

# 発話と推薦を柔軟に混合させる対話型推薦システムの実現

Dialogue based recommender system that flexibly mixes utterances and recommendations

積田 大介 <sup>\*1</sup>

Daisuke Tsumita

高木 友博 <sup>\*1</sup>

Tomohiro Takagi

<sup>\*1</sup>明治大学大学院理工学研究科情報科学専攻

Department of Computer Science, Meiji University

Many of the prior research in the recommendation through dialogue were designed separating dialogue and recommendation. However, since the accuracy of the recommendation itself is not necessarily high, rarely the recommendation result meets user needs. We human, however, can guide the solutions satisfying the user, by appropriately repeating the cycle of checking mismatch reason and making another recommendation in our conversations. In this paper, we proposed a system to leverage a dialogue strategy for reinforcement learning using recommendation results based on user's utterances. We realized a dialog system to perform adaptive behavior that naturally incorporates recommendations into conversation with users.

## 1. はじめに

接客や相談対応をするような目的指向型の対話システムの開発は、需要の高まりとともに発達が目覚ましい研究領域である。対話システムの設計において重要なのが、ユーザの発話に対してどういった応答を返すかを決定する対話戦略の定義である。従来の目的指向型の対話システムはルールベースによる対話戦略が主流であったが、多様なユーザの発話を想定し、膨大な数のルールを定義する必要がある。そのため、機能追加やルール修正に多大な労力を伴うという問題点がある。近年、これらのルールベースな対話戦略の課題に対し、バンディットアルゴリズムや強化学習などの枠組みによるモデルベースの対話システムが注目されている。また、一般的な目的指向型の対話システムでは達成すべき目標が決まっており、ユーザの求める解を提示するだけで目標を達成することができる。一方で、ユーザ自身が、ユーザニーズを明示的に理解していない状況下では、ユーザに対して推薦を行うシステムを構築する必要がある。しかし、推薦システムそのものの精度は必ずしも高くはないため、このような対話による推薦においては、対話の当初から推薦結果がユーザニーズに適合しない事が考えられる。その場合、適合しない理由を確認、補正し、別の推薦結果を繰り返し提示していくことで、潜在的なニーズを探索しながら、ユーザが満足する解まで誘導する必要がある。また、ユーザとの対話を通じた情報収集と推薦結果の提示とを混在させ、全体的にそのやりとりの回数を減らすなどの効率化が求められる。このような、人間のように会話の中で自然に推薦を織り交ぜるような発話をするシステムの構築は難しく、先行研究の多くは通常の発話と推薦を切り離れたシステム設計になっている。本稿ではユーザの発話情報を利用した推薦システムの設計と、その推薦に基づいた強化学習による対話戦略を定義し、人間のように通常の発話と推薦とを混在させた臨機応変な振る舞いを行う対話システムの提案をする。

## 2. 関連研究

モデルベースの目的指向型対話システム構築の試みは、近年、数多く行われており、中でも強化学習を用いた対話に関する研究は注目を集めている。

[Xiujun 17] らは映画のチケット予約に関する対話タスクにおいて、深層強化学習による対話戦略を実現した。ここでは、ユーザの発話理解からエージェント <sup>\*1</sup> の応答生成までを End-to-End に学習させる枠組みを提案している。しかし Xiujun らの研究では、映画予約というタスクの性質上、ユーザの目標とする映画はすでに決まっているものと仮定しており、エージェントによる推薦は行われていない。

[Yueming 18] らは対話における推薦問題に焦点を当てた研究をしている。レストラン予約のタスクに対し、深層強化学習による対話戦略に加え、ユーザの過去の行動履歴をもとに推薦をするモデルを提案し、高精度な推薦システムを用いて効率的なやりとりを行うことに成功している。しかしここでの対話戦略には、推薦システムの出力は考慮されていない。

