International Journal of Artificial Intelligence in Education (2000) 11, 79-105

Clarissa: A Laboratory for the Modelling of Collaboration

Mark Burton, Paul Brna, Rachel Pilkington *Computer Based Learning Unit, Leeds University, Leeds LS2 9JT, England, UK*
email: {mark,paul,rachel}@cbl.leeds.ac.uk

Abstract. There is a continuing debate about how to organise collaborative activities for them to be educationally valuable. This organisation can be analysed both in terms of how to set up the situation and how to arrange the interactions between participants. Here, we are interested in the different ways in which to organise the interaction, and in the consequences for effective collaboration. For example, what constraints might be usefully applied to 'unconstrained' collaboration? The work described here uses computational modelling to provide the tools for a systematic investigation. The approach taken is based on the assumption that students should learn through the adoption of different ways of using dialogue (dialogue roles). Very few, however, have sought to examine this assumption through the computational modelling of explicit dialogue role assignment within collaborative situations. The Clarissa (Collaborative Learning As Realised In Simple Simulated Agents) system allows the exploration of collaboration with respect to dialogue roles and different policies used for their allocation. A simple problem solving domain context is used, which exhibits many of the properties of more complex situations. The system is described, and selected results obtained from modelling a range of types of collaboration are presented. Findings from the analysis of a set of different collaborative arrangements indicate that there are more effective ways of organising collaborative situations than the free adoption of dialogue roles. In this paper, a pair of such policies are used to explore this issue using a baseline policy chosen as a representative of a commonly accepted form of collaboration. Clarissa itself provides a novel laboratory which has some implications for a new range of software agents capable of plausible collaborative behaviour.

INTRODUCTION

There is great interest in the study of collaboration, and in the study of how students learn to collaborate. Clarissa is a system designed for the study of collaboration, and more particularly, for the study of what makes for 'good' collaboration.

The development of Clarissa has entailed the adoption of a theory of collaborative activity. The system itself has been used to explore which kinds of collaborative situations are likely to induce good collaboration. This work has resulted in predictions which can be tested.

This paper starts with a brief outline of our underlying theory of collaboration, and how it fits in with other definitions of collaborative activity. This enables us to explain what is meant here by 'good' collaboration.

Central to the approach taken is the notion that utterances have an associated intention relating to the function of the utterance within the context of the dialogue. This function is termed a *dialogue role*.

The architecture and implementation of Clarissa characterises a subset of the collaborative situations that are of interest. The theory of collaboration embedded is quite general, and could be used to support a more 'powerful' version of Clarissa capable of generating better quality dialogue.

Results from a selected set of situations are presented, and indicate various ways in which the activities that are undertaken can influence the quality of the collaboration. The implications of these results are considered, and Clarissa's strengths and weaknesses discussed.

COLLABORATION, SHARED GOALS AND ROLES

Dillenbourg, Baker, Blaye and O'Malley have provided a good overview of the research into collaboration [Dillenbourg *et al*, 1994]. Though there has been a great deal of work in the area, many researchers - perhaps most - have focused on systems of agents (artificial or not) working together to achieve some shared goals while others have examined the processes of learning which are involved.

For example, Roschelle and Teasley define collaboration as "a coordinated, synchronous activity that is the result of a continued attempt to construct and maintain a shared conception of a problem" [Roschelle & Teasley, 1995]. In the context of the analysis of dialogues, they concentrated on the development and maintenance of a shared problem space. (They included domain-based goals in this space.) In the approach taken here, we regard the development of shared domain-based goals as an emergent property of collaborative activity.

However defined, collaboration entails both processes and goals. Exercising the processes associated with learning is viewed here as the key advantage that collaborative activity offers (this point is expanded on below). It is the quality of these processes that determines the pedagogical value of collaboration.

In order to examine collaboration we choose a way to measure collaboration that is indirect, one derived from the analysis of naturally occurring dialogue. There are numerous schemes for the analysis of human-human dialogue including classroom oriented ones where the teacher exercises a power relation in the educational context [Sinclair & Coulthard, 1975,Mehan, 1979] and ones which focus on more general dialogue [Stubbs, 1981].

Schemes for the analysis of dialogue, such as Pilkington's DISCOUNT scheme [Pilkington, 1999], seek to structure human-human dialogue in terms of a number of speech acts and their associated dialogue roles. The term *dialogue roles* refers to the linguistic functions of an utterance or related series of utterances, such as their declarative, interrogative, or imperative use. A given utterance may be used to state a fact, ask about a state of affairs, or issue a command, among other uses. In Clarissa, this basic set of common functions that utterances may perform has been expanded to include a wider set of roles.

The use of dialogue roles as defined here allows a direct connection to be made between dialogue roles and cognitive processes. It also allows us to regard the quality of collaboration in terms of these dialogue roles. In this paper, collaboration is only considered in terms of the dialogue which occurs between individuals.

THEORETICAL JUSTIFICATION FOR COLLABORATION

To be educationally beneficial, *useful* collaborative activities must promote learning. Collaborative activities are not always *useful*: the approach taken in this paper provides a means of discriminating between more and less useful collaboration.

The social interactions which occur during collaboration will promote learning; Teasley and Roschelle consider that "cognitive representations are built through social interaction" [Teasley & Roschelle, 1993, p230,]. They ground this belief in the theories of Vygotsky and Piaget, among others [1].

Boden [Boden, 1979] analysing Piaget's contribution, points out that although he offered a multi-stage model of development, he identified only one mechanism by which learning takes place. Essentially, learning (in Piaget's view) is achieved through the interaction between a learner and their environment.

Whilst Vygotsky would probably not have denied this basic mechanism, he emphasised the role of other people in the environment, especially adults [Vygotsky, 1978]. He suggested that by interacting with more mature peers or adults, students would be able to display abilities in advance of those displayed when on their own. In doing so, his suggestion is that learning takes place, not just through interactions with real objects, but also through verbal interactions which, as development progresses, become increasingly abstract. It is this conclusion, that language can mediate in learning through the social context, that has contributed to the increased interest in

collaborative activities. (The word 'collaborate' is used in Cole's translation of [Vygotsky, 1978, p86,].)

In Vygotsky's terms, the difference between what is achievable on one's own and with a more mature partner defines the 'zone of proximal development'. Skills exercised in this way, with a more mature partner, gradually 'mature' themselves, and are integrated into the child's understanding, so that the child eventually becomes able to perform these functions on their own. This offers the possibility that by allowing children (in different states of advancement) to collaborate, at least one of the children would learn at a much improved rate. They would be exposed to their 'zone of proximal development', exercising, practicing and learning the skills within it.

The original Vygotskian theoretical position has been developed to include *peer-peer* collaboration - see [Damon & Phelps, 1989] for an overview of the benefits of peer-peer collaboration, in comparison to peer-tutoring. Blaye, Light, Joiner and Sheldon provide some evidence of peer-peer collaboration being effective [Blaye *et al*, 1991]. Additionally, students that had been within a pair were then more effective at completing domain tasks on their own. This result is in line with ones that should be obtained as a consequence of the theory of collaboration outlined in this paper.

Their analysis of the possible reasons for the increase in skills includes the various 'roles' that are adopted by the individuals in a group, and the possible 'cognitive load sharing' [Blaye *et al*, 1991, p482,] that it affords. In their context, the roles arose from the distribution of physical resources (who controlled the mouse). In Clarissa, the control of the mouse also affects which roles can be adopted. Blaye *et al*'s interpretation of the results corresponds to a more coarse grained version of the theory underlying Clarissa.

The theoretical viewpoint taken here is that useful educational collaboration can and does occur between peers. While a partner's comparable cognitive processes may be developed to a similar degree, they may be developed in different ways.

This idea is exploited explicitly in Aaronson's Jigsaw approach to collaboration [Aaronson *et al*, 1978]. The Jigsaw method exploits different roles in peer-peer work, and typically involves dividing up a domain into some sub domains. These are then worked on by a number of individuals. They become 'experts' in these sub domains. The individuals then come together to collaborate with their partners. In this case, the individual students tutor each other, referred to as 'peer-tutoring' [Wood *et al*, 1995]. In terms of the sub-domain identified for the task, one peer is consistently in a superior state of development with respect to the other(s). This makes it very likely that cognitive processes will have been developed in quite different ways.

The Link between Learning and Collaboration

It is important, as motivation for our work, that people can learn to collaborate. There is empirical evidence that this does happen, and various people have speculated about the nature of the mechanism by which this might happen.

VanLehn, Jones and Chi break down learning mechanisms into two sets, those of acquisition and those of compilation [VanLehn *et al*, 1992]. They suggest that most theorists (e.g [Anderson, 1983] and [Newell, 1990]) "propose that there are only a few knowledge compilation mechanisms, such as chunking, proceduralization, and strengthening" (p3), and that these skills are "probably biologically determined".

VanLehn *et al.* go on to conclude that these methods are better understood than the broader set of acquisition learning skills [VanLehn *et al*, 1992]. However, it is these acquisition skills that are of interest to VanLehn *et al.* as they classify them as learnable - as distinct from those biological processes of learning which may themselves have to be 'trained' by suitable stimuli in the environment, but are acquired at an earlier age, and as part of a more general process of physical maturation.

One aim of the work reported here is to suggest ways in which collaboration can be modelled. In producing a framework for this model, it is unhelpful to make distinctions between the two learning processes that VanLehn *et al.* identify since both processes may play a part in collaborative activity, and it would be preferable to treat both in the same way. However, to

clarify the benefits that are perceived from collaboration, the distinction between 'acquisitional' and 'compilational' learning mechanisms will be maintained.

Providing a slightly different account of the same distinction, Salomon and Perkins classify learning in terms of 'high road' and 'low road' learning. In the latter, learning takes place through reinforcement, trial and error. This may be thought of as the instinctive, 'biological', mechanism that people use. However, humans also seem to be capable of displaying 'high' road learning, which entails "mindful management of the learning process" [Salomon & Perkins, 1998]. The intention of such learning is to set up situations in which 'low' road learning will take place. The learner manages the situation such that they can exercise 'low' road learning.

The ability to guide problem solving for instance, falls into this category. Simplistic problem solving may revolve around trial and error, but reflecting on this process and attending to those steps which seemed to contribute toward a problem solution in a given context allows strategies to be built up to deal with similar problems reducing error and time taken to solve new, similar, problems. These strategies allow more advanced problem solvers to arrive at solutions more rapidly. These strategies are 'high' road 'acquisitional' skills, although they employ 'low' level 'compilational' skills when exercised. Additionally, these high level skills are themselves subject to learning, in a bootstrapping process.

The activities of interest here are those that afford learners opportunities to *practise* and *observe* activities that are 'high' road learning, 'acquisitional', the meta skills that themselves enable and facilitate learning, 'learning to learn'. Ideally, pedagogic techniques such as collaboration are suitable to teach not just domain skills, but valuable *acquisitional* skills as well. Collaboration affords the opportunity to *practise* and *observe* 'acquisitional' skills.

