Technical Note PR-TN-2004/00370

Issued: 05/2004

Mass Customization in Ambient Narratives

Mark van Doorn

Philips Research Laboratories

Eindhoven

Arjen P. de Vries

Centre for Mathematics and Computer Science

Amsterdam

Unclassified

Authors' address			
	Mark van Doorn	WY 21	mark.van.doorn@philips.com
	Arjen P. de Vries		arjen.de.vries@cwi.nl

Title: Mass Customization in Ambient Narratives

Author(s): M.G.L.M. van Doorn; A.P. de Vries

Reviewers(s):

Technical Note: PR-TN-2004/00370

Additional Numbers:

Subcategory:

Project:

Customer: Philips Research

Keywords: Adaptive hypermedia, interactive storytelling, ubiquitous hypermedia, ambient intelligence, mass customization

Abstract: Ambient intelligence is a vision on the future of consumer electronics that refers to the presence of a digital environment that is sensitive, adaptive and responsive to the presence of people. This paper proposes a mass customization strategy to deliver tailor-made ambient intelligence environments on a mass basis. It introduces the idea of an "ambient narrative", defined as an interactive narrative situated in mixed reality, and discusses how ambient intelligence emerges out of the continuous interaction people have with an ambient narrative. More specifically, it describes an ambient narrative as consisting of a set of predefined, interrelated, modular parts called 'beats'. These beats are sequenced by an "ambient narrative engine" based on user's feedback, context and past experiences, into a unique mixed reality story we call ambient intelligence. Furthermore, it discusses how this concept of ambient narratives may be implemented within two existing extensions of the Dexter hypertext reference model, the Amsterdam Hypermedia Model and the Adaptive Hypermedia Application Model.

Conclusions: Ambient Intelligence is a vision on the future of consumer electronics that refers to the presence of a digital environment that is sensitive, adaptive and responsive to the presence of people. Since it is technologically not possible to mass produce Ambient Intelligence with the current state of the art in artificial intelligence and economically not feasible to manually develop tailor-made Ambient Intelligence products or services for each customer individually, we believe a different approach is needed to move Ambient Intelligence out of research laboratories into the real world. In this paper we described a mass customization strategy for Ambient Intelligence services offered over a collection of networked devices to customize Ambient Intelligence on a mass basis. In this approach modular parts of Ambient Intelligence descriptions are assembled based on user feedback and interaction history into a personalized flow of personalized, interactive media presentations in which multiple devices may participate simultaneously. The Ambient Intelligence service thus allows people to create their own Ambient Intelligence within the scope of possibilities set down by the designer or writer of an ambient narrative, an interactive narrative in mixed reality that is designed to support people at home or in a particular service encounter in performing their everyday activities. We also explained that this mass customization strategy for ubiquitous hypermedia applications can be implemented in the existing Amsterdam Hypermedia Model (AHM) and Adaptive Hypermedia Application Model (AHAM). Although we believe this approach of mass customization in ambient narratives is conceptually sound, and supported by economical and social-cultural drivers, it will need to be implemented and tested in practice. To this end we are working on an example ambient narrative situated in HomeLab, the usability and feasibility lab at Philips Research.

Contents

1. Introduction

Ambient Intelligence denotes a new paradigm in the history of computing that aims to improve the quality of people's life by making everyday activities more convenient and enjoyable with digital media. Technically, ambient intelligence refers to the presence of a digital environment that is sensitive, adaptive, and responsive to the presence of people [1]. Electronic devices are embedded in furniture, clothing or other parts of the environment; the technology recedes into the background of our everyday lives until only the function (i.e., the user interface) remains visible to people. At the same time, the Human moves into the center of attention, in control of the devices around him. These work in concert to support the performance of everyday activities in an intelligent manner. By moving technology and devices into the background and the user experience into the foreground, the Ambient Intelligence vision acknowledges the shift from a supply-driven industrial economy into a demand-driven, global, informational network economy. In the network economy, people pay access to personalized information and services that meet their widely-diverging individual needs and wants (instead of paying for ownership of standard, mass produced physical products) [9,33]. This leads us to view Ambient Intelligence as a personalized information service delivered over one or more networked devices simultaneously to individual users. The intelligence provided by the service is not bundled together with the ambience in a single and inflexible physical product, like a TV or set-top box, but offered separately as a service over simple network appliances that provide the ambience.