## 3. 問題設定

提案するシステムの学習、評価に [Chaitanya 17] らの作成した対話データセット (以下 Personalized Dialog dataset と呼ぶ) を使用した。このデータセットはユーザとシステムのレストラン予約に関するやりとりから、次の発話を予測するというタスクに対して作成されたものである。「性別、年齢層、好みの料理、菜食主義か否か」といったユーザのプロファイル情報が与えられた状態で、予約完了までに行った会話の全記録が提供されている。また、知識ベースとして 2,400 店舗のレストラン情報が用意されており、「料理ジャンル、場所、価格帯、人数」の条件、いわゆるファセット項目でレストランを絞り込むことができる。本研究では Personalized Dialog dataset を基に、レストランを探すユーザに対して、やり取りの中でユーザが満足する店舗を予測する推薦システムと、その予測を考慮した上で発話を行う対話システムを構築する。

連絡先: 積田大介, 明治大学大学院理工学研究科情報科学専攻, 〒214-8571 神奈川県川崎市多摩区東三田 1-1-1, tsumidai@cs.meiji.ac.jp

<sup>\*1</sup> 本稿では提案する対話システム全体に対し、対話戦略をとるモデル自体をエージェントと呼ぶ。

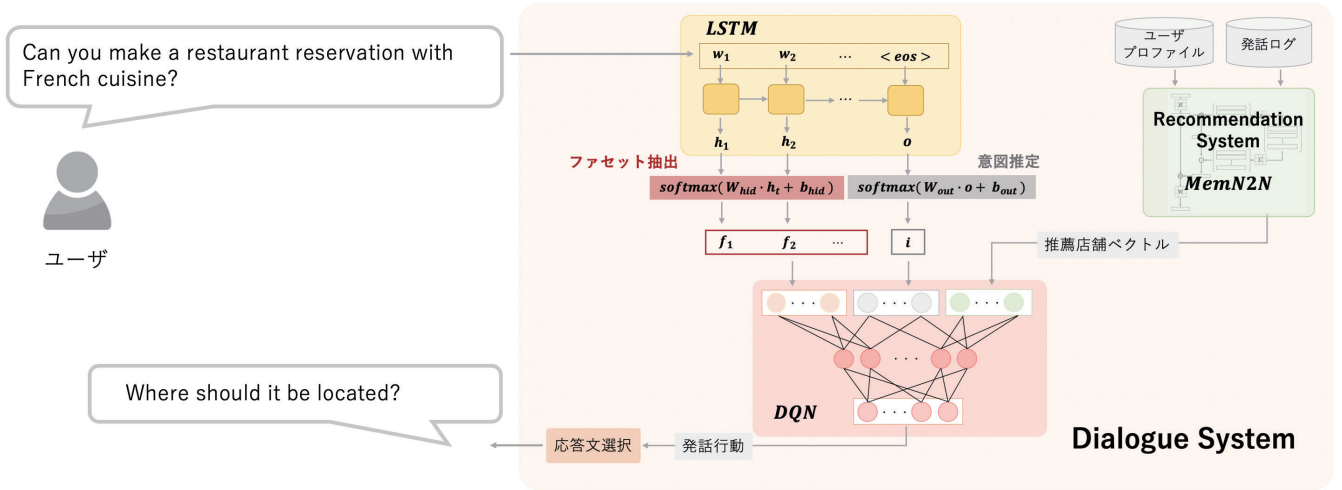


図 1: 提案システムの構成図

## 4. 提案システム

本研究では Xiujun らの対話システムをベースに深層強化学習の枠組みで学習を行うエージェントを設計し、その対話戦略に推薦システムによる予測を利用する。図 1 に提案システムの構成を示す。以下、本節では提案システムの各モジュールの説明を記述する。

