

ユーザの焦点を用いた POMDP による 音声情報案内システム

吉野 幸一郎¹ 河原 達也¹

概要: 雑談形式で情報案内を行う音声対話システムについて述べる。このシステムは Web 上のニュース記事を情報源とし、ユーザはシステムとの対話を通じてその日のニュースを知ることができる。このシステムのふるまいを、統計的制御手法である POMDP によって制御する。通常 POMDP による対話制御は対話のゴールやフローを想定するが、情報案内ではこうしたゴールやフローは定義されない。そこで本研究では、POMDP による制御を対話モジュールの選択問題として捉え、入力として音声言語理解の結果であるユーザ状態とユーザの焦点を用いる。音声言語理解は識別モデルによって行われ、信念状態の更新と強化学習により得られた政策関数によって、応答する対話モジュールを選択する。POMDP における報酬関数は、基本的な対話制御の適切性と情報案内が成功したかどうかの 2 種類により定義する。被験者との対話実験において、ユーザの焦点を用いた提案手法の有効性が確かめられた。

1. はじめに

近年、スマートフォンアプリケーションに代表される多くの音声対話システムが実世界で展開され、多様なユーザに利用されている。しかし、こうしたシステムの多くは天気案内 [1] や乗換案内 [2], [3] に代表されるあらかじめ作りこまれたタスクで、タスクゴールや必要なスロットを作りこむことによって成り立っている。こうしたシステムでは、ユーザはシステムに定義されたタスクゴールと、場所・日時などのタスク達成に必要なスロットといったシステムの想定に沿った発話を行うことが求められる。また、近年のシステムでは質問応答機能が併用されているが、こうした機能は「甲子園はいつできたんですか?」や「甲子園の銀傘の大きさは?」などのファクトイド型の質問や、「あなたの名前は?」などのあらかじめ定義された質問に対してのみ機能する。システムの想定外の質問がされた場合、システムは「その質問には答えられません」と答えるか、Web での検索結果を表示する。しかしこうしたふるまいは不自然であり、またそもそも唯一の回答が得られない情報要求も多いと考えられる。

本研究では、特定のタスクゴールを想定しない情報案内システムを実現する。システムはスポーツや旅行などのドメインを想定し、ユーザの質問意図が明確でない場合ドメイン知識や対話の履歴を利用して柔軟な応答を行う。こ

うした情報案内タスクは先行研究でいくつか検討されているが [4], [5], [6]、本研究ではニュース記事を情報源として情報案内を行うシステムについて述べる。

こうしたシステムに対する対話制御は、人手で作られた単純なルールによることが一般的であった [7], [8], [9] が、構築方法が確立されているわけではない。WikiTalk [6], [10] は Wikipedia の内容について話すシステムで、あらかじめ構築されたシナリオに沿って対話を行う。また、機械翻訳の手法を利用してマイクロブログの対話データなどから発話を生成する枠組み [11] も提案されているが、ユーザの意図や対話状態を利用してユーザが望む情報を提供できていないわけではない。これに対し本研究では、ユーザの意図と焦点を利用して対話制御を行うことでユーザが潜在的に求める情報を提供することを目指す。

近年、強化学習を用いた対話制御が盛んに行われている。マルコフ決定過程 (MDP) や部分観測マルコフ決定過程 (POMDP) を用いた対話制御は、タスク達成対話において成功を収めている [12], [13], [14], [15], [16]。しかし、強化学習は明確なタスクゴールとそれに到達することによって得られる報酬の定義が必要で、本研究で対象とする情報案内に適用することが難しい。

Pan ら [5] は文書集合の構造を用いて、目標となる情報を含む文書に到達することを報酬として定義している。こうした手法は構造化された文書集合に対しては有効であるが、どのような文書集合にも利用できるわけではない。Shibata ら [17] は雑談対話の最後にユーザが評価を行うこ

¹ 京都大学 学術情報メディアセンター
Academic Center for Computing and Media Studies, Kyoto
University, Sakyo, Kyoto 606-8501, Japan

