

Jeanne Cornillon · Duska Rosenberg

Dialogue organisation in argumentative debates

Received: 26 May 2004 / Accepted: 8 July 2004 / Published online: 5 August 2004
© Springer-Verlag London Limited 2004

Abstract This paper presents a conceptual framework for the study of social intelligence in a real-life environment. It is focussed on the dialogue organisation in argumentation, in particular how our understanding of dialogue phenomena in mediated communication may help us to support natural interaction in classroom debates. Dialogue organisation is explored in terms of the cohesive structure of dialogue that emerges as the result of information maintenance and change, specified locally by the adjacency pair and turn-taking, and globally by topic threads. We give a descriptive analysis of their interplay in real chat debates, based on lexical chains of repeated words in an attempt to clarify what feedback participants in the debate require in order to reflect on their contributions to the debate. We then show how this analysis can be used as a basis for conceptual design of a feedback adviser tool to automatically search for topics in student chat debates. Much of the data used in our analysis has been generated as part of the SCALE project¹ focussed on specifying tools for computer supported argumentation in the classroom.

Keywords Electronic chat dialogue · Lexical chain · Threaded chat · Topic · Knowledge co-construction · Adjacency pair · Interpersonal awareness · Feedback adviser

1 Introduction

One of the key challenges for computer-mediated communication is to develop tools such that the technology supports human cognitive and social processes in an effective and natural manner. In this paper we address three key issues in social intelligence design, namely the nature of a conceptual framework that will enhance our understanding of social intelligence in the context of the relationship between the technology and its human users, and in particular understanding of

J. Cornillon (✉) · D. Rosenberg
School of Management Royal Holloway, University of London, Egham, Surrey,
TW20 0EX, UK

¹SCALE (Internet-based intelligent tool to support collaborative argumentation-based learning in secondary schools) – EU Framework V project.

the group dynamics of knowledge creation. We apply these principles of social intelligence approach in the study of classroom debates that aims to inform the development of argumentation tools and their implementation in real-life use.

The motivation for this research comes from pedagogical considerations that suggest that student debates are only beneficial if they are conducted by active, well-functioning learning teams. Such teams require effective communication where each student is capable of reflexion leading to conceptual change (cf. Baker and Lund 1997; Pilkinton 1998; Soller et al. 1999). In order to meet the communication requirements of learning teams, we examine the significance of feedback and reflexion whereby students monitor their contribution to the debate and learn how to argue by reflecting on the value of individual contributions to the team effort. This understanding may contribute to the development of a multi-modal environment, developed in SCALE, with the specialised function of making the flow of the arguments visible to all participants. As the basis of dialogic argumentation or debate is to get to the core of the issue in question irrespective of the mode of expression, we explore how various claims are related to a topic by putting forward arguments in support of or against.

The multi modal environment—DREW 2003—supports dialogic argumentation by providing a range of freeware (<http://test.scale.emse.fr>) which include:

- A text editor which allows two users to update a shared text
- A graph editor where two users construe a graph together in real time (JigaDrew)
- Chat, free and structured facilities (ALEX).

One of the aims of the SCALE project was to develop a feedback device for ALEX, which would inform the participants on the state of the unfolding debate. In this paper, we consider in what ways and to what extent our understanding of the linguistic features of classroom debate could help towards an automatic discourse analysis that would underpin the design of the feedback advisor for ALEX and DREW.

1.1 Feedback in classroom debate

The feedback adviser's function in the context of learning in general is to take the role of the expert. Student–student dialogues and expert–student dialogues give rise to different learning styles. While peer collaborative interactions generate conflict and reflexion, asymmetric dialogues between an expert who takes the role of enquirer and the student that of the experimenter give rise to reflexion and conceptual change (Pilkington 1999). In particular, this approach takes learning as the acquisition of new, or more refined, mental models. New or refined models result from cognitive conflict, that is, two competing existing models are either reconciled, or one is chosen over the other. If we consider learning to be the successful outcome of a cognitive conflict, then we would expect an expert to help the student through the impasse with the right probe or question.

Another view of the role of the expert in classroom debate is to support collaborative learning (Soller et al. 1999). In collaborative learning the process by which students are jointly participating leading up to the solution *is* the central learning goal. By putting the emphasis on assessment of collaborative learning interactions, pedagogical studies in the field of collaborative learning

have lead to the definition of an ideal process of interaction to promote effective collaborative learning (Baker and Lund 1997). The ideal process of interaction is a reflective activity where both partners not only provide a solution to the problem, but also attempt to explain, evaluate, and/or justify solutions. It is also a symmetrical process for participants to regulate how they work together, how they respond to each other's feedback and how they co-ordinate their diverse perspectives.

Collaborative argumentation in the classroom with the aim of working towards the resolution of disagreements and the assessment of differing viewpoints is a special case of collaborative learning where the process of knowledge co-construction is the main object of the task-focussed activity. Expert-monitoring of the student interactions, as they happen, can help to achieve the goals of the pedagogically motivated debate by enabling the students to reflect on their participation in the debate, and possibly modify it as the debate progresses.

