Conversations Amidst Computing: A Study of Interruptions and Recovery of Task Activity

Shamsi T. Iqbal[1] and Eric Horvitz[2]

[1]Department of Computer Science, University of Illinois, Urbana, IL 61801
[2]Microsoft Research, One Microsoft Way, Redmond, WA 98052
siqbal@uiuc.edu, horvitz@microsoft.com

Abstract. We present results from a field study investigating the influence of conversations on the multitasking behavior of computer users. We report on several findings, including the timing of the resumption of tasks following conversational interruptions and on the nature and rate of computing activities that are performed concurrently with conversation.

Keywords: Interruption, disruption, recovery, conversation, cognitive models

1   Introduction

Interrupting a task at a computer user's focus of attention often leads to a switching of attention to the source of the interruption [9]. Conversations with other people, including face-to-face conversations, phone calls, and talk through walls have been found to contribute to 15-45% of switches away from the task at a user's focus of attention [2, 3]. Indeed, becoming engaged in conversational dialog may pose greater disruptions to users than alerts delivered within a computing system because social conventions on personal responsiveness may make it difficult to take the time and actions to prepare for the task switch [9, 10].

We report on a study exploring how conversations occurring during computing tasks affect computing activities. We employed a tool with the ability to log computing activities as well as track the occurrence of conversations by noting the acoustical fingerprint of conversations. The tool logged the start and end of conversations as well as sets of activity variables before, during, and after conversations. We describe a field study undertaken with the tool.

2   Logging Computing Activities and Conversations

We developed and fielded a tool named DART (for Disruption Awareness and Recovery Tracker) for studying the influence of interruptions on computing activities. DART was constructed on top of Eve, a set of user and system monitoring components developed at Microsoft Research [4]. DART runs as a background process and logs user activities, including engagement with software applications, switches among windows, and the presence of mouse and keyboard activity. To protect the privacy of study participants, only a subset of keyboard events were recorded. The latter events

centered on actions that could provide evidence of attempts to stabilize a task before switching attention to a conversation. These included the typing of periods and the input of carriage returns (evidence of an attempt to complete a sentence or paragraph), the use of shortcuts for saving (intent to save unsaved changes), and shortcuts for cutting and pasting (discharging volatile content stored in human memory).

In a related study [5], we employed DART to investigate disruption and recovery of tasks following email and instant messaging alerts. Here, we focus on conversational disruptions. We integrated into DART a conversation-detection component developed previously at Microsoft Research [4]. The conversation detector recognizes acoustical energy in the audio spectrum in the human-voice range. The component can distinguish live conversation from other acoustical signals, including voices coming from speakers. We note that DART only tracks the occurrence of a conversation; to protect the privacy of subjects, it does not capture conversation audio. Given interim periods of silence that occur naturally during conversations, we employed a heuristic policy for distinguishing a continuation of a conversation from the onset of a new one: quiescence in conversation lasting longer than 15 seconds was considered as indicating a conversation had ended. Similar thresholds have been used to define distinct conversations during instant messaging [1, 6].

A limitation of our study is that the system did not have the ability to distinguish in an explicit manner conversations associated with face-to-face interactions versus phone calls. Also, as the conversation detector simply identified the presence of conversation, we could not disambiguate whether a detected conversation was initiated by the user or by others. Such information would be useful in distinguishing self-interruptions from external interruptions, and for studying how visual cloaking provided by a traditional phone call—and potential reduction in social pressures to attend fully to a conversation—might influence the likelihood and nature of concurrent computing activities.

3 Analysis and Results

We deployed DART for a period of two weeks on the primary machines of 16 people at Microsoft whose job titles included program managers, researchers, and software developers. The participants had microphones installed on their computers as part of their default configuration (largely via embedded laptop microphones). The occurrences of conversations were logged only if the user was active on the computer.

We coarsely classified computing applications into two categories: *task-centric* and *peripheral* applications. We define task-centric applications as the use of software development or productivity applications, used typically within our organization to perform primary job responsibilities, *e.g.*, Visual Studio and Microsoft Office applications. We deemed communication applications, such as Microsoft Outlook for email and calendaring activities, and web search engines as peripheral.