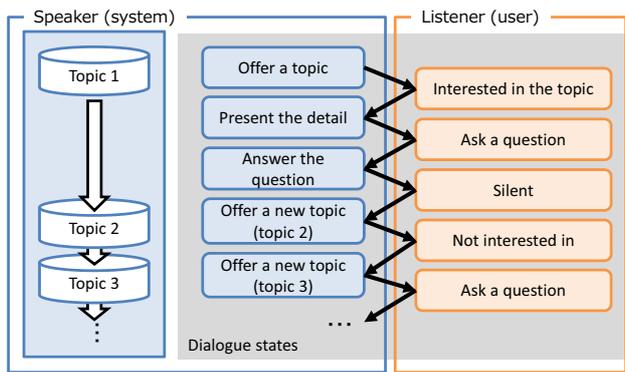


図 1 情報案内対話の例

とによって報酬を定義している。また、Meguroら [18] は聞き役対話における報酬を与える識別モデルを構築することで報酬を与えている。しかしこれらの手法は、ユーザの評価や識別モデルの学習データの構築コストがかかる。これに対し本研究では、システム側の対話制御の明白な適切さと、客観的に観測可能な情報案内の成功 [19] により報酬を定義し、POMDP の学習を行う。

2. 情報案内システム

自然な人対人の対話を考えた場合、参加者は何を話すか事前に考え、そのトピックを基に対話を進めていることが多い [20]。これを話し手と聞き手の役割に分けて表したものが図 1 である。話し手は聞き手の反応に応じて現在のトピックを継続したり、質問に答えたり、次のトピックを案内したりする。本研究ではこうした対話構造をモデル化し、システムが話し手の役割を持つことによって情報案内を実現する。

2.1 ニュース案内システムの概要

本研究では情報案内のタスクとして、ニュース案内を扱う。システムは Web 上から収集したニュース文書を、ユーザの興味に応じて案内する。ここで、トピックは 1 つのニュース文書として定義し、ドメインは野球や株価などそのニュース文書のカテゴリとして定義する。システムはドメイン文書集合から自動抽出した知識 [19] を用いて対話を行う。図 2 にシステムの枠組みを示す。システムは Topic Presentation (TP) モジュールによって対話を開始し、案内できるニュース (=トピック) を紹介する。このモジュールはニュースのヘッドラインを読み上げる。紹介したトピックにユーザが興味を示した場合、システムは Story Telling (ST) モジュールによってニュースのあらましを紹介する。このモジュールはニュース記事の冒頭にあるサマリを読み上げる。ユーザが詳細に関して質問を行った場合、システムは Question Answering (QA) モジュールによって応答を行う。ユーザの反応がなくなっても、ユーザがまだ興味がありそうな場合は、Proactive Presentation (PP) によ

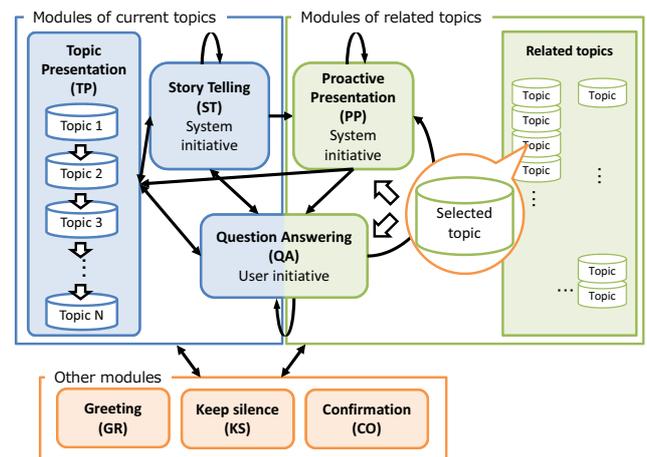


図 2 情報案内システムの概要

てシステム主導の情報推薦を行う。QA と PP のモジュールは、述語項構造を用いた情報案内システム [19] の機能を用いる。この他に Greeting (GR)、Keep Silence (KS)、Confirmation (CO) の各モジュールが実装され、図 2 で定義された遷移によって対話を行う。定義されるモジュールの種類は情報案内の種類に依存するが、ニュース案内タスクではこれらの 7 種類のモジュールを用いる。

2.2 情報案内におけるユーザの焦点

焦点は、「所与の時点における物体、状態、関係、発話意図などの情報への顕著な注意状態」として定義される [21]。ユーザ発話が焦点を含む場合、発話中のいずれかの対象に対して興味があると解釈できる。そこで本研究では、ユーザの焦点を「ユーザの興味の中心を示す特定の対象」として定義する。例えば、「イチローは活躍しましたか？」という質問では「イチローは」に焦点が当たっていると解釈する。本研究で対話制御に用いる POMDP では、ユーザ発話中にこの焦点が存在するかの情報を用いる。

