An Effective Conversational Agent with User Modeling based on Bayesian Network

Seung-Ik Lee, Chul Sung, and Sung-Bae Cho

Computer Science Department, Yonsei University, 134 Shinchon-dong, Sudaemoon-ku, Seoul 120-749, Korea
{cypher, sungc}@candy.yonsei.ac.kr, sbcho@csai.yonsei.ac.kr

Abstract. A conversational agent is an agent that interacts with users using natural language interface. Especially in Internet space, its role has been recently highlighted as a virtual representative of a web site. However, most conversational agents use simple pattern matching techniques for their answer without considering user goals and thus give user an unsatisfactory response. In this paper, we propose a conversational agent that utilizes user model constructed on Bayesian network for its responses consistent with user goals. The agent is applied to the active guide of a website, which shows that the user modeling based on Bayesian network helps respond to user queries appropriately and consistently with the user goals.

1 Introduction

As the Internet users are growing up rapidly, each site is requiring strong needs for interacting effectively with users. One of them is to provide a good tool for searching information on site. Until now, however, most Internet sites provide a simple keyword-based search engine, which is difficult to give the right information because of not considering the user goals.

As an alternative for the dumb interfaces of web sites, conversational agents are recently being developed because of it being able to have conversations with users by processing natural language. Eliza, one of the first conversational agents, was born at Massachusetts Institute of Technology in 1966. Eliza was contrived for the research on natural language processing. It uses a simple pattern matching technique [1]. ALICE (Artificial Linguistic Internet Computer Entity, http://www.alicebot.org) is written in a language called AIML (Artificial Intelligence Markup Language) that is based on XML. It enables other people besides the author to modify for their personalities. However, it has shortcomings of not being able to respond to users reflecting their intentions because of simple sequential pattern matching based on keywords. Tackling this problem requires much time and effort in constructing the response database.

One of the techniques for modeling user's intentions or goals is to use Bayesian network. It is an intuitive and parsimonious representation of probability distributions and effective in diagnosing user's needs and thus provides useful enhancements to legacy software applications when embedded within them. Pynadath [2] used probabilistic models for making inferences about the goals of car drivers in navigation. Albrecht [3] applied Bayesian models in action prediction in a multi-user computer game.

Horvitz [4] worked to build models for inference and decision-making under uncertainty about user's goals at each level of the task hierarchy with Bayesian networks.

In this paper, we propose a conversational agent that can have a more intelligent conversation by inferring diverse user goals through user modeling based on Bayesian network and apply it to the introduction of a website to show the usability and possibility.

2 Bayesian Network

A Bayesian network is a graphical specification of a probability distribution. Each Bayesian network is comprised of two distinct parts: A directed acyclic graph (DAG) and a set of conditional probability tables (CPTs). DAG nodes represent variables of interest and are typically discrete. DAG edges have both a formal and an informal meaning. Informally, edges represent direct causal influences between variables. The formal meaning of edges is stated in terms of conditional independence. In particular, a Bayesian network specifies that each variable is independent of its non-descendants given its parents.

Fig. 1 shows a simple Bayesian network. Variable E has only one parent, C, and its non-descendants are A, B and D. Therefore, the probability distribution of variable E is independent of variables A, B and D, once the state of variable C is known. In general, the chain rule of probability theory allows us to write

$$\begin{aligned} P(E,D,C,B,A) =& P(E|D,C,B,A) \cdot P(D|C,B,A) \cdot \\ & P(C|B,A) \cdot P(B|A) \cdot P(A) \end{aligned} \tag{1}$$

and the conditional independences specified by the Bayesian network structure allow us to simplify the above terms as follows [5].

$$P(E,D,C,B,A) = P(E|C) \cdot P(D|C,B) \cdot P(C|A) \cdot P(B|A) \cdot P(A) \tag{2}$$

If each variable X_i is independent of its non-descendants given its parents in the graph G, we denote this as:

$$\forall_i, I(X_i;\ \mathrm{NonDescendants}(X_i)|Pa(X_i)) \tag{3}$$

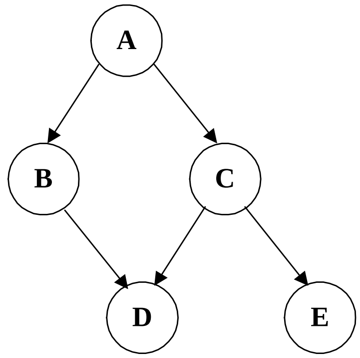
where $Pa(X_i)$ is the set of parents of X_i in G, and NonDescendants(X_i) are the non-descendents of X_i in G. By applying the chain rule of probabilities and properties of conditional independencies, any joint distribution that satisfies equation (3) can be decomposed in the product form [6].

