

POMDP を利用した思考喚起型対話の制御

南 泰浩[†] 澤木 美奈子[†] 東中 竜一郎[†] 堂坂 浩二[†]

[†]日本電信電話株式会社 NTT コミュニケーション科学基礎研究所

〒619-0237 京都府相楽郡精華町光台 2-4

E-mail: [†] {minami@cslab.kecl, minako@apollo3, rh@cslab.kecl, dohsaka@cslab.kecl}.ntt.co.jp

あらまし 我々は、会話エージェントが、状況に即した適切な働きかけを通して、人間の自発的思考を喚起し、人間の会話意欲を高める思考喚起型の会話システムの実現を目指している。本稿では、思考喚起型会話システムの一例であるクイズを対象とする対話システムにおいて、POMDP を利用し不確定な情報にも適切に対処する対話システムの制御手法を提案する。不確定な情報として、ここでは、人がユーザの表情と音声からユーザを感じる難易度を判定した結果を利用する。この制御手法では、実験データから強化学習により POMDP の方策を作成し、その方策を使ってヒントを適切にスキップすることにより、ユーザのヒントに対する心的状態（易、ニュートラル、難）を制御する。本稿では、提案手法の有効性をシミュレーションにより評価を行った。

キーワード 思考喚起型会話, POMDP, 音声対話システム

Controlling thought-evoking dialogue using POMDP

Yasuhiro Minami[†], Minako Sawaki[†], Ryuichiro Higashinaka[†], Kohji Dohsaka[†]

[†]NTT Communication Science Laboratories NTT Corporation

2-4 Hikaridai Seika-cho Soraku-gun, Kyoto, 619-0237 Japan

E-mail: [†] {minami@cslab.kecl, minako@apollo3.brl, rh@cslab.kecl, dohsaka@cslab.kecl}.ntt.co.jp

Abstract We are researching thought-evoking dialogue systems where conversation agents appropriately affect users and evoke their voluntary thoughts to motivate human communication. This paper proposes a thought-evoking quiz dialogue system using the Partially Observed Markov Decision Process (POMDP) that can treat such uncertain information as paralinguistic information. As uncertain information, we employ the user's level of difficulty in handling quiz hints. Another person detects this difficulty level by observing the user's facial and voice information. The system controls the user's difficulty levels (easy, neutral, and difficult) for the hints by skipping hints based on the POMDP policy that was learned by reinforcement training. This paper evaluates the proposed system in simulation experiments.

Keyword Thought-evoking dialogue, POMDP, Spoken dialogue system

1. はじめに

我々は、人と環境を仲介し、場に即した適切な働きかけを行うことで、人に知的な愉しみ、発想、安らぎをもたらす、ひいては、人どうしのコミュニケーションを活性化することを目的の一つとする環境知能プロジェクトの研究を行っている[1].

この一環として、広い話題領域の会話システムを自動的に構築し、システム側からの働きかけにより人間の自発的思考を高める思考喚起型会話の実現を目指している。この会話システムは、人間の要求に応じて情報案内をするといった会話の受け身的・道具的な側面

に着目するような対話ではなく、対話の積極的・社会的な側面に着目し、会話エージェントが、人間に対して場に即した適切な働きかけを行うことで、人間の自発的思考を喚起し、ひいては、コミュニケーションの場を活性化させる会話を狙いとしている。

この思考喚起型会話システムの一例として、我々はクイズ対話システムを作成した[2][3][4].

場を活性化させるという目的に目を向けると、クイズ対話のように完全に自動ではなく、人間がサポートする必要があるが、人と人との対話においてロボットが介在し、人と人とのコミュニケーションを活性化す

る研究も始まっている[5].

また、コミュニケーションの場をより活性化させるためには、パラ言語情報などにより人間の心的状態を観察することが重要であり、このような情報を用いたロボット上の音声対話システムが提案されている[6]. しかし、[6]では、パラ言語情報が不確定な情報であるにも関わらず、不確定な情報を使って対話を制御することについては十分に検討されていない.

