Development of the software infrastructure for a ubiquitous computing environment – the DSTO iRoom

Jason Scholz, Mark Grigg, Paul Prekop, Mark Burnett
Command and Control Division
Defence Science and Technology Organisation
Department of Defence, Fern Hill Park
Canberra ACT 2600
AUSTRALIA

{jason.scholz, mark.grigg, paul.prekop, mark.burnett} @defence.gov.au

Abstract

This paper describes the development of a ubiquitous computing environment called the iRoom. Built from a foundation of the MIT Metaglue multi-agent software, a base level of functionality has been established in the iRoom. We discuss the research issues, applications currently under development, and a broader theme called iEcology of which we see iRoom as an instance. [a]

Keywords: pervasive computing, multi-agent systems, military command and control.

1 Introduction

The pervasive computing vision: "instead of people learning how to work with computers, pervasive computing is about the computers learning how to work with people" has seen early instantiations in a number of "smart" rooms or spaces in research organisations. These include MIT's Intelligent Room (MIT 2000), Stanford's iRoom Project (Stanford 2001), NIST's Smart Space Lab (NIST 2001) and Georgia Tech's Aware Home Project (Georgia 2000).

For example MIT's Intelligent Room (MIT 2000) (which is part of the Oxygen project (MIT 2002)) hosts a variety of projects exploring collaborative tools, interfaces modalities and other innovative applications. Within the MIT's Intelligent Room, vision and speech are the dominant modes of interaction, with current research exploring how perception by the room -- in the form of speaker identification, face-recognition, and gaze tracking -- can be used to improve interaction within the room. A multi-agent development environment called Metaglue handles communication and co-ordination between devices and software in the room, and provides the digital glue that holds the room together. In contrast to the "computers everywhere" model of ubiquitous computing researchers at MIT have been motivated by the desire "to enable unencumbered interaction with non-augmented, non-computational objects (like chairs) and to do so without requiring that people attach high-tech gadgetry to their bodies" (Coen 1998).

While recognising that the optimal design of an intelligent environment is an open question, we have taken the approach that the issue will be easier to deal with in a domain-specific environment. The DSTO iRoom builds on the Metaglue infrastructure with the goal of exploring pervasive computing information systems and infrastructure to support military command and control (C2). In this paper we report on the current status of the iRoom, the research issues we see and the experiments underway.

The iRoom is an instance of an *iEcology* environment – our term for research into pervasive computing information systems. In pervasive computing, the ecology metaphor is used to evoke the picture of many different software and hardware artefacts, at varying levels of complexity, interacting in a complex environment to support human activities. The comparison with living systems can be pushed further to include competition or co-operation over resources, the need for species (purpose-specific collections of agents) to survive and so on. The research underway in iEcology applies to many different real world situations, both inside and outside. Currently the environment under investigation is the iRoom but in due course we expect to instantiate the ideas in iEcology in other ways.

2 Military Application and Issues

Pervasive computing information systems may improve decision-making and increase the tempo of Command and Control (C2)[1]. Specifically in supporting the "operational execution" dimension of C2. Operational execution is that part of the C2 process that matches plans to actions, providing command guidance to forces, and monitoring of activities to control the flow of the operation according to the plan. To achieve this, the iRoom agents, applications, and services will need to be made "aware" of the schema and specific language of military campaign plans as its digitised "context". From this baseline, activities and events that may be reasoned

[1] C2 is defined by NATO (NATO 1996) as "the organisation, process, procedures and systems necessary to allow timely political and military decision making and to enable military commanders to direct and control military forces".

to potentially affect achievement of the mission or desired end states may be identified and reasoned about as early as possible.

Command and Control centres are a well known and favoured target in modern warfare, and it is therefore desirable to distribute the C2 function in such a way as to reduce its vulnerability and improve survivability. We envisage a future of "ubiquitous C2" (Lambert 1999) in which all assets possess a similar and significant C2 capability, such as headquarters, aircraft, ships, missiles, UAV's, and so on. Such a system would gracefully degrade in its capability as parts of the total system are destroyed or compromised in combat. We believe that the design of an efficient distributed information system with these features would rely on coding principles, rather than simple duplication or replication of functionality.

