

自然な流れに沿って対話を進める アジェンダベース雑談対話システム

Agenda-based chat-oriented dialogue systems that talk along the dialogue flow

杉山 弘晃* 成松 宏美 水上 雅博 有本 庸浩
Hiroaki Sugiyama Hiromi Narimatsu Masahiro Mizukami Tsunehiro Arimoto

NTT コミュニケーション科学基礎研究所
NTT Communication Science Laboratories

Abstract:

Conventional dialogue system studies have specialized for the improvement of response at the single question level without planning the flow of dialogues. Even with an appropriate system utterance as a single response, if it does not match the dialogue flow, the user confused about the system's intention because it is shirking its primary duty of facilitating the dialogue. To overcome this problem, we previously proposed a novel chat-oriented dialogue system based on a dialogue management using hand-crafted finite state transducers and achieved the 1st prize in the dialogue system live competition 1; however, its effectiveness on proceeding the dialogues along the flow has not been not examined. In this research, we examine the effectiveness and report the result of the dialogue system live competition 2.

1 はじめに

近年、従来のタスク指向の対話システムとは異なる、雑談を行う雑談対話システムに注目が集まっている [Higashinaka 14, Vinyals 15]. タスク指向対話は、明確なゴールを持つタスクの効率的な達成を目的とする対話であるが、雑談はこれとは異なり、対話そのものから楽しさを得たり、満足を得ることを目的とする対話である [Cheepen 88]. すなわち、雑談対話システムは、対話を通して人を楽しませたり、満足を与えることを目的とするシステムといえる。

雑談対話の特徴の一つに、対話の話題が明確に定まっていなことが挙げられる。そのため、従来の雑談対話システム研究では、対話全体の流れの制御よりも対応可能な話題の幅を広げることを優先し、入力発話・履歴に対する応答文生成として問題を定式化することが多かった。対話全体に対する制御やプランニングを行わないプランレスなアプローチには、人同士の対話から得られた大量の細切れの発話列に基づく応答生成が可能となり、また評価も容易という利点がある [Jurafsky 18]. こうした背景から、近年はプランレスなアプローチでの応答生成に関する研究が非常に盛んであり、実際に

生成される発話も一問一答としては非常に自然なものとなっている。

しかしながら、プランレスなアプローチから生成されたシステム発話は、対話の流れを考慮せずに生成されているため、一問一答として見れば適切であっても、対話の流れとは整合しない場合が多い。対話の参加者は対話を円滑に進めることが期待されるため、自然な流れと整合しない不自然な発話を行うと、対話を円滑に進める意思がないと感じられ、対話破綻やユーザ満足度低下の原因となる。そのため、対話システムは本質的に対話の流れを考慮し発話することが求められる。

この問題に対し、我々是对話の流れと内容を分離し、対話の流れに特化したシナリオを手で定義することで、流れに沿って対話を進めるシステムを提案してきた [杉山 18]. 提案システムは 2018 年に開催された対話システムライブコンペティション 1 (以下、ライブコンペ 1) [東中 18] の本戦にてトップの成績を獲得しており、雑談においても流れに沿った対話が重要であることを示唆する結果となっている。しかしながら、提案システムが、実際に対話の流れに沿った発話を行っていたかについては検証されていない。本稿では、昨年度システムの構成についてより整理された形で議論するとともに、ライブコンペ 1 で公開された代表的なシステムと対話の流れについて比較評価を行う。合わ