$$P(X_1,\ldots,X_n) = \prod_{i=1}^{n} P(X_i|Pa(X_i)) \tag{4}$$

3 Conversational Agent

3.1 Overall Structure

Fig. 2 shows the overall structure of a conversational agent. After preprocessing user's queries to correct typos and supplant synonyms, *Goal Inference* module infers user's



$P(T_A | A{=}true) = 0.60 \quad P(T_A | A{=}false) = 0.40$

$P(T_B | A{=}true,B{=}true) = 0.20 \quad P(T_B | A{=}true,B{=}false) = 0.80$
$P(T_B | A{=}false,B{=}true) = 0.75 \quad P(T_B | A{=}false,B{=}false) = 0.25$

$P(T_C | A{=}true,C{=}true) = 0.80 \quad P(T_C | A{=}true,C{=}false) = 0.20$
$P(T_C | A{=}false,C{=}true) = 0.10 \quad P(T_C | A{=}false,C{=}false) = 0.90$

$P(T_D | B{=}true,C{=}true,D{=}true) = 0.95 \quad P(T_D | B{=}true,C{=}true,D{=}false) = 0.05$
$P(T_D | B{=}true,C{=}false,D{=}true) = 0.90 \quad P(T_D | B{=}true,C{=}false,D{=}false) = 0.10$
$P(T_D | B{=}false,C{=}true,D{=}true) = 0.80 \quad P(T_D | B{=}false,C{=}true,D{=}false) = 0.20$
$P(T_D | B{=}false,C{=}false,D{=}true) = 0.00 \quad P(T_D | B{=}false,C{=}false,D{=}false) = 1.00$

$P(T_E | C{=}true,E{=}true) = 0.70 \quad P(T_E | C{=}true,E{=}false) = 0.30$
$P(T_E | C{=}false,E{=}true) = 0.00 \quad P(T_E | C{=}false,E{=}false) = 1.00$

Fig. 1. Two parts of Bayesian network, a DAT and CPTs

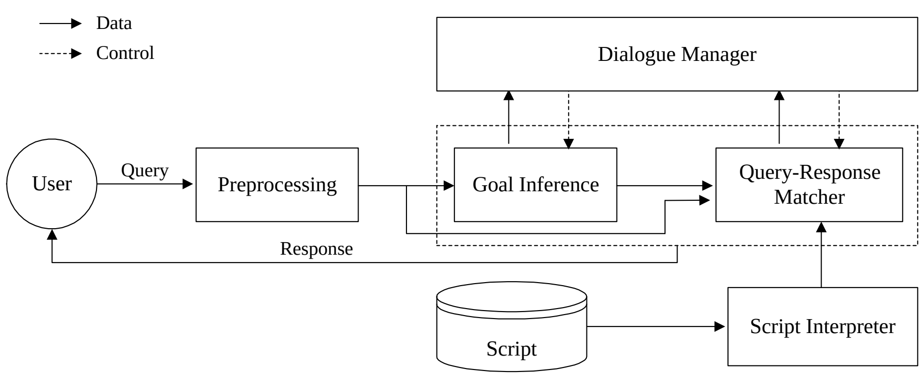


Fig. 2. The architecture of a conversational agent

goals. When this module cannot infer user's goals because of insufficient information, the Dialogue Manager requests the user to give more information. Once the goal is inferred, the agent presents an appropriate answer by searching the knowledge base or script. This requires the *Goal Inference* play an important role in finding the appropriate answer by reasoning goals during the mixed-initiative interactions with the user [7].