1.2 Dialogue organisation: an overview

Any form of discourse including chat debate has several dimensions, but we shall focus here on information flow which is the dimension in dialogue that makes "one stretch of discourse about something and the next stretch about something else" (Brown and Yule 1974) in order to capture the knowledge co-construction aspect of the debate. Initially in the project, we wanted to develop a prototype based on a set of argumentation rules with applications to a business environment (Saeedi and Sillince 1999) to analyse classroom debate, but, although the formulation of an argument is about content structuring, the level of granularity of the analysis did not provide us with any information as to how information is organised in the team, how information is passed on from one individual to the other, in short, how the information is shared within the interaction. Speech act analysis as applied to dialogue can track how people organise the interactive event, but its scope appears limited to the exchange structure, which is a highly local cohesive structure, and, in that sense, cannot capture the information flow of the debate. Lexical distribution in text is said to encode larger chunks of information, but it has only been applied to monologues. We show here however that lexical chains can successfully capture the co-construction aspect of knowledge building in dialogue. We also distinguish between lexical semantic relatedness and lexical repetitions, which we argue have anaphoric properties in order to further motivate our analysis of information flow in terms of lexical chains of repeated words.

2 Information flow in dialogue

In the literature we find two views about dialogue information flow. The first approach is to look at dialogue as texts (i.e. a monologic structure) where topics are encoded by lexical chains at text level. The second view proposes that dialogue structure is essentially different from that of the monologue and should be given a specific treatment in terms of moves, turns and exchange structures encoded by speech acts. In this paper we present a mixed model of chat debate made up of overlapping cohesive structures, the turn and the adjacency pair on

the local level, and, on the global level, topics encoded by lexical chains in order to offer a more precise characterisation of threading in computer mediated communication where multiple topics appear to coexist side by side.

2.1 Cohesion in discourse

2.1.1 Topics in monologic structures

If we view the chat debate as a monologue, then information content and its organisation is accounted for in terms of topic change and maintenance. In order to help the hearer find an interpretation and minimise his effort in accessing the right linguistic context, we hypothesise that the information flow is structurally regulated. There are many models of utterance interpretation, mainly used for pronominal resolution, because pronouns crucially rely on the identification of antecedents, the items of information necessary for the interpretation of pronouns. In general, information structure is said to be subject to a linear modification principle. Firbas (1986) describes linear modification as “a gradation in importance of information which, provided that there is no interfering factor, gradually increases with movement towards the end of the sentence; the last element of information becoming the most important because it completes the communicative purpose of the sentence”.

We focus here on the topic-focus articulation. The topic-focus articulation divides the sentence into a flat intra sentential structure made of background (theme/topic) and foreground (rheme/focus) information organised from left to right. This left to right structure is constrained by language grammar hence the linear modification principle does not always hold.² The evidence for such a flat information structure is based on the mirroring effects that occur in language such as repetitions and matching pairs. A matching pair is typically a sequence such as a question answer pair or a contrasting pair, where the structure of the second can be derived directly from a transformation (e.g. deletion, substitution, movement) of elements from the first. Matching pairs or contrasting pairs can also serve as diagnostic in order to identify the focus element. For instance the contrasting pair, *I turned left but you turned right* only allows the interpretation where both the subject (in violation of the linear modification principle) and the adverbs are focus elements.

- (1) I (focus) turned (topic) left (focus) but you (focus) turned (topic) right (focus) (Carlson 1983) Furthermore, each sentence’s functional structure is derived from an abstract question forming with it a matching pair.
- (2) Who would trust an idiot?
An idiot (focus) would trust (topic) an idiot (focus)
(Carlson 1983)

Matching pairs also show that repeated words are motivated by the underlying topic focus structure. In the example above, the second *idiot* has been retained and not pronominalised (i.e. *an idiot (focus) would trust him (topic)*). The

²Nevertheless Mc Carthy (1998) argues that linear modification is much more marked in dialogue.

repeated lexical item *idiot* is used instead of the pronoun *him* to put it into contrastive focus with the other element *idiot*. Such NP is modelled not as an entity but a variable referring to a set of entities (e.g. Rooth 1992), forcing the hearer to access an interpretation whereby the element in focus is not interpreted in isolation, but considered alongside a set of alternatives making its occurrence more likely in “communicative situations such as disagreements, assessments, particularly negative assessments and overt recognitionals (know + NP)” (Fox 1988). In general, repeated NPs are like pronouns in that they have antecedents, but are motivated by information structure requirements. A repetition can also be used instead of a pronoun on the basis of a resource-limitation on the set of (linguistic or non linguistic) referents that is stored in short term memory. Walker et al. (1998) argue, for instance, that the acceptability of overspecified NPs depends on whether the utterance is perceived as initiating a new discourse segment, as exemplified in the following list, where the segment (3c) and not (3b) is an acceptable continuation of (3a):

- (3a) Suzan gave Betsy a hamster
 She reminded her that hamsters were quite shy
 (3b) *?and then Suzan laughed
 (3c) and then Suzan left

Further, in the absence of an immediate non-linguistic context as in chat, we note that repetition has a higher incidence than in spoken dialogue with referential NPs replacing the use of the deictic pronouns, *it*, *this* and *that* in English which perform different functions in the organisation of topic and focus: “*it* signalling continued ongoing topics *this*, new or significant focus and *that*, a variety of distancing or marginalising functions” (Mc Carthy 1998).

