Introducing Third Party Objects into the Spatial Model of Interaction

Steve Benford and Chris Greenhalgh
Department of Computer Science, The University of Nottingham, UK

We introduce an extension to the spatial model of interaction for CVEs called third party objects that provides support for contextual factors in awareness calculations and that enhances scaleability. Third parties can have two effects on awareness: attenuation or amplification of existing awareness relationships; and the introduction of new aggregate awareness relationships We propose a range of applications for third party objects including: world structuring regions, aggregate views, common foci, representational and group services, and dynamic load management. We present an implementation, the MASSIVE-2 system, focusing on its network architecture of a dynamic and self-configuring hierarchy of multicast groups

1. Introduction and background

This paper extends the spatial model of interaction for Collaborative Virtual Environments (CVEs) as defined in (Benford, 1993) to include support for contextual factors in awareness manipulation and also greater scaleability. This is achieved through the introduction of third party objects, independent objects in a virtual world that can perform various awareness adaptations and aggregations. Third parties can be used to model a wide range of awareness manipulation mechanisms ranging from nested spatial structures such as rooms within buildings within regions, objects that act as common foci for interaction, through to mobile crowd aggregations. The paper also describes how this extended spatial model can be mapped onto an underlying network architecture based on a dynamic hierarchy of multicast groups. This mapping lies at the heart of the MASSIVE-2 system, a more scaleable and flexible successor to the MASSIVE system as described at ECSCW'95 (Greenhalgh 1995). The paper describes the implementation of MASSIVE-2 and mentions several example applications including an arena where events can be staged in front of crowds of participants; a poetry performance and

J Hughes et al (eds), Proceedings of the Fifth European Conference on Computer Supported Cooperative Work, 189-204

exhibition and a collaborative software visualisation tool.

However, before introducing the concept of third party objects, we first consider the background to this work in terms of the history of development of the spatial model, the current limitations of the model and also related work on spatially structured distributed virtual environments and multicast network architectures.

1.1. A history of the spatial model and its current limitations

The spatial model was first proposed as a way of managing awareness in CVEs through the mechanisms of aura, awareness, focus, nimbus and adapters (Benford, 1993) and was initially demonstrated in the DIVE system and then first fully implemented as the MASSIVE system (Greenhalgh 1995). MASSIVE has since been used to hold many virtual meetings and several papers have emerged from these experiences including studies of interaction within (Bowers, 1996a) and around (Bowers, 1996b) MASSIVE meetings and also mathematical analysis of system log data from an extended series of meetings (Greenhalgh, 1997). Recently, at CSCW'96, Tom Rodden proposed a generalisation of the spatial model to arbitrary "network" structured spaces that might map onto a range of CSCW applications (Rodden, 1996).

Having developed, implemented and extensively tested spatial model based CVEs, we have become aware of two major limitations:

Contextual factors in awareness - the original spatial model concepts of focus and nimbus (see section 2) were concerned with how an observer and observed could control the former's awareness of the latter. Beyond a limited notion of adapters (see below), there was no truly general consideration of how contextual (i.e. environmental) effects on awareness could be realised (e.g. the effects of boundaries in multiple nested spaces, crowds of participants or shared objects).

Scale - although the initial spatial model support scaleability in terms of awareness driven presentation of large volumes of information to an individual and the use of aura to limit connectivity, it still necessitated dealing with individual awareness relationships or aura collisions on a bilateral basis. This represented a major bottleneck to implementing CVEs that could support hundreds or thousands of simultaneous users. What was needed for some more fundamentally scaleable mechanism as a core part of the model.

The concept of third party objects, a much extended notion of the adapters in the original model, is intended to address both of these problems. Third party objects were first introduced in (Benford, 1997) which focused on their use for supporting dynamic crowds of participants in CVEs. This paper extends this work in two main ways. First, in section 3, it explores a broader range of possible applications of the third party concept. Second, in section 4, it describes how the extended spatial model has been implemented as the MASSIVE-2 CVE system with a particular focus on the mapping of the model onto an underlying multicast network architecture.

