Available online at www.sciencedirect.com

ScienceDirect

Information Processing and Management 43 (2007) 225–236

INFORMATION PROCESSING & MANAGEMENT

www.elsevier.com/locate/infoproman

A semantic Bayesian network approach to retrieving information with intelligent conversational agents

Kyoung-Min Kim, Jin-Hyuk Hong, Sung-Bae Cho *

Department of Computer Science, Yonsei University, 134 Sinchon-dong, Seodaemoon-ku, Seoul 120-749, Republic of Korea

Received 22 November 2005; received in revised form 3 April 2006; accepted 5 April 2006
Available online 5 June 2006

Abstract

As access to information becomes more intensive in society, a great deal of that information is becoming available through diverse channels. Accordingly, users require effective methods for accessing this information. Conversational agents can act as effective and familiar user interfaces. Although conversational agents can analyze the queries of users based on a static process, they cannot manage expressions that are more complex. In this paper, we propose a system that uses semantic Bayesian networks to infer the intentions of the user based on Bayesian networks and their semantic information. Since conversation often contains ambiguous expressions, the managing of context and uncertainty is necessary to support flexible conversational agents. The proposed method uses mixed-initiative interaction (MII) to obtain missing information and clarify spurious concepts in order to understand the intention of users correctly. We applied this to an information retrieval service for websites to verify the usefulness of the proposed method.

Keywords: Conversational agents; Pattern matching; Semantic Bayesian networks; User interface; Mixed-initiative interaction

1. Introduction

There has recently been increased interest in conversational agents that act as effective and familiar information providers. Conversational agents are representative intelligent agents that are capable of responding in an intelligent way (with natural language dialogue) to requests from users. They can understand the intention of users through conversation. After understanding, they are able to offer an appropriate service (Garcia-Serrano, Martínez, & Hernández, 2004; Jennings & Wooldridge, 1995; Macskassy, 1996; Maes, 1994; Symeonidis, Kehagias, & Mitkas, 2003).

Many researchers in the speech recognition community view "dialogue methods" as a way of controlling and restricting interactions. This reflects the persistent belief that spoken dialogue is the most natural and powerful user interface with computers (Allen et al., 2001). Most conversational agents lack flexibility in diverse situations because they are only able to respond repeatedly to users with the fixed answers that they

* Corresponding author.
E-mail addresses: kminkim@sclab.yonsei.ac.kr (K.-M. Kim), hjinh@sclab.yonsei.ac.kr (J.-H. Hong), sbcho@cs.yonsei.ac.kr (S.-B. Cho).

0306-4573/$ - see front matter
doi:10.1016/j.ipm.2006.04.001

have stored in the reply database in advance. Pattern matching, a popular method for constructing conversational agents, works well at sentence level, but it is not feasible when trying to understand dialogues in which context must be considered. Moreover, it is likely to fail to understand complex sentences that require deeper analysis. Recently, researchers have investigated flexible dialogue models using Bayesian networks (BN) (Hong & Cho, 2003; Horvitz, Breese, Heckerman, Hovel, & Rommelse, 1998).

When application domains are complex with many variables, it becomes very difficult to infer the inference of users. In this paper, we propose a conversational agent that uses semantic Bayesian networks (SeBN). This agent not only reduces the complexity of construction, but also infers user's intentions in more detail. Since conversation often contains ambiguous expressions, the ability to manage context or uncertainty is very important in the construction of flexible conversational agents. The proposed method uses mixed-initiative interaction to obtain missing information and clarify for spurious concepts in order to understand the intention of users correctly. This not only reduces the complexity of the networks, but also infers the intention of users more proactively (Allen, 1999).

The remainder of this paper is organized as follows. Section 2 discusses related works in terms of intelligent conversational agents and Bayesian network models in information retrieval systems. Section 3 presents the proposed approach, and Section 4 presents the results of our experiments. Finally, conclusions and suggestions for future work will be described.

2. Backgrounds

2.1. Intelligent conversational agents

Conversational agents can communicate with users with natural language dialogue. This method allows an understanding of intentions through conversation and helps the user by executing an appropriate action. Contrary to conventional interfaces like menus and keywords, the use of dialogue makes it possible to interact more naturally and to include information that is more complicated. Therefore, conversational agents can act as effective user interfaces in complex systems (Macskassy, 1996; Nugues, Godéreaux, El Guedj, & Revolta, 1996).

