

マルチモーダル情報を扱う確率的対話システムの開発

飯島采永 (指導教員：小林一郎)

1 はじめに

近年、家庭用ロボットが多く普及されてきている。ロボットと共に生活していく上で、ロボットのコミュニケーション能力のさらなる充実が、今後、益々必要と考えられる。そこで、本研究では家庭用ロボットの身体性を利用し、生活を見守るためのシステムの構築を目的とし、その実現に向けて、マルチモーダル情報を用いた部分観測マルコフ決定過程に基づいたロボットとの対話処理と、ロボットの観測に基づく1日の要約文やそれに基づく質問応答の作成に取り組む。また、ユーザが対象のドメインに関する知識を十分に有していないために、対話システムを利用して意思決定を行うために必要な情報をシステムに相談し収集する状況について考える。システムがユーザの潜在的な嗜好を推定した上で、ユーザが興味を持つ情報を推薦する必要がある。本研究ではユーザの嗜好を考慮したレシピ提案システムの構築も目指す。

2 ロボットとのマルチモーダル対話

2.1 マルチモーダル情報の観測

2.2 マルチモーダル情報の観測

本研究では、実環境での観測情報の不確実性を考慮するため、部分観測マルコフ決定過程 (POMDP: Partially Observable Markov Decision Process) の枠組みを用いる。

一般的に POMDP の観測状態は $\{S, A, T, O, Z, R, b_0\}$ で表される。 $s \in S$ はユーザ状態、 $a \in A$ はシステムの行動を表す。また、 T は行動 a によって状態 s が s' へと遷移する確率 (状態遷移確率 $P(s'|s, a)$) の集合であり、 $o \in O$ はユーザから観測される観測値を表す。 Z は行動 a によって状態が s' に遷移し、観測値 o' が観測される確率 (観測値出力確率 $P(o'|s', a)$) の集合である。 $r(s, a) \in R$ は状態 s で行動 a を行った時の報酬を表す。

POMDP では、観測値 o から直接観測できない状態 s を確率分布として推測し、その分布を信念状態 $b(s)$ とする。初期信念状態を b_0 と表す。信念状態 $b(s)$ が既知のとき、状態遷移確率と観測値出力確率により、次の時刻の信念状態 $b'(s')$ は式 (1) の漸化式で記述される。

$$b'(s') = k \cdot P(o'|s', a) \sum_s P(s'|s, a) b(s) \quad (1)$$

ここで係数 k は $\sum_s b'(s') = 1$ を満たす正規化項である。

2.3 マルチモーダル状態表現への拡張

ユーザとのインタラクションを想定して、3つのユーザ状態、心理状態 s^e 、物理状態 s^p 、言語状態 s^l を考える。

マルチモーダル対話例を表1に示す。

表 1: 対話例

話者	発話・行動	観測	$b(s)$
ユーザ Pepper	(遠くにいる) 近くにおいでよ	o^p (遠い)	
ユーザ Pepper	(近くにくる) 僕とお話しよう	o^p (近い)	
ユーザ Pepper	こんにちは こんにちは	o^l (こんにちは)	
ユーザ Pepper	(暗い顔) 疲れた顔をしてるね	o^e (悲しい)	

3 1日の報告文の生成

次に、ロボットが観測した内容を時系列データとしてデータベースに記録し、1日分のデータに対する報告文の生成を行う。今回は記録する情報として、{日付, 時間, ユーザ名, 距離情報, 表情情報, ユーザの発話内容, ロボットの発話内容}と設定する。ここで、ユーザ名とは対話、もしくは観測した一人の相手のことを指す。距離情報はロボットからの距離を記録し、表情情報は、5種類の表情 {無表情, 幸せ, 驚き, 怒り, 悲しみ} を数値で記録する。

3.1 重要情報の抽出

あるデータにおいてひとつ前のデータと比較し、変化のない場合にはそのデータを削除することにより、データベースの冗長性とデータサイズを削減する。データベースへ格納した時系列データに重要度を付与することで、重要な情報を効率的に伝えることができるようになる。

3.2 発話内容の特定

ユーザとロボットとの発話の内容の特定を行う。今回は会話の内容を2つのトピック {天気, ご飯} に分類する。トピックの分類のため、Wikipedia における「天気」¹ と「飯」² の概要と、時間で区切った範囲のユーザとロボットの対話内容を単語分割し、文書ベクトルを作成する。それらの文書ベクトルに対して cos 類似度を測ることで、ユーザとロボットの対話内容に最も近いトピックを推定する。

3.3 テンプレートによる文生成

重要度に対して適切な閾値を設定し、その閾値よりも値の高い文を選ぶことによって重要な情報を取り出し、その情報にテンプレートを適用することにより要約文を生成する。重要データが {14:55, iijima, 0.69, {0.02, 0.10, 0.51, 0.05, 0.32}, -, -} の時には「“いつ”、“だれ”と会いました。その人は“どんな様子”でした。」というテンプレートを使用し、「14時55分にも飯島さんと会いました。少し驚いたような様子でした。」という文が生成できる。

¹<https://ja.wikipedia.org/wiki/天気>

²<https://ja.wikipedia.org/wiki/飯>

4 ユーザの意図を考慮した料理推薦システムの構築

また、ユーザの満足度の高い対話を達成させるために、ユーザの発言には現れていない潜在的な考えを推測し、その考えに則した対話戦略の構築を目指す。今回はユーザの意図を汲み、適切なレシピを推薦することを目標とする。

