Building Assistive Technology for Breakdown Situations in Collaborative Communications via IM

Nik Nailah Binti Abdullah[1], Eric Tschetter[2], and Shinichi Honiden[1,2]

[1]Honiden Laboratory, Intelligent Research Systems Division,
National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-Ku, Tokyo 101-8430, Japan.
[2]Department of Computer Science, Faculty of Science, University of Tokyo,
7-3-1, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan.
{bintiabd, eric, honiden}@nii.ac.jp

Abstract

Current trends in scientific collaboration focus on developing effective Web-based communication tools, such as instant messaging (i.,e IM) or video-conferencing. One objective is to provide *informal communication opportunities* for collaborating scientists. This paper focuses on developing an assistive technology for predicting where breakdown situations have likely occurred in chat communications. The immediate goal of the assistive technology is to help users cope with discrepancies between their expectations and the tools they are using during scientific collaboration. The ideas in this paper also have far reaching implications, introducing a method that can be used by any human-interactive application to help the application understand when the user does not know what to do.

Introduction

Recent years have given way to an increasingly global society. Corporations have become multi-national, news travels around the world in a matter of seconds and researchers are increasingly seen working on collaborative research efforts that span national borders. These researchers must maintain a constant stream of contact with their collaborators while still remaining physically located in their own country. A number of different tools that capitalize on the power of the internet have been used to achieve this communicative goal, but none of these tools is perfectly intuitive (Binti Abdullah 2006). That is, there are things about the tool itself that are sometimes difficult to understand for the users, for example, the user might not know what to do if the tool suddenly disconnects. This paper discusses the implications of cognition theory to the creation of software and its application interface. As an example, this paper focuses on one specific software application used for scientific collaboration: BuddySpace IM (Eisenstadt , Komzak and Cerri 2004).

In this paper, we introduce the beginnings and initial motivations for an assistive tool that facilitates effective communication protocols among any group of collaborating individuals. Our ultimate goal is to build an assistive technology that can automatically detect and alert users about the probability of a breakdown situation during a particular context of communication. From here onwards, the assistive technology helps to indicate the next best action the user should take in order to anticipate such problematic contexts. Achieving this goal requires that our system first know how to properly indicate and classify a breakdown situation, which is the direction we are currently pursuing.

The paper discusses (i) some background information regarding the context of our research, it then goes into (ii) the nature of the EleGI joint project, followed by (iii) theoretical foundations and method of analysis. We next discuss the current (iv) results and (v) evaluation, ending with some perspectives on our work as well as future implications.

Background

The theoretical framework this paper is based on is concerned with how people *learn to induce communication protocols* (Binti Abdullah 2006, Binti Abdullah and Cerri 2005). The framework was based on actual communication protocol problems encountered by scientists working together during the Haughton-Mars Project (i.e, HMP) (HMP 2001, Clancey 2001). It refers to several significant events where users were thrown into different breakdown situations [1] that impacted the understanding of how communication and learning is taking place in respect to using *'located' tools* (Binti Abdullah 2006). We focused on several issues: (i) how human learning and communication mutually influence one another in a

[1] A *breakdown situation* is a situation in which a conversation "breaks down" due to some influence external from the actual conversation. A technical failure with the communication tool or some other peripheral device is one example.

situated *context engaged in an activity*. For example, a user using a tool conducting a survey would be engaged in the context of 'conducting survey". He is situated in *this context* engaged in the activity of 'using the sampling fossil tool' while communicating with a co-worker. They are both learning and communicating about using the tool and about the context of 'conducting survey' on that particular fossil. (ii) how do humans abstract and recognize similar contexts. That is, after some time, people handle *context of communications* as if they were *repeatable*. If we are to commit to this notion of repeatability, then we must ask ourselves, how the actors are able to recognize a *context* as being *context A* and act within it. Hence, we must know how the actors are aware of with what and with whom to communicate (and applying a particular communication protocol) in their situated context.

This paper uses data collected from a similar context of study: collaboratory efforts taken by the researchers working on the European Union Learning GRID Infrastructure. The European Learning Grid Infrastructure (EleGI[2]) project was started with the idea to radically advance the effective use of technology-enhanced learning in Europe through the design, implementation and validation of a pedagogy-driven, service-oriented software architecture based on GRID technologies (Breuker et al 2006).

