

# Improvement of Dialogue Efficiency by Dialogue Control Model according to Performance of Processes

*Hideaki KIKUCHI*      *Katsuhiko SHIRAI*  
School of Science and Engineering, Waseda University  
kikuchi@shirai.info.waseda.ac.jp

## ABSTRACT

In this research, we aim at the establishment of a method of controlling dialogues in accordance with the changes of performance in computation. So that the system takes the most suitable dialogue strategy according to performance in computation, it needs to calculate evaluation functions modeled with performance in computation in the inside all time. In this paper, we clarify dialogue efficiency is the most important evaluation function in controlling task-oriented dialogues, and model dialogue control centered on that evaluation function with computation time of system's processes. By the users' dialogue experiment, we confirmed that the proposed model is effective for improvement of dialogue efficiency.

## 1 INTRODUCTION

Most of spoken dialogue systems are designed on condition that controlling dialogue is done within performance of processes in fixed environments. But, performance of processes such as speech recognition and language processing improves certainly, and those of appearance depend on speakers and environments. Human can change his dialogue manners according to changes of condition and environment suitably. If a system could change a way of controlling dialogues according to changes of performance of processes, it can not only realize comfortable dialogues for users but also enhance essential merits of spoken dialogue. Therefore, we aim at the establishment of the method of controlling dialogues according to performance of processes.

Some methods of learning or designing dialogue strategies with cost functions have been proposed[1][2]. These kinds of methods can express models of adopting appropriate dialogue strategies in fixed environment. So that the system takes the most suitable dialogue strategy according to performance of processes, it needs to calculate evaluation functions expressed by performance of processes in the inside all time.

In this paper, we clarify important evaluation function in dialogue control by analyzing results of evaluations of dialogues. Then, we model dialogue control centered on that evaluation function with system's performance in computation. We also implement the model on spoken dialogue system and evaluate its effectiveness.

## 2 DIALOGUE EVALUATION FUNCTION

For evaluation of dialogues, there are various elements to be considered[3]. It seemed that some important evaluation functions should be used in dialogue control model. However, we focus on only the most important one of them at first in this paper.

In this section, we clarify what evaluation function is the most important in controlling task-oriented dialogues.

### 2.1 Experiment

At first, we analyze the most significant index by multiple regression analysis between total index and other indexes in evaluation of dialogues.

We collected three human-human dialogues in each of five subjects and made them evaluate their companion's utterances in their own dialogues after them. In this research, we are focusing on task-oriented dialogues, so that we set the academic paper retrieval task in the experiment, and made subjects speak freely as possible. For the evaluation, we set the twenty-five indexes such as dialogue efficiency, transparency, and so on. Subjects evaluated dialogues with five-steps in all of those indexes and total index.

### 2.2 Result

Table 3 shows the result of multiple regression analysis between total index and five main indexes in the evaluation result.

This table shows the almost of all subjects made much of the dialogue efficiency in evaluating dialogues. From this result, we focus on dialogue efficiency as the most important evaluation function in modeling dialogue control.

In the next section, we model the dialogue control centered on dialogue efficiency with computation time as a feature of performance of processes.

Table 1: Multiple regression analysis between total index and other indexes

(\*\*: $p=0.01$ , \*: $p=0.05$ , A-E:subjects.)

Index	A	B	C	D	E
Dialogue Efficiency	*		*	**	*
Transparency				*	*
Rhythm				*	
Subjectivity					
Variety					

### 3 DIALOGUE CONTROL MODEL

#### 3.1 Modeling

Generally, it is said that number of utterances or length of turns should be decreased to improve dialogue efficiency. However, by the analysis of correlation between subjective evaluation and such kinds of objective quantity on the results of last experiment, we confirmed rate of pause-length in dialogues was the most correlative with subjective evaluation (correlation coefficient is -0.45). So that, we can say that the dialogue control model should decrease pause-length to improve evaluation of dialogue efficiency.

Filler, confirmation-utterance, and meta-utterance has the role of decreasing pause-length between utterances. Dialogue system is expected to output these kinds of utterances which can present status of system in appropriate timing to improve evaluation of dialogue efficiency.

So, we model dialogue control which can improve dialogue efficiency by outputting utterances which present system status in appropriate timing of the system. In next part, we describe how to decide output timing of system's utterances presenting status of system.

#### 3.2 Output Timing

We suppose that the system is composed of some sequential processes and computation time of processes

can be modeled by a single dimensional Gaussian distribution. Value of average and variance of distribution of each process can be calculated in advance.

