

文脈に対して同調的応答を生成するニューラル会話モデルの検討

Toward Entrained Response Generation for Neural Conversation Model

河野 誠也 *¹

Seiya Kawano

水上 雅博 *²

Masahiro Mizukami

吉野 幸一郎 *¹

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 NTT コミュニケーション科学基礎研究所

NTT Communication Science Laboratory

Synchronization of words in conversation, called entrainment, is generally observed in human-human conversation. It is reported that entrainment has a high correlation with dialogue engagement. In this work, we consider using several entrainment scores based on the word, co-occurrence to evaluate the entrainment of system generation. We try to build a neural conversation model optimized for these entrainment scores by using reinforcement learning. Experimental results showed that the proposed neural conversation model optimized by the entrainment scores achieved a better entrainment score than general neural conversation models.

1. はじめに

エントレインメント (Entrainment) は、同調傾向やシンクロニー (Synchrony) とも呼ばれ、対話における話者間の話し方や声の調子などの振る舞いが同調、類似する現象を指す。この現象は、語彙 [Brennan 96], 統語構造 [Reitter 07], 文体 [Niederhoffer 02], 韻律 [Ward 07], ターンテイキング [Campbell 10], 対話行為 [Mizukami 16] など、対話における様々な要素で観測されることが知られている。エントレインメントは、対話のタスク成功率や自然性、対話意欲 (Engagement) と相関することが報告されており [Nenkova 08, Levitan 15, Nasir 19], エントレインメントの分析を通して対話システムの性能や対話の質を評価する試みが行われている。また、エントレインメントを対話中に考慮し、用例ベース対話システムにおける応答文選択に応用する取り組みが行われている [水上 17]。しかしながら、近年盛んに研究されている sequence-to-sequence モデル [Vinyals 15, Serban 16] を用いたニューラル会話モデルにおいて、エントレインメントを考慮したうえで応答生成を行う手法についての検討は行われていない。そこで本研究では、エントレインメントを考慮して応答生成を行うニューラル会話モデルについての検討を行った。具体的には、提案するニューラル会話モデルはユーザとの対話履歴を考慮して、ユーザに対して同調した発話の生成を行う。このようなモデルを構築するにあたり、本研究ではエントレインメント評価指標を報酬として用いた方策勾配型強化学習によりニューラル会話モデルの最適化を行った。

2. 提案手法

本研究では、ニューラル会話モデルがより文脈に対してエントレインメントしているような応答を生成できるように、強化学習を用いたニューラル会話モデルの最適化法の提案を行う。本章では、まず 2.1 節で本研究で用いるニューラル会話モデルの構成について述べる。また 2.2 節において、本研究で用いたエントレインメント評価指標について述べる。最後に 2.3 節で、強化学習に基づいたニューラル会話モデルの最適化法について述べる。

2.1 注意機構付き階層型 Encoder-Decoder モデル

ニューラル会話モデルにおける応答文生成では、対話履歴 $H = \{H_{i-1}, H_{i-2}, \dots, H_{i-N}\}$ が与えられたとき、応答発話 $R_i = \{w_{i,1}, w_{i,2}, \dots, w_{i,t}\}$ を生成する。ここで、 i はターン、 N は履歴長 *¹、 t は単語順である。本研究では、ニューラル会話モデルの構成として、与えた対話履歴を固定長の文脈表現に符号化する階層型 Encoder と、階層型 Encoder から受け取った文脈表現を用いて発話生成を行う Decoder からなるモデルを用いる [Serban 16]。さらに本研究では、より対話履歴を考慮した応答生成を実現するために、上記モデルに対して対話履歴を考慮した注意機構を新たに導入する。RNN セルには、Gated Recurrent Unit (GRU) を用いた。

2.1.1 階層型 Encoder

階層型 Encoder では、発話 Encoder を用いて対話履歴における各発話 H_i を固定長の発話表現 h_i に符号化する。ここで、 h_i は H_i の各単語 $w_{i,t}$ について、次式を再帰的に適用することで最終的に得られるベクトル $u_{i,t}$ である。また、Embedding は単語 $w_{i,t}$ を固定長密ベクトルに写像する線形変換関数である。

