Semi-structured interface in collaborative problem-solving

Patrick Jermann, Daniel Schneider
TECFA (unit of Educational Technology),
School of Psychology and Education Sciences
University of Geneva, Switzerland

Abstract

The aim of this research is to observe the usage of a semi-structured communication interface in collaborative problem-solving in the light of recent research on collaboration.

Ten mixed pairs have to set up the schedule of a working conference by taking several constraints into account. The subjects can act on a shared problem representation and communicate synchronously by two modes inspired by previous work on structured communication interfaces. In the 'free' mode, they can type a message in a bare text field. In the 'structured' mode, utterances are issued either by pressing a button or by filling in text fields preceded by sentence openers. Data was automatically collected by a MOO[1] which served as backend for a graphical user interface written in JAVA.

The results show that pairs who use the 'free' communication mode more than the 'structured' mode, produce more 'off-task' statements than the pairs who prefer the 'structured' mode. Another finding is that when strategic issues are discussed during the first third of total time, there is a tendency to be more accurate in placing the events at the right place at first trial.

Introduction

The distributed cognition label includes various approaches which give social and cognitive processes a different weight in the analysis of computer mediated activity. A common idea to these approaches is that the object we observe is no more restricted to 'knowledge in the head' as it was the case in traditional cognitive science, but includes several agents and the artifacts they use to mediate their activity. Authors like Hutchins (1995) and Smith (1994) analyse the ongoing activity in a group by looking at the knowledge flow across different artifacts.

Smith observes that access to and processing of knowledge in a computer-mediated environment happen either by individuals or by groups. He presents a knowledge typology based on the information's persistence of display, i.e how long it can be viewed on a particular medium (Dillenbourg & al., submitted). Tangible knowledge is persistent and manipulated by individuals. It comprises the final product (called 'target product') and 'instrumental products' which support the group's work on the target but are not part of it. Intangible knowledge is not physically accessible and constitutes the shared and private knowledge people have in mind. Between these two types of knowledge, lie ephemeral products which temporarily take a physical form. These allow the transformation of knowledge from one type into the other by the contributions of several members of the group. A semi-persistent media is associated to this type of knowledge. Following this terminology, the system we describe later on comprises a tool for persistent target knowledge, the so-called problem representation and another tool for semi-persistent ephemeral products, the so-called communication interface (See "Interface design" section).

The application of the distributed cognition approach to the design of collaborative applications consists in matching the characteristics of a computer tool to the knowledge type it should mediate. The notion of affordance (Pea, 1993) refers to the same kind of relations: *"a door knob is for turning"*. However, a particular tool can serve different communicative or problem solving functions as shown in Dillenbourg & al. (submitted). The allocation of tools to functions can vary across pairs and within a pair during the collaborative process.

1. MOO stands for 'Multiple User Dungeons Object-Oriented'. MOOs are multi-user, text-based virtual reality servers. For detailed information, see http://tecfa.unige.ch/tecfamoo.html.

Roschelle and Teasley (1995) defined collaboration as a "Coordinated, synchronous activity that is the result of a continued attempt to construct and maintain a shared conception of a problem". More precisely, the 'Joint Problem Space' (JPS) "[...] is a shared knowledge structure that supports problem solving activity by integrating *(a)* goals *(b)* descriptions of the current problem state, *(c)* awareness of available problem solving actions, and *(d)* associations that relate goals, features of the current problem state, and available actions". A computer supported problem solving environment can reify the JPS by associating physical representations to the elements which constitute it.

We defined a simple content typology with three categories: 'task', 'strategy' and 'interaction management' (See "Coding" section). The communication interface allows two modes of expression, 'free' and 'structured' (See "Interface design" section). Our hypothesis is that there is a preference to use a particular interface mode when expressing a particular content type. Previous work by Baker & Lund (1996) has shown that when using a 'chat' interface only subjects showed up more 'off-task' contributions than if they used a semi-structured, so-called dedicated interface. In our case, subjects have to choose a mode when producing an utterance. We have also computed a monotony index which reflects the style of task realisation (See "Monotonic reasoning" section). A major question is how can we relate a higher monotony (few tuning actions needed to accomplish the task) to the content type of utterances and consequently to the interface usage?

