

# Distant Supervision を用いた情報量を制御可能な雑談応答生成

高山 隼矢, 荒瀬 由紀

大阪大学大学院情報科学研究科

{takayama.junya, arase}@ist.osaka-u.ac.jp

## 1 はじめに

雑談応答生成とは、特定のタスクの達成を目的とせず、ユーザの発話に対して適切な応答を生成するタスクである。近年では、機械翻訳分野において広く用いられてきた Sequence to Sequence (seq2seq) [1] 型のモデルを応答生成に適用した手法 [2] が提案されている。seq2seq は入力文を基に End-to-End に出力文を生成するモデルであるため、学習データとして大量の発話応答対を用意すれば比較的流暢な応答文を生成することができる。しかし、seq2seq は機械翻訳のように入力文と意味的に類似した文を出力するタスクでは高精度な文生成ができるが、雑談応答のように入力文に対して可能な応答が多様なタスクにおいては「そうですね」や「わかる」などといった、可能な中で最も無難で高頻度な応答を生成してしまう問題を抱えている [3]。これは、seq2seq が入力発話文に対する応答として参照文が生成される確率を最大化するように訓練されるためであると考えられる。

本研究では無難な応答を抑制し、入力文に対する関連度が高く、かつ情報量の多い応答を生成することを目的とする。提案手法では単語間の正の自己相互情報量 (Positive Point-wise Mutual Information; PPMI) を用いて、ある入力発話文に対して共起しやすい特徴的な単語が出力されやすくなるように応答生成器を改良する。また PPMI に基づいて発話文と応答文の共起を測る MaxPMI という尺度を設計し、これを学習データに自動的に付与することで Distant Supervision を導入する。これによって入力文と MaxPMI に対する応答文の一意性が高まると期待される。推論時には任意のスコアを入力することで生成文の情報量を制御できる。また、MaxPMI の上限値は入力文によって一意に定まるため、この値を入力すれば入力文に対して情報量が最大となる文を出力することも可能である。

DailyDialog (英語学習者どうしの雑談対話) での実験においては、自動評価で既存手法を上回る結果を得られた。

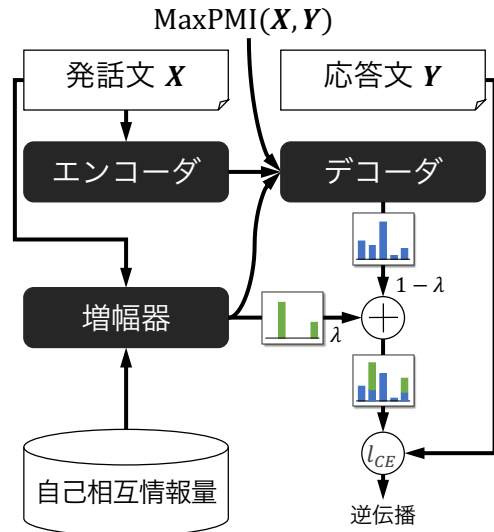


図 1: モデルの構造図

## 2 関連研究

雑談応答生成における無難な応答の抑制を図った手法の一つに、MMI [3] がある。この手法では発話に対する相互情報量が高い応答を返すために、発話文  $X$  と生成文  $Y$  の間の PMI を  $(1 - \lambda) \log p(\mathbf{Y}|\mathbf{X}) + \lambda \log p(\mathbf{X}|\mathbf{Y})$  で近似し、生成時にこれを用いてランキングを行う。しかし、この手法は訓練時に相互情報量を最大化しないため、 $p(\mathbf{Y}|\mathbf{X})$  による  $N$ -best 生成において妥当性の高い語が出力されなかった場合には効果が期待できない。

Zhang ら [4] は、学習データ中の応答文に文の特異性を示すスコアを自動で付与し Distant Supervision による学習を行うことで、出力文の特異性を制御できる応答生成手法を構築している。文の特異性の尺度としては、単語逆頻度が高い語を持つ文は特異性が高いという仮説から設計された Normalized Inverse Word Frequency (NIWF) を用いている。しかし、NIWF は  $\mathbf{Y}$  のみによって決定される変数であるため、特異であっても入力文に対する関連度が多いとは言えない応答がしばしば生成される。

### 3 提案手法

提案手法の概要を図 1 に示す。提案手法ではまず PPMI に基づく共起尺度 MaxPMI を用いて入力発話文と応答文の共起度合いを示すラベルを自動でアノテーションする (第 3.1 節)。モデルは事前に計算した PPMI と MaxPMI に基づいて文生成を行う (第 3.2 節)。学習時は Distant Supervision の枠組みを利用し、発話応答対と事前付与したラベルを基に訓練を行う (第 3.3 節)。推論時には任意の固定値をスコアとして入力する方法と、スコアの上限值を推定して入力する方法によって文生成を行う (第 3.4 節)。

