

オススメ度により対話シナリオを調整可能な 統計的対話システム

岩田 憲治^{1,a)} 小林 優佳¹ 吉田 尚水¹ 藤村 浩司¹ 赤嶺 政巳¹

概要：ショッピングセンターの案内などを行う対話システムをサービスとして提供する場合、セール情報の変化などサービス内容の変化に応じて対話システムの対話シナリオをカスタマイズすることが求められる。しかし従来の音声対話システムでは、カスタマイズのためには人手による複雑な対話フローの修正や対話データの追加収集・学習が必要となる。そのため、対話システムの中身を詳しく知らない人が対話シナリオをカスタマイズするのは困難であった。本研究では、そのような人でも対話シナリオを低コストにカスタマイズ可能とする技術の実現を目標とする。提案手法では、案内候補のオススメ度を変更することにより、対話システムの性能を維持しつつ対話シナリオをカスタマイズ可能とする。この実現方法として、オススメ度に基づく報酬とユーザの満足度に基づく報酬を用いて対話シナリオの強化学習を行う。これにより、ユーザが不満に感じない程度に積極的にオススメ度の高い候補をユーザに提示する対話シナリオを構築する。本手法をショッピングセンターの店舗案内システムを用いて評価し、その有効性を確認した。

Statistical Dialogue System with Capability of Controlling Dialogue Scenario by Degree of Recommendation

KENJI IWATA^{1,a)} YUKA KOBAYASHI¹ TAKAMI YOSHIDA¹ HIROSHI FUJIMURA¹ MASAMI AKAMINE¹

Abstract: When a dialogue system such as a guidance system of a shopping mall is deployed as a service, a dialogue scenario of the dialogue system needs to be customized as a change of contents of the service such as a change of sales. However, conventional dialogue systems require a modification of complex dialogue flows by hand or more dialogue data for the customization. Therefore, it is difficult for a person who doesn't know details of the dialogue system to customize the dialogue scenario. In this research, we aim to realize that such a person customize the dialogue scenario with low costs. Our proposed method controls the dialogue scenario without degrading a performance of the dialogue system by changing a degree of recommendation for candidates to provide to users. To realize this method, we exploit a reinforcement learning for the dialogue scenario referring rewards based on the degree of recommendation and user's satisfaction. Thus we create the dialogue scenario that the system provides the candidate with high degree of recommendation aggressively while not making users uncomfortable. Results of experiments using a dialogue system which guides stores of a shopping mall suggest that our method works effectively.

1. はじめに

近年、バス案内システムや旅行案内システムなど、様々なタスク指向型の音声対話システムの開発が行われている

る [1], [2]. これらのシステムではユーザの発話から希望する条件を理解し、条件に合う候補が多い場合は絞り込みを行い、最終的にいくつかの候補を提示する対話が行われる。従来、このような対話を行うために必要な対話シナリオは人手で作成しており、開発コストが高いという問題があった。この問題を解決するため、近年、強化学習を用いて対話シナリオを対話コーパスから自動的に学習する統計的対話制御技術の研究が進められている [3]. この技術を用い

¹ (株) 東芝 研究開発本部 研究開発センター メディア AI ラボラトリー

Media AI Laboratory, Corporate Research & Development Center, Research & Development Division, Toshiba Corporation

^{a)} kenji4.iwata@toshiba.co.jp

実際に対話システムをサービスとして提供する際には、サービス内容の変化に応じて対話シナリオをカスタマイズすることが求められる。例えばショッピングセンターでは、セールスの開始・終了などの変化に伴いオススメしたい店舗を変更したいなどの要望が考えられる。従来のルールベースや統計ベースの対話シナリオ生成手法では、カスタマイズのために人手で複雑な対話フローの追加開発を行ったり、システム応答が切り替わる閾値を調整したり、追加で対話の事例を集めて再学習したりする必要がある。そのため、サービス運営者のような対話システムの中身を詳しく知らない人が目的に合わせてカスタマイズを実施するのは困難である。また、対話システム開発者の数も限られているため、サービス内容の変化に対話システムが即座に対応できないという問題が発生する。これまで、対話シナリオを新しいドメインに低コストで転用する研究 [6] や、対話を重ねることでユーザ個人個人に適応していく研究 [7] はあったが、対話システムの中身をよく知らない人が対話シナリオをカスタマイズ可能とするための研究は行われてこなかった。