3.2 User Model based on Bayesian Network

We construct a user model for the introduction of a specific site with Bayesian network. Fig. 3 shows the structure of the model. The evidence variables of a node or a goal are represented as parent nodes. For each of the goal variable, evidence variables are defined as in Table 1. Words in curly braces are synonyms with the one in front of the braces. When the information from a query does not suffice for the transition from the

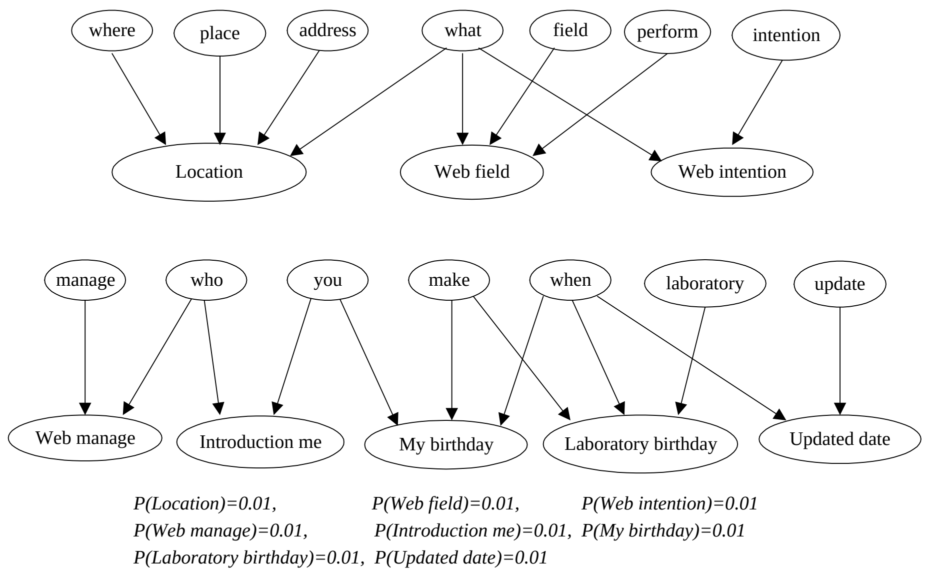


Fig. 3. Bayesian network architecture for modeling user goals

current goal to a more detailed one, the agent requests more information to the user. For example, when 'Location' is chosen as a goal as shown in Fig. 4, the agent asks the user to give more information on the four places, 'Yonsei univ.', 'Engineering building', 'Professor room', and 'Laboratory'. When the user gives more information, the agent infers the final goal.

3.3 Knowledge Representation

In order to response to a query, a script database should be constructed prior to the conversation. The script consists of a list of query (or condition) and response pairs. A part

Table 1. Goals and their evidences

Goal variable	Evidence variable
Location	where, place{position, location}, what, address
Web field	field{type, category, nature, style, range, variety, ground, personality, kind}, perform
Web intention	what, intention{goal, object, aim, target, purpose}
Web manage	who, manage{supervise, direct, lead, control, handle}
Introduce me	who, you
My birthday	when, you, make
Laboratory birthday	when, make, laboratory
Update date	when, update

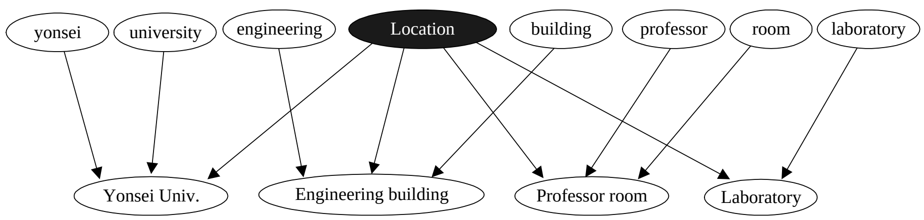


Fig. 4. More information is required to infer the final goal.

of the script grammar is illustrated in List 1 using BNF notation. Topic is the primary component that represents a query-response pair in the grammar. A topic begins with TOPIC keyword followed by name and one or more conditional statements followed by ENDTOPIC keyword. A conditional statement is represented as IF (condition) THEN (action). A condition is a Boolean expression composed of the operand and their operators, AND and OR. An operand is evaluated to true or not true and includes goal information and the results of comparison functions. List 2 shows an example script. When a user asks the location or direction of something and “lab#,” “softcomputing,” or “soft” & “computing” appear in the query, one of the items below the “SAYONEOF” is randomly selected and presented as a response to the user.

