NEXUS: Mixed Reality Experiments with Embodied Intentional Agents

G.M.P. O'Hare[1], B.R. DUFFY[2], A.G. CAMPBELL[1]
[1] *Dept. Computer Science, University College Dublin, Ireland*
[2] *Media Lab Europe, Sugar House Lane, Bellevue, Dublin, Ireland*
[1]Gregory.OHare@ucd.ie, [2]brd@medialabeurope.org

Abstract *This paper seeks to erode the traditional boundaries that exist between the physical and the virtual world. It explores mixed reality experiences and the deployment of situated embodied agents that offer mediation in the control of, and interaction between, avatars. The NEXUS system is introduced which facilitates the construction and experimentation with mixed reality multi-character scenarios. The behaviour of such characters or avatars is governed by a BDI agent architecture that can effectively sense both the real and the virtual world overlay. Within this paper we describe the NEXUS infrastructure together with the technology set that envelops it. We outline the design and results from a series of mixed reality experiments that have been conducted using the NEXUS system.*

1 Introduction

This paper describes NEXUS, a framework that supports the fusion of the physical and the virtual, creating a single world in which people can interact with virtual entities in their own space. We envisage NEXUS as a place in which virtual and physical information spaces become seamlessly entwined.

Virtual Reality environments presents the possibility to immerse oneself in digital information yet allows the judicious maintenance of aspects of such real world features as visual and aural stimulus, and spatial awareness for example. Work which aims to fuse the real and the virtual may help our navigation of the often overwhelming volume of information and its complexity found in the digital information space.

The degree of interaction between these two spaces is not simply a matter of people controlling the form of a virtual space, or even one of artificial entities that control the form of a physical space. Rather, NEXUS is viewed as a fusion of physical and virtual spaces in which a myriad of entities, virtual and physical interact and manipulate the shared space, dismantling the traditional barrier that has existed between the real and the virtual.

This work strives to extend the context-dependent integration paradigm between virtual spaces and physical worlds as found in current augmented reality research by investigating how a seamless functional transition (rather than primarily perceptual) can be achieved. This is achieved through employing intentional agent technology to develop a seamless human computer interface between the real and the virtual.

Agent Chameleons [8] [15] [16] lays the foundations for mobile lightweight agents which may inhabit various embodied states. NEXUS offers an architecture that facilitates the creation of autonomous characters and mixed reality experiences which are mediated by intentional agents.

Two key issues are developed in this paper. Firstly, the extension of the common approach to deal primarily with our *perception* of mixed reality environments to encompass stronger *functional* features facilitated through such environments and secondly, the development of intentional multi-agent technologies in this field of research.

2 Related Work

The first hybrid ride-and-film attraction, the Cinerama Air-Balloon Panorama appeared more than a century ago in Paris [2] and demonstrated the mixing of reality and fantasy. Later, Ivan Sutherland [20], who developed Sketchpad at MIT, a CAD program in which the user drew directly on the screen using a light pen, is often viewed as having introduced virtual and augmented reality in 1962.

Augmented or mixed reality research to date has primarily focused on the integration of the virtual space (either 2D or 3D objects) with the physical world through an overlay through such technologies as see-through head-mounted displays. Such application domains include the use of augmented reality for face-to-face collaboration [13] [17], Virtual Round Table [4] and IVE: Interactive Video Environment [22]. These examples provide a clear example of the problem in facilitating the perception of spatial and temporal representations. Participant's perceptions of concepts being explained through teleconferencing can similarly be facilitated through such technologies, particularly when the communication medium is primarily display based [1].

The advantages of such strategies are clear with the provision of visually augmented information transfer to the user, uniquely tailored to the context of the physical environment. If not uniquely tailored, then it acts as a distraction from the real and consequently, the justification for its integration into our perception space becomes questionable.

Existing work in this field primarily aims to seamlessly merge our perception of the two realities into one coherent stimulus. Our interaction and our understandings of this interaction can however be developed further. The NEXUS project strives to extend this context-dependent integration paradigm between virtual spaces and physical worlds by investigating how a seamless *functional* transition (rather than primarily perceptual) can be achieved.

The realisation of evolvable characters in virtual environments can be seen in such work as Synthetic Characters at MIT-ML [3], and work on agents as synthetic characters, [7], [19].