3. 音声言語理解

本章では、ユーザの意図と焦点の解析について説明する。音声言語理解モジュールは、尤度付きの解析結果を出力する。ユーザの意図としては、以下の 6 通りを仮定する。

- TP: TP モジュールへの要求
- ST: ST モジュールへの要求
- QA: QA モジュールへの要求
- GR: GR モジュールへの要求
- II: 音声認識誤りに起因する無効入力
- NR: 一定時間の無音の検出 (本研究では 5 秒)

また、焦点については当該発話中に焦点が存在するかを判定する。この音声言語理解の例を図 3 に示す。これを識別モデルにより実現した結果 [22] は、ユーザの意図に対する識別精度が 92.2%、ユーザの焦点が発話中に存在するかどうかの識別精度が 99.9% であった。

```

SYS, tp, "巨人松井臨時コーチ、ランチ特打で弾丸5発。桃井球団社長に、「契約して」"
USR, st, 0, "詳しく教えてください",
SYS, st, "巨人の松井臨時コーチがランチ特打に参加し、古巣巨人軍の若手にお手本を見せたよ。"
USR, qa, 1, "松井は監督になるんですか?",
SYS, qa, "桃井球団社長は松井に、「契約して」"
USR, tp, 0, "次のニュース",
SYS, tp, "スーパーエースへ。藤浪晋太郎はルーキーイヤーから進化しているのか"
...
    
```

図 3 対話データの例。USR がユーザ発話、SYS がシステム発話を表し、それぞれユーザの意図と応答したモジュールがアノテーションされている。また、ユーザ発話中に焦点が存在するかどうかは 2 値でアノテーションされている。

4. 情報案内のための対話制御

POMDP による統計的対話制御では、以下のような変数を用いて定式化が行われる。ターン t に対する変数は、

- $s \in I_s$: ユーザの状態 (ユーザの意図及びユーザの焦点)
- $a \in K$: システムの行動
- $o \in I_o$: 観測結果
音声認識や言語理解の誤りを含むユーザの状態の観測結果。
- $b_{s_i} = P(s_i | o^{1:t})$: 信念
各ユーザの状態に対する確率変数。
- π : 政策関数
信念 b が与えられたときにシステムの行動 a を出力する関数。 π^* は学習の結果得られた最尤の政策関数。
- r : 報酬
ユーザの状態 s とシステムの行動 a の組に対して報酬を与える関数。

統計的対話制御の目的は、時系列 (1 から t) で与えられた観測結果の系列 $o^{1:t}$ に対して最尤の行動 \hat{a}^t を出力することにある。

次に、信念の更新について述べる。信念は観測関数と状態遷移関数を含む以下の更新式から求められる。

$$b_{s_j}^{t+1} \propto \underbrace{P(o^{t+1} | s_j^{t+1})}_{\text{Obs.1}} \sum_{s_i^t} \underbrace{P(s_j^{t+1} | s_i^t, \hat{a}_k^t)}_{\text{Trans.1}} b_{s_i^t}^t. \quad (1)$$

式 (1) の中で、**Obs.1** は音声認識と音声言語理解によって与えられる。**Trans.1** は状態遷移関数で、1 ターン前のユーザの状態 s_i^t とシステムの行動 \hat{a}_k^t が与えられたとき次のユーザの状態 s_j^{t+1} を出力する。この状態遷移関数は対話コーパスから最尤推定によって推定する。信念が更新されると、学習された政策 π^* を用いて次の行動

$$\hat{a} = \pi^*(b) \quad (2)$$

が出力される。

4.1 POMDP の学習

最尤の政策 π^* の学習には Q 学習 [23], [24] を用いる。Q

学習は、割引率 γ で割引かれた将来の報酬を最大化する関数で、次の式で学習される。

$$Q(b, a) \leftarrow (1 - \varepsilon)Q(b^t, a^t) + \varepsilon[R(s^t, a^t) + \gamma \max_{a^{t+1}} Q(b^{t+1}, a^{t+1})]. \quad (3)$$