1.2. Related work on spatial partitioning for CVEs

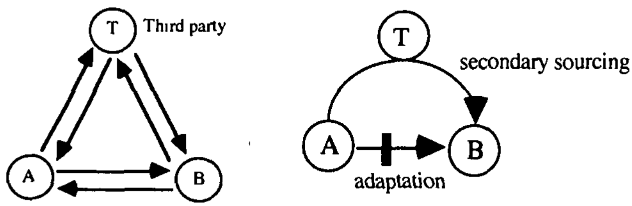
The concept of third party objects also relates directly to other work in the field of distributed virtual environments. In particular, several researchers have been considering how to introduce some form of spatial partitioning into shared virtual worlds (typically involving notions of zones or regions) and then map this onto a underlying network architecture in an efficient way. For example, NPSNET (Macedonia, 1995) tiles the world with hexagonal cells, each with its own multicast group, so that observers need consider only near-by cells. The Spline system (Barrus, 1996) composes the world from arbitrarily shaped regions or "locales" that localise interaction and that may be "stitched" together by arbitrary 3D transformations. In a slightly different vein, RING (Funkhouser, 1996) scopes interaction and communication according to potential visibility in densely occluded environments (e.g. within buildings) while localising interactions at servers responsible for specific regions of the world. Other researchers have focused on alternative ways of exploiting multicast protocols, including the association of multicast hierarchies with the hierarchical composition of objects and collections of objects in a virtual world as implemented in recent versions of the DIVE system (Hagsand, 1996). The work described in this paper is intended to be sufficiently general so as to subsume previous spatial partitioning approaches. At the same time, it offers an interesting contrast to the strict "object hierarchy" approach. In the following section, we begin our discussion with a general introduction to the third party concept.

2. Third party objects and the spatial model

In order to understand third party objects, one must first be familiar with the original spatial model. The spatial model defines mechanisms for the spatial management of awareness and communication in shared virtual spaces such that the quality of one object's perception of another in a given medium is represented as the quantifiable concept of awareness. In turn, this awareness is enabled through collisions between auras and then negotiated through two further spatial fields, focus (observer's control) and nimbus (observed's control). Auras, foci and nimbi are defined as general spatial (and possibly non-spatial) fields that can have arbitrary shape and extent and can assume different values across their extent. They are also medium specific. The further concept of adapters involves the manipulation of awareness, focus, nimbus and aura as a result of interacting with other objects in the space and can be used to model simple spatial boundaries and communication aids such as podia and conference tables. However, the original adapters could only perform limited awareness manipulations (i.e. replacing focus, nimbus and aura) and there was no general definition of how they were activated and whether they might operate recursively (e.g. they could not be used to model nested boundaries within boundaries as one might expect to find in a large virtual building).

2.1. The concept of third party objects

We now introduce the concept of third party objects as an extension to the spatial model (Benford, 1997). A third party object is an independent object that affects the awareness between other objects (see figure 1 (a)). Two points should be noted about third party objects from the outset. First, as with focus and nimbus, all aspects of their operation as described below may be medium specific (e.g. they may operate differently in the audio medium than in the graphical, textual or video media). Second, as they are objects in their own right they might be embodied, have a configurable spatial extent, be mobile or fixed in the world, be dynamically or statically created and might apply their effects recursively to one another. This potential for recursive application is a key feature that enables the construction of complex nested virtual world structures and has profound implications for the network architecture of any implementation as we shall see later. Their ability to be mobile and resizeable is also important as this enables support for dynamic crowds in CVEs.



(a) third parties (b) effects of third parties

Figure 1: Third Party Objects

There are two key aspects to the definition of third party objects: their effects (i.e. what they do to awareness relationships between other objects) and their activation (i.e. when and how these effects are brought into operation).

2.2. The effects of third party objects

Third party objects can have two general kinds of effect on awareness and these may be applied in different combinations (see figure 1 (b)).

Adaptation involves the manipulation of existing awareness relationships between objects. This includes attenuation (e.g. a third party acting as a barrier between objects) and amplification (e.g. increasing awareness between people who are accessing a common object).

Secondary sourcing involves the introduction of new indirect awareness relationships between objects in order to enable new transformed flows of information between them. Typically, secondary sourcing involves the consumption of information from an external object or group of objects, its transformation in some way and its subsequent re-transmission in order to provide an alternative view of the object or group (e.g. generating a single aggregate view of a crowd of people, a room or a region when seen from a distance). Various filters may also be applied at different stages of this process in order to reduce level of detail or to select key information. At the heart of secondary sourcing for groups lies the problem of creating a single aggregate view or stream of information from a number of sources. In general, there are three approaches to the aggregation problem: *selection* - switching between individual views or streams in some way (e.g. round robin, "loudest wins" etc.); *combination* - the direct composition of a new view from existing views (e.g. tiling multiple video windows); and *abstraction* - generating an entirely new representation based on statistical information describing the sources (e.g. mapping number of sources into size of representation, level of activity into colour etc.).

The combination of adaptation and secondary sourcing is the main reason why third party objects support scaleability in CVEs. Under appropriate circumstances, a third party can replace many individual awareness relationships with a single secondary source, thereby reducing the amount of information being transmitted across the network and processed by individual computers.