The iRoom infrastructure features layering, abstraction and information hiding or "encapsulation" as the key means to building greater complexity as an information system. Metaglue also supports a number of communication mechanisms between agents including publish/subscribe and event broadcasting, and maintains agent state so that agents may be re-started in the event of failure. However, a key issue for the survivability of the infrastructure is the danger of excessive encapsulation or "hyper-encapsulation" (Browne 2000). When a system is subject only to passive threats (for example hardware failures) such as those that commercial middleware systems are designed for, or even weak active threats (for example node loss in combat), in general, increased encapsulation does not harm survivability. However, strong active threats may exploit this. We believe that part of the multi-agent system ecology requires specific agents at each layer which are capable of analysis and reasoning about abnormal behaviour, and sharing this information with agents at other layers, in order to affect an over-riding of normal responses in likely threat situations.

3 Research Issues

Three main areas of research under exploration in the iRoom, with the following broad themes:

1. What are the computational needs of ubiquitous computing environments and the software infrastructure that supports them?
2. Support for specific human activities within the environment – what should the room know about the people and activities inside the room, and how best to represent this for computational purposes?
3. Interaction with the environment – how do people in the room interact most effectively with a largely hidden software infrastructure and what new types of interaction does the environment afford?

Although presented here as distinct, these themes have strong overlaps and interconnections. The next section describes these areas in more detail.

3.1 Computational needs of an iRoom environment

Metaglue provides a low-level communication medium for agents in the room and a useful multi-agent development environment. We are now beginning to investigate whether higher-level agent frameworks can help bridge the communicative gap between man and machine, and improve machine-machine performance. This is predicated on the belief that to build "intelligence" and awareness into the software infrastructure the software must be capable of responding autonomously to requests from its surroundings. The surroundings include software, people, sensors and other environmental inputs. Agent frameworks, with support for constructs such as plans, beliefs and desires, offer one avenue of attack to this problem.

3.2 Support for specific human activities within the environment

The notion of *context* as an overarching concept for ubiquitous computing environments is being explored in relation to its usefulness in supporting specific activities. Typical activities within an iRoom include meetings of various types, presentations and planning sessions. At a finer level of detail each of these activities involves information access, information sharing and often involves making sense of a complex issue. Context, which can be defined as any information that can be used to characterize the situation of an entity, may be a useful concept in this space, and one that allows activities at many levels to be modelled within a single framework. A *context aware* application is defined as one that uses the context of an entity to modify it behaviour to best meet the context of the user (Dey and Abowd 1999). Currently most attempts to use context-awareness within ubiquitous computing environments have centred on the physical elements of the environment, the user, or the user's device. While many authors acknowledge the importance of capturing the cognitive elements of a user's context (Schmidt et al. 1999) little work has been done to develop models to support this. (Prekop and Burnett In Press) have presented an activity-based model that focuses on the context that surrounds the performance of an activity by an agent. Based on the thinking in this paper and other realisable models of context (Dey et al. 1999) we have begun experimenting with context awareness in the iRoom to support complex cognitive activities.

3.3 Interaction with the environment

Language is the fundamental basis of interaction, and it is here that the age-old problem of "knowledge capture" surfaces. Some emerging technologies claim to solve the human-machine interaction problem, requiring some 3-6 months to capture the domain-specific knowledge of a user – but the issue here is how the system can learn and change its agent ontologies, goals and plans of its agents, in response to inevitable change in human activity, without having to spend an exorbitant amount of time retraining the machine?

Multi-modal interaction research examines how to obtain more accurate measurement of human activity by simultaneous instrumentation of many attributes of an interaction (eg. speech, visual gaze, body attitude, etc). Interaction is a well-researched area with good contributions in the domain of military C2 (Lambert 2000). In addition to speech as a ubiquitous interaction mechanism we are investigating the capabilities offered by handheld devices such as PDAs in iEcology. If the room is viewed as an intelligent device what capabilities does it present to a PDA for example? We are starting to look at how PDAs can use the iRoom as an extension of their utility, conversely how the room can make use of PDAs, and how the room can be controlled in a personalised way by a PDA.

4 The iRoom at DSTO

In this section, we introduce the iRoom and describe the initial implementation and applications that have been incorporated into the environment.

4.1 Introduction

The design of intelligent environments such as the MIT Intelligent Room and the DSTO iRoom have specific computational requirements that must be satisfied to support the types of interactions proposed in these environments. The environment consists of a potentially large number of computational resources in the form of hardware and software components. The effective use of the environment relies on the ability to manage the various computational resources. This is realized through the provision of capabilities such as managing the resource configuration, supporting collaboration between components, dynamic reconfiguration of the system and, from a practical standpoint, the ability to debug the operating environment. These requirements dictate that the system is built upon an infrastructure providing these various capabilities and further allows a relatively straightforward extension of the capabilities as the system is developed.