Verbalisation in Collaboration

VanLehn *et al.* propose an acquisitional skill set, that set associated with the 'self-explanation effect'. The effect refers to the observation that learners who are capable of generating appropriate questions and answering them for themselves, seem to learn more. VanLehn *et al.* have modelled this effect and compared the results with observations of learners. They conclude that this is an acquisitional skill that students are able to learn, with positive results.

This is supported by Bielaczyc, Brown and Pirolli who have looked as self-explanation in a collaborative setting [Bielaczyc *et al*, 1994]. Within collaborative situations, the skill of asking suitable questions, and generating explanations may be performed by different participants in the collaboration.

Collaboration takes place in a social environment and as a consequence some social skills must be acquired, as prerequisites, in order for learners to take part in effective collaboration. In Bielaczyc *et al.*'s collaborative situation, an example would be the skills of asking and answering questions. Salomon and Perkins categorise six forms of social learning: the fifth, learning to be a social learner, is of most interest here [Salomon & Perkins, 1998].

Verbalising within a social environment, uttering the questions, and/or answers, provides a learner with two crucial facilities. First, the learner can receive a critique of their skills. More appropriately, they can watch the ways in which their cognitive skills are applied, and accepted within the social context. Second, they can concentrate on only one skill (say that of question generating) while a collaborative partner concentrates on another (say that of explanation generation).

So there are good reasons for verbalising thoughts, and accepting that cognition is distributed implies that verbalising thoughts is a key aspect of learning even though the necessity to verbalise is reduced by a process of maturation.

The learning of (for instance, acquisitional) skills is initially exposed by an individual who invites others to help with the process of learning. However, as these skills mature, the necessity to verbalise such speech is reduced. It has also been argued that rather than the verbalisation being reduced, it is shifted [Dillenbourg, 1992], a view point which is agreed with here. The language used changes to reflect learning with different cognitive processes. The quality of the language improves, reflecting enhancements of acquisitional skills, but the intention of the language remains, to invite others to assist in learning. The implication is that collaborative

activity may be appropriate from a very early age of development. Indeed it may be most appropriate at such stages. This is found to be the case by Harwood who has used collaboration with groups of children aged as young as seven [Harwood, 1995]. He found that the collaborations were better with the presence of the teacher who had a guiding role to play. Effectively the teacher was assisting in the learning of the skills of collaboration, while the collaboration itself was being used to learn about a domain.

In this section, we have seen why there are good theoretical reasons to believe that collaboration may be educationally useful. This, taken together with the empirical evidence that collaboration can sometimes be effective, justifies further investigation. The discussion above has given a glimpse of how the theory may work in practice. But a theoretical link has been made to the concept of socially distributed cognition, which will give further theoretical justifications for the usefulness of collaboration.

DIALOGUE ROLES AND COLLABORATION

In this paper, the process of collaboration is examined in terms of the dialogue which occurs between individuals. We will now go on to relate this dialogue with cognitive processes, and view the collaborative situation as one in which these cognitive processes are considered as distributed dynamically between participants, which will be reflected in the observable dialogue.

Here, the unit of analysis is the *dialogue role*. The analysis focuses on the assignment of roles to participants, and the way these roles are exercised and reallocated.

Dialogue Roles and Cognitive Processes

In any learning context a number of cognitive processes are presumed to be involved. In addition, a number of (dialogue) roles are identifiable within the observed dialogues. It is assumed that there is a strong link between cognitive processes and related dialogue roles. Further, all cognitive processes must 'communicate' with each other by exploiting a dialogue role.

The dialogue roles themselves may be complex, encompassing a wide range of dialogue or they may be relatively simple and capable of generating only a very few utterances. The ones used later include both *reason* which is quite a complex role and *Finished* which is quite a simple one indicating that the agent has no more to say (the list of roles used can be found in appendix A).

Here, the theoretical position is taken that processes of learning exists in their own right within the individual. Their existence has been conjectured by many other researchers, for instance [Minsky, 1987].

Further, it is the inter-process communication that is interesting here (Minsky does not fully address this issue). This communication is assumed to take place using *language*. Individuals may choose to attempt to get others to perform processes on their behalf. The communication between the participants is then the communication between the processes. This sharing of cognitive processes can be interpreted as collaboration.

If all the cognitive processes are co-present (in the same 'mind'), their communications will be internal to that mind, and there will not be any observable dialogue. In contrast, when processes are distributed across participants, there will be observable dialogue.

The dialogue observed to be associated with a cognitive process is also associated with a *dialogue role*. Notice that, depending upon the granularity chosen for the characterisation of both the roles and cognitive processes, more than one cognitive process may utilise the same dialogue role, or vice versa. The current system exploits a one to one mapping as adequate for the work's purposes.

One important consequence of the relationship between dialogue roles and cognitive processes is that by altering the roles which a participant is able to use, the cognitive processes that they will choose to use are affected. By stipulating the dialogue roles, the way in which

participants collaborate can be adjusted. This general effect is a consequence of the architecture. However, empirical work is needed to determine both the precise effects of changing the availability of roles, and whether these effects are found in practice. This paper provides an interesting answer to the first of these issues and suggests certain effects that might be observable amongst small groups of students.

A decision has to be made about which cognitive processes are present and worth studying. This will depend on the cognitive processes that the students are expected to learn and use, and this will vary between student groups depending upon their abilities and previous knowledge. This is also true for the 'meta level' skill of learning. The selection of cognitive processes is therefore a pedagogical decision. In this work, the specification of the processes and roles considered to be present were derived from Pilkington's work on dialogue analysis [Pilkington, 1999].

The Distribution of Roles

The key issue in this paper is the study of different patterns of dialogue role availability and different policies for redistributing these roles during problem solving.

Different distribution policies will have different consequences. For example, if collaborative activities are expected to be time efficient then the relevant cognitive processes should be distributed to participants who are capable of exercising them most effectively, while minimising the communication overhead. While this is a common educational situation, the goal of the research reported here is different.

The aim is to use collaboration as a tool for allowing participants to concentrate on specific cognitive processes, and to practice/ observe [2] them to improve their performance. Collaboration, as used here, affords a participant the ability to practice specific cognitive processes while their partners may address different ones.

Educational collaboration is situated in a learning context. In such contexts it is assumed for simplicity that there are a number of cognitive processes which are of interest in all educational situations i.e. that cut across issues connected with the actual material being learned. This subset of the cognitive processes are here referred to as *learning processes* [3].

It is useful to draw the distinction between these and the processes involved in the domain itself. This allows the possibility of using collaboration as both a generic pedagogic tool for teaching domain-based material, and one capable of being used to develop generic learning skills. Here, the learning processes are of greater interest than the cognitive processes associated with problem solving in the domain.

In this light, good collaboration involves distributing cognitive processes according to pedagogic criteria, including, amongst other factors, the ability of participants. The communication overhead would not be of any great concern. With respect to ability, it is unlikely that pedagogic criteria would ever lead to the distribution of cognitive tasks only to those most able to perform them; pedagogic considerations conflict with achieving quick problem solutions. It is not possible to achieve a collaboration characterised by successfully and quickly completed problems, and for the same collaboration to be 'effective' at allowing participants to improve their learning skills.

The theory outlined here, and developed more fully elsewhere [Burton, 1998] predicts that uncontrolled collaboration would not necessarily produce 'good' results either in terms of the goal to achieve improvements in learning skills, or in the often utilised criteria of time to completion, or accuracy of solution.

The next section introduces Clarissa, and how collaboration between Clarissa agents is managed. The results of examining collaboration using Clarissa give some indications about different approaches to collaborative activity which may be beneficial.

CLARISSA: THE ARCHITECTURE OF A SINGLE AGENT

To determine what can be learned from Clarissa, it is first necessary to describe the architecture briefly. Clarissa utilises a number of techniques that could prove useful to researchers wanting to build systems that support collaborative activity (as defined here). The architecture is divided into two main subsystems: the cognitive and dialogue systems.

The Cognitive System's Architecture

The cognitive system is broken into separate units (cognitive processes), which then communicate with each other. All 'internal' communication that might be regarded as occurring within the cognitive system are made using the dialogue system. The decision about which utterances are observable, and which passed directly between cognitive processes within the agent are made within the dialogue system (and will be discussed below). Other than their communication, individual cognitive processes may be implemented in any way. The choice here is a system akin to a forward chaining production rule system.

Communication between the Cognitive and Dialogue Systems

The dialogue system must allow messages to be passed between the cognitive processes found in the cognitive system. To achieve this communication, a new abstraction is introduced, a *dialogue goal*. A dialogue goal is an expression of the communicative act [Maybury, 1993] that the cognitive system wishes to take place. The dialogue system can then choose how best to achieve this goal.

The cognitive process initiates a dialogue goal whenever it wishes to pass a message to another process. These are 'thrown over the wall' to the dialogue system. This mechanism is straightforward and clean. It requires that the goals the cognitive system generates are understandable by the dialogue system so there will be some dependency between the two systems.

Goals so delivered, and acted upon by the dialogue system may eventually be completed. The desired result is that the relevant messages are delivered back to the cognitive system. To do so requires that there is some mechanism in the dialogue system for passing information back to the cognitive system.

There is a second reason for requiring a mechanism for the dialogue system to consult the cognitive system. The dialogue system will process its dialogue goals when it is ready to do so. A goal may be processed some time after it was initiated. In this case, it is advisable to check that the cognitive system still desires the goal.

The Dialogue System's Architecture

The architecture of the dialogue system itself is considered in terms of an interactive-learning 'game' (dialogue game) and role mechanisms that it employs.

The dialogue system, considered in itself is a complex system. It can be seen as independent (to an extent) like an 'agent' in its own right. Viewed from a Beliefs, Desires and Intentions perspective [Rao & Georgeff, 1991], the dialogue system has:

- **Beliefs** about what has been said.
  A history of the dialogue is kept to maintain a notion of focus.
- **Beliefs** about what can be said.
  A mechanism is provided which allow the prediction as to what might be said next, and permit some planning as a result.
- **Desires** about what is wanted to be said.
  These will be provided by the cognitive system in terms of dialogue goals.
- **Intentions** to say specific things.

The set of desires is required to be filtered to a set of possible things that can be said, and then further work is done to distinguish between utterances which should be made observable and ones which are 'internal' to the agent.

The individual parts of the dialogue system are outlined starting with the mechanism which allows the focus to be maintained and a prediction to be made about what might be said next, the dialogue game. Then dialogue roles will be discussed. The combination of these mechanisms permits limited planning to take place. Finally the dialogue moves have to be turned into language utterances which are the observable output of any system following this architecture.

The Dialogue Game

The dialogue game is a structure that: enables the focus of the dialogue to be maintained; supports decision making about what moves are available (i.e. dialogue moves); and helps predictions to be made about what might be said next, and hence allows some planning.

A dialogue game is defined to be a state machine which represents the entirety of possible dialogue utterances and the order in which they can occur. Dialogue utterances are the observable outcomes, and are theoretically generated from the dialogue moves with which they are associated. This can either be viewed as a very tight set of general constraints in a style akin to that of Mackenzie [Mackenzie, 1985], or as a generalised form of commonly used smaller dialogue games as found in Blandford's Wombat system [Blandford, 1994].