Techniques such as pattern matching, finite-state-machines and frame-based models are used as popular ways of designing conversational agents. For simple tasks, they are good enough because they are based on a static process that predefines all possible types to match. However, performance is limited with conversation that is more realistic. Dynamic topic changing and problem solving present difficulties, and sometimes dialogue is needlessly long and sentences are repeated. In addition, the size of the database needs to increase when analyzing complicated queries, and information might be duplicated unnecessarily. The plan-based model is different from these approaches because it is able to consider the plans and deciding actions of the user. Every time partial information is gathered from each query, and the agent is able to predict intentions gradually (Perugini & Ramakrishnan, 2003).

Using these techniques, conversational agents have been implemented as guiders for web pages and programs, buying commodities, touring groups and so forth (Zue & Class, 2000). Commercial products include Nicole of NativeMinds, SmartBot of Artificial Life, Verbot of Virtual Personalities, and so on.

2.2. Bayesian networks for information retrieval

Bayesian networks provide graphical representations that explicitly represent the independency among the variables of a given domain as well as the concise specifications of a joint probability distribution (Pearl, 1998). It is a DAG (directed acyclic graph) model that evaluates the belief of hidden variables with evidences using the dependency between them based on the Bayes' rule. Nodes in the networks represent random variables, and the edges denote the dependency of them (parent nodes for causes and child nodes for results). The edge produces a joint probability distribution, so that the parent has prior probability $P(p)$ and the child has the conditional probability $P(c|p)$. Using the conditional independency, the joint probability distribution $P(x_1, x_2, \ldots, x_n)$ can be factored as follows:

$$P(x_1, x_2, \ldots, x_n) = \prod_i P(x_i|Parents(x_i))$$

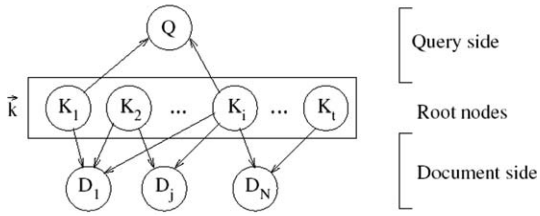


Fig. 1. Bayesian network in IR.

Bayesian networks were first used in information retrieval (IR) by Turtle and Croft, where they showed that the proposed IR model worked better than several traditional probabilistic models for ranking documents (Tutle & Croft, 1990). The model proposed by Ribeiro-Neto and Muntz (1996) not only provides probabilistic justification, but also uses evidence from past queries. More recently, Acid, de Campos, Fernández-Luna, and Huete (2003) presented a model in which network topology can be defined by an exact propagation algorithm, in order to efficiently compute the relevance probabilities of documents. Bayesian networks have been also applied to other problems such as automatic hypertext construction, information filtering, and document clustering and classification.

According to the fundamental model of Bayesian networks in information retrieval systems, queries, documents and keywords are regarded as events (Calado, da Silva, Laender, Ribeiro-Neto, & Vieira, 2004). In the network (Fig. 1), node D_j denotes a document, node Q represents the user query, and node K_i implies a term used in the domain. The vector $\vec{k}$ refers to any possible state of the root node K_i. The relationship between document D_j and query Q is interpreted as the probability of document D_j to occur in query Q. Using Bayes' law and the rule of probabilities, probability $P(D_j|Q)$ can be computed as:

$$P(D_j|Q) = \eta \sum_{k} P(D_j|\vec{k})P(Q|\vec{k})P(\vec{k})$$

3. Intelligent conversational agents

In previous studies, question–answering systems responded to queries from users by matching their pattern from a predefined knowledge base. For simple types of queries, the systems were able to offer correct answers, but people usually use difficult queries to understand the actual meaning like omitting important words based on their background or context (Meng, Wai, & Pierracinni, 2003). In this paper, we classify dialogue models according to the types of queries. Simple types of queries are dealt by using simple question–answering techniques, while the proposed inference model that analyzes semantic relationships between concepts in dialogue manages ambiguous queries.

3.1. Application domain

We developed a flexible conversational agent for virtual representation of websites using MFC, as shown in Fig. 2. It consists of a main window for displaying information, an input text box, and an avatar system with a speech generation engine. When the user types a query, the avatar responds in speech with a corresponding action. Q-avatar (www.qavatar.com) is employed as the avatar system, while Voiceware (voiceware.co.kr), a solution for speech generation, is used to provide the user with a realistic and convenient interface.

The target domain is mobile websites, which can be accessed with cellular phones, digital cameras, and MP3 players. Table 1 describes the attributes of each object in the target database. The database was built by extracting information from five websites: Naver.com (www.nshopping.naver.com), Samsung-mall (www.samsung-mall.co.kr), LG-eshop (www.gseshop.co.kr), Enuri.com (www.enuri.com) and DCinside.com (www.dcincide.com).

