

# EMCE: A Multimodal Environment Augmenting Conferencing Experiences

Arna Ionescu and Luc Julia

*Computer Human Interaction Center (CHIC!) - SRI International  
333 Ravenswood Avenue, Menlo Park, CA 94025 – USA  
ionescu@cs.princeton.edu and julia@speech.sri.com*

## Abstract

*To illustrate research regarding how to augment people's experiences in office meetings of different types, we have developed EMCE (Enhanced Multimodal Conferencing Environment), a prototypic conference room. EMCE assists meeting participants in performing a host of functions including but not limited to passing private messages, taking electronic notes, accessing personal files, writing both public and private annotations on projected objects, automatically creating meeting minutes and remote conferencing. Our multimodal approach allows people to interact with the conference room through spoken or handwritten commands and drawn information. The room reacts in a multimedia fashion, outputting sound, video, text, etc. To make EMCE's use as intuitive as possible, our interface consists of a virtual view of the conference room that users can access either with screens embedded in a conference table or on their personal laptops. Our goal is to create an augmented environment that is as natural and easy to use as existing, nonaugmented meeting environments.*

## 1. Introduction

Proposals for augmented spaces are everywhere. Researchers and developers are working hard to find ways to remove the tedious, repetitive aspects of everyday life by augmenting the world around us with embedded computers. A computer could easily perform many daily tasks thus freeing up valuable human time. On the Internet one can find articles about refrigerators that keep themselves stocked automatically, cars that “know” where they are going and that “perform” various office tasks while driving, televisions that can tailor programming to a user's tastes, and so on. All of these devices meant to augment a user's experience are subtly placed into an environment and work “behind our backs,” making many mundane tasks unnecessary, or at least a lot easier and less time-consuming.

One environment that has great potential for augmentation is a meeting room. There are many aspects

of meetings that we could all stand to have made a little easier. For instance, if a manager wants a summary of what occurs in a meeting, someone has to take minutes. This means that one person in the meeting cannot participate fully, because this person is busy preparing a document about what transpired. Another example is the series of actions that occur when one forgets to bring an important document to the meeting and must either waste valuable time to retrieve the document from an office or else proceed in the meeting without it. There are also times when two people may need to communicate privately, such as during negotiation proceedings, and there is presently no way to communicate subtly with only a single person during the course of a meeting. Another area that could be augmented is the task of taking notes. Few people enjoy taking notes, and notes often end up being very disorganized and unusable. Good archives of notes that are accessible years later are especially rare. A convenient method of keeping personal notes, not necessarily meeting minutes, would be very helpful.

Another essential aspect of meetings is the projection system. To switch between viewing documents on different media (such as from an overhead projector, to computer slides, to a VCR, to a whiteboard) usually requires a good amount of fiddling with hardware. Manipulating hardware should not be necessary, thus leading us to an obvious area that could be augmented. Furthermore, in most meetings only one person can write on the whiteboard or slides. This is often cumbersome when a participant other than the presenter wants to illustrate a question or a comment with a drawing. Unless that person loses valuable time by getting up and going to the whiteboard or computer, there is no way to include an illustration along with the spoken comments. This is another problem that could be greatly reduced by an augmented environment.

After looking at the features of a meeting that could be augmented and designing solutions for the problems we identified, we developed a prototypic conference room we call EMCE (pronounced “em-cee”), for Enhanced Multimodal Conferencing Environment. For each user, this room provides personal display devices that present a

virtual view of the physical conference room, allowing for the augmentation of many functions. Our main goal in creating this prototype was to develop an interface for the display device that is as natural and intuitive to use as possible; we believe that EMCE will not be taken advantage of if users must dramatically change their present method of participating in meetings. The purpose is to augment an existing environment, not to completely change it. We followed the virtual room metaphor to design the interface in hopes that it would lead to more intuitive interactions.

In this paper, we describe our solutions to the various problems that we have identified as arising in meeting situations. We have implemented many of our solutions with the hope of performing extensive user studies on our prototype.

## 2. Related Work

Several researchers have studied the concept of smart rooms. A group at Georgia Institute of Technology has worked on a project called Classroom 2000, which is a ubiquitous computing lecture hall that augments a student's lecture experience [1]. The class automatically takes notes for students so that they can focus more on the lecture rather than on writing it down. Another project looking at augmented rooms is MIT's Intelligent Room [2], in which researchers have created a system that tracks the actions of room inhabitants and discerns between commands spoken to the system and conversations between people. This room serves as a basis for a variety of specialized services. One example is an environment in which to give a presentation.