BuddySpace IM was tested for use in the EleGI project for two main reasons: (i) to re-engineer and harmonize the current tool (i.e., IM systems) and collaboration methods. Specifically, the tool developed is focused on contextualized, experiential and collaborative context that allows informal learning among the collaborators; and ultimately (ii) to support groups working in different collaborative contexts ragging from the self-organization of on-line virtual communities to the experimentation of communication and information management tools (Breuker et al 2006).

BuddySpace IM was installed with the initial goal to provide a quick and effective medium for discussing projects and accessing information. IM was frequently used as a "middle" communication protocol, that is, it was a part of some larger communicative process, like discussing a paper. It often became a stimulus to possibly long and complex protocols involved in paper writing and decision-making. It was as an effective and efficient method of giving step by step instructions to users/group members. The frequency of communications carried out on BuddySpace was extremely event-related. On the other hand, despite the merits of the BuddySpace system for the EleGI group members, chat logs revealed that *multitasking activities* caused *frequent pause situations* because the communication-enabling system became less flexible. In other words, when there is a discrepancy between how people think and behave with respect to computer systems a *breakdown* may occur.

[2] EleGI website: www.elegi.org

BuddySpace IM: a case study of two collaborators

We start this section with a discussion of common scenarios that take place during the collaboration. Then, from here onwards we will discuss the common problems that arise during the collaboration that inhibit the progress of scientific discussion. We show in Figure 1 the main window of BuddySpace user interface.

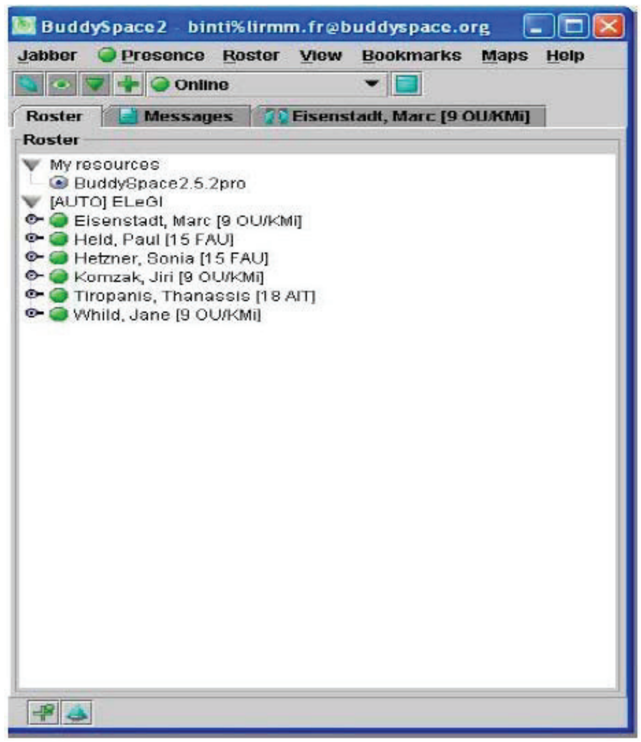

Figure 1: The main view of BuddySpace instant messaging.

Since the users (Figure 1) come from different institutions, e.g., "Eisenstadt, Marc [9 OU/KMI]" represents that user Eisenstadt is from the Open University (OU), Knowledge Media Institute (KMI). The number 9 denotes the institution number in the EleGI project. The EleGI project is composed of different research groups all working together (e.g., chemists, biologists, GRID specialist). Each of the members in the project has different roles. For example a project coordinator for the EleGI organizes meetings, collects contributions from group members, compiles the documents that have to be sent to the board of executives, and solves group members issues and budgeting within the group (e.g., traveling expenses for conferences, buying new equipment for running experiments etc). A project executive outlines the task of their affiliated laboratories and team members, as well as organizes everyone toward a desired goal for collaborating laboratories (i.e., that attempts to fulfill both the organization goal and EleGI project goals). In this section we only discuss the excerpted chat logs between a project coordinator and a project executive.

