Conventions in Human-Human Multi-Threaded Dialogues: A Preliminary Study

Peter A. Heeman and Fan Yang
Center for Spoken Language Understanding
OGI School of Science & Engineering
Oregon Health & Science University
20000 NW Walker Rd., Beaverton OR 97006
{heeman,fly}@cslu.ogi.edu

Andrew L. Kun and Alexander Shyrokov
Electrical and Computer Engineering
University of New Hampshire
Kingsbury Hall, Durham NH 03824
andrew.kun@unh.edu
shirokov@cisunix.unh.edu

ABSTRACT

In this paper, we explore the conventions that people use in managing multiple dialogue threads. In particular, we focus on where in a thread people interrupt when switching to another thread. We find that some subjects are able to vary where they switch depending on how urgent the interrupting task is. When time-allowed, they switched at the end of a discourse segment, which we hypothesize is less disruptive to the interrupted task when it is later resumed.

Categories and Subject Descriptors

H.5.2 [**Information Interfaces and Presentation**]: User Interfaces—*Natural Language*

General Terms

Experimentation, Performance, Human Factors

Keywords

Multi-Tasking, Dialogue, Speech Interface

1. INTRODUCTION

The purpose of this paper is to lay the groundwork for next-generation spoken dialogue systems. We envision that such systems will enable users to engage in multiple tasks at the same time in a hands- and eyes-busy environment, resulting in conversations in which *threads* of dialogues are interleaved. Such a capability would be useful in a number of domains. For example, it would allow a police officer, while running licence plate checks, to receive computer dispatches [2]. It would also allow a driver in a car to catch up on email and occasionally check upcoming traffic conditions or receive navigation instructions. In our definition, each dialogue thread is about a different topic. It is not about separate lines of reasoning about the same topic, such as comparing different furniture arrangements or different features of bank accounts [5, 4]. Our investigation of the use of multi-threaded dialogues is similar to cognitive load studies in which subjects switch between two separate manual-visual tasks [3].

Multi-threaded dialogues, even in the best circumstances, will be difficult for users, as users need to remember the details of two or more topics. Hence, we need to be careful in designing a spoken dialogue system that can switch between dialogue threads. If the system interrupts the current thread wherever it wants to and if it resumes a thread exactly where it left it off, users will probably become confused.

To discover what conventions are natural for people to use, we are running experiments in which pairs of subjects need to complete multiple tasks at the same time, and where the tasks require the two subjects to converse. Of course the reason why dialogue threads are interleaved is that there is some urgency in at least one of the tasks. In our experiment, we have subjects collaborating on an ongoing task, while occasionally having another task to solve, which has a time constraint. We varied how urgent the interrupting task is. We are examining the resulting conversations to discover the conventions that the subjects used when they interrupt, how they signal the interruption, and how they resume an interrupted thread.

In this paper, we focus on how subjects dealt with the varying urgency of the interrupting tasks when they interrupted the current thread. We found that some of the subjects altered their behavior based on the urgency. We found that when they had enough time, they would choose a point where the interruption would have a lesser effect on the dialogue. In the rest of the paper, we first discuss the setup of the tasks, then the setup of the experiment, and then present our analysis.

2. TASK SETUP

Users need to have multiple tasks that they need to complete, which require them to converse. There are a number of dimensions to distinguish the difficulty of tasks, such as:

- How long does the task take to finish?
- Can the task be efficiently done by single initiative?
- How much context is built up during the task?
- Can the outcome of the conversation be measured?
- How cognitively difficult is the task for the participants?

IUI'05, January 10–13, 2005, San Diego, California, USA.
ACM 1-58113-894-6/05/0001.

- Is the task being affected by changes in the environment?
- Do the participants' actions have different outcomes?

In this current study, we decided to keep things relatively simple and have the participants do two task types, an *ongoing task* that takes a long time to finish and a *momentary task* that just takes several turns to complete. We will study how the participants interrupt the *ongoing task* to complete the *momentary task* and then how they restore the *ongoing task*.