Some researchers have used a room metaphor in developing virtual environments. The TeamWave project, in which researchers developed an environment where participants use various electronic features to share information and work collaboratively, illustrates the use of this metaphor [3]. Features include a communal whiteboard, the ability to send electronic PostIt notes to each other, and the ability to share files.

A few groups have worked specifically on augmented meeting rooms. Kumo Interactive, for instance, is a smart conference room designed at FXPAL [4]. This project includes provisions for electronic note taking, automatic meeting recording, the creation of meeting archives, equipment control, and a simple projection system. Another collaborative conference room is the i-LAND project created at the German National Research Center for Information Technology [5]. In this system, an entire room becomes networked for shared group work. Large displays on a table or wall and personal display devices all merge to form an integrated system.

Work has also been done on various components that would belong in an augmented conference room. One

such project is Flatland, an interactive, electronic whiteboard [6]. While Flatland would not create a fully augmented space on its own, it should definitely play an important role in a smart conference room.

In designing our prototype, we looked at various pieces of related work and tried to synthesize them.

## 3. Design and Implementation of EMCE

### 3.1. Overview

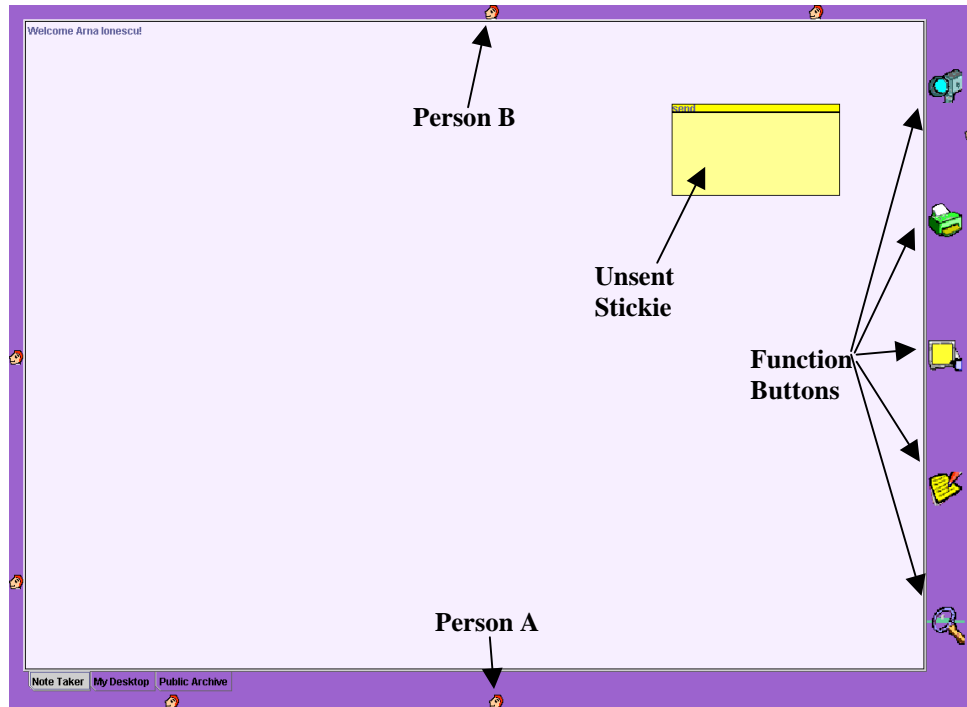
EMCE is implemented with the Open Agent Architecture (OAA™) [7], a distributed multiagent framework that enables rapid integration of component technologies. This architecture allows for quick integration of new features, and, by combining different sources of information, we end up with a system that is much greater than its parts.

Physically, an EMCE conference room has three main differences from traditional conference rooms: it has touchscreens embedded in the table to serve as personal display devices, it has an electronic whiteboard, and it has a microphone and earphone at each location at the table. The rest of the hardware is what one would find in a traditional conference room: a computer projector, VCR, printer, and so on.

