

批判的思考を促進する対話文生成に関する研究

吉田 聖[†] 馬 強^{††}

[†] 京都大学工学部情報学科 〒606-8501 京都府京都市左京区吉田本町 36-1

^{††} 京都大学院情報学研究科 〒606-8501 京都府京都市左京区吉田本町 36-1

E-mail: [†]yoshida.satoshi.62c@st.kyoto-u.ac.jp, ^{††}qiang@i.kyoto-u.ac.jp

あらまし 本論文では、話者の観点や思いを補える文章を自動生成する手法を提案する。既存手法の多くは、コンテキストに沿った流暢な対話文生成に着目しているが、話者の感情に対する考慮が不十分である。例えば、ネガティブの対話文のみを生成してしまったり、会話の内容が局所的になってしまったりする可能性がある。本論文では話者の感情を考慮して、その想いや観点を補える対話文を生成する手法を提案する。本提案手法によって生成された文章は、話者の感情を和らげたり、批判的思考を促進したりする効果が期待できる。

キーワード 対話文生成, 批判的思考, 感情分析

1 はじめに

近年、言葉でコミュニケーションを行う対話システムが普及している。Siri や Alexa など、スマートフォン、スマートスピーカー、カーナビゲーションシステムに搭載されるなど利用可能な対話システムが様々な形で提供されている。また対話システムに関する手法も多く提案されている。様々な情報が錯綜する現代において、情報を鵜呑みにせず、批判的に物事を捉える力が必要不可欠であるが、既存の研究は人間が書いたように自然な文章を生成することを目的としているのであって、使用者に再考させる機会を与え、批判的思考を促すものではない。

本研究では批判的思考を Ennis (1985) や Michael Scriven & Richard Paul (1987) の定義に従い、物事の問題を特定し適切に分析するための思考法とする [1], [2]。我々は感情的になって発言する際に、会話の方向性が偏ってしまい、視野が狭くなってしまふことがある。カウンセラーが相談者に対して行う会話を例に挙げる。カウンセラーは相談者にアドバイスを与えるといった形で会話をしない。あくまで相談者の立場から他にも見方があると提示し、相談者に再考の余地を与え、感情的に陥りがちな思考をその他の観点から考える機会を与えるのである。

批判的思考に関する研究も数多くある [3], [4]。しかし、既存の研究はメディアリテラシーなど一方的に提供される情報に対する批判的思考である。対話文という双方向性がある文脈において、相手の感情に注目しながらどのように文章が生成されるのかという研究は少ない。批判的思考を促進する可能な対話文生成する手法に関する研究が、我々の知る限りではまだない。

本研究では、批判的思考を促進する対話文生成手法を提案する。話者が発言した内容に関して、その文脈に即した内容でありながら、話者の感情的側面を推し量った上で再考させる文章を生成することを目的とする。例えば、"Room seemed clean but it smelled very musty." と部屋に否定的な発言に対して、"I want to leave a positive review, because the room had a neat look, and it was a nice quiet stay." と肯定的な対話文を返すのである。

提案手法では、まず、話者の発言から属性語と評判語を抽出する。属性語は話題の対象となるものである。例えば、主語や目的語などの名詞が挙げられる。評判語は属性語で表現する対象に対する話者の感情が表すものである。例えば、形容詞が挙げられる。この属性語と評判語の関わりや話者の感情を分析することによって、批判的思考を促す対話文を生成していく。

二つの単語の関係を測る尺度として、自己相互情報量を用いる。自己相互情報量は、二つの事象の間の関連度合いを測る尺度であるので、二つの単語がどれだけ共起しているかを計算することができる。本研究では属性語同士の共起する確率を Polarity Co-occurrence (P-Cooc)、属性語と評判語の共起する確率を Modification Co-occurrence (M-Cooc) と定義する。

流暢な文章を生成するのに寄与するのは P-Cooc である。P-Cooc では、ある感情の文章から異なる感情の文章を生成する属性語が共起する度合いを計算する。属性語の関係を異なる視点から生成される文章の検索に用いるのである。例えば、"Very good stay." "I guess you like breakfast." という会話があったとする。対象となる属性語は"stay"と"breakfast"である。これは"stay"に関して感想を述べる話者に対して、"breakfast"が関わっていると発言を返している。つまり、"stay"という発言があった時に"breakfast"に関する発言を行うと、同じホテルに関することを述べていながら別の視点から会話を進めることができる。

自然な文章を生成するのに寄与するのは M-Cooc である。属性語と評判語で共起率が高いということは、使用例の多い文章となるので、文章そのものを自然にする上で重要であると考えた。例えば、"room"という属性語に対して形容する言葉として"comfortable", "smoky", "small"などの評判語がある。その中で修飾する言葉として最も共起するものを選択すると、属性語を修飾するのに適当であると考えた。ある属性語を形容しやすい評判語の関係を分析することで文章生成に利用するのである。

対話文の感情極性の判定に Stanford Sentiment Treebank [5] を利用する。Stanford Sentiment Treebank では文章の極性を考えることができる。極性は Positive, Negative, Neutral の三つである。批判的思考は異なる視点から再考を促す思考法である。本研究

では異なる極性を持つ文章を生成して異なる視点を提供することによって、批判的思考を促すと試みる。

対話文生成が目的なので、生成した対話文が文脈上相応しいかを測る必要がある。そのため、二つの文章が連続したものであるかを測ることができる BERT の Next Sentence Prediction [6] を用いる。これにより、会話の流れを意識した文章を生成する。