*連絡先: NTT コミュニケーション科学基礎研究所
〒 619-0237 京都府相楽郡精華町光台 2-4
E-mail: h.sugi@ieee.org

SYSTEM こんにちは！お名前を伺ってもいいですか？
 USER USER といいます。
 SYSTEM USER さんとおっしゃるんですね。私はリコっていいです。よろしくお願ひします。USER さんは何県にお住まいなんですか？
 USER 埼玉県に住んでいます。
 SYSTEM ふむふむ。埼玉県ですか。埼玉いいですね。行ってみたいです。長瀨とか有名ですよ？
 USER 長瀨は近いので自転車で行くこともあります。
 SYSTEM いい桜があるのうらやましいなあ。私はお花見とか好きなんですけど、長瀨の桜ってどうですか？
 USER 荒川沿いの桜並木が見事で、春には桜のトンネルのような景色になります。
 SYSTEM 桜いいですよ。ところで、私青森県に住んでるんですけど、桜といえば弘前城とかもお勧めです。USER さんは行かれたことはありますか？
 :
 SYSTEM あ、嵐山ですか。京都府ですね。紅葉いいですよ。そういえば、こないだ三重県の番組見たんですけど、伊勢神宮とかおススメですよ。行かれたことありますか？
 USER 残念ながらまだ行ったことがありません。日本で最も格式高い神社ですから、一度は行ってみたいです。
 SYSTEM まあそんなもんですよね。伊勢神宮もぼちぼち有名なところだと思います。ですよ。あ、そろそろ終わりですね。お話ししていただいてありがとうございます！
 USER 名所の話ができて楽しかったです。
 SYSTEM ではでは。またお会いできるのを楽しみにしています。

表 1: ライブコンペ 1 予選での対話例。ユーザ名を「USER」に置換している。

せて、ほぼ同一のシステムを利用した今回のライブコンペティション（以下、ライブコンペ 2）[東中 19] の予選結果について報告する。

2 対話のデザイン

一般に雑談対話システムは、応答可能な話題の多様性を重視し、オープンドメインなシステムとして構築されることが多い。しかし、多数の雑談対話を収集したコーパスには多様な話題が現れるものの、各対話内で個々に出現する話題の幅はさほど広くなく、特定の話題に限定されることも多い。また、自然な流れに沿った対話を行うには、システムは自身が導入した話題を継続する上で最低限必要となる知識を持つ必要がある。これらを鑑み、本研究では、特定の話題に特化し十分な知識の基に対話を進める雑談システムを構築する。

具体的な話題として、提案システムでは、ユーザ・システムの居住地や出身地、最近の旅行先について、名所を紹介し合ったりその良さを共感し合う雑談を対象とする。これは、初対面の 2 者間の対話では、お互い自身についての話題の出現数が多いこと [杉山 16]、その中でも対話の話題として旅行の話題が多く現れること [Arimoto 19]、旅行の話題で現れる名所等の情報は、固有名詞として一意に特定しやすくユーザ発話からの抽出・同定が容易であること、Wikipedia 等からの知識の収集が容易であることによる。また、システムが受け身になりユーザが主導するような対話では、ユーザは対話を継続するために話題を提供し続ける必要があり、負担に感じるものが危惧される。ユーザの対話継続に対する負担を軽減するため、システムが主導する流れで対話を進める。

表 1 に、ライブコンペ 1 予選において提案システム

が実際のユーザと対話した例を示す。対話は挨拶と自己紹介から始まり、ユーザの居住地（埼玉）にある名所（長瀨）の話題に展開される。システムが名所のポイントとして桜に焦点を当て、ユーザに長瀨の桜についての印象を尋ねることで、ユーザの印象を引き出している。また、ユーザの印象に共感する形で話題を一段落させ、システム自身の居住地（青森）の名所（弘前）に話題を転換する、という流れで対話を続けている。

ここで、ユーザに具体的な印象を尋ねる上で、もしユーザがその名所を知らなかったり、そこに行った経験がなければ、長瀨の桜の印象のような、行った経験があることを前提とする質問は不適當なものとなる。そのため提案システムでは、ユーザへの共感を示す対話の流れの最小構成として、名所についての知識および名所である認識、名所に行った経験、名所の具体的な印象の 3 要素を順に引き出す流れを設定する。提案システムでは、大局的な話題（ユーザ・システムの居住地・興味のある旅行先・出身地）を切り替えながら、この 3 ステップを繰り返すことで、自然な流れに沿った対話を実現する。

