6X-09

POMDPを用いたマルチモーダル対話への取り組み

飯島采永 †
† お茶の水女子大学理学部情報科学科

小林一郎 ‡
‡ お茶の水女子大学 基幹研究院自然科学系

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たインタラクションの実現を目的とし，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 (POMDP) に基づく対話処理に取り組む．

2 POMDPを用いたマルチモーダル対話

2.1 Pepperを用いたマルチモーダル情報の観測

ロボットはソフトバンク社製感情認識ヒューマノイドロボット Pepper を使用する．Pepper の様々なセンサを利用して，マルチモーダルな情報を取得し，それに基づくコミュニケーションを実現する．具体的には，マイクから音声情報，RGB カメラから表情などの画像情報，タッチセンサから触覚情報，レーザセンサから距離情報を取得する．画像情報を用いた顔認識では，ユーザに対して，個体の識別，年齢の推定，笑顔度の判定，5 種類の表情 { 無表情, 幸せ, 驚き, 怒り, 悲しみ } の識別を行うことが出来る．

2.2 POMDP

本研究では, 実環境で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．図 1 に POMDP のグラフィカルモデルを示す．

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．

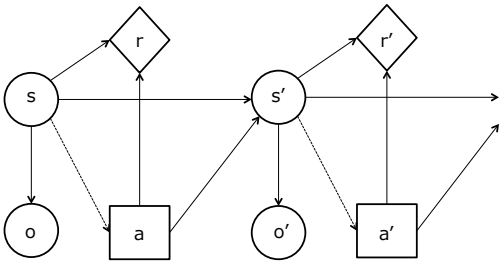

図 1: POMDP のグラフィカルモデル

POMDP では状態 s は直接観測できないので，観測値 o から推測した確率分布を扱う．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 $b'(s')$ は以下の漸化式で記述される．

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

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

3 マルチモーダル対話実験

Pepper との対話にマルチモーダル情報を使ったシナリオを POMDP の枠組みに沿って動かす．今回は文献 [1] を実装した，GitHub 上で公開されている Python のコード * を改良して用いた．

3.1 実験設定

Pepper の制御には専用のソフトウェアとして公開されている PythonSDK† を利用した．POMDP の設定は以下のように定義する．ただし今回の実験では状態遷移確率，観測値出力確率，報酬，初期信念状態は人手で設定をした．

- S：{ 挨拶, 悲しい, 楽しい, 嬉しい, 嬉しくない }
- A： { なにもしない, 挨拶する, 励ます, 楽しくなる, 照れる, へこむ }
- T：$P(s'|s,a)$
  表 1 に状態 s から s' への遷移確率を示す．本実験では，どの行動に対しても同じ遷移確率を与えた．

An Approach to a Multimodal Dialogue System using POMDP
†Sae IIJIMA(your-email@is.ocha.ac.jp),
‡Ichiro KOBAYASHI(koba@is.ocha.ac.jp)

*https://github.com/mbforbes/py-pomdp
†http://doc.aldebaran.com/1-14/dev/python/index.html

Copyright ©2016 Information Processing Society of Japan.
All Rights Reserved.

表 1: 状態遷移確率

s ＼ s'	挨拶	悲しい	楽しい	嬉しい	嬉しくない
挨拶	0.2	0.25	0.25	0.15	0.15
悲しい	0.2	0.15	0.15	0.25	0.25
楽しい	0.2	0.15	0.15	0.25	0.25
嬉しい	0.3	0.2	0.2	0.15	0.15
嬉しくない	0.3	0.2	0.2	0.15	0.15

- O： $\{o, o', ..., o\}$
  音声情報, 画像情報, タッチセンサ情報を観測する.
- Z：$P(o'|s', a)$
  音声情報を正しく観測する確率を 0.8，画像情報とタッチセンサ情報を正しく観測する確率を 0.7 とした．また，状態遷移確率と同様に，どの行動の後でも観測値出力確率は変わらないように設定した.
- R：$r(s, a)$
  各行動後に逐次的に与える報酬を表 2 に示す.