2.3. The activation of third party objects

Next we consider the circumstances under which different combinations of these effects are applied. The activation of a third party object is based on the awareness relationships between the third party and the other objects involved. Thus, referring to figure 1 (a), the activation of T depends on four possible awareness relationships: T's awareness of A and B respectively and their awarenesses of it. In figure 2 we identify three particularly interesting cases from among the various possible combinations of these awareness relationships.

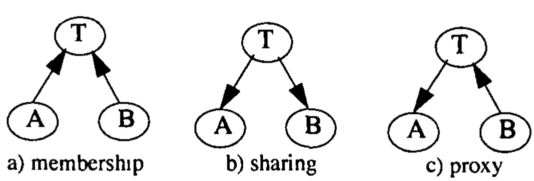


Figure 2: Activating third party objects

Membership - cases where the third party is activated according to how aware it is of other objects.The third party's awareness of an object represents and expresses the degree of membership that object has of the third party. For example, one might become a member of a room or region by crossing its boundary, resulting in various changes in awareness (see section 3.1).

Sharing - cases where the third party is activated according to how aware other objects are of it. The third party is, in some sense, shared by the objects which, in turn, influences their mutual awareness. For example, two people focusing on a common data item in a 3-D information visualisation or working with the same 3-D design might experience an enhanced awareness of each other caused by the object of common interest (see section 3.3).

Proxy - cases where the effects of the third party depend on how much one object is aware of it and how much it is aware of another object. The third party consumes information from its members and makes it available to external observers as an aggregate view at a low level of awareness. At higher awareness levels, this view is unfolded, to be replaced by information from each of the individual member objects. This turns out to be a useful case for crowds as we shall see in the following section.

This concludes our introduction to the general mechanism of third party objects in the spatial model of interaction. The following section explores a wide range of possible applications of this concept.

3. Possible applications of third party objects

We propose that the third party objects can be put to a wide range of uses, including:

- **World structuring and regions** - including nested spatial structures such as virtual rooms, buildings and zones that introduce awareness manipulations and aggregations across their boundaries.
- **Aggregate views** - these include dynamic and mobile crowds of participants and data districts in three dimensional visualisations.
- **Common foci** - objects that, when shared by a number of people, affect the mutual awareness between them (e.g. working with shared 3-D designs or chance encounters in a 3-D information space).
- **Representational and group services** - embodying general services such as support for anonymity, non-local communication, subjective presentation, brokering and introduction, group navigation and floor control.
- **Load management** - the automatic creation and destruction of third party objects by the system to cope with computational and network load.

We now consider each of these in turn, with a particular focus on the first two.

3.1. World structuring through regions

One of the key uses of third party objects might be in partitioning shared virtual worlds into regions that apply different awareness effects across their boundaries. A typical region would be a membership activated third party as in Figure 2, case (a) above, where membership would be determined by the act of crossing its boundary. In terms of its effects, awareness relationships between members would remain as normal (i.e. subject to individual foci and nimbi). However, awareness relationships between members and non-members might be subject to a combination of adaptation (typically, attenuation) and secondary sourcing (typically, the region might offer an abstract view of its contents to those outside of it). A virtual world might then be divided into an arbitrary number of such regions of varying size and shape.

This basic structure might be extended in several ways. First, the effects of a region might be applied differently across different media, resulting in the kinds of boundaries proposed in earlier spatial model papers (e.g. "windows" that would be permeable to visual media and impermeable to audio, or "curtains" that would have the reverse effect). Second, regions might apply different combinations of effects for member to non-member awareness that for non-member to member awareness. In other words, the boundaries of regions could be defined to be directional (e.g. "one way mirrors"). Third, and perhaps most importantly, regions could be nested, allowing for the definition of arbitrarily complex virtual world structures such as virtual cities, divided into districts, divided into buildings, divided into floors, divided into rooms, all applying different awareness manipulations across their boundaries.

An example of regions is provided by the recent poetry performance that we staged using our MASSIVE-2 implementation of the spatial model; see section 4 for an overview of MASSIVE-2 and (Benford, Greenhalgh, Snowdon and Bullock, 1997) for a description of the event. The performance world was structured as a top level region that contained a stage area where poets performed to a virtual audience. This region contained four sub regions, each of which contained graphic designs and fragments of poetry that could be explored during breaks in the poetry. These regions incorporated one-way boundaries such that those inside could see and hear what was happening in the stage area, but those outside couldn't see or hear what was happening inside, thus providing semi-private zones for social interaction.