[2004/04/19 10:38] <Executive> *just looking* at *the table* in Technical Annex v7.6, page 15

[2004/04/19 10:38] <Executive> with goals/objectives/results...

[2004/04/19 10:38] <Executive> *comparing* that with the *document sent last week*... about GQM..(edited)

[2004/04/19 10:39] <Coordinator> I am *looking for this one*

[2004/04/19 10:40] <Coordinator> now I have the two opened

[2004/04/19 10:40] <Executive> yeah, *email of 8th April* "setting up the G"

[2004/04/19 10:41] <Coordinator> first say what you have to do

....(*edited*)

[2004/04/19 10:48] <Coordinator> I had a technical problem, so I quitted

[2004/04/19 11:03] <Executive> heh no problems

[2004/04/19 11:03] <Coordinator> buddy space was suddenly frozen

[2004/04/19 11:03] <Executive> heh no probs

[2004/04/19 11:03] <Executive> WHILE TRUE DO: <deposit money into account of William Gates III>

[2004/04/19 11:04] <Coordinator> ha

[2004/04/19 11:04] <Executive> ah, buddyspace frozen... hmmm..

[2004/04/19 11:04] <Executive> WHILE TRUE DO: <deposit money into account of Bill Joy & Bill Gosling>

[2004/04/19 11:04] <Executive> (i.e. that's a Java bug)

[2004/04/19 11:04] <Executive> or rather Java Runtime Environment... nothing we can do

[2004/04/19 11:05] <Coordinator> anyway ,did the "Bill gates/ppt" reach you ?

[2004/04/19 11:05] <Executive> did you have an open 'dialog box' when the freeze occurred?

[2004/04/19 11:05] <Coordinator> Nothing, I had to kill the process with Task manager !

[2004/04/19 11:07] <Executive> aargh

[2004/04/19 11:07] <Executive> you mean you weren't doing anything other than typing to me?

[2004/04/19 11:07] <Coordinator> sorry

[2004/04/19 11:08] <Executive> (maybe you were opening powerpoint)?

[2004/04/19 11:08] <Coordinator> of course I had many other processes open

[2004/04/19 11:08] <Executive> just trying to diagnose the process...

[2004/04/19 11:08] <Executive> OK; well, let me know if it happens again... worth noting the context...

The IM messages shown above were part of a decision regarding the creation of the EleGI deliverable. The chat logs reveal that an e-mail and its attachment were referenced twice during the conversations. However, multitasking activities on the computer caused the system to disconnect for the coordinator. The project coordinator compared to the project executive had more pause situations due to system instability. He was also frequently multi-tasking his activities because of his role: he normally becomes a point for answering any questions about the project documents and e-mail. These observations suggest to us that role and duty also influences the behavior of IM users[3] and its activities on the system. The discussion had to be re-initiated after 4 minutes of discussing the breakdown. Unfortunately due to the breakdown situation, the executive was unable to discuss the initial business goals, instead, he *was drawn onto* the presentation slides sent through e-mail by the coordinator. In Isaacs et al (2002), the multitasking scenario is highly reported in various literature (Connell et al 2001, Nardi, Whittaker, and Bradner 2000). In (Binti Abdullah and Honiden 2007), the authors reported statistical findings on the relationship of roles to tool functionality and the frequent occurrence of pause situations. According to the report, there is a high level of correlation between a user's role and what conversational context (i.e. giving instruction) the tool is being used for. At the same time, there is a high degree of correlation between the number of concurrent activities (i.e., topics discussed over the IM) and the number of pause situations that occur.

Our observational unit starts from measuring "time" looking into pause frequency and duration of pause time. Based on the data, whenever chat reply is >= 2, it is somewhat likely that a user has (physically or mentally) become detached from the conversation at hand. Of course, not all pauses indicate a breakdown situation. An example is when a recipient is away from his office and does not reply to a chat immediately; where the reply may come in 8 minutes later. It is likely that whenever time pause is <= 5 minutes, this pause situation is a breakdown situation, i.e.