Producing Ambient Intelligence on a larger scale is problematic. First, it is technologically not possible in the foreseeable future to mass produce a product or service that generates Ambient Intelligence, given the current state-of-the-art in machine learning and artificial intelligence. Economically however, it is not feasible to manually design and produce Ambient Intelligence applications for each person individually. Our proposed solution is to introduce a different manufacturing approach. The paper describes ongoing work on a software product line approach towards Ambient Intelligence, designed to churn out Ambient Intelligence services on a mass basis, but customized through the individual user's input and feedback. At the moment we only provide a blueprint and partial implementation of such a system. We believe that it is a sound approach, and provide an argument by describing the envisioned system in terms of two existing extensions on the Dexter hypertext reference model.

The remainder of this paper is organized as follows. First, we look more closely at Ambient Intelligence from a social-cultural perspective to get a better idea of what it is exactly that needs to be mass customized. Next, we discuss the notion of ambient narratives and related work in interactive storytelling, hypertext and mixed reality applications. We will then describe how we can implement a mass customization strategy and map this strategy in terms of the Amsterdam Hypermedia Model (AHM) [22] and the Adaptive Hypermedia Application Model (AHAM) [6]. We end this paper with some preliminary conclusions.

2. Performances in Everyday Life

Put simply, Ambient Intelligence aims to help people in performing their daily activities better, by making these activities more convenient and enjoyable: by introducing interactive media. Notice the word ‘performing’ in this description. The word ‘performing’ is so commonly used in our language that we forget that performances are pervasive to every culture. Performances are not just seen on in the theater, music, dance and performance arts in general but also in our everyday lives: We perform the role of a father or son in our private lives but maybe also that of a doctor, judge or police agent for example in our professions. In order to understand where and how Ambient Intelligence can be applied to support these performances, we first need to gain a better insight into what performances are and what it means to perform.

Performance studies is an emerging yet wide-ranging interdisciplinary field that takes performances as the organizing concept for the study of a wide range of human behavior. It embraces research in the social sciences from anthropology to theatrical studies. Performance studies regards the prevailing division of the arts by medium and culture as arbitrary and therefore looks at performances from a holistic, post-modern point of view; a broader view towards literature and Western theater. Because performances vary so widely from medium to medium and culture to culture, it is hard to pin down an exact definition for performance. Schechner defines performance as “ritualized behavior conditioned/permeated by play” or “twice-behaved behaviour” [37]. When people are performing, they show behavior that is at least practiced once before in a similar manner. In traditional performance arts this behaviour can be detected easily: Actors in a theater play, opera or movie rehearse their roles off-stage and repeat this behavior when they are on stage. But this twice-behaved behavior can also be seen in a priest conducting a wedding ceremony, a surgeon operating on a patient or a McDonald’s service employee behind the counter. Even in our own homes, people show signs of this repeated behavior. This happens for example during everyday rituals, like brushing your teeth in front of a mirror in the morning, watching a soccer match with friends, or, coming home from work in the evening. Note that, here, the sending and receiving party in a ‘performance’ may be the same.

Performances in everyday life can be investigated from two different angles. Linguistic theorists begin with language. Linguistic philosopher J.L. Austin [3] coined the term ‘performative’ to indicate a word or sentence that does something (e.g. “I go home if he is not here in five minutes”). John R. Searle later expanded on Austin’s ‘performatives’ with speech act theory [38]. In speech act theory, a speech act is the production or issuance of a sentence token under certain conditions. According to Searle, speech acts and not symbols, words or sentences are the basic or minimal units of linguistic communication. In this linguistic approach towards understanding performances, behavior is to be ‘read’ as complex, interacting texts. ‘Text’ can be very broad in this view; according to French post-structuralist Jacques Derrida [15] there is nothing outside the text, text in his theory encompasses all of human culture. The second approach starts from the work of Ervin Goffman who views social life as theater. According to Goffman [19], people follow culturally specified social scripts that interact with each other. Social scripts may differ from culture to culture and from epoch to epoch, but no culture exists without social scripts. Although both approaches take a different perspective, they both agree that people are performing all the time, most of the time without knowing.