### 4.1 発話意図推定とファセット抽出

推薦を行う目的指向型の対話システムにおける機能として、ユーザの発話からそのユーザがどのような意図で発言をしたかを推定することが求められる。同様に、レストラン予約に求められるファセット項目を抽出する必要がある。本研究ではユーザの発話意図推定とファセット項目抽出を同時に行うモデルとして、生成モデルや時系列予測などの分野で成果を上げている Long short-term memory (LSTM) を使用する。Personalized Dialog dataset に含まれるユーザの発話文から、数種類のユーザの発話意図パターンとファセット項目を抽出し、このモデルの学習用の教師として使用した。以下に LSTM による予測の流れを示す。

1. 発話文を単語に分割し、単語ベクトルに変換したものを、LSTM への入力とする。単語ベクトルは総単語数の次元を持ったワンホットベクトルで、単語に対応した次元位置に 1 が入る形をとる。発話文の  $t$  番目に登場した単語を  $w_t$  とし、LSTM から得られる隠れ層の出力の内、最終出力のベクトルを  $o$ 、ベクトル  $o$  以前の中間出力のベクトルを  $h_1, h_2, \dots, h_n$  とする。また、最終セルには文の終了タグ ( $eos$ ) を入れる。

$$h_1, h_2, \dots, h_n, o = LSTM(w_1, w_2, \dots, w_n, eos) \quad (1)$$

2.  $h_1, h_2, \dots, h_n$  と  $o$  にそれぞれ異なる重み付けを行い、 $\text{softmax}$  関数に通すことで、それぞれベクトル  $f_1, f_2, \dots, f_n$  と  $i$  を得る。最終出力から得られたベクトル  $i$  がユーザの発話意図を、中間出力から得られた各ベクトル  $f_t$  が単語  $w_t$  のファセット項目を表現するように学習させる。

$$f_t = \text{softmax}(W_{hid} \cdot h_t + b_{hid}) \quad (2)$$

$$i = \text{softmax}(W_{out} \cdot o + b_{out}) \quad (3)$$

ここで得られたユーザの発話意図とファセット項目に関するベクトルを対話戦略の入力として利用する。

### 4.2 推薦システム

ユーザの過去の発話とプロフィール情報を基に、ユーザの嗜好に合った店舗を予測する推薦システムについて説明する。

本研究では、質問応答の分野で提案され、対話システムの発話予測問題においても、Recurrent neural network (RNN) や LSTM といった他の深層学習のモデルを上回る精度で注目されている End-to-End Memory Networks (MemN2N) [Sainbayar 15] の枠組みによる推薦システムを利用する。

質問応答の分野で使用される MemN2N は過去の発話を知識情報として記憶し、質問を受けると、質問文を基に記憶を参照して回答を導き出す。すなわち、過去の発話情報と質問文などの外部入力から応答文を出力するモデル設計になっている。

本研究における問題設定では、質問文の代わりにユーザのプロフィール情報を使用し、特定のプロフィール情報を持つユーザに対して、発話情報を参照した店舗推薦を行うシステムを構築した。MemN2N による推薦システムの出力は総店舗数の次元を持つベクトルであり、各次元位置に該当する店舗の推薦確率を表現した形をとる。

Personalized Dialog dataset から発話ログ情報とユーザプロフィール、実際にユーザが受け入れた店舗を抽出し、推薦システムのモデル学習に使用した。学習に用いた教師は各店舗に該当する次元位置が 1 となるワンホットベクトルであり、発話ログは対話システムが推薦を行う前までのものを使用した。

推定したユーザの発話意図とファセット項目同様に、推薦システムの出力ベクトルも対話戦略の入力として使用する。

### 4.3 対話戦略

本研究では Xiujun らの対話システム同様に、深層強化学習による対話戦略をとるエージェントを構築する。

強化学習における問題設定では、環境を定義し、その環境下で動作するエージェントに対し、状態を与える。エージェントは受け取った状態から、自身の持つ方策を基に次にとるべき行動を決定する。環境はエージェントの行動に対し、次の状態とその行動の良し悪しを定量化した報酬を与える。以上の流れを 1 ステップと定義し、エージェントが目標を達成する、あるいは

