部分観測マルコフ決定過程を用いたカーナビゲーションシステムにおける音声対話*

◎岸本康秀，滝口哲也，有木康雄 (神戸大・工)

1　はじめに

近年，音声認識をベースとしたインターフェースを備えたカーナビゲーションシステム（以下カーナビ）の実用化が進んでいる．しかし，車室内雑音、発話変形、発話誤りにより誤認識が起こり，その結果として，誤動作が生じ，自動車内で音声認識を利用するユーザが増えない理由となっている [1][2]．また，現在のインターフェースが未熟であるため，誤認識からの回復も困難となっている．

本稿では，不確定な情報に対しても対話を制御させるために，部分観測マルコフ決定過程（POMDP）を用いる．この手法によって，同じ雑音のある状況下で誤認識が起こった場合でも，自然な対話の中で回復することが可能となる．

また，POMDP の状態数の増加に伴い，強化学習の計算量が増大するため今まで大きなタスクを扱うことが困難であった [3]．本研究では複数の方策を組み合わせることにより，従来の POMDP よりも大きなタスクを扱うことが可能となった．本稿では，シミュレーション実験を行い，提案手法の有効性を示す．

2　POMDP の概要

一般的に POMDP は $\{S,\ A,\ T,\ O,\ Z,\ R\}$ で表される．$s \in S$ は状態を表す（システムとユーザの状態）．$a \in A$ はシステム側のアクションを表す．また T はアクション a によって，状態 s が s' へ変わる状態遷移確率 $P(s'|s,a)$ の集合である．$o \in O$ はユーザや環境から観測される観測値を表し，Z はアクション a によって状態が s' に遷移した後，観測値 o が観測される観測値出力確率 $P(o'|s',a)$ の集合である．$r(s,a) \in R$ は, 状態 s でアクション a を行った時の期待報酬を表す．それぞれの変数の関係をダイナミックベイジアンネットワークで表すと Fig. 1 のようになる。

POMDP は以下のように動作する．まず，システムがアクション a を実行すると，状態 s' へ遷移する．この時，s' は s と a のみに依存して決まる．その状態変化により，観測値 o' が観測される．そして，アクション a と状態 s で規定されている報酬 r を得る。これを対話が終了するまで繰り返す。POMDP では状態を直接観測出来ないので，確率分布として扱う．

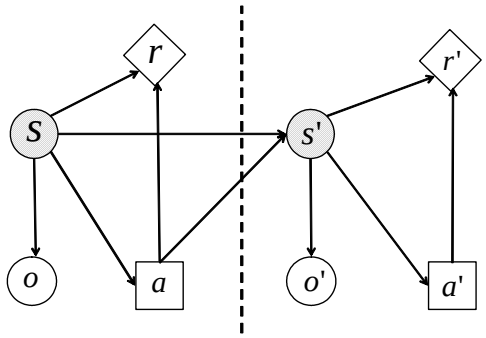

Fig. 1　DBN を用いた POMDP の影響図

その分布を $b(s)$ とする．この分布 $b(s)$ が分かっている時，アクションによって次の時刻の分布 $b'(s')$ は次式で表される．

$$b'(s') = k \cdot P(o' \mid s', a) \sum_{s \in S} P(s' \mid s, a) b(s) \qquad (1)$$

ここで k は $b'(s')$ の総和を 1 にするための正規化係数である．これを用いるとシステムが時刻 t までに得る割引報酬は次式で表される．ここで Υ（< 1）は割引率を表す．割引率は将来の報酬が現在においてどれだけの価値があるかを決定する．

$$V_t = \sum_{\tau=1}^{t} \Upsilon^{(\tau-1)} \sum_{s} b_\tau(s) r(s, a_\tau) \qquad (2)$$

POMDP の学習では，(2) 式を最大にするような方策を強化学習により求める．方策は，将来獲得できる報酬を最大にするアクション a を時間に独立に b のみから選択出来る．

3　POMDP を用いたカーナビにおける音声対話

カーナビでの音声操作において，入力における誤り（誤認識）が問題となることは周知である．この要因は，車室内雑音、発話変形、発話誤りの多さ等に起因する．その結果として，誤動作が生じ，自動車内で音声認識を利用するユーザが増えない原因となっている．

そこで本稿では，対話制御部に POMDP を導入した．システムはユーザのゴールを確率分布により求

*Spoken dialogue on car navigation system using partially observable Markov decision processes, by Yasuhide Kishimoto, Tetsuya Takiguchi, Yasuo Ariki(Kobe University)