Each participant must log in to the meeting upon arrival. Information about each user is stored in a text database and is retrieved by the system as necessary. Each user can either sit down at an existing console or bring in a personal laptop that connects to the system through a wireless Ethernet card, using the DHCP protocol.

EMCE is designed to work in two different types of meeting environments: collaborative and led. In a collaborative meeting each participant is on equal ground with equal permissions regarding placing objects into the projected space and writing public comments on objects. In a led meeting, the person who logs in as the leader has the permission to place items on public view and to write public comments for everyone to see. A leader is also able to temporarily pass leadership permission to a participant of choice.

The interface is designed to work with a pen on a touchscreen. For people using a laptop without touchscreens, a mouse and keyboard may be substituted as input devices. The main goal in designing our interface was to make it as intuitive as possible to use within the context of a meeting. Thus, functions performed electronically in EMCE are still executed in a manner similar to how they would be performed without EMCE. For features that are new to a meeting within the context of EMCE, we tried to retain the idea of the virtual room and to make the interface work in a manner relative to the physical space. We want people to have to adapt as



**Figure 1. EMCE interface for Person A**

little as possible when entering an EMCE conference room, as we do not believe that people will agree to dramatically change their manner of meeting participation.

### 3.2. Virtual View

Our idea of creating a physical representation of the space was influenced by a project at IBM in which objects can be dragged between computers in a room according to their physical relationships in space [8]. When users log into EMCE, they see a virtual representation of the meeting table in front of them. All meeting participants are represented as icons along the edge of the screen. As users log in and out, the icons appear and disappear. The icons always appear on the desktop relative to where a person is sitting. Thus, for person A's display, person A's icon appears at the bottom of the screen and the icon of person B, who is physically sitting directly in front of person A, appears at the top of person A's screen. The icons are flipped for person B: person B is at the bottom of the screen, and person A is at the top of the screen (Figure 1).

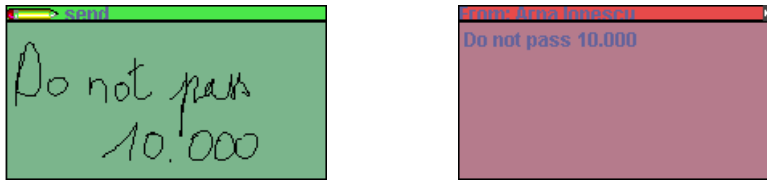
The icons representing the participants are useful in several ways. The first purpose is role identification. In a led meeting, for instance, a leader is represented by a different icon than the rest of the participants. When a user is temporarily given leadership permission, that user's icon changes as well. A second purpose is to obtain personal information. If a user passes the pen over an

icon, the name of the person whom the icon represents appears. If the user then presses down with the pen on the icon, a box with information about that person appears. This box contains that person's e-mail address, position, office location, telephone number, and so on. Thus, a meeting participant always knows the identity of all the other people logged into EMCE for a given meeting.

Another way in which the virtual space could be used is to represent a physical robot on the table. A small, independent robot moving around the meeting table could appear on the virtual table space as an icon relative to where it is actually located. In this manner, a user would just have to look down at the screen to know where the robot is physically located. Furthermore, a user could easily guide the robot around the table by simply dragging the icon around the screen.

The concept of the virtual space is also used regarding some of the functionalities directly linked to hardware. While a person does not need to know exactly where in a room the various pieces of hardware are located (i.e., printer, projector), each of these objects is given a virtual location represented by an icon on the screen. A user can interact with the hardware by dragging and dropping both the icons themselves and files onto the icons.

To create the virtual view, EMCE must know where each participant is located in the room. Eventually, we want to have a proactive system, with a database of voices that can use speaker identification methods that will make user login unnecessary. For now, however, the system is reactive and each location in the room is assigned a letter



**Figure 2. Stickies showing handwriting recognition; first image is of unsent stickie, second image is of received stickie**

that a user must enter upon logging in. A central agent checks for discrepancies (multiple leaders or multiple people in one location), and also provides a list of users and locations that can then be processed by an algorithm to determine each user's personal view of the meeting table.

### 3.3. Privacy and the Stickie Feature

As discussed earlier, it is virtually impossible to pass completely private notes in meetings. Handing a piece of paper to someone is usually distracting and often draws unwanted attention. We used the stickie pad metaphor in designing our solution to passing private messages. All users have a pad of stickies on their screens on which they can handwrite a note with a pen, and have the handwriting be recognized, turned into typed print, and sent to another meeting participant.

