The Fifth World Conference On
Integrated Design & Process Technology
Incorporating
IEEE International Conference on
Systems Integration
June 4-8, 2000
Crowne Plaza Hotel, Dallas Texas

Integrated Environments for Virtual Collaboration: an Empirical Modelling Perspective

Meurig Beynon
Department of Computer Science
University of Warwick, Coventry CV4 7AL, UK
Email: wmb@dcs.warwick.ac.uk

Soha Maad
Department of Computer Science
University of Warwick, Coventry CV4 7AL, UK
Email: soha@dcs.warwick.ac.uk

ABSTRACT

Collaboration is represented in many different kinds of activity: electronic management of document workflow for a group-work project, sharing information and knowledge across an enterprise, task scheduling in a networked environment, and group activities such as electronic trading and commerce or distance learning. Current trends are directed towards supporting such collaboration via multiple integrated communication channels, such as the internet, telephone, fax and video.

Building a collaborative environment might appear to be simply a technical task involving the set up of appropriate communication protocols between several applications running on different devices. However, for a variety of social and technical reasons, most integrated collaborative environments are failing to give adequate support to real collaboration. For instance, the challenges to be faced in building an integrated environment for collaboration include: resolving conflicts of interest in a group activity, maintaining a synchronized flow of collaborative tasks, and respecting the agency and dependency that operates in communication between human agents. It is also of particular practical importance both to ensure that the cost of deploying a distributed solution is in proportion to its benefits, and to integrate the media for front-end data capturing and communication with those for back-end persistent storage of information.

In order to construct more effective integrated environments for collaboration, new principles for systems representation and analysis that can cope with the complexity of distributed human computer interactions and their consequences are of critical importance.

Distributed Empirical Modelling offers a new paradigm that can be applied to human collaboration. It can potentially simulate interaction in the real world, where many people collaborate to accomplish a group work project, share knowledge to enhance their educational skills, or compete to maintain a favourable position in a challenging trading environment.

1. INTRODUCTION

This paper overviews several different kinds of collaborative activity. It highlights the major integration issues that arise in developing a collaborative environment from a technical, social and economic perspective. Virtual collaboration is discussed within the framework for analysing human information behaviour proposed by Sonnenwald (1998). This also provides the setting for an account of Empirical Modelling (EM), an approach to knowledge construction and communication that makes essential use of computer-based artefacts termed interactive situation models (ISMs). The potential applications of EM to virtual collaboration are described with reference to illustrations drawn from three contexts. These address the use of spreadsheets in the examination process, the design and use of a Virtual Electrical Laboratory and an agent-oriented analysis of online trading.

2. FORMS OF VIRTUAL COLLABORATION

In this paper the term *virtual collaboration* refers to collaboration via an electronic medium. The following paragraphs overview different forms of virtual collaboration, and discuss their uses and limitations.

a) Project/group work collaboration

In current group-ware, documents and document-related processes define the logical context for collaboration (Marshak, 1998). This does not necessarily mean that a document-centric approach is deemed sufficient to support collaborative work; it may just be that it is an approach that can be most simply assisted by technology. But in as much as current attempts to capture corporate knowledge rely heavily on static documents, it is important that document-centric collaboration should be more structured. When teams collaborate on a project, the documents that result from their collaboration need to be maintained and managed. Technologies for collaboration then centre around tools to create, share, and distribute documents. The internet is the most common platform for document-centric collaboration. However, as observed by Ciancarini et al (1999), the web in its current state does not provide support for document-centric applications like group-ware or workflow that require sophisticated agent coordination. In this context, the term *agent* refers both to entities which can act autonomously and can receive/send messages according to some pre-defined protocol, and to human agents with assigned roles in the group work activity or project.

b) Collaborative learning

Early attempts at online collaborative education were motivated by the desire to explore technical advances in networking and communication rather than by well-defined educational goals. Experience has shown that generic network tools, such as e-mail, computer conferencing, and newsgroups, are weak in supporting collaborative learning (Harasim, 1999). This is attributed to several factors. These include: the lack of standardized ways to organize educational material; the overhead work to manage and monitor students' performance; and the lack of models to support learning strategies that involve knowledge building and sharing. Current online educational tools support collaborative learning and course management to a greater extent. These tools are mainly web-based, and include personal workspaces for students, course structuring, grade management, file management, and system management utilities. The Virtual-U web based learning environment (see www.vu.vlei.com) is one example of online learning tools. In addition to online educational tools, research and assessment tools have also been developed to study and analyse the behavior and teaching/learning processes. Today the main challenge facing the development of web based collaborative learning environments is to support richer interaction with web pages than mere front-end access to static information.

c) E- business

The telecommunications revolution and the growth of internet activity have challenged traditional business models by offering direct routes to market, reducing barriers to entry and increasing the efficiency of trade activity (Taylor et al, 1999). E-business is an umbrella term for e-commerce, supply-chain collaboration, online trading, and business to business online communication. Business today is converging on the internet. This creates a great opportunity for organizations to communicate and share data over the web with customers, partners and suppliers. However, the growth and profitability of e-business activity is inhibited by many technical and social problems, including:

- technology integration (the integration of the back office and front office systems)
- the adoption of a common e-business model
- security
- the introduction of new national and international legislation that protects the rights of all business parties in executing cross-border transactions
- the deployment of a low cost efficient solution for true flexible business collaboration
- the high risks of system failure.

3. ISSUES FOR VIRTUAL COLLABORATION

Following D. H. Sonnenwald (1998), collaboration amongst individuals engaging with information resources in information creation, exploration, seeking, filtering, use, dissemination and communication is here viewed as an example of *human information behaviour*. Sonnenwald discusses human information behaviour with reference to three basic concepts: the context, the situation and the social network:

- the *context* is the general setting within which an individual's interactions take place. Academia, family life, citizenship, clubs etc. are examples of contexts. A context is defined by a set of past, present and future situations.
- a *situation* is a particular setting for an interaction within a context. Teaching a course or attending a committee meeting are examples of situations within academia.
- a *social network* is defined by characteristic patterns and resonances of interaction between individuals within a context. In academia, the social network associated with teaching activity might comprise professors, lecturers, teaching assistants, secretarial and technical support staff and students.

Sonnenwald's analysis of human information behaviour provides a useful perspective for the work described in this paper. In her view, the goals of a collaboration are the sharing of meaning and the resolution of a lack of knowledge condition. For each individual, collaboration within a given situation and context is bounded by their *information horizon*, as defined by the variety of information resources upon which they can draw. These resources can include individuals, such as colleagues, team members or information professionals, as well as papers, web pages, journal articles or direct observation and experiment in the world. In investigating virtual collaboration, there is an important distinction to be made between information resources that can be accessed electronically, and those that are accessed by other means. In effect, each individual has both an information horizon and a *digital information horizon*.