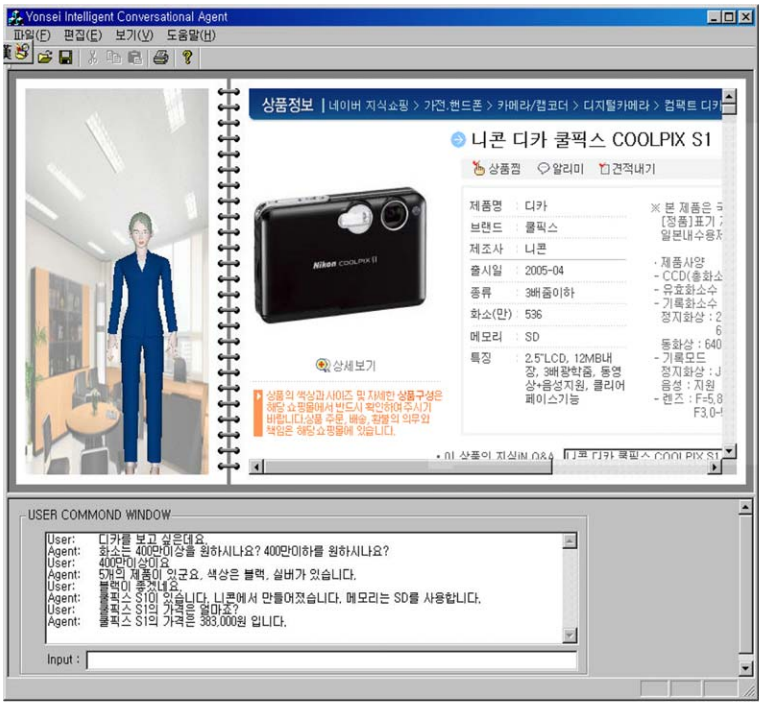


Fig. 2. The system interface.

Table 1
The attributes of objects

Object (#)	Attributes
Cellular phone (240)	Brand, product, model, image, bell, camera, pixel, size, weight, color, price, year
Digital camera (688)	Brand, product, model, image, memory, run-time, size, color, feature, weight, price, year
MP3 player (488)	Brand, product, model, image, pixel, memory, size, weight, feature, zoom, color, price, year

3.2. System architecture

As shown in Fig. 3, the proposed conversational agent is composed of two parts: the multi-modal dialogue interface and the inference modules. The multi-modal dialogue interface provides a familiar user interface as well as deals with general queries based on pattern matching, so system developers might easily construct

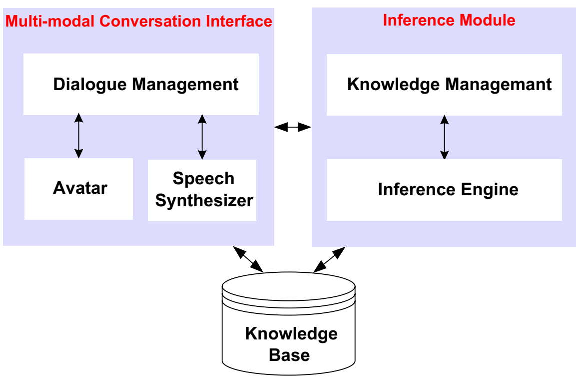


Fig. 3. The architecture of the proposed agent.

answer-scripts independent from the application domain. The inference module is composed of the inference engine and knowledge management module, where the inference engine analyzes what the user wants from ambiguous queries and the knowledge management stores information in the target domain by extracting specific content from web pages and accumulating this into the knowledge base. If there is not enough information to infer the intention, additional information is collected proactively from the user to provide proper responses to users. In the viewpoint of scalability of systems, the developers only construct the inference module according to the target domain.

In order to manage various queries, it is necessary to divide dialogue modules and set a hierarchical priority according to dialogue type. A subsumption architecture (Brooks, 1986), as proposed by Brooks, can be adopted to select one dialogue act per query. As shown in Fig. 4, the dialogue management module works in advance to respond to simple queries named "general dialogue." When it fails, the system regards the query as "information retrieval dialogue" and uses the inference module to manage it.

3.3. Dialogue management module

Fig. 4 shows the overall procedure of managing dialogue. In the preprocessing stage, keywords are extracted from the input query to match keywords in the answer-scripts. Responses can be output when scripts match. A set of candidate scripts are then sequentially matched to find appropriate responses, where the pattern of a given script is composed of keywords in the target domain. In pattern matching, a pattern–response pair can be selected by estimating the matched keyword frequency.

Traditional matching yields a high score when many keywords match, since it only considers the number of matches. However, it might fail because of the amount of information included in an input query as shown in Table 2.