We sought to explore the rate at which a user switched among tasks at different times, as captured by switches among windows of applications open on a user's computer, and the time spent on each application. We also were interested in how often users performed actions that could be characterized as stabilizing the state of a project, *e.g.,* saving files and completing sentences or paragraphs, pasting information

that had been previously copied, etc., as it would seem that such actions might be useful to perform so as to leave it in a more recoverable state before turning attention to something else. We were also interested in the influence of visual cues for suspended applications on the efficiency of task recovery; we hypothesized that visible windows could serve to remind users about suspended applications.

We found that computer users spent a mean time of 21 minutes, 49 seconds (s.d. 39m, 59s) per day on conversations that were initiated while they were performing tasks on the computer. Separating the results by job role, the breakdown is 30 minutes 51 seconds (s.d. 26m, 42s) for software developers, 42 minutes 27 seconds (s.d. 47m, 9s) for managers and 2 minutes 1 second (s.d. 2m, 42s) for researchers.

We compared computing activities prior to the conversation (PC) and after the conversation had started (AC). The PC phase was defined as starting 5 minutes prior to the conversation, an observational period enabling us to capture a representative sample of activities before the interruption. Following the initiation of a conversation, users were found to perform the predefined task-state stabilizing activities (saved file, completing sentence, etc.) at a significantly higher rate than during the PC period. These findings are captured as mean rates of task stabilizing activities in Table 1.

48.12% of conversations occurring while users performed computing tasks were associated with inactivity for at least part of the conversation. Out of these, in 0.9% of the cases, users became inactive *after* the conversation started (*i.e.*, activity on the computer persisted for some time into the conversation). In 30.7% of these cases, users were inactive just prior to the conversation and remained inactive throughout the conversation. For the remaining 68.4%, users were temporarily inactive prior to the conversation, but became active as the conversation progressed.

As we could only detect the occurrence of conversations, not the details about the initiating event, we hypothesize that the first breakout is where the user is interrupted by someone else and the second case is where users instigate conversations. For the third case, users decided to continue computing tasks concurrently with the conversation as the conversation progressed, potentially something that could be done more comfortably during telephone conversations than in person. Another explanation for computing activities during conversation is that the conversation somehow caused the user to become active (*e.g.*, check mail). We seek in future work to extend the logging so as to better understand the initiation of conversations.

When conversations led to inactivity for the entire duration of the conversation, 2 tasks (s.d. 1.51) on average were suspended and 16 minutes, 22 seconds (s.d. 14m, 24s) passed before activity was next seen on the computer. We hypothesize that in many of these cases, users left the computer during the conversation, accounting for the rather long break between suspension of and return to the computing tasks, confirmed by later interviews.

Table 1. Task state stabilizing activities/minute.

Activity	Pre-Conversation Mean (S.D.)	After Conversation Mean (S.D.)
Save	0.33(0.36)	0.85(1.16)
Paste	0.38(0.36)	1.19(2.63)
Sentence Completion	0.84(0.69)	2.88(4.09)
Paragraph Completion	0.78(0.81)	2.63(4.49)

Table 2. Peripheral activities/minute.

Activity	Pre-Conversation	Intra-Conversation
Mail open	0.39 (0.40)	1.13 (2.19)
Mail writes	0.44 (0.36)	1.49 (2.77)
Mail Sends	0.25 (0.14)	0.80 (1.79)
Web mail checks	0.37 (0.29)	1.39 (2.15)
Web searches	0.35 (0.31)	0.77 (1.00)

For cases where users were active on the computer during the conversation, they switched applications at a rate of 0.48/min (s.d. 0.67), significantly lower than the switch rate during the PC period (0.77/min, t(14)=7.88, p<0.001). Users spent on average 2 minutes and 24 seconds (s.d. 4m, 20s) on each application, significantly higher than during the PC period. Outlook was the most accessed application during this time. Other top applications were Internet Explorer, Visual Studio and Office Communicator. When accessing Outlook, users performed activities at a significantly higher rate (p<0.0001 for all actions), as well as higher rates of web mail checks and web search operations, as compared to the PC period (Table 2). This finding may indicate that the disruption of the focused task by the conversation offers users an opportunity to perform less attentionally taxing peripheral tasks, and that these are skimmed during conversation at a higher rate, potentially in a less focused manner. Visual Studio appearing in this list of otherwise peripheral tasks indicates that users were occasionally able to converse and continue working on their ongoing tasks.