$$u_{i,t} = \text{GRU}(u_{i,t-1}, \text{Embedding}(w_{i,t})) \quad (1)$$

次に、文脈 Encoder において発話 Encoder から得られた発話表現 h_i に対して次式を再帰的に適用することで対話履歴の文脈表現 c_i を得る。

$$c_i = \text{GRU}(c_{i-1}, h_i) \quad (2)$$

2.1.2 注意機構付き Decoder

Decoder においては、階層型 Encoder から得られた対話履歴 H の文脈表現 c_{i-1} を初期状態 h'_0 として用いて、デコーダの中間状態 h'_t と単語の生成確率 p_t を計算することができる。ここで、Linear は h'_t を語彙サイズ次元の密ベクトルに写像する線形変換関数である。また、 $w_{i,t}$ は p_t からサンプルされ次のステップの入力として使用される。

$$h'_t = \text{GRU}(h'_{t-1}, \text{Embedding}(w_{i,t-1})) \quad (3)$$

$$p_t = \text{softmax}(\text{Linear}(h'_t)) \quad (4)$$

連絡先: kawano.seiya.kj0@is.naist.jp

*¹ 本研究では、 $N = 6$ を設定した。

本研究では、エンタレインメントという会話現象を取り扱う性質上、より対話履歴を考慮したモデルを構築することが求められる。したがって、上記の Decoder に対して対話履歴における各発話の情報をより効率的に扱うための注意機構の導入を行う。具体的には、文脈 Encoder によって得られた文脈ベクトルの系列 $c_{i-1-N:i-1}$ 、 t ステップにおける Decoder の中間状態を h'_t としたとき、次のようにして各中間状態に対してアラインメントの重みを計算し、文脈ベクトル \bar{h}_t を求める。

$$\alpha_j = \frac{\exp(c_j \cdot h'_t)}{\sum_{\tilde{j}=i-1-N}^{i-1} \exp(c_{\tilde{j}} \cdot h'_t)} \quad (5)$$

$$\bar{h} = \sum_{\tilde{j}=i-1-N}^{i-1} \alpha_{\tilde{j}} \cdot c_{\tilde{j}} \quad (6)$$

さらに、文脈ベクトル \bar{h} を用いて、ステップ t における出力単語の予測を次のようにして行う。

$$\hat{h}_t = \tanh(\text{Linear}([\bar{h}, h'_t])) \quad (7)$$

$$p_t = \text{softmax}(\text{Linear}(\hat{h}_t)) \quad (8)$$

2.2 Local Interpersonal Distance

本研究では、自動エンタレインメント評価指標を報酬として用いた強化学習による最適化によって、ニューラル会話モデルが対話履歴に対してエンタレインメントした応答を生成する。本研究では、Nasir らが提案したエンタレインメント評価指標である Local Interpersonal Distance (LID) [Nasir 19] をベースとして、強化学習における報酬計算モデルの構築を行う。LID は、エンタレインメントの局所的な側面を発話レベルで捕捉することを目的としており、対話における各ターンの話者の発話のエンタレインメント度合いを Word Mover's Distance (WMD)[Kusner 15] に基づいて以下のように計算する。ここで、対話における各話者 A と B はそれぞれ交互に一発話ずつ対話 $(A_1, B_1, A_2, B_2, \dots, A_N, B_N)$ を行うものと仮定する。 N は発話ペアの総数、 k は直近の発話数であり、WMD は発話間の WMD 値を計算する関数である。

$$d_i^{B \rightarrow A} = \min_{i \leq j \leq i+k \leq N} \text{WMD}(A_i, B_j) \quad (9)$$

$d_i^{B \rightarrow A}$ は話者 B から A に対してのエンタレインメントを示す。同様に、 A から B に対してのエンタレインメントとして $d_i^{A \rightarrow B}$ を計算することができる。LID は、値が小さいほど発話がエンタレインメントしているとみなすことができる。さらに、Nasir らは、各ターンにおいて計算された LID の平均値から計算される対話セッションレベルのエンタレインメント評価指標である Conversational Linguistic Distance (CLiD) を提案しており、動機付け面接における患者の共感、カップルセラピーにおける患者へ与えた肯定的影響と有意な相関があることを確認している。

2.3 方策勾配型強化学習による最適化

前節で述べた LID を参考に、ニューラル会話モデルの生成応答のエンタレインメント度合いを評価するための報酬計算モデルを提案する。ニューラル会話モデルの最適化には、方策勾配型の強化学習の一種である REINFORCE アルゴリズム [Williams 92] を導入する。

