ISCA Archive
http://www.isca-speech.org/archive

6th International Conference on Spoken Language Processing (ICSLP 2000)
Beijing, China
October 16-20, 2000

The Use of Belief Networks for Mixed-Initiative Dialog Modeling

Helen M. Meng[1], Carmen Wai[1] and Roberto Pieraccini[2]
[1]Human-Computer Communications Laboratory
The Chinese University of Hong Kong
[2]SpeechWorks International Ltd.
{hmmeng, cmwai}@se.cuhk.edu.hk, roberto.pieraccini@speechworks.com

ABSTRACT

This paper describes the use of Belief Networks for mixed-initiative dialog modeling within the context of the CU FOREX system [1]. CU FOREX is a bilingual hotline for real-time foreign exchange inquiries. Presently, it supports two separate interaction modalities: a direct dialog (DD) interaction, which is system-initiated for novice users; as well as natural language (NLS) shortcuts, which is user-initiated for expert users. In this work, we propose to use Belief Networks (BNs) to automatically govern the model transitions for mixed-initiative interactions. Furthermore, we hope that our approach can reduce the amount of handcrafting involved in the development of current mixed-initiative dialog models, to ease portability across different applications domains.

1. INTRODUCTION

Spoken dialog systems have demonstrated a high degree of usability in many restricted domains. Among the various dialog strategies, the *mixed-initiative* model is deemed most desirable, since *both* the user and the system can influence the dialog flow over the course of interaction. This provides greater flexibility than the *system-initiated* model or *user-initiated* model. It is possible to build effective mixed-initiative interactions by *handcrafting* flexible transitions between the system- and user-initiated models. While handcrafting can produce a sophisticated dialog flow, the task is expensive, and may become intractable with increasingly complex domains. In this work, we propose to use Belief Networks to *automatically* govern the model transitions for mixed-initiative interactions. We have extended our Belief Network framework, previously used for natural language understanding [2,3], for the current task of mixed-initiative dialog modeling within the context of the CU FOREX system.

2. THE CU FOREX DOMAIN

2.1 Domain-Specific Constraints

We have chosen to investigate the feasibility of using BNs for dialog modeling, within the context of the CU FOREX system [1]. This is a spoken language system for the foreign exchange domain. It supports inquiries regarding the exchange rates between a currency pair, as well as the interest rates for various time durations for a currency. Hence the key domain-specific concepts include: CURRENCY, TIME DURATION, EXCHANGE RATE and INTEREST RATE. For database retrieval, there are two constraints: An inquiry regarding Exchange Rates mandates that the currencies to be bought and sold be specified. An inquiry regarding Interest Rates mandates that a currency and a time duration be specified.

2.2 Two Interaction Modalities

The CU FOREX has been made available for experimentation by the public since August 1999. Since then, we have received a few hundred calls per month on average. It currently supports two interaction modalities, based on SpeechWorks 4.0 [1]. The Directed Dialog (DD) is designed for novice users, which provides guidance for them to give a specific attribute at each dialog turn. The Natural Language (NLS) shortcut is designed for expert users, and allows the user to specify several information attributes within a single dialog turn. Example dialogs are shown in Table 1 and 2.

System:	Welcome to CU FOREX. Which language would you prefer, Cantonese or English?
User:	English
System:	Would you like 'exchange rates' or 'interest rates'?
User:	Interest rates
System:	Which currency are you interested in?
User:	Hong Kong dollar
System:	Would you like deposit durations for twenty-four hours, one month…. or one year?
User:	One month
System:	The quotes we provide are for reference only. The interest rate of the Hong Kong dollar for one month is XXX

Table 1. Directed dialog (DD) interaction between the system and the user, for an interest rate inquiry. Notice that at every dialog turn, the system guides the user to provide a specified information attribute.

System:	Welcome to CU FOREX. Which language would you prefer, Cantonese or English?
User:	English
System:	What kind of currency information are you interested in?
User:	I'd like to know the exchange rate between the US dollar and the HK dollar please.
System:	The quotes we provide are for reference only. Exchange rate. US Dollar to HK Dollar. The buying rate is XXX. The selling rate is XXX.

Table 2. Natural language (NLS) interaction between the system and the user, for an exchange rate inquiry. Notice that the user is able to specify *several* information attributes within a single dialog turn.