This application of third parties can also be related to other work in CVEs. Several existing CVEs support multiple disjoint virtual worlds linked by portals (e.g. DIVE and MASSIVE). These can be considered to be simple regions, although their only effect is to completely attenuate awareness across their boundaries. Other recent work has focused on spatial partitioning in virtual worlds. For example, the tiled cells of NPSNET correspond to simple regional third parties that have no adaptive effect, but that group their contents to give a unit of visibility or interest that can be mapped onto its own multicast group. Similarly, the localisation aspect of a locale in the Spline system also corresponds directly to a single region (although the spatial model does not attempt to address the issue of multiply-related coordinate systems as used in Spline). Scoping on potential visibility as in RING may be achieved by mapping rooms (or other basic open spaces) to regions, and either allowing each client to reason about what it might see according to a locally and hierarchically structured abstraction of the total environment, or by employing another third-party object (or objects) in a server-type role to determine inter-region visibility. Finally, regions need not be solely spatial in their determination, and regions with a non-spatial element (specifically "is-a-sub-object-of") can be used to realise object hierarchies, including the provision of different multicast groups for different sub-trees of the object hierarchy as prototyped in DIVE.

3.2. Aggregate views - crowds and data districts

Another application of regions is to define abstract views of dynamic collections of objects. The primary difference to regions is that these third parties might be mobile within a virtual world. We focus on two examples: crowds and data districts.

One of our first uses of third party objects has been to explicitly support crowds of participants in CVEs. This work is described in detail in (Benford, 1997). A crowd is a third party object that represents and manages a potentially large grouping of participants in a virtual world (e.g. audiences at virtual events or dynamic groupings of people in densely populated spaces). From the "outside" (e.g.when not a member or when perceived at low awareness) a crowd provides an aggregate image of its members across a range of relevant media. Typically, this aggregate would be generated dynamically and would provide information about the composition and current activity of the crowd. (Benford, 1997) proposes a range of crowd aggregation techniques for different media, including graphical visualisation of crowd statistics and the triggering of different media samples (e.g. audio and video fragments or texture mapped images). On the inside of the crowd, interaction with individual members becomes possible. Crowds are typically asymmetric; from the inside one can usually perceive the outside world in full detail (or even in amplified detail as in a crowd that amplifies awareness of a performer on a nearby virtual stage). Crowds may also be mobile (e.g. may follow or even pull their members around) or may be static (e.g. be attached to the virtual scenery in some way such as a bank of seating in a stadium).

We have implemented an application called the Arena to demonstrate crowds in the MASSIVE-2 system. The Arena provides a virtual world in which several participants may play a game of "Pong" (the classic computer tennis game) in front of a watching audience. Audience members are located in different crowd regions such that, those on the inside of a given crowd have normal awareness of one another, but those outside (e.g. the players or the members of other crowds) perceive an aggregate view. We have initially used two simple aggregation algorithms: a single graphical crowd body that grows in size according to the number of its members and the mixing together of audio streams coupled with the application of a low pass filter to produce a composite crowd rumble in the audio medium.

Outside of the Arena, we locate a number of mobile crowds. These are crowd objects (i.e. third parties) whose position in the world becomes the average position of their current members. People exiting the Arena typically fall into one of these crowds that then follows them and their neighbours around. These mobile crowds can be used to smooth the transitions of people around a world and to dampen down system load caused by many people moving at the same time. In order to test our application, we have also developed some simple "agents" that can be scripted to move through the world and that emit different utterances as they go.

Our second example of aggregate views is that of data districts. A data district corresponds to a dynamic grouping of information items in a three dimensional visualisation and the corresponding generation of an aggregate view that conveys the properties of the group as a whole. In addition, the aggregate might also convey summary information about the current users of the data items (e.g. might show how many other users are currently inside the data district). Our first example, implemented in MASSIVE-2, is a collaborative software visualisation. We have developed a three-dimensional hierarchical visualisation to show different relationships between the objects and classes in an object oriented software system (e.g. "inheritance", "using" and "aggregation" relationships). Collections of objects (e.g. class libraries or modules) can be defined as data districts whose detail is only expanded for those who are currently inside them or close to them. We are currently exploring other possibilities such as the use of "legibility techniques" applied to three-dimensional representations of the WWW (Ingram, 1996) to develop data districts of inter-linked WWW pages and the use of data-districts to provide aggregate views of relations (i.e. schema visualisations) in a relational database visualisation.

3.3. Common foci

Common foci are third party objects that affect the awareness between people who were using them. Unlike the above, they are activated according to how aware the participants are of the third party (case (b) from section 2.3). Thus, if two people both focus on the same object in the virtual world, then this object might have some effect upon their mutual awareness. Typically, this effect would involve amplification of their mutual awareness.