As the knowledge base increases, there will be many duplicated or similar patterns. Therefore, it is necessary that the matching process consider the amount of information. In this paper, matching scores are

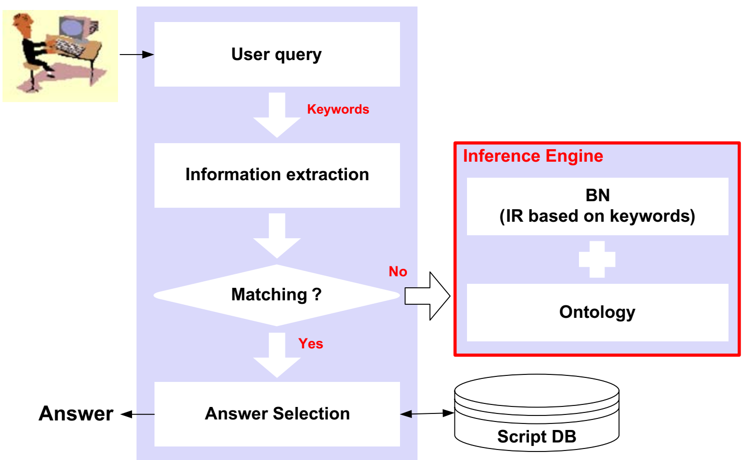


Fig. 4. Dialogue management of the proposed conversational agent.

Table 2
Examples for keyword matching

Query	Keywords **a b c d**	Traditional matching (precision)	Recall	F-measure (α:1)
Pattern A	a b c	0.75 (X)	**1.0** (O)	0.86 (X)
Pattern B	a b c d e	**1.0** (O)	0.8 (X)	0.89 (X)
Pattern C	**a b c d**	1.0 (O)	1.0 (O)	**1.0** (O)

calculated by the F-measure, which is a popular form of text classification. It sets up a weight of α as 1, considering both precision and recall equally.

$$\text{F-measure} = \frac{(\alpha + 1) \times \text{precision} \times \text{recall}}{\alpha \times \text{precision} + \text{recall}}$$

$$\text{precision} = \frac{A}{A + B}$$

$$\text{recall} = \frac{A}{A + C}$$

Pattern–response pairs	Input query	
	Included	Not included
Included	A	B
Not included	C	D

3.4. Inference module using SeBN

To obtain efficient inference, we design semantic Bayesian networks to be composed of the probabilistic inference and the semantic inference. This stepwise modeling helps to understand intentions of users in detail through conversation.

Fig. 5 shows a brief overview of the proposed semantic Bayesian network. It has three levels according to function: keywords, concepts, and targets. The keyword layer consists of words related to the user's query,

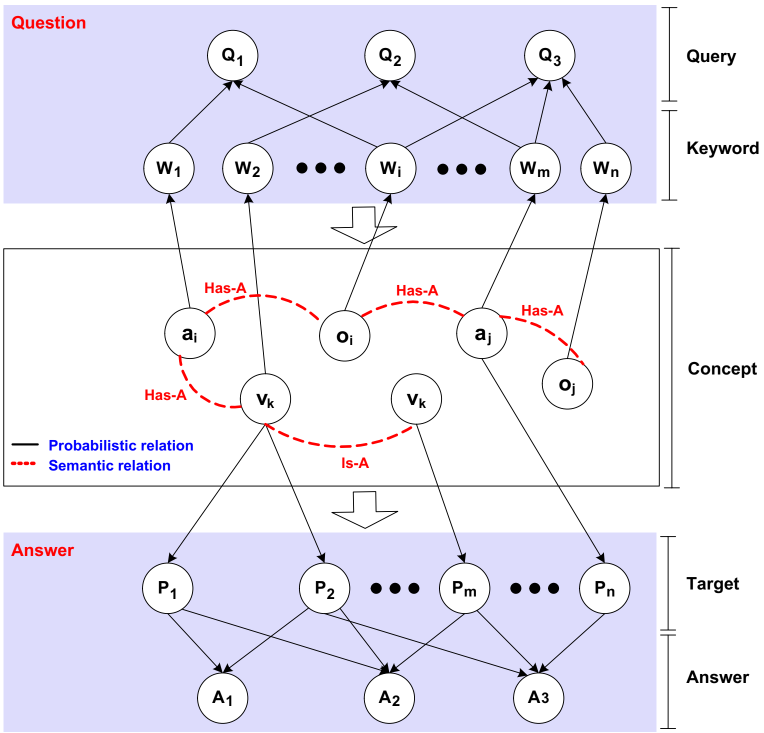


Fig. 5. The architecture of SeBN for inference.

while the concept layer is composed of entities of the domain and their semantic relationships. The target layer represents target information (products) whose attributes are defined. The concept layer is divided into three components: objects, attributes, and values. Each object is a set of attribute–value pairs, where node a_i is an attribute and node v_k is a value in the domain. A solid line represents the probabilistic relationship between nodes, while a dotted line signifies the semantic relationship between them. Especially, in the application domain, there are about 120 concepts and 1400 products used as nodes in the networks.