If social scripts are universal and we communicate with others by performing these scripts, it seems plausible that nature structures knowledge in our minds in a similar way, especially in the context of language understanding. Schema theory is not only studied in linguistics and anthropology but also in psychology and artificial intelligence. Interesting to mention here is the work of Schank & Abelson (1977) [36] who introduced the concepts of scripts, plans and themes to handle story-level understanding. In Schank's script theory, specific memories are stored as pointers to generalized episodes (scripts) plus any unique events for a particular episode. Scripts allow individuals to make inferences needed for understanding by filling in missing information.

Viewing life as social theater is interesting for us for two reasons: First, if people behave according to social scripts, we can codify interactive media applications to support people in carrying out these scripts. Just as lighting and sound effects add to the overall drama of a theater play, ambient intelligence may thus be applied to enhance the performance described by these social scripts. It leads to a concise and concrete definition of ambient intelligence: *media-enhanced environments designed to support everyday life performances*. Second, positioning Ambient Intelligence in performance theory gives us a well-studied and familiar frame of reference for the design of Ambient Intelligence environments and the underlying technology.

2.1. Private and Public Performances

This paper narrows down the scope of media-enhanced performances to Ambient Intelligence in the home and in professional service encounters. The home people live in can be seen as a stage on which perform everyday rituals, such as brushing your teeth, going to work or watching a soccer match. Let us give an example of how Ambient Intelligence could be applied to improve the experience of such a daily activity of tooth brushing.

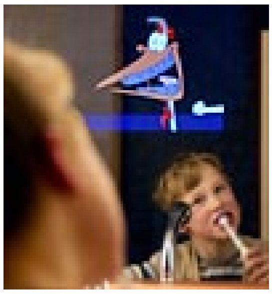

Figure 1: Enhancing everyday life performances with ambient intelligence

Figure 1 shows a cartoon projected on a mirror TV (a two-way mirror with an LCD screen behind) that invites the small child standing in front of the mirror to brush his teeth for two minutes. The cartoon carries the child through this daily task. It would be too easy to say that we can create an optimal experiences by making a task more effective or more entertaining. A better explanation of what optimal experiences are, is perhaps provided by psychologist Mihaly Csikszentmihalyi, who argues that happiness is not so much a result of finishing a task but more about being immersed and engaged in the process of performing the task. Only then do we get into a state of 'flow' [12] and optimal experience. In the mirror TV example, the cartoon shifts the attention of the child from achieving the end result to the process of getting there. The cartoon increases the flow of the activity by improving the level of engagement.

The central role of performances is also reflected in recent business literature about services. Pine and Gillmore [32] talk about an experience economy in which work is theater and every business a stage. They argue that staged experiences are valued higher by customers than delivered services, manufactured goods or extracted raw materials. Experiences fulfil in a larger subset of customer needs, so these are willing to pay a higher price. Rifkin [33] further argues that in this experience economy, culture itself is being pulled into the commercial sphere. Entertainment parks, themed shopping malls and organized tourist travel are just some examples where people no longer buy things, but pay to gain access to services and commodified culture. Earlier research on service marketing by Fisk and Grove [17] discusses a theater framework for service marketing, in which services are seen as performances, front-end service personnel as actors, the service setting or servicescape as the stage on which the service is performed, products used in the services as props and the business process of the service as the script. Bitner [5], Clark [11], and Hightower [23] provide empirical evidence that the 'servicescape', the environment of the service, plays an important role in how people perceive a service encounter. This suggests that Ambient Intelligence can also be applied in professional service encounters to enhance the atmospherics of a service encounter and thereby the perception or performance of a service in a positive way.

Consider for example a medical imaging room in a hospital. Many patients are frightened by the bulky equipment in the examination room. By enhancing this environment with immersive media, e.g. by projecting video clips on the walls and ceiling of the examination room, patients may feel more at ease, as illustrated in Figure 2. Recently, Philips Medical Systems demonstrated this concept in the "ambient experience pavilion" at the Annual Meeting of the Radiological Society of North America (RSNA) in December 2003.