[3] Perhaps a protocol should help coordinator to coordinate its multitasking activities? These are some scenarios we imagine an assistive technology can do to assist collaborators.

the user is actively seeking to rectify the detachment. Thus in the beginning we use the *time* as a rule of thumb to narrow down the conversations to only those that are belonging to breakdown contexts.

Therefore the preliminary issue we tackle is enabling the system to induce and recognize a breakdown context from a non-breakdown context. The subsequent sections offer an overview of the theoretical foundations used.

Theoretical foundation: Hierarchy of learning and communication

The conversations are analyzed and converted into formalized messages using the activity states framework (Binti Abdullah and Cerri 2006, Binti Abdullah 2006). The activity states framework is an integration of multiple cognition theories. The theories are (i) hierarchy of learning and communication (Bateson 1972); (ii) situated cognition (Clancey 1997) and (iii) activity theory (Leont'ev 1978). In this paper we only discuss the learning hierarchy by (Bateson 1972) that guides us for recognizing different contexts of communications.

The learning hierarchy was originally based on the work by Whitehead and Russell's type theory discussed in Principia Mathematica[4] (Bateson 1972). It is founded on the notion of *changes*, in fact, Bateson argues that the word learning in and of itself denotes an act of change.

For example, the changes from purely habitual to highly complex of learning can be assumed as *changes.* Consider that the changes from learning how to brush teeth (that has become habitual) to learning how to navigate an airplane. These acts of "change," or learning, have been ordered into levels 0, I, II, and III, shown below in Figure 2.

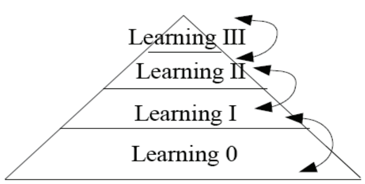

Figure 2. Hierarchy of learning and communication.

In Figure 2, the arrows represent two relationships. The arrow pointing one level up in the hierarchy describes the inductive nature of the learning process, i.e. that low level ideas and actions are the basis for higher level learning. The arrow pointing downwards represents the deductive relation among ideas and items of learning types. Put another way, they indicate that from complex learning of behavior could become at one point a merely habitual behavior. Hence the hierarchy can be read both inductively and deductively. The learning levels can be summarized as the following; *(i)* learning 0 describes learning something that is only right or wrong and is not subject to correction (e.g., When the lunch whistle blows, I learn that it is 12.00 o'clock). There exist *external event systems* (i.e., information from the environment) that inform people to select which item from which set of alternatives to respond to that particular stimulus; *(ii)* learning I is described by the change in the specificity of response by correction of errors of choice within a set of alternatives, where these set of alternatives are those that were habitual occurring in learning zero (e.g., I learn from 'items' *curtain* and *stage* that the hero discussing suicide is in a play. I do not immediately phone the police); *(iii)* learning II is described as a procedural change of learning I, a corrective change in the set of alternatives (i.e., different sets of alternatives indicating how to respond to some situation) from which a choice is made or a change in how the sequence of *experiences* is punctuated (e.g., I learn to *punctuate this event* as being 'debug the tool with procedure X' from my set of experiences); (iv) learning III is described as a change of process in learning II, specifically, it is a corrective change in the system of sets of alternatives from which choice is made (e.g., learning that occurs characterizing changes in character, or the selfhood).

In other words, learning zero is a simple classification of categorizing 'signals,' where signals are considered 'external event systems,' or a system that tells people what to select as a response and when. Bateson continued to categorize those signals that help to identify contexts, introducing the notion of repeatable context (i.e., two contexts A and B could be similar at times 1 and 2). There must be signals that tell people to discriminate one context to another. In learning II, these contexts are referred to as streams of events. The signals that mark or label a context are called 'markers' and punctuate the events of the context.

The learning hierarchy is used in order to allow the assistive technology to distinguish a breakdown context from a non-breakdown context. Specifically we focus on the idea of *items, signals* and *contexts*, which is discussed in the method of analysis.

Data for experiments

As mentioned above, the data for our experiments has been translated into a formal language designed for use in this type of conversational analysis. Specifically a message like:

> [2004/03/09 11:04] <Executive> if you press the 'sailboat' icon at the top of this window..