The NEXUS proposal resonates with work undertaken within the Equator project which explores mixed reality environments and user experiences of such [5]. Mobile mixed reality as exemplified by the "Can You See Me Now" project [11] which has pushed the mixed reality envelope yet further.

Recent work has investigated the control of avatars and the mediation of mixed reality experiences through the use of agents. The need for multi-agent techniques in the control of multi user virtual worlds has been recognised. Nijholt and Hondrop demonstrated this via the use of Deep Matrix, a VRML based multi-user environment system but, however, they did not deploy multi-agents. More recent work has actively sought to use agents in this regard and is typified by the research of Huang et al [12] who have again used VRML as the 3D technology but with an extended BDI architecture [18] achieved through Distributed Logic Programming [10]. Torres et al [21] similarly adopt BDI agents in the control and management of autonomous virtual characters where they utilise AgentSpeak(L) as their BDI delivery medium.

3 The NEXUS Experience

The NEXUS experience aims to contribute two fundamental aspects of augmented reality research, that of developing the capability set of the avatar itself through the use of Belief-Desire-Intention agent control strategies, and to seek to develop the *functional* aspects of human-in-the-loop augmented realities. The objective is a functional seamlessness in how users can exploit augmented realities.

While it is argued that the ultimate goal is to create a system such that the user is unable to tell the difference between the real world and the virtual augmentation of it, this may in fact constrain the role of the VR component to a *only* use the real as a reference.

When creating a mixed-reality experience, the physical world is generally controlled in conjunction with the virtual. This raises issues of monitoring the depth of field issues of objects in the foreground, middle ground and background. It is arguable whether the quality of the illusion necessitates a naturalistic approach where real-world textures and object control should be employed. The advantage of using the virtual overlay and it seamlessly merging with reality through context should allow the integration of those features that are otherwise difficult or impossible in reality (i.e. no gravity, morphing, cloning). This applies to both the perception of this mixed reality scenario and the functional attributes available to the participants within it. Limitations on achieving real-world-like degrees of resolution in virtual artefacts embedded in real environments through mixed reality becomes less of an issue with augmented functionality capabilities. The computational demands of real-like augmentation systems become significant when aiming for smooth transitions and movement. If the functional aspects of the virtual components are developed to embrace those features of digital information spaces and computational devices, their role in mixed realities becomes more evident. If it's not real, don't make it real. Let us exploit its virtuality.

The virtual world is viewed as a mere extension (not replication) of the physical and conversely the physical is viewed as an extension of the virtual. By way of characterising these notions let us consider a few examples. Envisage a scenario whereby an avatar within a Collaborative Virtual Environment picks up a torch and shines it toward the user in the physical. Imagine that the virtual light were to continue in the form of illumination within the physical world. Conversely the physical user looking into the dimly lit virtual world may retrieve a torch and shine it into the virtual. The light experiment in Section 5 takes the first step in achieving this. The functional interface between the two realities should become seamless.

In contrast, consider several computer screens juxtaposed. Imagine virtual characters that apparently were aware of entities and events beyond their world. One avatar may well gesticulate to another avatar present on the adjacent screen. Avatars may also point at, or make reference to, objects contained in the physical world. In so doing, they exhibit environmental awareness beyond the immediate parochial periphery of the world they currently inhabit.

Such projects as Agent Chameleons [8] [15] [16] aim to develop autonomous digital agent assistants which can act like a ghost friend and move between embodied containers such as robots, virtual reality avatars and animated agents on desktops and PDA's. NEXUS extends the functionality of such an agent by developing the reference of the agent being inherently linked to our reality. For example, gestures of an avatar in a VR space are fundamentally referenced in our physical reality. Similarly, the motion of the agent across numerous screens is based on realising a sense of mobility in physical space. The screen where we see the avatar represents a window through which the avatar can interact with us, not uniquely a window through which we can view the virtual space as is generally understood.

NEXUS participants are able to engage in human-computer collaborative activities that bridge multiple diverse digital information spaces. By imbuing artificial entities engaged in this collaboration with knowledge of their user and the user's environment, we strive to improve the quality of experience offered to the user.