3 アジェンダベース雑談対話システム

2 章で示した流れに沿った対話を実現するため、ライブコンペ 1 へ提出した提案システムは、対話の流れと内容を分離して定義するアプローチを採用している [杉山 18]。対話の流れは、有限状態トランスデューサ (FST) と、状態間の遷移に紐付いた発話テンプレートを人手で定義することで実現している。また、テンプレートに埋め込む具体的な発話内容は、ユーザ・システム自身についての情報（名前や居住地等）に加え、

Wikipedia から構築した名所に関する外部知識（名所の名称や種類，ポイント等）を利用している．我々は，対話の流れについて説明する上で，定義した FST とテンプレートを具体的に例示するという方法を取っていた．この説明は全く同一のシステムを再現する上では有用であるものの，必ずしも全体の設計が見通しのよいものではなく，同様のシステムを新規に構築する上での再現性に欠けるといふ欠点があった．

本稿では，タスク対話システムで用いられる対話制御手法の一つである，アジェンダに基づく対話制御の考え方に沿って，表 1 の流れを説明する．アジェンダに基づく対話制御とは，対話をサブゴールを持つチャンクに分割し，各サブゴールの達成に寄与するような大まかな対話の流れをデザインする方法である．提案システムでは，大局的な話題をチャンクと考え，それぞれで選定した詳細な話題（名所）について，ユーザから具体的な印象を引き出し，システムの共感を出発することをサブゴールとする．局所的なサブゴールを設定し，それに寄与するように対話を設計することで，一貫性のある自然な対話の流れを容易に実現できる．

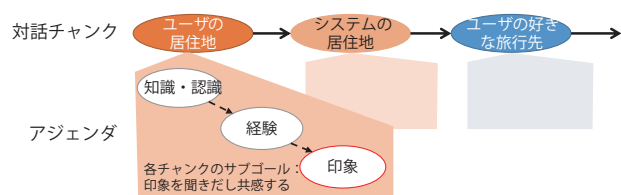


図 1: アジェンダに基づく対話制御

図 1 に，アジェンダに基づく対話制御の概観を示す．アジェンダ内の丸はチャンク内でユーザに表出させる要素を表している．システムは基本的にこれらを順に尋ね，それらに対するユーザの回答に対して称賛や共感を発話することで対話を進める．ただし，ある質問と対応しない別の要素がユーザ発話に含まれている場合は，その要素についてのシステムからの質問をスキップするとともに，ユーザから質問への回答が得られたものとして対話を進める．例えば表 1 の対話例では，「長瀬とか有名ですよ？」という知識・認識を問う質問により，ユーザの長瀬についての知識や，長瀬が名所だという認識についての表出を促している．それに対し，ユーザは「長瀬は近いので自転車で行くこともあります」と，長瀬へ行った経験を有することを含めて応答している．経験についてユーザが肯定的に発話した場合，その前提として，長瀬が名所という認識が肯定されたと推測される．そのため，「良い桜があるのうらやましいなあ」と長瀬のポイントである桜を（Wikipedia で構築した外部知識に基づいて）導入し称賛を表出する．合わせて，経験についての質問をスキップし，「私はお花見とか好きなんですけど，長瀬の桜ってどうですか？」と印象について質問する．ユーザが具体的な

印象を表出したことに合わせて，システムは「桜いいですよ〜」と共感を出して話題を一段落させ，次のチャンクであるシステムの居住地の話題に移る．

一方，流れの途中で否定された場合は，その後の質問が不適當になるため，チャンクを打ち切る必要がある．例えば対話後半では，システムの「伊勢神宮とかおすすですよ．行かれたことありますか？」という経験を問う質問に対し，ユーザは「行ったことがありません」と否定している．このとき，ユーザに具体的な印象を尋ねる質問は不適當になるため，システムは「まあそんなもんですよ〜」と，行った経験がないことについての共感を出し，別な話題に切り替えている．このようにチャンク単位で流れを持ち，適切に対話の進行を制御することで，自然な流れに沿った対話を実現できると考えられる．