Human information behaviour is both complex and variable. There are many technical challenges to be met in providing support for virtual collaboration. These include complex and dynamically evolving requirements, as motivated by several key issues:

Customisation: Electronic support for collaboration has to take account of the needs of the individual within the social network, situation and context. Identifying and developing algorithms and interfaces to support such behaviour typically requires a high degree of customisation. In particular, electronic support must be well-adapted to the information horizons of the participants.

Integration of the electronic and human activity: Human information behaviour necessarily involves a close interplay between human and automated activity. It is essential that virtual collaboration retains its situated character, so that the information processing activity is appropriately matched to the state of the external world.

Adaptation: The extent to which it is possible and desirable to augment or automate aspects of human information processing capabilities is highly dependent upon personal, technological and social factors. The information horizons of participants are typically neither static nor easily preconceived. They can also be influenced in a deliberate way by the actions of participants. Because of these factors, the requirements for a virtual collaboration are subject to continuous evolution.

A major obstacle to successful virtual collaboration is a fundamental mismatch between the roles that humans and electronic devices play in communication and interaction. This is well illustrated in current practice by products such as document-centric environments for collaborative work. For the human participant, a document is in general full of significance that eludes formal computational representation. Its meanings are rich, ambiguous and contextually determined. In communication about documents, the human interpreter generally exercises discretion, checking the integrity of interpretations with reference to external observation, or feedback from the person or device with which they are communicating. In contrast, an electronic device records and transmits information according to formal pre-conceived conventions, and - if it monitors external state at all - does so in ways that are highly constrained. Static conventions for representing information limit the extent to which the significance of external experiences can be electronically recorded and conveyed. An electronic device is subject to act without discretion, oblivious to its environment. This can lead to catastrophic failure should singular conditions arise.

The traditional approach to resolving these problems of mismatch is to constrain the interaction between humans and electronic devices to patterns for which a very high level of consistency can be guaranteed. As the above analysis has indicated, this approach is not well suited to the volatile practical demands of effective virtual collaboration. Its limitations are apparent in all three applications introduced in section 1. Unless these applications operate in stable environments where consistent patterns of interaction can be identified and exploited, the analysis of content and communication in document-centric collaborative work environments is primarily syntactic; the evaluation of user input in intelligent tutoring systems is stereotyped and semantically superficial; automated decision-making in e-business environments is inadequately guided by the high-level interpretation of actions.

Constraining interaction so as to guarantee reliable and consistent responses from devices affects the quality of human contributions to collaborative activity. Experiential and situational elements play a vital part in human interaction. In a virtual interaction, 'no response' admits quite different interpretations from 'no response' in a face-to-face encounter. Such issues motivate the integration of different communications technologies, such as telephones, computers and set-top boxes. To explore this integration effectively, it is not enough to view electronic devices and their interaction in abstract computational terms. The appropriate emphasis is on electronic components and software applications as mediators of state and experience. Successful integration in these terms entails the assimilation of devices and applications into their environment as instruments. Alternative principles suitable for studying automatic agency from this perspective are not only relevant for developing systems to support human information behaviour. They have an essential part to play in the evaluation of environments for virtual collaboration. They can also be used to assess the intrinsic limitations of existing systems and applications. It would be patently absurd to try to integrate computation with batch cards into a fly-by-wire system, but it is more difficult to assess whether (e.g.) current web and database technology is appropriate for virtual reality.

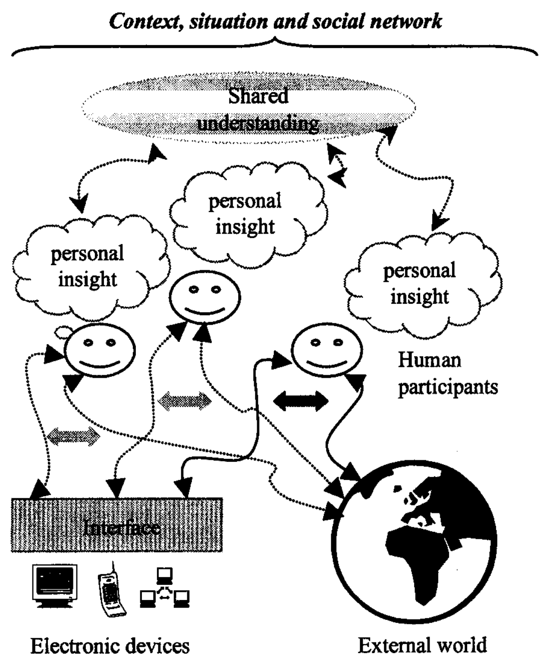

Figure 1. Correlating states in an Empirical Modelling framework

The principal aim of this paper is to introduce new principles for computer-based modelling – Empirical Modelling (EM) – that can potentially address these concerns. The key idea in EM is to establish a more intimate relationship between the computational activity and the human action and interpretation associated with each situation. To this end, participants must be able to engage with electronic components and the external world simultaneously, operating through interfaces that complement – rather than obstruct or detract from – interaction with the external environment. Close co-operation between human and computer is achieved through a computational framework in which electronically-mediated state and human perception of the external situation are readily correlated (cf. Figure 1). The aim is to complement the power of the computer to automate action by enabling participants to intervene intelligently when human interpretation and discretion is required as input, or singular conditions arise.

4. EM BACKGROUND

4.1 EM principles, notations, and tools

Empirical Modelling is a collection of *principles*, *techniques*, *notations*, and *tools* that have been developed at the University of Warwick.

Empirical Modelling principles are based upon *observation, agency, and dependency.* By adopting these principles, EM attempts to represent and analyse systems in a way that can address the complexity of the interaction between programmable components and human agents. The central concepts behind EM are definitive (definition-based) representations of state, based on the use of families of definitions called *definitive scripts*, and agent-oriented analysis and representation of state-transitions.

Empirical Modelling techniques involve an analysis that is concerned with explaining a situation with reference to agency and dependency, and the construction of a complementary computer artefact – an *interactive situation model (ISM)* – that metaphorically represents the agency and dependency identified in this process of construing. There is no preconceived systematic process that is followed in analysing and constructing an associated ISM. The modelling activity is open-ended in character, and an ISM typically has a provisional quality that is characteristic of a current – and in general partial and incomplete – explanation of a situation. The identification of agency and dependency often exploits previous knowledge and experience, but can be thought of as being derived in essence through observation and experiment. Unlike a closed-world computer model with a fixed interface and preconceived human interaction, an ISM is always open to elaboration and unconstrained exploratory interaction.