つまり、本研究での提案する手法は次のようである。はじめに、入力された文章から属性語を抜き出し、P-Cooc 上で関係の深い属性語を導く。次に、P-Cooc を用いて得られた属性語と、M-Cooc 上で関係の深い評判語を導く。結果として、入力された文章に対する応答文として用いられる属性語と評判語を得ることができる。さらに、この二つの単語を含む文章をデータベース上で検索し抽出する。また、抽出される文章は Stanford Sentiment Treebank を用いて入力された文章に対して異なる感情的視点によるものを選択する。ここで得られた文章は入力された文章に対して、自然な文脈とは限らないため、BERT の Next Sentence Prediction を用いて選択された文章をランキングし、より文脈にあった文章を出力する。

本研究の主な貢献を次に示す、

- 批判的思考を促すための対話文生成手法を提案する。
- 極性を考慮した共起 (P-Cooc) と修飾を考慮した共起 (M-Cooc) を提唱して、話者の感情極性を考慮して評価物の多様な側面とその評判語を発見する。
- 20 人によるユーザ評価実験を行い、提案手法の有効性を検証した。

本論文の構成は次の通りである。2 節では対話文生成に関する関連研究と批判的思考に関する関連研究を示し、3 節では本研究で提案する手法について記す。4 節では本研究の評価実験について、5 節では結論を述べる。

2 関連研究

2.1 対話文生成

近年、対話文生成という分野は目覚ましい発展を見せている。

対話文生成において state-of-art である OpenAI が公開する GPT-2 (Generate Pre-Training 2) [7] がある。ある一定の文章を入力すると後続の文章を膨大な学習データを用いて予測するというものである。生成される文章は、極めて人間に近い文章となる。800 万の Web ページで学習を行っており、これによりドメインの異なるデータセットでも学習なしで最高精度に近い記録を出している。

しかしながら、上記に代表する深層学習で生成される文章は一つの文章を取ると極めて自然な文章ではあるのだが、文脈上自然ではない文章が生成されることもある。

ニューラル対話モデル (Neural Conversation Model; NCM) [8] とは、機械翻訳において用いられる Sequence to Sequence (seq2seq) [9] を対話システムにおける応答生成に適用した手法である。このような学習方法では、出力した単語が発話に対する応答として、参照文と同じ単語を出力しなければ損失が下がらないため、頻度の高い (正解しやすい) 汎用的な語を

多く出力してしまう。この問題の解決を図った手法の一つに、MMI-bidi [10] がある。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、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。高山ら [11] は訓練時に目的関数を追加して、発話文と強く共起する単語を生成できるようにすることで、MMI-bidi を適用する効果を高めた。

本研究では自己相互情報量と深層学習を利用して検索するというアプローチで、文脈上自然な文章でありながら、その上批判的思考という点に着目して対話文生成を行う。

2.2 批判的思考

批判的思考というと、否定的に物事を見るというような誤解を受けるかもしれないが、そうではない。批判的思考の定義は学者によって異なる。例えば、Ennis (1985) や Michael Scriven & Richard Paul (1987) [1], [2] の両者の定義に共通しているのは思慮深く思考し、論理的に分析することである。

批判的思考力を持つ人物がどのような過程を経て文章から情報を得ているかという研究がある [3], [4]。批判的思考力の高い被験者は、テキスト間の対立や矛盾を理解するために、複数記事の内容を統合しながら閲覧すると考えられる。このような被験者は矛盾や対立といった意見に注目する可能性が高い、とされる。

これらの研究は批判的思考はどういった過程で行われているかを明らかにしている。批判的思考を促進するための対話文生成手法を検討するという点も本研究の特徴である。

3 提案手法

3.1 概要

全体の流れを図 1 に示す。

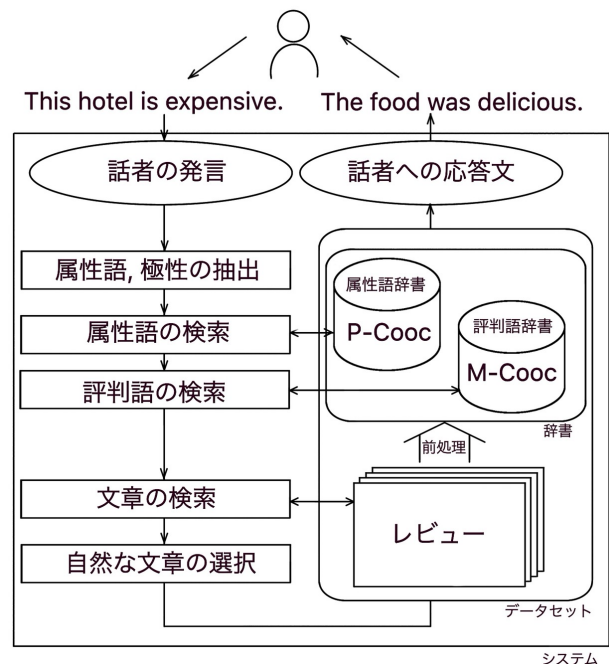


図1 全体図

最初の対話文を入力する。例えば, "This hotel is expensive" (図 1) とする。この対話文に対して, 別の側面から述べる文章を生成して, 批判的思考を促す文章を生成する。

次に話者の発言から属性語と極性を抽出する。入力する文章を "The hotel is expensive" とすると, 属性語は "hotel", 極性の結果は "Negative" である。なお, この時の評判語は "expensive" である。得られた属性語は異なる極性の文章の属性語の中で共起しやすいものを得るため, 属性語の検索へ進む。入力文の極性が "Negative" であるため, その逆の "Positive" 極性の文章を検索・生成して批判的思考を促す。