At textual level, information flow is mapped out through lexical strings. “Lexical strings with minimally two lexical words tell us what is being talked about in the text; each major lexical string indicating a topic or part of a topic” (Egins 1994). There are two types of lexical chains. The lexical chains composed of words related through semantic relations such as synonymy/antonymy (*but the world is changing rapidly/ communication would be too slow and inefficient*) and word associations (*invasion of privacy/employees rights*) are context dependent and additionally require knowledge about the level of granularity of the analysis. So, for instance, if we are tracking topics on *email privacy at work* on a set of random texts, then *invasion of privacy* and *email* vs. *invasion of privacy* and *IT* are highly related semantically, but if we look at a set of texts on *email privacy at work* it isn’t necessarily the case. On the other hand, lexical chains of repeated phrases (*reading our emails/read your emails*) or words (*our workers/working for an/monitor our work/quality of work/making my work levels drop*) form anaphoric chains, as motivated by the information structure requirements discussed above, and those can be utilised to track cohesive segments at text level.

Information structure theory using the topic focus articulation at utterance level and lexical chains at discourse level has been applied successfully to the analysis of monologues, but they do not take into account the dialogue exchange structure that is discarded or seen as a continuation of a monologue (see Walker et al. 1998). However, lexical word repetition is said to be especially common in spontaneous dialogue (de Beaugrande 1980) showing the importance and distinctiveness of the information structure in dialogue.

Table 1 Dialogue extract analysis (Eggs 1994)

Dialogue move	Speech function	Exchange structure
How did you have blood given before 36 times	Question Answer	Exchange
It makes me go funny just thinking about it You have never done it obviously	Statement Clarification	Exchange
Oh no	Response to clarification	

2.1.2 The exchange structure

If we view the chat debate mostly as a dialogue then information content and structure are usually defined in terms of speech acts framed by the exchange structure³ whose basic unit is the turn.⁴ A typical example of such dialogue analysis is given in Table 1:

Although the turn is an important structural boundary in dialogue, another local feature of dialogue is the local sequential effect or adjacency pair (Sacks et al. 1974). “An adjacency pair consists of two utterances produced by two different speakers and positioned adjacent to each other” (Tsui 1991). Crucially, the adjacency pairs are composed of speech acts, which locally trigger another and span over two turns. Speech act theory (Austin 1962; Searle 1969) has evolved from a description of a small set of verbs described as performative verbs (i.e. *pronounce*) to an abstract categorisation of the phenomenon of the sentence illocutionary force (e.g. question, answer, explanation, assertion, acknowledgement). Speech act theory was initially put forward by Austin (1962) in reaction to truth conditional semantics to account for the fact that ordinary language function is more than describing outside reality, it also performs an action on it. An example of such action performing verbs or performative verbs is given by Austin (1962) who shows that an utterance such as *I now pronounce you man and wife* rather than describing the world as it is, can dramatically change your life if a priest utters it under appropriate conditions. The attractive feature about speech acts in the context of dialogue analysis is that it captures the adjacency pair effect or the larger exchange structure and, ultimately, it is hoped the information flow in dialogue. From a cognitive perspective to human communication there is a sense in which a speaker cannot fail to be polite (Grice 1975) or exchange awareness of other conversation participants by answering a question, providing an explanation to a query and so forth as illustrated in bold in Table 2.

The exchange structure can appear under various guises, and we give a few examples of the variations that may be encountered. The exchange structure, for instance, can be broken down by inverting the second part of the adjacency pair with the rest of the sequence. Focussing on Jack’s contribution we have the acknowledgement of Oliva’s contributions in bold following Jack’s main arguments (Table 3). Line 8 refers to 6, and 12 to 10, while 7 and 11 are Jack’s new arguments.

³“The exchange structure is all that comes for instance between a question and answer” (Owen 1975).

⁴The turn corresponds to the entire contribution of a speaker. In our chat examples, this means that returns of carriages (press enter) are recorded as part of a participant’s turn.

Table 2 Extract from a chat dialogue history (Royal Holloway experiment 2002)

-
1. Oliva > Well anyway, e-mail might be one disadvantage but there are good things about IT and IS as well Oliva > for example what we are doing now
 2. Jack << **IU agree** Jack << But we have to think about all aspects of the system before we implement it
 3. Oliva > **Of course** Oliva > Some sort of system is however needed.
-

Table 3 Extract from chat dialogue history (Royal Holloway experiment 2002)

-
4. Oliva > An system can help managers in many ways: **time saving, it reduces costs**
 5. Jack << But we should be focussing on control aspect of the system
 6. Jack << **I am sure it saves**
 7. Oliva > control over correspondence with customers (not personal e-mails) > **A company who does not have a Web today, is in a quite weak position, do you agree**
 8. Jack << So do you think IT or IS does only good to an organisation?
 9. Jack << **Web is fine**
-

Table 4 Extract from a chat dialogue history (Royal Holloway experiment 2002)

-
1. Ceri > yes, it is very likely as these days many things are done on the Internet and saved in the server...
 2. Helen << once all the useless information is out of the way
 3. Helen << they can gain access the different information as well
 4. Ceri > which means that a manager sees virtually everything within his perspective...
 5. Ceri > implying tighter central control....?
-

Another departure from the adjacency pair is sentence continuation taking place over several turns as in a monologue. In this case, the dialogue structure of turn taking is somehow blurred and it may seem more natural to analyse the exchange in Table 4 as one turn.