ここで ε は学習率で、今回は 0.01 に設定した。また、割引率 γ は 0.9 に設定した。Q 学習の結果得られる政策は、

$$\pi^*(b) = \underset{a}{\operatorname{argmax}} Q(b, a) \quad (4)$$

となる。

しかし、取りうるすべての信念 b に対して Q 関数を学習することは不可能なため、Grid-based Value Iteration[25] によって限られた信念の点に対してのみ Q 関数を学習する。ここで信念の点は、

$$b_s = \begin{cases} \eta & \text{if } s = s_i \\ \frac{1-\eta}{|I_s|-1} & \text{if } s \neq s_i \end{cases} \quad (5)$$

によって与えられる。ここで η はユーザ状態 s が s_i である確率であり、シミュレーションの際には 0.0 から 1.0 まで 0.1 刻みの 11 点を用いる。それ以外の状態には一様な確率を与える。例えば、ユーザ状態の候補が 3 通り (s_1, s_2, s_3) あり、音声言語理解の出力候補である s_1 に対して η で 0.6 を与える場合、残りの 2 つに 0.2 を与える。

4.2 ユーザの焦点を用いた対話制御

提案する対話制御では、ユーザの状態をユーザの意図とユーザの焦点の 2 種類の情報で与える。ここで意図とは音声言語理解による 6 種類の出力 (o_s) であり、焦点はユーザ発話中に焦点が存在するかないかの 2 値 (o_f) である。信念の更新式は、これまでの焦点 f_l と次の焦点 f_m^{t+1} を導入して次のように与えられる。

$$b_{s_j}^{t+1, f_m^{t+1}} = \underbrace{P(o_s^{t+1}, o_f^{t+1} | s_j^{t+1}, f_m^{t+1})}_{\text{Obs.2}} \times \sum_i \sum_l \underbrace{P(s_j^{t+1}, f_m^{t+1} | s_i^t, f_l^t, \hat{a}_k^t)}_{\text{Trans.2}} b_{s_i^t, f_l^t}^t. \quad (6)$$

ここで、観測関数は以下のように近似できる。

$$\begin{aligned} \text{Obs.2} &= P(o_s | o_f, s_j, f_m) P(o_f | s_j, f_m) \\ &\approx P(o_s | s_j) P(o_f | f_m). \end{aligned} \quad (7)$$

この近似は、真のユーザの焦点 f_m から観測されたユーザの焦点 o_f の観測モデルと、真のユーザの意図 s_j から観測されたユーザの意図 o_s の観測モデルが互いに独立であるという仮定によって成り立つ。また、状態遷移確率は

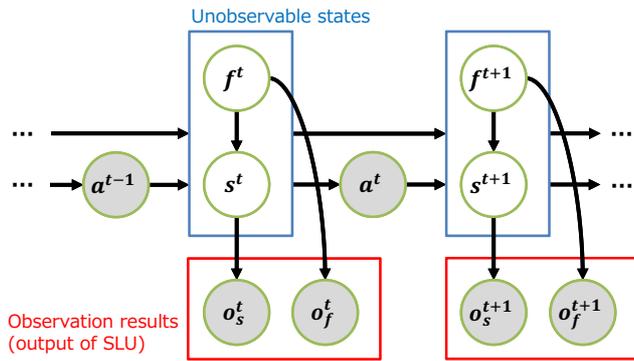


図 4 観測モデルと状態遷移モデルのグラフィカルモデル

$$\text{Trans.2} = \underbrace{P(s_j^{t+1}, |f_m^{t+1}, s_i^t, f_l^t, \hat{a}_k^t)}_{\text{state model}} \underbrace{P(f_m^{t+1} | s_i^t, f_l^t, \hat{a}_k^t)}_{\text{focus model}} \quad (8)$$

となり、ユーザの焦点 f が定まってからユーザの意図 s が定まるというように解釈できる。このグラフィカルモデルを図 4 に示す。

この結果、学習される政策は、

$$\hat{a} = \pi^*(b_{s,f}) \quad (9)$$

となる。この政策を獲得するための Q 学習は、

$$Q(b, a) \leftarrow (1 - \varepsilon)Q(b^t, a^t) + \varepsilon[R(s^t, f^t, a^t) + \gamma \max_{a^{t+1}} Q(b^{t+1}, a^{t+1})] \quad (10)$$

で定義される Q 関数で与えられる。

Q 関数の学習にはシミュレータを用いるが、これは式 (8) で与えられる

$$\text{focus model} = P(f^{t+1} | f^t, s^t, a^t), \quad (11)$$

$$\text{state model} = P(s^{t+1} | f^{t+1}, f^t, s^t, a^t) \quad (12)$$

をコーパスから最尤推定によって計算し、真のユーザの意図 s と真のユーザの焦点 f を出力する。観測結果 o_s と o_f は、任意の η を与えた式 (5) によってシミュレーションする。

