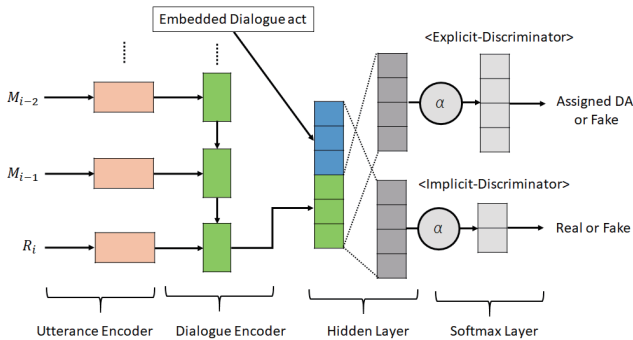


図 2: Implicit-Discriminator, Explicit-Discriminator の概要

4つの対話行為タグがアノテーションされている。本研究では、このコーパスを訓練 (11,118 対話)、開発 (1,000 対話)、評価用データ (1,000 対話) に分割し、対話行為による条件付き NCM の訓練と評価に用いた。語彙サイズは 2,500 に設定し、未知語は特殊記号 “UNK” に置き換えたうえで訓練を行う。対話行為による条件付き敵対的生成ネットワークの訓練方法については、Li らの手法 [Li 17a] を参考に以下の手順で行う。D-step, G-step におけるステップ数はそれぞれ、4, 20 に設定し、batch size は 32 とした。また、Discriminator の更新には SGD (学習率 $1e-3$)、Generator の更新には Adam (学習率 $1e-5$) を用いた *3。

Algorithm 1 条件付き敵対的生成ネットワークの訓練

```

1: for number of iteration do
2:   for number of D-step do
3:     sample  $(M, R, d)$  from training data
4:     generate response  $\hat{R}$  using  $G$  on  $(M, d)$ 
5:     update  $D$  using  $(M, \hat{R}, d)$  and  $(M, R, d)$ 
6:   for number of G-step do
7:     sample  $(M, R, d)$  from training data
8:     generate response  $\hat{R}$  using  $G$  on  $(M, d)$ 
9:     compute reward  $r$  for  $(M, \hat{R}, d)$  using  $D$ 
10:    update  $G$  on  $(M, \hat{R}, d)$  using  $r$ 

```

4.2 対話行為による条件付き応答生成の自動評価

対話行為による条件付き応答生成の評価については、関連性 (Relevance)、多様性 (Diversity)、制御性 (Controllability) の3つの観点から多面的に評価を行う。

- 関連性 (Relevance); 生成された応答発話とレファレンスの応答発話における単語分散表現の類似度により評価を行う。類似度の算出方法には Embedding Average, Greedy Matching, Vector Extreme の三つの評価指標 [Liu 16] を採用した *4。

*3 800 iteration 実行した後のモデルを評価に使用した

*4 これらの詳細については原論文 [Liu 16] を参照されたい。本研究

- 多様性 (Diversity); 対話システムの分野で広く採用されている Distinct-1 及び Distinct-2 により評価を行う。Distinct-1 (Distinct-2) は、テストデータに対して生成された応答発話集合におけるユニークな unigram (bigram) の数を、生成されたすべての unigram (bigram) の数でスケールした指標である [Li 16a]。

- 制御性 (Controllability); 入力として使用した対話行為と、実際に生成された応答文における対話行為の推定結果の一致率 (Accuracy, Precision, Recall, F1) に基づいて評価する。応答文の対話行為については、事前訓練した対話行為分類器 *5 による推定結果 (評価用データにおける推定精度は、0.8303) を正解として用いる。

5. 実験結果

表 1 に対話行為による条件付き NCM の自動評価結果について示す。ここで、条件付き NCM は SCE により訓練されたモデルであり、Adversarial-Implicit モデル、Adversarial-Explicit モデルはそれぞれ、Implicit-Discriminator, Explicit-Discriminator を使用した GAN により訓練されたモデルである。応答文の生成においては、貪欲的に出力単語を選択する場合 (greedy) と多項分布に基づいてサンプリングを行う場合 (sampling), beam search (ビーム幅 5) を行う場合 (beam) の 3 通りを試した。