The long term goal of our work is to design software agents capable of identifying a problem solving phase when observing the interface usage and interaction parameters. Such agents could then orient and supervise the usage of the interface in a way to facilitate the interaction. In a distributed cognition perspective, we don't model a student anymore, but a larger system including people and artifacts. To do so, we first have to identify characteristics of artifacts such as media persistency and correspondence between artifacts and knowledge types (See Dillenbourg & al., submitted).

Experimental setting

Task

The task consists in setting up a conference schedule by taking three constraints into account. Subjects receive a list of 16 talks, each one dealing with a particular theme ('Psychology of Art', 'The Proximal Zone of Development' and 'Speech and Cognition'). For every talk, the speaker needs some specific technical support (translation cabin and/or headlight projector). Three conference rooms are available amongst which only one has a translation cabin and a headlight projector. The two other rooms have only a headlight projector. Finally, every speaker has a status: 'Invited Speakers' talk before 'Keynote Speakers', who in their turn talk before the 'Short Presentations'.

Subjects were told to set up the schedule in such a way that a) two talks from a same theme should never be presented at the same time, b) talks which needed to be translated should be held in the appropriate conference room and c) the precedence of speakers should be respected.

Interface design

The interface comprises two distinct tools: the first is dedicated to the problem representation and the second is dedicated to the communication.

The problem representation tool

This tool consists of three windows, one for each conference room (See figure 1). Each window comprises 21 time slots which can be filled in with events. The subjects use command buttons to act on the problem representation. Once an event is created, it can be moved up and down, edited (shows the event window) or deleted. We had to implement a blocking mechanism to maintain consistency of data displayed to the subjects. As a consequence, some time-slots are temporarily inaccessible to other users when being edited by someone. We comment the consequences of this restriction on the collaborative process later in this paper.

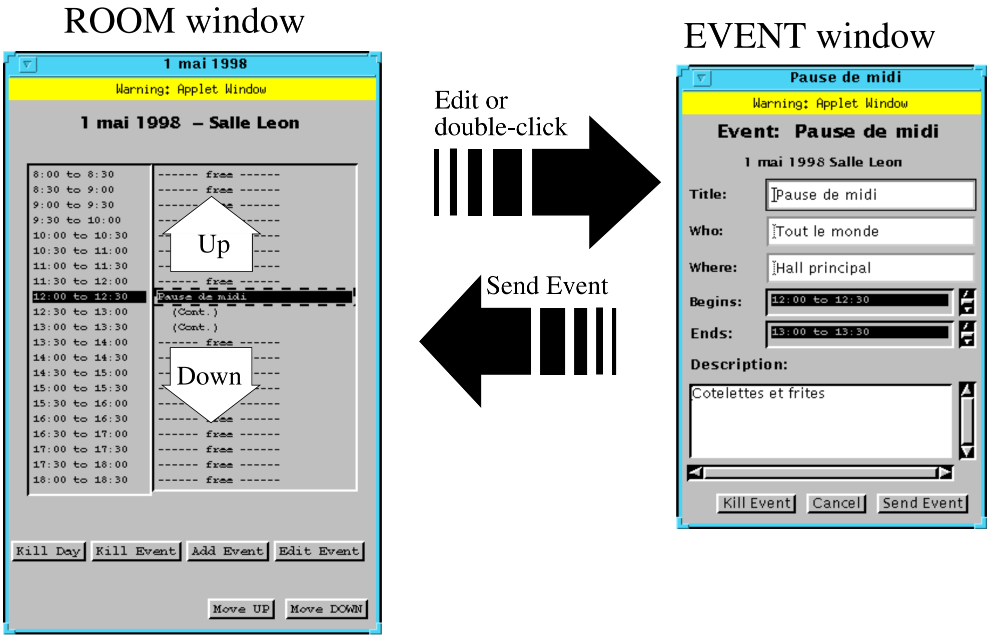

Figure 1: Problem representation tool

The communication tool

The second tool of the interface is dedicated to communication (See figure 2). Subjects either fill in text fields or use graphical buttons to issue utterances. A dialogue history contains the utterances made so far and allows the subjects to react to a particular one by selecting it. The answer is then indented and put after the utterance it refers to. When no item is selected in the dialogue history, the new utterance is put at the end of the list. This way of structuring discussion is common in WWW-based (See for example Hypernews[2]) and USENET discussion systems where messages are structured into threads.