4.3 報酬の定義

POMDP の学習のためには報酬が必要となる。そこで本研究では、ターンごとの基本的な対話制御の適切性に関する報酬である r_C と、情報案内の達成報酬である r_I を定義する。まず、表 1 に各ターンにおける対話制御の報酬 r_C を示す。この報酬は、正しい行動に対して +1、許容できる行動に対して 0、誤った行動に対して -1 の報酬を与えるものである。また、許容されるが正しい行動ではない行動 (0 の行動) をした場合 1、誤った行動 (-1 の行動) をした場合 2 のフラストレーションポイントが与えられ、フラスト

表 1 各ターンで与えられる対話制御の報酬 r_C

state s	focus f	action a						
		TP	ST	QA	PP	GR	KS	CO
TP	0	+1	-1	-1	-1	-1	-1	0
	1							
ST	0	-1	+1	-1	0	-1	-1	0
	1							
QA	0	-1	+1	1	-1	-1	-1	0
	1		-1		+1			
GR	0	-1	-1	-1	-1	+1	-1	0
	1							
NR	0	+1	-1	-1	-1	-1	0	0
	1	-1			+1			
II	0	-1	-1	-1	-1	-1	+1	0
	1							

レーションポイントが 10 に達すると強制的に対話が終了され-20 の報酬が与えられる。逆に、20 ターン以上対話を継続できた場合は+20 の報酬が与えられる。

次に、情報案内の報酬 r_I について述べる。ユーザの要求に従って QA または PP のモジュールが呼ばれた場合に、正しい応答をすると +1、曖昧な応答をすると 0、誤った応答をすると -1 の報酬が与えられる。先行研究 [19] で、質問応答 (QA) に関しては 64.1% が正解、6.5% が曖昧、28.9% が誤りであることがわかっているため、シミュレーションによる学習時はこの割合に従って報酬をそれぞれ、+1、0、-1 で与える。また、情報推薦 (PP) に関しても 63.6% が高いユーザ満足度、30.9% が中程度のユーザ満足度、5.5% が低いユーザ満足度を与えることがわかっているため、この割合に従って報酬をそれぞれ +1、0、-1 で与える。これらの対話制御の報酬 r_C と情報案内の報酬 r_I の和によって、学習時の報酬を定義する。

5. 実ユーザによる評価

POMDP による対話制御を学習するため、ルールベースの対話システム [19] によって 10 話者 606 発話、そこから構築した統計ベースの対話システムによって 8 話者 312 発話の対話データを収集した。また、評価用に 12 話者、626 発話からなる実ユーザとの対話を収集し、構築した対話制御を評価した。このうち、5 秒以上のユーザの無音区間は 58 回存在し、システムが行動を決定する点は 684 個であった。評価尺度としてはユーザの意図の捕捉精度 (DST)[26] とシステムの行動選択の精度 (ACT)、及び 1 対話あたりの報酬を用いる。

比較手法としては、ルールによって動作するシステム (Rule)[19] と、焦点を用いない POMDP による対話制御 (POMDP w.o. focus) を用いる。この結果を表 2 に示す。提案手法は比較手法より高い精度でユーザの意図を捕捉し、的確な対話モジュールを応答に選択できていることがわかる。特にシステムの行動選択の精度 (ACT) は、比較手法に対して有意な差 ($p < 0.01$) があつた。ユーザの意図の捕捉精度とモジュールの選択精度を各状態ごとに

表 3 ユーザの意図の捕捉精度 (DST) (適合率、再現率、F 値)

tag	count	Rule			POMDP w.o. focus			POMDP proposed		
		P	R	F	P	R	F	P	R	F
TP	239	0.985	0.812	0.890	0.908	0.828	0.867	0.951	0.820	0.880
ST	82	0.567	0.829	0.673	0.906	0.585	0.711	0.913	0.768	0.834
QA	299	0.798	0.819	0.809	0.825	0.963	0.888	0.843	0.950	0.893

