

対話エージェントにおける動向情報を用いた 話題転換に関する予備的検討

Preliminary Study on Topic Change by Conversational Agents Based on Trend Information

瀧口 慈勇¹ 高間 康史^{1,2}

Satoru Takiguchi¹, Yasufumi Takama^{1,2}

¹ 首都大学東京大学院システムデザイン研究科

¹ Graduate School of System Design, Tokyo Metropolitan University

² 首都大学東京システムデザイン学部

² Faculty of System Design, Tokyo Metropolitan University

Abstract: This paper proposes a method for conversational agents to change topic during conversation based on trend information. Detection of dialogue breakdown is one of important research themes for developing conversational agents. However, few studies have been done on how to do after detecting dialogue breakdown. The proposed method finds a topic that relates with current topic in terms of temporal trend by using context search engine. This paper also proposes a method to detect dialogue breakdown based on response time by a human. The results of preliminary experiments are reported to examine the possibility of the proposed methods.

1 はじめに

近年、介護や接客などへの利用目的で人間と対話を行うエージェントの研究が活発に行われるようになってきている[1]。対話エージェントの中でも、何らかの目的を持って対話を行うタスク指向形対話エージェントに関する研究は一定の成功を収めている[2]。しかし、特に目的を持たず雑談を行う非タスク指向形の対話エージェントは、NTTドコモの雑談対話API¹など手軽に利用可能なものも増えてきているが、研究はまだ発展途上の段階にあるといえる。対話エージェントが行う雑談においては幾つかの課題があり、その一つが適切に発言の意味を理解できないと不自然な応答を返してしまう場合(対話の破綻)が存在する事である。この原因としてタスク指向型においては限られた範囲での話題を扱えばよいのに対し、非タスク指向形においては多様な話題を扱う必要があるため、個々の話題に対する準備が不十分になってしまうことが挙げられる[3]。既存の対話エージェントを利用する研究で多く用いられている簡

易的な対話エージェントにおいて特にこの問題は大きい。

本稿では、簡易的な対話エージェントを対象として、対話が破綻した場合の検出手法、および話題を継続させるための対策として話題転換を行う手法について提案する。対話破綻の検出に関しては、話者の応答までの時間と応答の長さの変化に着目し、これらに基づく判断指標を提案する。話題転換については、元の話題とある程度の関連性を持つ話題に転換することが望ましいと仮定する。この考えに基づき、動向情報とソーシャルメディアのデータからある程度の関連性を持つ話題を発見する手法を提案する。予備実験を行い、提案する判断指標および話題転換手法の有効性について考察する。

2 関連研究

2.1 対話エージェント

対話エージェントには特定の事柄を目的としたタスク指向形の対話エージェント[4]と、雑談のような

¹ https://dev.smt.docomo.ne.jp?p=docs.api.page&api_name=dialogue&p_name=api_usage_scenario

対話を目的とする非タスク指向型の対話エージェント[5]がある。タスク指向型の対話エージェントには、観光案内を目的としたもの[6]やバスの運行情報の案内を目的としたもの[7]などが存在する。雑談を行う非タスク指向型ではもちろん、何らかの目的を持った対話を行うタスク指向型であっても、対話相手の信頼の獲得等に雑談は有用であるとされている[8]。雑談を行うエージェントは、あらかじめ用意したテンプレートとユーザ発話を照合して応答を生成する手法が一般的であり、NTT ドコモの雑談対話 API など簡単に利用可能なものが開発されている。この様な、発話の意味理解に基づかない簡易的な対話エージェントであっても、ロボットやバーチャルエージェントなどと組み合わせることで実用的なサービスを提供することが可能である。

対話エージェントの雑談における問題点の一つとして、発話の意味理解を行わずに応答を返すために、しばしば不自然な応答を生成してしまうことが挙げられる。不自然な応答をするとユーザの円滑な対話を妨げるため、その発生をなるべく抑制することが必要であり、研究が進められている。不自然な応答に関しては、対話が不自然な状況にあるパターンを分類した研究[9]や、LSTM (Long short-term memory) を用いた対話破綻の検出[10]のように対話破綻に関する研究が行われている。しかし、対話破綻を検出した後にどのように対応するかは著者らの知る限り少ない。