#### 3.2.1 State entropy as ambiguity of process status

A set of internal states of the system can be expressed as  $S = \{S_1, S_2, \dots, S_n\}$ , where the term  $S_i$  is the state when the system is executing the process  $p_i$  and the  $i$  is the order of appearance of the each state.  $S_1$  means the state when the system is executing the first process such as speech analysis and  $S_n$  means the state when the system is executing the last process such as speech synthesis.

$P_i(t)$  is the probability of being in the state  $S_i$  in time  $t$  after receiving user's speech. When  $D_i$  is the duration of the state  $S_i$  and can be modeled by a Gaussian Distribution  $N(\mu_i, \sigma_i^2)$ , and its probability density function can be expressed by  $f_i(t)$ , Figure 1 shows the relationship between these parameter.

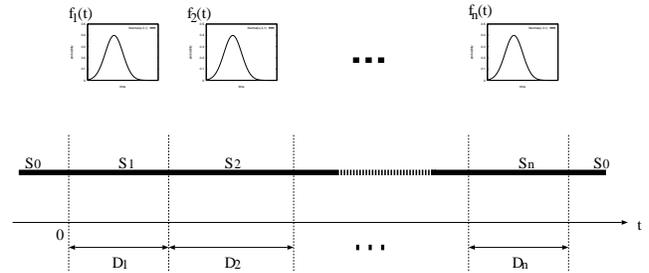


Figure 1: Process status and transition of states.

The probability of being in the state  $S_i$  in time  $t$  can be expressed as follows:

$$P_i(t) = \phi\left(\frac{t - \sum_{j=1}^i \mu_j}{\sqrt{\sum_{j=1}^i \sigma_j^2}}\right) - \sum_{j=1}^{i-1} P_j(t) \quad (1)$$

When the state which is not all other states is  $S_0$ , the probability of being in the state  $S_0$  in time  $t$  can be expressed as following equation:

$$P_0(t) = 1 - \sum_{i=1}^n P_i(t) \quad (2)$$

Finally, the entropy of state in time  $t$  can be calculated by using the probabilities of being in each state as follows:

$$H(t) = - \sum_{i=0}^n P_i(t) \log P_i(t) \quad (3)$$

We call  $H(t)$  "state entropy" as ambiguity of process status.

Figure 2 shows the transition of the state entropy on the two cases which the distribution of duration of three processes ( $D_1, D_2, D_3$ ) in the system are given. Average of state entropies are 0.69 and 0.51 in the each case. Here the numbers in the parentheses on explanatory notes mean average and standard deviation of the three given distribution of duration.

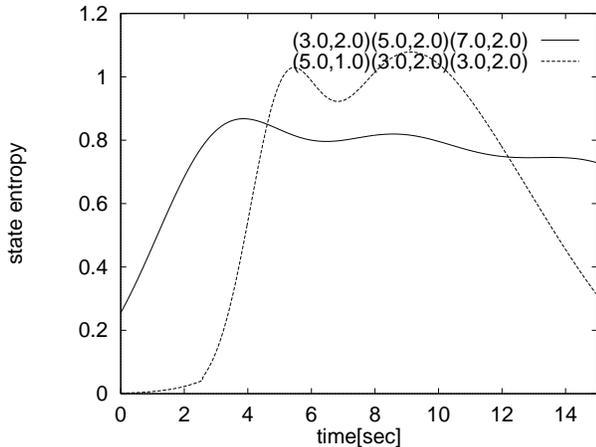


Figure 2: State entropy and distribution of duration.

As shown in this figure, the proposed "state entropy" can express the difference of ambiguity of system status in each environment. And we can say it is effective for adopting appropriate strategy in appropriate timing.

### 3.2.2 Relation between state entropy and dialogue strategy

For improvement of dialogue efficiency, a system should adopt appropriate dialogue strategies and execute them in appropriate timing to decrease pause-length. Filler, confirmation-utterance, and meta-utterance are the effective strategies in such a case. In this part, we describe how the model can decrease state entropy in decreasing pause-length by outputting meta-utterance.

Figure 3 shows the changes of state entropy by outputting a meta-utterance. Here this figure is the case that the system uttered that the process  $P_1$  was over in five seconds after starting processes, and the distribution of duration of three processes were given as

$N(5.0, 1.0^2)$ ,  $N(3.0, 2.0^2)$ ,  $N(3.0, 2.0^2)$ .

As shown in this figure, the ambiguity of process status for users can be decreased by outputting utterances which presents the status of the system before replying to the user's request. Figure 4 shows the relation between time of outputting a meta-utterance and an average of state entropy. This figure indicates that average of state entropy does not change even if the time of outputting a meta-utterance changes within a few seconds after starting processes. As like this, by calculating state entropy expressed by computation time of processes, it become possible to adopt and execute a suitable dialogue strategy in each environment.