To sum up so far, if we define a unit of meaning as a move, moves in dialogue can be cohesive around the turn structure or the adjacency pair. To give an example, in Table 5 we have an exchange structure made of three turns including turn 2 with two moves “*Hey Linda are you ready to debate?*” and two adjacency pairs with two moves each “*Hello Paul, Hey Linda*” and “*Are you ready to debate? Sure am!*”.

The turn and the adjacency pair characterised in terms of speech acts define the exchange structure, which is a complex cohesive structure unlike that of the monologue. However, it is questionable whether the exchange structure can capture efficiently larger chunks of information as evidenced by the threaded chats discussed in the next section.

Table 5 Extract from a chat dialogue history (Royal Holloway experiment 2002)

-
1. Linda > Hello Paul!
 2. Paul << Hey Linda are you ready to debate??
 3. Linda > sure am!
-

2.1.3 Threads in computer-mediated communication

In computed mediated communication, cohesive chunks of information are said to form threads. Studies have observed that “chat conversation in a chat medium flows and is intertwined as if there were multiple threads of conversations simultaneously occurring that results in multiple threads being interweaved” (Cadiz et al. 2000). In fact it is the existence of multiple threads running concurrently, often seen as a problem of coordination of perspectives due to computer mediated settings (Cadiz et al. 2000) which has resulted in chat facilities incorporating threading. In threaded chat, the chat dialogue history is not presented linearly, instead the participants are able to insert their posting directly underneath an earlier posting anywhere in the chat history by clicking on that specific dialogue turn (Table 6). The structure of the dialogue history is thus specifically designed to allow for several threads or topics (-1 underlined in text) to coexist side by side.

In threaded chat, unlike classical chat, the participants can simply index their contribution under the relevant topic heading. Thus, the conversational constraints of the exchange structure are now removed and we can encounter sequences such as the one *italicised* in Table 6 reproduced as follows in (4) and (5) which violate the conversational aspects of dialogue:

- (4) 10: Sorry, team. I crashed! Unless someone can convince me otherwise, I am ready to stop viewing Mark as a viable applicant.
 8: Notes on Mark Williams
- (5) 10: I’m picky about my coffee!
 -1: Discuss candidate #3

Table 6 Extract from a threaded chat dialogue history (Cadiz, Smith and Burkhalter 2000)

-1: Discuss candidate #2

10: I show “fair” on the creativity meter. Excellent verbal communications. Good visual and graphic skills.

7: Mark has good leadership and people skills

9: Mark’s decision making skills are so/so, and this position needs someone who is confident

7: I show his leadership and people skills to be very good. That is interesting to what you saw

10: U...

10: *Sorry, team. I crashed! Unless someone can convince me otherwise, I am ready to stop viewing Mark as a viable applicant.*

8: *Notes on Mark Williams*

8: Web Development: Good

8: Math and Number Skills: Fair

8: Very good with desktop processing and computing...

8: *Me thinks Mark is not a good candidate for this position.*

10: I agree

9: I agree..he can go make coffee with Steve

10: I’m picky about my coffee!

-1: Discuss candidate #3

10: I show that Emily is good with numbers, but her “composure” (that word again) skills are only fair. People skills are excellent.

7: Emily may make good decisions, but if she gets people really upset in doing so, we can’t have that

In (4), participant 8's contribution directly underneath participant 10 does not make any acknowledgement of 10's system crash. It is simply ignored. A similar juxtaposition of turns occurs in (5). These examples can be contrasted to a more traditional exchange structure as illustrated in (6) that enhances interpersonal awareness, and other social intelligence aspects of interpersonal communication.

- (6) Me thinks Mark is not a good candidate for this position.
 10: I agree
 9: I agree...he can go make coffee with Steve
 10: I'm picky about my coffee!

Overall, threaded chat, while keeping to its topic (discussion of candidates 2 and 3 in the extract), gives the impression of a fragmented conversation which normal chat does not exhibit due to its conversational requirements that enhance awareness of other dialogue participants.

To summarise threaded chat, dialogue histories provide us with evidence for distinguishing between the local exchange structure, enabling the exchange of awareness with others and the global topic (or thread) structure linked to information transmission, which in our dialogue extract corresponds to the organisation of the conversation according to the candidate reviewed to fill a job vacancy. Additionally, the relative lack of conversational effects in threaded chat can be related to the fact that competing topics can coexist side by side unlike in chat (or ALEX) which displays a linear dialogue history asking the speaker to attend to his or her conversational requirements.

2.2 Information flow in argumentative chat dialogues

We propose now to illustrate, using two argumentative dialogues, how lexical chains encode both the global and local cohesive structures of dialogue. We also propose that repeated words at turn level are subject to a principle similar to that of linear modification, and, furthermore, that mirroring in the linear sequencing of lexical items at turn level indicates a high level of consensus reached by two participants in a debate.