There are many possible examples of common foci. People working with a common 3-D design (or a common part of a large 3-D design), might experience heightened mutual awareness, perhaps resulting in the creation of an audio channel between them. People accessing the same data item in a visualisation might experience a similar effect. Thus, two people browsing the same page in a WWW visualisation might become aware of one another, or two software engineers who elect to modify the same component in our MASSIVE-2 software visualisation might automatically be put in contact with one another in order to resolve possible conflicts.

3.4. Representational and group services

Third parties might be used to embody and activate a range of other general services. Third party objects might provide transformed views of another object or objects to support, for example: anonymity - by concealing the identity of the indirectly viewed object; non-location communication - by generating the secondary source information at some distance from the object(s) being observer; or subjective views - via a personal third-party that provides a transformed view of the world. Third party objects might influence navigation, perhaps acting as group vehicles. On activating such a vehicle, one would be taken on a journey and would experience heightened awareness of the other passengers. As a final example, third parties might be used to implement more traditional floor control policies in group settings. In this case, one would loose direct awareness of others (adaptation) and this would be replaced by a secondary source channel where the aggregation algorithm implemented the floor control algorithm, switching between different participants (e.g. round robin, token passing or loudest wins).

3.5. Load management

Our final application of third parties operates at the system rather than at the application level. The scaleability offered by the replacement of individual awareness relationships with aggregate secondary sources might be exploited by the underlying system as a way of dynamically managing system load. Thus, when computational load or network traffic exceeds some threshold, the system might automatically introduce third party objects into the world, essentially re-grouping participants on the fly. The mobile crowds from the Arena application represent a pre-planned version of this. More generally, techniques have recently been proposed for pre-determining "social hot-spots" in virtual urban structures (Ingram, 1996). These might be used to automatically introduce third party objects at places in a virtual world where the population density might be expected to be high.

4. Implementation - MASSIVE-2

Having considered the definition and application of the third party objects, we now present one possible implementation of them in the form of the MASSIVE-2 system. MASSIVE-2 is, as its name implies, a successor to MASSIVE as described at ECSCW'95. Like MASSIVE, it provides a collaborative virtual environment within which a number of mobile and embodied participants can interact over a combination of graphical, audio and textual media, controlled by the spatial model. There are many important differences between MASSIVE-2 and MASSIVE, including MASSIVE-2's support for full interaction with the contents of the virtual world and provision of a general application development environment.

From the point of view of this paper, probably the most important difference between them is in their underlying network architectures. Whereas MASSIVE was based on a peer-to-peer unicast network architecture, MASSIVE-2 is based on a novel multicast network architecture. It is the mapping between the extended spatial model and the underlying multicast architecture that we explore in the remainder of this section. There are two main aspects to this mapping. First, enabling the transmission of information (e.g. state descriptions and state updates) between artefacts and observers in a virtual world. This is achieved through the introduction of a dynamic hierarchy of group objects, directly derived from the nested spatial structure of the world (as defined by third party objects), where each group provides one or more multicast addresses for distributing state and update messages for different media. Mechanisms are defined to enable transmitting artefacts to join and leave different groups as both move about the world, and for receiving observers to "page in" different groups as they move about the world. Second, the objects that are exchanging information in this way negotiate mutual awareness levels through additional protocol support for focus, nimbus and adaptation. Awareness is then used to tailor the presentation of the information received from the multicast groups (e.g. to control the volume of audio, quality of video or level of detail of graphics). We now examine each of these two aspects in more detail.

4.1. Spatially managed multicasting

As we saw above, the concept of third party objects allows us to define nested and evolving structures for virtual worlds through objects such as regions, bounded rooms and crowds, where these structures have direct effects on mutual awareness and information exchange. In MASSIVE-2, this spatial structure is mapped directly onto a hierarchy of multicast groups. The joining and leaving of these groups is automatically managed according to the movement of objects and structures in the virtual world. The aggregation effect of third party objects is supported by extending multicast groups to provide aggregate transmissions of their contents to their superiors in the hierarchy. There are four fundamental types of component involved in our architecture:

World - a single nominally infinite three-dimensional virtual space within which the other components can be situated. A world includes a configurable specification of media supported (e.g. text, audio, video, graphics) and also a mapping of these media onto multicast groups (e.g. all media sharing a single group, one group per medium or something in between).

Artefact - an object represented within the world that can be decomposed into medium-specific sets of attributes and streams of messages. An artefact is essentially a transmitter of information in the world.

Aura - is a device through which inhabitants perceive the world and takes the form of a scope or area of interest. Auras are defined by a world, location, spatial extent and set of relevant media. An aura is a receiver of information in the world. Thus, in order to participate in the world, a user will typically be represented by one or more artefacts and will observe the world through one or more auras.