最近、不確定な情報に対しても対話を制御するために Partially Observed Markov Decision Process (POMDP) を利用する手法が、提案されている[7][8]. POMDP は、隠れ状態を持つ統計モデルを定義し、その統計モデルをもとに方策と呼ばれる最適な行動選択規範を強化学習により獲得するものである. この手法はもともと AI やその延長線上にあるロボットの制御などの分野で研究されてきた[9]. しかし、対話などの制御では、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、今まであまり使われてこなかった. しかし、最近、高速な POMDP の強化学習手法[10][11] が提案され、対話処理に POMDP が利用し始められている.

しかし、今までの POMDP を用いた対話処理研究[7][8]では、音声認識などの不確定な情報に対する対話の制御に POMDP を利用してはいたが、言語以外のパラ言語情報によって対話を制御することは行われていなかった. 本報告では、この POMDP を使い、ユーザのヒントに対する難易度の認識という不確定な情報によって、ユーザの心的状態を制御することを試みる.

2. クイズ対話システム

本章では、クイズ対話システムについて説明する. 本システムはインターネット上にある任意の人物に関するクイズを出題することができ、クイズを通して会話意欲を向上できるという特徴を持つ[2][3][4].

2.1.対話例

この対話システムで実際に行われる典型的な対話例を以下に示す.

ふーふー: なぞなぞスタート

最初のヒント: 野球解説者だよ.

ユーザ: 長嶋茂雄?

ふーふー: 違う 日本の野球選手というのはあつてるよ.

第二のヒント: 妻は5歳年上の姉さん女房だよ.

ユーザ: 落合博満?

ふーふー: おいしい, 中日の選手というのはあつてるよ.

第三のヒント: 甲子園決勝では柳井高に敗れたよ.

ユーザ: 板東英二!

ふーふー: あたり! すごい! よく分かったね!

2.2.クイズ対話システム構成

図1にクイズ対話システムの構成を示す. システムは、主に、ヒント自動生成、ヒント文ランキング、超大語彙音声認識、ネットワークマイニング、ネットワーク可視化、音声合成の6個のモジュールから構成されている(詳しい内容については文献[4]を参照). これらの他に全体を統合するための対話制御モジュールも加えた.

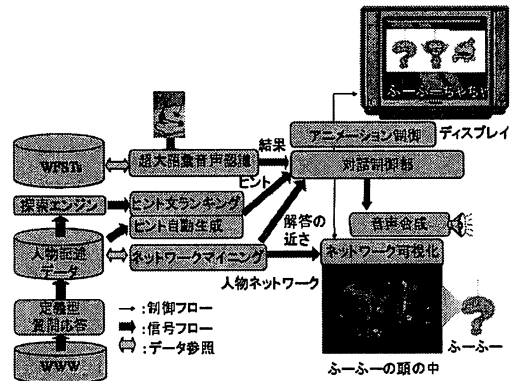


図1 クイズ対話システムの構成図

このシステム次のように動作する. まず、クイズが対象とする3万人の人名をあらかじめ、Wikipedia から選んだ. クイズの作成では、定義型質問応答技術により対象とする全ての人物の説明文を、インターネット上の大量の文書集合から集める. 次に、ヒント文ランキングでは、集めた説明文がどれくらいよく知られているか、またその人を特定するのにどれくらい役立つのか、などの観点からその説明文に点数をつける. 初めの方に簡単なヒントを出さないように、この点数によってヒントを順序づける.

ヒントに対する人の発声は大語彙連続音声認識システム SOLON によって認識される[12].

次に、認識結果がどれくらい正解に近いのかを調べるために、ネットワークマイニング技術を使って人名間の距離を計算し、二次元平面に射影することによって、任意の二つの人名間の距離を自動的に判定する[13][14].

ヒントや解答が正解に近いか遠いかなどのシステム側の情報は子供をイメージした音声合成[15]により提示される.