#### 3.1 発話文に対する応答文の情報量推定

本研究では入力発話文に対する応答文の情報量はそれぞれに含まれる単語どうしの共起の強さに相関するという直感に基づき、PPMI ベースの単純な共起尺度 MaxPMI を提案する。まず、訓練データ全体から単語どうしの PPMI を事前に計算する。単語  $x$  が発話文・応答文中に現れる確率をそれぞれ  $p_X(x)$ 、 $p_Y(x)$  とし、単語  $x, y$  がある発話応答対で同時に現れる確率を  $p(x, y)$  としたとき、PPMI は以下で計算される。

$$\text{PPMI}(x, y) = \max\left(\log_2 \frac{p(x, y)}{p_X(x) \cdot p_Y(y)}, 0\right)$$

ある発話文の単語列を  $X = \{x_1, x_2, \dots, x_{|X|}\}$ 、それに対する応答文の単語列を  $Y = \{y_1, y_2, \dots, y_{|Y|}\}$  としたとき、MaxPMI を以下で定義する。

$$\text{MaxPMI}(\mathbf{X}, \mathbf{Y}) = \max_{x \in \mathbf{X}, y \in \mathbf{Y}} \text{PPMI}(x, y)$$

ただし、モデルの学習時にはこれを  $[0, 1]$  の範囲に正規化したものを用いる。

#### 3.2 モデル構造

##### 3.2.1 エンコーダ

通常の seq2seq と同様、まず入力文中のトークンを Embedding 層でベクトル化した後、GRU によって入力文をエンコードし、ベクトル  $\mathbf{h}_{GRU}$  を得る。提案手法ではこれに加え、入力された  $\text{MaxPMI}(\mathbf{X}, \mathbf{Y})$  を多層パーセプトロン (Multi-layer perceptron; MLP) によって処理するスコアエンコーダを持つ。 $\mathbf{h}_{GRU}$  とスコアエンコーダの出力  $\mathbf{h}_s$  を連結したベクトル  $\mathbf{h}_e = \{\mathbf{h}_{GRU}; \mathbf{h}_s\}$  をエンコーダの出力とする。これにより、出力すべき文の情報量の目安をデコーダに与えることができる。

##### 3.2.2 デコーダ

デコーダの持つ語彙集合を  $V$  としたとき、ある単語  $v \in V$  と入力文  $X$  との単語共起スコア  $s_v$  を以下で定義する。

$$s_v = \sum_{x \in X} \text{PPMI}(x, v)$$

デコーダはまず、全ての語彙に対する単語共起スコアからなるベクトル  $\mathbf{v}_f \in \mathbb{R}^{|V|}$  を受け取り、これを多層パーセプトロン (Multi-layer perceptron; MLP) によって処理したベクトル  $\mathbf{h}_v$  を得る。これとエンコーダからの出力ベクトル  $\mathbf{h}_e$  を連結したものをデコーダの初期状態  $\mathbf{h} = \{\mathbf{h}_e; \mathbf{h}_v\}$  とする。これにより、デコーダは事前に入力と共起しやすい単語の情報を得ることができる。また、各タイムステップ  $i$  におけるデコーダの出力ベクトル  $\pi_i$  に対して  $\mathbf{v}_f$  を重み付きで足すことで、入力文との相互情報量の高い単語の出力確率を増幅させる (出力増幅器)。最終的なデコーダの出力  $\tilde{\pi}_i$  は以下の式で表される：

$$\tilde{\pi}_i = (1 - \lambda_i) \cdot \pi_i + \lambda_i \cdot \mathbf{v}_f.$$

ここで、 $\lambda_i$  は、デコーダの現在の中間状態  $\mathbf{h}_i$  に応じて以下のように計算される：

$$\lambda_i = \text{sigmoid}(W^{\text{gate}} \mathbf{h}_i + \mathbf{b}^{\text{gate}}).$$