Mackenzie defines a dialogue game as the set of rules that specifies a number of different dialogue moves in a formal argument which are possible between two people to produce a resolution. Blandford's dialogue games are based on Levin and Moore's notion of conversation as selecting a game from a set of games with each game being seen as providing a mechanism for achieving some goal [Levin & Moore, 1977]. If played correctly, to the end, then the goal (e.g. to inform the other player of a fact) is assumed to be successful.

A single dialogue utterance, from a specific position in the dialogue game, will cause the game to move to another state. A single dialogue utterances may be possible from a number of dialogue game states. A dialogue utterance from a specific dialogue game state is termed a *dialogue move*. The state machine holds information about which moves have been made, which can be made, and by whom. As participants make moves, a history is kept. State machine nodes may allow several dialogue moves to be made in parallel. In other words, it is possible for multiple utterances to be made in response to a previous dialogue move. Hence the dialogue history contains information about all of the possible moves that a player can make from all of the states they have visited.

When multiple moves are made from one state, the resulting dialogues are defined here as being *parallel dialogue threads*. A normal dialogue game would limit the number and type of moves that can be made from any given state. States from which no further moves can be made are said to be closed. If all the states on a parallel dialogue thread are closed (including the 'leaf' node), then the dialogue thread is said to be closed.

The next section examines how focus can be maintained utilising this tree.

Focus

The ability to maintain parallel dialogue threads allows agents to keep a number of topics active at the same time. This raises the issue of how to place new dialogue utterances into the dialogue history, in other words, knowing to what they refer. This is an aspect of the problem of 'maintaining the focus'. With parallel dialogue threads, the 'focus' of discussion is not simply stack based so maintaining focus becomes more complicated as a result. It allows agents to switch between a number of different threads, and to initiate new threads in accordance with the dialogue game.

Being able to keep track of focus is defined as knowing where to place new dialogue utterances in the dialogue history. The assumption in this architecture is that the dialogue game

should enable the focus to be tracked without needing to access the contents of the utterances [4].

In natural human dialogue, explicit dialogue markers are used, as described by [McCoy & Cheng, 1991], to enable shifts of focus which are more complex than depth-first tree search. These markers give simple signals about shifts in focus that help listeners maintain focus without resorting to the content. The current implementation of Clarissa can partly avoid the need for such a mechanism through one of two techniques:

- The history is kept in such a way as to assume that the most recently talked about dialogue thread is the most relevant. Hence, new dialogue utterances are placed on the most recent dialogue thread on which they are valid.
- The dialogue game mechanism is constructed to prohibit moves from being made which would cause confusion. In this case it is guaranteed that a new dialogue utterance would only be valid in one place in the dialogue game, making such a move will shift the focus to that point in the dialogue game history.

The first of these schemes, implemented in Clarissa, would handle errors made in the maintenance of focus. Such errors in natural language would be handled by repairs to the discourse. This repair mechanism, not currently implemented, could include dialogue goals that effectively repair the history to accommodate misunderstandings [5]. This results in a much more complex implementation, but has the advantage of being able to produce more natural dialogues.

The second mechanism places a restriction which significantly limits the parallelism which is possible. Previously, the concept of allowing multiple parallel threads of dialogue has been mentioned. Allowing parallel threads gives the dialogue great freedom. However, insisting that moves can not be made if they could be responded to in the same way as a move which has been previously made will reduce this freedom. One way around this problem is to increase the number of different dialogue move types. This will reduce the probability that two moves could be followed by the same response. At any point in the dialogue, there would be a greater chance of different moves being available which will not cause confusion. To increase the number of dialogue moves, the granularity of the dialogue moves should be reduced. This mechanism does not require significantly more complexity in the implementation. The dialogue will miss any self-referential statements intended to achieve dialogue repair, but this is judged as acceptable.

In addition to the focus mechanism, based on the dialogue game state machine, the dialogue system uses a role mechanism to embody the theoretical standpoint described in this paper.

The Dialogue Role Mechanism

Dialogue roles are used to control the communication between cognitive processes. In terms of dialogue, the extra information that roles provide can be used for planning.

For the purpose of the architectural description, the dialogue role is determined by the dialogue utterance types which make it up [6].

Dialogue utterance types classify the dialogue utterances which are used to make dialogue moves within a dialogue game. The definition of a dialogue game has been previously examined. For the architecture, it is one large state machine. A dialogue role can be seen as defining a (complex) zone within a dialogue game, a 'sub dialogue game'.

A dialogue utterance type may appear in several places in a dialogue game. To limit the scope of a role to one specific area of the dialogue game, the dialogue utterance type will need to be sub divided into distinct types.

In the case of a simulated agent, whose underlying cognitive processes are a known quantity, the dialogue utterance types generated by the different cognitive processes can be determined (being careful to make sure each process generates 'unique' utterance types by imposing fine enough classification criteria). These utterance types can then be used directly in the definitions of the roles. If the cognitive processes that have been used in the cognitive

system are too fine grained, the implementation may choose to group together some of the roles to form more complicated, but possibly pedagogically more useful roles. Such roles will exercise the underlying set of cognitive processes in a more complex manner.

A mechanism is utilised for tracking which agents are playing which roles. The dialogue utterances which make up roles are a known quantity, so by observing the dialogue utterances (and hence dialogue utterance types), it can be seen which roles they are playing. This is combined with rules about which roles can be combined with which others. Both these rules, and the make up of the roles are user definable within Clarissa. The role mechanism can then infer, for instance, that if some roles are being played by a participant, the participant cannot be playing some other roles.

One 'built in' restriction will be imposed, which is intended to reflect normal human behaviour. Roles that one agent is playing cannot be used by another. This may seem very restricting, but is not intended to be so. Roles will be swapped frequently, and the effect of the restriction will be to model the way people normally respond to each other, one asking a question, the other replying and so on. This is similar, but not the same as 'exchange structure'. To actually decide what is said next, both the role that individuals are playing, and the dialogue game is examined. The 'exchange structure' information is primarily held in the dialogue game, which defines what can possibly be said next, and by whom. The role mechanism mediates this.

In addition to these restrictions, user and system imposed, Clarissa also allows for a variety of dynamic mechanisms for distributing roles throughout the period of the ongoing dialogue. Together these form the various constraints on the collaborative situation that Clarissa can be used to investigate. This will be examined in more detail below.

CLARISSA: THE IMPLEMENTED SYSTEM

The full system is written in C++ in around 15000 lines of code, and runs on either Sparc or i486 platforms; there is also a primitive Java-based Web interface. Clarissa agents run within a specific computational context. This includes a software environment known as the 'room' in which agents can act; the various ways a Clarissa agent can be launched; and the system that manages the different processes (the 'task system').

Clarissa has over 20 flags that can be set at run-time. While many of these flags are useful for the study of Clarissa's notion of collaboration, some are designed to watch specific kinds of event and a few are primarily low level. The Java interface is currently under development, and is shown in Figure 1.

There are some important considerations about the mechanisms which allow the Clarissa agent to 'think' about a number of things concurrently (but maintain a consistent chain of thought). These details are not discussed here, as they rely on a technical discussion of a multi-tasking light weight process, or threaded, task system, [Burton, 1998]. Likewise, there are issues, especially to do with when Clarissa agents decide to finish their collaboration, which are not discussed here, as they also relate to this task system, in conjunction with the adopted mechanism for message transfer (which in this case is a TCP/IP connection).

In the remainder of this section, the domain is briefly described for which Clarissa is currently configured. The section is then finished with a 'run through' of a typical exchange within Clarissa.

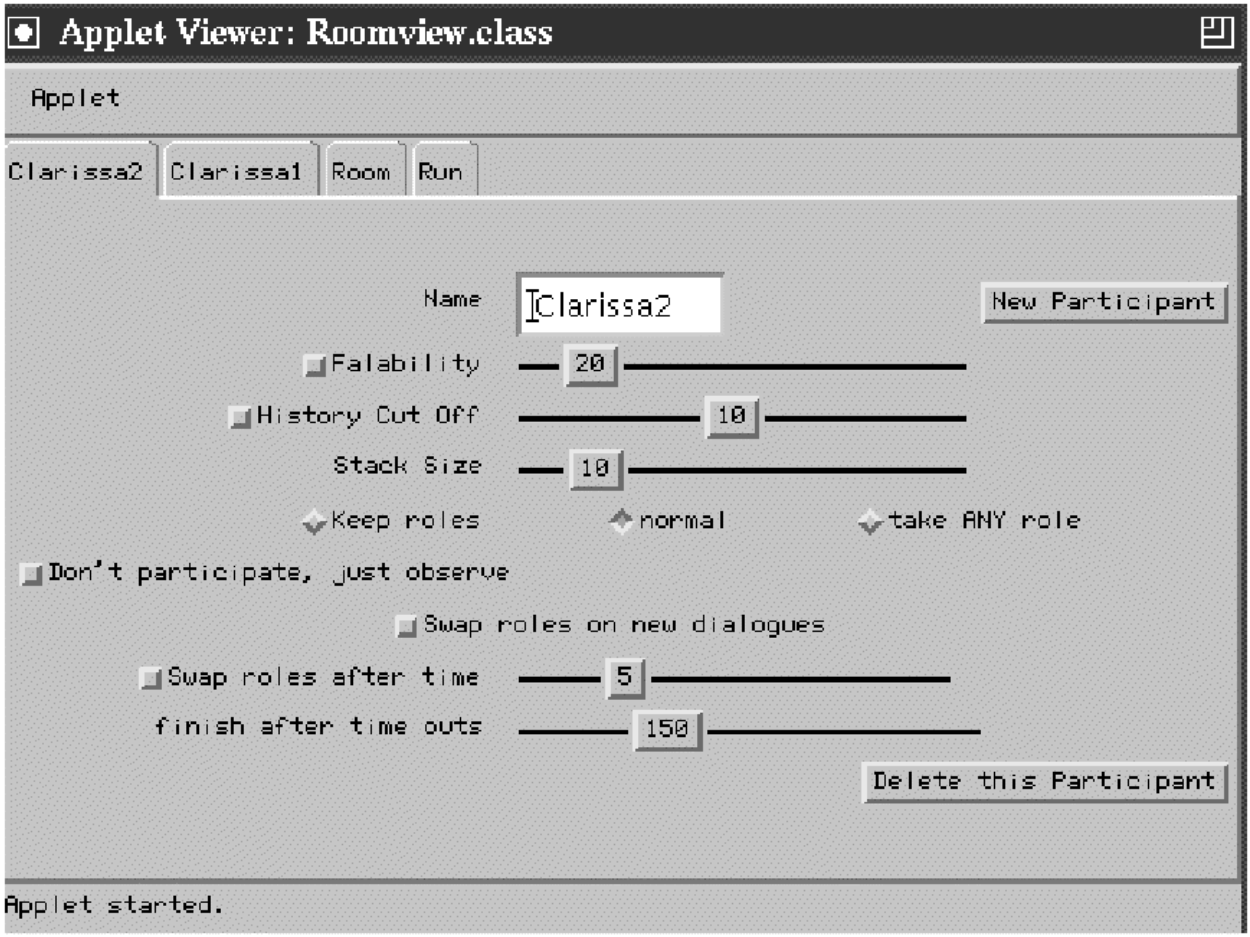


Figure 1: The Java Interface for Configuring Clarissa

The Domain

For an implementation, the cognitive system requires a domain to be adopted. Currently, Clarissa is implemented to work with a re-implementation [7] and extension of the ModelCHENE system developed by [Bental & Brna, 1995]. Bental and Brna and extracted the various cognitive processes from the use of CHENE, a system designed to support dyads collaborating in the construction of an energy chain [Devi *et al*, 1996]. This is combined with a multi-threaded environment using the dialogue mechanism as the 'message passing', or control, system.