属性語の検索で用いるのは P-Cooc である。P-Cooc では, 入力された属性語と共起率の高かった属性語を出力する。入力に対して上位複数の単語を出力する。図では "hotel" という入力に対して, 例えば "staff", "food" や "view" と言った共起率の高い属性語を出力する。こうして得られた属性語をその属性語を形容する評判語を得るため, 評判語の検索へ進む。

評判語の検索で用いるのは M-Cooc である。M-Cooc では, 入力された属性語と共起率の高かった評判語を出力する。入力に対して上位複数の単語を出力する。例えば, P-Cooc で 5 つ, M-Cooc で 5 つ出力されたとすると, この時点で属性語と評判語のペアは 25 個あるということになる。"food" に対して "delicious" や "cheap", "staff" に対して "helpful" や "welcoming", "view" に対して "beautiful" や "amazing"。ここで得られた属性語と評判語のペアを用いて文章を検索するため, 文章の検索へ進む。

文章の検索では属性語と評判語のペアと極性の結果を入力とする。提示されうる文章の候補は複数ある。入力の属性語と評判語のペア一つあたり, 複数の文章を出力する。例えば "food" と "delicious" を含む文章がデータの中にあり, "The food was absolutely delicious" という文章で, 極性を "Positive" であるとする。この文章は入力文章である "This hotel is expensive" の極性が "Negative" であるので得られる文章の一つとなる。また "staff" と "helpful" を含む文章がデータの中にあり, "The staff was not helpful" という文章で, 極性を "Negative" であるとする。この文章は入力文章の極性が "Negative" であるので得られる文章としない。こうして最大 (候補とする属性語の数) * (候補とする評判語の数) * (候補とする文章の数) だけ候補の文章が得られる。得られた文章を入力文章の次に来る文章として自然かどうかを判定するため, 自然な文章の選択をする。

自然な文章の選択では, BERT の Next Sentence Predictor を用いてランキングする。BERT では二つの文章が連続する文章であるかを表現する。1 に近いほど連続する文章である可能性が高く, 逆に 0 に近いほど連続した文章ではないとされる。ここでは連続する文章の度合いを全ての文章で計算し, 最も大きい値を出した文章を最終的な出力として提示する。

3.2 定義

この手法に用いる概念は次の 4 つである。

3.2.1 属性語

本研究では属性語は主に主語や目的語になりうる名詞とする。Stanford Sentiment Treebank における係り受けタグでは, obj

(目的語), nsubj (主語名詞), compound (複合名詞) となるものを使用している。属性語の極性を, それを含む文章の極性とする。ここでは極性を Positive/Negative/Neutral¹ としている。Pos, Neg, N はそれぞれ文章がポジティブなものか, ネガティブなものかニュートラルなものかを表す。

本研究において, 属性語は会話の中において物事の文脈であったり, 会話の対象になるのではないかと仮定した。

3.2.2 評判語

評判語は属性語にかかるもので, 主に形容詞を指す。自然言語処理における係り受けタグでは, amod (形容詞修飾語), conj (形容詞), root (文の根) となるものを使用する。

3.2.3 P-Cooc

P-Cooc (Polarity Co-occurrence) は逆の極性からよく言及されている属性を効率検索するための極性付き共起辞書である。P-Cooc において, 共起の対象となるのは, Pos と Neg のペア, つまり極性がポジティブとネガティブのペアになるものである。この辞書に単語を入力すると, 属性が異なるもののうち共起する確率の高い単語を返すことができる。ここでは一つの入力に対して共起率の高い単語を複数出力する。

文脈を維持したまま別の観点から会話を行うこと, これが P-Cooc を用いる目的である。例えば, ホテルの金額に対して不満に対して, そのサービスの良さを示すことで, その感情を和らげていく。

3.2.4 M-Cooc

M-Cooc (Modification Co-occurrence) は属性と評判語の修飾共起関係を調べられる辞書である。この辞書に単語を入力すると, 属性語を共起する確率の高い単語と共に出力する。ここでは一つの入力に対して共起率の高い単語を複数出力する。

ある属性語を修飾するのに最も適当な会話文を作成すること, これが M-Cooc を用いる目的である。

3.3 共起辞書の構築

3.3.1 P-Cooc

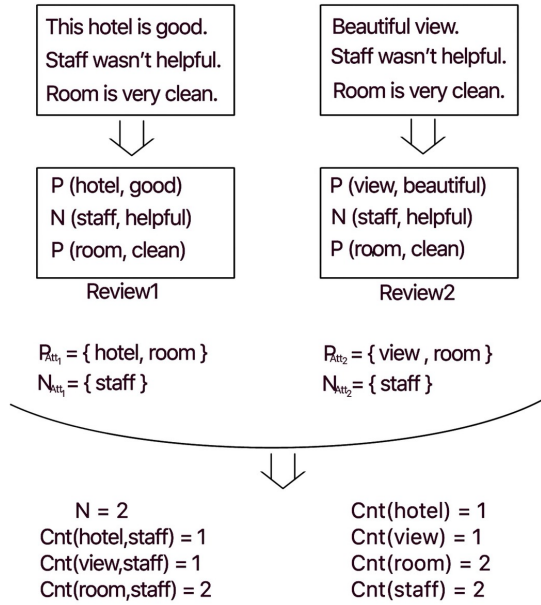
ここでは, P-Cooc の計算手法について述べる。P-Cooc で求めるのは異なる極性の文章に含まれる属性語が共起する自己相互情報量である。本研究では単語が共起する度合いを測る尺度として, 自然言語処理で用いられる自己相互情報量を使用した。

