MOTIVATION-BASED MODELING OF UTTERANCES IN DIALOGUE

Y. Yamashita *R.Mizoguchi*

I.S.I.R., Osaka University
8-1, Mihogaoka, Ibaraki-shi, Osaka, 567 JAPAN
{yama,miz}@ei.sanken.osaka-u.ac.jp

ABSTRACT

This paper proposes a new model of utterance emergence in dialogue based on the motivation chain. In general, we have motivations when we generate utterances and carry out dialogue. The motivation can be regarded as 'driving force' of dialogue which is indispensable to perform a goal-oriented dialogue. It is difficult to identify a unique motive which makes an utterance emerge. Multiple motives often exist for generating an utterance. We investigated motives for the utterances in dialogue corpus and categorized them into five different levels; communication establishment, problem solving, problem solving control, dialogue control, communication. In a proposed model, an utterance is driven by the motivation chain which is successive combination of several motivations at different levels. The patterns of the motivation chain are classified into six categories.

1. INTRODUCTION

Remarkable progress of speech recognition/synthesis based on stochastic approaches promotes development of spoken dialogue systems. Speech recognition and synthesis are indispensable techniques to the interface of the spoken dialogue systems. They can successfully work along with higher level information such as linguistic, semantic and pragmatic knowledge. The dialogue model plays a very important role to facilitate understanding spoken dialogues and generating spoken language appropriate to the dialogue context in an intelligent and sophisticated dialogue system.

In general, utterances in a dialogue are dynamically generated through interactions between two speakers as the dialog goes on. It is very difficult to design in advance the structure of a whole dialogue. The dialogue model is required to adapt to this dynamic characteristics and to be flexible enough to catch up with dynamic transitions in dialogues. Many studies have been conducted to capture general characteristics of dialogue. Grosz proposed a model of constructing the discourse structure based on the focus change and the intentional structure to catch up with dialog development[1]. Litman refined a plan-based approach of dialog modeling by introducing the discourse plan representing relationships between the utterances and the domain plans[2]. These studies address the dialog structure which is regarded as a result of performance of dialogues. However, there is another important aspect of dialogue which has not been discussed. The process of utterance emergence, that is the reason why an utterance is generated in dialogue, is a considerable aspect for modeling and understanding dialogues.

In this paper, we discuss motivation of the utterances in dialogue and propose a new model of utterance emergence based on the motivation chain which is successive combination of several motivations at different levels. Dialogue knowledge is important to spoken dialogue system and improves the accuracy of speech recognition and the naturalness of synthetic spoken language [3][5]. Knowledge of the utterance motivation accounts for why to utter and what to make an utterance emerge in dialogue, and is expected to make much contribution to the utterance prediction[4], which is an efficient technique for reducing the search space of words and improving speech recognition score of spoken dialogues.

2. MOTIVES OF UTTERANCE IN DIALOGUE

In general, an action is driven by motives. For example, an action of 'going to the restaurant' can be driven by motives, such as 'desire to go to the restaurant', 'desire to eat something', 'physiological condition of hunger', and so on. The motives make someone decide to do an action and they are often accompanied by a conscious desire and a purpose. However, it is very difficult to identify a unique motive for an action as some motives are enumerated from different aspects for an above example.

Similarly, an action of generating an utterance is driven by motives at some different levels. Imagine an action of 'telling "Could you tell me how I can get to the *Kiyomizu* temple?"'. (*Kiyomizu* is a name of the famous temple in Kyoto.) This utterance is motivated by several desires such as 'desire to inform something', 'desire to set up a goal in a dialogue', 'desire to go to the *Kiyomizu* temple'. The first desire is concerned with communication of information, not with contents of the utterance. On the other hand, the third one comes out of his problem solving in the simulation world for sight seeing and involves the contents of the utterance.

3. CLASSIFICATION OF MOTIVATION LEVELS FOR UTTERANCES

An action can be driven by several motives, and there are some levels where a motive occurs, which are called motivation levels. We investigated the motivation levels for uttering actions, that is utterances, using dialog corpus, and identified five kinds of the motivation levels as follows.

3.1. Communication Establishment

In dialogue, a speaker often makes acknowledgment by uttering 'uh-huh' or repeating a word or a phrase which his/her interlocutor said. These utterances are generated in order to inform receipt of information and to confirm information transmission, rather than to give new information. The acknowledgment utterance gives no direct contributions to problem solving to be carried out through the dialogue. This is driven by a motive for establishing a piece of information transmission.