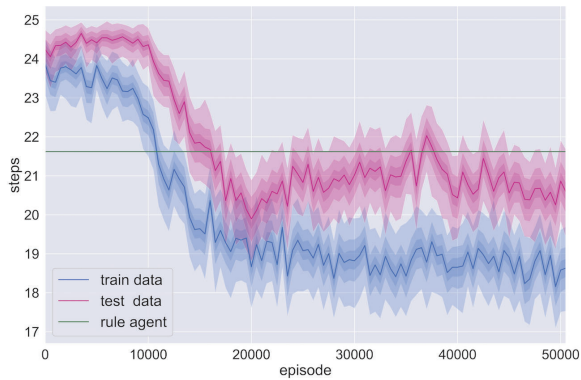


図 2: 対話システムの発話回数

は特定の条件で目標達成の失敗とみなされるまで繰り返す。このサイクルをエピソードと呼び、エージェントはエピソードを繰り返していく中で、自身の方策によって求められる行動価値を更新していく。Deep Q-Network(DQN)[Volodymyr 13] は行動価値を多層ニューラルネットワークで関数近似した深層強化学習の中でも基本的な手法である。本研究では、DQN の枠組みでの対話戦略を実現し、ユーザの発話意図とファセット項目、推薦システムの出力から次にとるべき発話の行動価値を最大化するような学習を行う。

## 5. 環境設計

強化学習の問題設定では、エージェントの発話行動に対して臨機応変にやりとりを行う環境が必要である。[Xiujun 16] らは映画予約に関する対話データをクラウドソーシングで収集し、ルールベースで応答文を選択するユーザシミュレータを作成した。

本研究では Xiujun らのユーザシミュレータを参考に、Personalized Dialog dataset から抽出したユーザ発話を用いてシミュレータを作成し、強化学習における環境として扱う。ユーザシミュレータの仕様を以下に示す。

- ユーザはプロフィール情報「性別、年齢層、好みの料理、菜食主義か否か」とレストランに求めるファセット項目を持ち、最終的に受け入れるレストランが必ず存在する。
- ユーザの持つファセット項目は会話の途中で移り変わる可能性がある。
- ユーザはエージェントに対して常に協力的な応答をする。(ファセット項目等の質問に対しては必ず答える)
- 一度のユーザの発話を状態、それに対するエージェントの発話を行動とし、ある状態において行動をとることを 1 ステップとする。
- ユーザはエージェントの行動に対し、適当な報酬を与える。
- エージェントによりユーザの満足するレストランが推薦された場合、エピソードを終了する。ステップ数が 25 に達しても終了する。

## 6. 評価実験

5. 章で説明したユーザシミュレータ環境を用いて、提案システムの検証を行う。ユーザのプロフィール情報、目的とする

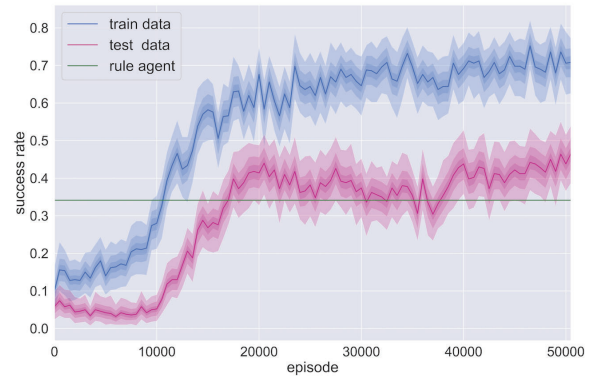


図 3: 対話成功率

ファセット項目を学習用、評価用に 12,000 件ずつ用意し、これらを基にシミュレータが発話を行うようにした。

### 6.1 評価方法

エージェントの学習が経過するにつれて、提案システムの発話回数と対話の成功率がどのように変化するかを確認する。

本研究では対話システムの発話回数が短いほど、ユーザにとって必要な労力を省くことができるため、良い対話戦略と考える。対話の成功率に関しては、店舗の推薦成功率と同義であり、ユーザがシステムの推薦を受け入れた時点で対話の成功とみなす。これらの指標に関して、ファセット項目を全て聞き出してから店舗推薦を行うルールベースのエージェントによる対話システムとの比較を行う。また、学習後の対話システムによるやりとりの内容を確認し、定性的な評価も行う。

