

ユーザの文法知識を状態に加えた POMDP に基づく音声対話システム

穂山 空道[†] 駒谷 和範[‡] 高橋 徹[‡] 尾形 哲也[‡] 奥乃 博[‡]
[†] 京都大学 工学部情報学科 [‡] 京都大学大学院 情報学研究科 知能情報学専攻

1. はじめに

近年、音声対話システムの対話管理に強化学習、特に Partially Observable Markov Decision Process (POMDP) を用いる研究が行われている [1]。この研究では、複数仮説を保持することによって音声認識誤りに対処した対話戦略を自動獲得できる。しかし、従来の POMDP を用いた音声対話システムでは、POMDP の観測は言語理解結果である。実際の対話においては言語理解結果が得られないことがあり、この場合従来手法では情報が得られない。本研究ではこのような場合に、言語理解結果だけでなく、発話全体概意を観測することで適切なシステム応答の生成を図る。本稿では発話の概意を、ユーザが意図した言語理解文法の種類と定義する。

我々は文法検証結果を POMDP に統合することで、言語理解結果が得られない場合にも発話の概意を得て対話の進行を可能にする。文法検証結果とは、システムの持つ全ての文法と音声認識結果の間の距離を、音声認識結果中の任意の単語が挿入・置換・削除誤りである可能性を考慮して計算する処理である [2]。文法検証結果は内容語が獲得できていなくても計算できるため、言語理解結果が得られない場合でも有益な情報を得ることができる。

2. POMDP の対話システムへの適用

2.1 POMDP とは

POMDP とは変化する系をモデル化する方法の 1 つである。POMDP では系を「完全に観測やコントロールが可能なエージェント」と「完全な観測は不可能な環境」の二者でモデル化する。環境は変化する状態 s を持っており、またエージェントは s に関する確率分布 $p(s)$ を持っている。これを信念状態と呼び $b(s)$ と書く。このときエージェントがとるべき最良のアクション a を強化学習によって決定することが POMDP 最適化問題である。最適化によって得られる b から a への写像 π を政策と呼ぶ。ここで「最良」とは環境がエージェントに与える報酬の和を最大化することなど、タスクによって設計者が設定する。

POMDP の動作を説明する。まずエージェントは学習済みの政策を用いて $a_0 = \pi(b_0)$ を選択し実行する。 b_0 は信念の初期値で、全ての s について $b_0(s) = \frac{1}{|S|}$ などとする。次に環境は a_0 に合わせて初期状態 s_0 から確率 $p(s_1|s_0, a_0)$ で次の状態へ変化する。最後にエージェントは s_1 と必ずしも一致しない o_1 を確率 $p(o_1|s_1, a_0)$ で観測し、また報酬 $r(s_0, a_0)$ を得て、

$$b'(s') = k \cdot p(o_1|s', a_0) \sum_s p(s'|a_0, s) b(s) \quad (1)$$

によって信念を更新する。ただし k は $\sum_s b(s) = 1$ となるような正規化係数である。以後、エージェントはアクション → 観測 → 信念更新を繰り返す。

Spoken Dialogue System based on POMDP having States about User's Grammatical Knowledge: Soramichi Akiyama, Kazunori Komatani, Toru Takahashi, Tetsuya Ogata, and Hiroshi G. Okuno (Kyoto Univ.)

エージェントをシステム、環境をユーザ、状態とその観測が一致しないことを認識誤りとみれば、音声対話システムを POMDP でモデル化できる [1]。本稿でも POMDP を音声対話システムのモデル化に用いる。

2.2 POMDP への文法検証スコアの統合

本研究では文法検証結果を POMDP へ導入する。その際、「距離」である文法検証結果を POMDP で用いるために「確率」に変換する。また POMDP の最適化が可能になるように設計する。全発話に対して文法検証結果を用いると内容語がの認識の成否に関わらず同じような値が観測される。そのため、状態の各要素を独立に見てアクションを決める最適化手法である Composite Summary Point-Based Value Iteration (CSPBVI) [3] が適用できなくなる。このため、言語理解結果の有無に応じて文法検証結果の使い方を切り替える。