5.1 応答生成結果の評価

GAN を導入した Adversarial-Implicit モデルでは、いずれの decoding 手法を用いた場合においても、GAN を用いない条件付き NCM と比較して、関連性尺度の Vector Extreme が向上していることが確認できる。一方で、多様性・制御性については改善が確認できなかった。また、sampling を使用する場合は、greedy に応答を生成する場合と比較して、多様性は改善する一方で、制御性は 0.0523 ポイントもの低下が確認できた。しかしながら、Adversarial-Explicit モデルでは、いずれの decoding 手法を用いた場合でも、GAN を用いない条件付き NCM と比較して関連性 (Greedy Matching & Vector Extreme) と制御性が一貫して改善している。特に、制御性においては、greedy に応答を生成する場合は 0.0199 ポイント、sampling を使用する場合は 0.0396 ポイント、beam search を使用する場合は 0.0157 ポイントもの改善が確認でき、本モデルの有効性が示唆される結果となった。

Adversarial-Explicit モデルでは、応答文が Generator か学習データのどちらに由来するものかを判別すると同時に、明示的に対話行為の推定を行っている。これに対して、Adversarial-Implicit モデルでは、生成された応答の対話行為を明示的に

では、wikipedia データで訓練された fastText による単語分散表現 (300 次元) を用いた。

*5 Explicit-Discriminator と同様のアーキテクチャを採用した。

推定しておらず制御性のスコア改善につながらなかったと考えられる。また、GANを導入した2つのモデルでは、GANを用いない条件付きNCMと比較してPerplexityの改善が確認できなかった。GANによるテキスト生成モデルが必ずしもPerplexityを改善しないことは先行研究 [Tevet 18] においても指摘されている。これは、敵対的生成学習は単語予測の最適化を目的としなかったためである。

5.2 各対話行為タグ毎における制御性の評価

表2, 表3はそれぞれAdversarial-Explicitモデルにおいて、beam searchとsamplingを使用して応答生成を行った場合における、各対話行為タグ毎の制御性の評価結果について示している。括弧内は、GANを用いない条件付きNCMにおけるスコアとの差分である。ここで、テストセットにおける各対話行為タグの出現数は、Informは2639, Questionsは1679, Directivesは938, Commisiveは484である。

表2: 各対話行為タグ毎における制御性 (beam search)

対話行為タグ	Precision	Recall	F1
Inform	0.8855	0.9231	0.9039 (+0.0125)
Questions	0.9699	0.9976	0.9836 (+0.0060)
Directives	0.9263	0.6695	0.7772 (+0.0622)
Commisive	0.5976	0.7211	0.6536 (+0.0187)
Macro Avg	0.8448	0.8278	0.8296 (+0.0249)
Weighted Avg	0.8926	0.8864	0.8854 (+0.0111)

表3: 各対話行為タグ毎における制御性 (sampling)

対話行為タグ	Precision	Recall	F1
Inform	0.8766	0.9075	0.8918 (+0.0256)
Questions	0.9086	0.9774	0.9418 (+0.0214)
Directives	0.7963	0.6375	0.7081 (+0.1126)
Commisive	0.6364	0.5930	0.6139 (+0.0587)
Macro Avg	0.8045	0.7789	0.7889 (+0.0546)
Weighted Avg	0.8526	0.8573	0.8530 (+0.0293)

表2, 表3より、GANを用いない条件付きNCMと比較して、Adversarial-Explicitモデルでは、すべての対話行為タグでF値が改善していることが確認できた。また、各対話行為タグにおけるF値のマクロ平均、加重平均も同様に改善していることが確認できた。特に、samplingを使用した場合には、各指標のより顕著な改善が確認できる。しかしながら、Questionsタグにおいては、各対話行為タグの中で最も小さなF値の改善にとどまった。これは、Questionsタグが疑問符や5W1Hのような単語を用いたある程度パターン化された文構造を持つため、問題が簡単(既にベースラインが比較的高い制御性を実現している)であることが原因だと推察される。一方で、Directivesタグについては、beam searchを使用する場合には0.0622ポイント、samplingを使用する場合には0.1126ポイントのF値の顕著な改善が確認できる。Commisiveタグにおいても同様に、beam searchを使用する場合には0.0187ポイント、samplingを使用する場合には0.0587ポイントのF値の改善が確認できた。

6. おわりに

本研究では、任意の対話行為に基づいて応答発話の生成を行う条件付き敵対的生成ネットワークを提案し、その評価を行った。その結果、提案モデルはベースラインモデルと比較して、関連性、制御性における客観評価尺度が改善されることを示した。しかしながら、本研究では客観評価指標による提案手法の