Empirical Modelling notations include LSD – a special-purpose notation that has been introduced to classify observables and dependency in agent interaction. An *LSD account* is a classification of observables from the perspective of an external observer, detailing where appropriate: the observables whose values can act as stimuli for an agent (its **oracles**); which can be redefined by the agent in its responses (its **handles**); those observables whose existence is intrinsically associated with the agent (its **states**); those indivisible relationships between observables that are characteristic of the interface between the agent and its environment (its **derivates**); and what privileges an agent has for state-changing action (its **protocol**).

The EDEN interpreter is the principal *modelling tool* that has so far been developed. It supports definitive scripts for line drawing and window layout and allows the user to establish dependency relationships between scalars, lists and strings using built-in user-defined functions. There is a distributed variant of EDEN. This allows several modellers to co-operate through communicating definitions and actions within a client-server configuration of EDEN interpreters.

Empirical Modelling has been applied to many areas including software and systems development, engineering design, concurrent systems modelling and simulation, and education. For more details, see the EM website at www.dcs.warwick.ac.uk/modelling/.

4.2 EM and Human Information Behaviour

Empirical Modelling concepts provide a useful way to elaborate upon Sonnenwald's conceptualisation of human information behaviour. Each context has its own families of characteristic observables. For instance, in academia, the degree programme, choice of module options and examination marks associated with a student are observables, whilst the weight and height of students are outside the scope of concern. A situation within a context typically includes other pertinent observables that reflect a special focus. For instance, in teaching a course, there is a curriculum, a relevant lecture schedule, and a current point that has been reached in its delivery. Dependencies amongst observables are crucial in shaping the semantics of a situation. For instance, examination marks attained and the current point in the semester may together determine the possible choices of module options, or the entitlement to transfer to another degree programme.

From an EM perspective, situations within a context will be represented by an ISM. The scope of this ISM will be open-ended, but will typically represent a situation from many different aspects, each developed to a greater or lesser degree of sophistication according to the knowledge and experience of the modeller, and their particular focus of interest. By way of illustration, an existing ISM for a railway comprises artefacts to represent observation of the track layout, the abstract interconnections between rail segments, the position and location of trains and stations, the interactions between railway personnel and passengers at stations, and an archetypal model railway interface (Beynon et al, 1994). Like a spreadsheet, an ISM is always apprehended in a particular state, but its significance can only be appreciated through intelligent interaction and interpretation.

In EM terms, the characteristic observables, dependencies and agents in a context are represented by script fragments and associated protocols for interaction. These can acquire the status of re-usable components as the integrity of a context is established, and as understanding is gained. The ISM that represents the current situation within a context is composed of such fragments, appropriately supplemented by observables relating to the current states of processes in progress (the use of observables of this nature is illustrated in the train arrival and departure protocol in Beynon et al (1994)). Within this framework, human information behaviour is typically in part implicit in interaction with the ISM, and in part made explicit via an agent-oriented analysis that leads to an LSD account. The agents in this account include the human participants in the social network. These agents have different kinds of observables according to their roles, and are privileged to observe and influence these observables according to different protocols. Automated information processing components are also treated as agents in the LSD account.

For Sonnenwald, the information horizon for each individual is determined by the information resources to which they can in principle access. In this definition, the promptness with which information can be retrieved is not the primary issue. In EM, in contrast, the notion of an *observable* is concerned with an agent's perception of current state (cf. Figure 2). Such a distinction is important both in relation to observation of the external world and to electronic interfaces, and has significant implications for administrative processes and organisation. For instance, a teacher typically has access to pupils' marks from some information resource, but how efficiently they can respond to a pupil's request for a mark is governed by specific observables. If the objective is to devise an effective and secure protocol for disclosing results to students, it matters whether the teacher is able to identify the pupil, has direct access to a written record of pupils' marks, or has to determine a pupil's mark by consulting a third-party. In a similar spirit, working practices can be influenced by whether or not the presence of unread e-mail is registered by an icon in a computer interface.

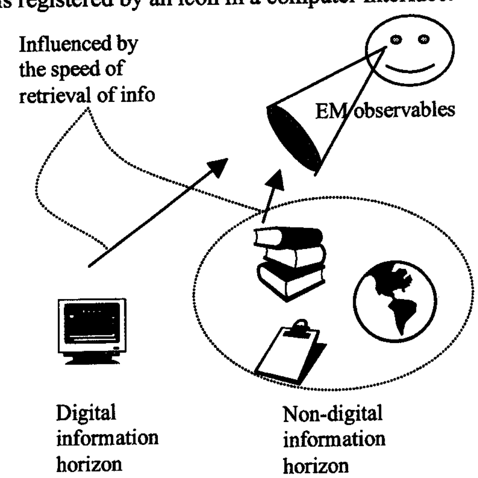

Figure 2. EM observables and the information horizon

Neither the information horizon nor the observables of an individual can be precisely prescribed. The interpretation of both is guided by situation and context. Where formal administrative processes are concerned, how a teacher gains access to a pupil's marks may be outside the scope of concern. Whether or not the presence of unread mail is deemed to be an observable for a user may depend upon their level of experience and their established precedents for use. An individual can develop skills and codes of practice that render immediately accessible what is at first difficult to ascertain. One of the principal motivations for expanding the digital information horizon is to enhance the observables and give more timely access to current knowledge of state.

4.3 Empirical Modelling and Virtual Collaboration

Empirical Modelling is centrally concerned with framing and communicating explanations for phenomena. An explanation here refers to a convincing way of accounting for perceived state-changes in terms of the interactions of agents. The key questions in this connection are: *What agents are deemed responsible for state-change? What are the cues for state-changing action on the part of agents? What are the direct effects of agent action upon the environments of other agents?* Contriving such an explanation requires evidence that is typically gathered from observation and experiment. Subjective and pragmatic judgements are involved in interpreting this evidence. It is not in general possible to give a comprehensive account of a phenomena in terms of agents and their interactions. Patterns of agency and dependency that can be reliably identified as part of an explanation can be framed as an LSD account. The evolving understanding of a phenomenon that eludes even such partial explanation is captured through developing an ISM. This ISM serves a similar purpose to the physical artefacts that an experimental scientist or engineer might construct in order to express their knowledge of a phenomena. The ISM can be regarded as representing the phenomena in the informal sense that experience of interaction with the ISM and with its referent are perceived as having characteristics in common. Creating an ISM is of its essence an open-ended activity in which the modeller takes account of ever richer perceptions of observables, dependency and agency. In the terminology introduced by Gooding (1990), an ISM serves as a *construal* of the phenomenon to which it refers.