The context being modelled includes students with no formal understanding of energy who are given an information sheet with limited instruction about energy. They are asked to draw an energy flow diagram to represent the experiment which consists of a charged battery connected to a bulb (with two wires), the bulb is shining. They are also reminded that energy flow diagrams must start with a reservoir, and end with a reservoir. They are not told about conservation of energy.

An advantage of using this domain is that there has been a fair amount of work carried out in terms of analysing real dialogues [Megalagaki & Tiberghien, 1995,Baker & Lund, 1996] and modelling the students' behaviour [Devi *et al*, 1996,Bental & Brna, 1995]. This modelling has mainly concentrated on the cognitive aspects, and therefore provides the raw material for Clarissa's cognitive system.

The ModelCHENE system, as implemented by Bental and Brna, described the behaviour of an ideal student, a 'typical' student as well as two individual students [Bental & Brna, 1995]. It

was necessary to 'break' the model of the ideal student and include some of the processes that, while erroneous, seem to be present in the case of real students.

Quite frequently, a typical student incorrectly identifies the second reservoir as being the same as the first (the battery). Clarissa uses a forward chaining production system based on a simpler conceptual model of problem solving compared to that used for ModelCHENE. The resulting model of student behaviour is not necessarily typical, nor is it necessarily more accurate than ModelCHENE, but it does produce similar energy chains to those generated by real students [de Vries, 1995]. It is therefore a candidate for the way in which some students behave. Further details of the precise implementation are available [Burton, 1998].

The cognitive system has been divided into two main cognitive processes. The first of these is responsible for generating 'new knowledge', the second section checks the validity of this knowledge. Clarissa provides run-time switches to allow agents to be restricted to one or other of these modes if required.

Two Clarissa Agents at Work

Typically, two 'student' agents [8] would be combined with a 'teacher' agent. The 'teacher' agent would take no part in the dialogue, but would collect and collate information about what has been said, and supply 'role swap' interventions (which will be discussed in more detail later). The 'student' agents would be normal Clarissa agents.

A typical sequence in the construction of an energy chain is illustrated in Figure 2. The rectangles are reservoirs, the rectangles featuring a cross are transformers and the arrows indicate a transfer of 'stuff' [9]. The 'unknown' tag indicates that the agents have not yet chosen a label describing the kind of 'stuff' transferred. As can be seen, step 4 involved the removal of a reservoir (new reservoir 1). The Java interface is shown in Figure 3, at the completion of the chain.

The implementation is illustrated with a small excerpt from a pair of Clarissa agents, and follow the program's execution. The following exchange is used as an example:

```
Agent A: Somebody should do Determine Theory Model on battery1
Agent A: I suggest we add a Reservoir to represent battery1
Agent B: I agree I suggest we add a Reservoir to represent battery1
Agent A: I have done Determine Theory Model on battery1
Agent B: I'll update the interface ADDReservoir:battery1(battery1)
```

The interpretation [10] of this exchange is that Agent A has identified the battery (battery1) as being a physical object for which it does not yet have a theoretical representation, or understanding. However, it possibly has some information that would help it come up with such an understanding. In this case, it knows that batteries store electrical energy, and therefore battery1 is likely to be an energy reservoir. Agent B agrees with this hypothesis, and also grabs the mouse and adds a reservoir to the interface, labelling it battery1.

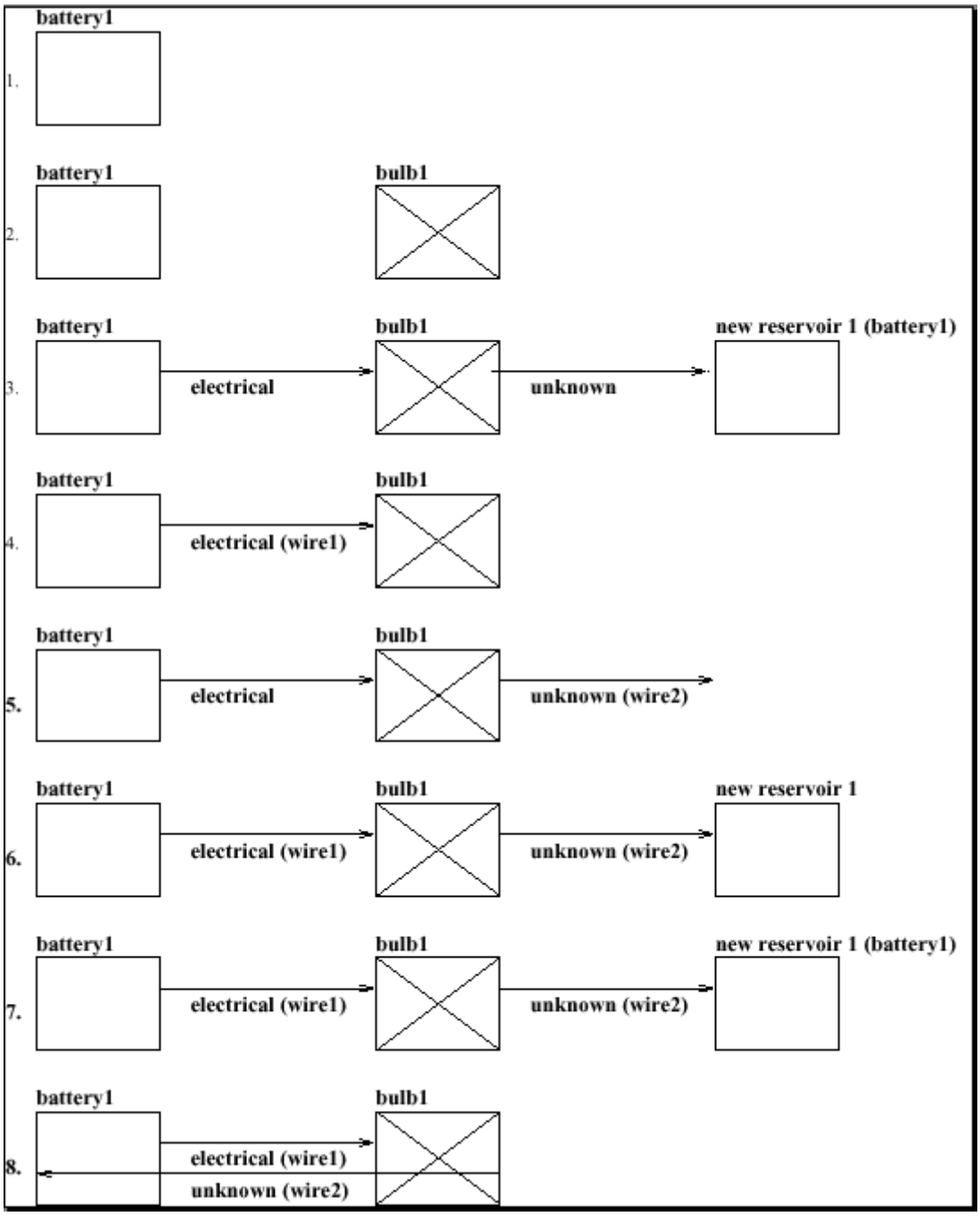
This exchange occurs very early in a dialogue; agent B mentions the fact that it will update the interface since until this point nobody had used the mouse. Agent B is in some real way 'grabbing' control of the mouse. The exchange is now examined in more detail.

Generation

Initially the agents (normally) think about a number of aspects of the problem including: the physical environment; deductions they can make from what they know; and finally whether their solutions are consistent, concise, match the physical environment, and match the requirements set by the teacher.

In this case, Agent A has been thinking about questions it may ask about the physical environment. It has noticed the object Battery1, and realises that it hasn't considered it before, or thought how it fits in. Effectively it realises that a new cognitive goal needs to be set up to incorporate the battery into the current understanding of the problem. This could have been done directly, but rather than doing so, it was a design decision to choose to 'expose' this type of

thought, and generate dialogue from it. The dialogue system is used to pass the message that a new cognitive goal is needed



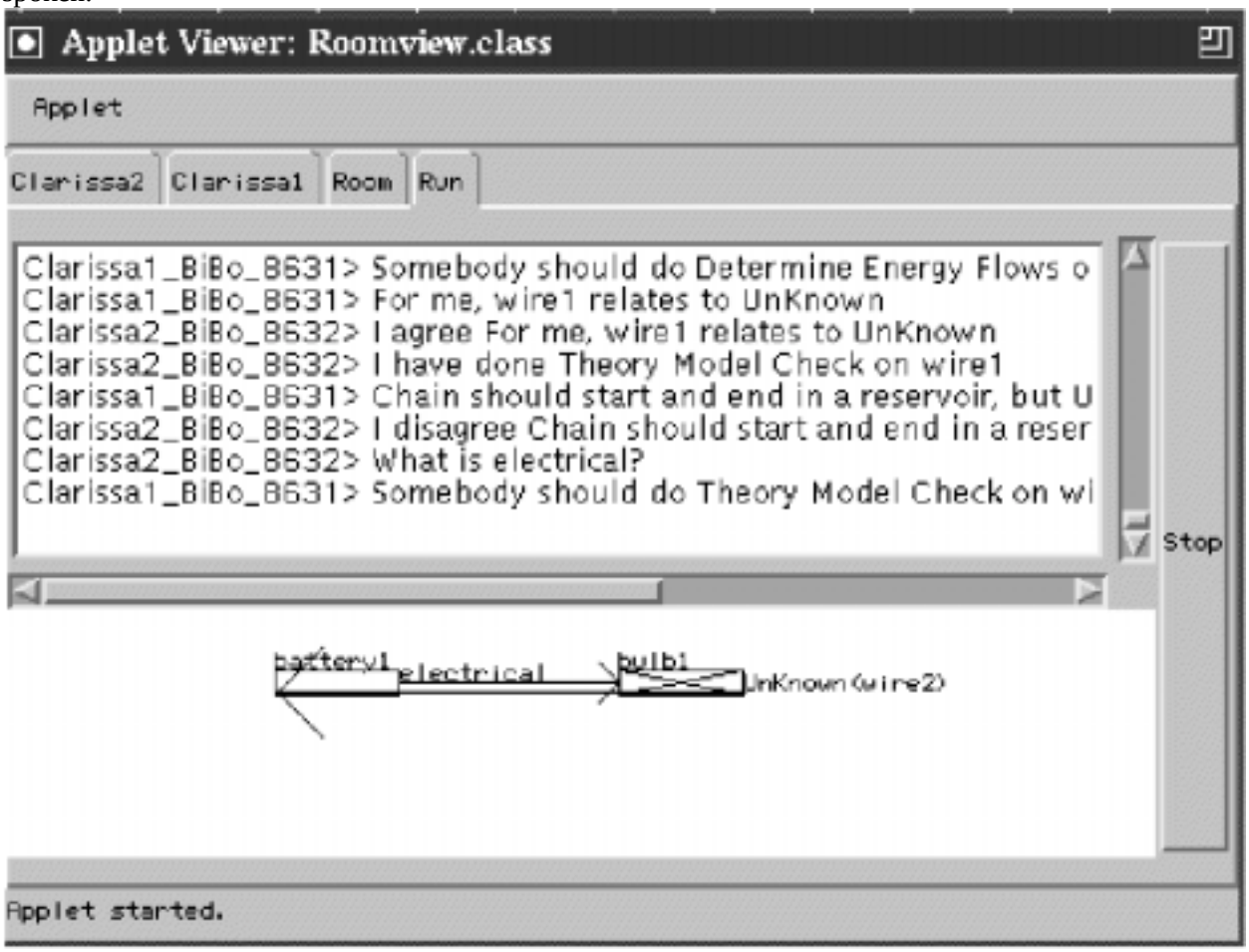
Eight steps in a solution path are shown. Reservoirs are indicated as empty boxes, Transformers as crossed boxes. Transfers are indicated by directed lines.

