

# Conversational system for information navigation based on POMDP with user focus tracking<sup>☆</sup>

Koichiro Yoshino<sup>\*</sup>, Tatsuya Kawahara

*School of Informatics, Kyoto University, Sakyo, Kyoto 606-8501, Japan*

Received 24 May 2014; received in revised form 21 November 2014; accepted 12 January 2015

Available online 20 January 2015

---

## Abstract

We address a spoken dialogue system which conducts information navigation in a style of small talk. The system uses Web news articles as an information source, and the user can receive information about the news of the day through interaction. The goal and procedure of this kind of dialogue are not well defined. An empirical approach based on a partially observable Markov decision process (POMDP) has recently been widely used for dialogue management, but it assumes a definite task goal and information slots, which does not hold in our application system. In this work, we formulate the problem of dialogue management as a selection of modules and optimize it with POMDP by tracking the dialogue state and focus of attention. The POMDP-based dialogue manager receives a user intention that is classified by a spoken language understanding (SLU) component based on logistic regression (LR). The manager also receives a user focus that is detected by the SLU component based on conditional random fields (CRFs). These dialogue states are used for selecting appropriate modules by policy function, which is optimized by reinforcement learning. The reward function is defined by the quality of interaction to encourage long interaction of information navigation with users. The module which responds to user queries is based on a similarity of predicate-argument (P-A) structures that are automatically defined from a domain corpus. It allows for flexible response generation even if the system cannot find exact matching information to the user query. The system also proactively presents information by following the user focus and retrieving a news article based on the similarity measure even if the user does not make any utterance. Experimental evaluations with real dialogue sessions demonstrate that the proposed system outperformed the conventional rule-based system in terms of dialogue state tracking and action selection. Effect of focus detection in the POMDP framework is also confirmed.

© 2015 Elsevier Ltd. All rights reserved.

MSC: 00-01; 99-00

Keywords: Spoken dialogue system; Dialogue management; Partially observable Markov decision process (POMDP); Focus in dialogue

---

## 1. Introduction

In the past decades, a large number of spoken dialogue systems have been investigated. Many systems are now deployed in the real world, most typically as smart phone applications, which interact with a diversity of users. In the

---

<sup>☆</sup> This paper has been recommended for acceptance by R.K. Moore.

<sup>\*</sup> Corresponding author. Tel.: +81 757537484.

E-mail addresses: [yoshino@ar.media.kyoto-u.ac.jp](mailto:yoshino@ar.media.kyoto-u.ac.jp) (K. Yoshino), [kawahara@i.kyoto-u.ac.jp](mailto:kawahara@i.kyoto-u.ac.jp) (T. Kawahara).

future, interactive robots will be deployed as a communication partner of users. However, a large majority of current applications, such as weather information systems (Zue et al., 2000) and train information systems (Aust et al., 1995; Lamel et al., 2002), are based on a specific task description which includes a definite task goal and necessary slots, such as place and date, for task completion. Users are required to share and follow these concepts; they need to have a clear task goal and specify it according to the system's capability. Some recent systems incorporate general question-answering capability, but it is usually limited to factoid questions such as "when" or "how tall", or pre-defined templates such as "what is your name?". When users ask something beyond the system's capability, the system replies "I can't answer the question", or turns to the Web search and returns the retrieval list in the display. This kind of dialogue is not natural in interaction with humanoid robots since people want to converse with them besides simple commands. A user-friendly conversational system should not reply with "I can't answer the question" even if it cannot find the result exactly matching the user query. Instead, it should present relevant information according to the user's intention and preference. Moreover, robots do not have a display to present a document. They must make a concise verbal reply.

The goal of this work is a conversational system with speech media only which can engage in information navigation. By information navigation, we do not assume a specific task goal, but assume a domain such as sports and travel. The system should present relevant information even if the user request is not necessarily clear and there is not a matching result to the user query. Moreover, the system can occasionally present potentially useful information even without the user's explicit request by following the dialogue context. In this work, we design and develop a news navigation system that uses Web news articles as a knowledge source and presents information based on the users' preference and queries.

There are several studies towards this direction (Kawahara, 2009), but there is not a clear principle nor established methodology to design and implement casual conversation systems. Dialogue management of this kind of systems was usually made in a heuristic manner and often based on simple rules (Bratman et al., 1988; Lucas and, 2000; Bohus et al., 2003). The companions project (Catizone et al., 2008; Cavazza et al., 1630) designed conversational agents that would engage elderly users in sustained conversations based on rules. Misu and Kawahara (2007) developed a Kyoto navigation system that conducts question-answering and proactive presentation by defining a topic structure based on Wikipedia articles. The information state approach to dialogue management (Traum and Larsson, 2003; Kronlid and Lager, 2007) allows for dialogue control to put a topic on hold and return to it later. WikiTalk (Wilcock, 2012; Wilcock and Jokinen, 2013) is a dialogue system that talks about topics in Wikipedia. This system works on the pre-defined scenario that is represented with an automaton, but it forces users to follow the system scenario. Moreover, developers need to implement a new scenario for a new domain or task. A data-driven approach based on phrase-based statistical machine translation (SMT) (Ritter et al., 2011) tries to train response generation from micro-blog data. This approach enables the system to output a variety of responses, but it does not track any user intention or dialogue state to fulfil what the user want to know.

In the past years, machine learning, particularly reinforcement learning (RL), has been investigated for dialogue management. Markov decision processes (MDPs) and partially observable Markov decision processes (POMDPs) are the most successful and are now widely used to model and train dialogue managers (Roy et al., 2000; Levin et al., 2000; Williams and Young, 2007; Young et al., 2010; Yoshino et al., 2013b). However, the conventional scheme assumes that the task and dialogue goal are clearly stated and readily encoded in the RL reward function. This is not true in casual conversation or information navigation addressed in this work.

Some previous work has tackled with this problem. Pan et al. (2012) designed a spoken document retrieval system whose goal is user's information need satisfaction, and defined rewards by using the structure of the target document set. This is possible only for well-defined document search problems. The strategy requires a structure of the document set and definition of user demand satisfaction. Shibata et al. (2014) developed a conversational chatting system. It asks users to make evaluation at the end of each dialogue session to define rewards for reinforcement learning. Meguro et al. (2010) proposed a listening dialogue system. In their work, levels of satisfaction were annotated in the log of dialogue sessions to train a discriminative model. These approaches require costly input from users or developers, who provide evaluation and supervision labels. In this work, we present a framework in which reward is defined for the quality of system actions and also for encouraging long interactions, in contrast to the previous approaches. Moreover, user focus is tracked to make appropriate actions, which are more rewarded.