Would be converted into

[4] The basic idea of Russell type theory is that we can avoid commitment to R (where R is the set of all sets that are not members of themselves) by arranging all sentences (or, equivalently, all propositional functions) into a hierarchy. The lowest level of this hierarchy will consist of sentences about individuals. The next lowest level will consist of sentences about sets of individuals. The next lowest level will consist of sentences about sets of sets of individuals, and so on. It is then possible to refer to all objects for which a given condition (or predicate) holds only if they are all at the same level or of the same "type." (Stanford 2003).

(26) (request-when :sender e :receiver c :content (if (c (press (icon sailboat))))

Which is read as:

(<*msg #*>) (<*act*> <*sender*> <*receiver*> <*content*>))

Where,
<*sender*> is represented as *':sender <actor>'*
<*receiver*> is represented as '*:receiver <actor>*',
<*content*> is represented as '*:content <msg>'*, and
<*msg*> is a representation of the human-analyzed message content

That is, (26) denotes the message number, the sender, e, is the executive and the receiver, c, is the coordinator. The <*msg*> in the <*content*> argument is written in predicates according to the model introduced in (Binti Abdullah 2006). This message is used as an example throughout the rest of the paper.

Method of analysis

There are three main notions used in our method of analysis for the assistive technology derived from the activity states (Binti Abdullah and Cerri 2005, Binti Abdullah 2006). They are (i) object; (ii) subject; and (iii) in-between process. The *object* and *subject* are formulated specifically as: (i) how people conceive what they "read," (ii) how people combine what they read with what they perceive and (iii) how actions are formulated through interaction (speaking or interacting with tools). An object can be defined as the 'perceiving' act of a speaker at some moment, for example, looking at a bowl. The bowl is the current focus of the speaker. The bowl is an object. A subject is defined as the referential process of the speaker in respect to their 'perception' (e.g., what ideas I come up with when looking at the bowl). The in-between processes deal with the occurrences taking place between the *object* and *subject* by relating it to (i), (ii), and (iii). We use the idea of signals and items that is has a transitivity property to an object.

For example, collaborator, Jenny, is engaged in drafting the first EleGI article. Jenny is focused on her *article, which is the object of her focus*. From here onwards, she is referring to the *subject* 'tools'. Perhaps the information 'deadline' may be a *signal* from sender Peter that prompted Jenny to focus on her article. *Instant messaging* is a *mediator* for Jenny to achieve her goal, it facilitates sending the drafts to her colleagues, receiving inputs from them and improving the paper from collective view points[5]. The *subject* in this example is *Jenny*. Thus, receiving instant messages from her colleagues are actually signals that mark a context for Jenny, allowing her to induce and punctuate her current situation. Jenny may learn from her experiences that, in this context it is best to send the technical issues regarding a tool to some person A for annotation.

This analysis is always 'situated' to the located tool, in this case BuddySpace IM.

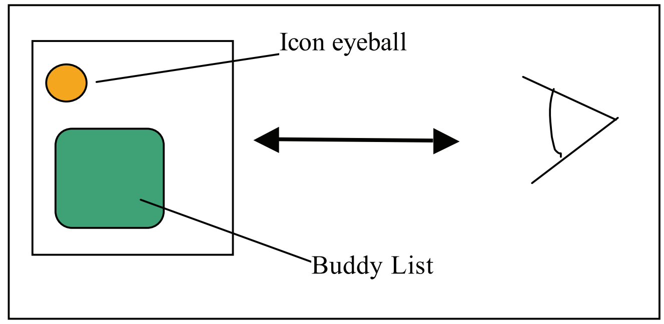

Figure 3: The relationship between focus and 'perception' tied to the IM.

Figure 3 shows an overview of Jenny's situation described above. First, we hypothesize that there are certain distinct features that appear on the tools and capture the attention or focus of a user. These features are signals that cause Jenny to focus on certain details that influence her choice of responses. Jenny is always actively remembering (Bartlett 1932) and responding situated in her context (Clancey 1997). Since we want to enable the assistive technology to learn from contexts, breakdown and non-breakdown, we observe patterns in re-sequencing the act and content predicates (refer to the previous section).