3. THE BELIEF NETWORK

The Belief network (BN) is a probabilistic causal network, and for our implementation, we have pre-defined a topology which is similar to the one in Figure 1 (without the dotted arrow). The black arrows indicate the causal relationships between the goal and the concepts. The predefined topology assumes that the concepts are independent with each other.

In this work, we adopt the use of Belief Networks in mixed-initiative dialog modeling. It involves the processes of (i) inferring the informational goal of a user's query, as well as (ii) verifying the input query against domain-specific constraints. To identify the user's informational goal, we have previously devised a framework that utilizes Belief Networks. Details can be found in [2,3]. A brief description is provided below.

3.1 Goal Identification

A BN is trained for each domain-specific informational goal. In this domain, there are two informational goals – Exchange Rates and Interest Rates. Hence we developed two BNs, one for each goal, using the NLS queries we have collected. Each BN receives as input all of the five domain-specific concepts:

CURRENCY1, CURRENCY2, DURATION, EXCHANGE_RATE AND INTEREST_RATE.[1] We have also enhanced the pre-defined topology by means of automatic learning using the Minimum Description Length (MDL) principle [3]. The resulting topology is illustrated in Figure 1. Notice that it captures not only the causal dependencies between the information goal and the corresponding concepts, it also shows the relation between the concepts (i.e. the dotted arrow).

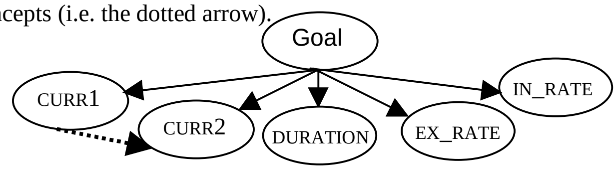

Figure 1. The predefined topology of our BNs is enhanced by the linkage (dotted arrow) learnt to capture dependencies among concepts. The arrows of the acyclic graph are drawn from cause to effect.

A trained BN is then used to make a binary decision based on the concepts present in the input query, regarding the presence or absence of the goal G_i. According to the topology shown in Figure 1, the network is divided into several sub-networks: {Goal, CURR1, CURR2}, {Goal, DURATION}, {Goal, EX_RATE} and {Goal, IN_RATE}. Then the updated joint probabilities are iteratively computed according to the Equation (1) by each sub-network, and the *aposteriori* probability $P^*(G_i)$ is computed by the marginalization of the updated joint probability $P^*(G_i,\mathbf{C})$. $P^*(G_i)$ is then compared to a threshold (θ) to make the binary decision – θ may be set to 0.5,[2] or an optimized value for each goal.

$$P^*(G_i,\vec{C}) = P(G_i \mid \vec{C})P^*(\vec{C}) \rightarrow P^*(G_i,\vec{C}) = \frac{P(G_i,\vec{C})}{P(\vec{C})}P^*(\vec{C}) \quad (1)$$

where $P^*(\mathbf{C})$ is instantiated according the presence or absence of the concepts; $P(G_i,\mathbf{C})$ is the joint probability obtained from training and $P^*(G_i,\mathbf{C})$ is the updated joint probability.

The decisions across all the BNs are combined to identify the output goal of an input query. We may label the query to a goal if the corresponding BN votes positive with the highest *aposteriori* probability. Alternatively, we may label the query with all the goals for which the BNs votes positive. Should all BNs vote negative, the input query is rejected as out-of-domain (OOD). Typical values of aposteriori probabilities that are obtained from goal inference are shown in Table 3. These values are compared with a theshold of θ = 0.5 for making the binary decision.

Query: "Can I have the exchange rate of the yen please?" BN *for Exchange Rates:* P(goal =Exchange Rate | Query) = 0.823 ➔ goal present BN for *Interest Rates:* P(goal =Interest Rate | Query) = 0.256 ➔ goal absent **Hence, the inferred goal is Exchange Rates.**
Query: "Tell me about stock quotes" BN for *Exchange Rates:* P(goal =Exchange Rate | Query) = 0.14 ➔ goal absent BN for *Interest Rates:* P(goal =Interest Rate | Query) = 0.13 ➔ goal absent **Hence, the inference result is OOD**

Table 3. Typical values of aposteriori probabilities obtained from goal inference using BNs in the CU FOREX domain.