### 6.2 実験結果

図 2 に 1 エピソードに要した対話システムによる発話回数の推移を示す。図中、*train data* と *test data* はそれぞれ学習データと評価データに対する提案システムの発話回数を表している。また *rule agent* はルールベースの対話システムによる、評価データに対する発話回数である。ルールベースのシステムがおおよそ 21.6 回の発話を要したのに対し、提案システムのエージェントでは、学習後の状態で、評価データに対してもおおよそ 20.5 回 (図 2 *test data* のグラフ右端の時点) と効率の良いやりとりを行うことができています。

図 3 に示した各グラフは図 2 と同様の条件での対話成功率を表している。エージェントの学習を繰り返すにつれて、対話成功率は向上し、ルールベースを上回る結果になっている。

表 1 は学習後の提案システムによる実際のやりとりの内容である。ルールベースによる対話システムの場合、全てのファセット項目を聞き出し、店舗の推薦行動に移ることになるが、表 1 左側に示した提案システムの対話では、推薦の確信度に応じて必要なファセット項目のみを聞き出し、効率良く対話を成功させていることがわかる。一度目の推薦のように上手くいかない状況になると、再度ファセット項目を聞き出し、ユーザが真に求める条件を探る働きかけを行う様子が確認できる。表 1 右側は対話失敗とみなされた例である。序盤からファセット項目を全て聞き出しているが、最大 25 回可能な発話の中で推薦が成功しなかったために、対話失敗となっている。

### 6.3 考察

本研究では、対話システムの発話数を短くすることに成功し、対話成功率を向上させた。一方で、ユーザの目的を満足させることのできないやりとりも多く存在した。今回のタスクにおい



表 1: 学習後の提案システムによる対話例

success dialogue example	failure dialogue example
<p><u>User</u>: Can you book a table in Madrid?</p> <p><u>Agent</u>: What type of cuisine would you like to eat?</p> <p><u>User</u>: With Italian food.</p> <p><u>Agent</u>: May I suggest this option: <b>resto_madrid_cheap_italian_7stars_2</b></p> <p><u>User</u>: No I don't like that.</p> <p><u>Agent</u>: Is there any change?</p> <p><u>User</u>: Instead could it be with British cuisine.</p> <p><u>Agent</u>: Is there any change?</p> <p><u>User</u>: No.</p> <p><u>Agent</u>: May I suggest this option: <b>resto_madrid_moderate_british_7stars_1</b></p> <p><u>User</u>: Let 's do it.</p>	<p><u>User</u>: Can you book a table with French cuisine in Bombay?</p> <p><u>Agent</u>: Would you mind telling me your price range?</p> <p><u>User</u>: I am looking for a moderate restaurant.</p> <p><u>Agent</u>: How many people would be in your party?</p> <p><u>User</u>: I'd like to book a table for two people.</p> <p><u>Agent</u>: Is there any change?</p> <p><u>User</u>: Instead could it be in a cheap price range?</p> <p><u>Agent</u>: Is there any change?</p> <p><u>User</u>: No.</p> <p><u>Agent</u>: May I suggest this option: <b>resto_bombay_cheap_french_7stars_2</b></p> <p><u>User</u>: Do you have something else?</p> <p><u>Agent</u>: May I suggest this option: <b>resto_bombay_cheap_french_3stars_1</b></p> <p><u>User</u>: No I don't like that.</p> <p><u>Agent</u>: Is there any change?</p> <p><u>User</u>: No.</p> <p><u>Agent</u>: May I suggest this option: <b>resto_bombay_cheap_french_8stars_1</b></p> <p><u>User</u>: No I don't like that.</p> <p><u>Agent</u>: Is there any change?</p> <p><u>User</u>: No.</p> <p><u>Agent</u>: May I suggest this option: <b>resto_bombay_cheap_french_4stars_2</b></p> <p><u>User</u>: Do you have something else?</p> <p><u>Agent</u>: May I suggest this option: <b>resto_bombay_cheap_french_3stars_2</b></p> <p><u>User</u>: No I don't like that.</p> <p><u>Agent</u>: May I suggest this option: <b>resto_bombay_cheap_french_1stars_1</b></p> <p><u>User</u>: No I don't like that.</p> <p>...</p>
推薦に必要なファセット項目のみを聞き出し、 効率の良いやりとりを行うことに成功している。	何度も推薦に失敗し、最終的にユーザの満足のいく 店舗を推薦することが出来ずに終了している。