The communication establishment (CE) is regarded as a motivation level which is completely independent of the other motivation level. An utterance driven by motives at this level is possible to emerge at any time in dialogue whenever information is transmitted between dialogue participants. Examples of motives at the CE level are 'desire to listen again', 'desire to confirm', 'desire to agree', and so on.

3.2. Problem Solving

In a goal-oriented dialogue, the speakers have problems to be solved through the dialogue, such as planning of sight seeing, tutoring, information inquiry, and so on. Many actions are necessary to solve the problems. They include actions which are actually performed through dialogue, such as explaining a solution in tutoring, and which are simulated in an imaginary world, such as visiting a famous temple in planning of sight seeing. Most utterances are originally derived from activities in the problem solving.

The problem solving (PS) level is identified as a primary level at which motives directly drives many actions for problem solving. Motives at the PS level are always dependent on each dialogue domain and have uncountable variety. For example, 'desire to visit the *Kiyomizu* temple', 'desire to take a train', 'desire to make a plan of sight seeing', 'desire to make a reservation of hotel', and so on, are classified into the PS level.

3.3. Problem Solving Control

A motive at the PS level is tightly connected to a specific action in the dialogue domain, and it rarely occurs in different tasks. However, some generalized motives which occur in the process of the problem solving are independent of the dialogue domain. We often compare several candidates and evaluate a provisional solution through the problem solving. The motivation level where these activities are driven is called the problem solving control (PSC). A motive at this level occurs for control of the process of the problem solving. It does not exist alone but works along with motives at the PS level. The independency to the dialog domain differentiates the PSC level from the PS level. Motives at the PSC level include 'desire to compare', 'desire to find a better solution', 'desire to get a solution', 'desire to evaluate', and so on.

3.4. Dialogue Control

Although the PSC level is independent of the dialogue domain, there is another generalization of problem solving, that is control of dialogue. Some activities in dialogue does not operate the problem solving in the dialogue task, but they deal with the dialogue itself. The dialogue participants are usually careful to accomplish smooth and efficient dialogue. For example, they sometimes kindly provide additional information before it is required, and ask a question related with continuation or interruption of dialogue.

The dialogue control (DC) can be identified as a level where motives occur by regarding dialogue as a special subject of the problem solving. Motives at the DC level occurs in order to control the flow of dialogue. These motives includes 'desire to begin/end dialogue', 'desire to move on the next topic', 'desire to carry out efficient dialogue', 'desire to suspend dialogue', and so on. Since these motives occur in order to affect the flow of dialogue, motives which can occur are involved with the position of the utterance in the dialogue structure. Motives at the DC level are independent of the dialog domain because control of dialogue can be dealt with independently of the dialogue domain.

3.5. Communication

The last level where motives occur is the most superficial and essential one in information transmission. In a goal-oriented dialogue, we usually want to inquire of the interlocutor something which we do not know, and to inform him/her of something which we think he/she does not know in order to carry out collaborative dialogue. These desires are found in all utterances except for those by the motives at the CE level.

The communication (COM) level is the most important for motivating utterances and contributes to the last stage in process of utterance emergence, mentioned 4.1. This level of motivation allows just two kinds of motives; 'desire to inform' and 'desire to know'. They depend on the direction of information flow for the utterance. The former drives an interrogative utterance and the latter drives an assertive utterance.

4. MOTIVATION-BASED MODEL OF DIALOGUE

4.1. Motivation Chain

Multiple motives can be identified for an action of generating an utterance in dialogue and they are classified into five levels, as mentioned in Section 3. An utterance emerges based on the combination of several motivations

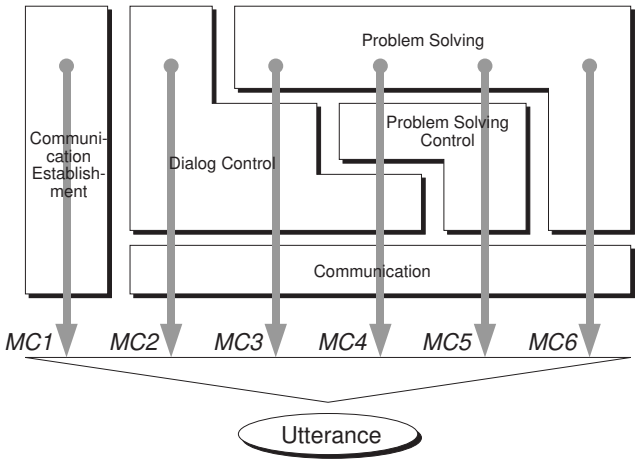

Figure 1. Utterance emergence based on motivation chains

at the different levels which successively occur, which is called *motivation chain*. All levels do not always contribute to the utterance emergence.