The probabilistic relationship in semantic Bayesian networks is similar to that of the traditional IR model. First, it infers probabilistically between the keyword layer and the concept layer. The user's query $U = \{k_1, k_2, \ldots, k_t\}$, where the keyword k_i is interpreted as an elementary word in the keyword layer. It sets a keyword node as 1 when the given word in the keyword layer is observed in query Q and otherwise, it is set as 0

$$P(w_i) = \begin{cases} 1, & w_i \in U \\ 0, & w_i \notin U \end{cases}$$

It then infers the probability of each node in the concept layer when all evidence variables associated with the keywords are set. The probability $P(c|W)$, using keyword W in the keyword layer as evidence, is defined as follows:

$$\begin{aligned} P(c|W) &= P(c|w_1, w_2, \ldots, w_N) = \frac{P(c) \times P(w_1, w_2, \ldots, w_N|c)}{P(w_1, w_2, \ldots, w_N)} \approx P(c) \times P(w_1, w_2, \ldots, w_N|c) \\ &= P(c) \times P(w_1|c) \times P(w_2|c) \times \cdots \times P(w_N|c) = P(c) \prod_{i=1}^{n} P(w_i|c) \end{aligned}$$

where N means the sum of the nodes in the keyword set, $c \in O \cup A \cup V$ (W: a set of keywords, O: a set of objects, A: a set of attributes, V: a set of values). After computing the probability of all the nodes in the concept layer, it infers the probability $P(p|C)$ of the product p in the target layer, using them as evidence.

$$\begin{aligned} P(p|C) &= P(p|c_1, c_2, \ldots, c_L) = \frac{P(p) \times P(c_1, c_2, \ldots, c_L|p)}{P(c_1, c_2, \ldots, c_L)} \approx P(p) \times P(c_1, c_2, \ldots, c_L|p) \\ &= P(p) \times P(c_1|p) \times P(c_2|p) \times \cdots \times P(c_L|p) = P(p) \prod_{i=1}^{L} P(c_i|p) \end{aligned}$$

where L means the sum of the nodes in the concept set, $C = O \cup A \cup V$ (C: a set of concepts, O: a set of objects, A: a set of attributes, V: a set of values). In conclusion, the probability $P(p|W)$ between the concept layer and the target layer is defined as:

$$P(p|W) = \eta \sum_{i=1}^{L} (p|c_i) P(W|c_i) P(c_i)$$

It selects a node in the target layer whose probability is higher than the threshold after the inference. It then provides information about the target product to the user when a proper number of nodes are selected. In this paper, we define successful execution as what happens when a product is selected.

This work also includes a preliminary examination of the portability of the BN-based framework across different application domains. Migration to new applications often implies a lack of domain-specific data to train the BN probabilities. Under such circumstances, the BN probabilities can be hand-assigned to reflect the "degree of belief" of the knowledge domain expert. The hand-assigned model requires human knowledge in order to decide the BN probabilities (Hix & Hartson, 1993; Meng et al., 2003). In this paper, we provide guidelines for assigning conditional probabilities manually in order to consider scalability. $P(p|c)$ and $P(w|c)$ are designed according to the same principles. We also present the designing guidelines of conditional probabilities with the use of $P(w|c)$. In the following w describes general principles for

Table 3
Guidelines for assigning values to $P(w=1|c_i=1)$ and $P(w=1|c_i=0)$

Condition	Probability of $P(w=1\|c_i=1)$
w must occur given c_i	0.95–0.99
w often occurs given c_i	0.7–0.8
w may occur given c_i	0.4–0.6
w seldom occurs given c_i	0.2–0.3
w never occurs given c_i	0.01–0.1
	Probability of $P(w=1\|c_i=0)$
w always occurs for concepts other than c_i	0.7–0.9
w sometimes occurs for concepts other than c_i	0.2–0.5
w seldom occurs for concepts other than c_i	0.01–0.1

assigning $P(w=1|c_i=1)$ and $P(w=1|c_i=0)$. The remaining probabilities can be derived from the following formula:

$$P(w=0|c_i=1) = 1 - P(w=1|c_i=1)$$
$$P(w=0|c_i=0) = 1 - P(w=1|c_i=0)$$

Table 3 presents guidelines by which we assign values to the joint probabilities $P(w=1|c_i=1)$ and $P(w=1|c_i=0)$. The assignment is based on the designer's judgment of the possible occurrence frequency of a keyword w and in the concepts of the goal c_i. If we identify a keyword w to be mandatory for a concept of goal c_i, we will hand-assign a high probability roughly from 0.95 to 0.99 for $P(w=1|c_i=1)$. For example, the assigned values of $P(w=1|c_i=1)$ are increased to the range 0.95–0.99 since there is close correlation between the keyword "hue" and a concept "color." Similarly, the assigned values of $P(w=1|c_i=1)$ are decreased to the range 0.7–0.8 because there is high correlation between the keyword "red" and a concept "color." The assigned values for $P(w=1|c_i=1)$ range from 0.2 to 0.3 since the keyword "blue" is not usually associated with the keyword "red." In the conditional probability $P(w=1|c_i=0)$, we assign a high probability for keywords that often occur for concepts other than C_i and a low probability for keywords that seldom occur for concepts other than C_i.

When there is no product selected, it executes the semantic inference of semantic Bayesian networks in the concept layer. There are two major relationships ('*Has-a*', and '*Is-a*') between nodes while '*Is-a*' has two different types ('*O–A*', and '*A–V*') as shown in Table 4.

Table 5 shows the semantic inference executed when the probabilistic inference fails to infer the user's intention. At first, it searches for an object node whose probability is higher than the threshold. Then, it looks up an attribute whose probability is lower than the threshold, which has an '*O–A*' relationship with the object node. It collects supplementary information on the attribute selected and carries out the inference again with information gathered from the user. It repeats the procedure until a target product is selected. In order to discover exactly what the user wants, it needs to gather enough information to infer target products. Traditional information retrieval systems work well only when the user's queries includes enough information for inference. When there is not enough information, however, the proposed method provides a suitable response to the user based on the mixed-initiative interaction.

Table 4
Semantic relations

Class	Sub-classification	Relationship	Examples
Has-a	Object–attribute	*O–A*	Phone-bell, MP3 player-price
	Attribute–value	*A–V*	Size-big, price-low
Is-a	–	Is-a	Size-volume

Table 5
Semantic inference in SeBN

```
[Concept]
    Object: O = {o1, o2, o3, ..., on}
    Attribute: A = {a1, a2, a3, ..., an}
    Value : V = {v1, v2, v3, ..., vn}
i = find_high_probability_object();
// Search an object over the threshold.
if (object(o) > α) {
    j = find_OA_attribute(o);
    // Search attribute 'a' whose probability is below the threshold which has an O–A relationship with node o.
    if (attribute(a) < β)
        response (a, v);
    else
        reject;
}
else {
    j = find_high_probability_attribute();
    // Search attribute 'a' over the threshold
    if (attribute(a) > α) {
    // Search attribute 'a' whose probability is below the threshold which has an O–A relationship with node o.
        i = find_OA_object(o);
        response(o);}
    else
        reject;
}
```

4. Experimental results

4.1. Qualitative analysis: illustration of MII for searching targets with insufficient information

In many cases, users have background knowledge in addition to the content of their conversations, so queries may not include all the information required to infer the user's intentions. The proposed conversational agent uses a mixed-initiative dialogue by requesting additional information from the user. Finally, information on the target product is provided to the user after inference.

As shown in Dialogue 1, the agent searches plural objects from the initial query. Since the agent needs additional information for correct intention inference, it outputs a supplementary query to the user, such as "Which color would you like? Red or Silver?" as the mixed-initiative interaction. The user responds "I'd like red." The agent then executes the probabilistic inference again using semantic Bayesian networks based on this response. Until it detects plural products as the result of prior inference, the agent keeps up the conversation by using mixed-initiative interaction. If a product is selected, the agent finishes the inference and provides information about the target product.

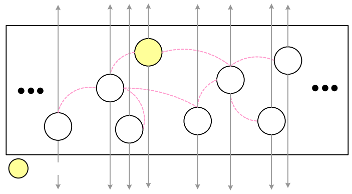
4.2. Quantitative analysis

4.2.1. Experimental designs

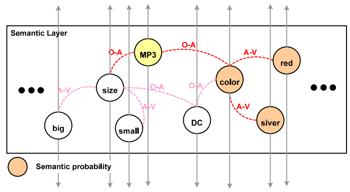
In order to evaluate the opinions of how satisfied younger adults are with the efficiency of the agent, we compared three conversational agents: script-based, BN-based and SeBN-based agents. The experiment aimed to estimate the speed and accuracy of the agent's responses. Thirty South Korean subjects aged from 22 to 33 evaluated the different kinds of agents. Table 6 shows the characteristics of these subjects. They had to perform ten tasks to search for information on several products, for example "find a small digital camera with a resolution rate of four million pixels." The users evaluate each system by posing questions constructed according to the QUIS (questionnaire for user interface satisfaction). Satisfaction scores were measured by single items on five-point Likert scales (1.0 = "not at all", 5.0 = "very much") for each task.