While the perceptual fusion of both real and virtual environments have been and continue to be investigated [9], it is not developed within this work. A number of off-the-shelf technologies are employed with the aim of facilitating the development of the NEXUS objectives. The following section introduces the NEXUS architecture which details the mechanisms for the experiments presented in Section 5.

4 The NEXUS Architecture

NEXUS constitutes an architecture, together with an associated methodology, which facilitates the design and delivery of experiments and experiences which fuse physical and virtual spaces into a coherent singular mixed reality employing agent based technologies. It embraces prior research conducted within the Agent Chameleons project [8] [15] [16]. This toolkit is comprised of a number of customisable hardware and software components that facilitate a rapid design and build cycle.

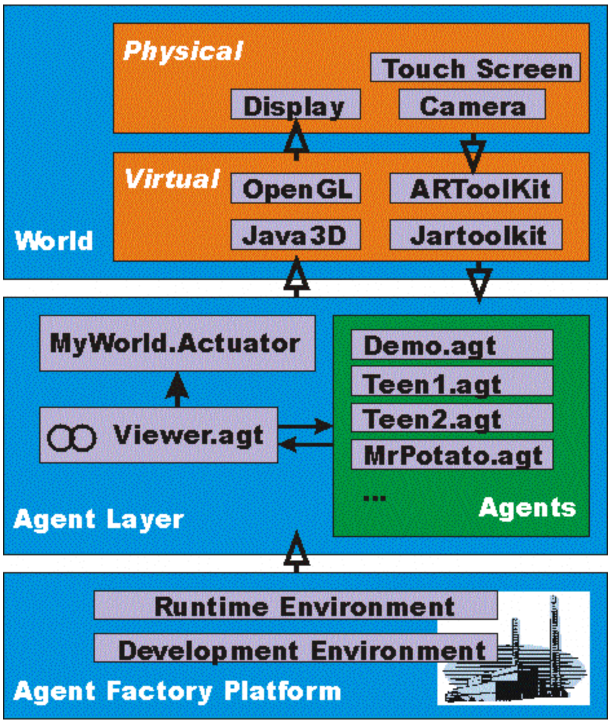

Fig 1: The NEXUS Architecture

The NEXUS Architecture is comprised of a stratification of technology layers. Figure 1 depicts this architecture and the three constituent layers. The underlying architecture consumes pre-existing of-the-shelf software systems with the innovation manifesting itself through the integration of these complex disparate software components.

The underlying agent apparatus is provided by Agent Factory [6] a cohesive framework that delivers structured support for the disciplined development and deployment of agent based systems. Two key components namely, the Agent Factory Development Environment and the Agent Factory Run Time Environment respectively support the design and debugging of agents and the subsequent execution of these agents. Agent Factory agent adhere to the broad class of Belief Desire Intention (BDI) agents.

The agent layer represents a collection of Agent Factory agents that have instantiated the generic agent template. We will illustrate this through two examples that are described in section 5. Residing on top of these agent technology layers is the world layer, which incorporates and integrates technologies for mixed reality experiences and various interaction modalities. Two substrata exist, that of the physical world and the virtual. The former supports the use of multiple juxtaposed touch sensitive screens, together with digital cameras, microphones, tracking/sensing technologies and micro head-up displays for augmented VR viewing. The latter supports the display of 3 Dimensional VR spaces. Again pre-existing open source software is consumed for these purposes. We use Java 3D API as a wrapper interface to the OpenGL API. We use Java 3D as it a natural way to interact with the Java Agent Factory agents. In addition Java 3D offers greater flexibility in terms of design rather than the more common place use of VRML and dynamic update via the External Authoring Interface (EAI). In order to support video input and the sensing of events in the physical world we utilise JARToolkit developed by C-Lab, Germany [23] which is in effect a Java wrapper allowing us to use the ARToolkit developed by Kato and Billinghurst [14]. Collectively this enables the following mixed reality experiments.

5 Two Mixed Reality Experiments

The Nexus Space is a physical space that is imbued with microphones, touch sensitive screens, speakers, cameras, and location sensing devices. The space contains a number of screens that comprise the various windows through which human and artificial entities are able to interact. These screens are augmented with CCD cameras that can be used to provide a looking-glass type of effect where required. Additional support for interaction is provided through the use of micro head up displays for augmented VR viewing.

In order to illustrate the interplay between these layers and to offer insights into the operation of the NEXUS architecture we now characterise two experiments.