3.4 Query-Response Matching

Scripts that are composed of according to the grammar in List 1 are interpreted and loaded into memory by script interpreter when the agent start running. The conditional part in each topic is transformed into a Boolean expression and the keywords listed in the conditional part are transformed into a regular expression as in Fig. 5. Fig. 6 show the overall procedure of matching a query with a response. For all the topics, the conditional part of a topic is compared with the goal information, the query, and the keyword list extracted during the goal inference. This returns scores of all the topics as a result. Different types of matching components, like goals, keywords, or Boolean operators, are assigned different scores as in Table 2. When all the topics are assigned

List 1 A part of BNF grammar for the query-response script database

```
<script>::=<topic_decl>|<pattern_decl>|<attribute_decl>|
           <script> <topic_decl>|<script> <pattern_decl>|
           <script> <attribute_decl>
<topic_decl>::=TOPIC QSTRING <cond_stmt_list> ENDTOPIC
<cond_stmt_list>::=<cond_stmt>|<cond_stmt_list> <cond_stmt>
<action_list>::=<action>|<action_list> <action>
<cond_stmt>::=<if_cond> <action_list> <continuation>
<continuation>::=DONE|CONTINUE|NEXTTOPIC
<action>::=<say>|<say_one_of>
<say>::=SAY <concat_string>
<say_one_of>::=SAYONEOF <items>|SAYONEOF <concat_string>
               <items>::=<item>|<items> <item>
<item>::=ITEM <concat_string>
<if_cond>::=IF <expr> THEN
<expr>::=<expr> OR <expr>|<expr> AND <expr>|NOT <expr>|
         '(' <expr> ')'|MEMORY|MEMORY MATCH <concat_string>|
         MEMORY CONTAIN <concat_string>|HEARD <concat_string>|
         MATCH <concat_string>|CONTAIN <concat_string>|ALWAYS
<concat_string>::=<concat_string> COMMA <concat_string>|
                  <concat_string> '+' <concat_string>|
                  <concat_string> '&' <concat_string>|
                  <s_string>
<s_string>::=QSTRING|PNAME|MEMORY| '(' <concat_string> ')'
```

List 2 A part of script

```
TOPIC "location of lab"
   IF ((?LOCATIONQUESTION  OR ?DIRECTIONSQUESTION)
      AND HEARD ("lab#", "softcomputing", "soft"&
                      "computing")))
   THEN
      SAYONEOF
         ITEM "It is located at the 3rd engineering building
                in yonsei university"
         ITEM "529, the 3rd engineering building, 134, yonsei
               university, shinchon-dong, seodaemoon-gu, seoul"
         ITEM "The 3rd engineering building in yonsei
                university"
   DONE
ENDTOPIC
```

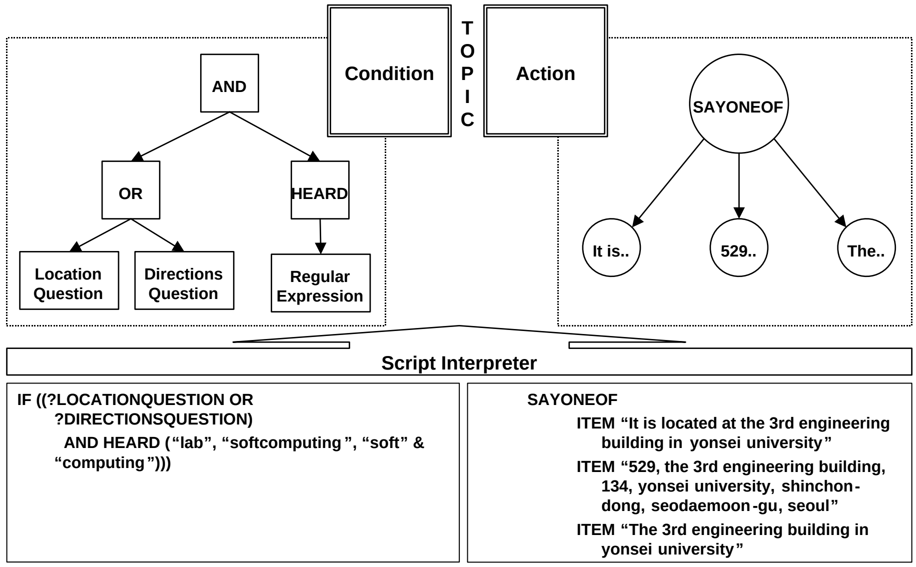


Fig. 5. Internal representation of the script of List 2

scores, the highest topic is selected to be presented to a user as a response to his query.