3.2 Backward Inference

Having inferred the informational goal of the query, the corresponding node (goal node) is instantiated, and we perform a backward inference to test the networks' confidence in each input concept. When the goal node is instantiated, the joint probability of $P(\mathbf{C}, G_i)$ will be updated for each sub-network by the formula below (Equation 2, which is similar to Equation 1):

$$P^*(\vec{C},G_i) = P(\vec{C} \mid G_i)P^*(G_i) \quad (2)$$

where $P^*(G)$ is updated and instantiated to 1, $P(\mathbf{C}|G_i)$ is the conditional probability obtained from training and $P^*(\mathbf{C},G_i)$ is the updated joint probability.

By marginalization, we can get $P(C_j)$. Again, the threshold θ = 0.5 is used to determine whether the concept should be present or absent.

$$P(C_j)\begin{cases} >=\theta \rightarrow C_j \text{ should be } \textit{present} \text{ in the given } Gi \text{ query} \\ < \theta \rightarrow C_j \text{ should be } \textit{absent} \text{ in the given } Gi \text{ query} \end{cases}$$

Backward inference *verifies* the validity of the input query against domain-specific constraints. In this way, we can test for cases of *spurious* and *missing* concepts,[3] and generate the appropriate systems response.

As an example, consider an interest rates query *"Can I have the interest rates of the yen for one month please?"*. We instantiated the goal node of the BN (for Interest Rates) to 1, and perform backward inference for each input concept. The associated probabilities and binary decisions are shown in Table 4. We see that the binary decision for each concept *agrees* with their actual occurrence. The semantic frame is thus ready to be processed for database retrieval.

Concept$_j$ (C_j)	***P(C_j)***	**Binary Decision for C_i**	**Actual Occurrence for C_i**
CURRENCY1	0.91	present	present
CURRENCY2	0.0058	absent	absent
DURATION	0.77	present	present
EXCHANGE_RATE	0.011	absent	absent
INTEREST_RATE	0.867	present	present

Table 4. Backward inference to test the BN's confidence for each input concept. The corresponding binary decisions obtained for each concept (using θ=0.5) agrees with the actual occurrences in the input query.

However, in situations where the binary decision for each concept *disagree* with its actual occurrence, further processing is necessary. The following shows two cases:

Case 1. Missing concepts

If the binary decision for *concept$_j$* is positive but it is absent in the input query, a missing concept is detected. The dialog model is designed such that the system will prompt for the missing concept. Table 5 illustrates the associated probabilities and binary decisions for an interest rates query *"Can I have the interest rate of the yen?"*. When we compare the results from backward inference with the actual occurrences in the input query, we detect that the concept <DURATION> is missing.

Concept$_j$ (C_j)	***P(C_j)***	**Binary Decision for C_i**	**Actual Occurrence for C_i**
CURRENCY1	0.91	present	Present
CURRENCY2	0.058	absent	Absent
DURATION	**0.77**	**present**	**Absent**
EXCHANGE_RATE	0.011	absent	Absent
INTEREST_RATE	0.867	present	Present

Table 5. Aposteriori probabilities obtained from backward inference. The actual occurrences of the concepts in the input query are indicated as well.

[1] Since our domain is relatively simple, we did not select the input concepts for the BN using Information Gain [2].

[2] We choose threshold at 0.5 since $P(G=1|\mathbf{C})+P(G=0|\mathbf{C})=1$.

[3] These may be due to speech recognition errors in an integrated spoken dialog system.

Case 2. Spurious concepts

Should a spurious concept be detected, i.e. the presence of *concept*$_j$ violates the binary decision, the system would automatically ask the user for clarification. For example, consider the query *"Can I have the interest rate of the lira against the yen"*, the inferred goal is Exchange Rates and the corresponding probabilities and binary decisions are shown in Table 6. It can be seen that the concept <INTEREST_RATE> is spurious. The BN for exchange rates indicates that this concept should not be present. Hence, the system will follow the inferred goal (Exchange Rates) to generate the clarification response: *"Are you referring to the exchange rate between the lira and the yen?"*.

***Concept**$_j$ (C_j)*	***$P(C_i)$***	**Binary Decision for C_i**	**Actual Occurrence for C_i**
CURRENCY1	0.91	Present	Present
CURRENCY2	0.92	Present	Present
DURATION	0.017	Absent	Absent
EXCHANGE_RATE	0.84	Present	Absent
INTEREST_RATE	**0.023**	**Absent**	**Present**

Table 6. Aposteriori probabilities obtained from backward inference. The actual occurrences of the concepts in the input query are indicated as well.

