From Rooms to Cyberspace: Models of Interaction in Large Virtual Computer Spaces

Steve Benford, Adrian Bullock, Neil Cook, Paul Harvey, Rob Ingram and Ok-Ki Lee
Communications Research Group
The University of Nottingham

University Park, Nottingham, UK

Abstract

Room metaphors have become increasingly popular as a basis for CSCW systems. This paper describes how such metaphors might be extended to support large scale communication through the introduction of a spatial model for meditating conversations in virtual computer spaces. The model is described in terms of an abstract mathematical framework and the paper then outlines how this might be applied to various kinds of CSCW system. As a next step, the combination of rooms into larger virtual structures is considered resulting in proposals for structuring, navigating and mapping a large virtual cyberspace for cooperation. Finally, the paper describes a current prototype application and reflects on some of the architectural issues involved in its realisation as a distributed system.[1]

1. Introduction

The primary goal of this paper is to present a model for supporting group interaction in large scale collaborative systems. This model aims to provide generic techniques for mediating communication in large groups and is intended to be applicable across a diverse range of CSCW systems from text conferencing, through multi-media to virtual reality. It is also intended to contrast with existing floor control and workflow modelling techniques by adopting a “spatial” approach which we believe is particularly applicable to systems based on the increasing popular metaphor of virtual rooms.

First, we would like to outline our general objectives in undertaking this work. Our major focus has been to consider mediation of conversations in collaborative systems that can scale to large numbers of people working over extended time periods. By large numbers we mean hundreds or maybe even thousands of people and by extended periods we mean weeks, months or even years. Further goals include a more seamless integration of synchronous and asynchronous modes of interaction than is evident in current systems as well as support for social and ad-hoc communication to complement more formal and planned work processes.

It is also necessary to say a few words about our underlying philosophy. Where possible, we have tried to encourage individual autonomy of action, freedom to communicate and minimal hard-wired computer constraints. In particular, we wish to avoid at all costs the kind of nightmare scenario where the computer actually disallows you from speaking (effectively what happens with some floor control policies) and replace it with a model of increasing constraints which may be

1. This paper will appear in Interacting With Computers (Butterworth-Heinmann) in 1993.

broken through increasing effort.

We have chosen to base our work around the metaphor of interaction within virtual computer spaces or “rooms” as they have come to be known. Variants of the rooms metaphor have already been adopted as the basis for a number of recent projects and we believe that the highly intuitive nature of the metaphor makes it a very powerful basis for the design of CSCW systems. The recent growth of public interest in Virtual Reality and Cyberspaces lends yet more weight to this argument.

In section two we briefly introduce of definition of a room as a virtual computer space. Section three then introduces the spatial model for mediating conversation within rooms. Key concepts are described and then defined using a mathematical framework. We also discuss how the model could be applied to a range of different systems. In section four, we consider a second aspect of scaling, namely connecting virtual rooms into larger structures. Here we identify basic concepts and also describe some recent work on mapping and navigating in virtual worlds. Finally, in section five we describe our current implementation, *CyCo* (Cyberspace for Cooperation), which aims to demonstrate the technical feasibility of our models. This section also describes some of the architectural work carried out, including a technique for federating individual “room servers” maintaining local virtual worlds into a large distributed system managing a global cyberspace.

2. Rooms as virtual spaces for interaction

A room is a virtual computer space within which people interact, either with each other or with the various tools which they find there. Put in a more abstract way, each room defines a particular domain of interaction. Rooms may represent task specific spaces (e.g. a room created for a specific project or meeting), more permanent organizational spaces (e.g. offices or seminar rooms) and also public spaces (e.g. foyers and corridors).

The use of virtual rooms as a metaphor for structuring computer systems is not particularly new. The Xerox Rooms system used the metaphor to structure user interfaces [Henderson 85, Clarkson 91] and this was later followed up with VRooms [Borning 91]; Cruiser explored the role of virtual rooms in supporting “social browsing” [Root 88]; several teleconferencing systems have utilised a metaphor of virtual meeting rooms (e.g. [Cook 91, Leevers 92]); the Audio Windows system was also influenced by the rooms metaphor [Cohen 91] and multi-user recreational environments have been available for some time (e.g. MUD [Smith 92] and Lucasfilm’s Habitat [Morningstar 91]. Not surprisingly, virtual rooms also feature heavily in discussions of virtual reality [Benedikt 91] and early multi-user VR systems are beginning to emerge [Codella 92, Takemura 92, Fahlén 92]. The rooms metaphor has also been considered as a basis for a large scale CSCW environment by at least two other recent projects [Michelitsch 91, Benford 92, Navarro 92]. Finally, spatial approaches to CSCW have also been explored in work on media-spaces [Gaver 92a, Gaver 92b]. Undoubtedly the main reason that the virtual room metaphor has proved to be so popular is its strong relation to physical reality and therefore its highly intuitive nature. Certainly our experience is that people easily grasp the concept of virtual rooms and very quickly enter into discussions about the possibilities that the metaphor opens up.

Virtual rooms play several key roles in our work. They provide an ether, a continually available communication channel, allowing occupants to interact freely at all times. This reflects a belief in the importance of supporting social interaction within groups (see [Root 88] and [Gale 91] for further discussion of this point). Rooms also provide awareness of the presence and comings and goings of all occupants and tools. This support for awareness in virtual environments is seen as a direct parallel to that already discussed for physical office environments [Lovstrand 91]. Finally, they provide access to tools. The main advantage of locating tools in rooms is that all tools become collaborative to some degree through the ether that surrounds them. For example, I might chat to another person while standing in front of a notice board. Thus, rooms also encourage a blurring of the distinctions between asynchronous and synchronous and formal and informal communication.

Our paper is particularly concerned with collaboration on a large scale. The following section begins discussion of this issue by considering how we can effectively scale a single virtual room to support many occupants.

3. Big rooms and the spatial metaphor

What happens as the number of occupants in a room increases beyond a small number? In a crowded room you are unlikely to cope with knowing every utterance of every occupant. Furthermore, the computational overhead involved in transferring utterances will soon become intolerably high. We refer to these problems as cognitive overload and computational overload respectively. The solution to both might be provided by some kind of medition mechanism. One approach is to introduce a "floor control" policy such as the use of communication roles (e.g. having speakers and chair-people with differing abilities to control the flow of communication), token-passing mechanisms (e.g. a "megaphone and gun" approach) or reservation systems [Crowley 90, Sarin 91, Cook 92]. Another approach has been evident in recent asynchronous CSCW projects which have proposed notations to model the workflow/process/activity within a group (e.g. the Coordinator [Winograd 86], Domino [Victor 91], Chaos [Bignoli 91], Cosmos [Bowers 88] and Amigo [Pankoke 89]).

Although these explicit floor control and coordination techniques may be useful in some circumstances, we believe that they are generally inappropriate for mediating the subtlety of human interaction in virtual rooms. More specifically, they suffer from trying to make explicit in the system what is essentially social convention and in so doing tend to be overly constraining. Put another way, a real-world implementation of explicit floor control conventions would be tantamount to gagging everyone at a meeting and then allowing them to speak by removing the gags at specific times. Our aim has therefore been to identify alternative mechanisms which reduce information overload while at the same time meeting the following goals:-

- ensuring individual autonomy of action as much as possible;
- maintaining a power balance between "speakers" and "listeners" (many communication systems, electronic mail for example, are biased towards speakers).
- starting with support for social mingling and adding more formal mechanisms later on if needed.