5.1 Light Beam Experiment

The light beam experiment allows virtual beams of light to be originated within the physical world and continued into the virtual. This illusion is achieved via a touch sensitive screen. When the user touches the screen a virtual beam of light is shone originating from that point into the virtual world. Movement of the finger on the touch sensitive screen will reposition the light beam. In this experiment an autonomous character that of the friendly deer is intimidated by the beam of light (see figure 2). Its associated behaviours will result initially in it remaining static but subsequently, if or when, the light becomes too close it becomes frightened and moves away.

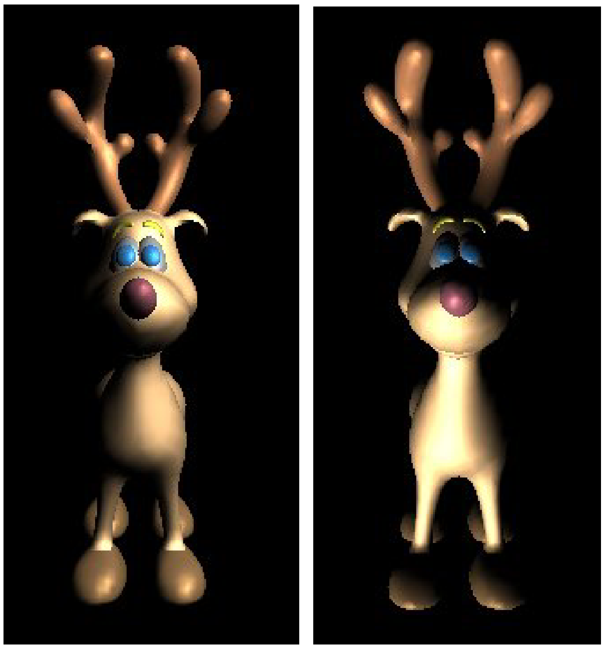

Fig 2: Showing light source moving from the upper right to the bottom left when a user touches the screen

The Agent Layer is populated with two agents. A *viewer agent* controls the virtual component of the world, while the *demo agent* informs it what to do with a 3D model representation of that agent within that particular world. The *viewer agent* runs its myWorld actuator to setup and control the virtual component within the world. The myWorld actuator also takes information from the touch screen; this information is then used to project a virtual light beam into the 3D world directed at the point of touch. This information is relayed back through the Viewer agent to the *demo agent* when it requests to know where it is within the world and the light's position.

If the light gets to close, the agent will request the viewers to move themselves away from the light. If it gets close to the edge of a given screen it will migrate to another viewer agent in another juxtaposed machine where possible.

The MyWorld Actuator calls Java 3D API, which acts as a high level interface for the OpenGL API. We use Java 3D as it a natural way to interact with Java-based AgentFactory agents. The AutoDesk *.3ds models are loaded into Java 3D using a loader developed by StarFireResearch [24].

The OpenGL world is thus projected out to the display but since it is also a touch screen it plays an active role in the program. We use a standard Touch screen developed by 3M. The information is thus fed back to the Agent Layer to project the light into the 3D world.

5.2 Hug Me Experiment

The second experiment is somewhat more complex involving a true mixed reality experience. The experiment involves a group of autonomous agents depicted in the form of avatars. Two avatar classes are depicted in fig. 3, the Mr Potato avatar and the teen avatar, with the latter being somewhat less rotund. Within this scenario, teens are attracted to each other and wish to hug each other. While they are somewhat concerned with any advances from Mr. Potato and take evasive action. The environment is comprised of a real world scene being acquired via a digital frame grabber. This image is then used as a backdrop for our mixed reality experiment and a serves as a canvas upon which the avatars are overlaid.

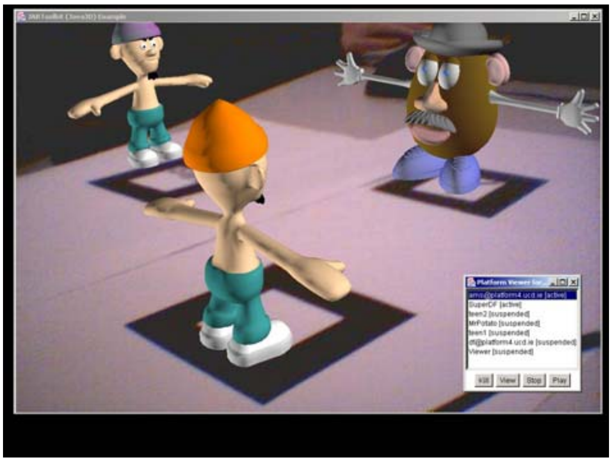

Fig 3: The NEXUS augmented reality agents.