対話システムはオススメ度の高い候補をユーザに提示できるように対話シナリオを変える必要がある。具体的には、候補を絞り込む対話が多くなるとオススメ度の高い候補が無くなってしまう可能性が高くなるため、オススメ度の高い候補が含まれている場合は候補数が多い段階でも提示をするように対話シナリオを変化させるのが望ましい。一方、ユーザから見ると、候補数が非常に多い段階で提示されると、候補全てを確認するのが困難なため煩わしいと感じてしまうと考えられる。そこで、対話システムとしての性能が低下しない程度に、オススメ度の変化を反映した対話シナリオを実現する必要がある。実現したい対話例を図1に示す。

Aストア（男性向けのバッグを扱っているお店）のオススメ度を調整

オススメ度が低い場合		オススメ度が高い場合	
 ユーザー バッグを買いたい 誰向けの商品をご希望ですか？ 男性向けで	 システム お店のジャンルについてご希望はありますか？ スポーツ用品かな	 ユーザー バッグを買いたい 誰向けの商品をご希望ですか？ 男性向けで	 システム Aストアなど61件見つかりました 積極的に提示するようになるが、ユーザが不満に感じないようある程度絞り込んでから提示する

提示タイミングが変化


Xストアなど18件見つかりました

Aストアがユーザーの要望に合う候補から外れる

店舗案内システム

お問い合わせがとうございます。
ご利用をお願ひします。

子供用の服がほしいたい

子供向けの服のお店として
アッシュ まで33件見つかりま
した。

もっと知りたいのですね

子供向けのリーズナブルな服の
お店として**クラシカルブラウン**
など4件見つかりました。

(詳しくはこちら)

 **案内終了**

ファッション＆フASHIONグッズ
クラシカルブラウン

品数11,710件



大人から子供まで家族全員に合うリアルカンユニファッशनを揃えます。

所在地 東京都 品川区 15-00-01
電話番号 03-5511-2061

ご来店リョウフプラン 4 店舗






クラシカルブラウン マーズ エクサトレート グループ

お問い合わせ日時 2022年 09月 12日 (火) 13:07

④ 来店履歴

店舗名	アッシュ	子供服、ファッション	着払い、返品	利用	お店の
アッシュ					
アッシュ					

手法によりオススメ度の高い候補が提示されやすくなり、また適切な報酬の重みの場合ユーザは不満に感じないことが確認できた。

2. 対話シナリオ学習フレームワーク

統計的対話制御における対話シナリオの学習フレームワークを図3に示す。対話状態に基づきシステムの行動を決定する際に用いる対話シナリオを強化学習 [8] によって生成する。強化学習は達成したい対話のゴールを報酬として定めておくとそのまでの最適な対話の流れを自動的に学習できるため、対話シナリオの学習に適した手法であると言える。

230

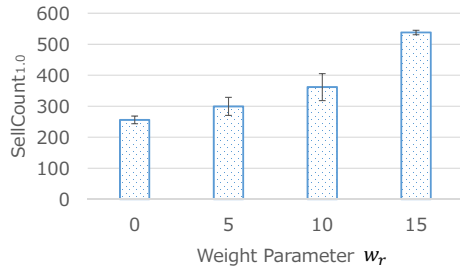


図 4 オススメ度 1 の候補を提示した対話数

計を反映した対話シナリオを学習するため、ドメイン非依存特徴量には提示する店舗の候補数、その候補の中で最も高いオススメ度も含む。報酬関数のパラメータは $\alpha = 20, \beta = 1, \mu = 25, \nu = 5$ に設定した。報酬の重みは w_u を 20 に固定し、 w_r を 0, 5, 10, 15 と変化させて評価を行った。 $w_r = 0$ がオススメ度を考慮しないシステム、 $w_r = 5, 10, 15$ がオススメ度を考慮するシステムに相当する。

発話理解とユーザシミュレータの学習に用いる対話コーパスはクラウドソーシングで収集した。クラウドソーシングの作業者がルールベースの対話システムと対話することで生成した。対話コーパス内の対話状態の正解ラベルなどは別途クラウドソーシングにより付与した。

発話理解のモデルには LSTM(Long Short-Term Memory)[10] を用い、現在のユーザの発話を理解するとともに、過去にユーザが伝えた条件も記憶するよう対話コーパスから学習した。ユーザシミュレータには Agenda ベースの手法 [11] を用い、あらかじめ定めたユーザゴールと直前のシステム応答に合うユーザの発話を学習した。