Descriptions of the proposed conversational information navigation system are provided in Section 2. In Section 3, details of dialogue modules based on the predicate-argument (P-A) structure are explained. In Section 4, we describe spoken language understanding (SLU) modules based on logistic regression (LR) and conditional random fields (CRFs).

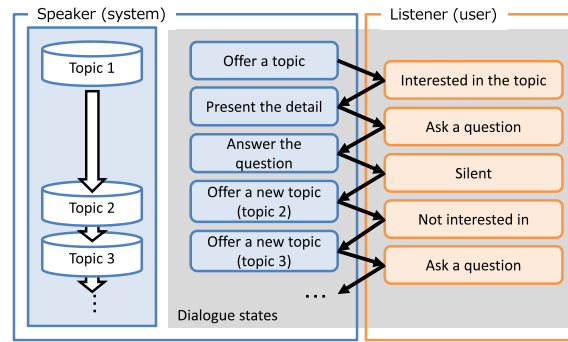


Fig. 1. An example of information navigation.

In Section 5, we give a belief explanation of POMDP and its extension by incorporating user focus. Experimental evaluations of the proposed POMDP-based system with dialogue sessions are reported in Section 6.

## 2. Conversational information navigation system

### 2.1. Task of information navigation

Information navigation does not assume a designed task and goal, but provides useful information according to the users' interest. When the user demands are not clear, the system clarifies the user demands through an interaction. The system presents relevant information even if there is not exactly matching result to the user query. Moreover, the system presents potentially useful information even when the user does not make any explicit request.

In natural human–human conversations, participants usually have topics they plan to talk about, and they progress the dialogue in accordance with the topics (Schegloff and Sacks, 1973). An example is shown in Fig. 1. First, the speaker offers a new topic and probes the interest of the listener. If the listener shows interest, the speaker describes details of the topic. If the listener asks a specific question, the speaker answers it. On the other hand, if the listener is not interested in the topic, the speaker avoids the details of that topic and changes the topic. We realize information navigation with this dialogue style.

In the proposed scheme, the system presents topics that it can talk about, describes the details of the current topic, or presents a topic related to the dialogue history when the system has an initiative. The user can explicitly designate the topic of his interest, or request the system to change the topic, or ask a question regarding the current or related topics. The initiative of dialogue comes and goes between the system and the user because it depends on the specification of the user demand. In this kind of dialogue, discourse features such as focus play a critical role on the decision of what to speak (Grosz and Sidner, 1986). The proposed scheme introduces the user focus as one of the parameters of dialogue management.

In this paper, we present a news navigation system (Yoshino et al., 2011) that is one realization of information navigation. The system provides topics collected from Web news texts, and the user gets information according to his interests and queries. Here, a “topic” is content of newspaper articles, and a “domain” is defined as the category of the news. For example, an article that describes the game of New York Yankees is one topic. The domain of the system in this work is baseball, and the system exploits the domain knowledge to conduct effective query matching. While the system proceeds a dialogue by following the today's topics, the user can ask anything, even related to the past events, within the baseball domain, at any time.

### 2.2. System overview

An overview of the proposed system is depicted in Fig. 2. The system has seven modules, each of which implements a class of actions. Each module takes as input a recognized user utterance, an analysed predicate-argument (P-A) structure and the detected user intention and focus.

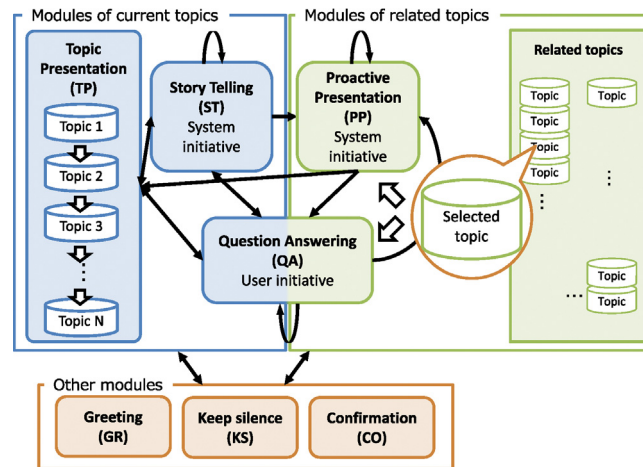


Fig. 2. Overview of the information navigation system.

The system begins a dialogue with the “topic presentation ( TP)” module, which presents a new topic selected from news articles. The system chooses the next module based on the user’s response. In this work, we assume that each news article corresponds to a single topic, and the system presents a headline of the news in the TP module. If the user shows interest (positive response) in the topic without any specific questions, the system selects the “story telling ( ST)” module to give details of the news. In the ST module, the system provides a summary of the news article by using lead sentences. The system can also provide related topics with the “proactive presentation ( PP)” module. This module is invoked by the system’s initiative; this module is not invoked by any user request. If the user asks a specific question regarding the topic, the system switches to the “question answering ( QA)” module to answer the question. This module deals with questions on the presented topic and related topics.

The modules of PP and QA are based on a dialogue framework which uses the similarity of P-A structure (Yoshino et al., 2011) between user queries and news articles, and retrieves or recommends an appropriate sentence from the news articles. This method searches for appropriate information from automatically parsed documents by referring to domain knowledge that is automatically extracted from a domain corpus. The details are described in Section 3.

Transitions between the modules are allowed as shown in Fig. 2. The modules “greeting ( GR)”, “keep silence ( KS)” and “confirmation ( CO)” are also implemented. The GR module generates fixed greeting patterns by using regular expression matching. The CO module makes a confirmation if the system does not have certainty about the user query. In terms of dialogue flow, these modules can be called at any time.

### 2.3. User focus in information navigation

“Focus” in discourse is “attentional state (that) contains information about the objects, properties, relations, and discourse intentions that are most salient at any given point” (Grosz and Sidner, 1986). The user has specific attention to an object if the user utterance contains the focus. In this work, we define the user focus as “the main piece of information of interest to the user.” It makes a central component when making a reply or selecting a relevant topic at the current dialogue state. For example, given “Did Ichiro perform brilliantly?,” user focus is “Ichiro” because the system reply should include information on Ichiro. This information is annotated on content words or named entities in user utterances. In the system proposed in this paper, we parameterize whether any user focus is detected in the user’s utterance instead of the focused object.