A user can move a stickie around the screen at leisure by simply dragging and dropping it. If the user wants to write a message to another participant, then the stickie must be locked into a writing mode. This is done by quickly tapping it with the pen. A pencil image appears at the top of the stickie to indicate that it is locked. In the writing mode, a close button appears when a stickie is moved and a new empty stickie is placed on top of the original stickie stack. A written message is recognized by Paragraph's Calligrapher handwriting recognition technology [9], which uses a dictionary and picks up mixtures of cursive and printed characters to achieve a high accuracy rate (Figure 2).

A user can keep the stickie on his or her desktop as a personal note or can drag the stickie onto the icon of another user (again using the functionality provided by viewing a virtual room). Once a stickie is sent, it appears in a different color on the recipient's desktop with the sender's name in the title bar.

An interesting issue regarding the stickies from an interface perspective is that we could not model the procedure after present meeting behavior since no such actions presently occur in meetings. Thus, we modeled the interface on a behavior that occurs in an office setting, namely the use of stickie notes. We do not know how the behavior will port to a meeting, but we hope that the action will be familiar enough for the functionality to be natural in the new setting. Because this is a useful feature,

especially in certain kinds of meetings (such as during negotiations), we hope that its foreignness to a meeting environment will pass quickly in light of the facility it brings.

### 3.4. Public Archive

People frequently want to review an image that was previously projected for all participants to see. Thus, EMCE archives all the objects that have been placed on public view. Participants, who can browse the archive at any time during the meeting, see a stack of tabs, with the most recently projected view on top. A whiteboard image can also be included in the stack. In addition, two checkboxes labeled "public" and "private" deal with annotation issues.

A participant can use the public archive to write private notes as well as notes for all participants to see. So that people do not accidentally write public notes they intend for private use, a private checkbox is given priority. If both public and private boxes are checked, notes being written are private. A user who wants to remove an annotation from view can scribble over it; this gesture will be recognized and the note will be deleted.

To avoid confusion about what is being viewed publicly, public annotations appear in a different color. Furthermore, the object on the top of the stack being viewed is slightly masked to let the user know that that object is being projected. The user may still write on the projected object, but the constant visual feedback is a reminder of what is being viewed on the projected screen. The public archive and projected view has the potential of leading to numerous blunders, and we designed this interface with the idea of minimizing such problems.

If a user decides to re-project an older view from the meeting, then that view must be dragged onto the projection icon, thus placing a second version on the top of the archive stack. The checkboxes can be used to state whether private and/or public annotations should be included in the new view.

One feature of the projected view is that any diagram or simple drawing is cleaned up. For instance, as a user draws a table, the lines of the table are straightened and the written text inside the table is changed to typed text [10]. Thus, cleaner documents are produced, and after the meeting there is no need for someone to process the ideas

and concepts discussed. The presentable form is created automatically as the participants work, thus saving valuable time.

### **3.5. Minutes**

The projected archive serves as a timeline for what occurred during a meeting. As an object is viewed, even if it was viewed earlier, it is placed on the top of the archival view. At the end of the meeting, all the images in the stack are processed, and minutes are created in HTML. The minutes, which can be viewed with any Web browser, build a timeline of the images from the projected view, with sound and video clips of the meeting available upon demand. Public annotations that were not removed are included, and private minutes files could be created to allow users to view their personal annotations as well. It might also be beneficial at some point to merge the projected items with a person's own notes to create a very personalized set of minutes.

In some cases, users may want a document not placed on the public stack to be included in the meeting minutes. EMCE has a virtual "minutes keeper" that appears in the form of an icon, and anything dragged onto it is included in the minutes document.

EMCE also archives the minutes for access during subsequent meetings. Discussion topics are tracked so that a list of past meeting minutes relevant to the topic at hand can be proactively created. At any time, a participant can bring up an old set of minutes either by selecting one from the list or by entering a keyword.

### **3.6. Note Taker**

Meeting notes are often cumbersome to take and often lost. Although laptops can be used for note-taking, most people learned to take notes with pen and paper, and they feel less comfortable doing so with a keyboard. Furthermore, many people include drawings as well as text in their notes, so a keyboard is insufficient. We solved the problem by creating an electronic note taker that is used much like pen and paper. Since the system is electronic, the notes can easily be archived for quick retrieval and the text and diagrams can be cleaned up for easier reading.