2.3.1 REINFORCE アルゴリズム

ニューラル会話モデルにおける応答生成では、対話履歴 H の文脈表現が与えられると、単語列 $R = \{w_1, w_2, \dots, w_T\}$ を

生成する。このような単語生成のプロセスは、マルコフ決定過程において、ある政策 π に従って実行される一連の行動系列とみなすことができる。本研究では、報酬計算モデルによって評価される生成応答のエンタレインメント評価値を報酬として利用し、この期待報酬を最大化するようにモデルのパラメータを REINFORCE アルゴリズムにより更新する [Williams 92]。目的関数 $J_{\text{REINFORCE}}$ と勾配は次式のように定義することができる。ここで、パラメータ θ を持つニューラル会話モデルを G_θ 、単語 w_t を生成する確率を p とする。また、パラメータ θ に従って、単語 w_t を出力することによって得られる報酬 r_t の期待値を Q^{G_θ} とする。

$$J_{\text{REINFORCE}}(\theta) = \sum_{w_{1:T}} G_\theta(w_t | w_{1:t-1}, H) \cdot Q^{G_\theta}((w_{1:t-1}, H), w_t) \quad (10)$$

$$\nabla J_{\text{REINFORCE}}(\theta) \simeq \frac{1}{T} \sum_{t=1}^T \sum_{w_t \in \mathcal{V}} Q^{G_\theta}((w_{1:t-1}, H), w_t) \cdot \nabla_\theta G_\theta(w_t | w_{1:t-1}, H) \quad (11)$$

$$= \frac{1}{T} \sum_{t=1}^T \mathbb{E}_{w_t \sim G_\theta} [Q^{G_\theta}((w_{1:t-1}, H), w_t) \cdot \nabla_\theta \log p(w_t | w_{1:t-1}, H)] \quad (12)$$

しかしながら、本研究で用いる報酬計算モデルは、終端状態まで生成された完全な発話系列に対してのみ計算されるため、中間ステップ t における報酬期待値 Q^{G_θ} を単純には計算することができない。したがって、各ステップ t からのロールアウトを用いた Monte Carlo Tree Search によって報酬期待値 Q^{G_θ} の見積もりを行う [Yu 17]。

2.3.2 エンタレインメントを考慮した報酬計算モデル

生成応答のエンタレインメント度合いを評価するための報酬計算モデルを定義する。最も単純な報酬計算モデルとして、WMD を用いて他方の話者による直前の発話 H_{i-1} と生答応答 R_i の類似度を報酬 r_{previous} として定義することができる。実際には、WMD は正規化されていない類似度指標であるため単純に強化学習における報酬として利用することは望ましくない。したがって、本研究では、WMD の計算結果を 0 から 1 に正規化するための関数 WMD_{norm} を定義した。

$$r_{\text{previous}}(H_{i-1}, R_i) = \text{WMD}_{\text{norm}}(H_{i-1}, R_i) \quad (13)$$

$$\text{WMD}_{\text{norm}}(H_{i-1}, R_i) = e^{-\text{WMD}(H_{i-1}, R_i)^2} \quad (14)$$

さらに、エンタレインメントは必ずしも他方の話者による直前の発話に対してのみ行われるものではなく、他方の話者による発話履歴から文脈に応じて適切にエンタレインメント対象の発話を決定することが望ましい。そこで、他方の話者の発話に対するエンタレインメント度合いを示す r_{LID} を次のように定義する。ここで、 R_i^{ref} は、対話履歴 H に対応するレファレンスの応答発話、 H' は対話履歴 H から、自身による発話をすべて除外したものである。また、 U_{ent} は与えた対話履歴 H' の中でレファレンスの応答発話と最も類似している発話を示す。

$$U_{\text{ent}}(H', R_i) = \arg \max_{H'_j \in H'} \text{WMD}_{\text{norm}}(H'_j, R_i) \quad (15)$$

$$r_{\text{LID}}(H', R_i, R_i^{\text{ref}}) = 1 - \frac{|r_{\text{ideal}} - \text{WMD}_{\text{norm}}(\text{U}^{\text{ent}}(H', R_i^{\text{ref}}), R_i)|}{\max(r_{\text{ideal}}, 1 - r_{\text{ideal}})} \quad (16)$$

r_{LID} では、応答発話がレファレンスから計算されたエントレインメント対象の発話と類似するように、理想的なエントレインメント値 r_{ideal} との相対値を考慮して報酬を与える。これにより、前節で述べた LID を間接的に最大化するようなモデルの学習を行うことが可能になる。本研究では、理想的なエントレインメント値 r_{ideal} として、訓練データにおけるすべての応答事例における実際のエントレインメント度合いを LID の定義を参考に以下の式を用いて計算し、上位 90%, 70%, 50% の値を理想的なエントレインメント度合い r_{ideal} としてモデルの学習を行う。

$$\text{LID}(H', R_i^{\text{ref}}) = \max_{H'_j \in H'} \text{WMD}_{\text{norm}}(H'_j, R_i^{\text{ref}}) \quad (17)$$