3. POMDP の基本概念

本章では、POMDP について簡単に説明する.

POMDP は $(S, O, A, T, Z, R, \gamma, b_0)$ で定義される。ここで、 $s \in S$ は状態を表す (システムとユーザの状態)。 $o \in O$ は、ユーザや環境から観測される観測値を表す (たとえば音声認識の結果など)。 $a \in A$ 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。 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 を実施した時の期待される報酬を表す。

ここで、 γ, b_0 の説明をする前に、POMDP のサイクルを図 2 を使って説明する。制御部からあるアクション a を実行すると、内部状態 s が変化し、その状態変化により、観測値 o が観測される。この時、そのアクション a と遷移状態 s で規定される報酬 r を得る。この動作をシステムが終了するまで繰り返す。

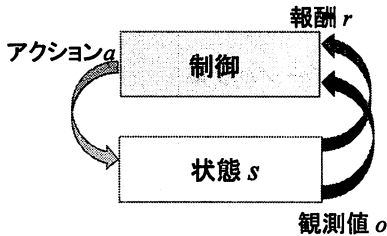


図 2 POMDP のサイクル

それぞれの変数の関係を、DBN の表記で書くと図 3 のようになる。

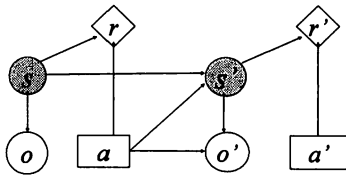


図 3 DBN を用いた POMDP の動作

POMDP では、HMM と同様に状態は、直接観測できないので、分布として扱う。その分布を $b(s)$ とする。

いま、状態 s の分布 $b(s)$ がわかっているとす。図 3 を用いると、アクションによって次の時刻の分布 $b(s')$ は以下のように求められる。

$$b'(s') = \eta P(o'|s', a) \sum P(s'|a, s) b(s). \quad (1)$$

ここで、 η は $b'(s')$ の総和を 1 とするための正規化係数である。(1)は、初期値が b_0 と設定されているので、再帰的に順次計算できる。この状態の分布を使うと、システムが時刻 t で得る割引報酬は以下のように表せる。ここで、 γ は割引率を表す。

$$V_t = \sum_{i=1}^t \gamma^{(t-i)} \sum_s b_t(s) r(s, a_t). \quad (2)$$

POMDP の学習では無限時間での報酬、すなわち、 V_∞ を最大にするように、アクション a を選択する方策を決定する。この方策は、強化学習により決定できる。方策は、将来獲得できる報酬を最大化できるアクション a を、時間に独立に b という分布だけから選択できる。

4. POMDP を用いたクイズ対話

2 章で述べたクイズ対話システムに、音声入力のために、ユーザのヒントに対する難易度という心的状態を観測する難易度認識部を導入する。それに伴いこの情報を扱うための制御モジュールに POMDP を導入した。この報告では、POMDP の制御に焦点を当てるため、実際に難易度認識部を作るのではなく、この部分には人がユーザの表情や音声から、ユーザのヒントに対する難易度を判定する仮定の難易度認識部を利用した。

4.1.クイズ対話でのアクション

POMDP で対象とする対話を制御するためのアクションには、きわめてシンプルなものを用いた。以下のリストは正解人名が「三船敏郎」の場合のヒントである。これらのヒントはユーザが簡単に答えに到達しないように、コンピュータで自動的に難しい順番に並べられている[4]。このヒントに対し、一つ以上のヒントをスキップして、より容易なヒントを提示するアクションを考える。ユーザにとってそのヒントが難しいとすると、ヒントをスキップすれば、より簡単な問題を提示することができ、ユーザのヒントに対する難易の心的状態を制御することができる。