3. 文法知識を反映させるための POMDP の拡張

POMDP を用いた音声対話システムではユーザの真の意図を環境の状態とみて推測する。[1] に従うと、 s はユーザの考える出発地と行先のペア (例えば $s = (\text{京都駅前}, \text{祇園})$) となる。これに対し我々は s_u をユーザの考える出発地と行先のペア、 g を直前の発話でユーザが意図した文法種類として、

$$s = (s_u, g) \quad (2)$$

とする。例えば、 $s = ((\text{祇園}, \text{京都駅}), \text{ToDest})$ は、ユーザが祇園から京都駅前に行きたいという意図を持っていて、かつ直前の発話で行先を言った可能性を表す。

アクション a に対して観測 o' が得られた時の b の更新式は、式 (1) を変形して、

$$b'(s') = k \cdot p(o'|s', a) \times \quad (3)$$

$$\sum_{s_u} \sum_g p(s'_u, g'|a, s_u, g) b(s_u, g)$$

$$= k \cdot p(o'|s', a) \times \quad (4)$$

$$\sum_{s_u} \sum_g p(s'_u|a, s_u, g) p(g'|s'_u, a, s_u, g) b(s_u, g)$$

$$= k \cdot p(o'|s', a) \sum_g p(g'|a) b(s'_u, g) \quad (5)$$

と定める。ただし、(5) ではユーザの考える出発と行先のペアは不変、すなわち $p(s'_u|a, s_u, g)$ は $s'_u = s_u$ のとき 1、それ以外のとき 0 であることと、ユーザの発話種類は直前のシステムアクションのみに依存することを仮定している。

次に観測確率について述べる。 s_u については [4] と同様に音声認識信頼度を用いる。 g について各想定文法との距離を表すスコアが得られるため、これを確率値に変換する。ある発話が文法 G_i を意図していた確率 p_i を次のように定める。ただし G_i は発話の音声認識結果の

N-Best 候補のそれぞれに最も近い文法， S_i は G_i のスコアである．

$$P_i = \frac{\exp(S_i)}{\sum_{j=1}^N \exp(S_j)} \quad (6)$$

2.2 節で述べたようにこれらを常に両方用いると内容語が獲得できているかどうか意図文法に関する信念 $\{\sum_{s_u} b(s_u, g)\}$ のみでは分からず，CSPBVI[3] が適用できない．そこで，内容語が獲得できた時には内容語の認識信頼度のみを用い，内容語が獲得できなかった時には文法検証結果から得られる信頼度のみを用いることとする．

4. 具体例

4.1 タスク

具体例として，京都市バス運行情報案内システム [5] を本稿の手法によって実装する．対話のタスクは以下の通りである．ユーザは京都市バスのバス停から任意の出発バス停，行先バス停を選び，音声を用いてシステムに伝える．システムは音声対話を進めながらユーザがどのバス停からどのバス停に行きたいのかを推測し，ある出発と行先のペアに確定したと判断すればその条件でバス時刻を検索し，ユーザに案内する．タスクの詳細な設定は以下の通りである．

バス停数 652

システムアクション 7 種類: 出発地/行先を両方尋ねる，出発地を尋ねる，行先を尋ねる，出発地/行先を両方確認する，出発地を確認する，行先を確認する，全て確定し時刻を検索する

想定文法 4 種類: どこからを言う，どこまでを言う，どこからどこまでを両方言う，バス停の名前のみを言う

この結果， s のインスタンス数は $652^2 \times 4 = 1700416$ 通りとなる．

4.2 改善される対話例

本稿の文法検証の統合によって発話ターン数が減少する対話例を以下に示す．ただしユーザの発話は実際の音声ではなく，単語とそれに対応する音声認識信頼度を手動で打ち込む形をとっている．また対話戦略 π の学習は行っておらず，システムのアクションは次の3つの値のみを見て，妥当であるものを主観で定めた．ただし，以下では出発地と行先をそれぞれ f, t とし， $s_u = (f, t)$ と書く． f, t には具体的にはバス停が入る．