Table 2. Score assignment

Target	**Score**
Goal	2
Keyword	Number of characters
AND operator	Sum of two operands's scores
OR operator	Max of two operands's scores
NOT operator	0 : if operand is true, else the score

4 Simulation Results

The conversational agent has been applied to the introduction of a website to show the usability and possibility. Following is an actual conversation with a user.

```
User : What is your address?
Agent: Which place out of Yonsei univ., Engineering
       building, Professor room, or Laboratory do you want?
User : Laboratory.
```

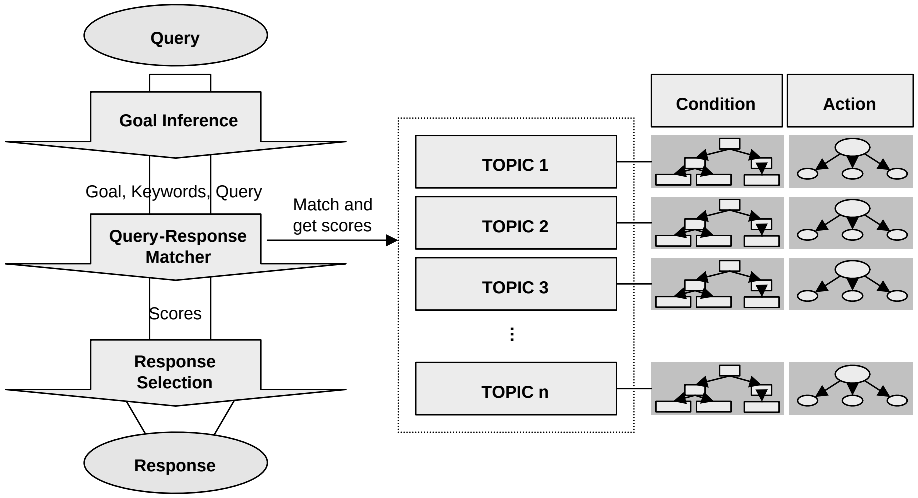


Fig. 6. Overall response matching procedure

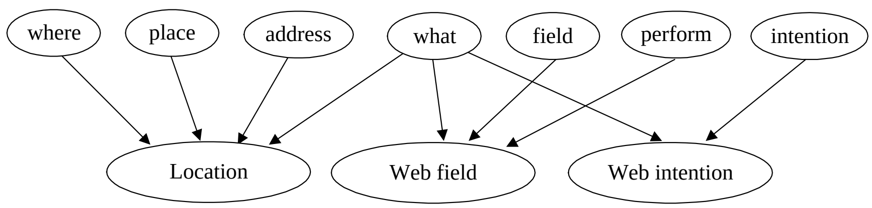
```
Agent: It is located at the 3rd engineering building in
       Yonsei university.
User : Tell me where Yonsei university is.
Agent: Yonsei university is 134 Shinchon-dong, Sudaemun-ku,
       Seoul 120-749, Korea.
```

When a user asks the agent “What is your address?,” the agent searches synonyms to be replaced with their representative words. In this case, no words are replaced. Then the agent compares some words of the query with the variables of the Bayesian network. As this query contains ‘what’ and ‘address’ as evidences in part or whole of ‘Location’, ‘Web field’, and ‘Web intention’, their probabilities are looked up from the predefined conditional probability tables as in Fig. 7.

Since the probability of ‘Location’ is over the threshold probability 0.8 and greater than the other goals, the agent can infer that user goal is ‘Location’. At this moment, the agent needs more information which of the four detailed goals the user actually wants to know. The agent asks the user “Which place out of Yonsei univ., Engineering building, Professor room, or Laboratory?” The answer, “Laboratory,” makes the ‘Laboratory’ variable true and thus the agent infers that user goal is ‘Laboratory’ as shown in Fig. 8. With this inferred user goal, the script is searched to produce the response “It is located at the 3rd building in Yonsei university.”

When a user asks the agent “Tell me where Yonsei university is,” the agent begins to compare with matching words. The agent takes ‘where’ word which is evidence variable of user modeling. $P(Location|where)$ is 0.85 as shown in Fig. 7, because ‘Location’ is conditionally independent given ‘where’. After inferring user goal, the agent gathers more information for the next sub goal. The agent takes ‘yonsei’ and ‘university’ words in query. As shown in Fig. 9 the agent infers ‘Yonsei univ.’ detailed goal with ‘yonsei’