ただし、 $W^{\text{gate}}$  は学習可能な重み行列であり、 $\mathbf{b}^{\text{gate}}$  はバイアス項である。

#### 3.3 Distant Supervision

学習時には発話応答対  $(\mathbf{X}, \mathbf{Y})$  に対して Distant Supervision を行うためのラベルとして  $\text{MaxPMI}(\mathbf{X}, \mathbf{Y})$  を計算し、これをモデルに入力することで学習を行う。なお、損失関数としては交差エントロピー損失を用いる。よって、モデルのパラメータを  $\theta$  とすれば、損失関数は以下で表される。ただし、 $\mathcal{D}$  は学習データセットである。

$$\mathcal{L} = \sum_{(\mathbf{X}, \mathbf{Y}) \in \mathcal{D}} \log P(\mathbf{Y} | \mathbf{X}, \text{MaxPMI}(\mathbf{X}, \mathbf{Y}); \theta).$$

#### 3.4 推論

$\text{MaxPMI}(\mathbf{X}, \mathbf{Y})$  は  $\mathbf{Y}$  が確定していない推論時には定義できない。そのため、 $[0, 1]$  の範囲の実数値  $s$  を入力して文を生成する (固定値デコード)。ここ

表 1: 実験結果

	BLEU-1	BLEU-2	NIST	distinct-1	distinct-2	length
提案手法 (最大化)	<b>22.30</b>	<b>17.62</b>	<b>2.87</b>	0.084	<b>0.41</b>	<b>11.89</b>
提案手法 ( $s = 0.5$ )	22.06	17.41	2.85	0.085	<b>0.41</b>	11.63
提案手法 w/o Score	19.53	14.91	2.62	0.085	<b>0.41</b>	10.19
提案手法 w/o Boost	18.05	13.34	2.16	0.098	<b>0.41</b>	11.51
Zhang ら	13.21	7.93	1.38	<b>0.099</b>	0.35	10.23
seq2seq	13.75	8.99	1.54	0.096	0.37	9.70

で, MaxPMI の性質上  $s$  が 1 に近ければ近いほど共起の高い語が生成されやすくなり, 情報量の高い文が出力されると期待できる. また, 入力文  $\mathbf{X}$  に対する MaxPMI の推定上限値  $s_{max}$  は事前計算した PPMI によって以下のように決定可能である:

$$s_{max} = \max_{x \in \mathbf{X}, v \in \mathbf{V}} \text{PPMI}(x, v).$$

よって,  $s_{max}$  を用いてデコードを行うことで  $\mathbf{X}$  に対してあり得る応答の中で情報量が最も大きい応答を生成できると考えられる (情報量最大化デコード).

## 4 実験

提案手法によって高頻度な応答を抑制し, 発話に対する妥当な応答が生成されることを検証するため, 雑談対話データを用いて応答生成の実験を行う.

### 4.1 実験設定

英語学習者どうしの対話からなる DailyDialog コーパスを用いてシングルターンの対話データセットを構築した. なお, train/validation/test データの分割は DailyDialog コーパス標準の設定に準拠している. データ数は train が 76,052, validation が 7,069, test が 6,740 である. また, 各発話・応答対は transformers<sup>1</sup> 内の BertTokenizer を用いてサブワード単位に分割した. PPMI の計算時はコーパス中で頻度が 50 以下の単語を除外した.

評価対象としては情報量最大化デコードを行った提案手法 (最大化) とスコアを  $s = 0.5$  に固定してデコードした提案手法 ( $s = 0.5$ ) を用いた. なお, スコアは validation データの BLEU が最大となる値を採用した. 比較対象としては通常の seq2seq モデルと, Zhang ら [4] の手法を用意した. なお, Zhang

らの手法においても推論時に任意のスコア  $s$  を入力する必要があるが, ここでは validation データでの BLEU が最大となった  $s = 0.7$  を用いた. また, 分析用に, seq2seq に対して提案手法のうちスコアエンコーダのみを適用した提案手法 w/o Boost と, 出力増幅器のみを適用した提案手法 w/o Score についても評価を行った.

学習時にはオプティマイザとして Adam [5] を用いており, 全てのモデルについて学習率は 0.0002 とし, その他のパラメータは推奨値を用いた. また, 勾配爆発の抑制のために gradient clipping を使用しており, 閾値は 5 とした. 全てのモデルについて, 隠れ層の次元数を 512, Embedding 層の次元数を 256 とした. 200 エポックにわたって訓練を行い, validation データでの BLEU が最も高かったモデルによって評価を行った.

### 4.2 実験結果

全体の評価は正解の応答文に対する BLEU, distinct [3], NIST [6] で行った. ここで distinct は, システムが生成した全ての生成文に含まれる  $n$ -gram のタイプ数を  $n$ -gram の総数で割ったものであり, システムが生成した  $n$ -gram の多様性を評価する指標である. NIST は BLEU をベースとした生成文と参照文との一致度を測る指標であるが, BLEU とは異なり  $n$ -gram に対して情報量の大きさを用いて重み付けを行うため, 機能語よりも内容語の一致が重視されるため, 発話に対する情報量と関連度を共に考慮した尺度であると言える.