評価のみにとどまり、生成文の品質についてより厳密に評価するためには人手による評価を行うことが必要であると考えられる。さらに、Generatorの更新の際に使用する報酬値の計算に多様性や発話長などの情報を導入することなど、敵対的生成ネットワークの訓練方法そのものの改善を行うことも重要であると考えられる。また、今後は提案手法の他タスク(感情やペルソナを考慮した応答生成など)における有効性の検討も進める予定である。

謝辞

本研究は、JST さきがけ (JPMJPR165B) および JST CREST (JPMJCR1513) の支援を受けた。

参考文献

- [Boyer 10] Boyer, K. E., Ha, E. Y., Phillips, R., Wallis, M. D., Vouk, M. A., and Lester, J. C.: Dialogue act modeling in a complex task-oriented domain, in *Proc. of SIGDIAL*, pp. 297–305 Association for Computational Linguistics (2010)
- [Lei 17] Lei, T., Zhang, Y., and Artzi, Y.: Training rnns as fast as cnns, *arXiv preprint arXiv:1709.02755* (2017)
- [Li 16a] Li, J., Galley, M., Brockett, C., Gao, J., and Dolan, B.: A Diversity-Promoting Objective Function for Neural Conversation Models, in *Proc. of NAACL-HLT*, pp. 110–119 (2016)
- [Li 16b] Li, J., Galley, M., Brockett, C., Spithourakis, G. P., Gao, J., and Dolan, W. B.: A Persona-Based Neural Conversation Model, *CoRR*, Vol. abs/1603.06155, (2016)
- [Li 17a] Li, J., Monroe, W., Shi, T., Jean, S., Ritter, A., and Jurafsky, D.: Adversarial Learning for Neural Dialogue Generation, in *Proc. of EMNLP*, pp. 2157–2169 (2017)
- [Li 17b] Li, Y., Su, H., Shen, X., Li, W., Cao, Z., and Niu, S.: Daily-Dialog: A Manually Labelled Multi-turn Dialogue Dataset, *arXiv preprint arXiv:1710.03957* (2017)
- [Liu 16] Liu, C.-W., Lowe, R., Serban, I., Noseworthy, M., Charlin, L., and Pineau, J.: How NOT To Evaluate Your Dialogue System: An Empirical Study of Unsupervised Evaluation Metrics for Dialogue Response Generation, in *Proc. of EMNLP*, pp. 2122–2132 (2016)
- [Tevet 18] Tevet, G., Habib, G., Shwartz, V., and Berant, J.: Evaluating Text GANs as Language Models, *arXiv preprint arXiv:1810.12686* (2018)
- [Tian 17] Tian, Z., Yan, R., Mou, L., Song, Y., Feng, Y., and Zhao, D.: How to make context more useful? an empirical study on context-aware neural conversational models, in *Proc. of ACL*, Vol. 2, pp. 231–236 (2017)
- [Wen 15] Wen, T.-H., Gasic, M., Mrksic, N., Su, P.-H., Vandyke, D., and Young, S.: Semantically conditioned lstm-based natural language generation for spoken dialogue systems, *arXiv preprint arXiv:1508.01745* (2015)
- [Williams 92] Williams, R. J.: Simple statistical gradient-following algorithms for connectionist reinforcement learning, *Machine learning*, Vol. 8, No. 3-4, pp. 229–256 (1992)
- [Yoshino 15] Yoshino, K. and Kawahara, T.: Conversational system for information navigation based on POMDP with user focus tracking, *Computer Speech & Language*, Vol. 34, No. 1, pp. 275–291 (2015)
- [Yu 17] Yu, L., Zhang, W., Wang, J., and Yu, Y.: SeqGAN: Sequence Generative Adversarial Nets with Policy Gradient., in *AAAI*, pp. 2852–2858 (2017)
- [Zhao 17] Zhao, T., Zhao, R., and Eskenazi, M.: Learning Discourse-level Diversity for Neural Dialog Models using Conditional Variational Autoencoders, in *Proc. of ACL*, Vol. 1, pp. 654–664 (2017)
- [Zhou 18] Zhou, H., Huang, M., Zhang, T., Zhu, X., and Liu, B.: Emotional chatting machine: Emotional conversation generation with internal and external memory, in *Thirty-Second AAAI Conference on Artificial Intelligence* (2018)
- [河野 18] 河野誠也, 吉野幸一郎, 中村哲: 条件付き敵対的生成ネットワークを用いたデータ拡張による対話行為分類法の検討, 情報処理学会研究報告音声言語情報処理 (SLP), No. 9 (2018-SLP-125), pp. 1–6 (2018)