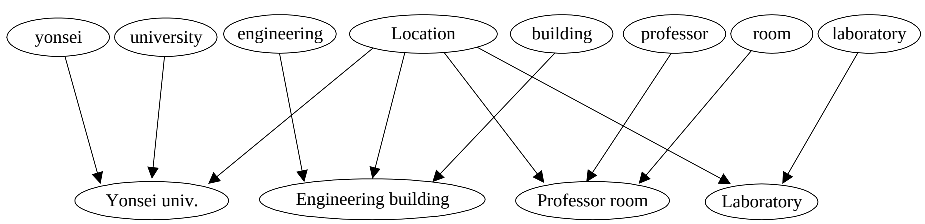
P(where=true) = 0.01 P(place=true) = 0.01 P(address=true) = 0.01 P(what=true) = 0.01
P(field=true) = 0.01 P(perform=true) = 0.01 P(intention=true) = 0.01



P(Location=true | where=false,place=false,address=true,what=true) = 0.83
P(Location=true | where=true,place=false,address=false,what=false) = 0.85
P(Web field=true | what=true,field=false,perform=false) = 0.38
P(Web intention=true | what=true,intention=false) = 0.36

Fig. 7. Bayesian network structure for presenting user goal 'Location'

P(laboratory=true) = 0.31
P(Location=true | where=false, place=false,address=true,what=true) = 0.83



P(Yonsei univ.=true | Location=true, laboratory=true) = 0.95

Fig. 8. 'Laboratory selection'

and ‘university’ words. After inferring ‘Yonsei univ.’, the agent finally answers a user “Yonsei university is 134 Shinchon-dong, Sudaemun-ku, Seoul 120-749, Korea.”

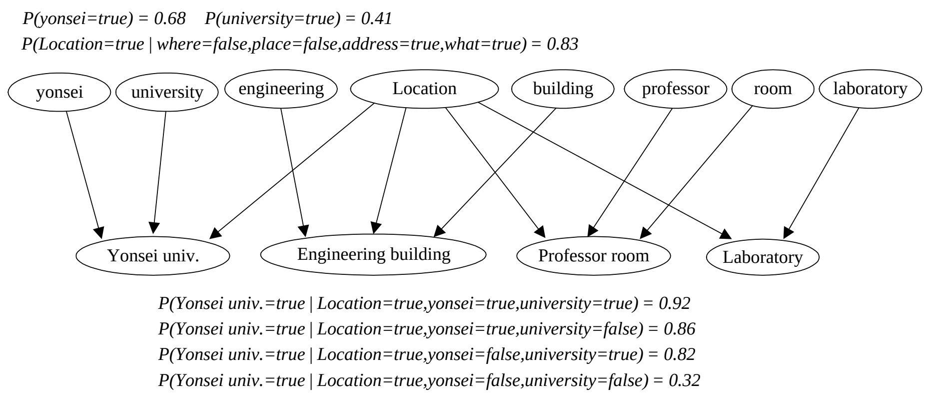


Fig. 9. Inferring detailed user goals

5 Conclusion

In this paper, we have described a conversational agent that can give responses more consistent with the user goals in a specific domain. A Bayesian network is used to construct the user model in the domain. Although we have constructed the user model of limited goals in a specific domain, the results show that this has the possibility of interacting with users more consistently with the goals. As a further research, we plan to construct user modeling on the independent domain with Bayesian network.

References

1. J. Weizenbaun: ELIZA - a Computer Program for the Study of Natural Language Communication between Man and Machine. *Communications of the ACM 9(1)*. (1965) 36–45
2. D. Pynadath and M. Wellman: Accounting for Context in Plan Recognition with Application to Traffic Monitoring. *Proc. of the Eleventh Conf. on Uncertainty in Artificial Intelligence*. (1995) 472–481
3. D. Albrecht *et al*: Towards a Bayesian Model for Keyhole Plan Recognition in Large Domains. *Proc. of the Sixth Int. Conf. on User Modeling.* (1997) 365–376.
4. E. Horvitz and T. Paek: A Computational Architecture for Conversation. *Proc. of the Seventh Int. Conf. on User Modeling.* (1999) 201–210
5. F.V. Jensen: An Introduction to Bayesian Networks, *Springer-Verlag*. (1996)
6. N. Friedman *et al*: Using Bayesian Networks to Analyze Expression Data. *Proc. of the Fourth Annual Int. Conf. on Computational Molecular Biology*. (2000) 127–135
7. J.F. Allen: Mixed-Initiative Interaction. *IEEE Intelligent Systems 5* (1999) 14–16