P-Cooc で用いる式について説明する。Cnt は引数が一つの時と二つ取る時とがある。引数は属性語である。引数が一つの時は引数となる属性語の出現回数を得る。引数が二つの時は引数となる属性語が共起した回数を得る。以上の定義を用いて, 自己相互情報量の定義をすると以下になる。数式内の A, B はいずれも属性語である。N はレビューの総数である。

$$P\text{-Cooc}(A, B) = \log_2 \frac{\text{Cnt}(A, B) \times N}{\text{Cnt}(A) \times \text{Cnt}(B)} \quad (1)$$

自己相互情報量を用いているため, 滅多に出ない単語の組み合わせを排除する必要がある。排除しなければ使用頻度の低いものを高い水準で推薦する可能性があるためである。

1: これ以降, Positive, Negative と Neutral をそれぞれ Pos, Neg, N で記す。



$$P\text{-Cooc}(\text{hotel}, \text{staff}) = \log_2 \frac{\text{Cnt}(\text{hotel}, \text{staff}) \times N}{\text{Cnt}(\text{hotel}) \times \text{Cnt}(\text{staff})} = \log_2 \frac{1 \times 2}{1 \times 2} = 0$$

図2 P-Cooc の計算例

図2に示すように、P-Coocではある共通の文脈を持つ属性語同士の共起率を求める。P-Coocではある程度の文脈を捉えるため、一つの文書に複数の文章を持つとする。その一つの文書単位をReviewとする。また、文章にはそれぞれ属性語が含まれる。文章の極性がPositiveのものはP、NegativeのものはNと表記している。一つのレビューで共起する単語のペアを数え上げる。図に示すCnt(A,B)では呼び出される回数を記録している。 P_{Attr} は極性がPositiveな文章に含まれていた属性語の集合、 N_{Attr} は極性がNegativeな文章に含まれていた属性語の集合である。Nはレビューの総数である。

3.3.2 M-Cooc

ここでは、M-Coocの計算手法について述べる。M-Coocで求めるのは属性語と評判語の共起する自己相互情報量である。本研究では単語が共起する度合いを測る尺度として、P-Coocと同様に自己相互情報量を使用する。

自己相互情報量を用いているため、滅多に出ない単語の組み合わせを排除する必要がある。排除しなければ使用頻度の低いものを高い水準で推薦する可能性があるためである。

図3に示すように、M-Coocでは全ての文章を属性語と評判語のタプルで表現する。そして全ての文章から得られたタプルを用いてM-Coocにおける自己相互情報量は求まる。

また、M-Coocの引数は属性語Aと評判語Bのペアであり、以下のように計算される。

$$M\text{-Cooc}(A, M) = \log_2 \frac{C(A, M) \times N}{C(A) \times C(M)} \quad (2)$$

Cは引数が一つの時と二つ取る時とがある。引数が一つの時は引数となる単語の出現回数を得る。引数が二つの時は引数と

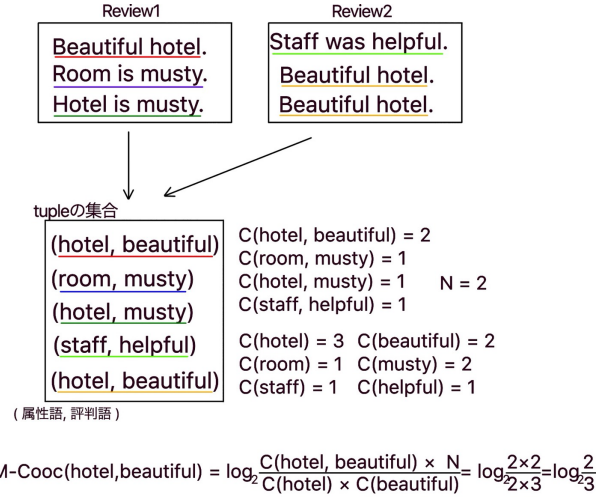


図3 M-Cooc の計算例

る属性語と評判語が共起した回数を得る。Nはレビューの総数である。

3.3.3 対話文生成

(1) 属性語と極性の抽出

文章の極性判定と属性語・評判語の抽出はStanford Sentiment Treebank (SST)を用いている。SSTは入力された文章を形態素解析したり、文章そのものにPos/Neg判定を行う事ができる。

(2) 属性語の問合せ

抽出した発言の属性語をP-Coocを通して、次の文章を作る生成文の属性語を得る。例えば、"hotel"という入力があれば、"staff", "room"などを出力で返す。

(3) 評判語の問合せ

得られた生成文の属性語それぞれに対して、M-Coocで共起する確率の高いものを選出する。属性語と評価語の共起率が高いものを選出することによって、属性語が用いられる際によく使用されやすい組み合わせを選ぶことができる。

例えば、"room"という言葉には"clean"や"comfortable", "spacious"など様々な評判語が付随するものの内、最も共起しやすい組み合わせをM-Coocで調べる。

(4) 文章の検索

本研究では、ホテルレビューのDataset in Datafinet [12]をコーパスとして利用する。データセット内にある文章数は147236個である。文章はピリオドを一区切りとしている。例えば"Nice and clean. Rooms are spacious with air conditioning. Good selection of breakfast."というレビューがあったとき、それぞれ一つの文章として捉えるのでこの例では三つの文章があることになる。