2.2 関連する情報の取得

専門用語集の自動編集や連想検索などを目的として、関連する単語や情報の取得に関する研究が行われている。佐々木らは、特定の 2 単語の関連度を Web 検索の結果から求める手法を提案している[11]。この手法では、Web 上の検索エンジンのヒット数や文中の用語の共起情報を利用し専門用語集を編集する。また、宮川らは、単語の意味に基づく連想検索を提案している[12]。この手法では対象のテキストからメタデータ空間を作成し単語の関連性を判断する。これらの手法を、本稿の目的である関連話題の発見に適用することを想定した場合、いくつかの課題が存在すると考える。

前者の手法では、指定した 2 単語について関連度を求めることができるが、ある単語に対する関連語を不特定多数の候補からリアルタイムで発見するような用途には適していない。また連想検索では、直接的な関連性を持つ連想が目的であるため、ある程度の緩い関連性を持つ話題の発見には適していない

と考える。

高間らは、動向情報を対象としたコンテキスト検索エンジンを提案しており、同時期に同様の変動を示したアイテムの発見などが可能であることを報告している[13]。時間的な関連性のあるアイテムは、意味が近いなどの直接的な関連を持たないため、転換話題候補の発見に利用可能と考える。

3 研究概要

3.1 提案システム

図 1 に提案するシステムの構成を示す。

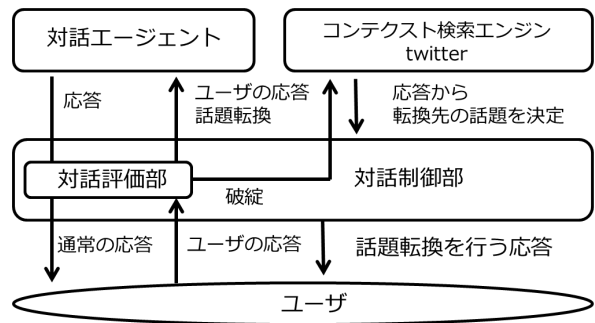


図 1 提案システムの構成

提案システムでは、通常は既存の対話エージェントを利用して応答を生成する。対話評価部ではエージェント、ユーザ双方の発話を評価し、対話が破綻状態にあるかを判断する。破綻していると判断した際に対話への介入を行い、それまで話していた話題を元にコンテキスト検索エンジン及び twitter の検索を行い、関連する話題への話題転換を行う。コンテキスト検索エンジンから得られたアイテムに関する話題を促す発話を生成し、エージェントの発話とする。この時、対話エージェント側にも話題転換を認識させる必要がある。この処理は利用する対話エージェントにより異なるが、今回利用した NTT ドコモの雑談対話 API では、これまでの対話に関連付けられている変数 context の中身を一度リセットした上で、「(コンテキスト検索エンジンから取得したアイテム)について話しましょう」という発話をユーザの発話として対話エージェントにも送信する。

3.1.1 話題転換

話題転換には以下の 3 通りのパターンがある[14]。

- ・ 関連性の強い話題への話題転換
- ・ ある程度の関連性を持った話題への話題転換

・無関係な話題への話題転換

この内、関連性の強い話題への話題転換は、小さな齟齬を修正する場合や、あるいは破綻の修正でなく同じ話題が連続した場合のアクセントとして話題転換を行う場合は有効であるが、本来の対話内容と大きくかけ離れている場合には間違った話題をそのまま続け、対話が破綻した状態を継続してしまう可能性が考えられる。また、無関係な話題への話題転換は、話題が一段落した時に別な話題を提示するときには適しているが、それ以外の場合においては破綻として感じられる場合がある[9]。一方、ある程度の関連性を持った話題への話題転換は人間同士の対話で最も多く行われる[14]ため、話題転換の必要性がない状態の時に話題転換してしまっても通常の対話と受け止められさほど悪影響を及ぼさないと考えられること、対象となる話題に近い話題を続けたい場合には適さないものの、齟齬があった場合の修正や話題が一段落した時の話題転換には適用が可能と考えられることから、提案システムではこのタイプの関連を持つ話題による話題転換を行う。

3.1.2 関連情報の検索

提案システムでは、近い時期に動向情報が類似した特徴的変動（急上昇など）を示したアイテム間は何らかの関係があると想定し、関連情報候補の絞込を行う。なお、ここでいう動向情報とは、例えば Wikipedia の閲覧数のように時間とともに変化するデータのことである。