4 実験 1: 対話の流れに関する評価

本節では，提案システムが対話の流れに沿った発話を生成できているかについて，従来のプランレスな一問一答型の雑談対話システムとの比較を通して検証する．

4.1 実験設定

本研究では，比較する雑談対話システムとして，ライブコンペ 1 の上位システムを利用する．具体的には，破綻検出チャレンジ 4 で公開された，上記システムとユーザとの対話データ [Higashinaka 19]，および初対面の人同士の対話データ [Arimoto 19] に第 3 者評価を付与し，システム間の比較を行う．利用したシステムは，Mariko bot [角森 18]，Zunko bot [阿部 18]，IRS [東中 18] の 3 種類である．システムごとに 10 対話をランダムに選び，23 名の日本語話者が 7 段階で評価を付与した．

我々の主な興味は，明確なゴールが話者間で共有されていない雑談であっても，提案するアジェンダベースの対話システムが流れに沿った対話の実現に寄与するか，またそうした対話によってシステムが対話を円滑に進めようとしていると感じられるか，の 2 点にある．そのため，以下を評価尺度とし，検証を行う．

- (a) 対話の流れ: S の発話は対話の流れに沿っている
- (b) 円滑に進める意志: S は対話を円滑に進めようとしている

4.2 結果

各システムおよび人の評価結果を図 2 に示す．尺度 (a) 対話の流れと (b) 円滑に進める意志において，提案システムが他のシステムを上回っていた．尺度 (a) で

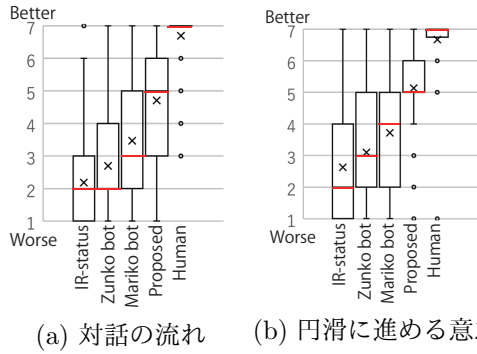


図 2: 対話ログに対する評価結果

は、提案システムの平均値 ($M=4.71$, $SD=1.71$) は次点の Mariko bot ($M=3.47$, $SD=1.63$) を上回った。また尺度 (b) においても、提案システムの平均値 ($M=5.13$, $SD=1.54$) は次点の Mariko bot ($M=3.72$, $SD=1.65$) を上回っていた。Bonferroni 法で多重比較の補正を行った調整信頼水準 $\tilde{\alpha} = 8.3e - 4$ ($\alpha = 0.05$ 、試行回数 $m = 60$) を有意水準とする Mann-Whitney U 検定を行ったところ、尺度 (a), (b) の双方で提案システムと次点のシステムとの間には有意な差が見られた。

一方、人間同士の対話の評価は、提案システムとの対話を有意に上回っていた。これらの結果は、対話スクリプトと外部知識の組合せが、一問一答形式の雑談システムよりも対話の流れに沿った雑談を実現する上で有用であることを示している一方、人間レベルの対話の達成には未だ課題が残っていることを示している。

5 実験 2: ライブコンペティション 2

提案システムをライブコンペ 2 [東中 19] のオープントラックに提出した。提案システムがライブコンペ 1 で優勝したシステムであることから、ある種のベースラインとなるものと考え、対話の流れや発話テンプレート等はライブコンペ 1 とほぼ同一の設定を用いた。前回の予選からの差分は、予選時に問題として指摘されていた、ユーザからシステムへの一部の問い返しへの返答について、一言だけ応じるようフローを調整したこと、ハンドルネームでも名前と認識できるよう、名前のチェックを緩和したのみである。

ライブコンペ 2 の予選において、44 名のクラウドワーカーと対話し、対話ログと各ログの評価値およびアンケートコメント 15 件を得た。平均評価値は 3.84 (5 が最高、1 が最低。点数分布は 5: 14, 4: 19, 3: 3, 2: 6, 1: 2) で、全 9 システム中 3 位の成績であった。コメントの内容をまとめると、会話が自然というコメントが 8 件と最も多かった一方、実際のユーザ自身の情報以外を入れたため、知識的に会話についていけなかったというコメントも 3 件あった。居住地や出身地等の、