[0] 日本の俳優だよ。

[1] 1961 年に初の海外作品「価値ある男」でメキシコ人役で主演したよ。

[2] 遺作は 1995 年の「深い河」だったよ。

[3] 東京の東宝撮影所撮影部にいる先輩である大山年治を訪ね、カメラマン助手の仕事に依頼したよ。

[4] 熊本の特攻隊基地で終戦を迎えたよ。

[5] 1961 年「用心棒」でヴェネチア映画祭主演男優賞、シネマヌーヴア金額賞だよ。

- [6] 1947年に黒澤監督の盟友・谷口千吉監督の映画「銀嶺の果て」でデビューしたよ。
- [7] 海外での受賞やノミネート歴は、1955年「七人の侍」で英国アカデミー賞主演男優賞ノミネートだよ。

4.2. POMDP を使った対話制御部の構成

図4に対話制御部の構成を示す。あらかじめ、状態遷移確率、観測値出力確率、報酬を設定しておく。その設定された確率、報酬から POMDP の強化学習により、方策を作成する。この方策は状態の分布 b から、適切なヒントを出題するアクションを決定する。対話時には、ユーザからの音声認識と難易度認識の結果から、次の状態の分布を決定する。次に、方策を用いて、この状態分布に最適なアクション（ヒントの番号）をヒントデータベースから選択する。ここで計算された状態の分布は次の状態の分布を計算するために、保存しておく。

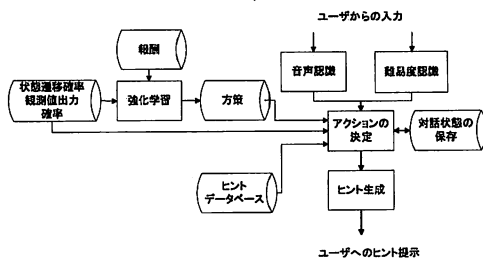


図4 対話制御部の構成

4.3. 状態と観測値の設定

POMDP では、システムとユーザの内部状態をどのように決定するかが極めて重要である。本報告では、状態 s を以下のようにシステムの状態を表す $s_{question}$ とユーザの状態を表す s_{knows} , s_{easy} との3種類に分解した。

$$s = (s_{question}, s_{knows}, s_{easy}) \quad (3)$$

これらの状態はそれぞれ以下のように設定した。

$s_{question}$: 出題しているヒントが何番目かを示す状態。

今回の設定ではヒントの数を8とし、0を初期状態として、0~8までの状態を作成した。

s_{knows} : ユーザが問題の答えを知っているか知らないかを表す状態。 $s_{knows} = \{\text{知っている}, \text{知らない}, \text{正解}\}$ の3値をとる。“正解”という状態は、対話が終了する状態を明確に記述する必要があるので設定した。

s_{easy} : ユーザのヒントに対する難易度を表す状態。問題の難しさによって、 $s_{easy} = \{\text{難}, \text{ニュートラル}, \text{易}\}$ となる。

また、観測値を、 $o = (o_{recog}, o_{easy})$ のように構成する。

ここで、各観測値は以下のように設定した

o_{recog} : 音声認識結果を分類し、 $o_{recog} = \{\text{正解}, \text{不正解}, \text{認識誤り}, \text{次のヒント}\}$ の4種類とした。

o_{easy} : 心的状態の観測値、問題の難しさによって、 $o_{easy} = \{\text{難}, \text{ニュートラル}, \text{易}\}$ の3種類を設定した。

すべての状態遷移確率、観測値出力確率を計算するのは変数が膨大になり難しい。そこで、以下の近似を用いた。

$$\begin{aligned} P(s' | s, a) &= P(s'_{question}, s'_{easy}, s'_{knows} | s_{question}, s_{easy}, s_{knows}, a) \\ &\approx P(s'_{question} | s_{question}, a) P(s'_{easy} | s'_{question}, s_{easy}, a) \quad (4) \\ P(s'_{knows} | s'_{question}, s_{knows}, a) & \end{aligned}$$

$$\begin{aligned} P(o' | s', a) &= P(o'_{easy}, o'_{recog} | s'_{question}, s'_{easy}, s'_{knows}, a) \\ &\approx P(o'_{easy} | s'_{easy}, a) P(o'_{recog} | s'_{knows}, a). \quad (5) \end{aligned}$$