2.3.3 強化学習によるニューラル会話モデルの訓練手順

強化学習によるニューラル会話モデルの訓練手順を以下に示す。まず、ニューラル会話モデルの事前学習として、Decoder の各ステップにおける単語予測結果と正解単語の負の対数尤度 $J_{\text{MLE}}(\theta)$ を最小化するようにモデルの訓練を行う。次に、強化学習を用いて目的関数 $J_{\text{REINFORCE}}(\theta)$ を最大化するようにモデルの訓練を行う。実際には、言語モデルの崩壊を防ぐために、各ステップにおいて $J_{\text{REINFORCE}}(\theta)$ によるモデルパラメータ更新後に、 $J_{\text{MLE}}(\theta)$ を用いたモデルパラメータの更新も行う。ただし、 $J_{\text{MLE}}(\theta)$ による更新の影響が支配的になることを防ぐために係数 $\lambda = 0.1$ を乗じたものを使用する。

Algorithm 1 ニューラル会話モデルの訓練手順

- 1: **for** number of steps **do**
- 2: sample (H, R) from training data
- 3: generate response \hat{R} by using G_θ on H
- 4: compute Q^{G_θ} for (H, \hat{R})
- 5: update G_θ based on $J_{\text{REINFORCE}}(\theta)$
- 6: update G_θ based on $J_{\text{MLE}}(\theta)$

3. 評価実験

3.1 データセット

ニューラル会話モデルの訓練と評価には、ConvAI2*2 で提供された PersonaChat データセットを使用した。データセットは、表 2 のように訓練・開発・評価用データセットに分割したものをを用いる。評価用データセットについては、非公開であったため開発用データを二分割し、それぞれ開発・評価用データセットとして用いた。

表 1: ConvAI2 データセットにおける対話数/発話数

	対話数	発話数
訓練	17,876	262,862
開発	498	7,798
評価	499	7,788

3.2 モデル設定

評価実験に用いるニューラル会話モデルとしては、Sequence-to-Sequence モデル (SEQ2SEQ) と、SEQ2SEQ に内積に

*2 <http://convai.io>

よるスコア計算を用いたアテンション機構 [Luong 15] を導入したモデル (Attention-SEQ2SEQ), 前章で述べた階層型 Encoder-Decoder モデル (HED), 注意機構付き階層型 Encoder-Decoder モデル (Attention-HED) の 4 つニューラル会話モデルを使用する。

いずれのモデルも RNN には GRU を使用し、単語埋め込み層の次元を 300, 中間層の次元を 300, 層数を 1 に設定した。また、使用する語彙サイズは 1,5000 とし、未知語は特殊記号 “UNK” に置き換えた。ニューラル会話モデルの事前訓練は交差エン트로ピー誤差を用いて行った後、報酬計算モデル $r_{\text{previous}}, r_{\text{LID}}^{90\%}, r_{\text{LID}}^{70\%}, r_{\text{LID}}^{50\%}$ のそれぞれを用いた場合で強化学習による最適化を行った。ここで、 $r_{\text{LID}}^{90\%}, r_{\text{LID}}^{70\%}, r_{\text{LID}}^{50\%}$ は訓練データにおける実際のエントレインメント度合いの上位 90%, 70%, 50% の値を r_{ideal} として r_{LID} の計算を行う場合に対応する。モデルの訓練においてはバッチサイズを 64, 学習率を 1×10^{-4} として、Optimizer には SGD を使用した。