EM principles can be used for constructing ISMs both as an individual and as a corporate activity. The principal tool that has been developed for this purpose is the EDEN interpreter, which has both standalone and distributed variants. Construction of an ISM by an individual has intimate connections with learning activities (Beynon, 1997), and corporate construction with the growth of shared understanding (Sun et al, 1998). The roles played by ISMs and LSD accounts in EM represent complementary aspects of experimental activity. The LSD account is a way of framing an explanation; the ISM provides an environment in which to explore and evaluate an explanation. EM activity may involve first framing an explanation in LSD, then generating an ISM as a test environment. Alternatively, it may involve constructing an ISM that can be used to explore possible explanations. In general it is appropriate in EM both to use prior knowledge and to seek experimental insight, and – to this end – to frame LSD fragments and incrementally construct ISMs concurrently.

From an EM viewpoint, sharing explanations and understanding is the key to effective virtual collaboration. There is evidence for this from our previous research, and from ongoing case-studies to be briefly outlined below. In relation to group project work, several previous papers (Adzhiev et al, 1994; Beynon et al, 1994) have examined the potential advantages of using EM both as a way of reaching consensus in design and resolving conflict in creative partnership and – simultaneously – as a playground for individual experiment. In computer supported education, ISMs can be used both to capture personal insights, and as vehicle for exploring and communicating understanding (cf. the models of heapsort discussed in Beynon et al (1998)). Our current research in e-business indicates ways in which EM can be used to investigate how human and automatic agents can co-operate through patterns of work flow and in decision support.

5. SOME ILLUSTRATIVE EXAMPLES

Three aspects of EM principles are particularly significant in connection with virtual collaboration: the flexibility of the processes that surround the use of models; the richness and directness of the communication between interacting agents; the scope for distributed model development that is open-ended, opportunistic and incremental. In this section, these merits will be illustrated with reference to three different contexts.

5.1 The Examination Spreadsheet

Many features of the use of ISMs in virtual collaboration can be illustrated with reference to spreadsheets. In the context of academia, the spreadsheet plays a key role in situations associated with the examination process. In a typical spreadsheet of examination marks, relevant observables include the names of students and modules, the details of which optional modules students have chosen, and the marks obtained by the students in each module. Other observables, derived from this data by dependencies, are the average marks for modules, and the overall average mark for the current year of the degree programme attained by each student. During the examination process, several agents are involved in determining the state of the final spreadsheet. Both the preparation and eventual application of the spreadsheet are integrated with rich human interpretative and administrative activity. In this activity, the norms and conventions of the social network are strongly in evidence.

The significance of the spreadsheet as an ISM stems from its role as a medium for communication between the agents involved in the examination process, the status it enjoys as a current representation of the state of play in the marking process, and the possibilities for dynamic interaction it admits. The qualities of the spreadsheet derive from its situated character and capacity for enhancing observation of state rather than the complexity of the computational task it performs. Its virtues are typically appreciated with reference to the technologies that preceded it. For instance, if it is deemed appropriate to scale the marks from one examination, the impact of this upon a student's overall mark is updated instantly. This renders directly observable what would previously have been accessible only through tiresome and time-consuming recalculation. In point of fact, such an interaction with the spreadsheet exploits its modelling power only to a very limited degree. It is also highly constrained by the external interpretation. There must be good reason to change a mark, and the procedures to correct marks must be authorised and audited.

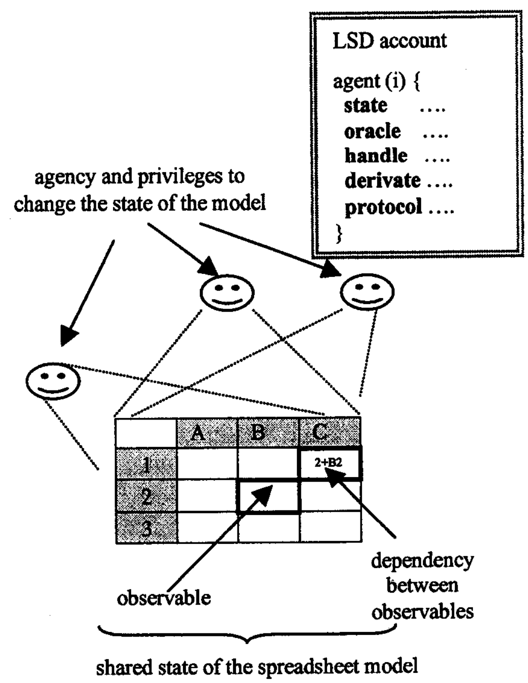

Figure 3. The spreadsheet as an ISM

A more illuminating perspective on the spreadsheet as an ISM examines the potential use to which it might be put, discounting the legacy of administrative practice established in relation to previous technologies. It should also reflect the powerful interaction that is supported by distributed EM tools (cf. Figure 3), rather than the indirect communication supported by such technologies as the pre-process-and-publish of spreadsheets across the web. To this end, it is appropriate to envisage that examiners are provided with spreadsheets networked in such a way that changes can be propagated from one to another.

Elements of discretion are often invoked in examination procedures, especially where there are extenuating circumstances to consider. Through interaction with a spreadsheet, it is easy to answer such questions as: *How many additional marks in a particular subject would be needed to improve the overall grade? What average mark was attained in examinations that took place after a significant event?* Decision-making is often guided by addressing questions of this nature, but the spreadsheet is not normally used interactively at an examination board as an experimental forum for evaluating performance. Individual examiners might use their spreadsheets in an experimental fashion, and present and defend their conclusions through interactive demonstration and argument. Such procedures would exploit the spreadsheet's capacity more fully, and arguably might improve the quality of decision-making in some cases. They would also raise many challenging and controversial administrative issues: how to moderate the fairness of judgements, what strategies to use for managing potential conflict between interactions, and what protocols for propagating changes of state.