5. シミュレーションの設定

4章で述べた制御部をシミュレーションにより評価した。シミュレーションに用いた確率はあらかじめ行った対話実験から得られたデータにより求めた。

5.1. 実験によるデータの収集

話者24人を使って、図1に示すシステムを使って、各種の統計情報を収録した。15分のクイズ対話セッションを一人8セッション行い、ヒント順毎の正答率を求めた。ヒントに対するユーザの難易度認識のデータはこのうちの一人の話者のみから作成した。これらのデータを用いて、 $P(s'_{question} | s_{question}, a)$, $P(s'_{easy} | s'_{question}, s_{easy}, a)$, $P(s'_{knows} | s'_{question}, s_{knows}, a)$, $P(o'_{easy} | s'_{easy}, a)$, $P(o'_{recog} | s'_{knows}, a)$ を決定した。システムやユーザの状態は実際に観測できないため、これらの確率の真の値を求めることは困難である。今回の報告では、以下の手順に従って、これらの値の設定を行った。

5.2. $P(s'_{question} | s_{question}, a)$ の設定

$s'_{question}$ はヒントの番号を表し、 a も質問するヒントの番号を表す。アクションが実行できるのは、遷移先の状態番号と a の番号が一致するときとした。さらに、ヒントのスキップは2つまで、かつ、ヒントの順番が後戻りしないという条件を加えた。この条件は以下のように表される。

$$P(s'_{question} | s_{question}, a) = 1 \quad \text{if } s'_{question} = a \ \& \ 1 \leq s'_{question} - s_{question} \leq 3$$

$$P(s'_{question} | s_{question}, a) = 0 \quad \text{otherwise.} \quad (6)$$

5.3. $P(s'_{easy} | s'_{question}, s_{easy}, a)$ の設定

ヒントの難易度に関する状態の遷移は、出題されたヒントの順番と、一つ前の難易度の状態のみに依存す

ると仮定した。この確率も $s'_{question}=a$ のときのみ存在する。この値は実験データから得ることができないので、徐々にヒントが簡単になるように著者の主観で設定した。 a のアクションが 1~8 までの各値での[難, ニュートラル, 易]から[難', ニュートラル', 易']への遷移確率行列を以下のように設定した。最初の行列の左上が, $P(\text{難}' | s'_{question}=1, \text{難}, a=1)$ の確率である。

$$\begin{aligned}
 a=1: & \begin{bmatrix} 0.8 & 0.2 & 0.0 \\ 0.0 & 1.0 & 0.0 \\ 0.0 & 0.1 & 0.9 \end{bmatrix} & a=2: & \begin{bmatrix} 0.7 & 0.3 & 0.0 \\ 0.0 & 0.8 & 0.2 \\ 0.0 & 0.1 & 0.9 \end{bmatrix} & a=3: & \begin{bmatrix} 0.6 & 0.3 & 0.1 \\ 0.0 & 0.7 & 0.3 \\ 0.0 & 0.1 & 0.9 \end{bmatrix} \\
 a=4: & \begin{bmatrix} 0.5 & 0.4 & 0.0 \\ 0.0 & 0.6 & 0.4 \\ 0.0 & 0.1 & 0.9 \end{bmatrix} & a=5: & \begin{bmatrix} 0.4 & 0.5 & 0.1 \\ 0.0 & 0.5 & 0.5 \\ 0.0 & 0.0 & 1.0 \end{bmatrix} & a=6: & \begin{bmatrix} 0.3 & 0.6 & 0.1 \\ 0.0 & 0.4 & 0.6 \\ 0.0 & 0.0 & 1.0 \end{bmatrix} \\
 a=7: & \begin{bmatrix} 0.2 & 0.7 & 0.1 \\ 0.0 & 0.3 & 0.7 \\ 0.0 & 0.0 & 1.0 \end{bmatrix} & a=8: & \begin{bmatrix} 0.1 & 0.8 & 0.1 \\ 0.0 & 0.2 & 0.8 \\ 0.0 & 0.0 & 1.0 \end{bmatrix}
 \end{aligned}$$