4.2 客観評価実験

客観評価実験ではオススメ度の高い候補が提示されやすくなっているか検証するため、以下の 2 つの指標で評価を行った。

- (1) SellCount_d: 提示した候補の中で最も高いオススメ度が d であった対話数
- (2) ReqNum: 候補を提示するまでに絞り込みを行った回数の平均

オススメ度に基づく報酬の重み w_r を変化させた各システムにおいて、初期値の異なる 5 種類の対話シナリオを 10,000 回のユーザシミュレータとの対話で学習し、それぞれの対話シナリオをユーザシミュレータとの 1,000 回の対話により評価した。5 種類の結果の平均を取ることによって SellCount_d と ReqNum を求めた。

図 4 に SellCount_{1.0}(オススメ度 1.0 の候補を提示した対話数) の結果を示す。重み w_r を大きくしていくことで、SellCount_{1.0} の値も大きくなっていくことが分かる。一方、SellCount_{0.5} や SellCount_{0.25} では重み w_r によって大きな

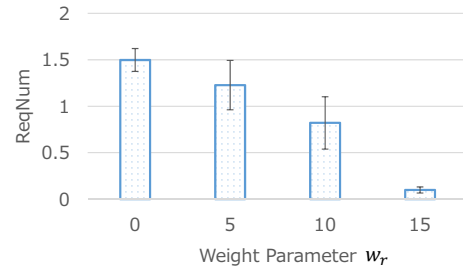


図 5 提示までに絞り込みを行った回数

変化は見られなかった。これはオススメ度を 0.5 や 0.25 に設定した候補とオススメ度を 1.0 にした候補で扱っている商品や価格帯などが重複しているものが多く、優先度 0.5 や 0.25 の候補が提示した候補の中で最もオススメ度が高くなることが少なかったためと考えられる。

ReqNum の結果を図 5 に示す。重み w_r を大きくしていくと ReqNum が減少していくことが分かる。 $w_r = 15$ では大抵はユーザの最初の発話に含まれる条件を受けてすぐに候補提示をしていることが分かる。これらの結果から、オススメ度に基づく報酬を入れて強化学習を行うことで、オススメ度の高い候補を積極的に提示するようになり、その積極さはオススメ度に基づく報酬の重みを強くしていくことで上がっていくことが分かる。

4.3 主観評価実験

オススメ度の高い候補を積極的に提示するようになった対話システムがユーザに許容されるか検証するため、主観評価実験を行った。被験者 12 名にオススメ度を考慮しないシステム ($w_r = 0$) とオススメ度を考慮するシステム ($w_r = 5$ または 10 または 15) それぞれの対話をペアで見せて、どちらの対話が良いかを評価してもらった。対話のペアは客観評価実験の評価に用いた 1,000 回の対話から作成した。オススメ度を考慮しないシステムと考慮するシステムでユーザシミュレータのゴールは同じものの、対話の進み方が異なっている対話をペアとした。オススメ度を考慮するシステムでの重み w_r ごとにペアを 10 ずつ抽出し、計 30 対話ペアを評価した。評価時にはどちらの対話がオススメ度を考慮するシステムかは分からないようになっている。また、初めて候補の提示を行うタイミングが好ましいかどうかを評価してもらうため、対話は候補の提示を初めて行うところまでに限定している。実際の対話ペアを図 6 に示す。

結果を図 7 に示す。 $w_r = 5, 10$ ではオススメ度を考慮した方が好ましいという結果になった。原因分析を行ったところ、オススメ度を考慮したシステムの対話の方が短いターンで候補の提示を行うことが多いこと、長いターンをかけて絞り込みを行っても結局同じくらいの候補数にしかないケースがあったこと、長いターンをかけることで

オススメ度を考慮しないシステム		オススメ度を考慮したシステム	
ターン	システム応答	ユーザ発話	システム応答
1	ご利用をどうぞ	バッグ	ご利用をどうぞ
2	ジャンルの希望はありますか？	ない	ジャンルの希望はありますか？
3	価格帯の希望はありますか？	ない	価格帯の希望はありますか？
4	対象者の希望はありますか？	女性向け	バッグでよろしいですか？
5	商品の希望はありますか？	バッグ	「Xストア」など79件見つかりました
6	バッグでよろしいですか？	はい	
7	女性か希望なしかどうかですか？	女性向け	
8	希望なしかおみやげかどうかですか？	希望なし	
9	「Xストア」など77件見つかりました		