The advent of the spreadsheet in principle makes it possible to modify and refine data at any stage and maintain the consistency of the current state automatically. A plausible application of this principle involves directly linking the spreadsheets for individual modules to the comprehensive spreadsheet of examination marks. The use of a spreadsheet gives unprecedented scope for adjusting module marks so that they conform to a perceived norm (e.g. to compensate for anomalous exam performance). This potential exposes significant educational issues regarding assessment: whether the results of a module should be determined entirely by preconceived marking conventions, whether to adapt marks to a norm irrespective of the quality of performance relative to other years, and whether to adjust marks only with reference to the module in isolation or with reference to the comprehensive spreadsheet to which it contributes. Other scenarios highlight the extent to which dynamic tracking of state is desirable. An adjustment of module marks should in fairness be propagated to all candidates, and so can have implications for several examination boards. If examination board spreadsheets were to be directly linked to module results, there is every possibility of disrupting examination decision procedures. For instance, if candidates' results are reviewed in sequence, the effect of moderating marks in one module could be to undermine decisions already taken.

This discussion reveals how conservative is current practice in academia in relation to potential spreadsheet use. As Nardi (1993) notes, the implications of adopting the spreadsheet paradigm can be very far-reaching where administrative structures in social networks are concerned. It is equally clear that the issues that are raised by such a paradigm shift are highly complex, and cannot easily be resolved without recourse to experiment and empirical studies. Formal rules (such as demanding that module marks are finalised before the decision procedures are initiated at an examination board) are intended to establish points of commitment in the interests of efficiency and clarity. Other practices (such as normalising module marks to achieve a standard average) make presumptions about the consistency of year-on-year experience. The role of techniques of this kind is to separate formal rule-based activities from the experience that informs them. In this way, the subjective elements in decision-making are reduced, as is the scope for unconventional and special justification in argument.

5.2 The Virtual Electrical Laboratory

The Virtual Electrical Laboratory (VEL) is a distributed EDEN model that was developed by C R Sheth and H P d'Ornellas as a joint MSc project in 1998.

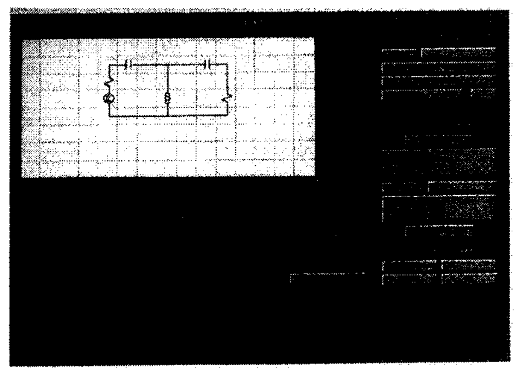

Figure 4. A snapshot of the VEL model

The VEL is primarily intended for interactive use by a teacher and a few pupils in a classroom or laboratory environment. The teacher sits at the server and the pupils at the clients of a client-server network. Figure 4 is a screenshot of the teacher's display. Many different modes of interaction between the teacher and pupils are possible, and these can be shaped interactively to suit different situations and pedagogical purposes. In one scenario, the teacher sets up a circuit that is broadcast to each client. The pupils then monitor the circuit set-up and take notes on demonstrations by the teacher of the effect of changing the values of the circuit components. In a variant of this scenario, the teacher can place the values of different components under the control of different pupils. In a third scenario, pupils build their own electrical circuit, either individually or in groups, perhaps so as to meet a function specified by the teacher. In the process of building such a circuit, pupils can experiment and record the effect of varying the values of the circuits components.

The teacher can monitor pupil work, importing example circuits from pupils for experimental use at the server if required, and selectively relaying interactions with the server in order to demonstrate instructive points.

The VEL model need not be used as a substitute for interaction with real electrical components and circuits. It has the advantage of requiring a smaller investment in time and money, and of permitting experiments beyond the scope of what physical equipment might withstand, but it can be used most instructively in conjunction with a real electrical laboratory. The main feature of the VEL is that, as an EM model, it is developed in a way that closely reflects the real-world situation to which it refers. The behaviour of an electrical circuit is construed in terms of particular observables and dependencies, and it is the representation of the current state of these observables and dependencies that is directly addressed in EM. This makes for much richer forms of interaction with the model than other paradigms support. It also makes it readily possible to adapt or extend the model to suit a different pedagogical style or environment of use.

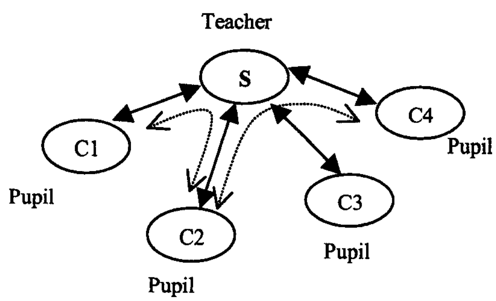

Figure 5. The star-type logical configuration for the network communication in the VEL model

Where interaction is concerned, the distributed EDEN interpreter allows transmission of re-definitions between scripts located at the server or the clients according to a star-type configuration for network communication (cf Figure 5). The semantics of such re-definition is wide open: it encompasses actions that assign different values to electrical components or create dependencies between the values of components, that introduce a new component or circuit, that simulate interface actions on behalf of the pupil, that reconfigure the interface for the pupil, that re-initialise the state of the pupil's model, or that revise the functionality of the VEL application. The VEL interface also uses the same paradigm to handle text-based interactions, for example, for transmitting messages between the VEL users and dealing with the introduction of new users. The most powerful and primitive mode of interaction with an EDEN interpreter is through entering definitions directly into an input window (not displayed in Figure 4). In principle, both teachers and pupils can exploit input of this nature. The teacher and pupil interfaces to the VEL are in effect adjuncts to the standard interface to the EDEN interpreter that serve to give GUI access to those re-definitions that are deemed to represent the most appropriate functionality for the VEL users. The teacher interface includes an additional feature that can be used to shape each pupil's environment: it is possible to specify which component values are **handles** and **oracles** in LSD terms, and accordingly can be changed and/or inspected by a pupil.

The quality of communication in distributed EDEN, and in the VEL in particular, stems from the fact that the representation of state is definitive, transitions are effected by transmitting re-definitions and each interpreter actively maintains and monitors the current state of relevant observables. All these features reflect the way in which agents are construed to act and interact in the real-world. They are also relevant to architectures for agency. Where present computing platforms are concerned, the need to distribute EDEN interpreters may be regarded as imposing a significant computational demand on each client. There are ways in which this issue could be addressed – in particular, by localising and customising dependency maintenance to balance the resources available for reconstruction of state and the bandwidth for transmission of state information.