A further goal of this work has been to try to describe these mechanisms in terms of a more formal framework thereby allowing their application to a variety of CSCW systems of different complexity ranging from text-based conferencing, through multi-media conferencing to virtual reality. The result of this work has been a so-called "spatial model" of interaction.

3.1. Key concepts in the spatial model

The spatial model, as its name suggests, uses spatial characteristics of rooms as the basis for mediating interaction. In the model, a room becomes a space and each person therefore has properties of position and orientation within the room. People can move around the room relative to each other and can also turn to face each other.

Next, we identify two key concepts which allow people to govern their interaction.

FOCUS - represents a sub-space within which a person focuses their attention. The idea is that a person is more aware of objects inside their focus and less aware of objects outside of their focus.

NIMBUS - represents a sub-space within which a person projects their presence. The idea is that an object within your nimbus is more aware of you than an object outside of your nimbus. Another way of thinking of nimbus is as a sub-space across which a person makes their activity available to others.

Together, focus and nimbus contribute to your **AWARENESS** of other people. It is this awareness level that dictates the information that you perceive. Focus and nimbus are manipulated implicitly by moving around the room and through orientation. They may also be manipulated explicitly (e.g. by focusing in on a particular object). The use of focus and nimbus in this way imply the following:

- Each person autonomously controls their awareness.
- Levels of awareness between two people need not be symmetrical.
- Just because you are speaking to me doesn't mean that I have to listen to you.
- I can't stop you speaking, but I can stop listening to you.

In order to support scenarios beyond mingling, we introduce various tools into rooms to manipulate focus and nimbus. We call these focus and nimbus adaptors. For example, a microphone tool might enhance my nimbus thus allowing me to project over many people and a table tool might fold together the foci of several people into a semi-permeable sub-space within a room. Room shape might play a similar role, for example, an alcove might act like the table tool.

Focus, nimbus, awareness and adapters are then the basic concepts that allow people to mediate conversation in large rooms. The following section outlines a more abstract definition of these concepts.

3.2. A Formal Model of Awareness, Focus and Nimbus

The concepts of focus, nimbus and awareness are intended to be useful across extremely diverse application domains. More specifically, they might be applied wherever a spatial metric can be identified (i.e. a way of defining and measuring space). In order to realize this generality, we introduce a more abstract framework which defines the spatial model in a consistent, application independent manner. This requires the introduction of one further concept, that of a communication medium. We assume that any room-based implementation will utilise one or more media for communication (e.g. sound, light or text). A model is required which will answer the following question:

"In a given room how aware is entity i of entity j via medium k?"

The Model

It is useful to represent the factors which influence awareness as mathematical functions. This affords us both the flexibility of definition we require as well as a method to manipulate awareness in the form of a numerical quantity.

Consider an arbitrary room in which communication takes place between n entities via m

media. The concepts of focus and nimbus are the most fundamental to the model. All entities define these quantities for each of the *m* media. We denote functions to represent them for entity *i* in medium *k* as f_{ik} and n_{ik}. Note that an entity may employ different focus and nimbus functions for different media. These functions evaluate to a measure of the focus (or nimbus) of a given entity on (at) another depending on certain *properties* of both of the entities concerned. The form of the functions and the properties used is a matter for a particular application of the model and may be chosen according to the media involved and the intended sophistication of the end system. Typical examples might be position and orientation. However, other properties could be of interest (e.g. focal lengths or even degree of activity within the system). Thus

$$f_{ik}: T_k \mathrm{x}\, T_k \to R$$
$$n_{ik}: T_k \mathrm{x}\, T_k \to R$$

Where f_{ik} is the focus function and n_{ik} is the nimbus function both of which map property values into the set of real numbers *R*. These properties are drawn from a set T_k for an entity *i* by the function t_{ik}. A measure of the focus of entity *i* on entity *j* in medium *k* is therefore given by:

$$f_{ik}(t_{ik}, t_{jk})$$

The form of the functions and the properties used is a matter for a particular application of the model and may be chosen according to the media involved and the intended sophistication of the end system. The focus and nimbus functions are used to describe how an entity interacts in a particular medium, hence they are grouped to form a 2-tuple known as a medium description. We define a medium description D_{ik} for entity *i* in media *k* as follows:

$$D_{ik} = (f_{ik}, n_{ik})$$

This represents how a particular entity influences its interaction in a given medium. We can extend this further to define an entity description, E_i as the combination of media descriptions across all media in a room.

$$E_i = (D_{i1}, D_{i2}, \ldots, D_{im})$$

Now we consider how to represent awareness in a given medium. Awareness in medium *k* is itself a function that combines focus and nimbus functions.

$$A_{kij}(f_{ik}, n_{jk}) : R^2 \to R$$

This function evaluates to a measure of awareness of a given entity *i* to another *j* based on values of the focus of entity *i* (f_{ik}) on *j* and the nimbus of entity *j* (n_{jk}) at *i*. We are now equipped to answer the question:

"In a given room how aware is entity i of entity j via medium k?"

$$\text{Level of Awareness} = A_{kij}(f_{ik}(t_{ik}, t_{jk}), n_{jk}(t_{jk}, t_{ik}))$$

It is worth emphasising that the awareness that entity *i* has of *j* is a combination of *i*'s focus on *j* and *j*'s nimbus at *i*.

Considering the tools which serve to adapt focus and nimbus in a room. A tool is itself an entity and so has an entity description specifying a set of focus and nimbus functions for different media. Upon selecting a tool a person acquires the awareness properties of the tool instead of their own. In other words, they replace their entity description with that of the tool they are using.

Applying the model

Our model allows the awareness that one person has of another to be calculated as a numeric quantity. This measure of awareness is then interpreted in some concrete way within an end system. The process of interpretation lies outside the domain of responsibility of the model and will typically be part of the functionality of a *User Agent* (UA). In other words, a user agent presents an interpretation of awareness which assigns a real world meaning to the numerical measure. For example, in the sound medium a particular UA may interpret the level of awareness as a measure of volume whereas in a text medium it may use it decide whether to display a given message or not. To apply our model to the design of an end system, several steps therefore need to be followed.

- 1. Identify the media involved.
- 2. Identify the elements of the property set for each medium.
- 3. Specify focus and nimbus functions for entities representing people in each medium.
- 4. Specify awareness functions for each medium.
- 5. Specify additional focus and nimbus functions for any tools for each medium.
- 6. Decide how awareness values are to be interpreted by the system for each medium.

In the following section we follow these steps with a simple example, applying the model to the design of a text conferencing system.

3.3. First example - applying the model to a text conferencing system

Now we consider how the spatial model might be applied to a text conferencing system (this is the example we are currently developing in the CyCo project). We will restrict rooms to be two dimensional spaces. It is important to remember throughout that this represents only one of possibly many applications of the spatial model.

Considering step 1, we have one medium, that of text messages.

For step 2, our first attempt at describing focus and nimbus functions might be to consider both of them as circular areas around a person such that for all points in this area the focus/nimbus value is 1 and for all points outside it is 0. In other words, if you are within a certain distance of me I will be focusing on you and if you are further way I will not. We can make a parallel definition for nimbus.

This is shown in figure 1 which also outlines the mathematical definition of these circular functions

Under step 3, we can define an awareness function which simply adds together the relevant focus and nimbus values. More specifically, the awareness that entity i has of entity j is calculated as the sum of i's focus *as measured at the position of j* and j's nimbus *as measured at the position of* i. Thus, $A = f_{ik} + n_{jk}$. The result of this function will always be one of three possible values for i's awareness of j.