Debate is a particular form of dialogue, which can be described as "a lucid discussion leading to the resolution of a disagreement". The two dialogues analysed below are both chat debates based on the topic of electronic privacy at work where each participant is asked to take a position for or against.

In the first debate, reproduced in Table 7, the repeated words lexical chains identified are: *IT, manage, invasion of privacy, employee rights, workers, emails*. Andrew talks about *inefficiency* twice but we may take this as rewording and remove it from our list. The lexical chains, as expected, similarly to monologues encode passages in the text that displays more cohesion (e.g. *IT* and *managing with or without IT* (lines 1–4) and *invasion of privacy/email* vs. *work* (lines 8–12)). Furthermore, in both passages at turn level, two lexical words belonging to different lexical chains occur as a regular pattern, matching the presence of what could be described as two threads of conversation in the chat dialogue below.

Second, at turn level we note a regular pattern in the lexical words distribution. Focussing on lines 8–12, *e-mail and work* appear in that order for Andrew's

Table 7 Extract from a chat dialogue history (Royal Holloway experiment 2002)

-
1. Andrew << I am for **IT**
 2. Liam > ok then. Liam > but think about it...years ago when **IT** was not so sophisticated **people managed to control** and run the businesses without it.
Liam > so if they can do it then why cant we
 3. Andrew << Yes but the world is changing rapidly and we have to keep up with competitors. Andrew << Who will **manage** better with **IT**
 4. Liam > yes but surely we can **manage** without it. Liam > **IT** has its disadvantages.....
 5. Andrew << no we can't, communication would be too slow, and inefficient.
Andrew << inefficient
 6. Liam > yes but what about **invasion of privacy**.....Liam > **employee rights**.....
 7. Andrew << yes they have **rights**, I never deny my **employees rights**. We as **managers have to manage** and keep tabs on our **workers**. Andrew << Otherwise nothing would get done. Andrew << Hello?
 8. Liam > I agree but surely you can do that without **invading our privacy**. Liam > well.....
 9. Andrew << Did I say anything about **invasion of privacy**, **working** for an organisation means **working**, what do you have to hide?
 10. Liam > nothing but surely you can monitor **our work** without **reading our emails**
 11. Andrew << I don't want to **read your emails** unless your quality of **work** falls and then the first thing to do is investigate why
 12. Liam > yes but what does **e mails** have to do with my **work**. Liam > I turn up early every day and **email** then and then get on with my **work**. Liam > so I cant be wasting time or making my **work** levels drop
 13. Andrew << Then you are a goody two shoes and have nothing to worry about, but what about Dossis your friend she is bad
 14. Liam > well talk to her then. Liam > and not to me that's not my problem...every case is different
-

Employers are entitled to invade the electronic privacy of their employees

turns and *work* then *email* in the case of Liam's turns. Up until line 11, the word ordering between *invasion of privacy/email* and *work* is thus inversed at every turn and retained at every other turn. This "retain" transition before "shifting" to a new topic can be thought as indicative of local sequential effects as exemplified below. Focussing on Andrew's contribution in (7), the repeated phrase "invasion of privacy" appears to be an acknowledgement of Liam's contribution rather than an elaboration on his topic. In fact, while exchanging awareness of Liam's contribution, the topic that Andrew proceeds to discuss is that "work comes before everything else".

- (7) Liam > I agree but surely you can do that without **invading our privacy**
Liam > well...

Andrew << Did I say anything about **invasion of privacy**, working for an organisation means working, what do you have to hide?

To sum up so far, lexical chains formed by repetitions are related to information flow as in monologue, but lexical chains of repeated words seem to encode different cohesive structures in the dialogue. The repeated lexical words either relate the second part of the adjacency pair whose function is to mark awareness of the other participant(s) in the conversation *or* indicate thread or topic continuation, and this, depending on their distribution at turn level.

Turning to our second example in Table 8, at the start of the dialogue (lines

Table 8 Extract from a chat dialogue history (Royal Holloway experiment 2002)

-
1. David << well **IT** will make it more **difficult for senior managers** because **it makes it harder to control**
 2. Sandra > I think **managers can have better control**, as **IT** can help them oversee **communication** flows within the organisation
 3. David << that is true but...
 4. Sandra > but what....
 5. David << **with IT** you have **the net** and other distractions, which are **harder to control?**
 6. Sandra > well you could employ people that check up on staff and see that they don't surf **the net** when they're meant to be working
 7. David << true what about **email** abuse?
 8. Sandra > well, that's a good point, but could you imagine a firm without **IT**, how would they **communicate?**
 9. David << there r laws that mean u cannot look at employees **e-mails** I think!
 10. Sandra > well, I guess you're right, but still I think the benefits outweigh the downsides
 11. David << true about **communication** but the argument is about **control to senior managers!**
 12. Sandra > well what about computer packages like **ESS** and **DSS**, think that's what they're called.....
 13. David << well, remind me I little please?
 14. Sandra > they're like **Executive Support Systems** which help managers make better decisions and help with long term **control**, which is I guess linked to **control**
 15. David << True, so for **senior managers to have control** they have to invest a lot in **other departments** then?
 16. Sandra > why do they have to invest a lot in **other departments.?**
 17. David << **IT** is essential for businesses but you cannot **control** everyone because there will always be someone clever enough to not get caught!. David > sorry I meant in **ESS** etc.
 18. Sandra << I guess you're right about that, but still its all about **controlling** the majority.....there will always be someone smarter or one step ahead....
 19. David > ok, looking at the sheet, I am meant to say that **IT** cannot help **senior manager central control** which I don't agree with totally. But I do think it **makes control harder** to achieve! David > Without **IT**, what is there to **control?**
 20. Sandra << well, ok, don't really understand why it makes **control harder** to achieve, but that's your opinion, the main thing is that its not really that clear cut of an issue, if you know what I mean....
-