A comparison with other programming technologies that might be applied to a VEL is helpful. Current web technology is highly document-centric. To transmit non-textual data requires techniques such as pre-process-and-publish that are far from delivering the direct influence over remote state that the transmission of a re-definition effects. The scope for interactive agency in a web network is inhibited by the standard net protocols: the state of a webpage is updated only when the viewer of the webpage initiates a request. With conventional software development methods, a Java implementation of the VEL would be targeted at a specific preconceived requirement for teacher-pupil interaction. In this connection, there is a trade-off between the narrowness of the requirement and the quality and efficiency of the solution. Similar considerations apply to the comparison between the VEL and a commercial product such as the PSPICE circuit simulator: though the VEL is clearly less sophisticated and polished as a software application, these drawbacks are for some purposes outweighed by its communication capabilities and potential functionality in an EM framework.

A brief account of how the VEL model is constructed illustrates how its functionality remains open-ended and imprecisely specified in a manner that resembles – and to some extent reflects – the information horizons of its designer-users. As in all EM models, there are many different ways in which the constituent definitive scripts, functions and actions can be organised into clusters. These can correspond to conceptual layers in the model, to submodels suitable for re-use, or to partitions into observables associated with specific subobjects for instance. The collaborative construction of the model was itself based on two complementary perspectives: d'Ornellas implemented the underlying mathematical model of electrical circuits, and Sheth developed the user interface. The mathematical model itself is framed in terms of a script of definitions relating complex matrices, and the VEL incorporates a family of data types and operators that can be used wherever such a family of linear relations can be exploited. A submodel suitable for raw circuit analysis lies beneath the user-interface. The particular functionality supported by the VEL user-interfaces masks the full potential for use and re-use of these components of the model. For instance, it is easy to introduce definitions to impose dependencies that relate the values attached to electrical components. It is also possible to link an abstract model of an electrical circuit to an EM model of a corresponding device, as d'Ornellas did for a pre-existing EM model of a set of traffic lights.

As an educational tool, the VEL model is more appropriately viewed as complementing an external activity and situation rather than meeting a prescribed functional specification. In due course, as EM principles and tools mature, it should become possible to shape such models to suit particular ways of assembling information resources for teaching purposes. The teacher and pupil interfaces to the VEL already reflect relevant ways of configuring resources within their respective information horizons. The teacher wishes to observe the structure of a particular circuit and the results obtained from mathematical analysis, to gain feedback from the pupils, and to inspect how the pupils interact with the models by setting up experimental situations that can expose their understanding of electrical circuits. They may well wish to combine these activities with the study of a relevant electrical device. The teacher may also have a deeper awareness of the underlying mathematical model that can be usefully exposed in certain situations. EM potentially makes it possible to integrate these activities within the digital environment. The power of EM principles, and the aspiration for EM tools, is to allow observables from such diverse sources to be linked up without obscuring the conceptual integrity of the model.

5.2 An EM model for online trading

The previous examples of EM have addressed virtual collaborations that are sufficiently modest in their scope and local in their impact that the wholesale adoption of a new paradigm is readily feasible. More difficult issues are raised when considering applications that presently incorporate many different work practices and pieces of traditional software, and cannot readily be re-engineered from scratch. Online security trading is a classical example of such a collaboration environment, where many information processing processes are involved, and competition takes precedence over knowledge sharing and group work. This section reviews some preliminary experiments towards developing EM models that can yield useful insights in this context. Such models serve a number of purposes. They can assist process comprehension and exposition. They can highlight critical logistic issues where computation and communication are concerned. They can also help to identify the requirements for software agents.

Advances in computer and telecommunication on electronic trading are promoting revolutionary rather than evolutionary change. The power of the PC – and the hardware in the back office – is close to overtaking markets that traditionally depended on bringing buyers and sellers together in one place. Geographical distance and layers of intermediation are becoming irrelevant (Langton, 1999). The major challenges for the global electronic marketplace are:

- offering straight through processing[1] (STP) of transactions. Technology integration is a key factor for the success of the STP initiative;
- integrating electronic trading systems in different markets;
- reducing the cost of transaction and increasing market liquidity;
- regulating cross border trading;
- providing intelligent management and delivery of trade information, reporting, and investment services.

A major problem is that the existing electronic trading networks are so diversified that it seems hardly possible to integrate them and to offer straight through processing of transactions whilst protecting the rights of the trading partners. The information processing mechanism adopted in different markets is a critical factor for an efficient electronic trading environment. This section illustrates how the development of an ISM and agent-oriented analysis using LSD can be used in conjunction to investigate such mechanisms.

In modelling an online trading environment, collaboration can be viewed as a workflow of interdependent tasks undertaken by human and electronic agents. Research in the area of workflow management systems in a business context attributes the difficulty of virtual co-operation between organizations to the lack of standard ways of representing an application's structure and sending and receiving work items (Shrivastava et al, 1999). Although the extended markup language (XML) and other similar standards for data exchange, such as Open Financial Exchange (OFX) and Open Trading Protocol (OTP), are significant advances in this area, full co-operation between organizations is still a long way off. Present day workflow systems are not scalable, as their structure tends to be monolithic and they offer little support for building fault-tolerant applications (Shrivastava et al 1999). Software development in the area of workflow management systems is directed towards developing concepts, methodologies, techniques, and tools to support workflow-process management (Sheth et al, 1999). The main challenge facing the networked economy is to design workflow processes that cross organizational boundaries. This is especially difficult

[1] STP refers to the fully automated, hands-free processing of security transactions from the fund manager's decision right through to settlement, reconciliation and reporting. It is also defined as the vertical integration of services from trading to settlement reconciliation (Reichardt, 1999)

when these boundaries are fluid and subject to continuous change.

A simple case study will be used to illustrate how the principles of agent-oriented analysis in EM can be applied to workflow in an online trading context. This study addresses the specific situation of a retail trade in the New York stock exchange (Harris 1998). The concepts introduced in section 4.2 above will be elaborated with reference to the embryonic ISM in Figure 7, and the fragment of an associated LSD account below.

In the online trading context, the social network comprises investors, brokers, dealers, arbitrageurs, and boards of trade. The trading marketplace may be a physical trading floor or an electronic system. In the retail trade situation, the relevant agents in the model are identified as: the *investor*, the *broker*, the *quote information system*, the *order entry system*, the *order routing system*, the *floor specialist*, and the *information reporting system*. Figure 7 is a screen snapshot of an ISM built using the EDEN interpreter. This display serves to animate the workflow associated with a retail trade.