4.1 ユーザの意図理解

相談型の対話を扱うために POMDP を用いたユーザの知識と嗜好を考慮する対話状態のモデルが提案されている [2]。その中で、ユーザの意図を推測するために、意思決定支援タスクでは代表的な手法として階層分析法 (AHP 法) [3] が利用されている。AHP 法では、問題の要素を「最終目標」「評価基準」「代替案」の三階層に分け、ユーザの各評価基準に対する局所重み (重要度) を推定することにより最適な意思決定を行う。レシピ推薦システムにおける最終目標は、ユーザ自身の嗜好に合ったレシピを決定することであり、代替案はシステムが提示できるレシピのリストである。評価基準には「料理を決める際に重視しているもの」を、実際に 20 代の女性 8 名にアンケートをとり、100 個の評価基準を選んだ。アンケートには評価グリッド法という、人間が何を知覚して、その結果どのような評価を下しているのかという認知構造を同定するための方法を用いた。図 1 に例を示す。評価基準 m が M 個、代替案 (レシピ) n が N 個あるとき、ユーザにとって最適なレシピの推薦を行うために、ユーザの評価基準に対する重み $P = (p_1, p_2, \dots, p_M)$ 、各代替案に対する各評価基準の観点からの重み $V = (v_{11}, v_{12}, \dots, v_{1M}, \dots, v_{NM})$ を決定する。ユーザにとって最適なレシピの候補 l は式 (2) に従い選択する。

$$\arg \max_l \sum_{m=1}^M p_m v_{lm} \quad (2)$$

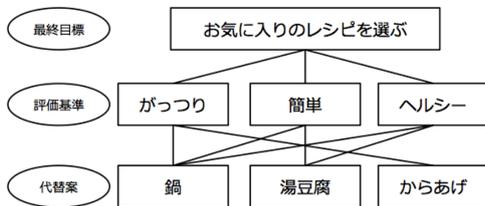


図 1: AHP 法における階層構造

4.2 ユーザシミュレータの構築

システムはユーザの内部状態を直接観測することができないためユーザとのインタラクションから推定する。システムが推定するユーザの嗜好の状態を表す確率分布として、 $P_{sys} = (Pr(p_1 = 1), \dots, Pr(p_M = 1))$ 、ユーザの知識の状態を表す確率分布として、 $K_{sys} = (Pr(k_1 = 1), \dots, Pr(k_M = 1))$ と表す。ユーザの状態 $S = \{s_1, \dots, s_M\}$ は嗜好の確率分布と知識の確率分布をかけたものである $P_{sys} K_{sys}$ を利用し、 $s_m = Pr(p_m = 1) Pr(k_m = 1)$ が成り立つ。システムが推定するユーザの嗜好 P_{sys} はベイズ則を適用することで更新される (3)。ここで $I^t = (a_{sys}^t, a_{user}^t)$ はユーザとシステム間のインタラクションを表す。

$$Pr(p_m = 1 | I^t) = \frac{Pr(I^t | p_m = 1) Pr(p_m = 1)}{Pr(I^t | p_m = 1) Pr(p_m = 1) + Pr(I^t | p_m = 0) Pr(p_m = 0)} \quad (3)$$

報酬関数は下式 (4) のように設定した。ここで l はユーザが決定したレシピを指す。

$$R = \sum_{m=1}^M p_m v_{lm} - \frac{1}{N} \sum_{n=1}^N \sum_{m=1}^M p_m v_{nm} \quad (4)$$

システムの行動は、以下の 6 つの推薦手法からソフトマックス方策によって選択する。

- 現在話題のレシピに関する情報推薦
- 現在話題の評価基準に関する情報推薦
- オープンプrompt
- 推定したユーザの知識が低い順に決定要因の提示
- 推定したユーザの知識が高い順に決定要因の提示
- 推定したユーザの嗜好に従ってレシピの推薦

また、ソフトマックス方策におけるパラメータ θ を Natural Actor Critic を用いて最適化する。パラメータ $\theta = (\theta_{11}, \theta_{12}, \dots, \theta_{1M}, \dots, \theta_{AM})$ は、 A (行動数、今回は 6 つ) $\times M$ (状態の特徴数) 個のパラメータと置く。

$$\pi(a_{sys} = a' | S) = Pr(a_{sys} = a' | S, \theta) = \frac{\exp(\sum_{i=1}^M s_i \theta_{a'_i})}{\sum_{a=1}^A \exp(\sum_{i=1}^M s_i \theta_{a_i})} \quad (5)$$

各試行におけるパラメータの更新した後の方策と報酬を以下の図 2 に示す。調整したパラメータを利用することで、対話に最適な行動を選択することができた。

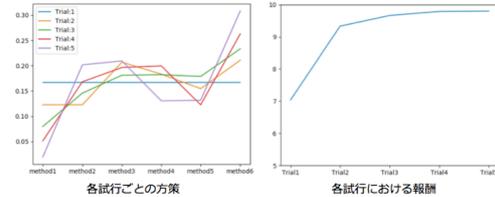


図 2: パラメータの更新と報酬の関係

5 まとめ

本研究では、Pepper を対象にした見守りシステムではマルチモーダル対話を POMDP の枠組みに沿って実装し、それに基づいた要約文生成を行った。更に、ユーザの嗜好を考慮した対話を達成するために、ユーザのレシピ決定を支援する対話システムのユーザシミュレータの構築を行い、方策を学習した。

参考文献

- [1] 遠藤充, 牛尾貴志, 山上勝義, 堀井則彰 “発話とコマンドの系列を制御する対話エージェント”, 第 31 回人工知能学会全国大会, 2017.
- [2] 翠輝久, 大竹清敬, 堀智織, 柏岡秀紀 “意思決定を支援する音声対話システム” 言語処理学会 第 18 回年次大会発表論文集, pp.658-661, 2012.
- [3] T. Saaty. The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation. McGraw-Hill, 1980.