Employers are entitled to invade the electronic privacy of their employees

1–10), and on a par with a fairly broad discussion of the topic, at every two turns the repeated words are distinct: *IT*, *hard to control for senior manager*, *email*, *communication* and *the net*. As in the previous dialogue, we also have local sequential effects.

- (8) 1. David << well **IT** will make it more difficult **for senior managers** because it **makes it harder to control**
2. Sandra > I think **managers can have better control**, as **IT** can help them oversee communication flows within the organisation

The items of vocabulary repeated at the next turn have again the opposite distribution within the turn. In other words, and as previously, in (8) lexical repetition enhances Sandra's acknowledgment of David's contribution while she expands on her own views about the benefits of IT within the organisation.

This is not the only feature of this dialogue however. There is a global consensus centred on the same topic of *control to senior managers* occurring from line 12 onward, and clearly indicated at turn level by lexical repetition:

- (9) About control to senior managers, long term control, which is I guess linked to control, for senior managers to have control, you cannot control everyone, etc.

From lines 12–14 onward in Table 8, Sandra's contribution in the debate has shifted from an adversarial stance towards debating the same topic as David.

In terms of lexical word distribution, we note that line 14 in particular is very distinct from Sandra's previous other contributions, in the sense that, the lexical item *control* appears twice in the second part of Sandra's turn. In particular, we want to contrast turn 14 where *control* appears in the mid and end of turn 14 to turn 2 where *control* appears in the first part of the turn.

- (10) 2. Sandra > I think **managers can have better control**, as **IT** can help them **oversee communication** flows within the organisation
 14. Sandra > they're **like Executive Support Systems** which help managers make better decisions and help with long term **control**, which is I guess linked to **control**

From that point onwards, the distribution of repeated words within turn is mirrored at every turn:

- (11) 21. David < < True, so for senior managers to have control they have to invest a lot in **other departments** then?
 22. Sandra > why do they have to invest a lot in **other departments**?

To sum up so far, just as in monologue, lexical chains formed by repeated lexical words in dialogue are related to information maintenance, however the dialogic structure means that the lexical chains may also encode adjacency effects which are indicative of interpersonal awareness rather than a shared topic of discussion. Lexical word chains not only mark threads or cohesive segments related to a main topic as in monologues, but they are also indicative of local coherence effects which demand an exchange structure to be completed although they do not inform us of the overall participant's dialogic goals.

The distribution of lexical items at turn level is significant too. Lexical words at the end of the turn appear to provide us with information as to which conversational threads are currently being pursued in the debate, if we assume tentatively that the lexical words at the end of the turn are more important than those at the beginning of the turn, in a way that is reminiscent of the linear modification principle discussed above.

In the next section, we discuss an application of how knowledge about local coherence and global coherence effects in dialogue can help us to motivate topic selection in students' debates.

3 An application: information chunking

The main goal is to inform the conceptual design of a feedback device which describes how information is organised and shared within the interactive event. The retrieval of the main threads or topics of the dialogue is the first step towards this goal. We have argued however that there are two types of information being passed down the chat history. Respectively, these are the threads

or topics of the debate and the local sequential effects that encode the speaker's awareness of the other participant's contribution, not his or her willingness to adhere to its aims.

Furthermore, the study of manual turn referencing and threaded chat indicates that, just as a single common topic of conversation can arise, multiple topics or threads can coexist side by side. Thus information chunking of dialogue must not only be able to segment the text from top to bottom, but also from right to left. Top-down chunking as in monologue is determined by chains of repeated lexical items, but this is not enough. Knowledge about the linear distribution of repeated lexical items at turn level may turn out to be what is also needed, as indicated by the results of the qualitative analysis of sample chat debates above.

3.1 Problems with information chunking

The number of texts analysed is very small so it is difficult to draw generalisations from them. But the observations made in the previous section can be considered as open to further research. Key issues identified at this stage include:

- The relationship between linear distribution of repeated lexical items at turn level and the information structure of the dialogue as a whole.
- The importance of repeated lexical items varies according to the degree of coalescence between speakers. The remarks seem to capture well a dialogue where speakers are taking opposing sides in a debate, but this analysis may not extend to dialogues other than debates.
- Boundaries for the identification of a topic or chunk of information may also be a matter of interpretation on the part of different annotators. It is not clear whether there is a shared understanding amongst annotators when asked to identify topics in dialogue, and this may be a matter for experimentation.