3.3 生成文の自動評価

応答生成の評価においては、これまでの既存研究と同様に、言語モデルの性能を測るための指標である Perplexity (PPL) による評価と、生成応答とレファレンス発話の類似度 (関連性) による評価を行う。関連性についての評価指標としては、BLEU による評価と、(14) 式で定義した正規化を適用した Word Mover’s Distance 値 [Kusner 15] の平均による評価 ($\overline{\text{WMD}}$) を用いた。また、エントレインメントに着目した応答生成の評価を行うために、各報酬計算モデルが応答発話に対して与える報酬値の平均 ($\bar{r}_{\text{previous}}, \bar{r}_{\text{LID}}^{90\%}, \bar{r}_{\text{LID}}^{70\%}, \bar{r}_{\text{LID}}^{50\%}$) を用いる。ここで、 $\bar{r}_{\text{previous}}$ は r_{previous} を用いた場合、 $\bar{r}_{\text{LID}}^{90\%}, \bar{r}_{\text{LID}}^{70\%}, \bar{r}_{\text{LID}}^{50\%}$ は訓練データにおける実際のエントレインメント度合いの上位 90%, 70%, 50% の値を理想的なエントレインメント度合いとして、それぞれ用いた場合を示す。WMD の計算においては、Twitter データで事前学習された 200 次元の単語分散表現ベクトル*3 をノルムが 1 になるように正規化したものを用いた。

4. 実験結果

表 5 に各ニューラル会話モデルにおける応答生成結果について示す。ここで、強化学習を使用しないモデルでは、開発セットにおいて最も Perplexity が下がったモデルを使用した。また、強化学習を適用したモデルでは、訓練に使用した報酬計算モデルが開発セットに対して生成した応答に与える報酬の平均が最も大きいモデルを使用した。

表 5 より、負の対数尤度の最小化による訓練を用いたベースラインモデル (MLE) は、人間による応答と比較してエントレインメント評価値 ($\bar{r}_{\text{previous}}, \bar{r}_{\text{LID}}^{90\%}, \bar{r}_{\text{LID}}^{70\%}, \bar{r}_{\text{LID}}^{50\%}$) が大きく劣ることが確認できる。これは、既存のニューラル会話モデルがエントレインメントを行う能力に乏しいのに加え、既存研究でも指摘されているように対話履歴の情報を有効に活用できていないことを示唆している。一方で、エントレインメント評価値を報酬として用いた強化学習を適用したモデルは、ベースラインモデルと比較して PPL がほぼ同等でかつ、モデルの訓練に用いた報酬計算モデルに対応するエントレインメント評価値が大きく向上していることが確認できる。

生成応答とレファレンス応答についての関連性評価については、強化学習を適用する場合は、 $\overline{\text{WMD}}$ が大きく向上していることが分かる。一方で、BLEU については低下傾向にあることが確認できた。今回訓練したモデルではどのような文脈においても必ずエントレインメントした応答を生成するよ

*3 <http://nlp.stanford.edu/projects/glove/>

表 2: 各ニューラル会話モデルにおける応答生成結果

訓練方法	モデル	PPL	BLEU	WMD	$\bar{F}_{\text{previous}}$	$\bar{F}_{\text{LID}}^{90\%}$	$\bar{F}_{\text{LID}}^{70\%}$	$\bar{F}_{\text{LID}}^{50\%}$
	Human	1.0	100	100	40.31	72.06	81.96	84.58
MLE	SEQ2SEQ	38.70	2.08	37.74	34.00	60.43	75.62	83.98
	Attention-SEQ2SEQ	38.25	1.81	36.51	32.89	57.92	72.99	81.57
	HED	38.34	1.47	36.4	31.62	56.99	72.71	82.14
	Attention-HED	38.23	1.9	38.16	33.36	60.52	76.42	85.22
REINFORCE w/ $r_{\text{LID}}^{90\%}$	HED	38.37	0.79	47.36	43.63	78.56	86.54	85.26
	Attention-HED	38.19	1.12	47.41	43.97	79.58	86.67	84.87
REINFORCE w/ $r_{\text{LID}}^{70\%}$	HED	38.35	0.27	44.70	41.94	76.41	87.10	87.22
	Attention-HED	38.13	1.12	45.43	41.48	76.28	88.30	88.72
REINFORCE w/ $r_{\text{LID}}^{50\%}$	HED	38.34	1.6	41.11	37.37	68.30	84.68	90.84
	Attention-HED	38.10	1.04	39.27	36.08	66.32	83.54	90.96