The momentary task can be kept quite simple as we will not be studying interruptions to it. Hence, this task does not need to build up much context and can be single initiative. We also want to vary the urgency of this task.

The ongoing task needs to build up significant context that participants will have to keep in mind. After an interruption, this context will be needed to finish the task, and so might need to be re-established. The task should also encourage both participants to equally participate as we feel this will be representative of tasks that spoken dialogue systems will be used to accomplish in the future.

2.1 Ongoing Task Definition

The ongoing task is a card game in which users work together to form a poker hand of a full house, flush, straight, or four of a kind. Each participant has three cards in their hand, which the other cannot see. Participants take turns drawing an extra card and then discarding one.

The card playing is done on computers. Both participants have a computer, which displays their cards. To discard, a participant clicks on the card. The card disappears from the screen, and an extra card is automatically dealt to the other player. When they have a winning hand, the player with the extra card needs to click the '**Done Poker Hand**' button; the computer then deals them new cards so that they can begin again. Figure 1 shows the user display.

2.2 Momentary Task Definition

The momentary task is for one participant to find out whether the other participant has a certain picture on the bottom of the display. In the screen image of Figure 1, the participant has to find out whether the other one has a blue circle.

To alert the user to the question, we flash a solid bar above and below the participant's cards. Thus the participant will know that there is a pending momentary task without having to take their attention away from the card game. The color of the flashing bar depends on how much time is remaining: green for 26-40 seconds, yellow for 11-25 seconds and red for 0-10 seconds. Participants can see the exact amount of time in the heading for the momentary (or picture) task.

The initial amount of time that the participant has is varied among 10, 25, and 40 seconds. All pairs of subjects are given the same sequence of questions, timing of the questions, and urgency.

2.3 Scoring

We want to encourage participants to accomplish both the ongoing tasks and the momentary tasks, and to use the different urgency levels so as to make as little disruption in the ongoing task as possible. Hence, we give the participants an overall score that takes both tasks into account. For the ongoing tasks, participants get 50 points for each successful poker hand they complete. To discourage participants from simply rifling through the cards to look for a specific card without talking, we subtract one point for each discard and ten points for a missed poker hand. We also subtract ten points for an incorrect poker hand. For the momentary tasks, the participants get five points for each correct answer.

Figure 1: Screen display for participants

3. EXPERIMENT

3.1 Collection Tool

To collect the conversations, we created a tool consisting of three processes, each running on a different computer. A central process deals the cards, decides the timing of the momentary tasks, records both channels of speech and creates a log file. The two other processes run the display shown in Figure 1 for each participant. The three processes communicate by sending messages over TCP/IP sockets.

3.2 Sessions

Sessions lasted about one hour. After a short orientation, the participants played the card game for about five minutes to become familiar with the rules. They then did a practice session with the card game and the momentary task for about 15 minutes, so that they could get use to managing

both tasks. Finally, they did two 15 minute sessions, which are the data that we report on.

3.3 Transcription and Annotation

We used several tools to annotate the data. We started with *SpeechView* [6] to transcribe the words that each speaker said. We then used DialogueView [1] to code utterance boundaries, speech repairs, and utterance tags. We also used it to group utterances into hierarchical discourse segments. We coded discourse segments for each card that was being discussed, each card game, and each momentary task.

From the log file of the collection tool, we created two extra text files for each conversation. The files contain the actions of the system and the users for each task with timestamps. For the momentary task, it contains each question that was asked and how it was answered; for the card game, it contains all of the cards that were dealt and what was discarded. We modified DialogueView so that it could display this information time-aligned with the word and utterance annotations. This setup gave us a full view of what was happening, and aided us in coding the discourse segments.