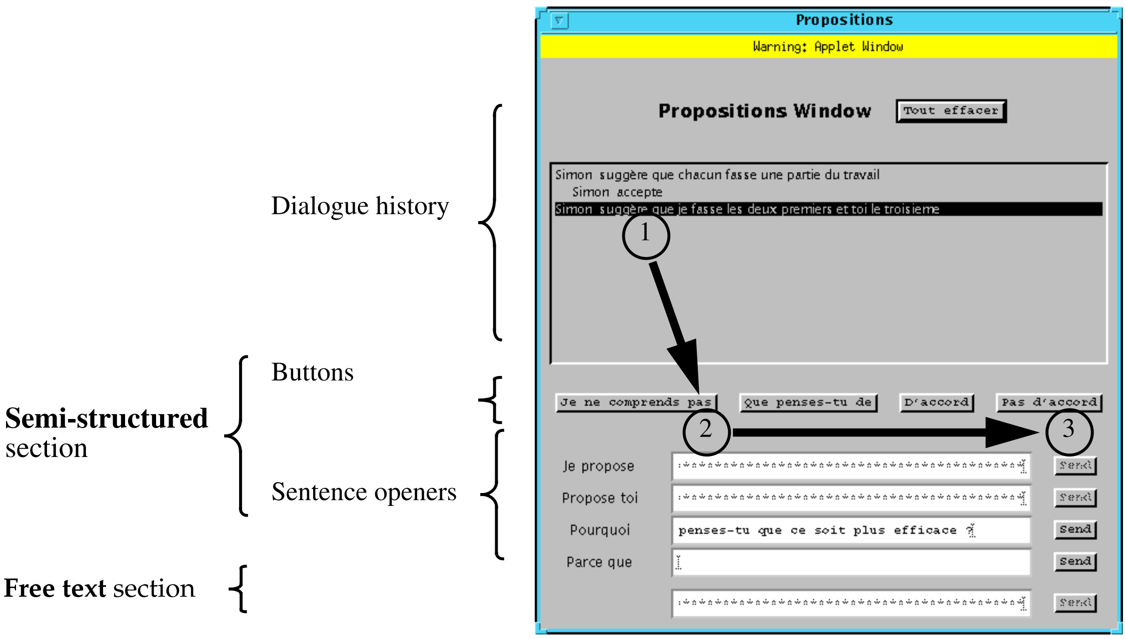

Figure 2: Communication interface
The arrows show the three operations necessary to produce an utterance. The subject has to:
1) Select the utterance he reacts to in the dialogue history 2) Fill in the text field after the sentence opener 3) Validate the contribution by typing 'Return' or by pressing on the 'Send' button.

When issuing an utterance, subjects have the choice between a so-called 'semi-structured' and a 'free-

2. Hypernews: http://union.ncsa.uiuc.edu/HyperNews/get/hypernews.html

text' mode[3]. The former consists of four buttons labelled '*I don't understand*', '*What do you think*', '*I agree*' and '*I disagree*' and of four text fields preceded by the labels '*I propose*', '*You propose!*', '*Why*' and '*Because*'. The latter is simply a text field where subjects type in a phrase.

All actions and utterances performed by the subjects are automatically recorded by the MOO-server. This makes the data collection much easier than coding a video transcript and above all, makes the data acessible to software agents which could compute various indices during the collaborative process.

Coding

We have defined three types of contents inspired by Dillenbourg and Baker's negociation spaces (1996) and by Bunt's (1995) distinction between task-oriented and dialogue control acts.

- *Task* related: utterances which contain a proposal for an event's time or room allocation, as well as reflections about the completion of the task.
- *Strategy* related: utterances concerning the distribution of roles and the negociation of a plan.
- *Interaction management* related: utterances used for requesting attention, marking the beginning and end of work as well as remarks about the interface and the way to use it. This type of utterances is sometimes called 'off-task'.