## 3. Presentation of relevant information based on P-A structure

In this section, we describe flexible matching of P-A structure on which the proposed question answering ( QA) and proactive presentation ( PP) modules are based (Yoshino et al., 2011). Text of news articles and user utterances are parsed to extract a P-A structure (an example is shown in Fig. 3). A P-A structure represents a sentence with a predicate,

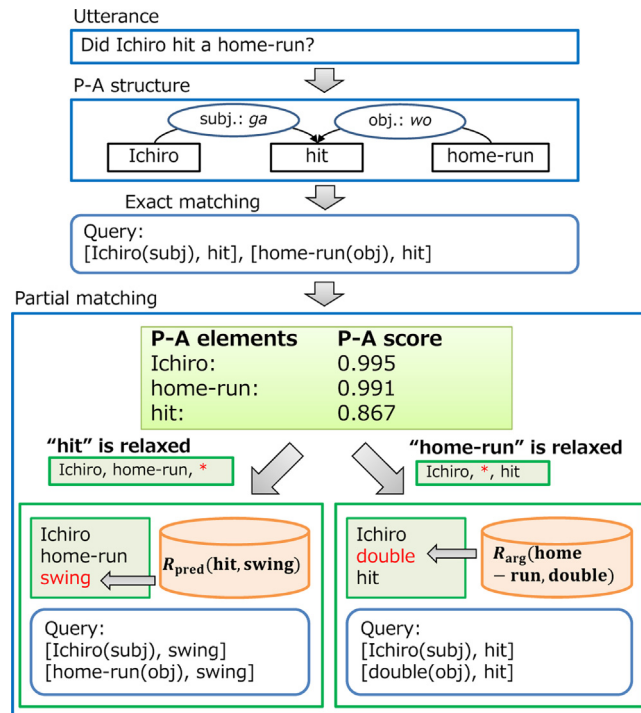


Fig. 3. The proposed scheme of partial matching.

arguments and their semantic role labels (Johansson and Nugues, 2008; Hajič et al., 2009; Matsubayashi et al., 2012). We used the Japanese text parser KNP<sup>1</sup> (Kawahara and Kurohashi, 2006) to obtain a P-A structure, a structure of the predicate phrase and a set of the argument phrases that depend on the predicate. In the example, the sentence “Did Ichiro hit a home-run” is parsed as, arguments, “Ichiro (subject)” and “home-run (object)”, depending on the predicate “hit” with respective semantic roles. This structure is a classic concept in natural language processing, but recently automatic semantic parsing has reached a practical level thanks to corpus-based learning techniques (Kawahara and Kurohashi, 2006).

In the QA module, the system searches for a sentence that matches the user utterance in the P-A structure. When the system fails to find exact information that matches the user query, it tries to find relevant sentences. In the PP module, when the user does not speak for a while, the system tries to make proactive information presentation. It is based on partially matched entries of the current or latest query. The fall-back is similar to collaborative response generation in the conversational spoken dialogue systems (Sadek, 1999), but it is intended for proactive information presentation using general documents of news articles.

### 3.1. Response generation based on partial matching

We describe the response generation used in the QA module. For preference among multiple components in the P-A structure of the user query, we define a significance score based on Naive Bayes. The score is used when relaxing search queries in the QA module, as a weight of the score in the PP module, and as a feature of spoken language understanding. The score is a probability of phrase  $p_i$  (either predicate or argument) being in the document set (news

<sup>1</sup> <http://nlp.ist.i.kyoto-u.ac.jp/index.php?KNP>, 2014/11/06

articles in this work) of a particular domain  $D$  (baseball in this work) by assuming the other set of documents  $\bar{D}$  of different domains,

$$P(D|p_i) = \frac{P(p_i|D) \times P(D)}{P(p_i)} \quad (1)$$

$$\simeq \frac{C(p_i, D) + P(D) \times \gamma}{C(p_i) + \gamma}, \quad (2)$$

where  $C(\cdot)$  stands for a word count and  $P(D)$  is a normalization factor determined by the size of  $D$  and  $\bar{D}$ .  $\gamma$  is a smoothing factor estimated with a Dirichlet prior using the Chinese Restaurant Process (CRP). The score is calculated for every element of P-A structures.

The proposed scheme of partial matching is shown in Fig. 3. First, the module searches for the exactly matched information. If it fails, the module searches for partially matched information. Specifically, we relax (ignore) the component of the least significance score, then search for relevant information. If any entry is not still matched, we relax the next less significant component. If multiple entries are found with this matching, the system selects the most relevant entry by using the relevance measure of the predicate or arguments. The relevance measure of an argument is based on the co-occurrence of words in a document. The relevance of argument phrase  $p_{arg_i}$  and  $p_{arg_j}$  is defined as,

$$R_{arg}(p_{arg_i}, p_{arg_j}) = \frac{\left\{ C(p_{arg_i}, p_{arg_j}) \right\}^2}{C(p_{arg_i}) \times C(p_{arg_j})}. \quad (3)$$

Here,  $p_{arg_i}$  is in the original query and relaxed (ignored) in the partial matching, and  $p_{arg_j}$  of the best relevance score is retrieved for response generation.

The relevance measure of a predicate is based on the distributional similarity (Harris, 1951; Pantel et al., 2009) of a vector of arguments that depend on the predicate. The relevance measure of predicates is calculated by cosine similarity,

$$R_{pred}(p_{pred_i}, p_{pred_j}) = \cos(p_{pred_i}, p_{pred_j}) = \frac{p_{pred_i} \cdot p_{pred_j}}{|p_{pred_i}| |p_{pred_j}|}, \quad (4)$$

where  $p_{pred_i} = (C(p_{arg_1 \& srl_1}), \dots, C(p_{arg_n \& srl_m}))$ .

Here,  $p_{pred_i}$  is the original predicate and  $p_{pred_j}$  is the compared predicate.  $arg_n \& srl_m$  is an argument phrase  $p_{arg_n}$  of the semantic role  $srl_m$ .

### 3.2. Proactive presentation based on similarity of P-A structure

The PP module also uses the significance score and the relevance measure of the predicate and arguments. The PP module selects information that is related to the dialogue history. The sentence-level relevance measure is defined as,

$$R_s(u_i, u_v) = R_{pred}(p_{pred_i}, p_{pred_v}) \times \frac{P(D|p_{pred_i}) + P(D|p_{pred_v})}{2} + \sum_j R_{arg}(p_{arg_j}, p_{arg_v}) \times \frac{P(D|p_{arg_j}) + P(D|p_{arg_v})}{2}. \quad (5)$$

Here,  $u_i$  is a sentence in the dialogue history and  $u_v$  is a candidate information to present.  $p_{pred}$  and  $p_{arg}$  are a predicate and arguments, included in  $u_i$  and  $u_v$ .  $R_{pred}(\cdot)$  and  $R_{arg}(\cdot)$  are relevance of a predicate and arguments. The scores  $P(D|p_{pred})$  and  $P(D|p_{arg})$  are included to weigh the important phrases. The relevance measure is defined to select the related information on the basis of an assumption that the priority order depends on the importance in the domain.



```

SYS, tp, "Matsui, a part-time coach for the Giants, demonstrated 5 homeruns.
        The Giants president told him "Come back as a player!"
USR, st, 0, " Give me the detail?", "Give:me/0 the:detail/0"
SYS, st, "Matsui joined the extra batting training in the Giants' camp, and
        demonstrated his batting to young players."
USR, qa, 1, "Will Matsui become the Giants' manager in the future?",
        "Matsui/0 will:become/0 the Giants' manager/1 in the future/0"
SYS, qa, "The director of Giants told him "Come back as a player!"
USR, tp, 0, "Next news, please.", "Next:news/0 please/0"
SYS, tp, " To be a ace pitcher, has Fujinami improved from the rookie year?
        ..."