An extra feature not possible with traditional pen and paper note taking is the ability to add pictures or objects that are presented during a meeting. A user can include an image in personal notes by simply dragging it into the note taker area. Eventually, the note archive should include the minutes created with a user's personal notes. We believe that both the personal note taker and the ability to write private annotations on the projected archive will be used because they serve slightly different purposes. One feature gives a user a blank slate on which

to write and draw whatever is needed, and the other provides an explicit background on which to add personal reactions to a given object.

### **3.7. Desktop Link**

It is not unusual for someone to enter a meeting and realize that an important document has been left on a desktop computer in an office. In most cases, neither of the two options available is satisfactory: the participant can either waste valuable time by going back to the office, or the meeting can continue without the document. EMCE solves this problem by providing an open, two-way connection with each participant's desktop. Each user simply needs to install a server on his or her computer in advance (an install wizard makes this process very simple), and then enter the password set for that computer upon logging in to EMCE. This gives participants the ability to work on their desktops remotely. During a meeting, a user can retrieve files from a desktop computer or run any application remotely. For instance, a person who wants a personal reminder note opens a word processing application and writes a note, which will be on the desktop computer exactly as it was written during the meeting.

This desktop feature uses open-source software called Virtual Computer Network (VCN) developed by AT&T labs in Cambridge, England [11]. While this package runs especially slowly on older machines, it is very reliable and easy to use and truly allows a user to work remotely. Every change is sent to the desktop so that it appears as if a ghost is working at the desktop computer. The key to the system is that it is intuitive, and people who have background applications running that track various activities do not lose the information collected while working remotely. After user studies, we may decide to allow users to configure their systems so that only rare, discrete packages of information are sent. This will make the connection faster, but the feature may not be as intuitive to use and will add hassle to the setup process. We want the setup and use of EMCE to be as simple as possible so that its use is not restricted to people in technical fields.

### **3.8. Remote Conferencing**

Many remote conferencing tools are available, and most of them use a Web interface. None of these tools, however, really combine two conference rooms into one virtual space and break down physical barriers. As each user's view consists of a virtual conference room, it is quite natural to add virtual people to represent remote users. The virtual people are placed around the outside of the entire view.

In the center of an EMCE conference table is a 360-degree camera that can view all participants. One of the virtual pieces of hardware on the screen is a video camera that a user can drag to a person icon and thus view a video stream of that person. If the user makes no selection regarding which remote users to view, then the person speaking is automatically brought into view.

With regard to sound, EMCE has an open speaker/microphone system. Everything that is said is transmitted, just as in the traditional conference environment where one hears essentially everything that is being spoken.

One major advantage that EMCE offers is that remote participants can project onto their wall whatever image is being projected in the other conference locations. The public archive allows remote users to participate in the same manner as local users because they all have access to the exact same archive. Remote users can also write notes on the projected view for everyone to see, and they can write on a whiteboard and have their writing appear on the whiteboards in all meeting locations. Most remote conferencing systems do not allow all users such easy access to writing on a major projected view. Furthermore, the stickie feature of EMCE allows users to send private messages to remote participants. There would really be no way to do this without EMCE, as a piece of paper cannot be passed between people in different locations.

### **3.9. Utilities: Printing, VCR, E-mail, Voicemail, and Environment**

It is often essential for all meeting participants to have hardcopies of a file that is being discussed. EMCE provides a virtual printer onto which a user can drag a file, causing the correct number of documents to be printed at each location of the meeting. A file is also printed at each remote location.

Some people making a presentation find themselves in need of a VCR. Since no one should have to get up and fiddle with a separate piece of hardware, we include the ability to work a VCR from a user's virtual screen.

Furthermore, if people need to dim the lights for their presentation, they should be able to accomplish this from their seat at the table. The ability to open and close windows and doors may also be a useful feature to add at some point in the future.

A voicemail feature is also included to notify participants of important messages they receive during a meeting. When an important (as defined by the user) message arrives, either a note appears on the participant's screen or a sound plays in the earphone, depending on the user's preference.