We analyzed dialogue corpus and classified the patterns of the motivation chain into six categories. Figure 1 illustrates the hierarchy of five kinds of the motives and the six motivation chains. MC1, ... MC6 indicates each pattern of the motivation chain. The details of motivation patterns are described below together with a dialogue example shown in Figure 2. In this figure, the pattern of the motivation chain is indicated in square brackets for each utterance. Only initiate utterances were labeled with the motivation chain and response utterance indicated by the different label 'Res.' was out of analysis of motivation.

MC1: (CE)

The first motivation chain consists of only a motive at the CE level. In the U4 utterance in Figure 2, the speaker U requested to repeat a long station name which is not familiar to him, because he could not catch it in S3. U7 and S7 make an acknowledgment by repeating a part of the preceding utterance and giving a simple reply, respectively, in order to inform the interlocutor of their understanding and confirm the information transmission. These utterances are driven by the simple motivation chain, MC1, and are tightly related to the preceding utterance.

MC2: (DC + COM)

In a goal-oriented dialogue, the domain activities contribute to the emergence of most utterances. However, some utterances which are originally driven by a motive at the DC level, such as greeting, are independent of the domain activities. The motivation of the DC level is directly chained to the COM level and this chain makes these utterances emerge in dialogue. This pattern of motivation chain is MC2.

The motivation at the COM level contributes to the

U1: '*Chotto suimasen.*' [MC1]
(Excuse me.)

U2: '*Tokyo made dou ittara ii ka osiete moraitai no desu ga.*' [MC3]
(Could you tell me how to get to Tokyo?)

S2: '*Sinkansen to hikouki to dochira ga ii desu ka?.*' [MC6]
(Which do you prefer, Sinkansen or airplane?)

U3: '*Sinkansen de iku to dou narimasu ka?*' [Res.]
(How about by Shinkansen?)

S3: '*Hankyu de Minamikata made dete, Nishinakajima-minamikata kara chikatetsu de Sin-osaka made ikimasu.*' [Res.]
(Please take Hankyu line to Minamikata station, and go to Shin-osaka station from Nishinakajima-minamikata station by subway.)

U4: '*E, doko kara chikatetsu ni norun desu ka?*' [MC1]
(Pardon, what station should I take subway?)

S4: '*Nishinakajima-minamikata desu.*' [Res.]
(Nishinakajima-minamikata.)

U5: '*Shinkansen no kippu wa doko de kaemasu ka?*' [MC6]
(Where can I get a ticket for Sinkansen?)

S5: '*Ryoko-dairiten ya midorinomadoguchi de kaemasu.*' [Res.]
(Travel agency or JR ticket counter.)

S6: '*Ryokin wa 13,480en desu.*' [MC3]
(The fare is 13,480 yen.)

U7: '*13,480en.*' [MC1]
(13,489yen.)

S7: '*Hai.*' [MC1]
(Yes.)

U8: '*Hikouki dato ikura kakari masu ka?*' [MC4]
(How much by airplane?)

S8: '*Katamichi 14,600en, oufuku de 26,500en desu.*' [Res.]
(14,600yen for one way, 26,500yen for round trip.)

U9: '*Hoka ni chigai ga arimasu ka?*' [MC5]
(Are there any other differences?)

S9: '*Hikouki no hou ga 30pun kurai hayaku tsukeru to omoi masu.*' [Res.]
(You can get there 30 minite earlier by airplane.)

U10: '*Sou desu ka. doumo arigatou.*' [MC2]
(I see. Thanks a lot.)

S10: '*Douitashimashite.*' [MC2]
(You are welcome.)

Figure 2. An example dialogue.

emergence of all utterances, except for the utterances driven by MC1. Since the main purpose of dialogues can be regarded as information transfer, it naturally follows that utterances are finally driven by this motivation. In Figure 2, U1 is an utterance opening the dialogue, and U10 and S10 are greetings to close the dialogue.