発言の極性と、評判語の問合せで得られた属性語と評判語を用いて検索する。例えば、極性の入力として"Negative", 属性語と評判語のペアを"food", "delicious"とする。入力の極性が"Negative"なのでその反対である"Positive"な文章かつ"food"と"delicious"という単語を持つ文章を検索する。その結果、"The food was absolutely delicious"という文章を候補として選び出される。また、属性語と評判語のペアとして"staff"と"helpful"が挙げられた

としても、文章の極性の結果が同じものは候補文章として出力されない。例えば、”The hotel staff was not helpful.”は候補の文章として得られていない。

(5) ランキング:自然な文章の選択

(4) の検索によって、複数の候補文が得られ、批判的思考の文章候補が選出された。しかしながら、文章の流れを鑑みているとは言えない。そのため自然な文章の選択に投入して話者の発言に対して次に来る文章として相応しいものを選択する。

BERT の next sentence prediction [6] を用いることで、どの文章が入力された文章の次に来る文章かを測る。BERT は Masked Language Model と Next Sentence Prediction のタスクを学習した言語モデルである。本研究では主に Masked Language Model を使用している。これは二つのセンテンス間の関係性を言語モデルに学習させるものである。連続する二つの文章とランダムな二つの文章を半分ずつ入力に入れることで学習する。

Next Sentence Prediction を用いる事により、出力された文章はその自然さや妥当さを保証していることとなる。評価値は 0 から 1 までの値を取り、1 に近い程次に来る文章として自然な文章である。BERT を通し、最も自然度の高い文章を出力する。

4 評価実験

4.1 概要・目的

提案手法を用いて批判的思考を促進する文章が生成できているかを評価するため、アンケートを用いて評価を行った。今回の手法で適用している言語は英語であるため、英語に普段から慣れ親しんでいる英語研究会²に所属する日本人 20 名に協力していただいた。アンケートでは次の例で示す二人による会話として出題した。

• Mike: This hotel is way over rated, it's nothing like the description or the reviews.

• Jane: This is a good mid priced placed to stay.

Mike の文章はレビューの中から適当に取り出した文章で、Jane の文章が提案手法を用いて出力した文章である。

評価する項目は次の三つである。

- 会話が会話として成り立っているか
- 二人目の文章が文章として成り立っているか
- 批判的思考が来ているか

以上を本研究を評価する項目とする。また、対照実験を行うため入力に当たる Mike の文章を用いて GPT-2 で生成されたものを出力した。入力に用いた文章は 13 個あり、そのうち Positive な文章は 7 つ、Negative な文章は 6 つである。実際に用いた文章は以下の表 (1) にある。最低値が 1、最高値が 7 の七段階評価を行ってもらった。

4.2 データセット

Datafiniti's Business Database が公開している、1000 ものホテ

表 1 実験に用いた文章

Room was well outfitted and bed was very comfortable.
The resort was beautiful, well kept, and the room was luxurious without being opulent.
Good experience and very convenient to where I needed to be.
Beautiful hotel.
Staff was great and rooms were clean and very comfortable.
We only stayed one night but this is a good place to stay.
Room seemed clean but it smelled very musty.
Very loud rowdy customers staff would not do anything about the frat like partying.
Room was not clean.
Nice hotel , with very friendly staff and helpful - great choice for breakfast.
The king size bed was the most uncomfortable bed I had slept in for decades.
This hotel is way over rated, it's nothing like the description or the reviews.
The breakfast was adequate and the housekeeping staff were friendly.

表 2 属性語と評判語

単語	総数	単語数	使用する総数	使用する単語数
属性語	181483	9063	170986	2337
評判語	162722	6751	154912	1830

表 3 Neg/Pos の自己相互情報量

Neg, Pos	自己相互情報量
room, inconveniences	0.6449663977057334
hotel, ports	1.0317250913592815
staff, sabrina	0.6833845190889121
breakfast, shores	1.0783373263749623

ルのレビューをまとめたものをデータセットとして用いている [12]。Datafiniti's Business Database のデータセットでは、住所やホテル名を始め、評価点数やホテルのレビューを含んでいる。数ある項目の中でも今回はレビューデータを用いる。用いるレビューは 35912 にも及んでいる。レビューに含まれる文章は 147236 ある。この節では、前処理、P-Cooc、M-Cooc、でどのような処理をしているかを述べる。

4.2.1 前処理

用いているレビューの中で属性語として得られた単語は 9063 個、総数は 181483 個存在した。属性語の中で出現頻度の上位は次の表に挙げる。また同様に評判語として得られる単語は 6751 個、総数は 162722 個存在した。評判語の中で出現頻度の上位は次の表に挙げる。

自己相互情報量を用いているため、滅多に出ない単語の組み合わせを排除する必要がある。排除しなければあまり使用されていないものを高い水準で推薦する可能性がある。よって、頻度の少ない単語を除外する。今回は出現回数が 5 回未満のものを除外した。結果、属性語は 2337 個、評判語は 1830 個となった。

ここで、属性語と評判語の数の差を表にしてまとめると以下のようなになる。

4.3 P-Cooc の構築に関する考察

この節では Polarity Co-occurrence において、データセットが実際にどのような値を出したかを確認する。今回、P-Cooc に用いた自己相互情報量の求め方は二つ。P-Cooc の図 (1) にて表したように、極性を組み合わせで Neg/Pos としたもの、Pos/Neg としたものである。以下に前処理で述べた単語の中で自己相互情報量の最も高いものを掲載する。