Another desirable feature is the sending of a document or a message by e-mail to a group of people at once. EMCE satisfies this need by allowing a user to verbally

request that document X be sent by e-mail to persons A, B, and C. The printing and VCR functions can also be accessed with natural-language voice commands. A user can simply ask the conference room to print a document or play a video that is in the VCR. A user signifies that he or she is speaking to the computer by touching a button on the microphone. Eventually, however, we would like to incorporate MIT's intelligent room technology and have EMCE be able to automatically discern when it is being spoken to [2].

The printing, VCR, e-mail, and voicemail features each provide a function that, while not central to EMCE, we hope will be useful. They can all be accessed by different modalities (pen and voice), and we believe that they will further enhance our augmented space.

## **4. Future Work**

Many of our solutions for augmenting a meeting environment are presently implemented. The note taker, remote conferencing, proactive topic tracking and some of the utilities remain to be implemented, but the heart of the project -- the public archive and virtual view of the room -- are in working order. We have designed EMCE to be used only by direct meeting participants. However, the functionality could be expanded to include more people in different roles. For instance, a person whose role is to bring snacks to meeting participants could be presented with a different EMCE view containing information on each person's location and food preference.

As soon as full implementation is completed, our next step will be to perform extensive user studies. There are many questions to answer regarding how people will use EMCE. For instance, will EMCE truly be natural and intuitive to use? Our hope in creating this augmented environment is that it will not cause significant changes in a participant's behavior. We tried to model our electronic features on actual behaviors. We want to determine whether we have succeeded in our goal. As with any new technology, while we do not desire behaviors to change drastically, we do expect them to at least change slightly. Thus, user studies looking at the amount and permanency of behavior change are also in order. Our hope is that using EMCE will become natural to the degree that subsequent use of nonaugmented meeting environments will become cumbersome because of the lack of electronic aids. A user study evaluating how well people adapt to EMCE and how necessary it becomes during the course of meetings would be very interesting.

## **5. References**

- [1] Abowd, G., Atkinson, C., Feinstein, A., Hmelo, C., Kooper, R., Long S., Sawhney, N., and Tani, M. Teaching and

Learning as Multimedia Authoring: The Classroom 2000 Project. *In the Proceedings of the ACM Conference on Multimedia '96*, pp. 187-198, November 1996.

[2] Brooks, R. A. with contributions from M. Coen, D. Dang, J. DeBonet, J. Kramer, T. Lozano-Perez, J. Mellor, P. Pook, C. Stauffer, L. Stein, M. Torrance, and M. Wessler. The Intelligent Room Project. *Proceedings of the Second International Cognitive Technology Conference (CT'97)*, Aizu, Japan, August 1997.

[3] Roseman, M., and Greenberg, S. TeamRooms: Network Places for Collaboration. *Proceedings of ACM CSCW'96 Conference on Computer-Supported Cooperative Work*, pp. 325-333, 1996.

[4] Chui, P., and Wilcox, L. Kumo Interactive: A Smart Conference Room. *DARPA/NIST/NSF Workshop on Research Issues in Smart Computing Environments Position Paper*, July 1999.

[5] Streitz, N., Geibler, J., Holmer, T., Konomi, S., Muller-Tomfelds, C., Reischl, W., Rexroth, P., Seitz, P., and Steinmetz, R. i-LAND: An interactive Landscape for Creativity and Innovation. *CHI'99 Conference Companion*. ACM Press, pp. 120-127.

[6] Mynatt, E., Igarashi, T., Edwards K., and LaMarca, A. Flatland: NewDimensions in Office Whiteboards. *CHI'99 Conference Companion*, ACM Press, pp. 346-353.

[7] Martin, D., Cheyer, A., and Moran, D. The Open Agent Architecture: A framework for building distributed software systems. *Applied Artificial Intelligence: An International Journal*, volume 13, number 1-2, January-March 99, pp. 91-128.

[8] Rekimoto, Jun, and Saitoh, Masanori. Augmented Surfaces: A Spatially Continuous Work Space for Hybrid Computer Environments. *CHI'99 Conference Companion*. ACM Press, pp. 378-385.

[9] Paragraph: <http://www.paragraph.com/>

[10] Julia, L. and Faure, C. Pattern Recognition and Beautification for a Pen Based Interface. *In the Proceedings of ICDAR'95 : Montreal (Canada)*, pp. 58-63, 1995.

[11] Virtual Computer Network:  
<http://www.uk.research.att.com/vnc/>