Group - is a representation of a third party object from the spatial model. Groups are structuring tools for virtual space that introduce awareness adaptations and aggregations and that may contain one or multicast addresses so as to transmit information between artefacts and auras that are associated with the group. There are two separate classes of group membership: sending members (artefacts) and receiving members (auras). MASSIVE-2 groups are therefore a generalisation of multicast groups, in that they define their own membership mechanisms, support aggregation and may map onto several underlying multicast addresses.

We now discuss how these components work together, focusing on the mechanisms defined for managing artefact and aura membership of groups.

4.2. Groups and artefacts

Figure 3 shows how an example virtual world maps onto an underlying hierarchy of group objects. In our example, the virtual world is structured into an outer room that contains a further inner room. There are two dynamic crowds, one currently in the top-level world and one in the outer room. In this example, the rooms are fixed both in size and spatial location but the crowds may be mobile and resizeable. This structure maps onto to the hierarchy of group objects shown in figure 3(b). The eight participants in the world are represented by the artefacts a to h.

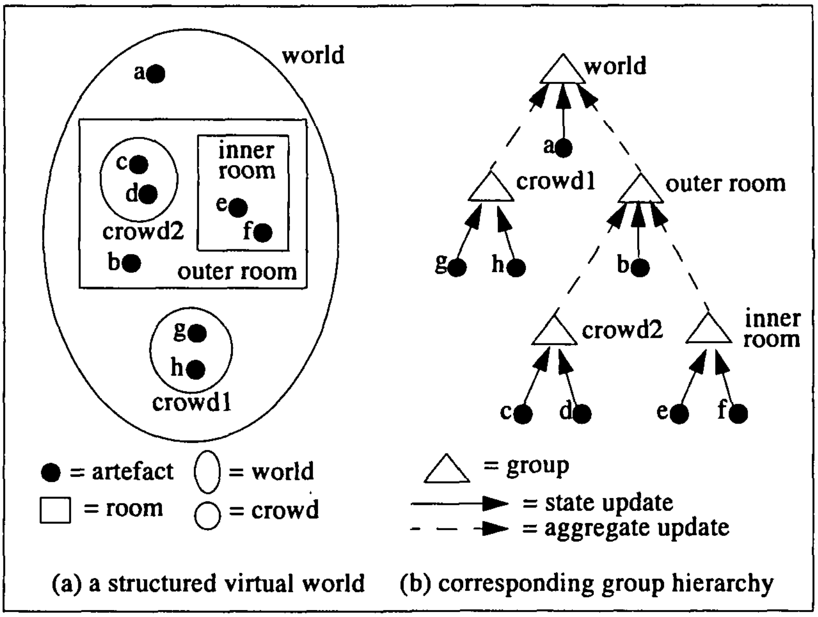


Figure 3: Spatial structure, group structure and sending updates

At any given moment in time each participant is deemed to be a sending member of its most local containing group. This means that the artefact associated with this participant sends its information (e.g. movement updates, audio packets etc.) to the multicast addresses associated with this particular group. As the participants move around the world, so they will pass through different spatial structures and, consequently, will be invited to join and leave different multicast groups as senders. Each group may also send an aggregated stream of information to its superior group in the hierarchy. Each artefact or group may only be a sending member of one other group at a time and membership as a sender is determined by strict spatial containment (i.e. an artefact or group has to be completely contained within a surrounding group in order to become one of its sending members).

As with participants, the movement of third party objects such as crowds around the world may cause the hierarchy of group membership to be automatically updated as spatial containment relationships change. Our architecture therefore results in there being an automatically self-configuring hierarchy of groups driven by movements in the virtual world.

4.3. Groups and auras

We now turn to the converse problem, that of dynamically managing the membership of auras within groups. In order to observe the world, a user (or other process) registers one or more auras in the world. An aura can have arbitrary (and changing) spatial extent and at any one time will overlap one or more third party structures. The effect of overlapping a group in this way is to join the aura as a receiving member of the group. Unlike artefacts, aura membership is based on spatial overlap not strict containment. Thus, an observer will receive information from potentially many groups depending upon the location and the extent of their auras. They may also dynamically resize their auras so as to increase or decrease the volume of this information. This is essentially a form of "spatial paging" of group objects - as one moves across the word, so different groups are paged in by one's aura and parts of the world are unfolded to different levels of detail. Thus, in figure 4, we see that two observers, X and Y, have each registered an aura in the world. X's aura currently overlaps the world, the outer room and crowd 2 and hence X receives updates from their associated group objects. Y's, on the other hand, overlaps the world and crowd 1, so they receive information from their associated groups.