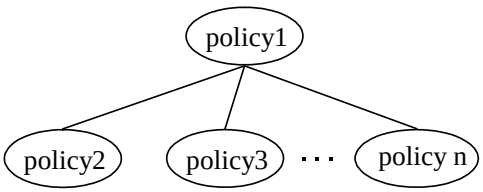

Fig. 2　方策の組み合わせ

めているので，たとえ誤認識を起こした場合でも，回復が可能である．

また，従来の音声認識システムでは，あらかじめ定められた定型句からなる音声コマンドしか受理できず，ユーザが前もって音声コマンドを正確に記憶していることを前提としており，音声入力に不慣れなユーザが直観的に音声だけでカーナビを操作することは困難である．そこで，POMDP において様々な発話を想定したユーザモデルを作成することにより、ユーザの多様な発話にも頑健に対応することが出来，直観的に使うことが可能となる．

本稿では，エアコン，オーディオ，近隣の店舗検索を扱うタスクを設定した．対話の制御において，POMDP の状態数の増加に伴い，強化学習の計算量が増大し、大きなタスクを扱うことが困難であったが，今回 Fig. 2 のように複数の方策を組み合わせることにより，大きなタスクを扱うことが可能となった．この時，policy 1 はユーザの求めているゴールが属しているタスクの分類を行い，その下の方策でエアコンの調整など具体的なユーザのゴールを求める．

システムのアクションはユーザの発話について，聞き直す，確認を取る，ユーザのゴールを決定するの 3 種類がある．この時の報酬の設定は，ユーザのゴールを正しく求めることが出来れば大きな正の報酬を与え，誤ったゴールを求めた場合には大きな負の報酬を与える．対話が早く収束するために，聞き直す，確認を取るというアクションを取った場合にも小さな負の報酬を与える．

4　評価実験

従来のカーナビと POMDP との比較実験を行うために，カーナビの音声認識システムを想定した方策（HC）を作成した．この方策は，ユーザの発話がシステムのアクションに対して適切でない場合にはもう 1 度聞き直すというアクションを取るが，それ以外の時にはユーザの発話に対して確認を取らずにユーザのゴールを決定する．

この方策と POMDP の方策を用い，ユーザシミュレーションによる試行を 1000 回繰り返し、その平均をとって，1 回の試行で得られる平均的な報酬とした．

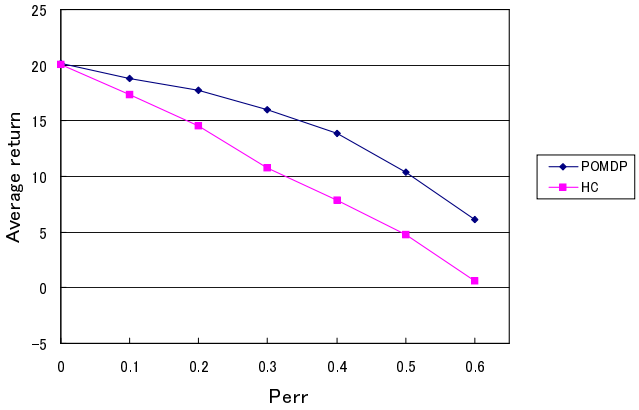

Fig. 3　実験結果

Fig. 3 にコンセプト誤り率 Perr（100∗{1-(正解発話数)/(総発話数)}）を 0.0 から 0.6 まで変えながらシミュレーション実験を行った結果を示す．結果から，コンセプト誤りが無いときには等しい報酬を得ているが，コンセプト誤りが増加すると POMDP の方策の方がより多くの報酬を得ていることが分かる．これは音声認識誤りがない時には，同等の性能を持っているが，認識誤りが増えると POMDP の方が頑健に動作していることを示している．それは，POMDP が常にユーザのゴールについて複数の仮説を保持しており，誤認識が起こった場合でもユーザのゴールを求めることが出来るからである．

このことから，POMDP は音声認識誤りが多い環境において，従来のカーナビよりも頑健に動作すると考えられる．

5　おわりに

本稿では，カーナビの対話制御部に POMDP を導入して，シミュレーション実験を行った．その結果，POMDP の方がより多くの報酬を得られた．このことから，従来のカーナビの音声認識よりも POMDP を用いた対話制御部の方が誤認識に対して頑健であると考えられる．今後の課題として，タスク拡大を図るとともに，より誤認識に強く，直観的に使える音声認識システムの実現に取り組んでいく．

参考文献

[1] 西村雅史,“ 音声認識ビジネスの現状と将来展望 ”，IPSJ，2005-SLP-55-3，pp. 13-15(2005).

[2] 神沼充伸,“ 自動車用音声インターフェースへの期待 ”，IPSJ，2006-SLP-63-9-2，pp. 47-48(2006).

[3] Williams,J.D.，Young,S.J.,“ Partially observable Markov decision processes for spoken dialog systems ”，Computer Speech and Language 21 (2),231-422.2009