There are a number of 'middleware' systems available that support these capabilities to varying degrees, however the level of abstraction is generally such that development with these systems is at a relatively low-level. The MIT Metaglue system, upon which we have based our initial system developments, provides the computational glue to tie the various system components together within a framework that provides these desired capabilities. Metaglue is basically an extension of the Java language that allows the construction of interactive, distributed computations that form the basis of the intelligent environment. Specifically it provides the ability to interconnect and manage large numbers of computational resources, allows for the control of assemblies of interacting software agents, provides the ability to dynamically modify a running system without having to restart the system, manage allocation of resources to specific tasks and provides mechanisms for persisting state information.

4.2 Initial implementation

Our initial implementation of the iRoom has been built to support some relatively simple scenarios in order to explore some of the issues with the practical realization of an intelligent environment. The first scenario we have explored is simply to provide support to allow a user to make a presentation. In this scenario an individual indicates to the room their desire to make a presentation. The interaction with the system is via voice with all commands being spoken and the room responding appropriately during the interaction. The user first identifies themself to the environment and indicates their desire to make a presentation. The system configures the room suitably to make a presentation by marshalling and configuring the various resources required to support the presentation. This involves configuring the displays, modifying the room lighting and creating a virtual directory of the user's presentation files. The virtual directory consists of all resources that the user would normally use to store presentation files. The virtual directory is then displayed showing the most recent files at the top of the display hierarchy. The user can then indicate the file to be displayed at which point a suitable application to display the presentation is launched. Interaction with the display application mediated via the Metaglue framework through the agent associated with the display application. Voice commands are entered to the system using IBM's Via Voice that are interpreted using grammar files and subsequently sent to the application agent as a control message. The agent then controls the application through, in the case of the presentation application, the Windows COM interface. During the presentation queries can be posed relating to elements of the presentation. These queries are automatically passed to a suitable natural language query system such the MIT START system (MIT 1999) with the responses being displayed again in a suitable application.

Figures 1 and 2 show an example of this question-answering in action. Figure 1 shows the speech recognized by the dialog system, and Figure 2 shows START's response to the question.

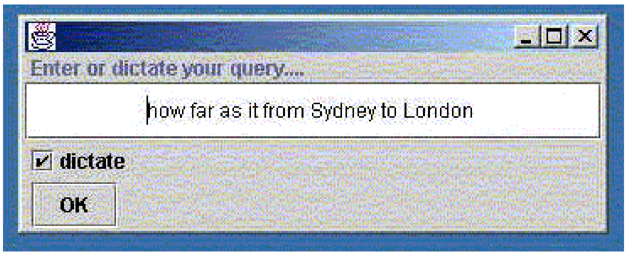

Figure 1 Dictated query

While this scenario is relatively simple it has allowed us to explore some of the issues with the design of Metaglue agents to deal with the various computational resources available within our environment. Metaglue had provided much of the required infrastructure to manage the