Results

General observations

Ten mixed pairs took part in the experience. The average time needed to complete the task is 75 minutes (σ=29.5). All the pairs completed the task.

The subjects were allowed to use a sheet of paper in order to write down whatever they liked. In most of the cases, the subjects used the instruction sheet containing a list of 16 events by adding a checkmark near the events they had already put on the screen. They also wrote down the time and room they had chosen for a particular event. It seems that a clear representation of work completion was lacking in the problem representation window. The system could be enhanced in this respect by adding an automatically updated checklist which following Smith' terminology corresponds to tangible instrumental knowledge.

Co-action

The co-action index reflects the amount of shared manipulation of an event, e.g. S1 creates an event and S2 moves it later on to match a new constraint of the task. This index is very low due to the blocking mechanism we had to implement for technical reasons. The subjects were presented an error message when attempting to modify an event at the same time which of course discouraged them to act on the same events. As a consequence, negotiation could hardly happen through actions on the problem representation in our experiment. A whiteboard would better fit the needs of co-manipulation of the events because these could be placed one over the other, put aside for a moment without hampering the computational representation of the task. People would still manipulate one event at the time, but they could do it simultaneously on two events belonging to the same room.

Interface usage and content type

The average distribution of all utterances (N=1039) across content types is as follows: 58% belong to the task category, 22% concern strategy and 20% concern management. The structured section of the interface is more frequently used than the free section. (2-tailed Student test: $p<.05$). It appears that the usage of the free versus structured interface is not independent from the type of content being uttered (χ^2 $p<.000$). The management is more often expressed by using the free section while the task and strategy contributions are more often expressed by the structured one. (See figure 3).

The independence test doesn't tell us anything about relations between specific values of the content

3. These will be called *free* and *structured* from now on.

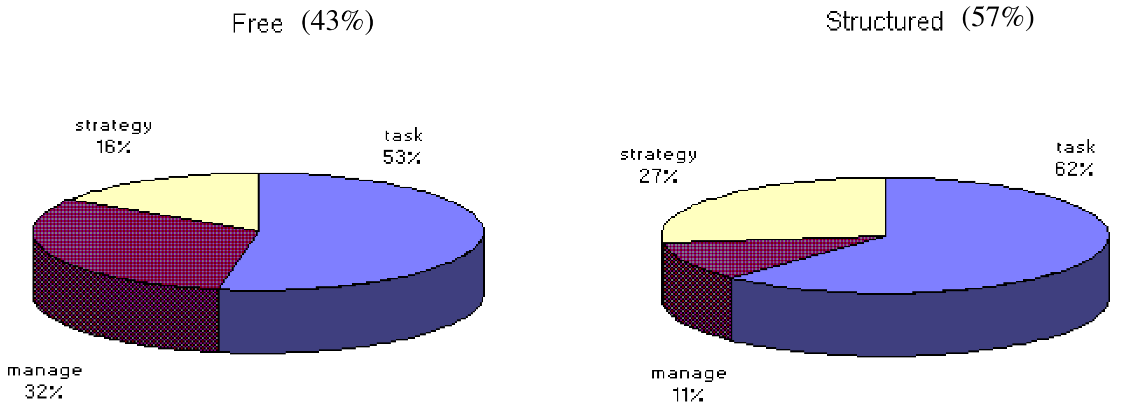

Figure 3: Expression mode and content type

and the interface usage frequencies. The question is: does a higher freqency of use of the structured interface go along with a higher frequency of task-related utterances?

We can notice following correlations:

Table 1: Expression mode and content type [a]

	Task frequency	Strategy frequency	Management frequency
Structured frequency	**.7420** ($p<.05$)	**.8284** ($p<.05$)	.6185 (NS)
Free frequency	.3393 (NS)	.6022 (NS)	**.7455** ($p<.05$)

a. Numbers indicate Pearson correlation coefficients. 2-tailed significance is indicated in brackets.