2 : (<https://kuess.rakusaba.jp/wordpress/>)

表 4 Pos/Neg の自己相互情報量

Pos, Neg	自己相互情報量
room, controls	0.9080008035395273
hotel, radios	1.7946858940584323
staff, omelet	1.2683470198100684
breakfast, adrenaline	2.6632998270961186

表 5 属性語同士の共起回数

入力単語	単語	出現回数	入力単語との共起回数
room	inconveniences	15	3
hotel	ports	40	22
staff	sabrina	5	2
breakfast	shores	6	2
room	controls	28	3
hotel	radios	15	14
staff	omelet	5	3
breakfast	adrenaline	6	6

P-Cooc では入力単語に対して上位に位置する単語を出力とする。例えば, breakfast という入力に対して shores, available, innkeepers, residence, furnishingなどを挙げる。Pos/Neg も同様である。

表 3 について, 一見 inconveniences という文章を Negative に導きそうな単語がある。これは Stanford Sentiment Treebank の極性を分析する結果が正しくない場合が考えられる。例えば, "In my opinion, all those things arent too bad, but more so inconveniences." という文章は Positive なものとして判定された。使用した Stanford Sentiment Treebank の論文によると極性判定の正答率は 8 割程度であるためである。勿論文章に Never など否定語が付いて文章が Positive であると判定されることもある。

文章の感情判定結果が Negative から Positive を導く属性語の共起率, Positive から Negative を導く属性語の共起率を表 4 に記した。例えば"hotel"によく付随する単語は"radios"という直感とは少し異なる結果となった。ここで P-Cooc, M-Cooc から選出された単語の出現数と共起回数を表 5 にまとめる。

半数は出現頻度の下限である 5 に近い回数のもので, 2 桁回数のもので共起するものの候補として挙がっていた。出現頻度が少なすぎる単語を排除するため, 出現回数が 5 未満のものを排除したが, 出現頻度の下限をもう少しあげるのが良いと思われる。しかしながら出現頻度の下限を, 例えば 10 回とすると対象となる単語が 1156 個になり, 20 回とすると 685 個になる。

出現頻度の下限を引き上げた場合, 例に出した 4 単語と共起する単語がどのように変化するかを考察する。

下限を引き上げるほど共起率は負の方向へ大きくなる傾向があるが, 共起するものとして直感的に妥当だと思えるものもある。しかしながら, 出現頻度が下限に近い単語が選出されやすいという問題は残っている。自己相互情報量は, 共起する確率を求めるが, 出現頻度を考慮できているとはいえない。例えば, 単語"staff","room"の出現頻度が共に 10 であり, 共起する回数が 10 である場合と, 単語"hotel","breakfast"の出現頻度が共に 1000 であり, 共起する回数が 1000 である場合を考える。自己相互情

表 6 Neg/Pos で出力される属性語出現回数の下限を変化させた時

入力単語 (下限)	出力単語	共起率	出現回数	共起回数
room (5)	inconveniences	0.64	15	3
room (10)	art	-0.09	10	3
room (20)	recommendation	-0.36	24	6
hotel (5)	ports	1.03	40	22
hotel (10)	ports	1.03	40	22
hotel (20)	ports	1.03	40	22
staff (5)	sabrina	0.68	5	2
staff (10)	cake	-0.11	13	3
staff (20)	welcome	-0.42	27	5
breakfast (5)	shores	1.08	6	2
breakfast (10)	innkeepers	0.93	10	3
breakfast (20)	love	-0.51	27	3

表 7 属性語, 評判語の自己相互情報量

属性語, 評判語	自己相互情報量
room, stunk	0.6449663977057334
hotel, weary	1.0317250913592815
staff, accomadating	0.6833845190889121
breakfast, worldwide	1.0783373263749623

表 8 属性語と評判語の共起回数

属性語	評判語	評判語の出現回数	属性語との共起回数
room	stunk	5	5
hotel	weary	30	22
staff	accomadating	496	301
breakfast	worldwide	6	6

報量の式 (1) に当てはめると, 後者の方が小さい値となる。

また, 負の値が出てしまっているが負の値は自己相互情報量において共起しにくいという意味合いを持つ。単語によって特定の単語と共起しにくいものもありうるという結果になった。

今回の実験ではより多様な表現を持たせるために出現頻度の下限を 5 回にしたものを使用した。

4.4 M-Cooc の構築に関する考察

この節では Modification Co-occurrence において, 実際にどのような値を出したかを確認する。M-Cooc の図表にて表したように, 属性語と評判語が共起する度合いである。以下に前処理で述べた単語の中で自己相互情報量の最も高いものを掲載する。

P-Cooc 同様, 直感的に共起するとは思えない単語群が提示された。これらを更に分析するため, 選出された単語の出現数と共起回数を表 8 にまとめる。

半数は下限である 5 に近い回数のもので, 2 桁以上回数のもので共起するものの候補として挙がっていた。出現頻度が低いものを提示する傾向にあるのは P-Cooc と同様の傾向にある。しかしながら, 出現回数に比べて共起回数が属性語同士よりも多いようである。

これは, 表 3 にあるように属性語と比べて評判語の使用されるものが限られている可能性があるからだと言える。例えば, 属性語 1 単語辺りに対して用いられた回数の期待値は, 約 73.2 回なのに対して, 評判語では 84.7 回である。