表 4 システムの行動選択の精度 (ACT) (適合率、再現率、F 値)

tag	count	Rule			POMDP w.o. focus			POMDP proposed		
		P	R	F	P	R	F	P	R	F
TP	259	0.884	0.822	0.852	0.908	0.764	0.830	0.946	0.807	0.871
ST	90	0.600	0.800	0.686	0.878	0.400	0.550	0.893	0.744	0.812
QA	290	0.782	0.828	0.804	0.798	0.966	0.874	0.848	0.941	0.892
PP	38	0.929	0.342	0.500	0.000	0.000	0.000	0.833	0.921	0.875

表 2 ユーザの意図の捕捉精度、システムの行動選択の精度、及び 1 対話あたりの報酬

	Rule	POMDP w.o. focus	POMDP proposed
DST (1-best)	0.812 (=508/626)	0.853 (=534/626)	0.867 (=543/626)
ACT (1-best)	0.788 (=539/684)	0.751 (=514/684)	0.854 (=584/684)
報酬	33.9	26.6	40.0

求めたものを表 3 と表 4 に示す。各状態に対する適合率 (P)、再現率 (R) とその調和平均 (F) が示されている。なお、挨拶 (GR) と無効入力 (II) に関しては、数が少なかったためここには示していない。まず、ユーザの意図の捕捉精度 (DST) では、提案手法は焦点を用いない比較手法よりも音声認識の誤りに頑健で、音声言語理解の精度を向上できていることがわかる。実際の例を見ると、特に Topic Presentation (TP) や Story Telling (ST) であるものが Question Answering (QA) に誤認識されていた例が修正されていた。また、POMDP の信念の更新を用いることで、特に QA や TP に誤分類された ST が修正されていた。

次にシステムの行動選択の精度 (ACT) について述べる。ここでも KS と GR は数が少なかったため示していない。また、CO は正しい行動としてはラベル付与されていない。この結果から、提案手法はすべてのモジュールの選択精度において比較手法を上回っていることがわかる。特に、Story Telling (ST) と Proactive Presentation (PP) では有意な差が確認された ($p < 0.01$)。提案手法は、ユーザの意図が No Request (NR) であった場合にユーザの焦点に応じて適切に Topic Presentation (TP) と Proactive Presentation (PP) を選択できていた。

また、提案手法は音声言語理解の結果が誤っている場合に、効果的に Confirmation (CO) を行うことができていた。具体的には、提案手法で行われた 18 回の CO のうち、15 回が音声言語理解が誤った場合に行われたものであった (15/18=83.3%)。これと比較して、焦点を利用しない既存手法の POMDP で行われた 2 回の CO は、いずれも音声言語理解が正しい場合であった (0/2=0.0%)。

Dialogue 1

Sys: こんにちは。	
Usr: こんにちは。今日の野球のニュースを教えてください。	$s=TP, P(o s)=0.8, b(s=TP)=0.800, f=0$
Sys: マー君、脅威のスプリット。一緒に練習のサブシア、「うお！何だこれ！」	$a=TP, b(s=TP, f=0)=0.8$
Usr: 脅威のスプリットってなんですか、	$s=QA, P(o s)=0.4, b(s=QA)=0.934, f=1$
Sys: 田中のスプリットは直球の軌道で、ホームプレート直前で沈み込むよ。	$a=QA, b(s=QA, f=1)=0.9$
Usr: 田中はどこで練習してたの。	$s=QA, P(o s)=0.8, b(s=QA)=0.973, f=1$
Sys: 田中が20日、ニューヨークヤンキースのブルペンで練習を行ったよ。	$a=QA, b(s=QA, f=1)=1.0$
Usr: (無音)	$s=NR, P(o s)=1.0, b(s=NR)=1.00, f=1$
Sys: ちなみに18日の立ち投げでは、従来のものよりも沈むツーシームを練習したよ。	$\alpha=PP, b(s=NR, f=1)=1.0$
Usr: ツーシームが沈むの？	$s=QA, P(o s)=0.6, b(s=QA)=0.867, f=1$
Sys: 18日に田中のブルペン投球を受けた捕手は「ツーシームが従来の横の変化だけでなく沈んだ」と証言していたよ。	$a=QA, b(s=QA, f=1)=0.9$
Usr: ふーん、他にニュースは？	$s=TP, P(o s)=0.8, b(s=TP)=0.984, f=0$
Sys: ヤンキースイチロー、猛練習！	$a=TP, b(s=TP, f=0)=1.0$
Usr: それ教えて。	$s=ST, P(o s)=0.9, b(s=ST)=0.871, f=0$
Sys: ニューヨークヤンキースのイチローが普段より一時間早く球場入りし、練習に汗を流したよ。	$a=ST, b(s=ST, f=0)=0.9$
...	