4. RESULTS

We have developed a dialog system using Belief Networks. The example dialogues in Table 7 show that the BN-based dialog model can handle pure DD or NLS as well as mixed-initiative interactions. It can prompt for missing concepts, ask for clarification from the user when spurious concepts are detected, and reject OOD input.

Dialogs 1 and 2 in Table 7 show that the new model can produce pure DD or NLS interactions identical to the previous dialog models illustrated in Tables 1 and 2. In the system-initiated DD interaction the BN-based dialog model detects the missing concepts and prompts the user until all concept occurrences agree with the domain-specific constraints. In the user-initiated NLS interaction the concepts agree with the domain-specific constraints and hence the system is able to produce a response directly.

Furthermore, our new BN-based dialog model can freely transition between the DD and NLS models. Users are no longer constrained to inputting one information attribute per dialog turn. In addition, the system is capable of prompting for missing information as shown in Dialog 3, or clarifying confusion due to spurious concepts in Dialog 4.

The system is also able to provide a mixed-initiative interaction with either novice or expert users on a single spoken interface, as well as reject OOD input automatically as shown in Dialog 5 and 6.

Dialog 1. Pure DD, asking for exchange rate information.	
System	What kind of currency information are you interested in?
User	Exchange rates
System	Which currency would you like to sell?
User	Hong Kong dollar
System	Which currency would you like to buy?
User	US dollar
System	The exchange rate of the Hong Kong dollar against the US dollar is XXX.

Dialog 2. Pure NLS interaction, asking for interest rate information.	
System	What kind of currency information are you interested in?
User	I would like the interest rate of the Yen for one year.
System	The interest rate of the Yen for one year is XXX.
Dialog 3. Prompting for missing information.	
System	What kind of currency information are you interested in?
User	I would like the interest rate of the Yen. *(NLS interaction)*
System	How long would you like to deposit?
User	Two months *(DD interaction)*
System	The interest rate of the Yen for two months is XXX.
Dialog 4. Asking for clarification upon the detection of spurious concepts.	
System	What kind of currency information are you interested in?
User	**Interest rate** of the Hong Kong dollar against the Yen
System	Are you referring to the **exchange rate** of the Hong Kong dollar against the Yen?
User	Yes please
System	The exchange rate of the Hong Kong dollar against the Yen is XXX.
Dialog 5. Handling mixed-initiative interaction.	
System	What kind of currency information are you interested in?
User	Exchange rate *(DD interaction)*
System	Which currency would you like to sell?
User	Hong Kong dollar against the US dollar *(NLS interaction)*
System	The exchange rate of Hong Kong dollar against the US dollar is XXX.
Dialog 6. Handling OOD query.	
System	What kind of currency information are you interested in?
User	Tell me about stock quotes. *(OOD query)*
System	Sorry we don't have the information you want. Please input service, exchange rates or interest rates.

Table 7. Example dialogs produced by the BN-based dialog model.

5. EVALUATION

Our evaluation is based on 550 dialog sessions collected using the CU FOREX system during the period between November and December 1999. Approximately 17% were rejected manually as the users were clearly attempting to break the system. Of the remaining queries, 285 calls were obtained from the DD while 170 calls were obtained from NLS hotlines. The task completion rates of the DD and NLS models were shown in Table 8. Failures in the DD model are mainly caused by queries with multiple information attributes or OOD input. Failures in the NLS model are due to missing concepts, spurious concepts or OOD queries. Detail statistics are tabulated in Table 9. In comparison, our BN-based mixed-initiative dialog model can automatically reject OOD input, and successfully handle *all* the dialogs. Table 10 shows the typical causes of failure in the original DD and NLS models respectively.

	Task completion rate	
	CU FOREX	**BN Dialog model**
DD (total: 285 calls)	85%	100%
NL (total: 170 calls)	63%	100%

Table 8. Task complete rates of the CU FOREX and BN dialog model.

	Failures for the CU FOREX dialog model			
	Multiple attributes	**Missing concepts**	**Spurious concepts**	**OOD input**
DD	11%	---	---	4%
NL	---	30%	4%	3%

Table 9. Causes of failure for the CU FOREX dialog model. Percentages refer to the proportion of the evaluated queries.