3.2 ALEX feedback adviser

ALEX is a structured collaborative chat tool where users produce controlled discourse. Similarly to the ELIZA tutoring approach, ALEX feedback adviser makes inferences based on simple pattern recognition. The inferences drawn by the feedback adviser, unlike in the ELIZA tutoring approach, are based on patterns that occur at the macro level of the conversation rather than those of individual words. Its function is also distinct from that of the ELIZA system in that the feedback adviser does not participate directly to the conversation, but informs the participants on how their conversation is developing. In particular, the algorithm aims to draw the participants' attention to the elements that best summarise their discussion, and how each participant interacts within the discussion. Another aim of the feedback adviser is to be able to determine the degree of coalescence between the partners in the discussion.

At the moment, the feedback adviser relies on the non-linguistic cues of discourse recorded in the ALEX knowledge base in order to identify topics. These cues are the manual references made to previous turns, which, we argue, play the same role as lexical repetitions in chat discussions (Table 9). Here, we want to

Table 9 Extract from ALEX dialogue history (Royal Holloway experiment 2003)

0	Employers are entitled to invade the electronic privacy of their employees
1	I agree with 0 because This will prevent employees scheming against their boss or supervisors, or even the company itself.
2	I don't agree with 1 because I feel that employees should have the same privacy as they have in general.
3	In support of 1 I give the following example For example, this could deter certain employees from disclosing information to rival companies, such as profit figures, etc.
4	3 is true to an extent but would be better stated as Private e-mails shouldn't be allowed, but employee privacy is still an important fact as it shows trust in your employees.
	Can you give an example in support of 4
5	In support of 4 I give the following example If we don't give our employees trust there will be a them and us atmosphere and therefore having a private e-mail ban will increase productivity.

abstract from the template categories such as *I agree with* in line 1 or *In support of* in line 3 of the dialogue extract below and focus instead on the numbering system and what users do with this referencing system. As part of debate structuring, the ALEX user makes an explicit choice about the statement he/she wishes to support or dispute by choosing the number of that specific contribution among those that identify each the contributions already posted. ALEX thus functions similarly to threaded chat facilities where participants must index their turn to a specific topic.

Unlike the issue based information system (IBIS) (Rittel 1972) and gIBIS (Conklin and Begeman 1988) which provide a visualisation of debate by building a graphical map, reasoning is based on the reference system that encodes topics. ALEX also emphasises the conversational aspects of debate by providing a dialogue linear history unlike the discussion forums that present a tree representation of the debate. In ALEX, however, the linear dialogue history means that the participant is faced with several choices. As discussed above, participants can choose to express awareness of the previous participant's contribution by referring to his/her contribution or pursuing an overall dialogue goal, such as providing support for his/her own argument. Both goals need to be fulfilled, and in a linear dialogue that does not visually define a conversation as a set of topics, a reference to the previous turn is often made while possibly pursuing a distinct thread of conversation. So, for instance, if we assume a definition of topics as "the elements that best summarise a conversation" (Teufel and Moens 2002), then looking at the dialogue history above, we would rather select turns 1 and 2 (or 4) which summarise the main two conversational threads of the debate rather the four threads (0, 1, 3, 4) created by the participants over six turns. In order to account for the local sequential effects which do not bear relevance to the overall goals of the speaker, we propose thus to exclude references to an immediately preceding turn which does not refer to the players' own turn, and reason over the rest of the sequence. This allows us now to retrieve as required the two topics from the dialogue in Table 6. Whereas the local sequential turn references suggested that turns 0, 1, 3 and 4 represented the topics of the overall linear dialogue, non-local sequential referencing now reduces the set to 1 and 4 as conversational threads (*privacy and trust vs. scheming against the company*) coincidentally representing here the two sides of the debate.

If no topic has been detected, the debater looks at the individual moves made by each player (i.e. post a question on the basis of the calculation of initiative).

In collaborative situations where “one of the goals is to allow for maximum synergy in the cooperation work, balancing contribution between members and promoting inter group awareness” (Soller et al. 1999), we find that appropriate regulation of the dimension initiative is needed. ALEX messages also provide the participants with some feedback about this dimension. This is possible because the referencing system not only indicates the topics of the debate as in the tree representations of threaded chats and discussion forums, but also unravels the participants’ other communicative aims which are present in standard dialogue. The algorithm in ALEX therefore can use the turn referencing patterns in order to determine the speaker’s engagement with the other participants’ and his own contributions in a conversation where there is no elected chair.

To summarise, in this approach the algorithm does not rely on the argumentation templates that emphasise “cumulative argument construction and critiquing”, but the turns the participants wish to refer in the conversation to assess the level of knowledge co-construction in the debate. In other words, the emphasis is on the aspect of knowledge co-construction aspect of a debate in a dialogue context rather than on the argumentative strength of the debate in a graphical environment as expounded in the IBIS approach implemented in gIBIS.

4 Conclusion

We looked at the conceptual design of a feedback adviser that would tell us about the knowledge co-construction aspect of a debate. We argued that this knowledge co-construction aspect of dialogue is encoded in lexical chains. But, while tracking for the information content of the conversation, we noted that aspects of social intelligence are coded into the dialogue. Most specifically, it appears that repetitions have several functions in dialogue, and one of them is to encode awareness of connectedness. The lexical chains encode some kind of automatic acknowledgements which may not be related with actual knowledge co-construction (i.e. building a common goal or solution), but, rather, are conversational aspects which have been investigated under the term of adjacency pairs or local sequential effects (i.e. if someone asks a question you answer it, etc.). Whether there is indeed some linear precedence principle at work which allows us to distinguish between the social intelligence mechanisms at work in conversation and the genuine knowledge co-construction aspect of the dialogue is still a question open for further research.