Conversational contexts are defined using the following pseudo-algorithm:

1) Start by collecting all items, called "primitives," in the content argument of the start of a context (e.g., 'if', 'c', 'press', 'icon', and 'sailboat' from message (26) above).
2) Current message = next message
3) If a primitive in the now current message has been repeated in the current context,
   (a) The primitive is labeled a dominant primitive for that context.
   (b) The current message is added to the context
   (c) Current message = next message
   (d) GOTO Rule 3)
4) ELSE
   (a) Context has ended

Dominant primitives signify the *focus/object* of that particular context. Each response is equivalent to the theoretical idea of a *signal*. The non-dominant primitives are those that are 'evoked/produced' from associating the *focus* to the actor's memory/experience. This observation

5 These occurrences are referred to as activity theory (Leont'ev 1978), which is not discussed in this paper. For details see (Binti Abdullah 2006, Leont'ev 1978).

of the re-sequencing of primitives and its communicative act helps us mark the *beginning* and *end* of a context. In other words, it is an algorithm for identifying conversational contexts, which can subsequently be *punctuated*.

Experiments

We decided to use a rule-based analysis method as a springboard for determining the best possible method of context recognition and *punctuation*. We also decided to limit our current activities to the analysis of breakdown situations. We started with three breakdown situations that we found through human analysis of our conversations data. We then set out to label a general pattern for breakdown situations and came up with the following algorithm.

1) IF 5 min >= time difference between communication >= 2 min
   (a) Message begins a breakdown situation, tag it
2) Current message = next message
3) IF primitive in current message appears in previous message
   (a) Add current message to breakdown situation
   (b) Add all recurrent primitives to set of *dominant primitives*
   (c) Add all other primitives to *breakdown history*
4) ELSE
   (a) Tag situation as *irregular*
   (b) END
5) Current message := next message
6) IF a *dominant primitive* appears in current message
   (a) Add current message to breakdown situation
   (b) IF any primitive in *breakdown history* appears in current message
      I. ADD it to set of *dominant primitives*
   (c) GOTO 5)
7) ELSE
   (a) CHECK <*act*>
   (b) IF <*act*> = ACK
      I. Add Message to breakdown situation
      II. Update *dominant primitives* and *breakdown history*
      III. GOTO 5)
   (c) IF <*act*> = INFORM-REF && <act> of next message = CONFIRM
      I. Add both messages to breakdown situation
      II. Update *dominant primitives* and *breakdown history*
      III. Current message := message after CONFIRM
      IV. GOTO 6)
   (d) IF <*act*> = INFORM && <*act*> of next message = INFORM-REF
      I. Add INFORM message to breakdown situation
      II. Update *dominant primitives* and *breakdown history*
      III. GOTO 5)
   (e) ELSE
      I. Current message := next message
      II. IF dominant primitive appears in current message
         1. Associate current message and previous message
         2. GOTO 5)
      III. ELSE
         1. END

Basically, the algorithm starts using the pause situation discussed above to determine that a breakdown situation has begun. It then iterates down the conversation adding communicative events that share an *object* or *focus* (i.e., primitive) with the breakdown situation. When the messages no longer share primitives, the breakdown context has ended. One particular point of note is Rule 4, this rule basically states that if the message after the beginning of a breakdown situation does not share a primitive with the breakdown situation, it is labeled as *irregular* and conversational context analysis is broken.

Results

We applied this algorithm to 20 conversation files[6], took the results and then looked back at the conversations to determine if the output truly was a breakdown situation, the results of which are shown in Table 1.

	# breakdowns	# output	% correct
w/ *irregulars*	6	26	23.1
w/o *irregulars*	4	7	57.1

Table 1. number of correctly analyzed breakdown situations

The first row in Table 1 represents all outputs while the second row represents all outputs that are not labeled as *irregular* (see Rule 4 in the algorithm above).

As is readily apparent, the rule-based algorithm above finds many situations that are not actually breakdown situations. However, it also found three breakdown situations that we had not previously known about, which indicates that the rules do actually generalize, though to a very limited extent.