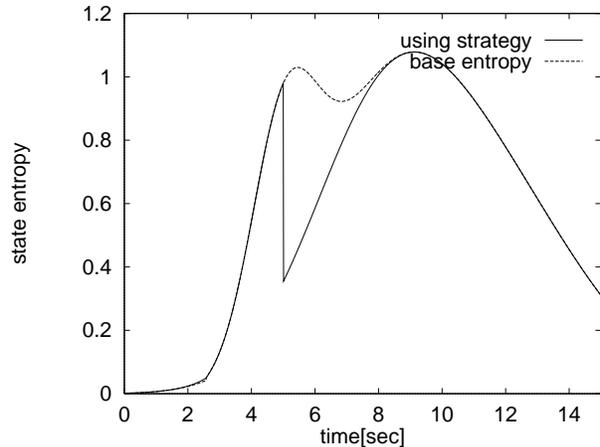


Figure 3: Decrease of state entropy by outputting meta-utterance.

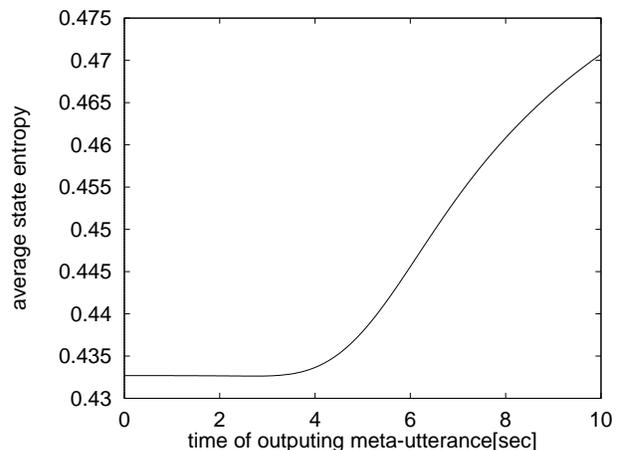


Figure 4: Time of outputting meta-utterance and average state entropy.

## 4 EVALUATION EXPERIMENT

We implemented the proposed model on the dialogue system and evaluate it by the dialogue experiment.

### 4.1 Prototype of spoken dialogue system

We constructed the prototype of spoken dialogue system which task is retrieval of academic papers from the database. The system is mainly composed of three processes such as speech recognition, paper retrieval, and utterance generation. There are almost five hundred papers in the academic paper database and the system can recognize five thousand words of speech.

In the experiment, the system changed output timing of meta-utterance which describes the first process is over for five subjects. The subjects evaluated dialogue efficiency of their own companion's utterances.

### 4.2 Experimental result

Table 2 shows the subjective evaluation result of dialogue efficiency. As shown in this table, we can see improvement of dialogue efficiency in the case of (d) "Calculated timing with state entropy" from the case of (a) that the system didn't output any meta-utterances. Moreover, we can see (d) is the best in the three cases of environment.

From these results, we confirmed the proposed model can decide appropriate timing of meta-utterances according to changes of system performance in computation.

Table 2: Evaluation result of dialogue efficiency in each models of meta-utterance control.

(Average time for paper retrieval: (1)2.4, (2)5.4, (3)10.4[sec].

Model of meta-utterance control: (a)No output, (b)Right after speech recognition, (c)After average time for paper retrieval, (d)Calculated timing with state entropy.)

Model	(1)	(2)	(3)
a)	2.5	2.0	1.8
b)	3.3	3.8	3.0
c)	3.1	4.1	3.3
d)	3.6	4.3	4.3

## 5 CONCLUSION

In this research, we aim at the establishment of a method of controlling dialogues according to performance of processes, and we modeled dialogue control centered on dialogue efficiency as an dialogue evaluation function with state entropy. In evaluation experiment, we confirmed that the proposed model is the best for deciding output timing of meta-utterance. By using this model, the system can decrease pause-length between utterances by outputting meta-utterances in appropriate timing, so that evaluation of dialogue efficiency can be improved.

Now we are modeling other evaluation functions such as the reliability of the results of processes and the rhythm of dialogues. We will model them, clarify the relationship between all models and develop the method of controlling dialogue strategies with considering trade-off between them.

## References

- [1] Esther Levin and Roberto Pieraccini, "A stochastic model of computer-human interaction for learning dialogue strategies," Proc. of ESCA, Eurospeech97, p-p.1883 - 1886, 1997.
- [2] Sadaaki Furui and Koh'ichiro Yamaguchi, "Designing a multimodal dialogue system for information retrieval," Proc. of ICSLP98, 1998.
- [3] R.K. Moore, "Understanding speech understanding," Proc. ICASSP, pp.1049-1052 (1998).
- [4] B. Shneiderman, "Designing the User Interface 2nd edition," Addison-Wesley Publishing, 1992.