5.4. $P(s'_{knows} | s'_{question}, s_{knows}, a)$ の設定

s'_{knows} はユーザが問題を“知っている”, “知らない”, “正解”の3つの状態を仮定している。“正解”の状態では, ユーザは 100%の確率で正解を発声し, 認識誤りもしないと仮定している。[知っている, 知らない, 正解]から[知っている', 知らない', 正解']への遷移確率を以下に示す。 A を 1~8 まで変化させた時の値である。

$$\begin{aligned}
 a=1: & \begin{bmatrix} 0.35 & 0.0 & 0.65 \\ 0.07 & 0.76 & 0.17 \\ 0.0 & 0.0 & 0.0 \end{bmatrix} & a=2: & \begin{bmatrix} 0.35 & 0.0 & 0.65 \\ 0.06 & 0.79 & 0.15 \\ 0.0 & 0.0 & 0.0 \end{bmatrix} & a=3: & \begin{bmatrix} 0.35 & 0.0 & 0.65 \\ 0.08 & 0.77 & 0.15 \\ 0.0 & 0.0 & 0.0 \end{bmatrix} \\
 a=4: & \begin{bmatrix} 0.35 & 0.0 & 0.65 \\ 0.09 & 0.7 & 0.21 \\ 0.0 & 0.0 & 0.0 \end{bmatrix} & a=5: & \begin{bmatrix} 0.35 & 0.0 & 0.65 \\ 0.05 & 0.81 & 0.19 \\ 0.0 & 0.0 & 0.0 \end{bmatrix} & a=6: & \begin{bmatrix} 0.35 & 0.0 & 0.65 \\ 0.06 & 0.8 & 0.14 \\ 0.0 & 0.0 & 0.0 \end{bmatrix} \\
 a=7: & \begin{bmatrix} 0.35 & 0.0 & 0.65 \\ 0.08 & 0.72 & 0.20 \\ 0.0 & 0.0 & 0.0 \end{bmatrix} & a=8: & \begin{bmatrix} 0.35 & 0.0 & 0.65 \\ 0.05 & 0.84 & 0.11 \\ 0.0 & 0.0 & 0.0 \end{bmatrix}
 \end{aligned}$$

これらの状態の遷移確率も実際には実験から直接観測できないが, 以下の手順により設定した。最初の行には, 文献[16]で得られた実験結果の音声認識性能を反映させた。すなわち, “知っている”状態から正解への遷移に音声認識性能を設定し, 知っていてかつ答えが正解のときは, “正解”という状態に遷移するようにした。これに伴い, “知っている”の状態では必ず不正解の観測値を出力するようにし, 音声認識率を観測値の出力確率ではなく状態遷移に反映させた。2行目は, “知らない”から“知らない”, “知っている”“正解”への遷移を表している。この部分には実験で得られた各ヒントに対するユーザの回答の正解率を設定した。最初の行列($a=1$)の2行目は[0.07, 0.76, 0.17]とな

っている。これは, 7%のクイズで, 最初のヒントを誤認識し, 17%で正解を返し, 76%が“知らない”状態にとどまることを表している。

5.5. $P(o'_{easy} | s'_{easy}, a) P(o'_{recog} | s'_{knows}, a)$ の設定

$P(o'_{easy} | s'_{easy}, a)$ に関しては, 一人のユーザを別の人間がそのユーザがヒントをどう思っているかを観察し, “難”, “ニュートラル”, “易”の3つに分類した。これを o'_{easy} とした。また, このときクイズをしたユーザが, 実験の映像を見て, 同じクラス分けした結果を s'_{easy} とした。この結果から観測値の出力確率を求めた。