```

Fig. 4. An example of annotation for collected dialogue. System utterances have a tuple of three elements separated by a comma: utterer, called module and utterance. User utterances have a tuple of four elements: utterer, the module the user request falls in, binary information of user focus, utterance and user focus annotation on each phrase or P-A element. Here, words in a phrase is joined by a colon and the existence of user focus (1/0) is annotated on each phrase.

The system considers the dialogue history of both the user and the system, and calculates the dialogue-level relevance measure by using the latest  $h$  utterances,

$$R_{pp} = \sum_{i \in h} R_s(u_i, u_v). \quad (6)$$

In this work we set  $h = 2$ .

The PP module presents information that has the highest score of Eq. (6).

#### 4. Spoken language understanding (SLU)

In this section, we present the spoken language understanding (SLU) components of our system. It detects the user's focus and intention and provides them to the dialogue manager. The SLU modules are formulated with a statistical model to give likelihoods which are used in POMDP.

##### 4.1. Training data and annotation

We collected 606 utterances (from 10 users) with a rule-based dialogue system (Yoshino et al., 2011), and 312 utterances (from 8 users) with a preliminary statistical-based dialogue system which was constructed by using the data collected with the rule-based system. An example of annotation is shown in Fig. 4. We highlighted annotation points in the bold font.

To prepare the training data, each utterance was labelled with one of the six modules, indicating the best module to respond. In addition, each phrase or P-A element is labelled whether it is the user's focus or not. The user focus is determined by the attributes (=specifications of words in the domain) and preference order of phrases to identify the most appropriate information that the user wants to know. For example, in the second user utterance in Fig. 4, the user focus is on the phrase "the Giants' manager" instead of "Matsui" as the user's interest is whether Matsui will be a manager.

##### 4.2. User focus detection based on CRF

To detect the user focus, we use a conditional random field (CRF).<sup>2</sup> The problem is defined as a sequential labelling of the focus labels to a sequence of the phrases of the user utterance. Features used are listed in Table 1. ORDER features are the order of the phrase in the sequence and in the P-A structure. We incorporate these features because the user focus often appears in the first phrase of the user utterance. POS features are part-of-speech (POS) tags and their pairs in the phrase. P-A features are the semantic role of the P-A structure.

We also incorporate the P-A significance score defined in Section 3 (P-A Score). The score is discretized to 0.01, 0.02, 0.05, 0.1, 0.2, and 0.5.

Table 1  
Features of user focus detection.

Feature type	Feature
ORDER	Rank in a sequence of phrases Rank in a sequence of elements of P-A
POS	POS tags in the phrase POS tag sequence
POSORDER	Pair of POS tag and its order in the phrase
P-A	Which semantic role the phrase has
P-AORDER	Which semantic roles exist on the utterance
P-A score	Pair of semantic role and its order in the utterance P-A significance score
WORD	Words in the phrase Pair of words in the phrase Pair of word and its order in the phrase

Table 2  
Accuracy of user focus detection.

	Accuracy
Phrase	78.5%
Phrase (with WORD features)	80.8%
Sentence (focus exists or not)	99.9%
Sentence (focus exists or not) (with WORD features)	99.9%

Table 2 shows the accuracy of user focus detection, which was conducted via five-fold cross-validation of the entire training data. “Phrase” is phrase-level accuracy and “sentence” indicates whether the presence of any user focus phrase was correctly detected, regardless of whether the correct phrase was identified. This table shows that WORD features are effective for detecting the user focus, but they are not essential in the sentence-level accuracy. For portability across domains, we adopt the sentence-level focus information, and do not use the WORD features.

The CRF gives a probability of whether any user focus is detected for each phrase  $h_l$ . The probability of focus existence  $o_f$  for a sentence  $h$  is calculated by combining the sequence of probabilities of user focus label 0;

$$P(o_f|h) = 1 - \prod_l P(o_{f_l} = 0|h_l). \quad (7)$$

#### 4.3. User intention analysis based on LR

The module classifies the user intention from the user utterance. We define six intentions as below.

- *TP*: request to the TP module.
- *ST*: request to the ST module.
- *QA*: request to the QA module.
- *GR*: greeting to the GR module.
- *NR*: silence longer than a threshold.
- *II*: irrelevant input due to ASR errors or noise.

<sup>2</sup> CRFsuite (Okazaki, 2007).



Table 3  
Features of user intention analysis.

Feature type	Feature
POS	Bag of POS tags
	Bag of POS bi-gram
P-A	Bag of semantic role labels
	Bag of semantic role labels bi-gram
P-A score	Pair of semantic role label and its rank
TYPICAL	P-A significance score
BC	Occurrence of typical expressions
	Backchannel

Table 4  
Accuracy of user intention analysis.

	All features	Without TYPICAL
<i>TP</i>	98.7%	98.7%
<i>ST</i>	81.1%	65.3%
<i>QA</i>	95.1%	95.5%
<i>GR</i>	97.7%	97.7%
<i>II</i>	31.3%	31.3%
All	93.6%	92.2%

We adopt the logistic regression (LR)-based dialogue act tagging approach (Tur et al., 2006; Higashinaka et al., 2006). The probability of user intention  $o_s$  given an ASR result of the user utterance  $h$  is defined as,

$$P(o_s|h) = \frac{\exp(\omega \cdot \phi(h, o_s))}{\sum_i \exp(\omega \cdot \phi(h, o_{s,i}))}. \quad (8)$$

Here,  $\phi(h, o_s)$  is a feature vector and  $\omega$  is a feature weight. We use POS, P-A and P-A score as a feature set. In addition, we add a typical expression feature (TYPICAL) to classify *TP*, *ST* or *GR* tags. For example, typical expressions in conversations are “Hello” or “Go on,” and those in information navigation are “News of the day” or “Tell me in detail.” Backchannel of the user is a sign for the system to continue the explanation. We add the binary feature of backchannels (BC). Features for the classifier are listed in Table 3.