Figure 2: Enhancing a medical examination room with ambient intelligence.

The medical examination room and bathroom mirror TV cases are clear examples of media-enhanced environments that are designed to support everyday life performances in private or public spaces, there are many more of such examples of Ambient Intelligence.

3. Ambient Narratives

To model media-enhanced performances in the home and commercial service encounters in a machine understandable way, we choose to represent the structure and interrelationships of a set of related media-enhanced performances as an interactive or episodic narrative. Interactive narratives allow readers to affect, choose or change the plot of a story. Most interactive narratives are situated either in the real world (e.g. live-action role playing games, improvisational theater) or in some virtual reality (e.g. hypertext novels, computer games), see e.g. Murray [31] for an overview of different forms of interactive narrative. The media-enhanced performances we are concerned with are however 'real' nor virtual; they occur in *mixed reality*. The behavior of people and the use of products in a media-enhanced performance may take place in the real world, but the digital media to augment the performance originate from a virtual dimension.

Another difference is that these media-enhanced performances are not really 'read' like a book or hypertext novel, but enacted like a theater play. We introduce the term *ambient narratives* to denote these kinds of dramatic, interactive narratives that play in a mixed reality setting. We can look at ambient narratives from a consumer (reader) point of view or a producer (writer) perspective. From a reader point of view, interaction with the ambient narrative creates an ambient intelligent environment. Interaction should be taken very broadly here as an ambient narrative can span both virtual and physical dimensions at the same time. Media-enhanced performances in different rooms may be linked to each other in one narrative structure, allowing people to influence the plot of the ambient narrative (the evolving Ambient Intelligence); for example, by walking around, or, touching a glowing lamp. From the writer's perspective, the ambient narrative describes all possible media-enhanced performances and their interrelationships. The ambient narrative is written in advance and enables consumers to create their own personal story, their own Ambient Intelligence. Unless explicitly mentioned, we will look at ambient narratives from this writer's point of view in the remainder of this paper, as our focus is on the production of Ambient Intelligence.

4. Related Work

Ambient narratives can be related to interactive storytelling, hypertext, and mixed reality applications.

Many interactive storytelling systems designed so far in the interactive fiction domain are based on Artificial Intelligence (AI) techniques that control either the plot generation or real-time character behavior, see e.g. [4,10,27,28,41]. Because AI planning techniques deal with action selection, these techniques lend themselves also well to sequence plot elements into a compelling dramatic story. The work on interactive drama dates back to the work of Laurel [25], who defined the following notion of interactive drama: "*a first-person experience within a fantasy world, in which the user may create, enact, and observe a character whose choices and actions affect the course of events just as they might in a play. The structure of the system proposed in the study utilizes a playwriting expert system that enables first-person participation of the user in the development of the story or plot, and orchestrates system-controlled events and characters as to move the action forward in a dramatically interesting way".* One of the main issues in interactive drama systems is the balance between player interactivity and plot structure: Too much structure imposed, and users may get the impression that their actions do not matter, because the important choices have been made anyway. Too much focus on interactivity on the other hand causes the impression of feeling 'lost'. Galyean therefore writes in her dissertation [18] that interactive narratives should offer "narrative guidance". She used a river flow analogy to explain how interaction through the plot structure can be guided without disrupting the dramatic structure or restricting the freedom of the player too much. Believability is another important factor [4] in interactive drama (as well as in human computer interaction design in general [26]). The characters in the story world may seem alien, but their actions should be believable at all times within the context of the story world in which they live. Murray [31] proposes three categories for the analysis of interactive story experiences: immersion, agency and transformation. Immersion is the feeling of being present in another place and part of the action, it relates strongly to believability. Agency is the feeling of empowerment that comes from being able to take actions that relate to the player's intention. Transformation is the effect that happens if a player's self is altered by the multitude of possible perspectives offered by an interactive story experience.

