Title	
Author(s)	,
Citation	
Issue Date	2001-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1446
Rights	
Description	Supervisor: , ,

A construction of the dialog system to control the animated agents which execute multiple actions

Mizuki Fujisawa

School of Information Science,
Japan Advanced Institute of Science and Technology

February 15, 2000

Most of the existing natural language processing technology have mainly dealt with the documents written in natural language, such as machine translation, information retrieval, classification, summarization. So they lacked the viewpoint of actions. But the actions of the humans are closely related to understanding natural language. For example, through natural language, people ask a partner to do actions and are asked to do actions. Therefore it is important to research to connect natural language with actions. This research enables to control the agents on a computer and the robots in the real world. It is not necessary to learn the linguistic interface because people use it usually. And the interface has the advantage that people can briefly explain complex instructions. And people can modify the previous instructions and use it again.

The first research to make computers understand natural language and execute some actions to the 3-dimentional virtual world is SHRDLU by Winograd. In SHRDLU the users can instruct the system to manipulate the world of blocks which is shown on the display. But the world of blocks which is limited and small world enables realization of SHRDLU. In spite of having pointed out the importance of the research to control the action of robots by using natural language, it did not get results after SHRDLU. The reasons of it is that the function of robots in real world and agents on a computer was limited, inputting the instruction on terminal spent time, natural language processing technology was unskilled, and the power of a computer was not enough. But as computer graphics technology made progress in recent years, we can make a 3-dimentional software robot which has a plenty of functions, and move it without the mechanical restriction. And as speech-recognition technology became to be practical use, we can use a instruction with a voice, and the performance of a computer made a remarkable progress. In this

situation, Shinyama developed the system "Kairai" which the user can instruct the agents in the virtual world by using natural language with a voice. Though Kairai is a so useful system, it has a problem which the user can instruct the agents in the virtual world only in a simple sentence. When we ask a partner to execute some action, we usually ask to execute more than two actions on a statement. Namely we usually use a complex sentence to ask. So the system which can deal with only a simple sentence is not enough useful interface.

So the puropse of this research is to construct a more useful system which controls the actions of animated agents. I develop the system by improving Kairai which solves the problem that Kairai can not deal with a complex sentence.

To make the system deal with the instruction with a complex sentence means to instruct agents to execute multiple actions. The system specifies the time relation between the actions and creates the animation corresponding to it. Then the system can easily deal with the expressions to execute actions sequentially , such as "..after..", "..before..", and the expressions to execute actions simultaneously , such as ".. while ..". But it is difficult to deal with the expression ".. and .." because the system can not find out either to execute sequentially or to execute simultaneously. In this paper, I describe the following technique to deal with it.

It is necessary to handle the aspect of verbs to specify the time relation between multiple actions. So in this research I use lexical conceptual structure which adopts aspect information. Using lexical conceptual structure enables to handle the aspect of verbs and to reason in a more general way because of making rules to not each verbs but each classes based on lexical conceptual structure. At first I classified verbs into some classes, and I made rules of the animation corresponding to the combinations of the classes. I showed the rules of the combinations such as the combination of two move verbs, the verbs that preceding verb expresses the manner of following verb, two causative verbs that preceding verb expresses the method of following verb, a move verb and a causative verb, the verbs that preceding verb expresses the change of the condition of the subject, and I showed how the system create the animations of each rules. And there are instructions whose interpretations are different depending on the condition of the virtual world. I also showed the case of the combinations of verb classes and how the system create the animation.

When the system actually deals with the instructions in natural language, at first the system specifies the actions included in the instruction, and specifies the classes including the actions. Next, the system specifies the time relation of the combination of the found classes based on the rules, creates an animation. In this way the system became to be able to deal with the instructions including a complex sentence.