Figure 2. The Development of a Shared Energy Chain

At this point, the cognitive system generates a dialogue goal passing it to the dialogue system. The goal can loosely be expressed as "Somebody should do Determine Theory Model on battery1" [11]. This goal is then added to the list of dialogue goals Agent A is working on if the goal does not already exist. In this example, it may be the case that this dialogue goal was generated some considerable time before the utterance was observed. The dialogue system will continually re-evaluate all of its goals and only 'process' the most relevant. To achieve this re-evaluation, the language system will also be continually consulted as to the most appropriate

dialogue utterance which would assist in completing this goal. In this case, this is simply the utterance "Somebody should do Determine Theory Model on battery1".

Having chosen to process this dialogue goal, with the chosen utterance, a calculation is made as to whether the utterance should be 'out loud', or private. In this case the utterance is spoken.



The applet features a run of two Clarissa agents. In the window the agreed solution to the problem is featured. It is equivalent to step 8 in the solution path shown in Figure 2.

Figure 3: The Java Interface onto the Room

The message "somebody should... " is directed as much at Agent A as it is Agent B. In fact, in this specific situation (but not all), the agent which proposes that something should be done is the one that does the action. Hence Agent A will set up a new cognitive process, the object of which is to look at the battery and suggest ways of incorporating it into the agent's understanding. This process is now explained.

Receipt

At some point after the utterance was generated, it will be received by all agents. An agent is not allowed to have more than one utterance 'outstanding', so between the utterance being 'said' and being 'heard' there will not be any other utterances heard which have originated from the same agent (Agent A in this case). However, it is perfectly possible that other agents may have said things. Indeed they may have said exactly the same thing, or even something which makes the utterance irrelevant or redundant.

People seem to have a number of verbal and non-verbal ways of dealing with this. Clarissa also handles this situation, and guarantees that all agents 'hear' the same things. Agents can't say

things at exactly the same time, and the order is important and maintained. This is less flexible than the human system, but substantially more tractable.

In the example there were no complications, and the received utterance is recorded in the dialogue history. To do so, firstly the dialogue for which this move is most relevant is found, and then this utterance is added to the end of that dialogue. Dialogues augmented in this way are then indicated as being the most relevant. This simply maintains the focus tree.

The architecture indicates that the language utterance should first be interpreted as a dialogue goal (for the speaker), and then any cognitive goals should be generated from that dialogue goal. However, this issue is simplified in the implementation: the language utterance is identified, and added to the dialogue game history, and as it is being added, the language system is consulted for the last time. The language system is given the responsibility of generating cognitive goals.

In this case, the language system is asked to generate the cognitive goals for the utterance "Somebody should do Determine Theory Model on battery1". This is fairly cleanly achieved by generating a new cognitive goal where the task is to "Determine Theory Model on battery1". The new task (cognitive goal) will then execute when the agent is ready. Its priority will be set quite high, since it is something that is clearly relevant to the current discussion. However, it is also something that the agent might possibly forget to do [12].

Cognitive Processing

The new task which has been set up by the language system may now run at any time. However, its priority will be set such that it is likely to run fairly soon. The associated code is simply an induction step taken directly from the ModelCHENE model (even the name "Determine Theory Model" has been taken from that model). In the example, the rule states that an electrical object which is capable of storing energy should be identified as a reservoir.

The agent does not make this connection in its knowledge base straight away, first it is suggested; if (as is likely to be the case) the suggestion is accepted, the update will be made. Hence the only action that the cognitive goal needs to do is set up a new goal for the dialogue system, namely "I suggest we add a Reservoir to represent battery1". Once this dialogue goal successfully executes, the cognitive goal has completed. The language system, in this case, was waiting for the cognitive goal to complete, so that it could initiate a new dialogue goal ("I have done Determine Theory Model on battery1").

In the meantime, the utterance "I suggest we add a Reservoir to represent battery1" will be received. Both agents will process this in the same way as the previous example. However, in this case, the language system starts a task whose goal is to evaluate whether the agent believes the statement, and if so, then updates its knowledge base. Both agents perform this in the same way, except that Agent B, having received the utterance from Agent A will 'reply', in this case by 'agreeing'. This is achieved by initiating a new dialogue goal.

It should be noted that while an "I Agree" statement may appear in the dialogue, it is somewhat misleading. The cognitive change will have occurred for both agents on receipt of the initial statement (the "I suggest" in this case); prior to the agent verbalising the agreement. It may be better to view the "I Agree" statement as "I have agreed". This is fine in the agreement case, but is more problematic in the disagreement case. In the example both agents have agreed, but a digression is now made to examine the case of disagreement briefly.

Disagreements are not fully handled by Clarissa. The development of a more general purpose argumentation system would be part of further work needed to develop Clarissa. However, there is a pragmatic reason why the mechanism was not considered a priority: much dialogue and, in particular, the human-human dialogue which is modelled by Clarissa does not feature much argumentation.

If an agent receives information which conflicts with its knowledge base, it will utter a move such as "I disagree". Other agents will not respond this move. However, the agent would also be expected to give some reason for its disagreement, and this would normally take the form of information that it believes to be correct, and justifies its position. There is a crude system for deciding whether a piece of information is 'better' or 'worse' than another. On receipt

of new information, each agent decides whether it is better or worse than what it currently believes. All agents have the same crude system, which is based upon whether the objects concerned can be 'seen' in the environment. This is totally domain specific. Once all the participants have offered their information, the expectation is that all the participants will end up agreeing upon the 'best' information.

The final utterance in the excerpt is not generated as a direct result of the previous utterances. Rather, both Agent A and B will also be thinking about making sure that the diagram that they are generating (the external representation) matches their own private understanding of the problem. Both Agent A and B will have noticed that they now have more information in their private knowledge bases than there is on the interface. Both would wish to update the interface, i.e. take the mouse and add the reservoir. In this case, Agent B says "I'll update the interface ADD--Reservoir:battery1(battery1)" before Agent A. By making this statement Agent B is effectively doing two things. Firstly, evidently, it is updating the interface; secondly, the move "I'll update..." is defined as being in the 'interface' role, by making the move, Agent B is indicating that it is playing that role. All dialogue moves carry these two pieces of information with them.

Subsequently, Agent B will continue to move the mouse until, either explicitly asked by the teacher agent to stop using the role, or they decide to use a role that is defined (by the user-definable role rule set which has been mentioned above) as conflicting with the interface role. This is the case for all roles.

The outline presentation of how Clarissa works is finished. We now want to turn to the issue of how different policies for distributing roles affect collaboration. Before doing so, there is a need to look more clearly at how 'good' collaboration can be defined.

AN EXAMINATION OF EFFECTIVE COLLABORATION

The results to follow, and their interpretation, provide the basis for a discussion of the predicted relative value of different forms of collaboration. In order to evaluate the results, specific characteristics of collaboration have to be examined both in terms of the issues that are important to the research and in terms of how these issues are addressed through experiments with Clarissa.

It has been argued that collaborative activity may improve the ability of participants to learn by affording them the opportunity to practise individual elements of the learning process. At a less theoretical level, the 'skill' that the participants practise may be at the domain level. Collaborative activity may not be the only, nor indeed the most appropriate, mechanism for enabling participants to practice individual elements of a generic domain.

For a computer model to enable us to evaluate the change in the ability to learn would involve not only modelling learning, but also a second order effect, the change in this ability. Clarissa currently features a very simple problem solving algorithm. So it is not possible to measure any changes in learning.

Clarissa does provide the opportunity to examine one very important issue. Hitherto it has been assumed that the application of roles to a situation is possible, and acceptable to the participants. The metric for collaboration (with respect to Clarissa) is a function of the degree to which agents can use all the roles available, and the degree to which they perceive a benefit from the collaboration.

It is possible to directly measure the perceived usefulness of the collaboration in terms of the number of utterances received by an individual from a partner that cause that individual to change their knowledge base.

Likewise it is possible to measure the degree to which an individual in a collaboration exercises a wide range of roles in a relatively uniform manner. This relates to the fundamental claim in this paper: that the exercising of a range of roles is advantageous. Thus a goal in the kind of collaboration argued for is that all (or most) of the available roles are exercised to a reasonable extent during the collaboration, and that the usage is balanced between the participants.

The degree to which a role is used will be approximated as the number of utterances which are made by participants while an individual is 'playing' that role. This is an approximation in the following respects. The number of moves an individual makes 'within' a role is not considered, since to do so would not include information about how long the role was used for, and to what degree the other party was active during that time. The time that the role is used for is not useful, since this does not necessarily reflect the amount of activity the participants are engaged in (especially if the machine the simulation is run on is subject to varying loads). Hence the amount of 'collaboration time' is approximated by the number of utterances made by either participant. This gives us some measure of the length of time a role has been used for, and hence some measure of the degree to which an agent is 'using' a role.

While this measure has been designed with a quite specific definition of good collaboration in mind, it is necessary to bear in mind that much more can be written about the different kinds of 'good collaboration' that might be of interest in the classroom. The argument here is that the practice of a range of roles has both pedagogical validity and cognitive validity as a means of fostering the forms of long term learning that are desirable and that take advantage of both relevant domain content and social conditions.

Having developed a measure of collaboration, we need to return to the main use to which such a measure can be put - namely, to compare different ways of collaborating so that we can make a provisional determination about the ways of collaborating which yield good collaboration in terms of the above metric. In the next section we concentrate on two specific policies for managing dialogue roles in collaborative situations. One is as close to a commonly accepted form of collaboration as possible. The other is an example of a combination of policies which, taken together, turn out to produce more effective collaboration.

EXPERIMENTS WITH CLARISSA

Clarissa has been introduced as a test-bench for examining collaborative activity [13]. Above, Clarissa's architecture has been examined.

The four fundamentally different environmental situations that have been defined include *Free*; *Multi*; and *Swap* [14].