コンテキスト検索エンジン[13]では対象のアイテムの動向情報が特徴的な変動を見せた時期、または逆に特定の時期に特徴的な変動を見せたアイテムを検索することができる。また、特徴的な変動のタイプとして、最大値、最小値、ピーク、ボトム、急上昇、急降下の 6 種類のうちいずれかを指定して検索できる。本稿ではそれまでの話題に対応するアイテムの動向情報が最大値を迎えた時期と同時期に、同様に最大値を迎えたアイテムを検索し、それを関連アイテムの候補とする。

コンテキスト検索エンジンの検索結果より得られた関連アイテムの候補から、実際に関連性を持つ可能性があるものに絞り込むために、twitter の情報を利用する。具体的には、同様の特徴的変動が見られた時期に、直前までの話題と関連アイテム候補のどちらもキーワードとして含むツイートが最も多いアイテムを関連性がある単語と判断し、そのときに得られたツイートを整形して話題転換のための応答文とする。ここで、キーワードとなる単語の判定には

yahoo のキープレーズ抽出²を用いた。

3.2 対話破綻の検出指標

話題転換を伴わない、対話エージェントとの通常の対話を行う予備実験を行ったところ、破綻した応答をエージェントが返した際はユーザが応答を返すまでの時間が長く、また応答そのものも短くなる傾向が観測された。そこで本稿では、対話が破綻していたかの判断には、ユーザが対話エージェントの発話に対し応答するまでの時間及びその応答の長さを用いる。ユーザの応答までの時間を t 秒、応答の長さを l 文字としたとき、 t/l で得られる値が大きい場合に対話が破綻していると判断する。なお、この値を本稿では文字当たり応答時間と呼ぶ。

4 評価実験

4.1 実験概要

提案システム及び対話破綻の検出指標の有効性や課題を調査することを目的として、予備実験を行った。本実験では、既存の対話エージェントとユーザとの対話に介入し話題転換が可能であること、及び提案する指標により対話破綻がどの程度検出可能であるかを確認する。

4.2 実験設定

対話の開始前に、実験協力者にコンテキスト検索エンジンの検索結果のうち、2006 年から 2014 年の間に動向情報が最大となっているアイテムを提示し、その中から話題としたいものを選択してもらい、対話エージェントと対話を行ってもらった。この時、一般的でない話題だと対話エージェントとの対話が最初から破綻してしまう場合があるため、なるべく一般的と思われる単語を話題として選択すること、固有名詞や名詞以外のアイテムは避けることを指示した。

対話時には、基本的には文献[9]の対話 10 ヶ条に従ってもらったが、話題転換の効果を見るための実験であるためなるべく実験協力者の側からの話題転換は行わないように指示した。

²

<http://developer.yahoo.co.jp/webapi/jlp/keyphrase/v1/extract.html>

対話時間は1回あたり5分で、途中で不自然な発話を行っていたと判断した場合話題転換を行った。話題転換は、1度だけ行うようにし、その後は話題転換を行わないように設定した。また、5分が経過しても対話破綻を検出しなかった場合は、話題転換を行わず終了するように設定した。なお、3.2節で提案した指標の有効性検証に必要なデータを収集することも本実験の目的であるため、提案する指標ではなく、ほぼ同じ内容の発話が連続した場合に対話が破綻したと判断し話題転換を行っている。

また、それぞれの対話終了後に、自然と感じた発話・不自然と感じた発話、及び話題転換が行われたことに気付いたか否かに関するアンケートを行った。実験は工学系学部生・院生8人に協力してもらい、1人あたり3回で計24回の実験を行った。

4.3 実験結果

4.3.1 対話破綻検出指標の評価

対話の記録とアンケートの結果から、対話エージェントの発話を自然な発話と不自然な発話に分類した。それぞれのタイプの発話について、3.2節で提案した文字当たり応答時間の平均値を実験協力者毎に求めた結果を表1に示す。表において、括弧内に示した値は、外れ値を除去して求めた平均値である。ここで、実験協力者が実験時に使用したキーボードに不慣れだったために必要以上に入力に時間をかけてしまう場面が見られたため、第3四分位数に四分位範囲の1.5倍を加えた値を越えている要素を外れ値と判断している。

表1：文字当たり応答時間の平均値 ([s/字])

実験協力者	自然	不自然
A	2.16 (1.65)	2.32 (2.32)
B	2.11 (2.11)	3.98 (2.94)
C	3 (3)	7.6 (7.6)
D	3.57 (2.32)	3.25 (3.25)
E	1.93 (1.93)	4.87 (3.27)
F	2.01 (2.01)	2.75 (2.43)
G	5.52 (1.76)	3.17 (2.70)
H	1.53 (1.53)	2.15 (1.60)