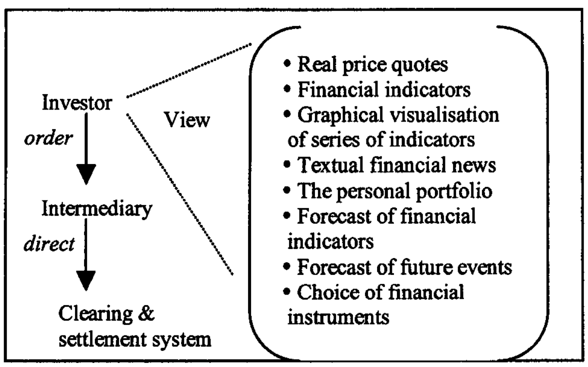

Figure 6. A typical information horizon of an investor

The roles of the various agents in the NYSE have to be understood in terms of the relevant observables. Some of these observables (such as the current status of a BUY/SELL order) are particular to the retail trade situation, but the actions of agents also relate to observables generic to the online trading context. The information horizon for the investor includes many of the most significant of these observables (cf. Figure 6).

Today's investor is looking beyond receiving delayed financial indicators. Trading in a sufficiently liquid and cost efficient market is becoming a major concern for investors (Langton, 1999), and this motivates a better understanding of the trading environment and the layers of intermediation. The support of a large range of instruments, the quality and timeliness of information feed, the functionality of the front-end, and the scalability and performance of the system are important factors in designing digital information resources for an investor.

Technology is opening up new avenues for investors to cut out the layers of intermediation and talk to one another directly. This places a question mark over what value can be added to the trading process by the stock exchanges and their constituent brokerages (Langton, 1999). Current online trading networks provide a huge amount of static information for the investor to interpret and analyse. Online trading web sites have been created by brokerage firms with the aim of extending the digital information horizon of an investor. To this end, these web sites are currently delivering free access to delayed prices, portfolio management services, and graphic visualisation of financial indicators.

In the retail trade situation, the relevant observables for the participating agents comprise:

- Order information, including: investor name, ID, BUY/SELL order, share name and symbol, quantity of shares, type of order (such as market, stop loss, limit order, etc.), price (if needed), expiry date of the order, the date and time of the order.
- Stock quotes, including: stock symbol, bidder, BID/ASK, price, size, time and date.
- Stock information, including: stock symbol, stock name, last trade price, change from previous day close, time last traded, place last traded, highest day price, lowest day price, day volume.
- Order indication from dealers and brokers, including: the stock name, the name of the broker/dealer, the time, and the date.

The above information involves live data that is continuously changing. Creating a realistic computer representation is challenging for a variety of reasons: the relevant data is large in volume, the data requires a live feed from an external information source, and the significance and status of the data changes with time. This is problematic for the EDEN interpreter, which is not well suited to data-intensive applications and does not deal gracefully with the expiry of data. Commercial databases and single-user spreadsheets offer partial solutions to these problems, but do not deal effectively with the dynamic behaviour of the trading environment, especially when data is captured via the stateless web front end. To build an ISM for the purposes of demonstration, it suffices to encode a small amount of representative data in EDEN, and it is on this basis that the ISM depicted in Figure 7 has been devised.

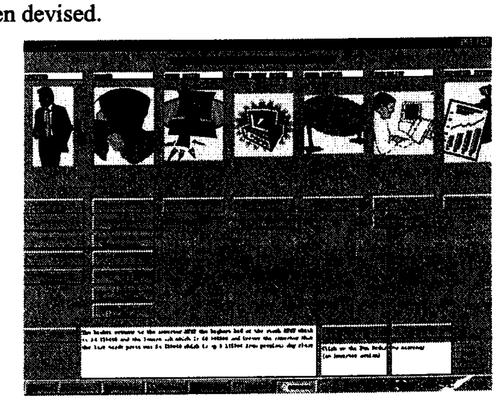

Figure 7. A snapshot from the ISM for a retail trade in NYSE

Constructing an ISM for an NYSE retail trade is a way of modelling an external observer's explanation of the retail trade process (RTP). In its most naïve form, such an explanation explicitly relates the actions of agents to the stages of the trading protocol. This simply involves identifying the actions for which each agent is responsible, and identifying the preconditions under which each action is performed. As Figure 7 illustrates, the major roles in a retail trade are played by the investor and the broker. The *investor* requests information on a particular stock from the broker, puts a trading order, confirms his order, pays for his transaction, and acquires or releases share ownership following the execution of his order. The *broker* requests quotes from the quote information system, returns this information to the investor, enters any received order in the order entry system, reviews the order details prior to its release in the order entry system, reports the trade execution to the investor, receives payment including charges fees, and mediates the exchange of share ownership. Each of these actions on the part of investor and broker is performed at a specific stage in the RTP.

To formulate this simple explanation in LSD, it suffices to interpret the current stage reached in the RTP as an observable for the participating agents, and to formulate each agent action in terms of re-definitions of observables. For instance, in the initial stages of the RTP, the broker requests quotes from the quote information system when an investor has requested information on a particular stock. The LSD account of the broker is then as follows:

```
agent broker {
    state       info_requested, quotes_info_requested, ...
    oracle      stage_in_retail_trade
                info_requested
                ....
    handle      quotes_info_requested=0
                ....
    derivate    stage_in_retail_trade = F(info_requested, ...)
                ....
    protocol    (stage_in_retail_trade = init_trade)
                and (info_requested) → quotes_info_requested=1
                ....
}
```

The animation in Figure 7 can be derived from such an LSD account. The explanation given by the external observer is very superficial in this case: it does not take the actual character of the transactions and interactions into account, but merely registers the pattern of the workflow. In practice, the possible scenarios that can arise in the RTP are much more subtle than the workflow alone indicates. Transactions may be disrupted by communication failure, by human error, or by the dishonest dealing. The RTP also takes place in a setting where other kinds of observation pertain. There will be a stage at which the investor is legally commited to complete, for instance. These broad issues regarding the RTP will have to be reflected in devising a useful ISM. The potential subtlety of the RTP is mirrored in the possible interpretations that can be given to the LSD account above, and the elaborations that these motivate. The LSD account refers to info_requested as both a **state** and an **oracle**. This highlights a potential ambiguity concerning a particular information request. As an **oracle**, info_requested refers to an observable that is associated with an investor. This can be interpreted as saying that the broker is – or at any rate can be – aware that an investor is requesting information. For the purpose of giving a routine account of the workflow, how such a request is mediated to the broker is irrelevant, and the possibility that the broker may be too preoccupied to note the request is discounted. As a **state** for the broker, info_requested refers to an observable whose status is private to the broker. Recording info_requested as a **state** potentially admits discrepancies between what the broker believes or recalls and what the investor has declared. The consequences of such discrepancies are implicit in the interpretation of the broker's protocol. The precondition for action on the part of the broker can be read as: the broker *believes* that a particular stage in the RTP has been reached and that information has been requested.