The constraints on the way in which these environmental conditions can be combined are given in Table 1. Not all possible combinations are useful. It makes no sense to combine *Free* with *Multi*, since *Multi* is a superset of *Free*. However either can be combined with *Swap* [15]. Table 1 shows the allowable combinations.

Table 1. Sensible Combinations of Environmental Conditions

{PRIVATE}	Free	Multi	Swap
Free			Yes
Multi			Yes
Swap	Yes	Yes	

The environmental conditions which have been identified here are examples of what can be achieved with Clarissa. They are not the only constraints which could be used.

As an example of the types of environment that Clarissa can be used to investigate, two situations are examined here. Further work on other environments can be found in [Burton, 1998]. However, the two selected environments provide an indication of the extremes obtained - i.e. *MultiSwap* produces better results (according to the metric supplied) than either *Multi* or *Swap*.

The Context

In the first instance, Clarissa is set up to behave in what is regarded as a 'normal', or *Free* way. In this mode, Clarissa normally expects only one agent to use any one role at any time.

This form of 'normal' collaboration is then compared with a second, which is referred to as *MultiSwap*. This latter collaborative situation is constructed from two separate constraints. Firstly, the social norms used in the 'normal' mode are relaxed. For *MultiSwap*, agents are allowed to use the same role at the same time. The resulting conversations will be strange in that the dialogue will seem to lose its coherence to some degree as questions are asked together, and agents may 'reply' with other questions. But Clarissa agents are assumed to have a very good memory, and they will return to these questions, and answer them more directly when they can. Memory is a key feature of this collaborative situation, and must not be overlooked in the interpretation of any results.

Secondly, constraints are added that were suggested by Soller who noticed that in successful collaborative groups (in terms of her metrics), participants tend to swap roles at the beginning of a new episode [Soller, 1997].

Soller identified a number of dialogue markers to define 'episodes', and defined her own set of roles which are related to those of Clarissa [Soller *et al*, 1998]. The expectation seems to be that, while an episode boundary may be a reasonable place to expect people to drop the roles they are playing, it does not follow that they would choose different ones for the next episode. She has found that in the cases where people do choose different roles, collaborative activity is more beneficial (according to her measures).

A very similar situation can be modelled by Clarissa using the opening of a new dialogue game as an indication of a new dialogue episode. These episodes will be somewhat smaller than Soller's, but a similar positive effect can be expected [16].

The participants will stop using the roles they have been using every time an individual starts a new dialogue game, i.e. at the beginning of a dialogue game which stems from the root node of the dialogue game tree. The participant who uttered the initial move which started this new dialogue game, will keep the role that they used to play that move. For them to drop that role would make no sense, since they have just started using it!

Method of Analysis

For all the situations, the system was set up to run 100 runs of two Clarissa agents collaborating on the standard test problem.

It requires a careful analysis to find out whether or not, in a sample of 100 pairs, one agent consistently uses a role significantly more than another. We need to look at the distribution of the differences between the agents in terms of role usage.

The distribution of interest is the difference in the number of utterances i.e. that of |A1-A2| where A1 and A2 are the number of utterances made by an agent for a given role. We can more easily examine the distribution of the sum of the number of utterances and use the results to compute the required properties of the positive difference of the number of utterances.

The analysis of this non-standard distribution can be found in [Burton, 1998]. The 1% level at which a difference becomes so great that we can assume the agents are not getting an equal share has been calculated. The conclusion is that the mean of |A1-A2| should not be greater than $2.61797 \times \frac{1}{\sqrt{2}} SD_{\mathrm{A1+A2}}$ where $SD_{\mathrm{A1+A2}}$ is the standard deviation of the distribution of sums of utterances made by an agent for a given role.

In addition to knowing the degree to which agents are 'split' by the collaborative environment within a pair, so that one performs differently from the other, we can also discover something about the differences in their performance by looking at the correlation between the number of utterances of the agents.

For good collaboration, it is hoped that agents in a pair perform equally well. The above statistic will tell us whether it is likely that the differences between the participants in a pair can

be accounted for simply by the variation in the population's behaviour. Ideally, we would like an agent in a pair that uses a given role a great deal to share this usage with the other agent.

In other words, we would like a positive correlation between the two agents for all of the roles investigated. Negative correlations are likely to result in significantly split role usage. A negative correlation implies that the collaborative environment, of itself, is inducing the agents to divide the roles unevenly. For populations of 100, the Pearson Rank Coefficient at the 1% level is 0.254. In other words, the correlation is significantly greater than zero, at the 0.01 level when the coefficient is calculated, any value above 0.254, or below -0.254.

Free Collaboration

In order to examine the *Free* collaboration situation, the system was set up to run two Clarissa agents collaborating on the standard test problem for a hundred runs for each of the conditions.

For the sake of brevity, the effects on each role will not be examined, but we concentrate on one (representative) role, and the number of utterances each individual agent received from their partner which caused them to change their knowledge base. These are referred to as 'interesting' events.

The following roles are defined: 'generate'; 'reason'; 'check'; 'question'; 'response'; 'argue'; 'IWasThinking'; 'interface' and 'Finished'. The utterances defined to be within the 'generate' role permit the agent to initiate cognitive goals to begin to look at parts of the problem. The 'reason' role includes moves which agents use to describe their findings as a result of these cognitive goals being achieved. The check role is used by agents to indicate that they think there is some aspect of the problem that their current solution does not address. The question role is different from the check role in that it is more general, and allows the agents to raise questions about various aspects of the domain. The respond role allows agents, naturally enough to respond to these questions. The argue role consists of moves used either to indicate that the agent agrees or disagrees with these responses. Subsequently the agent may then go on to use the question or response role to follow up disagreements. The use of the mouse is associated with the interface role, and agents declare themselves happy with the solution with the Finished role. 'I Was Thinking' moves are allowed in the IWasThinking role.

The rules in Table 2 are then combined with these role definitions to complete the dialogue role definition in this first case.

Table 2. Dialogue Game Role Constraints
Definition of which dialogue roles can/cannot co-occur. Note that != means 'cannot co-occur', and == means 'may co-occur'.

{PRIVATE**}**question != response	Question != argue
generate == reason	Generate != check
generate != interface	Generate != question
generate != response	Generate != argue
reason != check	Reason != interface
response != generate	Response != reason
response != question	Response != argue
argue != reason	Argue != check
argue != generate	Argue != question
argue != response	
Finished != question	Finished != generate
Finished != reason	Finished != argue
Finished != response	Finished != check
Finished != interface	Finished != IWasThinking

In the *Free* case, 5 out of the 7 roles, and the number of interesting events (this is included in the metric, so this gives 6 out of 8) are significantly negatively correlated. As one agent used a role more, the other uses it correspondingly less, and similarly for the degree of interest. Correspondingly, all those roles which are significantly negatively correlated are also significantly split (both measures are significant at the 1% level). The second of these two measures indicates that not only does one agent use a role more when the other uses it less, but that in any one pair of agents, one agent uses the role significantly more than the other. Pearson's correlation coefficient indicates a propensity to do so, the 'split' measure indicate that they actually do divide up the way in which they use the role. This is demonstrated by the 'reason' role with statistics represented in Figure 4.

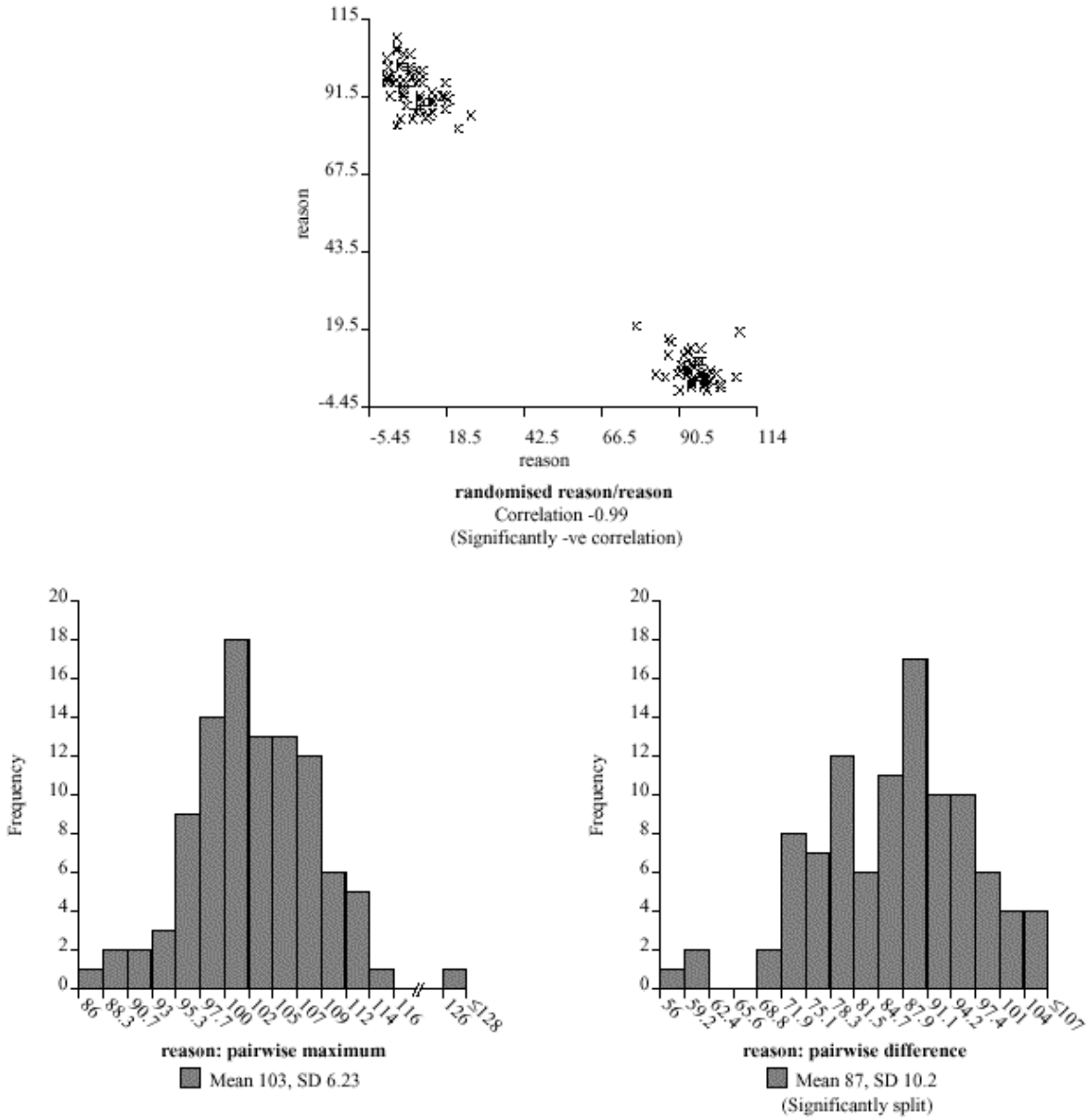


Figure 4: Statistics for the 'reason' role for *Free* Collaboration

The scatter-graph (randomised reason/reason) plots the degree to which one agent uses the 'reason' role against the degree to which the other agent uses the same role.