In this demonstration four agents comprise the agent layer, the *viewer agent* again controls the Virtual component of World. The three remaining agents are based on a agent class entitled model Agent and are simply given different names. Based upon different markers/models which the viewer agent can perceive model avatars are mapped accordingly to these and projected accordingly on to the surface. The three *model agents* request the viewer to tell them what they can see. If the viewer tells them that the agent they can see is similar they move forward, if the agent is different they simply move backwards. As with the previous experiment we utilise Java 3D to produce the virtual component of our world, but in this example we take harvest of information about the physical world by way of a digital video image and using the JARToolkit developed by C-Lab [23]. The MyWorld Actuator takes in the camera images and performs a pattern match exercise interrogating them for the pre-defined makers. Using the ARToolkit it can subsequently place the 3D representations of the agents on top of the relevant marker. Furthermore it keeps note of what agents exist in is plane of vision so that agents can request this information when they want to know who is around them.

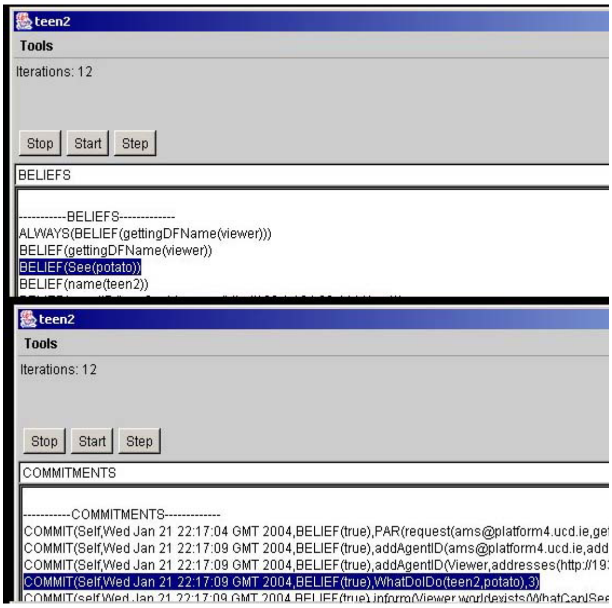

Fig 4: The beliefs and commitments leading to Teen2 agent's intention to move away from the Mr Potato agent

As before, the MyWorld Actuator calls the Java 3D API which in turn calls OpenGL. The JARToolkit calls the ARToolkit which captures the information from the camera. It searches the image for any markers and any that it finds it outputs as a translation matrix, which JARToolkit feeds back to the MyWorld Actuator. We have modified the original JARToolkit in two important respects. Firstly we ensure that it outputs the current image back to MyWorld Actuator and secondly that its behaviour now takes cognisance of the agents current position such as has it moved forward or back, and by how much.

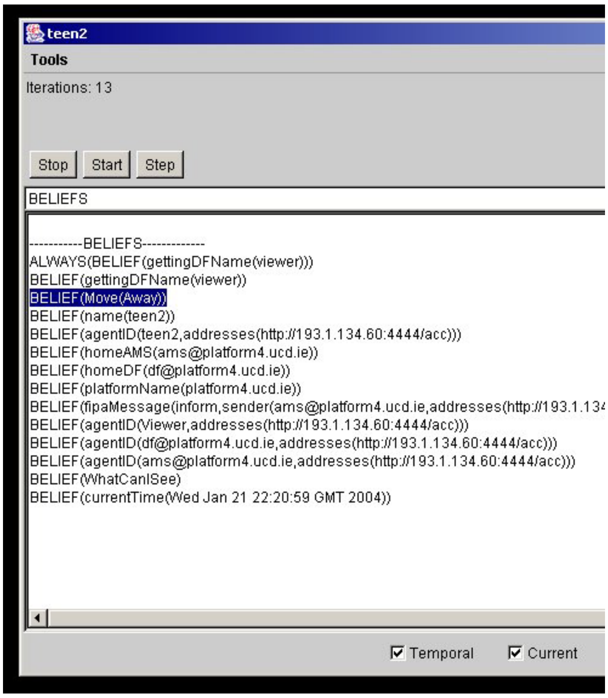

Fig 5: Teen2 belief set showing agent's actions to move_away from the Mr Potato agent