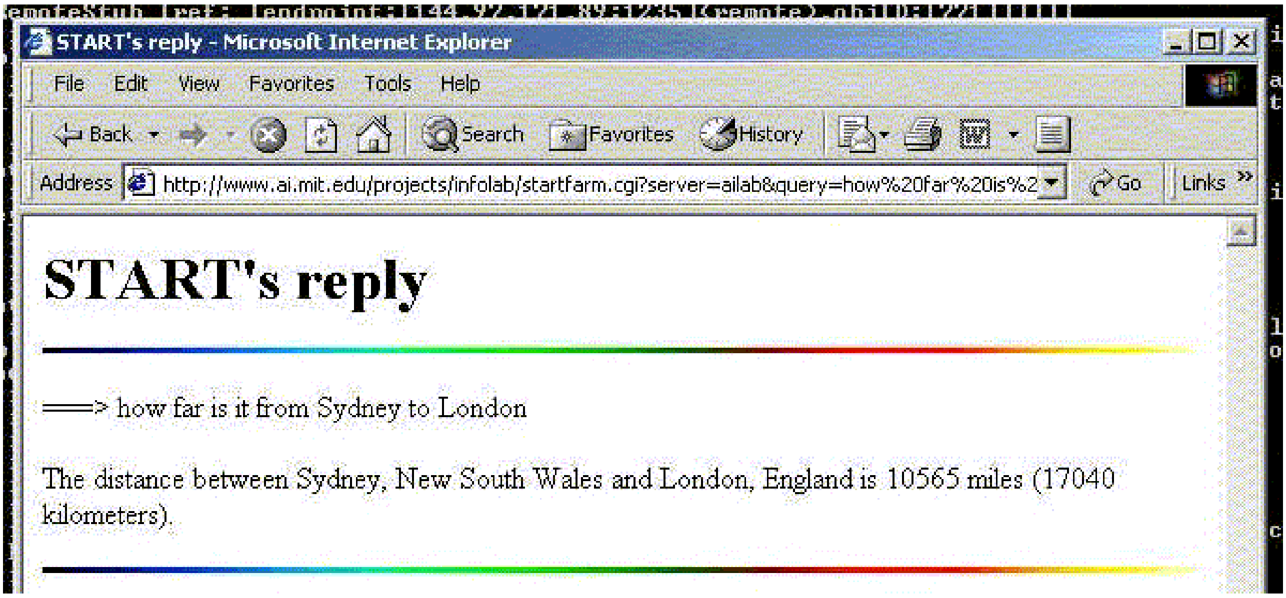

Figure 2 START's reply

computational resources in the iRoom. Incorporation of these resources into the framework is achieved through an extension of the Agent interface that provides much of the underlying plumbing to make any resource an integral part of the environment. Metaglue supports the concept of societies of agents which are logically related resources. This means that it is possible to dynamically group resources together to support various tasks with each of these resources having the ability to communicate via channels associated with the society. Figure 4 shows the current configuration of the iRoom, and for our simple scenario, shows the various agents, message channels and supporting objects involved in the interaction.