4. ANALYSIS

We have run five pairs of speakers, for a total of ten speakers. After we annotated all of the data, we examined whether participants varied how quickly they started the momentary task in response to the urgency of the task. The results are reported in Table 1. We found that six of the speakers did not vary how quickly they started the momentary task after it was presented. We refer to these speakers as *constant responders*. The four other speakers did vary how quickly they started the momentary task. We refer to these subjects as *variable responders*. When time permitted, the variable responders waited with initiating the momentary task. Interestingly enough, two of the pairs contained one variable responder and one constant responder.

Table 1: Responding to momentary task

	Number	10s	25s	40s
Constant Responders	6	2.59	6.38	6.32
Variable Responders	4	3.49	12.20	21.61

To determine how the variable responders used the extra time, we examined where the variable responders interrupted and contrasted this with where the constant responders interrupted. The results are shown in Table 2. The first row reports the number of times that participants interrupted during either their own or the other participants utterance. The second row reports the number of times they interrupted after an utterance, but inside of discussing a card. The third and fourth rows report the number of interruptions that occurred after either the participants finished talking about a card or finished a card game. We see that for the constant responders, there is not much difference between where they responded when they had 10 seconds versus 25 or 40 seconds. This is consistant with Table 1. For the variable responders, we see that there is a shift in where they respond. When they have 25 or 40 seconds, they tend to interrupt more after a card game and less after an utterance. The variable responders also tend not to interrupt their own utterance or their partner's utterance.

Table 2: Where interruptions occurs

	Constant Responders				Variable Responders			
	10s		25-40s		10s		25-40s	
Interrupt	9	27.3%	5	8.6%	0	0.0%	2	4.2%
Utterance	16	48.5%	37	63.8%	11	63.6%	14	29.2%
Card-Topic	6	18.2%	15	25.9%	6	27.3%	13	27.1%
Game-Topic	2	6.1%	1	1.7%	2	9.1%	19	39.6%

5. DISCUSSION AND FUTURE WORK

In this initial study, we see that some of the participants reasoned about the amount of time that they had to complete the momentary task. When time permitted, they tried to wait until the end of a card game, or at least the end of discussing a card, to interrupt. We hypothesize that these spots are much less disruptive than interrupting during an utterance or during the middle of a discourse segment. This should make it easier for the participants to resume the interrupted thread. In future work, we will quantify this effect.

These results are a first step in determining how a spoken dialogue system should interact with a user in managing multiple dialogue threads. The results suggest that the system should use the strategy employed by the variable responders, which should minimize the effect of thread-switching on the human user.

6. ACKNOWLEDGMENTS

This work was funded by the National Science Foundation under IIS-0326496. The authors wish to thank Kristy Hollingshead, Rachel Coulston, and Rebecca Lundsford for helpful conversations.

7. REFERENCES

[1] P. A. Heeman, F. Yang, and S. E. Strayer. DialogueView: An annotation tool for dialogue. In *Proceedings of the 3rd SIGdial Workshop on Discourse and Dialogue*, pages 50–59, Philadephia, July 2002.

[2] A. L. Kun, W. T. Miller, and W. H. Lenharth. Project54: Introducing advanced technologies in the police cruiser. In *Proceedings of the 2002 Spring IEEE Vehicular Technology Conference*, Birmingham, AL, May 2002.

[3] D. C. McFarlane. Comparison of four primary methods for coordinating the interruption of people in human-computer interaction. *Human-Computer Interaction*, 17(1):63–139, 2002.

[4] L. A. Ramshaw. A three-level model for plan exploration. In *Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics*, pages 39–46, Berkeley CA, June 1991.

[5] C. P. Rosé, B. D. Eugenio, L. S. Levin, and C. van Ess-Dykema. Discourse processing of dialogues with multiple threads. In *Proceedings of the 33rd Annual Meeting of the Association for Computational Linguistics*, pages 31–38, Cambridge MA, June 1995.

[6] S. Sutton et al. Universal speech tools: the CSLU toolkit. In *Proceedings of the 5th International Conference on Spoken Language Processing (ICSLP-98)*, pages 3221–3224, Sydney Australia, November 1998.