The OpenGL environment is displayed to a monitor. We use a simple USB web cam made by Logitech. It can operate up to 640 x 480 in 32-bit at 15 frames a second or 352 x 288 in 32-bit at 30 frames a second. The markers are simple A4 sized pages with makers printed in the centre. Using the ARToolkit, we can create markers from any pattern, but the patterns should be invertible so as to enable the viewer agent to track rotation and lock avatar orientation. Due to the nature of using patterns to mark position in the real world, lightening conditions are very important. Environment with a single light source and no shadows work best for detection. The Camera has a auto contrast ability and thus can work in varying light levels within reason.

The beliefs and commitments leading to Teen2 agent's intention to move away from the Mr Potato agent are illustrated in figure 4. Based upon the belief that Teen2 can see a Mr Potato agent it adopts a commitment to move away. This will result in the move away actuator being invoked and the viewer agent effecting a redisplay of the avatar in the virtual world. Figure 5 depicts the Teen2 belief set showing agent's actions to *move_away* from the Mr Potato agent, while figure 6 illustrates this movement across subsequent time frames.

Fig 6: Showing Teen2 agent moving away from the Mr Potato agent

These experiments are sufficient to illustrate how intentional agents may be incorporated and used as the control apparatus for autonomous characters. Rather than the avatars being merely empty vessels or containers they are viewed as membranes that embody a rich reasoning machinery that can perceive, reflect upon and subsequently effect the environment. The mental states illustrate how avatar behaviour is driven by the interplay between perceptions, beliefs and commitments.

6 Conclusions

Augmented reality provides a compelling innovative means of integrating the real and the digital in order to facilitate our accessing the digital information space. To achieve a successful design, the system must incorporate those features which facilitate rather than confuse, focus rather than distract.

The work presented here argues for the development of a more functional interface between the user(s) and the augmented reality through the use of intentional agent-based deliberative systems and mechanisms which provide for a more seamless integration of the real and the virtual. It does not advocate the design of mixed reality systems which seek to disguise the virtual through such instruments as very high resolution, high processing power and sophisticated rendering strategies. While undoubtedly useful, the inherent advantage of fusing two quite different worlds should be exploited rather than lost.

This paper has introduced NEXUS, a layered architecture that supports the rapid construction of mixed reality experiences. It encapsulates BDI agent machinery which may be used to manage autonomous virtual characters together with a rich collection of visualisation technologies.

Current work incorporates the use of Micro Head-Up displays to provide a richer augmented reality experience. Rationalisation of the agent designs will imbue each and every agent with environmental perception capabilities, rather than employing a single viewer agent.

Acknowledgements

The work undertaken as part of the Agent Chameleons project (http://chameleon.ucd.ie) is a collaborative project between the Department of Computer Science, University College Dublin (UCD) and Media Lab Europe (MLE), Dublin. We gratefully acknowledge the financial support of the Higher Education Authority (HEA) Ireland.