$P(o'_{recog} | s'_{knows}, a)$ は音声認識に関する出力確率を設定した。ここで, $o'_{recog} = \{\text{正解}, \text{認識誤り}, \text{不正解}, \text{次のヒント}\}$ である。これらは, 音声認識に得られた結果を, 分類したものであり。実際の値は, 文献[16]の実験結果から以下のように設定した。

$$\begin{bmatrix} 0.0 & 1.0 & 0.0 & 0.0 \\ 0.0 & 0.1 & 0.2 & 0.7 \\ 1.0 & 0.0 & 0.0 & 0.0 \end{bmatrix}. \quad (7)$$

5.3 で述べたように, 音声認識性能を, 状態遷移確率に反映させたので, 1行目のように“知っている”ときは必ず誤認識をように設定した。2行目は, ユーザの状態が“知らない”という時には, 70%は次のヒントと発声し, 20%は不正解の解答(認識誤りは含まない)を発声し, 10%は認識誤りとなる発声をするということを意味している。

6. シミュレーションによる実験

POMDP では, 報酬の設定により, 全体の系を様々な状態に統計的に導くことができる。それを示すために, シミュレーションによる実験を行った。今回の実験では, s_{easy} をできるだけニュートラルの状態にるように, 報酬を以下のように設定した。

$$r(s, a) = 5.0 \quad (\text{if } s_{easy} = \text{ニュートラル}). \quad (8)$$

この時の割引率 γ は 0.9 と設定した。

難易度の認識器を以下の3種類に変化させて実験を行った。

- (1) 等確率認識器 (各確率 1/3)

難易度認識器の性能をすべて等確率とした。具体的には, 全ての $P(o'_{easy} | s'_{easy}, a)$ を 1/3 に設定した。

- (2) 実験値を使った認識器

人が見てユーザのヒントへの難易度を認識する手法を難易度認識器とした。この $P(o'_{easy} | s'_{easy}, a)$ を計算するために一人のユーザの実データを使用した。

- (3) 理想認識器

理想的な認識器での性能を実現した。すなわち,

$$P(o'_{easy} | s'_{easy}, a) = 1 \quad \text{if } o'_{easy} = s'_{easy}$$

$$P(o'_{easy} | s'_{easy}, a) = 0 \quad \text{otherwise.} \quad (9)$$

この条件でシミュレーションを行った。シミュレーションでは、5章で述べた確率を設定し、その確率だけから評価を行った。実データは使っていない。評価では、式(2)で得られる報酬を計算した。報酬の計算では10,000回の試行を5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。表1にその結果を示す。

表1 各種難易度認識器に対する平均報酬

	等確率認識器	実験値認識器	理想認識器
平均報酬	3.91	3.99	4.14

この結果から、理想的な難易度の認識器を用いると、等確率で答えを返す認識器に比べて、より多くの報酬を獲得できている。しかし、その差はそれほど大きくない。これは、今回のタスクでは、選択できるアクションが、ヒントのスキップと限られており、アクションの選択によるユーザの心的状態の変化がそれほど大きく変わらないことによると考えられる。

実際のデータから求めた、人による難易度認識器を使うと、等確率の難易度認識器より、高報酬を獲得できることがわかる。このことから、難易度の認識器が不確定であっても POMDP という枠組みにより、ユーザのヒントに対する難易度という心的状態を制御できる可能性があることが確認された。

7. まとめ

不確定な情報に基づいて対話処理を行うための手法として POMDP を利用するクイズ対話シミュレーションシステムを作成した。このシステムでは、人の心的状態を認識する装置として、人がユーザの表情を見て、ユーザのヒントへの難易度を観測する難易度認識器を用いた。実験により得られたデータに基づき POMDP を利用するシミュレーションシステムをパラメータを設定し、難易度認識器の不確定な結果を用いて、ユーザの心的状態をどの程度制御できるかの評価をおこなった。シミュレーションの結果、平均報酬という評価において、ユーザの難易度の心的状態を制御できることがわかった。

謝 辞

本研究の評価データは、京都大学大学院情報学研究科美濃研との共同研究下で得られた。

文 献

[1] 前田, 南, 堂坂, “妖精・妖怪の復権—新しい「環境知能」像の提案—”, 情報処理, vol. 47, no. 6, pp. 624-640, 2006.