The classification accuracy in five-fold cross-validation is shown in Table 4. The result for no request (*NR*) is not shown here, because the state is invoked by a prefixed time interval (5 s in this system). The TYPICAL feature improves the classification accuracy especially for *ST* while keeping the domain portability.

#### 4.4. SLU for ASR output

ASR and the intention analysis involve errors. Here,  $s$  is a true user intention and  $o_s$  is an observed intention. The observation model  $P(o_s|s)$  is given by the confidence measure score of the ASR result  $P(h|u)$  (Komatani and Kawahara, 2000) and the likelihood of the intention analysis  $P(o_s|h)$ . The ASR confidence measure is a smoothed posterior probability given  $N$ -best hypotheses.

$$P(o_s|s) = \sum_h P(o_s, h|s) \quad (9)$$

$$\approx \sum_h P(o_s|h)P(h|u). \quad (10)$$

Here,  $u$  is an utterance of the user. We combine the  $N$ -best ( $N=5$ ) hypotheses of the ASR result  $h$ .

The user focus detection also needs to consider ASR errors. The probability of user focus is given by the likelihood of the ASR result  $P(h|u)$  and the likelihood of the user focus detection  $P(o_f|h)$ ,

$$P(o_f|f) = \sum_h P(o_f, h|f) \quad (11)$$

$$\approx \sum_h P(o_f|h)P(h|u). \quad (12)$$

## 5. Dialogue management for information navigation

The POMDP-based statistical dialogue management is formulated as below. The random variables involved at a dialogue turn  $t$  are as follows:

- $s \in I_s$ : user state  
User intention.
- $a \in K$ : system action  
Module that the system selects.
- $o \in I_o$ : observation  
Observed user state, including ASR and intention analysis errors.
- $P(o|s)$ : observation probability  
Output of SLU with its confidence score, which is defined in Eqs. (10) and (12).
- $P(s_j^{t+1}|s_i^t, \hat{a}_k^t)$ : state transition probability  
Model to predict the next user state  $s_j^{t+1}$  given the previous user state  $s_i^t$  and system action  $\hat{a}_k^t$ .
- $b_{s_i} = P(s_i|o^{1:t})$ : belief  
Stochastic variable of the user state.
- $\pi$ : policy function  
This function determines a system action  $a$  given a belief of user  $b$ .  $\pi^*$  is the optimal policy function that is acquired by the training.
- $r$ : reward function  
This function gives a reward to a pair of the user state  $s$  and the system action  $a$ .

The aim of the statistical dialogue management is to output an optimal system action  $\hat{a}^t$  given a sequence of observation  $o^{1:t}$  from 1 to  $t$  time-steps.

Next, we give the belief update that includes the observation and state transition probabilities. The belief update of user state  $s_i$  in time-step  $t$  is defined as,

$$b_{s_j^{t+1}}^{t+1} \propto P(o^{t+1}|s_j^{t+1}) \sum_{s_i} P(s_j^{t+1}|s_i^t, \hat{a}_k^t) b_{s_i}^t. \quad (13)$$

The state transition probability  $P(s_j^{t+1}|s_i^t, \hat{a}_k^t)$  is calculated from the training corpus annotated with user states  $s$  and system actions  $a$ . Once the system estimates the belief  $b_{s_i}^t$ , the policy function outputs the optimal action  $\hat{a}$  as follows:

$$\hat{a} = \pi^*(b). \quad (14)$$

### 5.1. Training of POMDP

We applied Q-learning (Monahan, 1982; Watkins and Dayan, 1992) to acquire the optimal policy  $\pi^*$ . Q-learning relies on the estimation of a Q-function, which maximizes the discounted sum of future rewards of the system action  $a^t$  at a dialogue turn  $t$  given the current belief  $b^t$ . Q-learning is performed by iterative updates on the training dialogue data:

$$Q(b^t, a^t) \leftarrow (1 - \varepsilon)Q(b^t, a^t) + \varepsilon \left[ R(s^t, a^t) + \gamma \max_{a^{t+1}} Q(b^{t+1}, a^{t+1}) \right], \quad (15)$$

where  $\varepsilon$  is a learning rate and  $\gamma$  is a discount factor of a future reward. We experimentally decided  $\varepsilon = 0.01$  and  $\gamma = 0.9$ . The optimal policy given by the Q-function is determined as,

$$\pi^*(b) = \underset{a}{\operatorname{argmax}} Q(b, a). \quad (16)$$

However, it is impossible to calculate the Q-function for all possible real values of belief  $b$ . Thus, we train a limited Q-function given by a grid-based value iteration (Bonet, 2002). The belief is given by a function,

$$b_s = \begin{cases} \eta & \text{if } s = s_i \\ \frac{1 - \eta}{|I_s| - 1} & \text{if } s \neq s_i \end{cases}. \quad (17)$$

Here,  $\eta$  is an approximated likelihood of  $s = s_i$ , that is the best hypothesis of SLU is  $s_i$ . It is quantized into 11 discrete points from 0.0 to 1.0 by 0.1. For example, if  $\eta$  is 0.6,  $s_i$  is  $s_1$  and there are four user states, the function selects a point of  $b_{s_1} = 0.600$ ,  $b_{s_2} = 0.133$ ,  $b_{s_3} = 0.133$ ,  $b_{s_4} = 0.133$ . We also added the case of uniform distribution, that is  $b_{s_i} = 0.25$  for all four user states.

## 5.2. Dialogue management using user focus

The proposed POMDP-based dialogue management refers two kinds of belief information: the user intention  $s$  and the user focus  $f$  (0 or 1). Accordingly, two observations come from SLU: result of user intention analysis  $o_s$  and result of user focus detection  $o_f$ .

The equation of the belief update (Eq. (13)) is extended by introducing the previous focus  $f_l^t$  and current focus  $f_m^{t+1}$ ,

$$b_{s_j^{t+1}, f_m^{t+1}}^{t+1} = P(o_s^{t+1}, o_f^{t+1} | s_j^{t+1}, f_m^{t+1}) \sum_i \sum_l P(s_j^{t+1}, f_m^{t+1} | s_i^t, f_l^t, \hat{a}_k^t) b_{s_i^t, f_l^t}^t. \quad (18)$$

The observation probability is approximated as,

$$\begin{aligned} P(o_s, o_f | s_j, f_m) &= P(o_s | o_f, s_j, f_m) P(o_f | s_j, f_m) \\ &\approx P(o_s | s_j) P(o_f | f_m). \end{aligned} \quad (19)$$