- 0 - when i is not within j’s nimbus and j is not within i’s focus;
- 1 - when i is within j’s nimbus and j is not within i’s focus or i is not within j’s nimbus and j is within i’s focus;
- 2 - when i is within j’s nimbus and j is within i’s focus.

Now we consider the issue of possible interpretation of these values. An awareness level of 0 might imply *peripheral awareness* - you are aware that the other person is present in the room and you know where they are standing. An awareness level of 1 might imply *semi-awareness* meaning that you are also notified when ever they communicate. However, you don’t hear what they say. Finally, an awareness level of 2 might imply that you actually hear what they say (i.e. receive their messages).

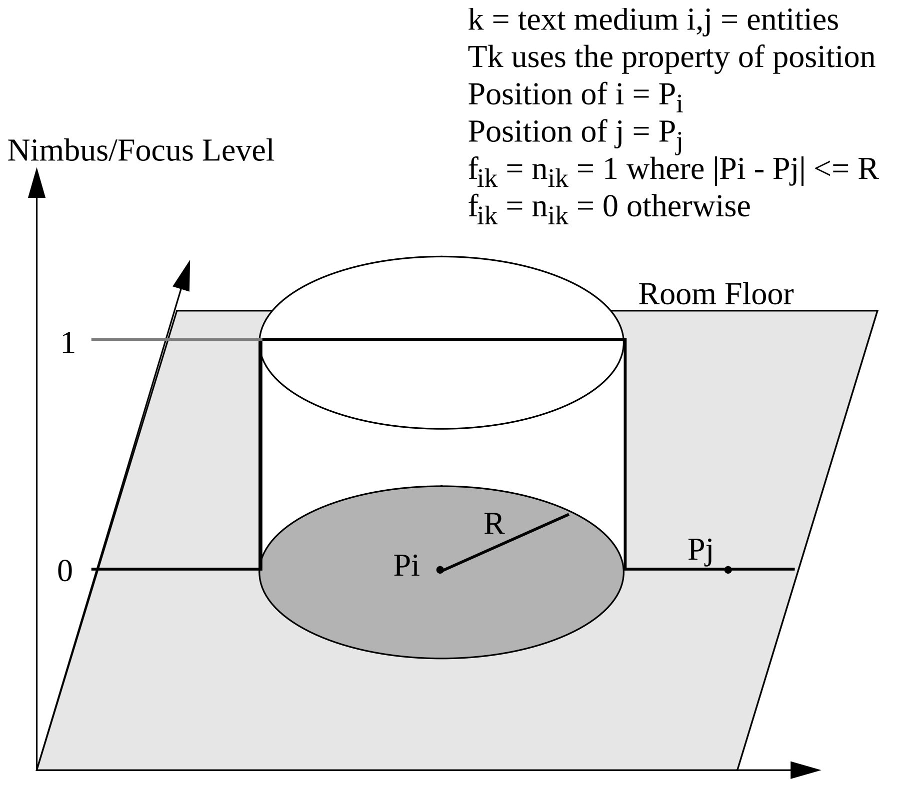

Figure 1: A simple representation of focus and nimbus

In terms of implementation, this might be represented by three windows at an interface: a map window showing the positions of everyone in the room; a notify window informing you of communication acts taking place and a speech window where you receive messages. The implementation would then let you move around the room in order to get close to people so that you could speak to them and hear them.

To finish this particular mapping, we briefly consider some other functions that we could have chosen for focus and nimbus which either represent more sophisticated mechanisms (refinement of step 2) or which might represent specific adapters (step 4).

First, our circular function doesn’t take account of the orientation of either party. A better function might be one in which both focus and nimbus are represented as segments of a circle. In this case, the orientation of the segments would be significant and the radius of the segments could be altered (e.g. to represent focusing in or out) provided that the area remained the same. Figure 2 shows the function used to model this second scenario. The advantage of this refined function is that both parties can use their orientation to influence a conversation.

Next we consider support for a more formal scenario, say a lecture or presentation where a speaker wishes to address a large crowd. We can model this by replacing the speaker's current nimbus function with one which spans the entire room. In the simplest case this is a function which delivers the value 1 everywhere (see figure 3). This might be triggered by the speaker picking up a tool such as a virtual microphone. Notice that people can still ignore what the speaker is saying although they will always at least be notified that they are speaking.

k= text medium, i,j = entities
Tk= uses the properties position and orientation
Position of i = Pi
Position of j = Pj
$f_{ik} = n_{ik} = 1$ where |Pi -Pj| < R and ^(Pi-Pj) < α
$f_{ik} = n_{ik} = 0$ otherwise

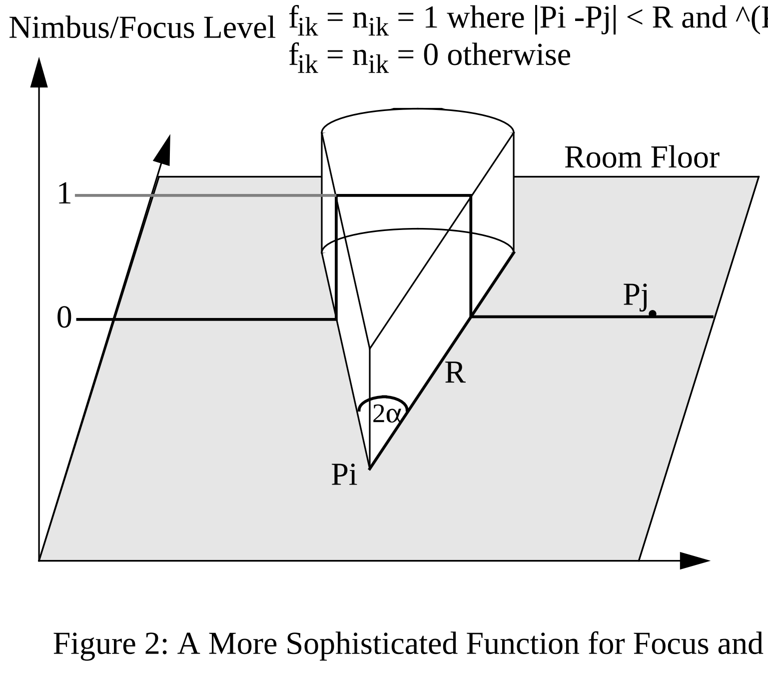

Figure 2: A More Sophisticated Function for Focus and Nimbus

Finally, consider the case where a group of people wish to form a sub-group within the room in order to have a (semi)private discussion. We assume that they don't want an entirely isolated discussion which could be achieved by moving to a separate private room, but instead wish to maintain some degree of awareness of what is happening around them. We can realise this scenario through a table tool (see figure 4). When a person sits at the table their focus and nimbus are replaced by a common focus/nimbus function relative to the centre of the table. In effect, the focus and nimbus of all people around the table is folded into a common area. The implications of this are that the details of the conversation can only be perceived by those sitting at the table but that others are aware that there is a conversation taking place and that the sub-group are equally aware of people elsewhere in the room. In summary, we have seen how the formal model of awareness, focus and nimbus might be applied to a text conferencing system by following each of the five steps mentioned above. Even at this relatively simple level, it should be possible to see the difference between the spatial metaphor and more explicit forms of floor control. The next few paragraphs very briefly discuss how the model might be applied to a more sophisticated system.