1. $\max_f \sum_t \sum_g b((f, t), g)$
2. $\max_t \sum_f \sum_g b((f, t), g)$
3. $\max_g \sum_f \sum_t b((f, t), g)$

第3の値は直前の発話で意図された可能性の最も高い文法で，本手法で新たに導入したものである．これらはCSPBVIで最適化を行う際に必要な情報の全てである．すなわち，以下の対話例は自動的に獲得した対話戦略でも実現できると考えられる．

従来法での対話例を以下に示す．ただし sys はシステム発話，usr はユーザ発話を表し，ユーザ発話の後の括弧内は音声認識結果とその信頼度を表す．

sys: ご利用になる停留所をおっしゃってください．
usr: 四条烏丸から [から 0.5]
sys: ご利用になる停留所をおっしゃってください．
usr: 四条烏丸 [しみずちよー 0.5 まで 0.5], [しじょうからすま 0.4]
sys: 清水町までですね．

この例は，ユーザの一発話目の文頭が発話区間検出誤りにより音声認識できていない状況である．人間ならば「から」のみ聞こえれば次には「どこからか」と質問できるが，従来のPOMDPを利用した対話システムでは観測として言語理解結果のみを得るため，この場合何も観測されず，信念更新もされずに最初の質問を繰り返す．これに対し本手法では以下のような対話が可能である．

sys: ご利用になる停留所をおっしゃってください．
usr: 四条烏丸から [から 0.5]
sys: どちらからでしょうか？
usr: 四条烏丸 [しみずちよー 0.5 まで 0.5], [しじょうからすま 0.4]
sys: 四条烏丸からですね．

文法検証を統合したことにより，内容語が獲得できない場合でも出発地を言いたいというユーザの概意が理解でき，次にシステムは「どこから」と尋ねることができる．この結果，さらに次に認識結果が上記のように音声認識誤りを含み，かつ正解である認識結果の信頼度が低かったとしても，「どこから」に対する回答なので「まで」のついた清水町よりも四条烏丸の方が確からしいと判断できる．

5. おわりに

本稿では，POMDPを用いた音声対話システムに文法検証スコアを統合することによって

1. 内容語が認識できない場合でも，想定文法の中で最もユーザの意図している可能性の高いものを推定し
2. その結果対話のターン数が減少する可能性を示した．

実用的なサイズの対話を扱うPOMDPの正確な最適化は計算量が多く不可能である [1][3] が，本稿では既存の近似最適化手法CSPBVI[3]が適用できるように「意図文法に関する信念は内容語が理解できている時には更新しない」と設計したため，CSPBVIによる対話戦略の自動生成が可能である．今後，実際に [3] と組み合わせる本稿で示した可能性を確認する．

謝辞

本研究の一部は，科研費，GCOE，三菱電機の支援を受けた．

参考文献

- [1] J. D. Williams. Partially observable markov decision processes for spoken dialog systems. *Computer Speech and Language*, Vol. 21, No. 2, pp. 393–422, 2007.
- [2] 福林他. 音声対話システムにおける動的ヘルプ生成を指向したWFSTに基づく文法検証によるユーザ知識推定. *SIG-SLUD-A703-09*, pp. 45–50, 2008.
- [3] J. D. Williams and S. Young. Scaling POMDPs for spoken dialog management. *IEEE Transactions on Audio, Speech and Language Processing*, Vol. 15, No. 7, pp. 2116–2129, 2007.
- [4] J. D. Williams et al. Statistical approaches to dialogue systems. *Tutorial Material, InterSpeech*, 2009.
- [5] K. Komatani et al. User modeling in spoken dialogue systems to generate flexible guidance. *UMUAI*, Vol. 15, No. 1, pp. 169–183, 2005.