As mentioned above, each MASSIVE-2 group object actually maps onto several underlying multicast groups. First, a world designer may choose to utilise different multicast groups for different media if they wish on a per-world basis. Second, different multicast groups are provided to handle different phases of joining and leaving a group as a receiving member. More specifically, the dynamic joining and leaving of groups raises an interesting problem. On joining a group, an aura will start to receive updates from a multicast group (e.g. movement notifications). However, without an initial snapshot of the whole state of the group these may not make much sense. In other words, a new aura joining a group has to first receive a (potentially large) state transfer before it can begin processing updates. This is handled by the introduction of a second multicast address that can be used to address the artefacts who are currently sending to the group (i.e. a back channel from auras to artefacts). This address is used by incoming auras to request a state snapshot from these artefacts. Thus, on joining a group, an aura first multicasts a state snapshot request to its artefacts. Once this has been received, these artefacts can then multicast updates to this state to the aura. As a further twist, the group may choose between unicasting and multicasting the state snapshot depending on how many requests are received within a given time (e.g. if many people enter a region of a world at one time, it may make sense to multicast the state update to them).
With reference to the underlying protocols, MASSIVE uses reliable multicasting

for all communication except for "real-time" media (e.g. audio and video). The protocols are receiver reliable in that sequence numbers and heart-beats on messages allow receivers to detect errors and to request retransmission. Thus, MASSIVE-2 has been designed to be run over generally reliable networks.

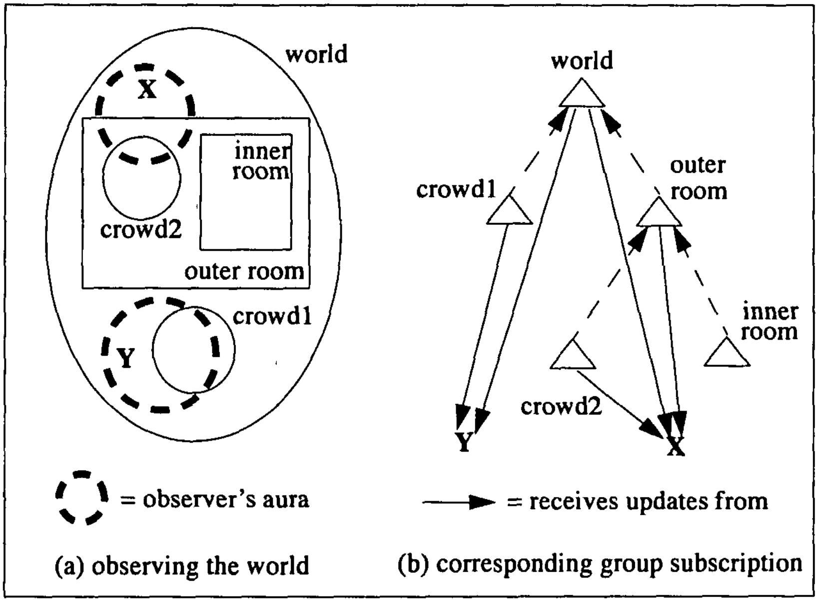


Figure 4: Aura and subscription to groups

4.4. Awareness negotiation

Having discussed in detail how groups, artefacts and auras combine to dynamically manage the multicasting of virtual world information, we now consider how objects negotiate detailed levels of awareness using focus and nimbus in order to manage the presentation of this information. This is achieved through specially defined awareness protocols.

Each artefact contains attributes that declare its specific focus and nimbus functions for different media. Default functions are provided based around a central region of maximum focus/nimbus that tails off according to a configurable distance and angle. However, other focus and nimbus functions can be added. Each artefact also contains attributes that list the groups of which it is currently a member, including a definition of the awareness adaptation effects of these groups. There are three such effects: a multiplier (that can amplify or attenuate awareness), a maximum awareness limit and a 'gate' threshold (an observer only becomes aware of the artefacts within a group once awareness of the group exceeds this threshold).

Each artefact provides methods that allow other artefacts to inspect its focus and nimbus functions, its current group memberships and also to sample its nimbus on another artefact. A further method is provided to interrogate an artefact about the value of its awareness of any other artefact. Evaluation of this method proceeds as follows. If we ask A about its awareness of B in some specific medium then:

- A samples the nimbus of B on A in that medium;
- A then asks B about its group memberships and compares them with its own, applying adaptation effects accordingly (again, taking the medium into account);
- Finally, A applies its local focus on B in that medium in order to calculate a final awareness level.

Awareness is then used to modify the presentation of the information that is delivered through the underlying multicast structure. For example, awareness may be used to control the local volume of an audio channel, the quality of service of video or the level of detail of graphics. This concludes of discussion of the MASSIVE-2 implementation of third party objects. As noted in section 3, MASSIVE-2 has already been used to develop a variety of CVE applications, including the public poetry performance described in (Benford, Greenhalgh, Snowdon and Bullock, 1997).