Dialogue 1, 4th turn without user focus

Usr: (無音)	$s=NR, P(o s)=1.0, b(s=NR)=1.00$
Sys: (Keep silence)	$\alpha=KS, b(s=NR)=1.0$

Dialogue 2

Sys: こんにちは。	
Usr: こんにちは。今日の野球のニュースを教えてください。	$s=TP, P(o s)=0.8, b(s=TP)=0.800, f=0$
Sys: 巨人松井隆時コーチ、ランチャ特打で弾丸5発。桃井球団社長に、「契約して」	$a=TP, b(s=TP, f=0)=0.8$
Usr: 何があったの？	$s=QA, P(o s)=0.8, b(s=QA)=0.532, f=0$
Sys: 吾崎でついにゴジラ弾が復活したよ。	$\alpha=ST, b(s=QA, f=0)=0.5$
Usr: ゴジラ弾って何？	$s=QA, P(o s)=0.8, b(s=QA)=0.806, f=1$
Sys: ゴジラ先生がホームランのお手本を見せ、若手には貴重な経験となったよ。	$a=QA, b(s=QA, f=1)=0.8$
Usr: 次のニュース	$s=TP, P(o s)=0.8, b(s=TP)=0.986, f=0$
Sys: スーパーエースへ。藤浪晋太郎はルーキーイヤーから進化しているのか	$a=TP, b(s=TP, f=0)=1.0$
...	

図 5 対話例

加えて、提案手法は 35 回の Proactive Presentation (PP) を行ったが、そのうち 17 回で次のターンにユーザからのさらなる質問が行われた。提案手法がニュース案内におけるインタラクションを活性化させることができたといえる。

システムとの対話例を図 5 に示す。この 1 つ目の例では、提案システムではユーザが何も話さなくなった場合に適切に PP によって情報推薦を行っているが、焦点を用いない既存システムでは何も話さない (KS) 政策が学習されている。2 つ目の例では、音声言語理解の誤りから、ユーザ状態 $s=QA$ が低いスコアで与えられたが、学習された政策により正しい行動である ST を選択している。

6. まとめ

本研究では日々動的に更新される Web のニュース記事が

ら情報案内を行う対話システムについて述べた。このシステムはユーザの焦点を捕捉することで、ユーザの興味に適応的な情報案内を行うことができる。具体的には、ユーザの焦点を POMDP における対話制御に組み込み、これが適切な対話モジュールの選択に寄与することを実ユーザとの対話実験で示した。また、提案手法は比較手法であるルールによる対話制御や、ユーザの焦点を用いない POMDP より有意に高い精度で対話モジュールの選択を行うことができた。

このシステムはドメイン拡張性のある枠組みで実装されており、より一般的な情報案内に拡張できることが期待される。本研究ではユーザの焦点が情報案内における対話制御に有効であることを示したが、他の談話に関する素性の利用についても検討したいと考えている。