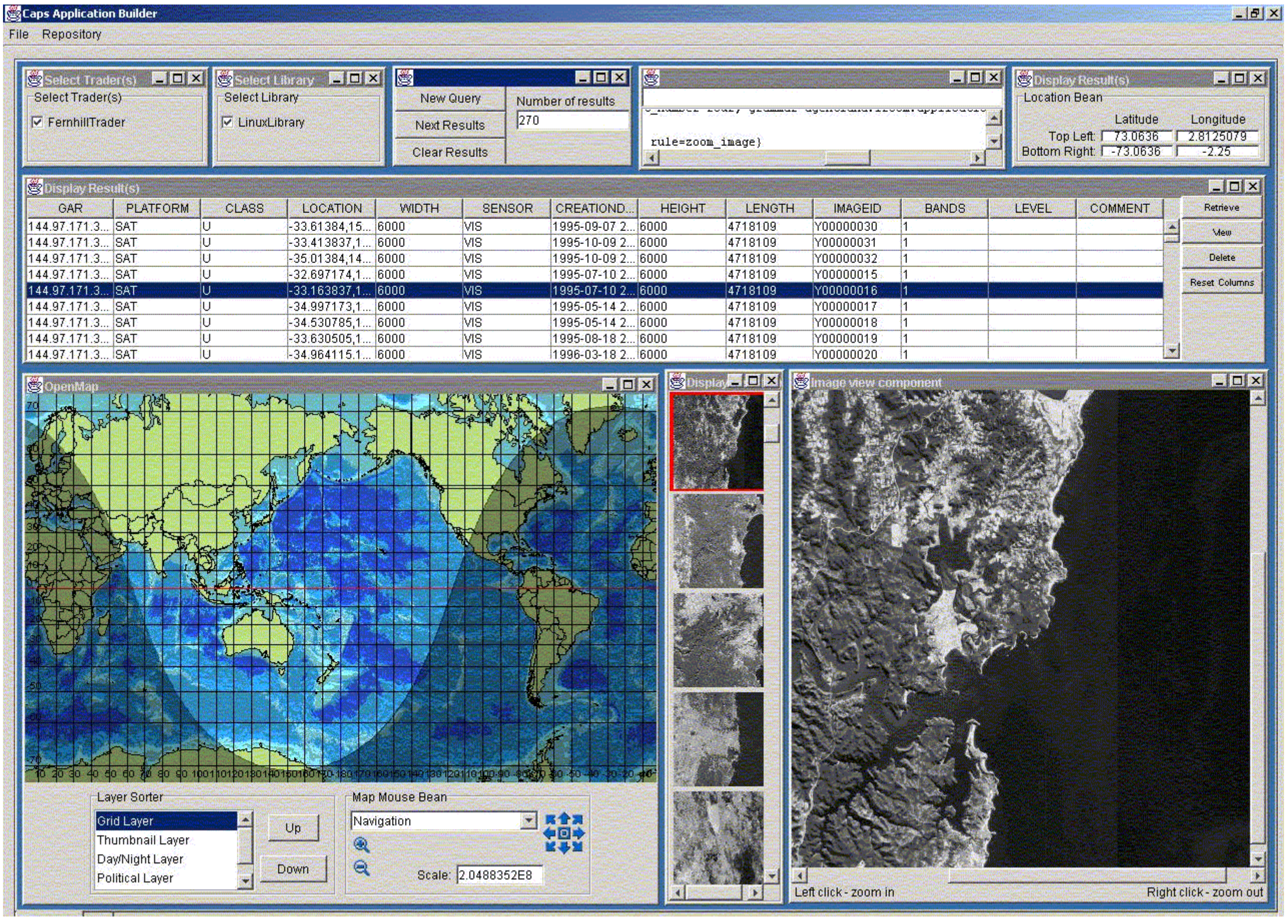

Figure 3 The iMedia application in the iRoom

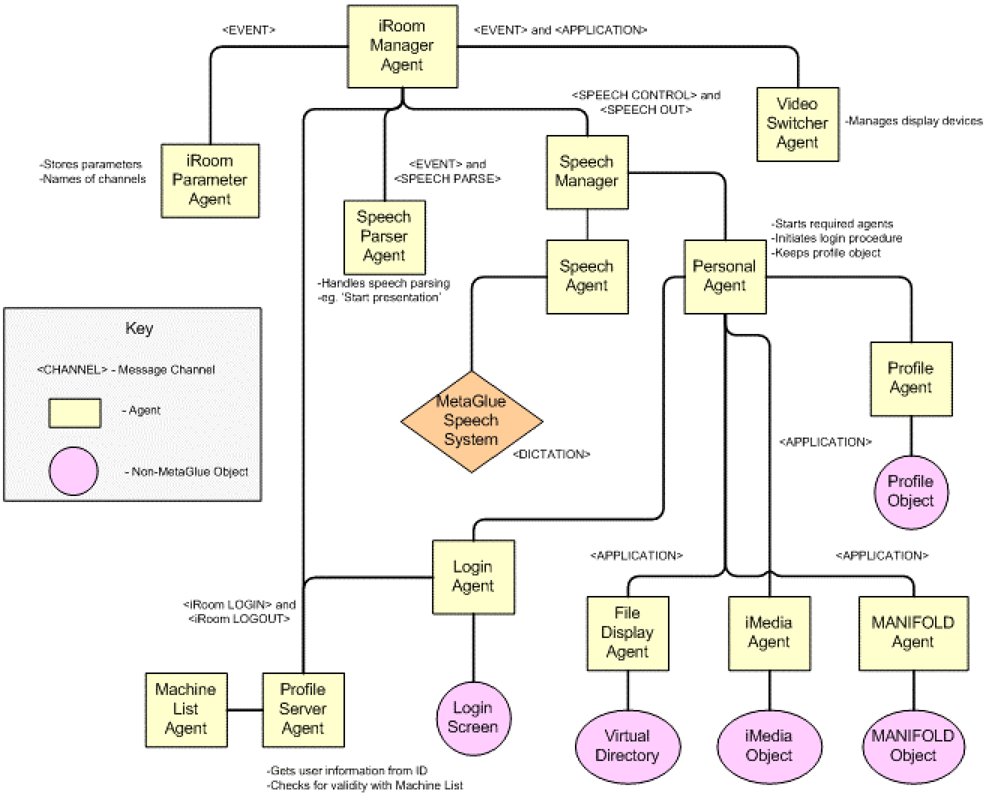

Figure 4 The iRoom configuration

4.3 Interoperability issues

We are now extending the work to support more complex scenarios such as the execution of a operation where we use the iRoom to provide information as the operation progresses. To this end we are incorporating some of the information resources which form part of an experimental command and control testbed that has been built to explore the issues associated with the use of service based architectures and their impact on future command support systems. The services we have incorporated are iMedia, (Grigg et al. 2000) which is a distributed media management system, and MANIFOLD, (Hawthorne et al. 2001) that allows for the access to heterogeneous data sources through a single interface. The incorporation of these resources has been achieved by essentially making their external interfaces extend the Agent interface. Grammars for each system have been designed allowing querying of each system through the voice Agents. Figure 3 shows the view of the iMedia system that has been integrated with the iRoom.