MC3: (PS + DC + COM)

The utterances driven by the MC3 patterns emerge with the source motivation at the PS level and two kinds of the successively chained motivation, the DC and the COM level. These utterances spontaneously give the additional information to perform effective information transmission, set up a goal/sub-goal of the dialogue, or make the topic return. U2 and S6 in Figure 2 are classified into MC3. In S6, the speaker kindly presents the fare which are necessary to buy a ticket before he is asked.

MC4: (PS + PSC + DC + COM)

MC4 has the longest chain of motivations and consists of four kinds of motivations at the PS, PSC, DC, and COM levels. An utterance which emerges with this pattern seeks information which are necessary to evaluate something in the problem solving, and promote the dialogue. In Figure 2, U8 is driven by the MC4 chain. In this utterance, the speaker U asks an airplane fare to compare it with a railroad fare, because he needs to evaluate the fares in order to make a decision of how to get to Tokyo. Furthermore, he introduced a new topic of the airplane as an alternative way to promote the dialogue.

MC5: (PS + PSC + COM)

The MC5 pattern for the utterance emergence is a combination of motivations at the PS, PSC, and COM level. The difference between MC4 and MC5 is the existence of the DC level. The utterance driven by MC4 neither introduces a new topic nor changes the dialogue flow. In Figure 2, U9 is an example of the utterance driven by MC5. The DC level motivation does not contribute to the emergence of U9, because the speaker continues to compare the two ways without introducing any new topics.

MC6: (PS + COM)

In a goal-oriented dialogue, the speaker has some problems to be solved through the dialogue. When he is in an impasse of problem solving in the simulation world or the real activity through the dialogue because of information shortage, he asks necessary information to his interlocutor. This situation sometimes straightforward causes an utterance to emerge. This pattern is classified into the last motivation chain, MC6. In Figure 2, the simulation result of traveling made the speaker be conscious of not knowing where he can get a ticket, and it caused U6.

4.2. Motivation Chain and Topic Transitions

One of the important views of the dialogue structure is the topic structure[1][4]. A proposed motivation-based model can account for some characteristics of topic transitions in dialogue.

The DC level has much effect on the topic because it controls the dialogue development. The utterances driven by the MC3 and the MC4 motivation chain which include motivations at the DC level often introduce a new topic into the dialogue and make the topic return. The motives at the PSC level are tightly related to the status of the problem solving. When they contribute to the emergence of the utterance, comparison or evaluation is sometimes necessary in the problem solving. It tends to guide the topic into detailed one and make a sub-dialogue by requesting additional information.

5. CONCLUSION

This paper proposed a model which accounts for the process of utterance emergence. It is based on an idea that there should be a sort of force driving dialogue, that is 'motivation' for the utterance. The motivation chain which is a combination of some motivations at different levels cause the utterance to emerge. The motivation levels were classified into five categories and six patterns of the motivation chain were identified.

Application of this model in the engineering viewpoint is not sufficiently discussed although the relationships between the motivations and the topic is briefly described. Rough analysis of the surface sentence of the utterance suggests that some sentence patterns are characterized by the motivation level or the motivation chain. This model may predict the next utterance semantically and syntactically to improve speech recognition performance. A model of the utterance emergence is necessary to understand the spoken language in dialogue and estimate the mental state of the speaker.

REFERENCES

[1] B. Grosz and C. Sidner : "Attention, Intentions and Structure of Discourse", *Computational Linguistics*, **12**, 3, pp.175-204 (1986).

[2] D.J.Litman and J.F.Allen : "A Plan Recognition Model for Subdialogues in Conversations", *Cognitive Science*, **11**, pp.163-200 (1987).

[3] A.G. Hauptmann, et al.: "Using Dialog-Level Knowledge Source to Improve Speech Recognition", *Proc of AAAI '88*, pp.729-733 (1988).

[4] Y. Yamashita, T. Hiramatsu, O. Kakusho and R. Mizoguchi : "An Utterance Prediction Method Based on the Topic Transition Model", *IEICE Trans.*, **E78-D**, 6, pp.622-628 (1995).

[5] J. Hirschberg and B. Grosz : "Intonation Features of Local and Global Discourse Structure", *Proc. of Speech and Natural Language Workshop*, 441-446 (1992)