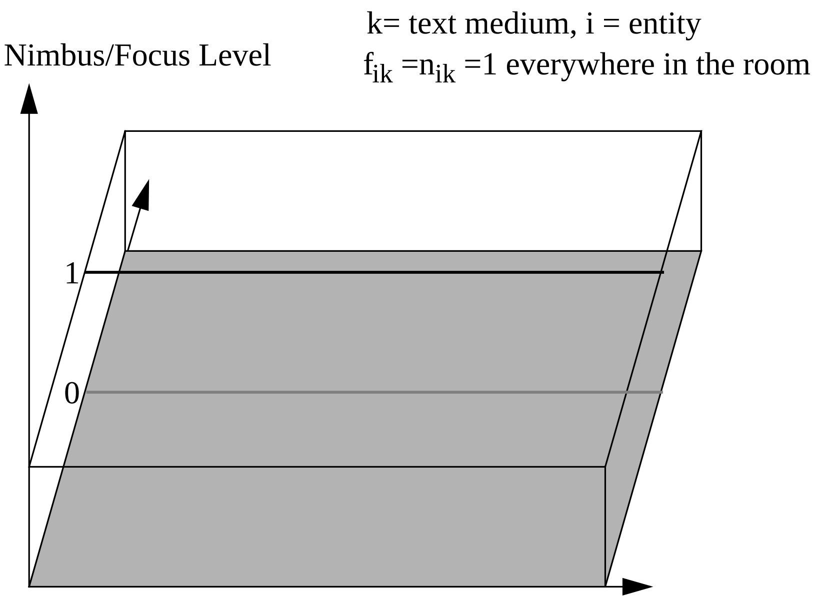

Figure 3: A Nimbus Function Representing a Virtual Microphone

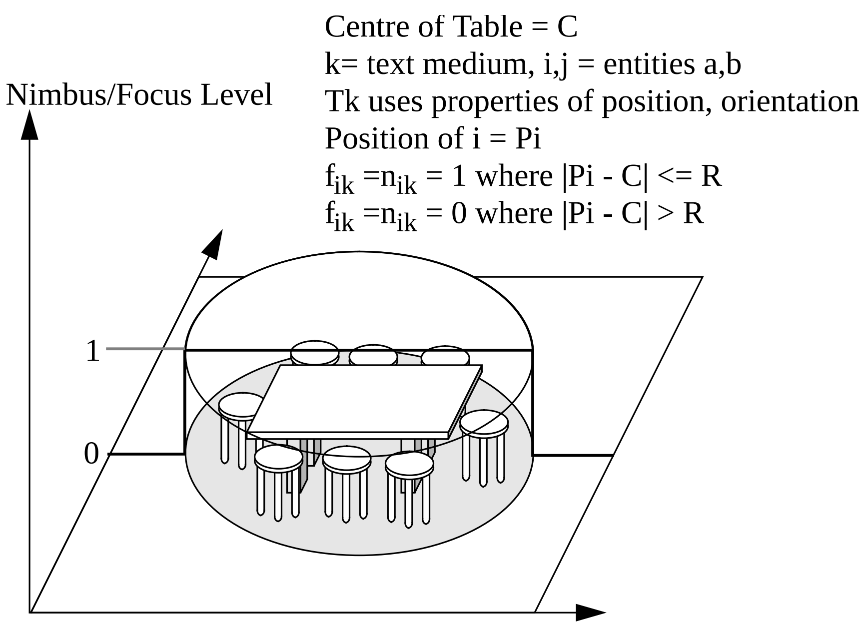

Figure 4: Functions Representing a Virtual Table

3.4. Second example - Applying the model to a virtual world with audio capability

Imagine now that you are constructing a more sophisticated virtual world with audio capabilities. Let us still consider two dimensional rooms, but in this case choose functions which deliver more continuous values of focus and nimbus in contrast to the discrete functions chosen above. A good function might have some kind of local maximum representing the peak of focus and nimbus, with its value asymptotically decreasing towards the edge of the room. Figure 5 shows the focus of one person and the nimbus of a second along one dimension of the room using such a function. As before, we can calculate the awareness that i has of j by summing the value of i's focus at j and j's nimbus at i to get some real number. However, in this case we have a continuous range of possible values as opposed to just three levels of awareness. These values might be interpreted in several ways. For example, they might be used to control the value of an audio channel such that as people moved towards you they would gradually become louder as you turned away they would become quieter. In terms of choosing specific functions, it should be possible to find functions with a range of interesting properties. For example, quadratic functions with one local maximum (i.e. one point of focus), cubic functions with two, quartic functions with three and so on.

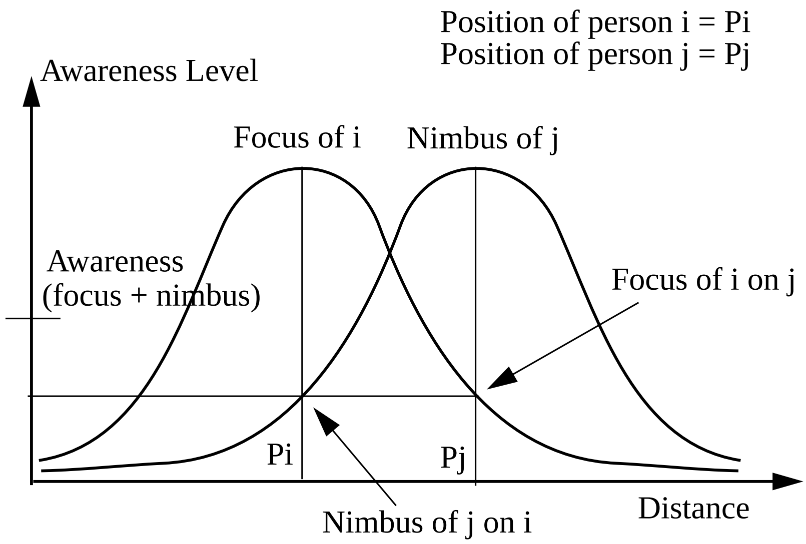

Figure 5: Continuous Functions for Focus and Nimbus

3.5 Other applications

So far we have discussed the application of the spatial model to virtual worlds based on text or audio media. There are many other possible applications of the model. The concepts of focus and nimbus naturally apply to virtual reality systems (related concepts have already been demonstrated in the DIVE system developed at the Swedish Institute of Computer Science as part of the Multi-G programme [Fahlén 92, Fahlén 93]). Furthermore, existing VR implementations often model a bounding space around an object which defines its sphere of interaction for detecting impacts and collisions.

To go a stage further, the virtual world and spatial models can be applied to any kind of system where a spatial metric can be identified. For example, we might measure the distance between people in shared hypermedia in terms of number of links between them. The spatial model could then allow people to become more aware of each other as they moved closer within a hypermedia document (e.g. automatically opening up conferencing facilities if they came within a certain distance). The same approach might be applied to the joint editing of more conventional documents or even databases, perhaps resulting in large collaborative information spaces, or *Populated Information Terrains*.

Finally, note that the model need not only support interaction between people, it could also support interaction between people and objects. For example, as I approach a notice board in the virtual world it might become more aware of me. As a result it might redisplay itself to give a more detailed view of its notices.

4. From virtual rooms to virtual worlds

The next level of scaling is achieved by connecting distinct rooms together into larger structures analagous to virtual buildings, organisations and even virtual cities.

We start with the assumption that there may be many rooms in the world and that all rooms are potentially connected to all others. However, there would seem to be two major problems with such a flatly structured world. First, navigation and location become increasingly difficult as the number of rooms to be considered at any one time increases. In effect, the world will become too complex for people to perceive and map once it grows beyond a certain point. Second, effective management of a virtual world will require some kind of management structure (e.g. the ownership rights required to create and destroy rooms). These observations imply the need for a structured virtual world. The remainder of this section considers two aspects of this structuring in more detail: connecting rooms to form larger organisational structures and the support required for people to effectively navigate around such a structured world.