In the description above we have referred to the Metaglue system as a framework supporting the management of resources through Agent interfaces, where the implication is that the entities within the system are agents. In reality, the computational resources are presented through an interface which allows their management in a uniform manner. That is looking out towards the various resources they are seen by the infrastructure in a relatively uniform manner – a computer resource presents the same interface as a projector resource although obviously their capabilities are different. The interfaces to the various resources can then been viewed as agents for each of the resources.

4.4 Multi-agent system layers

From an information systems perspective the system we have developed currently can be viewed as essentially two layers. At the lower layer we have the various information services that provide information to support the various tasks within the iRoom. The next layer up is provided by Metaglue that provides the management framework for the various resources in the environment including the information services such as iMedia and MANIFOLD in addition to providing some support to user interaction through the voice agents and other input devices.

We are now adding a multi-agent system layer over Metaglue. The multi-agent system layer will allow us to explore the use of cooperating intelligent agents for the support of more intelligent interaction with the environment and more intelligent behaviours from the environment. For example the multi-agent system may have agents that act on behalf of the user or other systems to pre-emptively provide access to information resources as the scenario in the iRoom develops. We see this as essentially different to what the Metaglue agents provide as the multi-agent system can be used to execute goal directed behaviour.

4.5 Applications of the iRoom

In this section we describe at a number of supporting applications that have been built in the iRoom.

4.5.1 Black Box Recorder

The black box recorder imitates the functionality of an aircraft black box. It is intended to record every aspect of the use of the room. This includes:

- every word spoken
- any security violations
- every website visited
- identities of people in the room
- a track of people entering and exiting the room
- changes in the room security clearance level

The intention is to use the Black Box as a tool to aid in development of the iRoom, rather than an end-user application. Figure 5 shows a sample log of the black box.

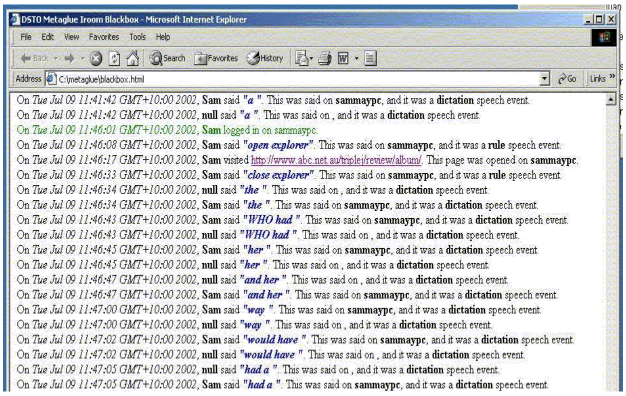

Figure 5 Black Box Recorder

4.5.2 Ambient Information Seeking

One research area being explored within the iRoom is the development of ambient information seeking applications, applications that automatically and pro-actively gather and display information relevant to the activities being performed within the room. For example, during a planning meeting, an ambient information seeking system would pro-actively gather information relevant to points being discussed by the meeting participants, without the participants actively directing the information search.

Key to developing ambient information seeking is building and maintaining an understanding of what activities are occurring within the room and what information may be relevant to them. One approach to building and maintaining this understanding is though the use of context (Dey and Abowd 1999), and in particular, the use of informational (Burnett and Chapman 2001) and activity-centred context (Prekop and Burnett In Press).

So far we have developed an initial Context Aware Information Manager application. This application essentially listens to what is said in the room and identifies trigger phrases that describe what information may be useful to the activities being performed in the room. These phrases are then passes to a context aware information retrieval engine (Burnett and Chapman 2001), which then surfaces links to information that may be relevant to the discussion taking place within the room.

One of the key difficulties with this approach is that natural conversations and discussions seldom follow a simple linear structure, (Robertson et al. 1996) but will jump from topic to topic, often with no clear relationship between them. To address this, we are beginning to devise ways of triangulate speech input (gathered by the Context Aware Information Manager) with other information, for example meeting agenda documents, to improve the quality of the trigger phrases identified.