表9 M-Cooc で出力される属性語出現回数の下限を変化させた時

入力単語（下限）	出力単語	共起率	出現回数	共起回数
room（5）	stunk	3.32	15	3
room（10）	non-smoking	2.92	133	101
room（20）	non-smoking	2.92	133	101
hotel（5）	weary	1.03	30	22
hotel（10）	weary	1.03	30	22
hotel（20）	weary	1.03	30	22
staff（5）	accomadating	0.68	496	301
staff（10）	accomadating	0.68	496	301
staff（20）	accomadating	0.68	496	301
breakfast（5）	worldwide	4.33	6	6
breakfast（10）	fruit	4.23	15	14
breakfast（20）	cereal	4.16	26	23

表10 実験結果

手法	流暢度	自然さ	批判的思考
GPT-2	3.93	4.70	3.27
提案手法	4.07	4.32	3.91

評判語も同様に下限を引き上げた場合例に出した4単語（room, hotel, staff, breakfast）と共起する単語がどのように変化するかを考察する。

属性語と評判語の場合は下限を引き上げた方がより出現頻度も共起する回数が多い単語を選出できるので、こちらの下限を5ではなくある程度引き上げるべきだと言える。

また、表の中で“fruit”のように、評判語に来るべき所に属性語と受け取れる単語が存在する。これはレビューがそもそも会話文ではなく、短文かつ文構造をなしていないものがあるためである。例えば、fruit を評判語として取られる文章が次のようなものが存在する。

“Continental breakfast was nice and fresh, fruit and baked goods with.”(3)

この文章で“fruit”は“nice”, “fresh”に並列なものとして扱われており、評判語として扱われてしまっているのである。今回、属性語として出現している“fruit”は119個あり、誤って評判語に加えられているのは表9の通り15個である。いずれも例のように“breakfast”の並列として挙げられているものが誤って評判語として数え上げられているため、共起率が極めて高い値を出してしまったと思われる。対策としては次のことが考えられる。前処理で単語を判定する時、属性語と評判語どちらにも取る可能性がある。そこで、属性語、または評判語として回数が多く数え上げられている方を正しく使用されているものとして登録しておき再び数え直す。例えば、上記のように“fruit”は属性語として多く用いられているので、例(3)の文章が出て属性語として扱うようにするという方法である。

4.5 提案手法の結果・考察

全ての質問と各項目にて平均をとると表10のようになる。

有意差を検定するため、対応ありの両側 t 検定を用いた。帰無仮説は、比較手法と提案手法に差は無いとする。有意水準は $\alpha = 0.05$ で用いるのは自由度19の t 分布である。次の表に載せ

表11 t 検定の p 値

流暢度	自然さ	批判的思考
0.433	0.013	0.010

表12 実験結果

手法	入力文章の極性	流暢度	自然さ	批判的思考
GPT-2	positive	3.86	4.74	3.14
提案手法	positive	4.13	4.33	3.99
GPT-2	negative	4.01	4.66	3.43
提案手法	negative	3.98	4.30	3.83

表13 p 値

入力文章の極性	流暢度	自然さ	批判的思考
Positive	0.225	0.043	0.008
Negative	0.561	0.119	0.346

るのは p 値である。p 値が 0.05 以下で帰無仮説を棄却でき、提案手法とベースライン手法の有意差が得られたことが分かった。

文章の流暢さはどちらの手法も変わらないが、文章の自然さと批判的思考の項目は大きく差異が出る結果となった。t 検定の結果も自然さと批判的思考において有意差が現れる結果となった。また、それぞれ入力の記事にあたるものを positive と negative とで分けた場合は以下になる。

文章の自然さはいずれも同様の差が出ているが、批判的思考に関しては入力が positive な文章の方が大きく差が出る結果となった。また入力が positive なものに関しては t 検定による、有意差が認められたが入力が negative なものには有意差が見られなかった。これは先に述べたように、入力がもともと negative なものにも関わらず、誤って positive だと判定してしまい、対話文の結果が positive から positive となり批判的思考をするに至らなかった可能性がある。

各項目について詳細に考察していく。

4.5.1 流暢度

いずれの表においても流暢度には大きな差は見られなかった。これは提案手法は GPT-2 と同様の手法である、BERT の Next Sentence Prediction がうまく作用している結果だと判断できる。

今回提案手法で出力された文章が入力される文章の度合いの平均はほぼ1に近い値であった。これは、BERT の Next Sentence Prediction においては次に来る文章として相応しい文章を選出できていることを表す。また、P-Cooc, M-Cooc を通して得られた属性語と評判語のペアが入力された文章に対する回答として、流暢なものとなる一助になっていると言えるのではないだろうか。

4.5.2 自然さ

本研究で使用されている文章は全て人間が文章によるものであるにも関わらず、文章の自然さはいずれも GPT-2 が有用であるという結果になった。この結果はデータセットが全てレビューによる事が起因していると考えられる。

通常、文章とは日本教育において主語動詞名詞などといった定型的文章がある。しかしながら、レビューではしばしば砕けた表現が見つかる。例えば、“Excellent view.”など端的に書かれていたり、“Nice hotel, with very friendly staff and helpful - great

choice for breakfast , something for everyone.” というように、意図は伝わるが文章であるかという問いに簡単に答えられないものがある。特に今回の実験で最も評価の低い点数を取ったものは”Hotel was great staff was happy n positive energy everywhere.”という、文章であった。この文章がレビュー文として入力された経緯は推察できないが、このように文章の体をなしていないようなものは省けるようにするべきである。

4.5.3 批判的思考

本研究の目的である、批判的思考を促す対話文を生成するという観点では GPT-2 より良いという結果が出た。

ある項目では 1 から 7 まで偏りなく値をとるものがあった。これは、被験者が読む文章は各項目 2 つであるため背景を被験者が補完したのではないかと考えられる。批判的思考は、今までの文脈を正しく理解した上で別の観点による意見が出る事で対話に寄与するものであるが、今までの文脈を無視した上でなされる対話は批判的思考とはならない。

また、今実験で最も高い評価を得た文章は以下のようである。

Room seemed clean but it smelled very musty.