うに訓練されており、実際の人間の応答傾向とは大きく異なっていることが関連性評価指標に一貫した改善が確認できない要因であると考えられる。

5. おわりに

本研究では、自動エンタインメント評価指標を報酬として用いた方策勾配型強化学習によるニューラル会話モデルを提案した。評価実験の結果、提案モデルは、言語モデルの性能を維持しつつ、従来モデルと比較して大幅に生成応答のエンタインメント度合いを改善することを示した。今後の課題としては、客観評価だけでなく、実際の人間による対話を通じた主観評価により提案法の有効性を検証する必要がある。また、本研究で提案したモデルは、どのような文脈に対しても必ずモデルの訓練時に与えた理想的なエンタインメント値に近づくような応答の生成を行っている。しかしながら、対話において常にエンタインメントすることが必ずしも有効でない場合がある。したがって、文脈に応じた適切なエンタインメント度合いの調整や制御を行えるようにモデルの拡張を行う必要がある。また、応答発話と対話履歴における単語表層の単純な類似性を考慮するだけでなく、対話行為を考慮した応答生成手法 [水上 17, Kawano 19] の検討も行う。

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

- [Brennan 96] Brennan, S. E. and Clark, H. H.: Conceptual pacts and lexical choice in conversation., *Journal of Experimental Psychology: Learning, Memory, and Cognition*, Vol. 22, No. 6, p. 1482 (1996)
- [Campbell 10] Campbell, N. and Scherer, S.: Comparing measures of synchrony and alignment in dialogue speech timing with respect to turn-taking activity, in *Eleventh Annual Conference of the International Speech Communication Association* (2010)
- [Kawano 19] Kawano, S., Yoshino, K., and Nakamura, S.: Neural Conversation Model Controllable by Given Dialogue Act Based on Adversarial Learning and Label-aware Objective, in *Proceedings of the 12th International Conference on Natural Language Generation* (2019)
- [Kusner 15] Kusner, M., Sun, Y., Kolkin, N., and Weinberger, K.: From word embeddings to document distances, in *International conference on machine learning*, pp. 957–966 (2015)

- [Levitan 15] Levitan, R., Benus, S., Gravano, A., and Hirschberg, J.: Entrainment and turn-taking in human-human dialogue, in *2015 AAAI spring symposium series* (2015)
- [Luong 15] Luong, M.-T., Pham, H., and Manning, C. D.: Effective approaches to attention-based neural machine translation, *arXiv preprint arXiv:1508.04025* (2015)
- [Mizukami 16] Mizukami, M., Yoshino, K., Neubig, G., Traum, D., and Nakamura, S.: Analyzing the Effect of Entrainment on Dialogue Acts, in *Proceedings of the 17th Annual Meeting of the Special Interest Group on Discourse and Dialogue* (2016)
- [Nasir 19] Nasir, M., Chakravarthula, S. N., Baucom, B. R., Atkins, D. C., Georgiou, P., and Narayanan, S.: Modeling Interpersonal Linguistic Coordination in Conversations Using Word Mover’s Distance, in *Proc. Interspeech 2019*, pp. 1423–1427 (2019)
- [Nenkova 08] Nenkova, A., Gravano, A., and Hirschberg, J.: High frequency word entrainment in spoken dialogue, in *Proceedings of the 46th annual meeting of the association for computational linguistics on human language technologies: Short papers*, pp. 169–172 Association for Computational Linguistics (2008)
- [Niederhoffer 02] Niederhoffer, K. G. and Pennebaker, J. W.: Linguistic style matching in social interaction, *Journal of Language and Social Psychology*, Vol. 21, No. 4, pp. 337–360 (2002)
- [Reitter 07] Reitter, D. and Moore, J. D.: Predicting success in dialogue (2007)
- [Serban 16] Serban, I. V., Sordani, A., Bengio, Y., Courville, A. C., and Pineau, J.: Building End-To-End Dialogue Systems Using Generative Hierarchical Neural Network Models., in *AAAI*, pp. 3776–3784 (2016)
- [Vinyals 15] Vinyals, O. and Le, Q.: A neural conversational model, *arXiv preprint arXiv:1506.05869* (2015)
- [Ward 07] Ward, A. and Litman, D.: Measuring convergence and priming in tutorial dialog, *University of Pittsburgh* (2007)
- [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)
- [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)
- [水上 17] 水上 雅博, 吉野 幸一郎, Neubig, G., 中村 哲: エンタインメント分析に基づく応答文選択モデルの評価, 言語処理学会第 23 回年次大会 (NLP2017), 茨城 (2017)