4.5.3 Infrastructure for User Interaction

Within ubiquitous computing environments, such as the iRoom, existing user interface approaches designed for a single user interacting with a single desktop computer are not really suitable (Mark 1999). New interaction metaphors and combinations of interactions modalities (Robertson, et al. 1996) that account for the interactive nature of structure and complexity of the ubiquitous computing environment are needed.

Currently within the iRoom we are exploring different user interface approaches, and combinations of interaction modalities. We are using the concept of *pointing*, and pointing combined with speech, and other GUI interfaces elements as one method for interacting with the iRoom. We are using iPaqs as our pointing device. Each iPaq is fitted with an IR extender (which improves the range of the iPaq's IR port from a meter to several meters) that broadcasts the iPaq's unique ID code. Each fixed devices with the room is fitted with an IR receiver. To interact with the fixed devices within the iRoom, the user simply points his iPaq at the device, and clicks.

Depending on the type of interacting being supported, pointing and clicking on a fixed device may cause the download (via wired LAN connection) of a customised interface applet for the application running on the selected device. This applet allows direct control of the application. For example, if a large screen is running PowerPoint, pointing and clicking the iPaq at the screen will download a PowerPoint control applet onto the iPaq, which can be used to control the slide show. We currently have applets to support PowerPoint, Word, iMedia (a sophisticated defence image repository), and a generic mouse controller.

The iRoom is also a collaborative, interactive space, where users need to be able to interact with the room, each other, as well as with their own private information (Greenberg and Boyle 1998). The iPaqs used within the iRoom provides each user with a private information space, where information can be manipulated and created. When using an iPaq to select a fixed device, a user can move information from their private space, the iPaq, to the collaborative space, the room, and from the collaborative space to their private space. We are also exploring methods for tighter integration between the iPaq and the iRoom's devices to provide seamless movement of information between the two spaces (cf (Rekimoto 1997)).

A major limitation with our current limitation of point and click is that granularity of pointing is limited to a whole device. We are beginning to explore the use of other pointing devices, in particular laser pointers, to provide the ability to point to specific elements within an iRoom device (Myers et al. 2001).

4.5.4 Privacy and Security

The ubiquitous computing goal of invisible, intelligent, ambient computing requires the collection and processing of detailed information about an individual and their activities. Collecting this kind of information, often invisibly, raises important privacy issues; what rights does an individual have to this information, and what control do they have over what information is collected (Langheinrich 2001)?

To address these privacy issues, we have extended meta-glue to provide basic support user privacy. To use the iRoom, a user needs to first login to the system. Each logged in user has an associated user profile, which includes basic user preference information, as well as what rights they give to the room and to other users in the room to record their activities, and to access their information. This system provides users with the ability to opt-out of any information collection or information access activity they don't feel comfortable with, as well as the ability to see what information the iRoom collects. Visibility of collection, and the ability to opt-out is the corner stone of any privacy system (Langheinrich 2001).

Within the defence domain, information security is a major concern. Using the iRoom, we are beginning to identify what kinds of information security issues ubiquitous computing rases, as well as how ubiquitous computing may be used to enhance information security.

Currently the iRoom is able to determine the highest level of information classification that can be accessed or displayed within the room. The users profile information includes their *security clearance* a code that describes the highest level[2] of classified information they can access.

Based on the logged in user's security clearances and the current state of the room, the iRoom is able to determine the highest common security level of information that can be accessed or viewed in the room. When the state of the room changes, for example, the door is opened, or a public phone line is opened, or when someone enters or leaves the room, the room will reassess the highest common security level of information that can be accessed or viewed.

5 Conclusion

In this paper we have reported on the initial development of a ubiquitous computing environment, the DSTO iRoom. The infrastructure aspects are currently being addressed with the Metaglue multi-agent system, and a number of applications have already been built using this infrastructure. Higher-level agent systems are under consideration for future developments. We are looking to creating an environment to support the execution of military command and control scenarios through the provision of infrastructure components allowing seamless and context aware interaction between users and the environment. A number of diverse research issues are being addressed in this work relating to infrastructure development, how best to support human activities in the room and how people interact with the environment.

6 Acknowledgements

We'd like to acknowledge the development and programming support of the following IBL[3] (Industry Based Learning) students, Ken Scambler, Michael Gannon, Andrew Gilette and Sam Tilbrook, from Swinburne and University of Technology, Sydney.