表1から、自然な発話より不自然な発話の方が、どの実験協力者においても文字当たり応答時間が大きくなっていることがわかる。

文字当たり応答時間の違いが有意なものであるかを確かめるためにt検定を行った。この時、実験協力者Cについては、自然、不自然の差が大きくなっているが、アンケート結果から自然・不自然と判断された応答がそれぞれ1つしかなく、外れ値であるか否かの判断ができないため、除外して検定を行った。

外れ値を除外した文字当たり応答時間の平均値に関してt検定を行った結果、p値は $0.0030 < 0.01$ となり、有意水準1%で差が認められた。以上より、文字当たり応答時間は対話破綻の検出指標として用いることができると考える。

4.3.2 話題転換

計24回の実験のうち、18回で5分間に話題転換が行われず、また話題転換が行われた6回のうち2回は文章が意味を持たない発話になってしまった。

話題転換を行うことができた4回に対して文字当たり応答時間を計算した結果を図1に示す。

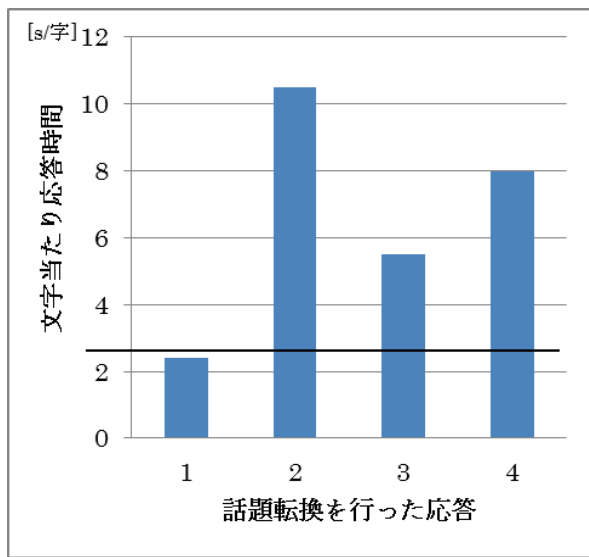


図 1: 発話転換実施時の文字あたり応答時間

図に示した横線は、不自然と判断された発話の文字あたり応答時間の平均値である。4回の発話うちこの値以下のものが1回であったのに対し、残り3回は大きく上回った。この結果から、文字あたり応答時間を基準とした場合にはこれらは話題転換が不自然な発話に分類される。その理由として考えられることとして、この3つの話題転換はいずれも、対話内容が固有名詞について言及していた事があげられる。一般的ではない話題への話題転換を行ってしまったため、実験協力者が反応に困り、応答までの時間がかかってしまったことが考えられる。このとき、実験協力者の応答は3回中2回が疑問文、残り1回は続きを促すような応答となっていた。このケースにおいては、その後の対話も含めアンケートで不自然に感じた対話と回答されていたが、一方でエージェントが話題を理解できておりその話題を続けることができれば不自然な対話ではないとの回答もあった。現状では、対話エージェントが話題を理解するのが困難であり、その後の対話の継続に支障があると考えられるため、転換の際は固有名詞を避けて話題を選択する事が好ましいと考える。しかし、対話エージェントの能力が向上し、事前に対応可能な話題かを判断することが可能となれば、より高度な対話の実現に貢献することが期待できる。

また、話題転換が行われなかった18回の対話の中には、5分の間に応答が破綻していたと判断されなかった場合と、破綻していたと判断されたが適切な話題が見つけれなかった場合の2通りがあった。18回のうち適切な話題を見つけれなかった場合

は10回あった。この原因として考えられるのは、コンテキスト検索エンジンのデータ不足が挙げられる。これに関しては検索可能なアイテムを追加することで、改善していくことが期待される。

話題転換を行ったが意味のない発話を返してしまった場合を確認したところ、単語を羅列したツイートがどちらのアイテムもキーワードとして含むツイートと判断され応答文が生成されていた。この問題は、文章として意味をなさないツイートを除外することで改善が可能と考える。