表 2: 報酬

s ＼ a	なにもしない	挨拶する	励ます	楽しくなる	照れる	へこむ
挨拶	-1	5	-10	-10	-10	-10
悲しい	-1	-10	5	-10	-10	-10
楽しい	-1	-10	-10	5	-10	-10
嬉しい	-1	-10	-10	-10	5	-10
嬉しくない	-1	-10	-10	-10	-10	5

- b_0：初期信念状態
  初期信念状態は，はじめに挨拶する確率が高いとして，$b_0 =$ (挨拶：0.3, 悲しい：0.2, 楽しい：0.2, 嬉しい：0.15, 嬉しくない：0.15) とした.

3.2 対話シナリオ

マルチモーダル対話シナリオを表 3 に表す.

表 3: マルチモーダル対話シナリオ

話者	発話内容	観測情報	行動
ユーザ	こんにちは	音声情報	
Pepper	こんにちは		挨拶する
ユーザ	(暗い顔)	画像情報	
Pepper	元気ないですね 僕が励まします		励ます
ユーザ	ありがとう (頭をなでる)	音声情報 センサ情報	
Pepper	照れるなあ		照れる

3.3 実験結果

実験結果を表 4 に表す．観測 o より，状態 s を推測する．1 段目では, ユーザが「こんにちは」と発話したという観測に対し，$b(s) =$ (挨拶：0.873, 悲しい：0.036, 楽しい：0.036, 嬉しい：0.027, 嬉しくない：0.027) を得た．これにより確率の一番大きい「挨拶」という状態を推測した.「挨拶」状態から行動遷移確率に従い,「挨拶する」行動を選択している．ここで，割引率 γ は 0.95 として設定した.

表 4: 実験結果

話者	発話内容	実行結果
ユーザ Pepper	こんにちは こんにちは	o:音声 [こんにちは], $b(s) = (0.873, 0.036, 0.036, 0.027, 0.027)$, a:挨拶する, $r = 6.97653155$
ユーザ Pepper	(暗い顔) 元気ないですね 僕が励まします	o:画像 [暗い顔], $b(s) = (0.045, 0.73, 0.157, 0.034, 0.034)$, a:励ます, $r = 6.95527813429$
ユーザ Pepper	ありがとう (頭をなでる) 照れるなあ	o:音声 [ありがとう] & センサ [頭に触れる], $b(s) = (0.045, 0.034, 0.034, 0.73, 0.156)$, a:照れる,$r = 8.23924008711$

3.4 考察

実験結果より，想定したシナリオを POMDP の枠組みに沿って動かすことが出来たことを確認した.
しかし，今回の実験ではマルチモーダル情報の観測を逐次的に処理することしかできていない．例えば，ユーザが「ありがとう」と発話し,「頭をなでる」という行動を観測するのに，まず音声情報を観測し，その後タッチセンサ情報を観測している．異なる複数のモダリティの情報が，同時に与えられることで意思決定がされ，行動につなげるという枠組みが必要と考えられる.

4 まとめと今後の課題

本研究では，Pepper を対象にしたマルチモーダル対話を POMDP の枠組みに沿って実装を行った．今後の課題として，今回は人手で設定した状態遷移確率，観測値出力確率を強化学習を用いて自動で獲得することが挙げられる．これには Q 学習を取り入れていくつもりである．また，複数のモダリティが入力情報として同時に与えられた時の行動選択に対しても考察を行うつもりである.

参考文献

[1] Jason D. Williams, Steve Young“Partially observable Markov decision processes for spoken dialog systems”, Computer Speech and Language, Volume 21 , Issue 2, pp. 393-422, 2007.

[2] 盧迪, 久保 伸太郎, 深山 覚, 中沢 正幸, 西本 卓也, 嵯峨山 茂樹 “マルチモーダル入力と強化学習による擬人化エージェントの対話制御の検討”, The 24th Annual Conference of the Japanese Society for Artificial Intelligence, 2010.

[3] 吉野幸一郎, 河原達也 “ユーザの焦点に適応的な雑談型音声情報案内システム”, 言語処理学会 第 20 回年次大会発表論文集, pp. 761-764. 言語処理学会, 2014.

Copyright ©2016 Information Processing Society of Japan.
All Rights Reserved.