The imprecision of the algorithm introduced is primarily caused by the fact that the only method of determining that a breakdown situation has started is by looking at the time difference between messages. However, upon examination there was no readily apparent pattern for determining the start of a breakdown situation that uses communicative acts and primitives alone.

Instead, we will need to look deeper into the re-sequencing of the communicative acts *and* primitives, employing the full gambit of the activity states theory. The challenge we face is to allow the program to identify at what point the breakdown situation has occurred. Timing shows that it is not sufficient to capture this information. The time mechanism used as a start is too ambiguous, we need another operation to refine our search. On the other hand, a conversation that is categorized solely based on the element of 'focus' (i.e., dominant primitives) may be useful as a guide, regarding the topic of conversation, possibly allowing us to classify context (i.e., giving

[6] The 20 files include the conversations that were originally analyzed for breakdown situations.

instructions, discussions, tool debugging). However, again this is insufficient; it does not allow us to know which categorized context houses the breakdown contexts. Specifically, the problem lies at knowing how to differentiate these events into punctuated contexts.

Even during a breakdown situation, we found that there were some pause situations that the program was unable to capture, resulting in the program performing naïve search and indicating that a breakdown situation has ended prematurely. We propose that in the next direction is to give more definition to the program. Let us consider the sentence "Would you be interested in using a version map", after conversion it is (following the syntax introduced in Data for Experiments):

(3) (request-whenever :sender m :receiver p :content (interested (m, using (p, version (map))))

We focus only on the content layer.

[1] what (who, act (who, what (what)))
[2] interested (m, using (p, version (map)))
[3] a1 a2 a3 a4 a5 a6

The primitives in the content of message number (3) followed the syntax shown in [1] defined in (Binti Abdullah 2006). The primitives in the content shown in [2] are labeled with corresponding numbers shown in [3] to indicate that they are different. To explain further what we can do with these structures, consider that the speaker says "I showed it in Barcelona". We have:

<(4) (inform-ref : sender m :receiver p :content (showed (m, it (p, Barcelona))))) >

We look at the content.

[4] act (who, what (who, where))))
[5] showed (m, it (p, Barcelona)))
[6] a7 a8 a9 a10 a11

Refer to [3] and [6], the structure a5 (i.e., version) and a6 (i.e., map) is referring to a9 (i.e., it). If we refer to the syntax defined in [1], it means that that the primitive in its abstraction form; what (what) has been referred to a specific place in [4]; what (who, where) which is; it (p, Barcelona). Through tagging these changes, we can assume that speaker m in context of 'proposing' regularly communicates with the communicative act request-whenever (i.e., message number (3)) followed by inform-ref (i.e., message number (4))). On the other hand, its primitives in content in association with these communicative acts commonly refers first to *what* is proposed (i.e., version map), followed by associating this to location/place.

In (Bartlett 1932) it is suggested that people are always actively remembering, and some findings and reports (Clancey 1997, Binti Abdullah and Cerri 2005) suggest that in remembering, we are always re-sequencing and re-enacting information situated in our punctuated context. If we apply this to the tagging of the associations and changes of the re-sequencing of both the communicative acts and its content, it should enable us to categorize contexts. For example, through observing the generality that emerges through tagging the re-sequencing of these structures, we can find some properties that inform us that this is where the context of breakdown has started in the conversations. Another potential refinement to the categorization follows the hierarchy on learning and communication by (Bateson 1972) on understanding how contexts are punctuated based on the ideas of context marker, items and sets of alternatives. These primitives are then categorized from items to set of alternatives. For example, the primitive "version map" could be labeled as an item in the set of alternatives "discuss documents" and so on (i.e., based on observation and regularities).

Much of the future work would be geared towards this direction for analysis.

Perspectives

In this paper we introduced the beginnings of a methodology and framework for understanding *breakdown* situations. Understanding these situations has broad implications for many areas of computer science. On an abstract level, understanding the state-of-mind of a user can be directly applied to user interfaces that can adjust themselves to the expectations of each specific user. On a specific, targeted level, understanding when faults occur in a communicative tool such as BuddySpace IM, allows the application client to help the user focus on what is important. At present, in order to model specifically the assistive technology, we are analyzing more corpuses to precisely guide us what and how the system can help the discrepancies between user and the computer system.