Overall, these findings provide evidence that subjects used the break to switch to concurrent peripheral applications. The lower mean switch rate during conversations may be due to crosstalk among related cognitive resources for having conversations and executing computing work, given, *e.g.,* the need to share verbal and other skills simultaneously [7, 8]. It is also possible that users consciously decide to perform only certain tasks (*e.g.*, checking email) in parallel with the conversation as they can effectively share resources without drastic degradation in the performance of either.

Following the completion of a conversation, or becoming active on the computer again if the conversation caused inactivity, users took on average 11 minutes 20 seconds to *resume* their suspended applications. Windows that were less than 25% visible took significantly longer to resume than windows that were more than 75% visible (t(16)=3.259, p<0.005), suggesting that the visibility of the suspended application windows served as a cue to return to the suspended applications. This observation was validated later through user interviews.

We further explored the relationship between applications that users were focused on before the interruption and the time to resume suspended applications. Our analysis showed that there was a 0.2 probability of not resuming activity on active windows as such within 2 hours of the end of the conversation, which may indicate that users had forgotten about these tasks.

We also explored whether the time spent on active windows before the suspension had an impact on the resumption time. Figure 1 illustrates the findings. Active windows where users spent < 1 minute before suspension, had a 60% probability of being resumed within the first minute of return, but also had a 2.1% probability of not being resumed at all during the session. Windows on which 5-15 minutes were spent before suspension were almost certain to be resumed within 5 minutes.
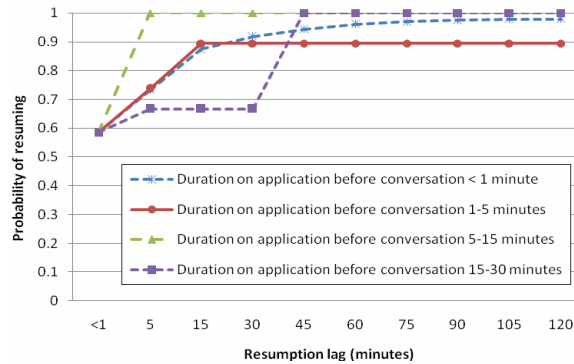
Fig 1. Cumulative probabilities of resuming work in a suspended active window as a function of resumption lag.

4   Conclusion

We performed a field study of user task execution behavior before and following conversational interruptions. We found that, following an interruption of tasks by conversation, users often suspend their ongoing computing tasks to participate in the conversation, may embark on peripheral tasks such as email correspondence and web searches, and show a slowing of computing activities. The time until resuming a task after a conversation was influenced by the duration of activity on the task before the interruption and increased visibility of suspended application windows was associated with faster resumption.

References

1. Avrahami, D. and S. Hudson, Communication characteristics of instant messaging: Effects and predictions of interpersonal relationships. CSCW 2006, 505-514.
2. Czerwinski, M., E. Horvitz and S. Wilhite, A diary study of task switching and interruptions. CHI 2004, 175-182.
3. Gonzalez, V.M. and G. Mark, Constant, constant, multi-tasking craziness: managing multiple working spheres. CHI 2004, 113-120.
4. Horvitz, E. and J. Apacible, Learning and reasoning about interruption. ICMI 2003, 20-27.
5. Iqbal, S.T. and E. Horvitz, Disruption and recovery of computing tasks: Field study, analysis, and directions, CHI 2007, to appear.
6. Isaacs, E., A. Walendowski, S. Whittaker, D.J. Schiano and C. Kamm, The character, functions, and styles of instant messaging in the workplace. CSCW 2002, 11-20.
7. Treisman, A.M. Contextual cues in selective listening. Quarterly Journal of Experimental Psychology, 12. 242-248.
8. Wickens, C.D. Multiple Resources and Performance Prediction. Theoretical Issues in Ergonomic Science, 3 (2). 159-177.
9. Wickens, C.D. and J. Hollands Engineering psychology and human performance. Prentice Hall, Upper Saddle River, NJ, 2000.
10. Yantis, S. Stimulus-Driven Attentional capture and attentional control settings. Journal of Experimental Psychology: Human Perception and Performance, 19 (3), 676-681.