The correlation statistic associated with this graph is extremely large, -0.99, and highly significant [17]. It indicates that as one agent uses the role more, the other uses it

correspondingly less. This graph also indicates that there is a very great difference between the two agents. One *always* uses the role a lot more than the other which *always* uses the role a lot less.

The third graph (bottom right) gives more information about this split. It indicates how big the difference is between the two agents. For pairs who use a role approximately the same amount as each other, it would be expected that this should be a non normal distribution as described above. The result suggests a significant split in the distribution of the 'reason' role. This indicates that while the average for the pair considered together might be fine, that within the pair, one agent is using the role significantly more than the other, and this is consistent across all the runs. The inference is drawn that it is the collaborative situation which is encouraging the agents to behave in this way.

While the analysis of the reason role provides a clear example, it demonstrates the analysis which is being carried out in order to arrive at the overall statistic that 6 of the 8 roles are both significantly split and negatively correlated. Included in the metric is the degree to which agents find each other's utterances useful (not itself a dialogue role).

This 'free' collaborative situation is characterised by a tendency for participants to use dialogue roles unevenly, and hence unevenly distribute their cognitive processes. The results outlined suggest that if all other factors are equal, students that collaborate together, with no control on their dialogue, are likely to benefit unequally from the experience. The first student who takes control of the conversation will remain in control, while the other participants adopt a more subservient role. This is undesirable, but possibly common in small group work. The other collaborative situation is now considered.

MultiSwap Collaboration

Two Clarissa agents were set up to collaborate using the *MultiSwap* mode of collaboration. A hundred runs produced results shown in Figure 5 which shows that there is an improvement over *Free* collaboration. The data set show no split or negative correlation, and this is common across all the roles. In the case of the 'reason' role there is a significant positive correlation (the coefficient is greater than 0.254). This is highly desirable as it implies that if one participants uses the role more, the other will follow suit.

Although there has been no detailed discussion of the other experiments undertaken, it is worth giving some idea of how *MultiSwap* relates to *Free* and some of the other possible situations for the number of interesting events and 5 of the 7 roles [18]. As can be seen *MultiSwap* gives distinctly better results relative to the chosen metric than *Free*, *Swap* and *Multi*.

{PRIVATE}	Interesting	Reason	Generate	Response	Interface	Question
Free						X
Swap				X		X
Multi		X	X	X		X
MultiSwap	X	X	X	X	X	X

X indicates that there is neither a significant split nor a significant negative correlation for the given role and policy

MultiSwap collaboration is much better according to the metric used here, especially as it seems to encourage the even distribution of all dialogue roles. This implies that all the participants in the collaboration will have an opportunity to practice all of the dialogue roles that are available, and by doing so, the participants should have the opportunity to practice and hopefully improve their ability to execute those underlying processes.

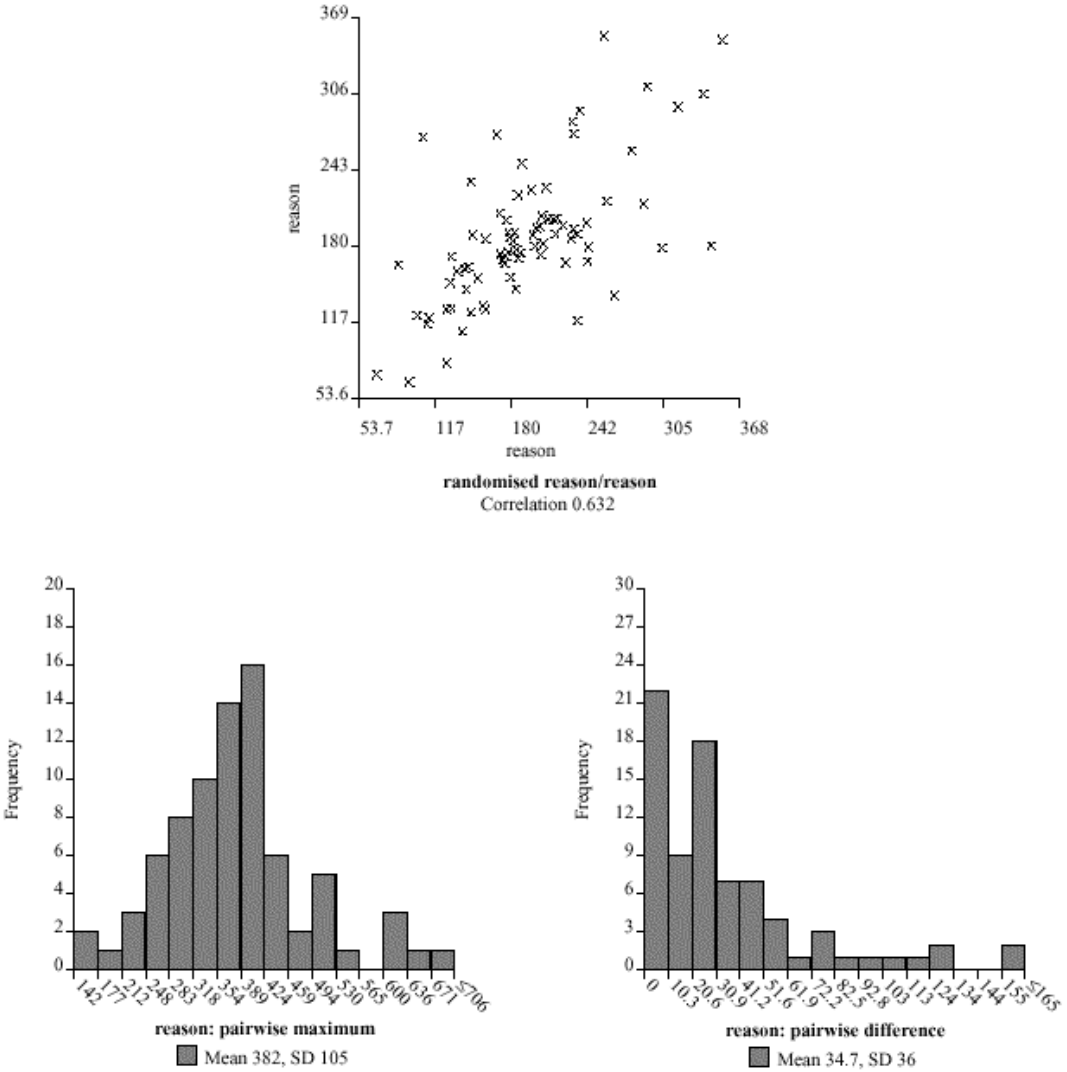


Figure 5: Statistics for the 'reason' Role, in a *MultiSwap* Collaboration

DISCUSSION: FREE VS MULTISWAP COLLABORATION

Clarissa is intended to provide a workbench with which to investigate collaboration. Two example forms of collaboration have been demonstrated and compared. The interest is in collaborative activity which promotes an even distribution of cognitive processes. To measure this, the roles that are played have been examined. The roles that have been played have not been enforced directly. Nor has it been specified when the role is to be played and by whom; rather, collaborative situations have been set up which have various characteristics in terms of the way roles are swapped. *Free* collaboration does not promote an even distribution of roles, but *MultiSwap* does.

This result should be examined in terms of dialogue within a real educational context. If *MultiSwap* collaborative activity is to be recommended, a clear understanding is needed of what this entails in practice.

MultiSwap introduces two sets of constraints. The first involves a change in the social norms. In an educational context this could conceivably be implemented by a small amount of prior training. Collaborators should speak their mind, ask all the questions they have, say

everything they can about a problem. While this suggests they may be disregarding their partners it is crucial that they pay careful attention. Questions, comments and suggestions their partners make must be noted down. The human participants would need some additional assistance to allow them to keep track of what everybody had said, (this is common in most meetings, and often some such device, like paper and pencil, are encouraged in educational collaboration).

The second set of constraints implies that participants should keep asking questions, if that's what they start doing; keep making statements about what is known about the problem; keep proposing deductions, suggesting things to do next. Participants should continue concentrating on one 'role' until the beginning of a new 'dialogue episode'. As a new dialogue episode begins, they should stop playing that role, and possibly take up a new one.

A dialogue episode may not an easy thing to identify from the perspective of a collaborative participant. Furthermore, the rules described are not so easy to follow. To address this problem, further tests have been carried out with a form of collaboration referred to as *Polite*. The effects have been very similar to those reported above. 'Polite' collaboration simply involves participants 'dropping' their roles at the *end* of episodes. Rather than swapping roles at the beginning of a new episode, the participant who has lead an episode stands back for the next. In other words, and simply put, if you have been taking the lead for a while, stop, and let somebody else take the floor.

It must be emphasised that the finding reported here is derived from analysis of the simulation of a model. The model is cognitively plausible subject to the many simplifications that have been made. For example, the architecture of human cognition is not as cleanly delineated. Nor does Clarissa currently include any modelling of the learning of an individual agent. Thus the communicative processes that agents take part in are not internalised, and therefore the model would need to be extended to provide a fuller account of the processes involved - especially to model the kind of collaborative activity that takes place over a much longer period of time.

Further work might also be productively coupled to research into some of the affective issues that apply to small groups. For example, the work on how the perceived status of group members can influence their degree of participation. Cohen and Lotan have indicated that seeking to raise status in quite a direct manner can result in more equal participation in dialogue [Cohen & Lotan, 1995]. There has, as yet, been no deep consideration of whether this effect could be modelled within Clarissa. Indeed, it may be that the results with Clarissa throw a little light on how the modification of perceived status may be related to the dialogue mechanisms that lead to evenly distributed patterns of dialogue role within small groups.

Research with Clarissa therefore suggests ways in which others may take this research further, having decided upon their pedagogic requirements, and the dialogue roles that they hope will fulfil these.

CONCLUSION

A theoretical basis has been provided for a model of collaboration which has been implemented, and used to both explain some of the previous findings about collaboration, and to make predictions.

The work has concentrated on the *dialogue* between a number of students, and the roles associated with this dialogue. Collaboration has been defined in terms of the distribution of these roles and it has been shown how this definition is both computationally elegant and sufficient to generate collaborative dialogue.

A system, Clarissa, has also been presented which embodies this approach. The system provides a workbench for modelling collaboration in terms of dialogue roles and a set of dialogue role allocation mechanisms.

The architecture of the system has been described, and key implementation issues discussed. Examples of how the system can be used have been presented.

In future work, Clarissa agents may be redesigned for use as agents that can be embedded into such learning environments. While not primarily intended for such use, they have been designed with this in mind. The technologies used to build Clarissa are directly applicable to others working in this field, especially at the level of adopting roles as the primary modelling mechanism, removing the need to maintain shared beliefs.

Clarissa supports any number of agents, including teacher/ facilitator agents who manages the role allocation process dynamically and monitoring agents who can collect data on the collaboration. Many more possibilities exist for productive interactions than have yet been explored.