7 References

BROWNE, R. (2000): C4I Defensive Infrastructure for Survivability Against Multi-mode Attacks. *Proc. 21st Century Military Communications Conference (IEEE MILCOM 2000).*

BURNETT, M. and CHAPMAN, C. (2001): Context-Aware Internet Access. *Proc. The 6th Australasian Document Computing Symposium*, Coffs Harbour, NSW, Australia.

COEN, M.H. (1998): Design Principles for Intelligent Environments. *Proc. Fifteenth National Conference on Artificial Intelligence (AAAI-98) & Tenth Conference on Innovative Applications of Artificial Intelligence*, Menlo, CA, USA.

DEY, A.K. and ABOWD, G.D. (1999): Towards a Better Understanding of Context and Context-Awareness. Technical Report GIT-GVU-99-22. College of Computing, Georgia Institute of Technology. Atlanta, Georgia, USA.

DEY, A.K., ABOWD, G.D. and SALBER, D. (1999): A Context-Based Infrastructure for Smart Environments. *Proc. 1st International Workshop on Managing Interactions in Smart Environments (MANSE '99)*, Dublin, Ireland., 114-128.

GEORGIA, T. (2000): Georgia Tech Aware Home, http://www.cc.gatech.edu/fce/ahri.

GREENBERG, S. and BOYLE, M. (1998): Moving Between Personal Devices and Public Displays. *Proc. Workshop on Handheld CSCW. ACM Conference on Computer Supported Cooperative Work.*

GRIGG, M.W., LUI, A.K., JAMES, S.P., OWEN, M.J. and LO, E.H.S. (2000): Distributed Imagery Library System Using Java and CORBA. *Proc. Evolve*, Sydney, Australia.

HAWTHORNE, P., COOMBER, G. and BURNETT, M. (2001): Delivery of Enriched Information to Deployed Forces. *Proc. 6th International Command and*

[2] Within the defence domain, security classifications are hierarchical codes that ranging from unclassified, through several increasing levels of classification, to the highest classification level.

[3] http://www.swin.edu.au/corporate/ili/ibl/current/about_ibl

Control Research and Technology Symposium, Annapolis, MD, USA.

LAMBERT, D. (1999): Ubiquitous Command and Control. *Proc. Information, Decision and Control Conference*, Adelaide, South Australia.

LAMBERT, D. (2000): FOCAL. *Australian DEFENCE SCIENCE* **8**(4).

LANGHEINRICH, M. (2001): Privacy by Design -- Principles of Privacy-Aware Ubiquitous Systems. *Proc. International Conference on Ubiquitous Computing Conference (Ubicomp2001),* Atlanta, Georgia, USA.

MARK, W. (1999): Turning Pervasive Computing into Mediated Spaces. *IBM Systems Journal* **38**(4):677-692.

MIT (1999): START, http://www.ai.mit.edu/projects/infolab/.

MIT (2000): MIT's Intelligent Room, www.ai.mit.edu/projects/iroom.

MIT (2002): Oxygen Project, http://oxygen.lcs.mit.edu/.

MYERS, B.A., PECK, C.H., NICHOLS, J., KONG, D. and MILLER, R. (2001): Interacting at a Distance Using Semantic Snarfing. *Proc. International Conference on Ubiquitous Computing Conference (Ubicomp2001)*, Atlanta, Georgia, USA.

NATO (1996): Annex B to MC Guidance for Defence Planning. MC-299/5. North Atlantic Treaty Organisation (NATO).

NIST (2001): NIST's Smart Space lab, http://www.nist.gov/smartspace/theLab.

PREKOP, P. and BURNETT, M. (In Press): Activities, Context and Ubiquitous Computing. *Journal of Personal Communication (Special issue on Ubiquitous Computing)*.

REKIMOTO, J. (1997): Pick-and-Drop: A direct Manipulation Technique for Multiple Computer Environments. *Proc. ACM Symposium on User interface Software and Technology*, Banff, Alberta, Canada.

ROBERTSON, S., WHARTON, C., ASHWORTH, C. and FRANZKE, M. (1996): Dual Device User Interface Design: PDAs and Interactive Television. *Proc. Conference on Human Factors in Computing Systems (CHI96)*, Vancouver, British Columbia, Canada.

SCHMIDT, A., BEIGL, M. and GELLERSEN, H. (1999): There is more to context than location. *Computers & Graphics* **23**(6):893-901.

STANFORD (2001): Stanford's iRoom project, http://graphics.stanford.edu/projects/iwork.