We observed that pairs tend to prefer one section to the other and keep this preference during the whole interaction. The correlations we just presented show that pairs which prefer the free section of the interface also tend to produce management utterances with a higher frequency and that pairs who tend to prefer the structured interface, produce task and strategy related utterances with a relatively higher frequency (See Table 1: “Expression mode and content type”). Baker & Lund (1996) showed that more ‘off-task’ utterances were produced by subjects using a chat interface than by subjects using a structured, dedicated interface.

Monotonic reasoning

We computed a monotony-index which reflects the amount of ‘tuning’ actions necessary to put an event in the right place once it has been created[4]. The more tuning actions are necessary, the lower will the monotony index be. At first sight, a better accuracy of placement (a high monotony) mainly should be due to a better and more intense negociation between the subjects. Several constraints have to be discussed when choosing a time slot for an event: are there events belonging to the same theme positioned at the same time in another room? Is the speaker precedence respected? Is the room equipped with the necessary devices? Who will create the event? If all these issues are discussed before any action on the problem representation takes place, the probability of having to modify the events afterwards should be lower than if subjects don’t negotiate. In statistical terms, the correlation between monotony and frequency of task or strategy related utterances should be positive.

Now this is not true in our experiment: monotony is negatively correlated to the frequency of strategy related utterances ($r=-.5686$, 1-tailed $p<.05$), to the frequency of task related utterances (-.8485, 1-tailed $p<.05$) and to the frequency of use of the structured interface ($r=-.5867$, 1-tailed $p<.05$). This indicates

4. Monotony is computed as follows: (100 - %move -%kill-%update) where %move corresponds to the use of the ‘*move up*’ and ‘*move down*’ buttons, %kill is the use of the ‘*kill event*’ button an update is the use of the ‘*edit event*’ button. (See figure 1)

that some pairs did not verbally coordinate their actions when positioning the events on the problem representation. Two hypothesis could explain this unexpected phenomenon. Either only one subject was acting while the other was watching or the subjects split up the task before starting to act on the problem representation.

1) Collaborative activity implies tuning activities, i.e. solutions are constantly reelaborated and renegociated in the interaction whereas an individual executes his plan once he has fixed it up without further questioning. Pairs in which a strong leader emerges and takes control of the problem solving should then be less monotonous because one individual does most of the work. We computed an asymmetry index[5] on actions and communication to test this possibility but monotony is not significantly correlated to any asymmetry index. However, the action asymmetry index is negatively correlated to the frequency of actions (r=-.695, 2-tailed p<.05) which means that the more symmetric a pair is, the more actions are produced.

2) Collaboratively acting on the problem representation was difficult, almost impossible due to a blocking mechanism (See "Co-action" section). Pairs who did not split up the task and tried to do the positioning of events together must have had more coordination problems. Pairs who split up the task and distribute the roles could then act individually on the problem representation and be more accurate. Our assumption is that strategy has to be discussed in the beginning of the interaction. We chose to partition the interaction into three phases in order to check if early strategic negociation is related to monotony. We characterize these phases as follows:

a) a management and strategic oriented phase where subjects initiate contact and distribute roles and work. (high management and high strategy proportions)

b) a task realisation phase where we have a high proportion of task related utterances

c) a control and congratulations phase where task and management proportions are high ("au revoir" phase)

This distribution is supported by our data when looking at the mean proportions of utterances by type (strategy, management and task, see figure 4). Of course, representing the mean values hides the specificity and variability of pairs.

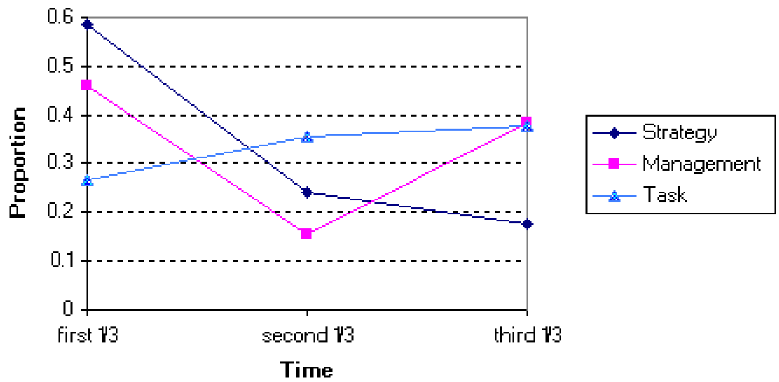

Figure 4: Content type across time

How can we explain that some pairs have a lower monotony? We should have two types of interaction: a first one where strategic utterances are spread across the whole interaction time and another where strategy is discussed mostly at the beginning of the interaction.