The system has been used to examine various relationships between the available roles, the role allocation mechanism and the resulting properties of the dialogue generated. The following results obtained from examining the computational model have been found and outlined in this paper:

- The 'best' collaborative situation involves participants having the ability to communicate in a 'non-normal' way, and have some form of aide-memoire which would allow this abnormal communication to be effective.
- Socially 'normal' role usage is not conducive to educationally beneficial collaboration. It seems that one agent takes control of the conversation, and the social norms are such that the agent remains in control. This problem seems to be addressed by the multi-role distributions. Allowing participants to adopt the same roles at the same time is not socially normal, but seems to have a very significant effect on the quality of collaboration.

The results reported here suggest that *normal* collaboration is not very effective. More conducive to learning would appear to be collaborative situations in which roles are kept until a new topic is initiated or the focus is changed, participants have access to external representations of key aspects of the dialogue, and finally, participants are polite and allow each other a turn at leading the discussion.

The finding that *MultiSwap* and *MultiPolite* are better than the *Free* mode goes against a certain 'folk' view of collaboration as being more desirable if there is only one questioner at a time, only one informer and so on. The challenge is to examine these results in terms of real human-human collaboration.

Clarissa has also proved to be a complex but powerful workshop. While some potentially useful features have not been incorporated, it has proved useful in investigating the generation of collaborative dialogue. There is now a challenge to apply the techniques used in Clarissa in educational environments, including computer supported collaborative systems.

Acknowledgements

Mark Burton was supported in this work through an EPSRC PhD Studentship. The authors wish to thank the many insightful comments of the reviewers.

References

Aaronson, E., Blaney, N., Stephan, C., Sikes, J. and Snapp, M., (eds.). (1978). *The Jigsaw Classroom*. Sage Publications, Beverly Hills, CA.

Anderson, J.R. (1983). *The Architecture of Cognition*. Harvard University Press, Cambridge, MA.

Baker, M. and Lund, K. (1996). Flexibly structuring the interaction in a cscl environment. In Brna, P., Paiva, A. and Self, J.A., (eds.), *Proceedings of the European Conference on Artificial Intelligence in Education*, pages 401-407. Edicões Colibri, Lisbon.

Bental, D. and Brna, P. (1995). Enabling abstraction: Key steps in building physics models. In Greer, J., (ed.), *Proceeding of the World Conference on Artificial Intelligence in Education*, pages 162-169, Charlottesville, VA. AACE.

Bielaczyc, K., Brown, A. and Pirolli, P. (1994). Collaborative explanations and metacognition: Identifying successful learning activities in the acquisition of cognitive skills. Xerox Parc Technical Report CSM-8, Xerox PARC, Palo Alto, California.

Blandford, A.E. (1994). Teaching through collaborative problem solving. *Journal of Artificial Intelligence in Education*, 5(1):51-84.

Blaye, A., Light, P., Joiner, R. and Sheldon, S. (1991). Collaboration as a facilitator of plannning and problem solving on a computer-based task. *British Journal of Developmental Psychology*, 9:471-483.

Boden, M. (1979). *Piaget*. The Harvester Press, Brighton.

Burton, M. (1998). *Computer Modelling of Dialogue Roles in Collaborative Learning Activities*. Unpublished Ph.D. thesis, Computer Based Learning Unit, The University of Leeds.

Cohen, E.G. and Lotan, R.A. (1995). Producing equal-status interaction in the heterogenous classroom. *American Educational Research Journal*, 65(1):99-120.

Damon, W. and Phelps, E. (1989). Critical distinctions among three approaches to peer education. *International Journal of Educational Research*, 13:9-19.

de Vries, E. (1995). Hypermedia for physics learning: Activating prior knowledge? Rapport No. CR-12/95, Equipe COAST, UMR GRIC, CNRS-Université Lyon 2.

Devi, R., Tiberghien, A., Baker, M. and Brna, P. (1996). Modelling students' construction of energy models in physics. *Instructional Science*, 24(4):259-293.

Dillenbourg, P. (1992). The language shift: a mechanism for triggering metacognitive activities. In Winne, P. and Jones, M., (eds.), *Adaptive Learning Environments: Foundations and Frontiers*, volume 85 of *NATO ASI Series F: Computer and Systems Sciences*, pages 287-315, Springer-Verlag, Berlin.

Dillenbourg, P., Baker, M., Blaye, A. and O'Malley, C. (1994). The evolution of research on collaborative learning. In Spada, H. and Reimann, P., (eds.), *Learning in Humans and Machines*, Pergamon, Amsterdam.

Harwood, D. (1995). The pedagogy of the world studies 8-13 project: The influence of the presence/absence of the teacher upon primary children's collaborative group work. *British Educational Research Journal*, 21(5):587-611.

Levin, J. and Moore, J. (1977). Dialogue-games: Metacommunication structures for natural language interaction. *Cognitive Science*, 1(4):395-420.

Mackenzie, J.D. (1985). No logic before friday. *Synthese*, 63:329-341.

Maybury, M.T. (1993). Planning multimedia explanations using communicative acts. In Maybury, M.T., (ed.), *European Conference on Hypertext'94*. AAAI Press/The MIT Press, Menlo Park, CA.

McCoy, K.M. and Cheng, J. (1991). Focus of attention: Constraining what can be said next. In Paris, C.L.. Swartout, W.R. and Mann, W.C. (eds.), *Natural Language Generation in Artificial Intelligence*, chapter 4, pages 103-124. Kluwer, Amsterdam.

Megalagaki, O. and Tiberghien, A. (1995). Apprentissage de la modelisation par la rèsolution successive de problemes. In *Proceedings of European Conference on Cognitive Science ECCS-95*, Saint-Malo, France.

Mehan, H. (1979). *Learning lessons: Social organization in the classroom*. Harvard University Press, Cambridge, MA.

Minsky, M. (1987). *The Society of Mind*. MIT Press, Boston, MA.

Newell, A. (1990). *Electronic Information Handling*. Harvard University Press, Cambridge, MA.

Pilkington, R.M. (1999). Analyzing educational discourse: The DISCOUNT scheme. Technical Report 99/2, Computer Based Learning Unit, Leeds University.

Rao, A.S. and Georgeff, M.P. (1991). Modelling rational agents within a BDI-architecture. In Filkes, R. and Sandwell, E., (eds.), *Proceedings of Knowledge Representation and Reasoning (KR&R-91)*, pages 473-484. Morgan Kaufmann, San Fransisco.

Roschelle, J. and Teasley, S. (1995). The construction of shared knowledge in collaborative problem solving. In O'Malley, C.E., (ed.), *Computer Supported Collaborative Learning*, pages 69-97. Springer-Verlag, Heidelberg.

Salomon, G. and Perkins, D.N. (1998). Individual and social aspects of learning. *Review of Research in Education*, 23:1-24.
Sinclair, J. M. and Coulthard, R. M. (1975). *Towards an Analysis of Discourse: The English used by teacher and pupils*. Oxford University Press, Oxford.
Soller, A., (1997). Personal Communication.
Soller, A., Goodman, B., Linton, F. and Gaimari, R. (1998). Promoting effective peer interaction in an intelligent collaborative learning system. In Goettl, B.P., Halff, H.M., Redfield, C.L. and Shute, V.J., (eds.), *Intelligent Tutoring Systems: 4th International Conference, ITS'98*, pages 186-195. Springer-Verlag, Berlin.
Stubbs, M. (1981). Motivating analyses of exchange structure. In Coulthard, M. and Montgomery, M., (eds.), *Studies in Discourse Analysis*. Routledge, London.
Teasley, S. D. and Roschelle, J. (1993). Constructing a joint problem space: The computer as a tool for sharing knowledge. In Lajoie, S.P. and Derry, S.J., (eds.), *Computers as Cognitive Tools*, pages 169-197. Erlbaum, Hillsdale, NJ.
Tiberghien, A. (1994). Modeling as the basis for analyzing teaching-learning situations. *Learning and Instruction*, 4:71-87.
VanLehn, K., Jones, R.M. and Chi, M.T.H. (1992). A model of the self-explanation effect. *The Journal of the Learning Sciences*, 2(1):1-59.
Vygotsky, L. (1978). Mind in Society: The Development of Higher Psychological Processes. Harvard University Press, Cambridge, MA.
Wood, D., Wood, H., Ainsworth, S. and O'Malley, C. (1995). On becoming a tutor: Toward an ontogenetic model. *Cognition and Instruction*, 13(4):565-581.

APPENDIX A: CLARISSA'S DIALOGUE ROLES

A list of the main dialogue roles utilised in the studies with Clarissa follows. The roles are sketched in terms of what kinds of activities are associated with each role.

{PRIVATE} Dialogue Role	**Brief Description**
question	raise issues of all kinds
response	seek to satisfy questions
argue	challenge or support statements
generate	examine the current problem in terms of possible approaches
reason	explain chains of reasoning
check	see if the current partial solution is adequate
interface	manage interactions with the external representation of the problem
finished	indicate that there is nothing more that can be done
IWasThinking	bring (possibly old) topics back into consideration

Footnotes

[1] Roschelle and Teasley have defined a joint problem space in which participants work together, and (depending upon the state of maturity of the partners) where they may become exposed to a zone of proximal development. However, while this suggests that constructing a joint problem space may be important, it does not directly suggest the processes by which participants should achieve this. For that, a better understanding of the processes involved in learning is required.

[2] Throughout, practise includes observation as well.

[3] Not to be confused with the processes connected with learning new knowledge etc. The term encompasses a wide range of possible processes. It is not the concern of this work to discuss these further.

[4] This is clearly an assumption that only approximates reality.

[5] Indeed, very often arguments revolve around how dialogue moves should be interpreted.

[6] Dialogue moves are associated with dialogue utterance types. Given a dialogue move then a dialogue utterance type needs to be selected. Then an instance of the type has to be generated.

[7] The differences between the implementation of the cognitive system and ModelCHENE will not be discussed in any great detail here.

[8] There is no theoretical upper limit on the number of such agents except one imposed by the physical limitations of the computer's hardware and software.

[9] Energy, heat... students are very creative in labelling these arrows!

[10] While the language of the interpretation is deliberately anthropomorphic for the purposes of the exposition, the mapping between the anthropomorphism and the corresponding system actions is exact.

[11] The term 'Theory Model' is terminology adopted from Tiberghien's work [Tiberghien, 1994]. A human student is more likely to say "what [kind of thing] is battery1?".

[12] Forgetfulness can be simulated in Clarissa.

[13] Clarissa was partly derived from a study of human-human dialogue on the same task and using effectively the same 'free' policy of collaboration described below. Burton has provided a comparison between the dialogue of two such human students and two Clarissa agents [Burton, 1998]. This analysis could be taken further.

[14] It is also possible to arrange that the agents have different cognitive capabilities, which is referred to as 'cognitively split' (*Split*). In the current case for Clarissa, one agent does not operate the checking reasoner, while the other does not operate the generate reasoner.

[15] 'Cognitively split' environments can be used with any of the others.

[16] This has been tested with Clarissa. The results are not examined here, but they do indicate that this form of collaboration is better, both in terms of the way in which roles are distributed between partners, and marginally in terms of the degree of interest found in the dialogue by participants.

[17] Such a high correlation suggests that some design decision is responsible.

[18] The discussion of the check and argue roles is more complex.