て、対話の成功には推薦システムの性能が大きく関係しているため、推薦システムがユーザの嗜好をうまく汲み取れない状況において、対話の成功は非常に困難になる。また、本研究では一度の対話で行うことができる発話回数が限られている。そのため、ルールベースの対話システムのように、対話の序盤にファセット項目を全て聞き出し、その後、立て続けに推薦する対話戦略を用いたとしても、対話成功は難しいと考えられる。以上のことから、対話に失敗したエピソードが対話戦略として悪いものであるとは、一概に断定することはできない。事実、表 1 右側の失敗した対話例のように、全てのファセット項目を聞き出すことが出来ていても、推薦が原因で失敗した対話が多く存在した。この対話例は、推薦の確信度の低さや立て続けの推薦失敗という状況から、ファセット項目の修正が無いかを再度確認している点で、対話戦略としては、むしろ優れていると考えられる。

## 7. おわりに

本稿では、ユーザの目標が明示的でない状況での目的指向型対話を取り上げ、ユーザに対する推薦が求められる問題設定において、発話情報を利用した推薦システムの予測に基づいた対話戦略をとるシステムを提案した。さらに、ファセット項目の収集と推薦を混在させ、状況に応じて柔軟に発話をするような、人間らしい効率的なやりとりを行う対話システムの実現に成功した。一方で、本研究ではユーザ、システム共に非常に限ら

れた発話行動の中でしか検証ができていない。また、レストラン検索という限定的な問題設定下での対話システムになっている。今後はレストラン検索等の単一のタスクだけでなく、オープンドメインで多様な発話行動を想定した対話システムを構築する必要がある。

## 参考文献

- [XiuJun 17] XiuJun Li, Yun-Nung Chen, Lihong Li, Jianfeng Gao, Asli Celikyilmaz : End-to-End Task-Completion Neural Dialogue Systems, IJCNLP (2017)
- [Yueming 18] Yueming Sun, Yi Zhang : Conversational Recommender System, SIGIR (2018)
- [Chaitanya 17] Chaitanya K. Joshi, Fei Mi, Boi Faltings: Personalization in Goal-oriented Dialog, NIPS (2017)
- [Sainbayar 15] Sainbayar Sukhbaatar, Arthur Szlam, Jason Weston, Rob Fergus: End-To-End Memory Networks, NIPS (2015)
- [Volodymyr 13] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, Martin Riedmiller: Playing Atari with Deep-Reinforcement Learning, NIPS Deep Learning Workshop (2013)
- [XiuJun 16] XiuJun Li, Zachary C. Lipton, Bhuwan Dhingra, Lihong Li, Jianfeng Gao, Yun-Nung Chen : A User Simulator for Task-Completion Dialogues, (2016)