[2] R. Higashinaka, K. Dohsaka, and H. Isozaki, “Learning to Rank Definitions to Generate Quizzes for Interactive Information Presentation,” In Proc. of the 45th Annual Meeting on Association for Computational Linguistics (Poster Presentation), pp. 117-120, 2007.

[3] R. Higashinaka, K. Dohsaka, S. Amano, and H. Isozaki, “Effects of Quiz-style Information Presentation on User Understanding,” In: Proc. of Interspeech 2007, pp. 2725-2728, 2007.

[4] Y. Minami, M. Sawaki, K. Dohsaka, R. Higashinaka, K. Ishizuka, H. Isozaki, T. Matsubayashi, M. Miyoshi, A. Nakamura, T. Oba, H. Sawada, T. Yamada, and E. Maeda, “The World of Mushrooms: Human-Computer Interaction Prototype Systems for Ambient Intelligence,” In Proc. of ACM Int. Conf. Multimodal Interfaces (ICMI) 2007, pp. 366-373, 2007.

[5] 松山洋一, 谷山輝, 藤江真也, 小林哲則, “人-人コミュニケーションの活性化支援ロボットの開発”, 人工知能学会研究会資料, SIG-SLUD-A801, pp. 15-22, 2008.

[6] 藤江真也, 江尻 康, 菊池英明, 小林哲則, “肯定的/否定的発話態度の認識とその音声対話システムへの応用,” 電子情報通信学会 論文誌 D-II, Vol. J88-DII, No.3, pp. 489-488, 2005.

[7] J. Williams, P. Poupart, S. Young, “Partially Observable Markov Decision Processes with Continuous Observations for Dialogue Management”, In Proc. of Sigdial workshop 2005.

[8] J. Williams, “Applying POMDPs to dialog systems in the troubleshooting domain,” In Proc. of NAACL-HTL workshop on Bridging the Gap, 2007.

[9] S. Thrun, W. Burgard, “Probabilistic Robotics,” The MIT Press, Cambridge, pp. 513-566, 2006.

[10] T. Smith and R. G. Simmons, “Focused Real-Time Dynamic Programming for MDPs: Squeezing More Out of a Heuristic,” In Proc. of Artificial Intelligence (AAAI), 2006.

[11] P. Poupart, “Exploiting Structure to Efficiently Solve Large Scale Partially Observable Markov Decision Processes,” Ph.D. thesis, Department of Computer Science, University of Toronto, 2005.

[12] T. Hori, C. Hori, and Y. Minami, “Fast on-the-fly Composition for Weighted Finite-State Transducers in 1.8 Million-Word Vocabulary Continuous Speech Recognition,” In Proc. of Interspeech 2004, pp. 289-292, 2004.

[13] 斉藤和巳, 山田武士, 風間一洋, “k-dense 法によるネットワークのコア部抽出,” 第2回 ネットワーク生態学シンポジウム, 2006.

[14] T. Matsubayashi, and T. Yamada, “A Force-directed Graph Drawing based on the Hierarchical Individual Timestep Method,” In Proc. of IJECIS, pp. 116-121, 2007.

[15] 水野秀之, 磯貝光昭, 長谷部未来, 浅野久子, 阿部匡伸, “コーパスベースアプローチによるテキストからの音声合成,” NTT 技術ジャーナル, vol.16, no.1, pp. 23-26, (2004).

[16] 南泰浩, 東中竜一郎, 澤木美奈子, 堂坂浩二, 山田武士, 松林達史, 磯崎秀樹, 前田英作, “クイズ対話システムの構築と音声認識性能による評価,” 日本音響学会秋季研究発表会, 2-3-11, pp. 82-83, 2007.