Here, we assume that information of user focus  $f_m$  and  $o_f$  does not affect the observation of the user state  $o_s$ , and the user intention  $s_j$  does not affect the observation of the user focus  $o_f$ . These probabilities are calculated from the probability of the user intention analysis (Eq. (10)) and the probability of the user focus detection (Eq. (12)). The resultant trained policy is,

$$\hat{a} = \pi^*(b^t) = \pi^* \left( \left\{ b_{s_i, f_l}^t \right\} \right). \quad (20)$$

To train the policy, the equation of Q-learning (Eq. (15)) is modified as,

$$Q(b^t, a^t) \leftarrow (1 - \varepsilon) Q(b^t, a^t) + \varepsilon \left[ R(s^t, f^t, a^t) + \gamma \max_{a^{t+1}} Q(b^{t+1}, a^{t+1}) \right]. \quad (21)$$

The state transition probability in Eq. (18) is developed as,

$$P(s_j^{t+1}, f_m^{t+1} | s_i^t, f_l^t, \hat{a}_k^t) = \underbrace{P(s_j^{t+1} | f_m^{t+1}, s_i^t, f_l^t, \hat{a}_k^t)}_{\text{state model}} \underbrace{P(f_m^{t+1} | s_i^t, f_l^t, \hat{a}_k^t)}_{\text{focus model}}. \quad (22)$$

Thus, the observations  $o_s$  and  $o_f$  are controlled by hidden states  $f$  and  $s$  that are decided by the state transition probabilities,

$$\text{focus model} = P(f^{t+1} | f^t, s^t, a^t), \quad (23)$$

$$\text{state model} = P(s^{t+1} | f^{t+1}, f^t, s^t, a^t). \quad (24)$$

A graphical model of the proposed model is shown in Fig. 5.



Table 6

Summary of results comparing with the rule-based system and POMDP without user focus.

	Rule-based	POMDP w.o. focus	POMDP proposed
Accuracy of DST	0.812	0.853	0.867
(1-best)	(=508/626)	(=534/626)	(=543/626)
Accuracy of actions	0.788	0.751	<b>0.854</b>
(1-best)	(=539/684)	(=514/684)	(=584/684)
Average reward	33.9	26.6	40.0

Table 7

Performance of dialogue state tracking (DST) (precision, recall and F-measure)

Tag	Count	Rule-based			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	<b>0.834</b>
QA	299	0.798	0.819	0.809	0.825	0.963	0.888	0.843	0.950	0.893

the dialogue if the frustration points accumulate more than 10. Reward of +20 is given if 20 turns are passed to reward a long continued dialogue.

The second reward is a reward for the success of information navigation ( $r_1$ ). The system will receive a positive reward +1 if the modules **QA** and **PP** succeed in presenting appropriate information. If they fail, the system will receive a negative reward −1. For the training of POMDP, the function that gives  $r_1$  randomly outputs a positive or negative reward with an accordance of the success rate of each module (Yoshino et al., 2011).

The total reward that the system will receive is,

$$r_C + r_1. \quad (25)$$

## 6. Experimental evaluations

For evaluation of the system, we collected additional 626 utterances (12 users, 24 dialogues; 2 dialogues by each user) with the proposed dialogue system with speech input (Yoshino et al., 2013a). There are 58 cases regarded as no request (*NR*) when a user did not say anything for longer than 5 seconds. The gold-standard is annotated by two annotators. The agreement for the user states was 0.958 and Cohen's kappa (Carletta, 1996) was 0.932. The agreement for the system actions was 0.944 and Cohen's kappa was 0.915. We reprioritized the first annotator who is familiar with the task if the annotation was not agreed. For training, we used the data set described in Section 4.1 (918 utterances in total).

Table 8

Performance of action selection (precision, recall and F-measure)

Tag	Count	Rule-based			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	<b>0.812</b>
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	<b>0.875</b>

<b>Dialogue 1</b>	
Sys: Hello.	
Usr: Hello. Tell me today's news.	$s=TP, P(o s)=0.8, b(s=TP)=0.800, f=0$ $\alpha=TP, b(s=TP, f=0)=0.8$
Sys: Tanaka showed a crazy split. His colleague Sabathia was stunned.	
Usr: What is the crazy split?	$s=QA, P(o s)=0.4, b(s=QA)=0.934, f=1$ $\alpha=QA, b(s=QA, f=1)=0.9$
Sys: His split comes just like a four-seam, but it drops quickly in front of the home plate.	
Usr: Where was he training?	$s=QA, P(o s)=0.8, b(s=QA)=0.973, f=1$ $\alpha=QA, b(s=QA, f=1)=1.0$
Sys: Tanaka practiced at the bullpen in the New York Yankees' camp on Feb. 20th.	
Usr: (pause)	$s=NR, P(o s)=1.0, b(s=NR)=1.00, f=1$ $\alpha=PP, b(s=NR, f=1)=1.0$
Sys: By the way, he also practiced a drop two-seam on Feb. 18.	
Usr: Does his two-seam drop?	$s=QA, P(o s)=0.6, b(s=QA)=0.867, f=1$ $\alpha=QA, b(s=QA, f=1)=0.9$
Sys: The catcher said "Tanaka's two-seam not only slides, but also drops."	
Usr: OK, any other news?	$s=TP, P(o s)=0.8, b(s=TP)=0.984, f=0$ $\alpha=TP, b(s=TP, f=0)=1.0$
Sys: Yankees' Ichiro is training strenuously!	$s=ST, P(o s)=0.9, b(s=ST)=0.871, f=0$ $\alpha=ST, b(s=ST, f=0)=0.9$
Usr: Tell me more.	
Sys: Ichiro came to the field an hour earlier than usual and trained strenuously.	
...	
<b>Dialogue 1, 4th turn without user focus</b>	
Usr: (pause)	$s=NR, P(o s)=1.0, b(s=NR)=1.00$ $\alpha=KS, b(s=NR)=1.0$
Sys: (keep silence)	
<b>Dialogue 2</b>	
Sys: Hello.	
Usr: Hello. Tell me today's news?	$s=TP, P(o s)=0.8, b(s=TP)=0.800, f=0$ $\alpha=TP, b(s=TP, f=0)=0.8$
Sys: Matsui, a part-time coach for the Giants, demonstrated 5 homeruns. The Giants president told him "Come back as a player!"	
Usr: What happened?	$s=QA, P(o s)=0.8, b(s=QA)=0.532, f=0$ $\alpha=ST, b(s=QA, f=0)=0.5$
Sys: The godzilla cannon came back to the Giants' spring camp in Miyazaki.	
Usr: What is the Godzilla cannon?	$s=QA, P(o s)=0.8, b(s=QA)=0.806, f=1$ $\alpha=QA, b(s=QA, f=1)=0.8$
Sys: The Godzilla coach demonstrated his home-run batting to young players.	
Usr: Next please.	$s=TP, P(o s)=0.8, b(s=TP)=0.986, f=0$ $\alpha=TP, b(s=TP, f=0)=1.0$
Sys: To be a ace pitcher, has Fujinami improved from the rookie year?	
...	