User: I'm looking for a MP3 player.

- Infer probabilistically using the SeBN.
- Search multiple objects.
- Acquire additional information using MII.
  a. Search the objects: MP3 players
  b. Search the O-A attributes: colors



<Probabilistic Inference>

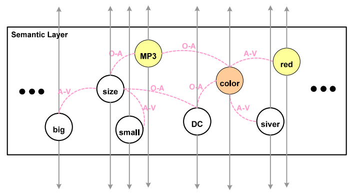


<Semantic Inference>

Agent: Which color would you like? Red or Silver?

User : I'd like red.

- Infer probabilistically using SeBN.
- Search the target product. → Inference completion
- Provide the information.



<Probabilistic Inference>

Agent: Here it is. (Samsung-Anycall SCH-S130).

The price is \383,000.

Dialogue 1. The target retrieval using MII.

Table 6
Comparative results in efficiency

Characteristics	Value	Proportion (%)
Occupation	Workers	16.7
	Graduate students	33.3
	Undergraduate students	50
Sex	Male	50
	Female	50
Experience on the internet	Less than 1 year	10
	More than 1 year	90
Experience with 'messenger' services	Yes	80
	No	20

Table 7
Comparative results in efficiency

Retrieval rate (PR)	Script		BN		SeBN	
Average interactions (AI)	PR (%)	AI	PR (%)	AI	PR (%)	AI
Average	87.51	3.53	92.15	3.18	94.42	2.96

Table 8
Comparative results of user satisfaction

User satisfaction	Script		BN		SeBN	
	Mean	SD	Mean	SD	Mean	SD
Easy	2.9	.7379	4.0	.4714	4.6	.5164
Friendly	2.7	.6749	3.8	.4216	4.7	.4830
Informative	3.1	.5676	3.7	.6749	4.4	.5164
Repetitive	3.9	.8756	2.3	.4830	1.6	.5164
Interesting	3.1	.5676	3.8	.7888	4.5	.5270

4.2.2. Analyses of results

The results (see Table 7) show that the proposed method ($M = 94.42$) is superior to the others ($M = 92.15$, 87.51). SeBN-based agents can manage various types of dialogues while script-based and BN-based agents fail to respond. SeBN-based agents also show good performance in providing suitable responses for users with only a few interactions ($M = 2.96$).

As shown in Table 8, satisfaction was very high when using the proposed method. The effects of the proposed method were evaluated in terms of the following criteria: ease of use, friendliness, informativeness, repetition and level of interest. These criteria were statistically measured by means of a one-way ANOVA with a variant of the SeBN as the among-systems factor. Post-hoc tests were also conducted, whenever one or more significant factor entailed more than two of the criteria. The emotional state measure revealed significant differences among the systems ($F(2,27) = 21.581$, $p < .05$). It showed that ease of use with SeBN-based agents was much higher ($M = 4.6$, SD = .5164) than with script-based agents ($M = 2.9$, SD = .7379) and BN-based agents ($M = 4.0$, SD = .4714). In terms of friendliness, the average score of the SeBN-based agents ($M = 4.7$, SD = .4830) was significantly higher than that of the script-based agents ($M = 2.7$, SD = .6749) and the BN-based agents ($M = 3.8$, SD = .4216), ($F(2,27) = 34.731$, $p < .05$). There was a significant difference ($F(2,27) = 12.160$, $p < .05$) in informativeness for script-based agents ($M = 3.1$, SD = .5676), BN –based agents ($M = 3.7$, SD = .6749) and SeBN-based agents ($M = 4.4$, SD = .5164) and the SeBN-based agents rated higher than all the other systems. The most noticeable result was in terms of repetition. The value of the SeBN-based agents ($M = 1.6$, SD = .5164) was significantly lower than that of the script-based agents ($M = 3.9$, SD = .8756) and the BN-based agents ($M = 2.3$, SD = .4830), ($F(2,27) = 32.921$, $p < .05$). In other words, the proposed method minimized unnecessary information in conversations. Finally, as for level of interest, SeBN-based agents ($M = 4.5$, SD = .5270) produced a significantly higher score than script-based agents ($M = 3.8$, SD = .7888) and BN-based agents ($M = 3.1$, SD = .5676), ($F(2,27) = 12.027$, $p < .05$).