図 6 オススメ度を考慮しないシステムとするシステムの対話のペア

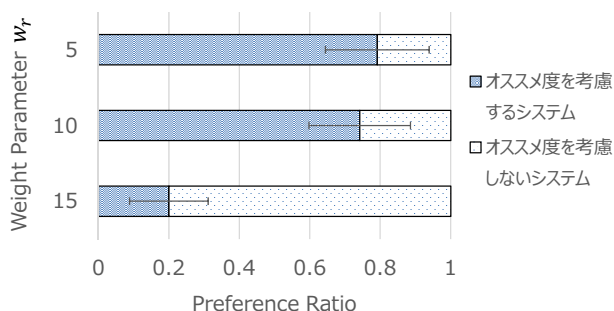


図 7 オススメ度を考慮するシステムとしないシステムの比較評価結果

発話理解の方で誤りが含まれる可能性が上がり、結果として対話が冗長になってしまったこと、などが原因として挙げられた。一方、 $w_r = 15$ ではオススメ度を考慮しない方が好ましいという結果が得られた。これは、オススメ度を考慮したシステムの方が候補数が非常に多い状態でも候補提示を行ってしまったためと考えられる。この結果から適切に重みを設定することで、ユーザが不満を感じることなく積極的に高いオススメ度の候補を提示することが可能になったと言える。

5. まとめ

対話システムの中身を詳しく知らない人でも対話シナリオをカスタマイズ可能とする技術の一つとして、オススメ度を変更することで対話システムの性能を維持しつつ対話シナリオを調整可能な対話制御技術を実現した。オススメ度に基づく報酬とユーザ満足度に基づく報酬を適切なバランスで重み付け和を取り、対話シナリオの強化学習をすることで、ユーザが不満に感じない程度に積極的にオススメ度の高い候補を提示する対話システムとなることが確認できた。今後は、実際のユーザで今回のシステムの性能評価を行っていく。また、対話シナリオをカスタマイズ可能とする技術として本手法以外の検討も進める。

参考文献

[1] Antoine Raux, Brian Langner, Alan W Black, and Maxine Eskenazi, “Let’s go: Improving spoken dialog systems for the elderly and non-natives,” in *Eurospeech03*. Citeseer, 2003.

[2] 翠輝久, 水上悦雄, 柏岡秀紀, “音声対話による観光案内システム AssisTra,” 研究報告ヒューマンコンピュータインタラクション (HCI), vol. 2013, no. 8, pp. 1–2, 2013.

[3] Jason D Williams and Steve Young, “Partially observable markov decision processes for spoken dialog systems,” *Computer Speech & Language*, vol. 21, no. 2, pp. 393–422, 2007.

[4] Milica Gasic and Steve Young, “Gaussian processes for pomdp-based dialogue manager optimization,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, no. 1, pp. 28–40, 2014.

[5] 岩田憲治, 吉田尚水, 小林優佳, 藤村浩司, “統計的対話技術を活用したショッピングセンター案内システム,” *SIG-SLUD*, vol. 5, no. 02, pp. 33–34, 2017.

[6] Zhuoran Wang, Tsung-Hsien Wen, Pei-Hao Su, and Yannis Stylianou, “Learning domain-independent dialogue policies via ontology parameterisation,” in *16th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, 2015, p. 412.

[7] Chaitanya K Joshi, Fei Mi, and Boi Faltings, “Personalization in Goal-Oriented Dialog,” in *31st Conference on Neural Information Processing Systems*, 2017.

[8] Richard S Sutton and Andrew G Barto, *Reinforcement learning: An introduction*, MIT press Cambridge, 1998.

[9] Blaise Thomson and Steve Young, “Bayesian update of dialogue state: A pomdp framework for spoken dialogue systems,” *Computer Speech & Language*, vol. 24, no. 4, pp. 562–588, 2010.

[10] Sepp Hochreiter and Jürgen Schmidhuber, “Long short-term memory,” *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[11] Jost Schatzmann, Blaise Thomson, and Steve Young, “Statistical user simulation with a hidden agenda,” *Proc SIGDial, Antwerp*, vol. 273282, no. 9, 2007.