5. Summary

This paper has introduced an extension to the spatial model of interaction called third party objects, intended to support contextual factors in awareness and also greater scaleability of CVEs. Third parties are independent objects in a virtual world that can adapt existing awareness relationships between other objects and can also introduce new indirect awareness relationships corresponding to secondary sources of information. As objects in their own right, third parties can be mobile, resizeable, dynamically created and apply recursively to one another. These characteristics make them a powerful tool for defining CVEs. Our paper has explored a number of applications of third parties including: world structuring (e.g. defining nested rooms, buildings, regions and zones that apply different combinations of adaptation and secondary sourcing across their boundaries); providing aggregate views, especially of crowds and data districts; acting as common foci for shared interaction; embodying a range of more general representational and group services; and supporting dynamic load management.

We have discussed MASSIVE-2's implementation of this extended spatial model, focusing on how it maps the structure of a virtual world as defined by third party objects onto a underlying network architecture of multicast groups. Given the notion of secondary sourcing and also the fact that third parties may be mobile and resizeable, this multicast architecture has a number of highly novel features, namely, an automatically self-configuring hierarchy of multicast groups, with aggregated transmissions from children to parents in the hierarchy and with a "spatial paging" mechanism used to manage the joining and leaving of multicast groups.

As with the first version of MASSIVE, we intend the next few years to see a range of applications of this system, eventually leading to organised trials and evaluation of the system and underlying model. In terms of applications, we anticipate a strong focus on "citizen oriented" applications in areas such as entertainment, arts and performance. Indeed, a companion paper to this describes an early experiment with this kind of application (Benford, Greenhalgh, Snowdon and Bullock, 1997). In the longer term, we anticipate the development of CVEs that will be capable of supporting hundreds or thousands of simultaneous participants and believe that the mechanism of third party objects may provide one of the techniques required to achieve this.

References

Barrus, J and Anderson, D. (1996) "Locales and Beacons." in *Proc. 1996 IEEE Virtual Reality Annual International Symposium (VRAIS'96)*, San Jose, 1996.

Benford, S. D. and Fahlén, L. E. (1993): "A Spatial Model of Interaction in Virtual Environments", *Proc. ECSCW'93*, Milano, Italy, September 1993.

Benford, S. D., Greenhalgh, C. M. and Lloyd, D (1997): "Crowded Collaborative Virtual Environments", *Proc. CHI'97*, Atlanta, Georgia, US, March 22-27, 1997 (in press).

Benford, S.D., Greenhalgh, C. M., Snowdon, D. N, and Bullock, A N. (1997) "Staging a Poetry Performance in a Collaborative Virtual Environment", also submitted to ECSCW'97.

Bowers, J., Pycock, J. and O'Brien, J. (1996a): "Talk and Embodiment in Collaborative Virtual Environments", *Proc. CHI'96*, ACM Press, 1996.

Bowers, J , O'Brien, J. and Pycock, J (1996b): "Practically Accomplishing Immersion: Cooperation in and For Virtual Environments", *Proc. CSCW'96*, pp 380-389, November 1996, Boston, ACM Press.

Funkhouser, Thomas A. (1996): "Network Topologies for Scalable Multi-User Virtual Environments", in *Proc. 1996 IEEE Virtual Reality Annual International Symposium (VRAIS'96)*, San Jose, CA, April, 1996.

Greenhalgh, C. M and Benford, S. D. (1995): "Virtual Reality Tele-conferencing: Implementation and Experience", *Proc. ECSCW'95*, Stockholm, September, 1995, North-Holland

Greenhalgh, C, M. (1997): "Analysing movement and world transitions in virtual reality tele-conferencing, submitted to ECSCW'97.

Hagsand, O. (1996) "Interactive Multiuser VEs in the DIVE system", *IEEE Multimedia*, Spring 1996.

Ingram, R. J., Benford, S. D. and Bowers, J. M. (1996) "Building Virtual Cities: Applying Urban Planning principles to the Design of Virtual Environments", *Proc. VRST'96*, Hong Kong, July 1-4 1996, pp. 83-91, ACM Press.

Macedonia, Michael R , Zyda, Michael J., Pratt, David R., Brutzman, Donald P., and Barham, Paul T. (1995): "Exploiting Reality with Multicast Groups: A Network Architecture for Large-scale Virtual Environments", in *Proc. 1995 IEEE Virtual Reality Annual International Symposium (VRAIS'95)*, 11-15 March, 1995, RTP, North Carolina.

Rodden, T. R. (1996): "Populating the Application: A Model of Awareness for Cooperative Applications", *Proc. CSCW'96*, pp 67-76, November 1996, Boston, US, ACM Press.