5 おわりに

本稿では、対話エージェントの対話破綻を検出する評価指標として文字あたり応答時間、及びコンテキスト検索エンジンを利用して関連話題を発見し、話題転換を行う手法を提案し、予備実験を行いその有効性について検討した。本稿で提案した文字あたり応答時間については、自然な発話と不自然な発話の差が有意であることを示した。また、対話破綻時の話題転換による対応については話題転換の対象となる話題が見つからない、一般的でない話題への話題転換を行った場合対話が破綻してしまうなどの課題が見つかったものの、それぞれ改善されることが期待できる事柄であり、より適切な話題転換を行うことができるようになることが期待できる。また、一般的でない話題への話題転換によって生じる問題については、現状では話題として固有名詞を避けることで対応するが、今後対話エージェントの性能が向上すれば、固有名詞を話題とした話題転換をより高度な対話を実現するために応用することも可能であると考えられる。今後はコンテキスト検索エンジンのデータ量を追加し、より多くの話題に対して話題転換の対象となる話題を発見できるようにするとともに、話題転換を実行する際の基準として文字あたり応答時間を用いた実験を行う予定である。また、話題転換の用途としては対話破綻だけでなく通常の対話時のアクセントとしての使用も期待できるため、破綻していない通常の対話時においても、話題転換を行うための適切なタイミングを判断する指標の検討も重要であると考えられる。

参考文献

- [1] 中野 有紀子, 比企野 純太, 安田 清, 認知症患者のための語りかけエージェント -プロトタイプシステムによる実証研究-, The journal of the Faculty of Science and Technology, Seikei University 49(2),

- pp.105-108, 2012
- [2] 西村 竜一, 原 直, 川波 弘 道, 李 晃伸, 鹿野 清宏,
10 年間の長期運用を支えた音声情報案内システム
「たけまるくん」の技術, 人工知能学会誌, Vo. 28, No.
1, pp.52-59, 2013
 - [3] 太田 知宏, 鳥海 不二夫, 石井 健一郎, 発話生成を
目的とした Wikipedia からの文抽出, JSAI2009,
2G1-NFC5-11, 2009
 - [4] 渡辺 卓夫, 大澤 一郎, 米澤 明憲, 対話領域の独立
性を指向した日本語対話理解システム, 人工知能学
会誌 Vol.2, No.2, pp.206-213, 1987
 - [5] 藤倉 将平, 小川 義人, 菊池 英明, 非タスク指向対
話システムにおけるユーモア応答生成手法,
JSAI2015, 2L5-OS-07b-1, 2015
 - [6] 翠 輝久, 河原 達也, 正司 哲朗, 美濃導彦, 質問応
答・情報推薦機能を備えた音声による情報案内シ
ステム, 情報処理学会論文誌 48(12), pp.3602-3611,
2007
 - [7] 安達 史博, 河原 達也, 奥乃 博, 岡本 隆志, 中嶋
宏, VoiceXML の動的生成に基づく自然言語音声対話
システム, 情報処理学会研究報告音声言語情報処理
(SLP) , 2002(10(2001-SLP-040)), pp.133-138, 2002
 - [8] Bickmore, T. and Cassell, J., Relational Agents: A
Model and Implementation of Building User Trust,
ICSLP2001, Vol.2, pp.729-732, 2001
 - [9] 東中 竜一郎, 船越 孝太郎, 荒木 雅弘, 塚原 裕史,
小林 優佳, 水上 雅博, テキストチャットを用いた
雑談対話コーパスの構築と対話破綻の分析, 自然言
語処理, Vol.23, pp.59-86, 2016
 - [1 0] 稲葉 通将, 高橋 健一, Bidirectional LSTM-RNN
を用いた対話破綻検出のエラー分析, 言語処理学会
第 22 回年次大会 発表論文集, pp.441-444, 2016
 - [1 1] 佐々木 靖弘, 佐藤 理史, ウェブを利用した関
連用語の自動収集, 情報処理学会研究報告 自然言語
処理(NL), pp.57-64, 2003
 - [1 2] 宮川 祥子, 清木 康, 特定分野ドキュメントを
対象とした意味的連想検索のためのメタデータ空間
生成方式, 情報処理学会論文誌データベース
(TOD) , Vol. 40, No. 5, pp.15-28, 1999
 - [1 3] 高間 康史, 加藤 優, 桑折 章吾, 石川 博, 動
向に関する問いを対象とした検索エンジンの提案,
人工知能学会論文誌, Vol.30, No.1, pp.138-147, 2015
 - [1 4] 藤本 英輝, 概念的関連性に基づく雑談の話題
転換点分析, JSAI2004, 2G3-01, 2004.