4.1. Connecting rooms

Our model constrains rooms to be connected into certain world topologies. In other words, it may not always be possible to move directly from one room to another. There are two further motivations for constraining connectivity in addition to the ones given above.

- Facilitating security and privacy. Examples might be the constraint that an organisation could only be entered through a public foyer or reception area or that a person's office could only be entered via their secretary's.
- Promoting social and un-planned interaction through the introduction of public or socially neutral areas through which people frequently move en-route to other destinations.

This second idea was demonstrated in the Cruiser system which explored the use of virtual rooms to promote social interaction [Root 88]. We share their view that the kinds of conversation that occur in hallways, coffee rooms and lounges in the physical world will also need to be supported through the provision of appropriate space in the virtual world.

We propose two basic extensions to our rooms model to facilitate the creation of structured worlds: nested rooms and links between rooms.

Nested rooms, as the name implies, are rooms within rooms. The main effect of this nesting is to render the internal rooms invisible from the outside, thus hiding their complexity. This mechanism allows us to build rooms within buildings and buildings within organisations and so on. We propose to avoid strictly hierarchical structures by allowing each space to be nested within many others, not just within a single parent.

The second structuring device, links, may be used to connect rooms on a much more ad-hoc basis. The presence of a link between two spaces will allow direct movement from one to the other. Links may be created in an ad-hoc manner to connect rooms that may initially be far distant in the nesting structure. Links may also be created and used on an individual basis and may not necessarily be visible to all occupants of the world. The notion of links is reminiscent of a number of systems including the “worm holes” of the Alternate Reality Kit [Smith 87].

The distinction between nesting and links represents a separation between the requirements for group management and personal navigation. Nesting allows the construction and management of a commonly visible world structure whereas links allow the structure to be tailored on a personal basis.

This section has proposed the two mechanisms of nested rooms and links between rooms for structuring the virtual world. The following section turns its attention to the problem of navigating and mapping such a structure.

4.2. Mapping and navigation

Future virtual worlds will consist of very large numbers of rooms interconnected in a highly complex way. In order to navigate around this environment the user will need to know the details of nearby locations while still retaining an idea of the structure of the ‘whole’ world. To aid this navigation we see some sort of map as being invaluable. We define a map as essentially some representation of the world which allows the user to visualise the relationships between different spaces and so navigate. An extreme example is that of a three dimensional graphical fly through model. A more familiar example might be that of a two dimensional abstraction of the world structure.

One of the aims of nesting is to render invisible the internal structure of a given space. Consequently, mapping need not focus on the problems of representing complex nested structures and instead should primarily consider representing links between spaces. A useful representation might be that of a mathematical graph, with rooms as vertices and links as edges.

This representation allows us to manipulate the map using the mathematics of graph theory. Thus, we could reduce the complexity of a large world by mapping it as a spanning tree of the graph, with the root being the node currently occupied. A breadth first algorithm for constructing the spanning tree would allow all the immediate connections to be shown but would present only a single route to each node that is connected. Furthermore, the spanning tree could be limited to a certain depth to prevent the map getting too large. Although this method does lose some information it can be very useful in reducing the complexity produced by a large number of concentrated connections.

We could also reduce the complexity of the map by the use of Fisheye Views [Furnas 86, Sarkar 92] which show nodes near to a ‘focus’ in great detail but those further from the focus is gradually reduced detail. A yet more radical technique for reducing complexity might be to turn the problem around and use the map as a device for dynamically constructing the world. In this way, the world designer would be implicitly encouraged to construct worlds that could easily be mapped.

This concludes our brief discussion of structuring and mapping the virtual world. In the following section we turn our attention to describing the results of our recent prototyping work.

5. CyCO - a distributed implementation

So far we have presented a number of models for building virtual worlds. This section describes some of the implementation work that has been carried out in order to demonstrate the feasibility of these models. We start with a general description of the demonstrators constructed so far, including a number of screen images. We then proceed to discuss some of the architectural issues involved in implementing the models. In particular, we describe the communication protocols which operate between our basic room server and client agents and outline the possible use of the X.500 Directory service in allowing room agents to federate and so bring about a potentially global scale virtual world.

5.1 The current demonstrators

Our implementation is called CyCo, *Cyberspace for Cooperation*, and at present consists of a set of related demonstrators grouped around the theme of a rooms based text conferencing system.[1] The core of CyCo is a distributed environment of rooms through which people can move and within which they can hold conversations through the medium of real-time text messages. The system maintains room descriptions and is able to notify occupants of the comings and goings of other people as they wander about the world. In the most basic implementation, the spatial model is reflected through support for three different kinds of interaction.

- *Talkin*g addresses the occupants of a room.
- *Whispering* is directed at an individual but other occupants may be aware that the whisper occurred.
- *Shouting* may be transmitted across several rooms.

CyCo is implemented on top of the ANSA Distributed Processing Platform [ANSA 89] and currently supports two user interfaces, an X Windows interface using the Motif widget set and a teletype interface based on the UNIX Curses C library. Figure 6 shows an example screen from the X Windows interface.

CyCo can be configured to support a specific world design by creating new room descriptions and topology information. Future plans include a management interface to allow the dynamic creation and deletion of rooms. CyCo also employs a random world generator system for testing purposes.

A second demonstrator implements a simple map tool which can be used to aid navigation in conjunction with the basic CyCo system. The current map tool takes the rooms data file and constructs a spanning tree from the given topology. An example screen from the map tool is shown in figure 7. It is important to note that, even though the map may present a spanning tree drawn from the perspective of the users current room, CyCo does not restrict the world to be hierarchical in structure. Thus, the map is only one simplified view onto the underlying world.

A third demonstrator shows a more sophisticated implementation of the spatial metaphor in which focus and nimbus are represented as areas around each occupant of a two dimensional room. So far, only a teletype version has been implemented (based on the style of the UNIX Rogue family of adventure games!). This demonstrator has yet to be integrated with the core CyCo system.

1. Originally developed for the internal project “Conference World”, which supported multiple rooms, and introduced the simple interaction model used by CyCo.

5.2. Distributed architecture

Arguably the most critical issue addressed by the implementation of CyCo has been the design of a suitable distributed architecture that is both scalable and extensible. We now proceed to outline the current CyCo architecture in terms of the software agents involved and the protocols operating between them. The CyCo system is implemented by two kinds of software agent: *room servers* and *room clients*. A room server manages an area of the virtual world. This means that it knows the topology and descriptions of the rooms in this area, manages movement and navigation through these rooms and routes text messages between occupants. Each occupant is represented by a room client which manages their local interface and makes requests of various room servers. The protocol between the room client and room server is defined by a set of interfaces, each of which supports a set of remote operations. Figure 8 shows the four interfaces between the client and server. The arrows indicate the direction in which operations are invoked (e.g. does the client request the operation from the server or vice versa).