5. Conclusions and future works

We have proposed a conversational agent that uses semantic Bayesian networks in order to be more flexible and considerable in terms of inferring intentions. If the information in the queries is insufficient, the agent asks the user to provide more information in order to infer the intention correctly. Finally, answering performance is improved when using SeBN-based agents. It is presumed that the design of networks will become easier and more comprehensible, since designers will be able to use more intuition.

The manual design of networks requires overhead operation, so research on the automatic construction of semantic Bayesian networks remains necessary, which improves the scalability of the proposed method. Several works on automatic learning of Bayesian networks (Yang & Chang, 2002) and semantic networks (Shamsfard & Barforoush, 2004) might be helpful for developing the learning technique for semantic Bayesian networks.

Acknowledgement

The work was supported by the Korea Research Foundation Grant (KRF-2004-005-H00005).

References

Acid, S., de Campos, L., Fernández-Luna, J., & Huete, J. (2003). An information retrieval model based on simple Bayesian networks. *International Journal of Intelligent Systems, 18*(2), 251–265.

Allen, J. (1999). Mixed initiative interaction. *IEEE Intelligent Systems*, 14–23.

Allen, J., Byron, D., Dzikovska, M., Ferguson, G., Galescu, L., & Stent, A. (2001). Towards conversational human-computer interaction. *AI Magazine, 22*(4), 27–38.

Brooks, R. (1986). A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation, 2*(1), 14–23.

Calado, P., da Silva, A., Laender, A., Ribeiro-Neto, B., & Vieira, R. (2004). A Bayesian network approach to searching Web databases through keyword-based queries. *Information Processing and Management, 40*, 773–790.

Garcia-Serrano, A., Martínez, P., & Hernández, J. (2004). Using AI techniques to support advanced interaction capabilities in a virtual assistant for e-commerce. *Expert Systems with Applications, 26*(3), 413–426.

Hix, D., & Hartson, H. (1993). *Developing user interface*. John Wiley & Sons.

Hong, J.-H., & Cho, S.-B. (2003). A two-stage Bayesian network for effective development of conversational agent. *LNCS* (vol. 2690, pp. 1–8).

Horvitz, E., Breese, J., Heckerman, D., Hovel, D., & Rommelse, K. (1998). The Lumiere project: Bayesian user modeling for inferring the goals and needs of software users. In *Proc. of the 14th conf. uncertainty in artificial intelligence* (pp. 256–265).

Jennings, N., & Wooldridge, M. (1995). Applying agent technology. *Applied Artificial Intelligence, 9*(4), 351–361.

Macskassy, S. (1996). A conversational agent. Master Essay, Rutgers University.

Maes, P. (1994). Agents that reduce work and information overload. *Communications of the ACM, 37*(7), 31–40.

Meng, H., Wai, C., & Pierracinni, R. (2003). The use of belief networks for mixed-initiative dialog modeling. *IEEE Transactions on Speech and Audio Processing, 11*(6), 757–773.

Nugues, P., Godéreaux, C., El Guedj, P. O., & Revolta, F. (1996). A conversational agent to navigate in virtual worlds. In Universiteit twente, Enschede, Luperfoy, S., Nijholt, A., Zanten, G. (Eds.), *Proc. of the 11th workshop on language technology* (pp. 23–33).

Pearl, J. (1998). *Probabilistic reasoning in intelligent systems: Networks of plausible inference*. Morgan Kaufmann Publishers.

Perugini, S., & Ramakrishnan, N. (2003). Personalizing web sites with mixed-initiative interaction. *IEEE IT Professional, 5*(2), 9–15.

Ribeiro-Neto, B., & Muntz, R. (1996). A belief network model for IR. In *Proc. of the 19th annual int. ACM SIGIR conf. on research and development in information retrieval* (pp. 253–260). ACM Press.

Shamsfard, M., & Barforoush, A. (2004). Learning ontologies from natural language texts. *International Journal of Human–Computer Studies, 60*(1), 17–63.

Symeonidis, A., Kehagias, D., & Mitkas, P. (2003). Intelligent policy recommendations on enterprise resource planning by the use of agent technology and data mining techniques. *Expert Systems with Applications, 25*(4), 589–602.

Tutle, H., & Croft, W. (1990). Inference networks for document retrieval. In *Proc. of the 13th int. conf. on research and development in information retrieval* (pp. 1–24).

Yang, S., & Chang, K. (2002). Comparison of score metrics for Bayesian network learning. *IEEE Transactions on Systems, Man and Cybernetics, Part A, 32*(3), 419–428.

Zue, V., & Class, J. (2000). Conversational interfaces: Advances and challenges. *Proc. of the IEEE, 88*(8), 1166–1180.