We computed the relation between total monotony and proportions of strategic utterances per thirds of total time. The critical correlation with respect to our hypothesis is between the proportion of strategy during the first third of time and the monotony (See Table 2: "Monotony and proportions of strategy"). It indicates that when a high proportion of all strategic utterances are produced in the first third of the total time, the total monotony will be higer too. The proportion of strategic utterances in the two next thirds of time are negatively and non significanty correlated to monotony (This lack of significance is partly due

5. The asymmetry index is computed as follows: ABS ((contributions S1-contributions S2)/(contributions S1 + contributions S2))

to the small number of pairs we have at our disposal).

Table 2: Monotony and proportions of strategy[a]

	Strategy first 1/3	Strategy second 1/3	Strategy third 1/3
Monotony	**.6463** ($p<.05$)	**-.5365** (NS)	**-.5308** (NS)

a. Numbers indicate Pearson correlation coefficients. 2-tailed significance is indicated in brackets.

As presumed, the pairs which spoke about strategy in the beginning end up with a lower monotony index. We don't think that these pairs are better collaborators than those which did not split up the task or the roles. However, given the properties of the system which hampered joint manipulation of events, splitting up the task seems to be optimal.

Conclusion and further work

In this contribution, we showed that the content type of utterances is related to the interface type used to produce them. Task and strategy related utterances are produced more often through the so-called structured interface while management contributions are more often produced through the so-called free interface. Pairs which tend to prefer the free interface produce more management utterances than pairs which prefer the structured one (See "Interface usage and content type" section).

We then computed an indicator of a pair's solution path. When events are put correctly at first try into the schedule, monotony is high. On the contrary, when tuning actions like moving the event up and down or deleting it are necessary, the monotony is low. Monotony is related to the proportion of strategic utterances produced during the third of interaction time. We propose the following interpretation: when splitting up the work and distributing the roles at the beginning, people accomplish the positioning as an individual would do it, by facing less coordination problem. This seems to be an optimal behaviour in our experiment as the simultaneous manipulation of events belonging to the same room was difficult for technical reasons (See "Monotonic reasoning" section).

As an AI&Ed application of the 'knowledge flow' approach presented in the introduction, we could detect phases in problem solving by observing a switch in the interface usage. Advice could then be given on the best way to use artifacts for a particular sub-task. Next step in this direction are Moobservers, non-intentional agents which consist of a recognising module which information is more or less displayed in a raw graphical manner to the collaborating subjects. A major question is whether the socio-cognitive behaviour of the peer changes when it gets information about its interaction.

References

Baker, M. & Lund, K. (1996) - Flexibly structuring the interaction in a CSCL environment. Proceedings of EuroAIED 96, Lisbon, September-October 1996.

Bunt, H. (1995) - Modeling Interaction in Intelligent Tutoring Systems. In R.-J. Beun, M. Baker, M.Reiner (Eds.), Dialogue and Instruction, NATO ASI Series. Berlin: Springer.

Dillenbourg, P., Jermann, P., Schneider, D., Traum, D. and Buiu, C. (submitted to AI&ED 97) - The design of MOO agents: Implications from a study on multi-modal collaborative problem solving.

Hutchins, E. (1995) - How a Cockpit Remembers Its Speeds. Cognitive Science 19, 265-288.

Jermann, P. (1996) - Conception et analyse d'une interface semi-structurée dédiée à la co-résolution de problème. Master thesis, TECFA, School of Psychology and Education Sciences, University of Geneva, Switzerland.

Roschelle, J. & Teasley, S.D. (1995) - Construction of shared knowledge in collaborative problemsolving. In C. O'Malley (Ed.), Computer-supported collaborative learning. New York:Springer-Verlag.

Smith, J.B. (1994) - Collective Intelligence in Computer-Based Collaboration. Hillsdale, NJ: Lawrence Erlbaum.