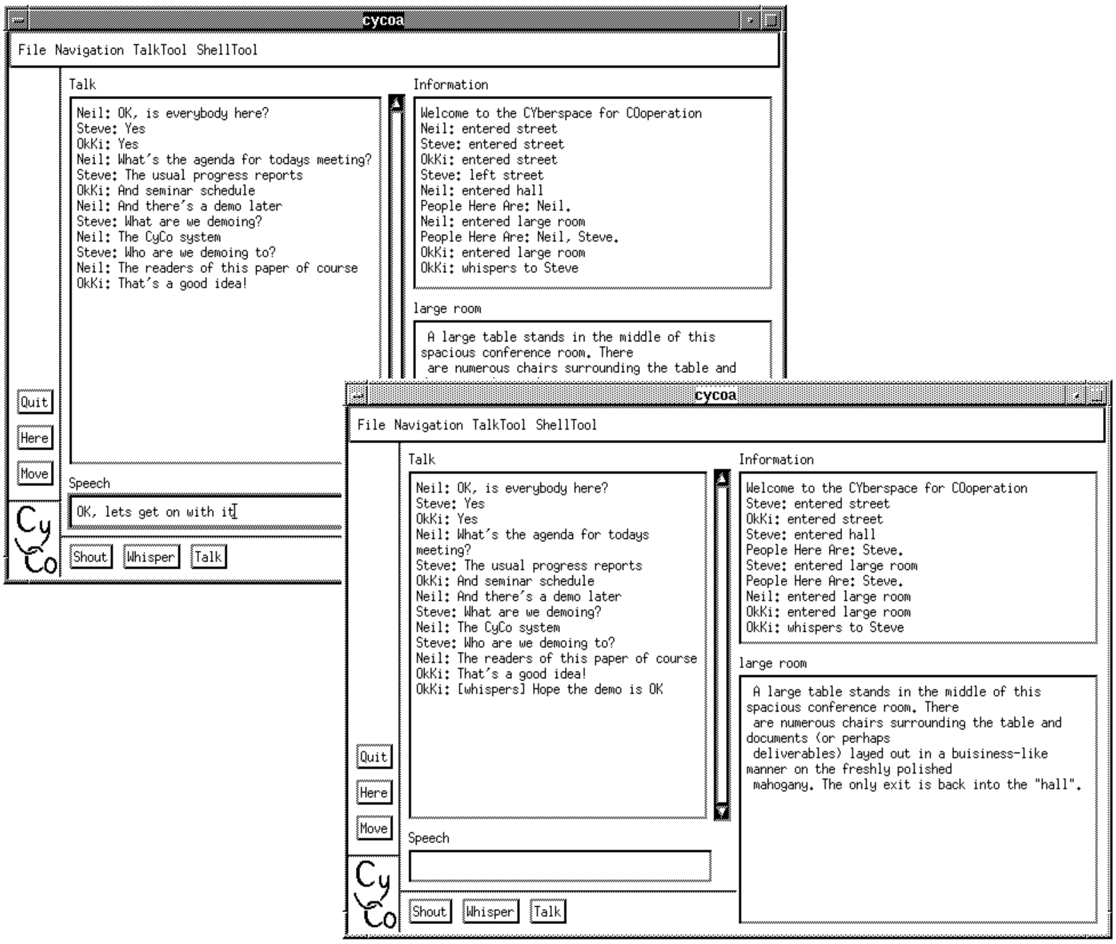

Figure 6: The Fundamental CyCo Interface

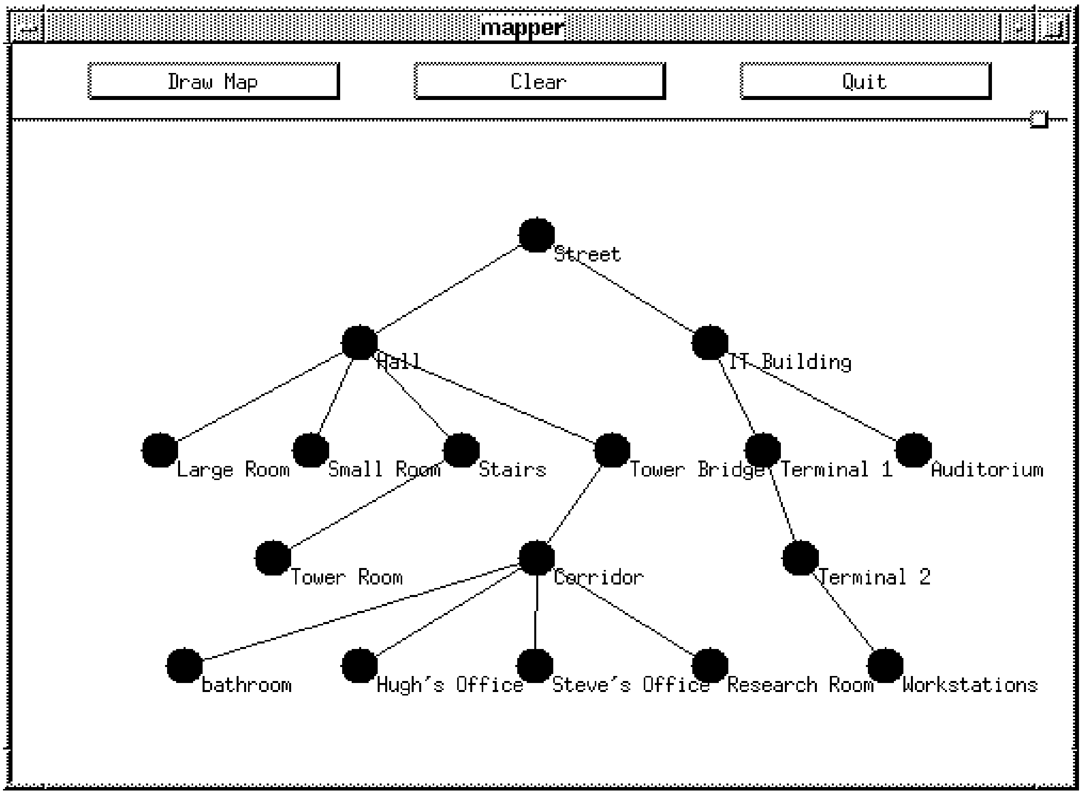

Figure 7: The CyCo Mapper Tool in Operation

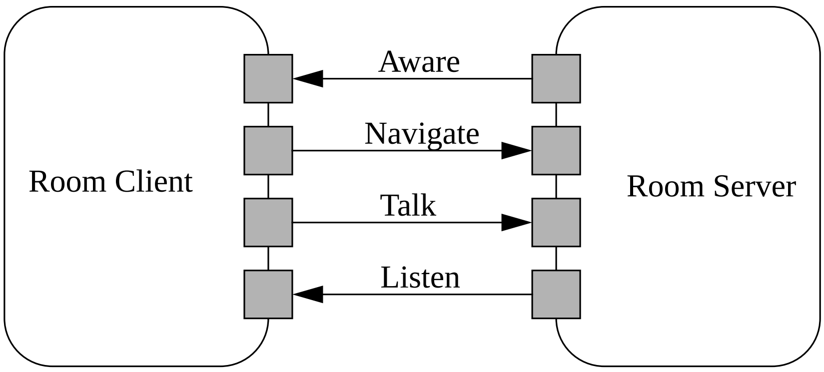

Figure 8: The CyCo Distributed Protocol

The *navigate* interface allows a room client to bind to a room server through the *attach* and *detach* operations and then to move between rooms via the *move room* operation and retrieve room information via the *show one room* and *show all rooms* operations. Attach and detach allow the client to specify which media types it supports thus allowing extensibility to multi-media implementations in the future (only text is available at present).

The *aware* interface is used by the room server to notify the room client about events taking place in a room. This currently utilises the *notify ente*r, *notify leave* and *notify whisper* operations to notify of when people enter the room, leave the room and whisper respectively.

The *talk* interface is used by the room client to transmit utterances to the room server. The three current operations *speak*, *shout* and *whisper* reflect different levels of nimbus derived from the spatial model.