実験結果を表 1 に示す. Zhang らの手法が distinct-1 を除いて seq2seq よりも低いスコアとなっているのに対し, 提案手法 (最大化) は distinct-1 以外の全ての評価尺度において最も高い評価値となっている. 特に, NIST においては提案手法 (最大化) が seq2seq に対して 2.3 ポイント上昇している. この結果から, Zhang らの手法で用いられている NIWF が単語の逆

<sup>1</sup><https://github.com/huggingface/transformers/> bert-base-uncased モデルを使用

表 2: 提案手法と比較手法の応答生成例

入力	mexican, japanese or chinese would be great.
提案手法 (最大化)	actually, i saw garlic ice cream on the rocks.
提案手法 (s = 0.5)	actually, it's my chinese dish.
Zhang ら	how about a facial cake?
seq2seq	well, there are lots of things.
参照文	oh, i know a place where you might like to eat.
入力	nonsmoking, please.
提案手法 (最大化)	nonsmoking. now, sir, does a single queen - size bed meet your approval?
提案手法 (s = 0.5)	nonsmoking. now, sir, does a single queen - size bed meet your approval?
Zhang ら	all right, ma'am.
seq2seq	all right. may i have your name?
参照文	next question, is a queen - size bed okay?

頻度ベースのスコアに基づいており、単語の多様性を直接的に向上させる効果を持つことがわかる。一方で、提案手法では発話文と応答文を共に考慮した MaxPMI を用いているため、単純に単語の多様性を向上させるだけでなく、発話文との関連度が高い応答を出力できていることがわかる。

提案手法 w/o Score と提案手法 w/o Boost については共に seq2seq や Zhang らに比べ NIST や BLEU, distinct-2 において高い評価値となっている。スコアエンコーダと出力増幅器はいずれも単体で発話に対する関連度と情報量を向上させる効果を持つことがわかる。また、これらと比べて提案手法 (s = 0.5) が高い BLEU, NIST 値であることから、両者を組み合わせることでさらに精度が向上することもわかる。しかし、提案手法 w/o Score の distinct-1 が seq2seq に比べて 0.014 ポイント低下している。これは PPMI の性質上コーパス上で出現頻度が低い語同士の共起が不当に高く評価されてよく出力されるようになった一方で、出現頻度が中程度の語があまり出力されなくなったためであると考えられる。

表 2 に生成例を示す。上段の例においては入力文が料理の話題であることに対応して提案手法 (最大化) は “garlic ice cream” や “chinese dish” などといった語を生成できている。下段の例においても入力文に対応した “nonsmoking” という単語を生成できているが、応答として不自然な文を生成してしまっている。生成文の前半で共起の強い単語を生成してしまったために、それ以降発話を無視してしまったものと考えられる。

## 5 おわりに

本研究では入力発話文に対して共起しやすい特徴的な単語が出力されやすくなるような応答生成器の構築と、文同士の共起度合いを示すスコアを学習データに自動的に付与する Distant Supervision の導入によって、入力文に対してより情報量の多い応答を生成する対話システムを構築した。今後は人手評価によって生成結果の総合的な評価を行う予定である。また、双方向デコーダ等の導入によって文脈を無視せずに共起の強い語を生成する手法を構築する予定である。

謝辞 本研究は JSPS 科研費 JP18K11435 および株式会社コトバデザイン の助成を受けたものです。

## 参考文献

- [1] Ilya Sutskever, Oriol Vinyals, and Quoc V Le. Sequence to sequence learning with neural networks. In *Proceedings of NIPS*, 2014.
- [2] Oriol Vinyals and Quoc V Le. A neural conversational model. In *Proceedings of ICML*, 2015.
- [3] Jiwei Li, Michel Galley, Chris Brockett, Jianfeng Gao, and Bill Dolan. A diversity-promoting objective function for neural conversation models. In *Proceedings of NAACL-HLT*, pages 110–119, 2016.
- [4] 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*, pages 1108–1117, 2018.
- [5] Diederik P Kingma and Jimmy Lei Ba. Adam: A method for stochastic optimization. In *ICLR*, 2015.
- [6] George Doddington. Automatic evaluation of machine translation quality using n-gram co-occurrence statistics. In *Proceedings of HLT*, 2002.