The specific method in this paper grants credence to the idea that something more than a static, rule-based method of recognizing and *punctuating* contexts is needed. The authors look forward to exploring the direct application of the activity states theory to this problem. As discussed in our results, we are also actively exploring integrating activity states theory with probabilistic learning methodologies for categorization, which could be used to categorize various conversational contexts that arise.

Acknowledgment

We would like to thank the European Community under the Innovation Society Technologies (IST) programme of the 6th Framework Programme for RTD - project ELeGI, contract IST-002205 for the permission to publish the excerpted chat logs. Also to the EleGI WP6 group members for their cooperation and constructive opinion on the research direction; Stefano A. Cerri and Philippe Lemoission from the LIRMM-University Montpellier II, France. To Marc Eisenstadt from the Knowledge Media Institute, Open University, The UK; and Joost Breuker from SWI, University of Amsterdam, The Netherlands.

References

Bateson, G. 1972. Steps to an ecology of mind. Chandler and Publications Co.

Bartlett, F.C. 1932. Remembering. A Study in Experimental and Social Psychology. Cambridge University Press.

Binti Abdullah, N.N, and Honiden, S. 2007. A topic categorization approach for understanding IM activities on the Web. Multiple and Ubiquitous Interaction- book of abstracts. Christina Brodersen, Susanne Bodker, and Clemens N. Klokmose (Editors). DAIMI PB-581, University of Aarhus, Denmark.

Binti Abdullah, N.N. 2006. A theoretical framework for the analysis of actual human collaboration on the Web. Ph.D. diss., Dept. of Computer Science, Univ. of Montpellier II, Montpellier, France.

Binti Abdullah, N.N., and Cerri, S.A. 2005 Analysis and Synthesis of Agent Communicative Behaviors. Towards the Learning GRID: advances in Human Learning Services. Edited by P. Ritrovato et al. Frontiers in Artificial Intelligence and Applications. Taylor and Francis, New York, USA. 19: (9-10): 1015-1041.

Breuker, J.; Cerri, S.; Dugénie, P.; Eisenstadt, M. and Lemoisson, P. 2006. D20 Conceptual and Organisational Framework for Conversational and Collaboration Processes. Intern Document ELeGI Project. Unpublished

Clancey, W.J. 2001. What we are learning about communication protocols. http://www.spaceref.com/news/viewsr.html?pid=3280

Clancey, W.J. 1997. Situated Cognition: On Human Knowledge and Computer Representations. Cambridge University Press.

Connell, J., Mendelsohn, G, Robins, R. and Canny, J. 2001. Effects of communication medium on interpersonal perceptions: Don't hang up on the telephone yet!, In Proceedings of the 2001 International ACM SIGGROUP Conference on Supporting Group Work, 117-124. Boulder, CO.: ACM Press.

Eisenstadt, M., Komzak,J., and Cerri, S. 2004. Enhanced Presence and Messaging for Large-Scale e-Learning. In the Proceedings of the 3rd International Conference Symposium on the Tele-Education and Life Long Learning (Teleduc'04).

Fipa-Acl. 2002. Foundation for Intelligent Physical Agent. Agent communication Language: Communicative Acts Specification Library.

HMP. 2001. NASA Haughton-Mars Project. Link:http://www.marsonearth.org/about.

Isaacs, E., Walendoswki, A., Whittaker, S., Schiano, D.J., and Kamm, C. 2002. *The Character, Functions, and Styles of Instant Messaging in the Workplace*. In the Proceedings Computer supported cooperative work. CSCW`02, 11-20 New Orleans, Louisiana.: ACM Press.

Leont'ev, A.N. 1978. Chapter 2: Psychic Reflection. In the Activity, Consciousness, and Personality. Psychology in the USSR, Problems of Dialectical Materialism. Progress Publishers.

Nardi, B., Whittaker, S. and Bradner, E. 2000. Interaction and outeraction: instant messaging in action. *In the Proceedings of Computer supported cooperative work.* CSCW '00, 79-88. Philadelphia, PA.: ACM Press.

Stanford. 2003. Link: http://plato.stanford.edu/entries/type-theory/