Another aspect of information flow in computer mediated communication is the importance of repetitions probably due to the scarceness of contextual reference that is overspecified in face to face communication (as opposed to the written text) together with the resource limitations on memory imposed by dialogue (stronger than the written text). These spontaneous expressions of the person in dialogue may provide another key to the design of intelligent adviser systems.

As far as building a manual referencing system with a linear history is concerned, we found that users do not refer to turns in the same way as in threaded

chat, which offers a visual representation of topics. Instead, some ambiguity remains. Users may choose to refer to the topic they want to discuss, but they may also refer to the previous turn as a matter of fact.

This provides additional evidence that exchanging awareness with other members is part of the inbuilt mechanism of conversation, but distinct from the creation of knowledge in natural interaction. Thus we do not want to have threaded chats or tree representations which rob the speaker of this dimension of dialogue. On the other hand, when searching for genuine knowledge co-construction in a linear dialogue one must be aware of these apparent traces of agreement which are merely an inbuilt mechanism for exchanging awareness of the other participants' contribution in a conversation.

References

- DREW tools (2003) <http://test.scale.emse.fr/Drew/Sujet/>. Cited 22 July 2004
- Austin J (1962) How to do things with words. In: Urmson J, Sbisà M (eds) *The William James Lectures*. HUP, Cambridge
- Baker M, Lund K (1997) Promoting reflective interactions in a CSCL environment. *J Comput Assist Learn* 13:175–193
- Beaugrande de R (1980) *Text discourse and process*. Longman, London
- Brown G, Yule G (1974) *Discourse analysis*. CUP, London
- Cadiz JJ, Smith M, Burkhalter B (2000) Conversation trees and threaded chat. In: *Proceedings of the CSCW'00 conference*, Philadelphia
- Carlson L (1983) *Dialogue games: an approach to discourse analysis*. Reidel, Dordrecht
- Conklin J, Begeman ML (1988) gIBIS: a hypertext tool for exploratory policy discussion. *ACM Trans Office Inf Syst* 6.4:303–331
- Eggs S (1994) *Introduction to systemic functional linguistics*. Pinter, London
- Firbas J (1986) On the dynamics of written communication in the light of the theory of functional sentence perspective. In: *Studying writing linguistic approaches*. SAGE, London
- Fox B (1988) *Discourse structure and anaphora*. Written and conversational English. CUP, London
- Grice HP (1975) *Logic and conversation*. In: Cole P, Morgan J (eds) *Syntax and semantics*, vol 3. Academic, New York, pp 41–58
- Kahn M, Fisher T, Shuler L, Wu T, Pottenger W (2002) *Mining chat room conversations for social and semantic interactions*. Lehigh University, Bethlehem, PA
- McCarthy M (1998) *Spoken language and applied linguistics*. CUP, London
- Owen M (1975) *Conversational units and the use of well*. In: Werth P (ed) *Conversation and discourse*. Croom Helm, London
- Pilkinton R (1998) *Dialogues in support of qualitative reasoning*. *J Comput Assist Learn* 14(4)
- Pfister H, Mühlpfordt M (2002) Supporting discourse in a synchronous learning environment: the learning protocol approach. In: *Proceedings of CSCL 2002*, Colorado
- Rittel HWJ (1972) *Second generation design methods*. Interview in: *Design methods group 5th anniversary report: DMG occasional paper*, 1: 5–10. Reprinted in: Cross N (ed) (1984) *Developments in design methodology*. Wiley, Chichester
- Rooth M (1992) A theory of focus interpretation. *Nat Lang Semantics* 1:75–116
- Sacks H, Schegloff E, Jefferson G (1974) A simplest systematics for the organization of turn-taking for conversation. *Language* 50:696–735
- Saeedi M, Sillince J (1999) Incorporating rhetorical and plausible reasoning in an electronic conferencing system. *Knowl Based syst* 12:113–127
- Searle JR (1969) *Speech acts: an essay in the philosophy of language*. Cambridge University Press, Cambridge
- Soller A, Linton F, Goodman B, Lesgold A (1999) *Towards intelligent analysis and support of collaborative learning interaction*. In: Lajoie SP, Vivet M (eds) *Artificial intelligence in education*. IOS Press, Amsterdam
- Teufel S, Moens M (2002) Summarizing scientific articles: experiments with relevance and rhetorical status. *Comput Linguistics* 28:410–445

- Todman J, Elder L, Alm N, File P (1994) Sequential dependencies in Computer aided conversation. *J Pragmatics* 21:141–169
- Tsui A (1991) Sequencing rules and coherence in discourse. *J Pragmatics* 15
- Walker MA, Joshi AK, Prince EF (1998) Centering in naturally occurring discourse: an overview. In: Walker MA, Joshi AK, Prince EF (eds) *Centering theory in discourse*. Oxford University Press, UK