Similar considerations apply to the **derivate** that determines the stage reached in the RTP. The definition

stage_in_retail_trade = F(info_requested, ...)

is used to indicate that the current stage in the RTP can be construed as functionally dependent on the status of transactions. In a naïve account of the RTP, this can be seen as reflecting the fact that, once the investor has requested information from the broker, a new stage of the RTP has been entered. Introducing such dependencies in the ISM for the RTP gives the assignment of a new value to the observable info_requested the quality of a re-definition – an action that potentially has indivisible effects on the state of other observables. Such a mechanism could also be used to take account of whether an action had some legal consequences, such as might express a commitment or obligation. From this perspective, it might also be appropriate to deem stage_in_retail_trade as also dependent upon the precise contents of a transaction: if shares were paid for using counterfeit money for instance.

One motivation for embellishing the LSD account and the ISM for the RTP is that many different communication technologies and information strategies can be used in the RTP. As an observable, an information request placed by telephone has quite different characteristics from a web request. What observables a broker uses to determine the current stage of the RTP may be hard to ascertain. A precise procedural account of how a broker processes a request from an investor might not resemble a re-definition, and could quite easily involve creating and then resolving inconsistent states. A possible account of the broker's response to an information request might be:

1. check status of the investor's information request;
2. get the investor's information request;
3. direct the request to the quote information system;
4. update the current RTP status.

By implication, this is not a simple re-definition, but a sequence of related assignments. Conceptually, it is much harder to guarantee the integrity of the state which it creates. It might on the other hand be necessary to formulate the broker's role in this much detail to capture the true situation more faithfully. For instance, if updating the current RTP status involves some explicit book-keeping on the part of the broker, it is possible for this action to be accidentally omitted.

The above discussion indicates the kind of analysis that accompanies the development of an ISM and LSD account for the RTP. The precise scope and nature of the ISM is open: it could be oriented towards a high-level account of workflow, or to a specific framework for implementation. The modelling process may be helpful in addressing the integration of human and automatic activities, and could be directed specifically towards related goals. It might also indicate that, in some situations, effective integration is infeasible with current technologies and paradigms, where computation is too far abstracted, and the potential agency of automated components is too limited.

6. CONCLUSION

This paper describes and illustrates how an Empirical Modelling approach can be applied to environments for several different forms of virtual collaboration. The principal features of EM are:

- it exploits communication that is centred on artefacts rather than documents;
- it supports the exploratory activity needed to identify and specify requirements for agent interaction;
- it supplies a framework in which model-building is intimately linked to learning and shared understanding;
- it provides a conceptual framework in which to examine issues of feasibility and human-computer integration;
- it allows concurrent examination of technical and social perspectives.

Current tools for EM are sufficient for proof-of-concept, but will require further development before they can be effectively used for applications that are data-intensive and require large-scale collaboration. The research reported in this paper indicates that, in principle, the EM approach can deliver more powerful distributed environments for collaboration than alternative technologies. Web-enabled EM tools have great potential interest in this respect, but web protocols are currently too restrictive to give the support for interaction and agency that is ideally required.

7. ACKNOWLEDGEMENTS

We are indebted to P-H Sun, C Sheth and H d'Ornellas for major technical contributions, and to D Sonnenwald and members of the EM group for many relevant ideas. The first author is supported by EPSRC grant M02675.

REFERENCES

Adzhiev, V.D., Beynon, W.M., Cartwright, A.J., Yung, Y.P., "An Agent-oriented Framework for Concurrent Engineering", *Proceedings of* IEE Colloquium: Issues of Cooperative Working in Concurrent Engineering, October 1994, Digest 1994/177, 9/1-9/4.

Beynon, W.M., Adzhiev, V.D., Cartwright, A.J., Yung, Y.P., "A Computational Model for Multiagent Interaction in Concurrent Engineering", *Proceedings of* CEEDA'94, Bournemouth Univ., 1994, 227-232.

Beynon, W.M., Maad, S., "Empirical Modelling of Real Life Financial Systems: The need for Integration of Enabling Tools and Technologies", to appear in *Proceedings of IDPT 1999.*

Beynon, W.M., Rungrattanaubol, J., Sun, P.H., Wright, A.E.M., "Explanatory Models for Open-ended Human-Computer Interaction", CSRR#346, Warwick Univ., 1998

Ciancarini, P., Rossi, D., Vitali, F., "A case study in designing a document-centric coordination application over the internet", *Proceedings of The Active Web Conference*, January 99, Staffordshire University, UK.

Gooding, D., "Experiment and the making of meaning", Kluwer Academic Publishers, 1990.

Harasim, L., "A Framework For Online Learning: The Virtual-U", IEEE Computer Society, Computer, September 1999, pp. 45-49

Harris, L.E., "Trading and Exchanges", Draft copy of the text book, December 4, 1998, University of Southern California, Marshall School of Business.

Langton L. J., "The impact of the electronic marketplace: The transformation of securities trading", *Proceedings of* On-Line Securities Trading 1999, AIC Worldwide Ltd, September 1999, London.

Marshak, R. T., "Open Text Livelink Intranet: Expanding Document Management to Support Collaborative Team Projects", Patricia Seybold Group's Workgroup Computing Report Vol. 21, No. 7, 1998.

Nardi, B.A., "A small matter of programming: perspectives on end-user computing", MIT Press, 1993.

Reichardt P., "Clearing and settlement online: implications for securities setllement arising from straight through processing", *Proceedings of* On-Line Securities Trading 1999, AIC Worldwide, September 1999, London.

Sheth, A.P., Aalst, W., Arpinar, I.B., "Processes Driving the Networked Economy", IEEE Concurrency Journal, July-September 1999, pp.18-31.

Shrivastava, K.S., Wheater, M.S., "Workflow Management Systems", IEEE Concurrency Journal, July-September 1999, pp.16-17.

Sonnenwald, D.H., "Evolving perspectives of human information behavior: contexts, situations, social networks and information horizons", *Exploring the Contexts of Information Behaviour*, 1999, Taylor Graham, 176-190

Sun, P.H., Beynon, W.M., "Empirical Modelling: A New Approach to Understanding Requirements", *Proceedings of ICSEA Conference*, Vol. 3, 1998, Paris.

Taylor, C., Walter, D., "E-commerce directions", Internet Business Magazine, May 99, Issue 28, pp 36.