Finally, the *listen* interface is used by the room server to transmit utterances from other people to a room client. It defines *talk, shout* and *whisper* operations which correspond to those in the talk interface.

This protocol defines the interaction between a single room client and room server. Clearly, a large virtual world will need to be partitioned between many room servers both for technical reasons (the limitations of centralised systems) and political reasons (few organisations will be willing to let some other organisation control their part of the virtual world). This implies that we need to consider the interaction between many clients and many servers. Our view is that a future global cyberspace might be realised as a federation of local cyberspaces each of which is maintained by a local room server. The key question then becomes: how does a client know the network location of the correct server for a particular room? For example, when I choose to step through a door between two rooms my client may need to transparently locate and bind to a new server.

We believe that the solution to this problem lies in the development of a global organisational database which maintains entries describing all organisations, people and rooms in the virtual world. A room client would then be able to query this database with the name of a room to determine the room server responsible for its management (stored as an attribute of the room's entry in the database). The development of such a global organisational database is not as far fetched as it might at first seem. Currently, the best candidates for developing a *large scale* organisational database would seem to be the network nameservers such as Xerox's Clearinghouse [Birrell 81], the Internet BIND nameserver [Mockapetris 83] and the OSI Directory Service (X.500) [ISO DS 88, Kille 88]. Our choice has been the X.500 Directory because we believe that it represents the most flexible of the systems available. In short, X.500 maintains a globally distributed database of entries which are made transparently available from any point on the network. Thus, each room in the virtual world will be registered in X.500 Directory with an attribute identifying the relevant room server. Room clients could then query the Directory in order to bind to the correct servers as people moved around the world. This is shown as a "server address look up" in figure 9. This approach seems to open up the possibility of global scale connectivity for virtual worlds and we are currently mapping the information contained in room data files into the Directory.

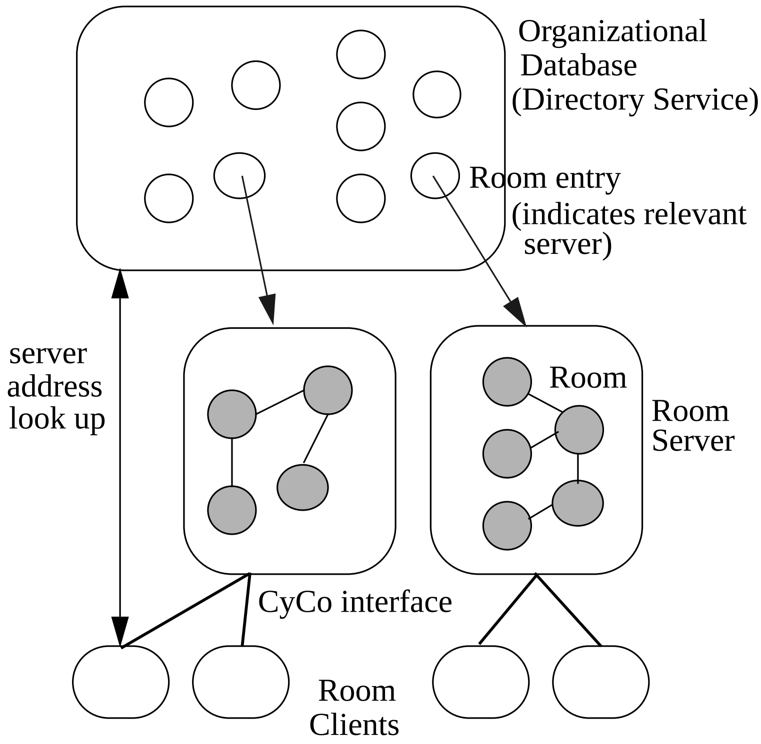

Figure 9: Federated Room Servers Via the Directory Service

6. Summary

Our paper has presented a number of models of group interaction within a distributed cyberspace of connected virtual rooms. The primary goal of these models has been to support cooperation on a *large scale*. The models themselves aim to provide generic mechanisms for promoting and managing interaction which could be applied across a diverse range of systems from text conferencing to virtual reality. In order to illustrate the models we have included discussions of how they might be applied, particularly with respect to our current demonstration system called CyCo. Our overall approach has been based on the increasingly popular "rooms metaphor" where interaction occurs within virtual computer spaces called rooms. We have then extended this metaphor in several novel ways.

First, we considered mechanisms for mediating communication in large virtual rooms and proposed a "spatial model" of interaction where a virtual room acquires properties of size and shape and people use parameters such as position and direction to control what they hear and see. We then described the concepts of *focus* and *nimbus* as two key parameters for controlling interaction. Furthermore, we introduced a formal mathematical framework for modelling focus and nimbus in terms of mathematical functions applying across a domain typically representing the "floor" of a room and showed how they might be combined to determine one person's awareness of another in virtual space. This framework allows different functions to be used to represent focus and nimbus in different systems. The paper discussed two applications of the model, a discrete two dimensional representation applicable to something like a text conferencing system followed by a continuous representation more suited to audio conferencing. It also outlined the representation of various tools which could enhance and alter focus and nimbus.

The second extension of the basic rooms metaphor was to allow the combination of virtual rooms to form more complex virtual structures such as buildings, organisations or even cities. To this end, we proposed two further extensions: allowing nested rooms and linked rooms. The difference between nesting and linking is intended to reflect a distinction between the commonly visible management structures within an organisation and the more personal ad-hoc connections required for navigation. The paper also included a discussion of the role of topology in virtual worlds including the suggestion that we need to avoid certain topologies such as strictly hierarchical worlds. We also presented some work into mapping virtual worlds in order to support user navigation. The final part of the paper described or current CyCo demonstrator which illustrates the technical feasibility of the models through their application to a distributed text conferencing system. In particular, we outlined our distributed CyCo architecture including a description of how the X.500 Directory might support communication between large numbers of room servers in a distributed realisation.

In conclusion, we believe that the increasingly popular “rooms metaphor” is appropriate to the design of large scale collaborative systems. In fact, to go a step further, we believe that the metaphor is a sensible way to approach interfaces to distributed systems in general. We hope that the models that we have proposed can be applied across a wide range of systems to enable them scale effectively. Our long term goal would be to move towards a globally connected cyberspace rich in communication potential. Perhaps in a few years it may be possible to step through the window of your computer into the densely inhabited virtual city that is the global computer network.

References

[Alexander] Christopher Alexander, A City is Not a Tree, in Zone 1/2, Michel Feher and Sanford Kwinter (eds), Urzoen Inc. pp128-157.

[ANSA 89] The ANSA Reference Manual, Architecture Projects Management Ltd, Poseiden House, Castle Park, Cambridge CB3 PRD, UK, 1989.

[Benedikt 91] Michael Benedikt, Cyberspace: Some Proposals, in Cyberspace: First Steps, Michael Benedikt (ed), MIT Press, 1991, pp 273-302.

[Benford 92] Steve Benford, Wolfgang Prinz, John Mariani, Tom Rodden, Leandro Navarro, Elsa Bignoli, Charles Grant Brown and Torbjørn Naslund, MOCCA - A Distributed Environment For Collaboration, In Proc. Telepresence’93, Lille, France, March 1993..

[Borning 91] A. Borning and M. Travers, Two Approaches to Casual Interaction over Computer and Video Networks, In Proc. CHI’91, New Orleans, April 27-May 2, 1991, pp13-19.