I want to leave a positive review, because the room had a neat look, the staff was friendly, and it was a nice quiet stay.

以上の文章は文章同士の対比が明確でありながら、簡潔である。一つ目の文章と二つ目の文章の感情の対比が明確であり、二つ目の文章がカンマ区切りでそれぞれの観点から簡潔に述べている。批判的思考の平均点が 3.91 であったのに対し、この文章は 5.55 とかなり高い水準の値が出た。

4.5.4 考察

提案手法ではレビューの文章から相応しいものを選択して出力する。文章を属性語と評判語を持つと仮定した本研究の手法を用いて得られた文章は比較手法である GPT-2 よりも批判的思考力のある文章を出力できるという結果を出した。しかしながら、用いたデータセットがレビューの文章であるため中には会話文に入るものとしては不適切と思われるものがあった。

また、対話文という形をとるのであれば入力された文章に関して少し言及し、その上で別の観点を述べれば話題の対象となるものの対比関係がわかりやすく比較をしやすくなるのではないだろうか。中には複数のレビュー文をただ貼り付けただけの印象を受ける項目も見受けられた。

また、批判的思考をする上で重要なことは文脈を正しく理解する事が重要であるが、本手法でうまくいったのはホテルのレビューという概ね文脈を共通する前提のもと行われている可能性がある。そのため、更に様々な分野や状況で本研究で提案した手法が有用であるかを検証する必要がある。

5 結 論

本研究では文章の属性語と評判語に着目し批判的思考を促進する手とを提案した。提案手法では、文章からその主体となる属性語とそれに対する評判語を抽出して、それらの関係を分析することによって批判的思考のある文章を生成した。英語研究会に所属する被験者 20 名に対し、各手法について実証実験を行い、

批判的思考を導く事が示唆された。

今後の課題としては大きく分けて以下の 3 点が挙げられる。

(1) 文章の流暢度に関する課題

実験を進めていく中で複数のレビュー文を貼り付けただけに思えるものがあった。これを解決するには文章 1 と文章 2 において話題に挙げられているものが明確に共通している必要がある。そのため、入力される文章で提示されているものに少し言及するなどといった工夫が求められる。

(2) 批判的思考に関する課題

今回用いたデータセットも実験に用いた文章もホテルに関する文章を用いたものである。今回批判的思考が出来ていると結果が出たのは、文脈が共通しているものを用いたことが影響している可能性がある。そのため、更に多くの話題や状況で用いられる文章を用いて提案した手法が有用であるかを見直す必要がある。

(3) 極性の方向に関する課題

今回属性語の関係を考えるときに、極性が反対の文章に含まれる属性語を結ぶだけでなく、その方向性も意識したものとなっている。つまり、Negative な文章に含まれる”hotel”と Positive な文章に含まれる”hotel”は別の単語としたのである。追加実験として極性が反対であることのみを考慮した共起辞書を作成し、比較する。

謝 辞

本研究の一部は科研費（19H04116）による。

文 献

- [1] Robert H Ennis. A logical basis for measuring critical thinking skills. *Educational leadership*, 43(2):44–48, 1985.
- [2] Michael Scriven and Richard Paul. Defining critical thinking. In *8th Annual International Conference on Critical Thinking and Education Reform, Summer*, 1987.
- [3] 平山るみ and 楠見孝. 批判的思考態度が結論導出プロセスに及ぼす影響. *教育心理学研究*, 52(2):186–198, 2004.
- [4] 楠見孝 and 松田憲. 批判的思考態度が支えるメディアリテラシーの構造. In *日本心理学会大会発表論文集 日本心理学会第 71 回大会*, pages 1PM085–1PM085. 公益社団法人 日本心理学会, 2007.
- [5] Richard Socher, Alex Perelygin, Jean Wu, Jason Chuang, Christopher D Manning, Andrew Y Ng, and Christopher Potts. Recursive deep models for semantic compositionality over a sentiment treebank. In *Proceedings of the 2013 conference on empirical methods in natural language processing*, pages 1631–1642, 2013.
- [6] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.
- [7] Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. Language models are unsupervised multitask learners. *OpenAI Blog*, 1(8):9, 2019.
- [8] Oriol Vinyals and Quoc Le. A neural conversational model. *arXiv preprint arXiv:1506.05869*, 2015.
- [9] Ilya Sutskever, Oriol Vinyals, and Quoc V Le. Sequence to sequence learning with neural networks. In *Advances in neural information processing systems*, pages 3104–3112, 2014.
- [10] Jiwei Li, Michel Galley, Chris Brockett, Jianfeng Gao, and Bill Dolan. A diversity-promoting objective function for neural conversation models. *arXiv preprint arXiv:1510.03055*, 2015.
- [11] 荒瀬由紀 高山隼矢. 自己相互情報量を用いた特徴語彙予測に基づく雑談応答生成. pages P3–34, 2019.
- [12] Dataset in datafiniti, access 12/20.