Fig. 6. A dialogue example.

### 6.1. Evaluation of dialogue state tracking and action selection

Dialogue state tracking (DST) is a task of tracking the correct user state with noisy inputs including ASR and SLU errors (Williams et al., 2013). It tries to maximize the probability of the belief of the correct states, but we evaluated the accuracy of the 1-best result of the belief update. The user state and the best system action were annotated by human. We also evaluated accuracy of the system action and average reward for dialogue sessions. The accuracy of the system action shows not only the effect of the belief update but also the effect of the trained policy. We have two baseline systems. The first baseline system is a rule-based dialogue manager that is operated by a score of the question-answering module using P-A structure (Yoshino et al., 2011) and regular expressions for the TP and GR modules. The other baseline system is operated by the POMDP-based dialogue manger that does not refer to user focus (POMDP w.o. focus).

The DST accuracy, accuracy of action selection and average reward are summarized in Table 6. This result shows that the proposed method tracks the dialogue state of the users with high accuracy, and responds with more appropriate modules. A breakdown of the DST accuracy is shown in Table 7. The table shows precision (P), recall (R) and F-measure (F) of each intention tag. The performance for greeting (GR) and irrelevant input (II) is not shown because the number of these tags was very small ( $\#GR=2, \#II=4$ ). Our proposed framework improved the SLU accuracy and robustness against ASR errors, especially reducing misclassification to question answering (QA) that were actually topic presentation (TP) or story telling (ST). Moreover, the belief update can detect the ST state even if the SLU incorrectly predicts QA or TP.

The proposed method also improved the accuracy of action selection. A breakdown is shown in Table 8. We show the results of TP, ST, QA and PP because the number of KS and GR was very small ( $\#GR=2, \#KS=4$ ), and CO was not labelled as a correct action. The proposed method outperformed the baseline systems for all actions. The overall

improvement over the other two systems shown in Table 6 is statistically significant ( $p < 0.01$ ). In detail, there are significant differences for ST and PP module selection ( $p < 0.01$ ). The proposed method improved the accuracy for topic presentation (TP) and proactive presentation (PP) especially when the user intention was no request (NR). The POMDP without user focus always selected the keep silence (KS) module if the user said nothing (NR).

The proposed method also made more effective confirmations (CO) when the SLU result was not correct. It made confirmations (CO) 18 times, and 15 times of them were done when SLU was incorrect ( $15/18 = 83.3\%$ ). The POMDP without user focus made only two confirmations, when the detected user intention was correct ( $0/2 = 0.0\%$ ).

The proposed method made 35 proactive presentations (PP) and 17 times of them ( $17/35 = 48.6\%$ ) invoked new user questions. This result demonstrates that the proposed system encouraged interactions in news navigation.

## 6.2. Discussion of trained policy

An example dialogue is shown in Fig. 6. In the example, the system selects appropriate actions even if the observation likelihood is low. At the 4th turn of Dialogue 1 in this example, the system with the user focus responds with an action of proactive presentation  $a = PP$ , but the system without the user focus responds with an action of topic presentation  $a = KS$ . At the 2nd turn of Dialogue 2, the user asks a question without a focus. The confidence of  $s = QA$  is lowered by the belief update, and the system selects the story telling module  $a = ST$ . These examples show that the trained policy reflects our design described in Section 5.3: it is better to make a proactive presentation when the user is interested in the topic.

## 7. Conclusions

We have designed and implemented a spoken dialogue system for information navigation of Web news articles updated day-by-day. The system presents relevant information according to the user's interest. We have introduced a user focus detection model, and developed a POMDP framework which tracks the user focus to select the appropriate action module of the dialogue system. In the experimental evaluations, the proposed dialogue management approach determines the state of the user more accurately than the existing rule-based system and the POMDP-based system without user focus information. The proposed method also improved the accuracy of action selection.

In the future, it is expected to refine the system with large-scale data. When we collect more data with the current system, we can incorporate them with the automatic annotation as the annotation accuracy (Tables 2 and 4) is very high. We should also investigate other discourse features for the dialogue management of information navigation such as a more elaborate dialogue topic model.

We plan to apply the system to other domains such as football and stock market in order to confirm the domain portability of the proposed scheme. Then, we can also design a multi-domain navigation system that can converse about a variety of news. Furthermore, other types of document sets that have a category or cluster such as Wikipedia can be explored as the source of information navigation.

## References

- Aust, H., Oerder, M., Seide, F., Steinbiss, V., 1995. The Philips automatic train timetable information system. *Speech Commun.* 17 (3–4), 249–262. [Interactive Voice Technology for Telecommunication Applications](#).
- Bohus, D., Rudnicky, A.I., 2003. Ravenclaw dialog management using hierarchical task decomposition and an expectation agenda. In: *Proceedings of the 8th European Conference on Speech Communication and Technology*, pp. 597–600.
- Bonet, B., 2002. An e-optimal grid-based algorithm for partially observable Markov decision processes. In: *Proceedings of International Conference on Machine Learning*, pp. 51–58.
- Bratman, M.E., Israel, D.J., Pollack, M.E., 1988. Plans and resource-bounded practical reasoning. *Comput. Intell.* 4 (3), 349–355.
- Carletta, J., 1996. Assessing agreement on classification tasks: the kappa statistic. *Comput. Linguist.* 22 (2), 249–254.
- Catizone, R., Dingli, A., Pinto, H., Wilks, Y., 2008. Information extraction tools and methods for understanding dialogue in a companion. In: *LREC, Citeseer*.
- Cavazza, M., De La Camara, R.S., Turunen, M., 2010. How was your day?: a companion ECA. In: *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: Volume 1 – Volume 1*, International Foundation for Autonomous Agents and Multiagent Systems, pp. 1629–1630.
- Grosz, B.J., Sidner, C.L., 1986. Attention, intentions, and the structure of discourse. *Comput. Linguist.* 12 (3), 175–204.