[Bignoli 91] C. Bignoli and C. Simone, AI Techniques to Support Human to Human Communication in CHAOS, in Studies in Computer Supported Cooperative Work: Theory, Practice and Design, J. Bowers and S. Benford (eds), Elsevier Science Publishers, 1991.

[Birrell 81] Andrew D. Birrell, Roy Levin, Roger M Needham and Michael D. Schroder, Grapevine, Xerox Palo Alto Research Center report, April 1981.

[Bowers 88] J. Bowers and J Churcher, Local and Global Structuring of Computer Mediated Communication: Developing Linguistic Perspectives on CSCW in COSMOS, In Proc. CSCW’88, Portland, Oregon, September 26-28, 1988.

[Clarkson 91] Mark A. Clarkson, An Easier Interface, BYTE, February 1991, pp227-282.

[Codella 92] Christopher Codella, Reza Jalili, Lawrence Koved, J. Bryan Lewis, Daniel T. Ling, James S. Lipscomb, Favid A. Rabenhorst, Chu P. Wang, Alan Norton, Paula Sweeney, Greg Turk, Interactive Simulation in a Multi-Person Virtual World, In Proc. CHI'92, ACM, 1992.

[Cohen 91] Michael Cohen and Lester F. Ludwig, Multidimensional Audio Window Management, in Computer Supported Cooperative Work and Groupware, Saul Greenberg (ed), Harcourt Brace Jovanovich, 1991, ISBN 0-12-299220-2.

[Cook 91] Cook,S., Birch,G., Murphy, A., and Woolsey, J., Modelling Groupware in the Electronic Office, in Computer-supported Cooperative Work and Groupware, Saul Greenberg (ed), Harcourt Brace Jovanovich, 1991, ISBN 0-12-299220-2.

[Cook 92] Neil Cook and Graeme Lunt, XT-Confer: Dynamic Desktop Conferencing, In Proc. European X User Group Fourth Annual Conference, September, 1992.

[Crowley 90] Crowley, T., Milazzo, P., Baker, E., Forsdick, H. and Tomlinson, R., MM Conf: An Infrastructure for Building Shared Multi-media Applications, In Proc. CSCW'90, October 1990, ACM Press.

[Fahlén 92] Fahlén, L. E. and Brown, C.G., The Use of a 3D Aura Metaphor for Compter Based Conferencing and Teleworking, In Proc. 4th Multi-G Workshop, Stockholm-Kista, May 1992, pp 69-74.

[Fahlén 93] Lennart E. Fahlén, Charles Grant Brown, Olov Ståhl, Christer Carlsson, A Space Based Model for User-Interaction in Shared Synthetic Environments, The Swedish Institute of Computer Science (SICS) (to appear in Proc.-InterCHI'93).

[Furnas 86] Furnas, *Generalised Fisheye Views*, Porceedings of the ACM SIGCHI'86 Conference on Human Factors in Computing Systems, 1986, pp16-23.

[Gale 91] Stephen Gale, Adding Audio and Video to an Office Environment, in Studies in Computer Supported Cooperative Work: Theory, Practice and Design, J. Bowers and S. Benford (eds), Elsevier Science Publishers, 1991.

[Gaver 92a] Gaver W., The Affordances of Media Spaces for Collaboration, In Proc. CSCW'92, Toronto, November 1992, ACM Press.

[Gaver 92b] Gaver, W., Moran, T., MacLean, A.,Lovstrand, L., Dourish,P., Carter, K. and Buxton W., Realising a Video Environment: EuroPARC's RAVE System, In Proc. CHI '92 Human Factors in Computing Systems, Monterey, Ca., USA, May 1992.

[Henderson 85] Henderson and Card, Rooms: The Use of Multiple Virtual Workspaces to Reduce Space Contention, ACM Transactions on Graphics, Vol. 5, No. 3, July 1985.

[ISO DS 88] ISO and CCITT, *Information Processing Systems - Open Systems Interconnection - The Directory*, ISO 9594-108, CCITT X.500-X.521, 1988.

[ISO ODP 90] ISO, *Information Procesing Systems - Open Systems Interconnection - Open Distributed Processing (ODP)*, ISO/TC 97/SC21/WG7, 1990.

[Kille 88] Steve Kille, T*he QUIPU Directory Service*, Proceedings of the IFIP 6.5 Conference on Message Handling Systems and Distributed Applications, Costa-Mesa, CA, USA, 1988, North-Holland, pp 173-186.

[Leevers 92] David Leevers, Prototyping Multimedia Communications for Manufacturing SME's, Presented at the CI Europe Seminar, 2-3rd July 1992, and also an internal report of BICC Central Development, Maylands Avenue, Hemel Hempstead, UK.

[Lovstrand 91] Lennart Lovstrand, *Being Selectively Aware with the Khronika System*, Proceedings of ECSCW'91, Amsterdam, Hollandm Kluwer Academic Publishers, pp265-278.

[Michelitsch 91] Georg Michelitsch, Providing a Shared Environment to Distributed Work Groups, In Proc. GLOBE-COM’91, Pheonix, Arizona, December 2-5th, 1991.

[Mockapetris 83] Paul Mockapetris, *The Domain Name System*, in Computer Based Message Services (Proceedings of the IFIP 6.5 Conference), Nottingham, UK, 1984, H. T. Smith (ed), North-Holland.

[Morningstar 91] Chip Morningstar and F. Randall Farmer, The Lessons of Lucasfilm’s Habitat, in Cyberspace: First Steps, Michael Benedikt (ed), MIT Press, 1991, pp 273-302.

[Navarro 92] L. Navarro, M. Medina and T. Rodden, Environment Support for Cooperative Working, In Proc. IFIP 6.5 ULPAA’92 Conference, Vancouver, Canada, May 1992, North-Holland.

[Pankoke 89] Uta Pankoke-Babatz (ed), Computer Based Group Communication - The Amigo Activity Model, Ellis-Horwood, 1989.

[Root 88] R.W. Root, Design of a Multi-Media Vehicle for Social Browsing, In Proc. CSCW’88, Portland, Oregon, Spetember 26-28 1988, pp25-38.

[Sarin 91] Sunil Sarin and Irene Greif, Computer-Based Real-Time Conferencing Systems, in Computer Supported Cooperative Work: A Book of Readings, Irene Greiff (ed), Morgan Kaufmann, 1988, pp397-421.

[Sarkar 92] Manojit Sarkar and Marc H. Brown, Graphical Fisheye Views of Graphs, In Proc. ACM SIGCHI’92 Conference on Human Factors in Computing Systems, May 3-7 1992, pp 83-91, ACM Press.

[Smith 87] R.B. Smith, The Alternate Reality Kit: an example of the tension between literalism and magic, Proc. CHI+GI87 Conference on Human factors in Computing Systems, (April 5-9, Toronto), ACM Press, 1987.

[Smith 92] Jennifer Smith (ed), Frequently Asked Questions: Basic Information about MUDs and MUDing, Posting to the alt.mud USENET News Group, July 1992.

[Takemura 92] Haruo Takemure and Fumio Kishino, Cooperative Work Environment Using Virtual Workspace, In Proc. CSCW'92, Toronto, Nov 1992, ACM Press.

[Victor 91] F. Victor and E. Sommer, Supporting the Design of Office Procedures in the Domino System, in Studies in Computer Supported Cooperative Work: Theory, Practice and Design, J. Bowers and S. Benford (eds), Elsevier Science Publishers, 1991.

[Winograd 86] T. Winograd and F. Flores, Understanding Computers and Cognition: A New Foundation for Design, Norwood, New Jersey, Ablex 1986 and Addison-Wesley 1987.