Dialog 1: DD interaction	
System	Would you like 'exchange rates' or 'interest rates'?
User	Exchange rates
System	Which currency you would like to sell?
User	Yen to Hong Kong please.
System (CU FOREX)	*Failed (Due to multiple information attributes in a single dialog turn)*
System (BN-based Dialog System)	The exchange rate between the Yen and the Hong Kong dollar is XXX. *(Concepts pass the domain constraints)*
Dialog 2: NLS interaction	
System	What kind of currency information are you interested in?
User	Tell me about interest rates.
System (CU FOREX)	*Failed (Due to the absent of concepts: <CURRENCY1> and <DURATION>)*
System (BN-Based Dialog System)	Please tell me the currency you are interested in. *(Detects the missing concepts and prompts the user for it)*

Table 10. Comparison of the DD/NLS interactions between the original CU FOREX system and the BN-based dialog model.

6. EXTENSIBILITY AND PORTABILITY

Our BN model is also extensible for integration with a simulated speaker verification (SV) component. We have implemented it in the CU FOREX domain, in which the SV component will randomly output a value ranged from 0 to 1 for each user's dialog turn. By using threshold of 0.6, a binary decision for speaker verification is determined. This binary decision can drive the dialog flow.

Also we have tried to test our framework on another domain – ATIS (Air Travel Information System), using Class A (context independent) and Class D (context dependent) queries. 11 BNs were trained for each goal in the ATIS domain, using Class A queries from training set. Then the goal for each testing query (Class A or Class D) was inferred and domain constraints are verified by using the BNs. Our experiments show that in a complex domain with more semantic keys, some concepts are sparsely trained, so binary decisions to determine whether concepts should be *present* or *absent* in the query are no longer accurate. To improve performance we defined two thresholds for determining whether in a given G_i query a concept is *mandatory*, *optional* or *forbidden*.

$$P(C_j)\begin{cases} >=\theta_{upper} \rightarrow C_j \text{ should be } \textit{present} \text{ in the given } Gi \text{ query} \\ <\theta_{upper} \textit{ and } >=\theta_{lower} \rightarrow C_j \text{ is } \textit{optional} \text{ in the given } Gi \text{ query} \\ <\theta_{lower} \rightarrow C_j \text{ should be } \textit{absent} \text{ in the given } Gi \text{ query} \end{cases}$$

In our experiments, we used 0.7 and 0.2 as the upper and lower thresholds. Also, in order to handle Class D (context dependent) queries, we enhanced our BN-based dialog model with the capability of context inheritance. Example dialogs are illustrated in Table 12. Table 13 shows the binary decisions for the corresponding Class D FARE_ID query.

System	What kind of flight information you are interested in?
User	Please list all the flights from Chicago to Kansas city on June seventeenth. (Class A query)
System	*Goal Inference: Flight ID (Concepts pass the domain constraints)*
User	For this flight how much would a first class fare cost. (Class D)
System	*Goal Inference: Fare ID. (Though the system detects the missing concepts <CITY_NAME1> <CITY_NAME2>, it automatically retrieves these concepts from the dicourse context.)*

Table 12. Examples of ATIS dialogs produced by the BN-based dialog model with the capability of context inheritance.

Concept$_j$ (C_j) (Part of concepts)	***P(C_j)***	**Binary Decision for C_i**	**Actual Occurrence for C_i**
AIRPORT	0.0000	absent	absent
CITY_NAME1	**0.9629**	**present**	**absent**
CITY_NAME2	**0.9629**	**present**	**absent**
CLASS_NAME	0.2716	optional	present
FARE	0.8765	present	present
ROUND_TRIP	0.3703	optional	absent

Table 13. Aposteriori probabilities obtained from backward inferencing for the Class D query "*For this flight how much would a first class fare cost.*" in Table 12.

7. CONCLUSIONS AND FUTURE WORK

This paper describes our first attempt in applying Belief Networks for dialog modeling in the foreign exchange domain. The topologies of our BNs are designed to capture domain-specific constraints. While the BNs are used to infer the informational goals of the user's query, we also attempt to verify the validity of the input query against the domain-specific constraints by using backward inference. The BNs can thus detect missing concepts as well as spurious concepts, invoke the dialog model to prompt for missing information, and ask for clarification. Future work includes investigating domain portability issue as well as testing within the framework of more complex domains.

REFERENCES

[1] Meng, H., S. Lee and C. Wai, "CU FOREX: A Bilingual Spoken Dialog System for the Foreign Exchange Domain," Proceedings of ICASSP-2000.

[2] Meng, H., W. Lam and C. Wai, "To Believe is to Understand," Proceedings of Eurospeech, 1999.

[3] Meng, H., W. Lam and K. F. Low, "Learning Belief Networks for Language Understanding," Proceedings of ASRU 1999.