- Hajič, J., Ciaramita, M., Johansson, R., Kawahara, D., Martí, M.A., Màrquez, L., Meyers, A., Nivre, J., Padó, S., Štěpánek, J., Straňák, P., Surdeanu, M., Xue, N., Zhang, Y., 2009. The conll-2009 shared task: syntactic and semantic dependencies in multiple languages. In: *Proceedings of the Thirteenth Conference on Computational Natural Language Learning: Shared Task, CoNLL 2009*, pp. 1–18.
- Harris, Z.S., 1951. *Methods in Structural Linguistics*. University of Chicago Press.
- Higashinaka, R., Sudoh, K., Nakano, M., 2006. Incorporating discourse features into confidence scoring of intention recognition results in spoken dialogue systems. *Speech Commun.* 48 (3), 417–436.
- Johansson, R., Nugues, P., 2008. Dependency-based syntactic–semantic analysis with propbank and nombank. In: *Proceedings of the Twelfth Conference on Computational Natural Language Learning, CoNLL’08*, pp. 183–187.
- Kawahara, T., 2009. New perspectives on spoken language understanding: does machine need to fully understand speech? In: *Proceedings of IEEE workshop on Automatic Speech Recognition and Understanding*, pp. 46–50.
- Kawahara, D., Kurohashi, S., 2006. A fully-lexicalized probabilistic model for Japanese syntactic and case structure analysis. In: *Proceedings of the Main Conference on Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics*, pp. 176–183.
- Komatani, K., Kawahara, T., 2000. Flexible mixed-initiative dialogue management using concept-level confidence measures of speech recognizer output. In: *Proceedings of the 18th conference on Computational linguistics*, pp. 467–473.
- Kronlid, F., Lager, T., 2007. Implementing the information-state update approach to dialogue management in a slightly extended SCXML. In: *Proceedings of the 11th International Workshop on the Semantics and Pragmatics of Dialogue (DECALOG)*, pp. 99–106.
- Lamel, L., Bannacef, S., Gauvain, J., Dartigues, H., Temem, J., 2002. User evaluation of the Mask Kiosk. *Speech Commun.* 38 (1–2), 131–139.
- Levin, E., Pieraccini, R., Eckert, W., 2000. A stochastic model of human–machine interaction for learning dialog strategies. *IEEE Trans. Speech Audio Process.* 8 (1), 11–23.
- Lucas, B., 2000. VoiceXML. *Commun. ACM* 43 (9), 53.
- Matsubayashi, Y., Miyao, Y., Aizawa, A., 2012. Framework of semantic role assignment based on extended lexical conceptual structure: comparison with verbnet and framenet. In: *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics. Association for Computational Linguistics*, pp. 686–695.
- Meguro, T., Higashinaka, R., Minami, Y., Dohsaka, K., 2010. Controlling listening-oriented dialogue using partially observable markov decision processes. In: *Proceedings of the 23rd International Conference on Computational Linguistics*, pp. 761–769.
- Misu, T., Kawahara, T., 2007. Speech-based interactive information guidance system using question-answering technique. In: *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. IV-145–IV-148.
- Monahan, G.E., 1982. State of the art? a survey of partially observable Markov decision processes: theory, models, and algorithms. *Manage. Sci.* 28 (1), 1–16.
- Okazaki, N., 2007. CRFsuite: A Fast Implementation of Conditional Random Fields (CRFs). <http://www.chokkan.org/software/crfsuite/>
- Pan, Y.-C., yi Lee, H., shan Lee, L., 2012. Interactive spoken document retrieval with suggested key terms ranked by a Markov decision process. *IEEE Trans. Audio Speech Lang. Process.* 20 (2), 632–645.
- Pantel, P., Crestan, E., Borkovsky, A., Popescu, A.-M., Vyas, V., 2009. Web-scale distributional similarity and entity set expansion. In: *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 2, Stroudsburg, PA, USA*, pp. 938–947.
- Ritter, A., Cherry, C., Dolan, W.B., 2011. Data-driven response generation in social media. In: *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pp. 583–593.
- Roy, N., Pineau, J., Thrun, S., 2000. Spoken dialogue management using probabilistic reasoning. In: *Proceedings of the 38th Annual Meeting on Association for Computational Linguistics*, pp. 93–100.
- Sadek, D., 1999. Design consideration on dialogue systems: from theory to technology – the case of Artemis. In: *Proceeding of the ESCA Workshop on Interactive Dialogue in Multi-modal Systems*, pp. 173–187.
- Schegloff, E.A., Sacks, H., 1973. Opening up closings. *Semiotica* 8 (4), 289–327.
- Shibata, T., Egashira, Y., Kurohashi, S., 2014. Chat-like conversational system based on selection of reply generating module with reinforcement learning. In: *Proceedings of the 5th International Workshop Series on Spoken Dialog Systems*, pp. 124–129.
- Traum, D.R., Larsson, S., 2003. The information state approach to dialogue management. In: *Current and New Directions in Discourse and Dialogue*. Springer, pp. 325–353.
- Tur, G., Guz, U., Hakkani-Tur, D., 2006. Model adaptation for dialog act tagging. In: *Proceedings of IEEE Workshop on Spoken Language Technology*. IEEE, pp. 94–97.
- Watkins, C.J., Dayan, P., 1992. Q-learning. *Mach. Learn.* 8 (3), 279–292.
- Wilcock, G., 2012. WikiTalk: a spoken wikipedia-based open-domain knowledge access system. In: *COLING-2012 Workshop on Question Answering for Complex Domains*, pp. 57–69.
- Wilcock, G., Jokinen, K., 2013. WikiTalk human–robot interactions. In: *Proceedings of the 15th ACM on International Conference on Multimodal Interaction*, pp. 73–74.
- Williams, J.D., Raux, A., Ramachandran, D., Black, A., 2013. The dialog state tracking challenge. In: *Proceedings of the 14th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pp. 404–413.
- Williams, J.D., Young, S., 2007. Partially observable Markov decision processes for spoken dialog systems. *Comput. Speech Lang.* 21 (2), 393–422.
- Yoshino, K., Mori, S., Kawahara, T., 2011. Spoken dialogue system based on information extraction using similarity of predicate argument structures. In: *Proceedings of the 12th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pp. 59–66.
- Yoshino, K., Mori, S., Kawahara, T., 2013a. Incorporating semantic information to selection of web texts for language model of spoken dialogue system. In: *Proceedings of IEEE International Conference on Acoustic, Speech and Signal Processing*, pp. 8252–8256.

- Yoshino, K., Watanabe, S., Roux, J.L., Hershey, J.R., 2013b. Statistical dialogue management using intention dependency graph. In: *Proceedings of the 6th International Joint Conference on Natural Language Processing*, pp. 962–966.
- Young, S., Gašić, M., Keizer, S., Mairesse, F., Schatzmann, J., Thomson, B., Yu, K., 2010. The hidden information state model: a practical framework for POMDP-based spoken dialogue management. *Comput. Speech Lang.* 24 (2), 150–174.
- Zue, V., Seneff, S., Glass, J., Polifroni, J., Pao, C., Hazen, T., Hetherington, L., 2000. Jupiter: a telephone-based conversational interface for weather information. *IEEE Trans. Speech Audio Process.* 8 (1), 85–96.