個人を全く特定できない程度の情報にも関わらず、こうした情報を発話しながらいない人が一定数存在するという事は、実際のサービスとして展開する際の障害となりうると思われる。

6 おわりに

本研究では、ライブコンペ 1 で提出したシステムについて、アジェンダベースの対話制御としての解釈を加えるとともに、他のプランレスな雑談システムに対して対話の流れが有意に向上していることを検証した。また、本システムをライブコンペ 2 へ提出し、また、本システムをライブコンペ 2 へ提出した結果、評価値は全 9 チーム中 3 位であり、コメントからも、本システムがある程度自然な流れの対話を実現できていることが示された。今後はアジェンダの構成自体のコーパスからの学習や、ユーザ発話の認識性能の向上に取り組む。

参考文献

[Arimoto 19] Arimoto, T., Sugiyama, H., Mizukami, M., Narimatsu, H., and Higashinaka, R.: Analysis of satisfaction and topics in repeated conversation through days, in *Proceedings of SEMDIAL* (2019)

[Cheepen 88] Cheepen, C.: *The predictability of informal conversation*, London: Pinter (1988)

[Higashinaka 14] Higashinaka, R., Imamura, K., Meguro, T., Miyazaki, C., Kobayashi, N., Sugiyama, H., Hirano, T., Makino, T., and Matsuo, Y.: Towards an open-domain conversational system fully based on natural language processing, in *Proceedings of the 25th International Conference on Computational Linguistics*, pp. 928–939 (2014)

[Higashinaka 19] Higashinaka, R., D’Haro, L. F., Shawar, B. A., Banchs, R., Funakoshi, K., Inaba, M., Tsunomori, Y., Takahashi, T., and Sedoc, J.: Overview of the dialogue breakdown detection challenge 4, in *Proceedings of WOCHAT* (2019)

[Jurafsky 18] Jurafsky, D. and Martin, J. H.: *Speech and Language Processing*, Pearson, 3rd edition (2018)

[Vinyals 15] Vinyals, O. and Le, Q.: A Neural Conversational Model, in *ICML Deep Learning Workshop* (2015)

[阿部 18] 阿部香央莉, 佐藤志貴, 佐藤拓海, 藤井諒, 松田耕史, 鈴木正敏, 山口健史, 赤間怜奈, 大内啓樹, 鈴木潤, 乾健太郎: Zunkobot: 複数の知識モジュールを統合した雑談対話システム, 人工知能学会 言語・音声理解と対話処理研究会第 84 回 (第 9 回対話システムシンポジウム) (2018)

[角森 18] 角森唯子, 大西可奈子, 藤本拓, 角野公亮, 吉村健, 磯田佳徳: カスタマイズ可能なオープンドメイン雑談対話エンジンの開発, 人工知能学会 言語・音声理解と対話処理研究会第 84 回 (第 9 回対話システムシンポジウム) (2018)

[杉山 16] 杉山弘晃, 日黒豊美, 東中竜一郎: 対話システムのパーソナリティを問う質問の大規模な収集と分析, 人工知能学会論文誌, Vol. 31, No. 1, pp. 1–9 (2016)

[杉山 18] 杉山弘晃, 成松宏美, 水上雅博, 有本庸浩: 文脈に沿った発話理解・生成を行うドメイン特化型雑談対話システムの実験的検討, 人工知能学会 言語・音声理解と対話処理研究会第 84 回 (第 9 回対話システムシンポジウム) (2018)

[東中 18] 東中竜一郎, 船越孝太郎, 稲葉通将, 角森唯子, 高橋哲朗, 赤間怜奈: 対話システムライブコンペティション, 人工知能学会 言語・音声理解と対話処理研究会 第 84 回 (第 9 回対話システムシンポジウム) (2018)

[東中 19] 東中竜一郎, 船越孝太郎, 稲葉通将, 角森唯子, 高橋哲朗, 赤間怜奈, 宇佐美まゆみ, 川端良子, 水上雅博: 対話システムライブコンペティション 2, 人工知能学会言語・音声理解と対話処理研究会 第 87 回 (第 10 回対話システムシンポジウム) (2019)