参考文献

- [1] Zue, V., Seneff, S., Glass, J., Polifroni, J., Pao, C., Hazen, T. and Hetherington, L.: JUPITER: a telephone-based conversational interface for weather information, *IEEE Transactions on Speech and Audio Processing*, Vol. 8, No. 1, pp. 85–96 (2000).
- [2] Aust, H., Oerder, M., Seide, F. and Steinbiss, V.: The Philips automatic train timetable information system, *Speech Communication*, Vol. 17, No. 3-4, pp. 249–262 (1995).
- [3] Lamel, L., Bennacef, S., Gauvain, J., Dartigues, H. and Temem, J.: User evaluation of the Mask kiosk, *Speech Communication*, Vol. 38, No. 1-2, pp. 131–139 (2002).
- [4] Misu, T. and Kawahara, T.: Bayes Risk-based Dialogue Management for Document Retrieval System with Speech Interface, *Speech Communication*, Vol. 52, No. 1, pp. 61–71 (2010).
- [5] Pan, Y.-C., Yi Lee, H. and Shan Lee, L.: Interactive spoken document retrieval with suggested key terms ranked by a Markov decision process, *IEEE Transactions on Audio, Speech, and Language Processing*, Vol. 20, No. 2, pp. 632–645 (2012).
- [6] Wilcock, G.: WikiTalk: A spoken Wikipedia-based open-domain knowledge access system, *COLING-2012 Workshop on Question Answering for Complex Domains*, pp. 57–69 (2012).
- [7] Bohus, D. and Rudnicky, A. I.: RavenClaw: Dialog Management Using Hierarchical Task Decomposition and an Expectation Agenda, *Proceedings of the 8th European Conference on Speech Communication and Technology*, pp. 597–600 (2003).
- [8] Bratman, M. E., Israel, D. J. and Pollack, M. E.: Plans and resource-bounded practical reasoning, *Computational Intelligence*, Vol. 4, No. 3, pp. 349–355 (1988).
- [9] Lucas, B.: Voicexml, *Communications of the ACM*, Vol. 43, No. 9, p. 53 (2000).
- [10] Wilcock, G. and Jokinen, K.: WikiTalk Human-robot Interactions, *Proceedings of the 15th ACM on International Conference on Multimodal Interaction*, pp. 73–74 (2013).
- [11] Ritter, A., Cherry, C. and Dolan, W. B.: Data-driven Response Generation in Social Media, *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pp. 583–593 (2011).
- [12] Roy, N., Pineau, J. and Thrun, S.: Spoken dialogue management using probabilistic reasoning, *Proceedings of the 38th Annual Meeting on Association for Computational Linguistics*, pp. 93–100 (2000).
- [13] Levin, E., Pieraccini, R. and Eckert, W.: A stochastic model of human-machine interaction for learning dialog strategies, *IEEE Transactions on Speech and Audio Processing*, Vol. 8, No. 1, pp. 11–23 (2000).
- [14] Williams, J. D. and Young, S.: Partially observable Markov decision processes for spoken dialog systems, *Computer Speech & Language*, Vol. 21, No. 2, pp. 393–422 (2007).
- [15] Young, S., Gašić, M., Keizer, S., Mairesse, F., Schatzmann, J., Thomson, B. and Yu, K.: The hidden information state model: A practical framework for POMDP-based spoken dialogue management, *Computer Speech & Language*, Vol. 24, No. 2, pp. 150–174 (2010).
- [16] Yoshino, K., Watanabe, S., Roux, J. L. and Hershey, J. R.: Statistical Dialogue Management using Intention Dependency Graph, *Proceedings of the 6th International Joint Conference on Natural Language Processing*, pp. 962–966 (2013).
- [17] Shibata, T., Egashira, Y. and Kurohashi, S.: Chat-like Conversational System based on Selection of Reply Generating Module with Reinforcement Learning, *Proceedings of the 5th International Workshop Series on Spoken Dialog Systems*, pp. 124–129 (2014).
- [18] Meguro, T., Higashinaka, R., Minami, Y. and Dohsaka, K.: Controlling listening-oriented dialogue using partially observable Markov decision processes, *Proceedings of the 23rd International Conference on Computational Linguistics*, pp. 761–769 (2010).
- [19] 吉野幸一郎, 森 信介, 河原達也: 述語項の類似度に基づく情報抽出・推薦を行う音声対話システム, *情報処理学会論文誌*, Vol. 52, pp. 3386–3397 (2011).
- [20] Schegloff, E. A. and Sacks, H.: Opening up closings, *Semiotica*, Vol. 8, No. 4, pp. 289–327 (1973).
- [21] Grosz, B. J. and Sidner, C. L.: Attention, Intentions, and the Structure of Discourse, *Computational Linguistics*, Vol. 12, No. 3, pp. 175–204 (1986).
- [22] Yoshino, K. and Kawahara, T.: Information Navigation System Based on POMDP that Tracks User Focus, *Proc. of SIGDial*, pp. 32–40 (2014).
- [23] Monahan, G. E.: State of the Art? A Survey of Partially Observable Markov Decision Processes: Theory, Models, and Algorithms, *Management Science*, Vol. 28, No. 1, pp. 1–16 (1982).
- [24] Watkins, C. J. and Dayan, P.: Q-learning, *Machine learning*, Vol. 8, No. 3, pp. 279–292 (1992).
- [25] Bonet, B.: An e-optimal grid-based algorithm for partially observable Markov decision processes, *Proceedings of International Conference on Machine Learning*, pp. 51–58 (2002).
- [26] Williams, J. D., Raux, A., Ramachandran, D. and Black, A.: The dialog state tracking challenge, *Proceedings of the 14th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pp. 404–413 (2013).