References

1. Billinghurst, M., “Real World Teleconferencing”, Hirokazu Kato, In proceedings of CHI ’99, Conference Companion, ACM, NY, 1999 (Pittsburgh, PA, USA, May 15-20)
2. Brandsford, W., “The Past Is No Illusion,” Digital Illusion, Entertaining the Future with High Technology, Addison-Wesley, Reading, Mass., 1997, pp. 49-57.
3. Blumberg, B. M. 1996. Old Tricks, New Dogs: Ethology and Interactive Creatures. Ph.D. Dissertation, MIT.
4. Broll, W., Meier, E., Schardt, T.: "The Virtual Round Table - A Collaborative Augmented Multi-User Environment". Proc. of ACM CVE 2000: The 3rd Inter’l Conf. On Collaborative Virtual Environments, E Churchill, M. Reddy (eds.), ACM, New York, 2000, pp39-46
5. Benford, S., Schnadelbach, H., Koleva, B., Gaver, B., Schmidt, A., Boucher, A., Steed, A., Anastasi R., Greenhalgh, C., Rodden, T., and Gellersen, H., “Sensible, sensable and desirable: a framework for designing physical interfaces”. Technical Report Equator-03-003, Equator, February 2003
6. Collier, R.W., “Agent factory: An environment for the engineering of agent-oriented applications”, Ph.D. Thesis, Department of Computer Science, University College Dublin, 2002
7. Doyle, P., and Hayes-Roth, B. 1998. Agents in annotated worlds. In Proceedings of the Second International Conference on Autonomous Agents. Minneapolis, MN: ACM Press.
8. Duffy, B.R., O’Hare, G.M.P., Martin, A.N., Bradley, J.F., Schön, B. “Agent Chameleons: Agent Minds and Bodies”, The 16th International Conference on Computer Animation and Social Agents - CASA 2003 Rutgers University, New-Brunswick, New Jersey, 7-9 May 2003.
9. Drascic, D., Milgram, P., “Perceptual Issues in Augmented Reality”, SPIE Volume 2653: Stereoscopic Displays and Virtual Reality Systems III, Editors: Mark T. Bolas, Scott S. Fisher, John O. Merritt, San Jose, California, USA, January - February 1996, pp 123-134
10. Eliens, A., DLP, A language for distributed logic programming, Wiley, 1992
11. Flintham, M., Anastasi, R., Benford, S., et al. "Where on-line meets on-the-streets: Experiences with mobile mixed reality games", CHI 2003, to appear.
12. Huang, Z., Eliens, A., Visser, C., “Programmability of Intelligent Agent Avatars” (Extended Abstract), Proceedings of the BNAIC2001, 2001.
13. Ishii, H., Underkoffler, J., Chak, D., Piper, B., Ben-Joseph, E., Yeung, L., Kanji, Z. Augmented Urban Planning Workbench: Overlaying Drawings, Physical Models and Digital Simulation, in Proceedings of Conference on IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR '02), (Darmstadt, Germany, Sept. 30 - October 1, 2002)
14. Kato, H., Billinghurst, M. “Marker Tracking and HMD Calibration for a video-based Augmented Reality Conferencing System”. In the Proceedings The 2nd International Workshop on Augmented Reality (IWAR 99). San Francisco, USA. October 1999
15. O'Hare, G.M.P., Duffy, B.R. “Agent Chameleons: Migration and Mutation within and between Real and Virtual Spaces”, The Society for the Study of Artificial Intelligence and the Simulation of Behaviour - AISB 2002, Imperial College, England, April 3-5, 2002.
16. O'Hare, G.M.P., Duffy, B.R., Bradley J.F., Martin, A.N., "Agent Chameleons: Moving Minds from Robots to Digital Information Spaces", 2nd International Symposium on Autonomous Minirobots for Research and

Edutainment (AMiRE), February 18-21, 2003, Brisbane, Australia
17. Ohshima, T., Sato, K., Yamamoto H., Tamura H., AR2 Hockey; A Case Study of Collaborative Augmented Reality. In Proceedings of VRAIS 98, 1998, IEEE Press, pp. 268-295.
18. Rao, A.S. and Georgeff, M.P., Modelling Rational Agents within a BDI Architecture, Prin. of Knowl. Rep. & Reas., San Mateo, CA., 1991
19. Reilly, W. S. 1996. Believable Social and Emotional Agents. Ph.D. Dissertation, CMU.
20. Sutherland, I., "Sketchpad: A Man-Machine Graphical Communication System," Proceedings of the Spring Joint Computer Conference, 1963, 329-346
21. Torres, J.A., Nedel, L.P., Bordini, R.H., "Using the BDI architecture to produce autonomous characters in virtual worlds", 4th International Working Conference on Intelligent Virtual Agents (IVA) 2003, 15-17 September 2003 at Kloster Irsee, Germany.
22. Wren, C.R., Sparacino, F., et al, "Perceptive spaces for performance and entertainment: Unthethered interaction using computer vision and audition", Applied artificial intelligence, Vol.11, No.4, pp267-284, June 1997
23. http://www.c-lab.de/jartoolkit/
24. http://www.starfireresearch.com/services/java3d/inspector3ds.html