The non-linear nature of interactive narratives has attracted the attention of both hypertext writers and researchers. Michael Joyce's "Afternoon, a story" (1987) and Stuart Moulthrop's "Victory Garden" (1991) are some successful examples of hypertext novels. During the nineties, hypermedia narratives were also demonstrated, HyperSpeech [2] and HyperCafe [35] to name just a few. Recent publications have expanded models and techniques into mixed reality environments [20,34,39]. With respect to interactive storytelling, Weal et al [42] describe an approach that records the activities of children in an outdoor environment and represent these recorded activities using adaptive hypermedia techniques to children later so they can later relive the experience. Romero and Correia [34] discuss a hypermedia model for mixed reality and describe a mixed reality game that takes place in a gallery environment.

In our notion of ambient narratives, we view the mixed reality experience itself as a story that emerges when people interact with the mixed reality environment. But, the everyday life performances do not necessarily have to be games; they are more likely to be ordinary tasks like brushing your teeth, or, services encounters in a hotel or hospital. The ‘reader’ of a mixed reality hypermedia narrative transforms into a player in a story where the physical world becomes the medium for interaction. A key design principle of ambient narratives is this immediacy of the physical world that is also seen in live role-playing games [16]. The fact that the experiences, that a character is subjected to, happen immediately to the player him- or herself, creates a more sophisticated form of sensory engagement during the activity augmented with Ambient Intelligence.

The focus on action (enactment) implies that the writing style of mixed reality hypermedia narratives will be different than that of ‘traditional’ hypertext novels and hypermedia documentaries, in which descriptions are the main ingredient to tell a story. Laurel [26] makes a useful distinction between dramatic and literary stories that may help to describe how traditional hypertext novels and hypermedia documentaries differ from mixed reality hypermedia stories that are performed in real-time. Laurel states that dramatic stories have properties of enactment, intensification and unity of action, where literary stories have properties of description, extensification and episodic structure. This classification is interesting because it implies that ambient narratives are not literary but dramatic narratives, designed to improve everyday life performances in private or public spaces that use the world as a medium for interaction. Ambient narratives are performed, acted in real time (enactment). The architecture, interior design and immersive media must work in concert with each other for an optimal experience (intensification). Because the goal of ambient narratives is to improve everyday life performances, all incidents should be causally related to a central action (unity of action).

5. A Dexter-based Ambient Narrative System

The ambient narrative approach outlined above enables mass customization of Ambient Intelligence because it enables each individual user to mix-and-match the desired Ambient Intelligence from predefined plot material. In general, for a mass customization strategy to work there must be [13]:

- Modularity present in the product or service,
- Customer interface through which customers can easily enter product requirements or specifications,
- Flexible product line designed to churn out customized products on a mass basis and a
- Network to connect the customer interface to the flexible product line.

In the ambient narrative approach, modularity is achieved by the ambient narrative itself which constitutes the modular parts of all the possible Ambient Intelligence experiences that can be assembled. The customer interface is implemented by the *ambient browser* that collects user input and context information from multiple devices and sends this information over a network to the flexible product line. The flexible product line is implemented by the *ambient narrative engine* that determines what the next best part is given this information and the current position in the ambient narrative plot. The ambient narrative engine returns the presentation description inside the newly selected part to the ambient browser that renders this presentation. So together, the ambient browser and ambient narrative engine assemble the customized product, Ambient Intelligence and deliver it to the actor.

The remainder of this chapter describes how the high-level ambient narrative system architecture above can be implemented in terms of (extensions of) the Dexter hypertext reference model. The Dexter hypertext reference model [21] was the result of two small workshops the first one held in 1988 at the Dexter Inn in New Hampshire to provide common terminology tied to a formal model of the important abstractions commonly found in hypertext systems. Although newer models have been proposed in recent years, the Dexter reference model remains widely used. We will use the same terminology for describing an ambient narrative system. The Dexter model introduced three conceptual layers for hypertext systems: The run-time layer, the storage layer and the within-component layer. The run-time layer deals with the presentation of hypertext components and the interaction with the user. The focus of the Dexter model is on the storage layer which models the node/link network of the hypertext system. The within-component layer finally contains the content and structure within the nodes. The Dexter model also describes the interfaces between these layers; the presentation specifications between the run-time and storage layer and the anchoring interface between the storage and with-in component layer. Figure 3 shows the mapping of our high-level ambient narrative system architecture on the Dexter model. The following chapters explain how the run-time layer is implemented by the ambient browser, the storage layer by the ambient narrative engine and the within-component layer by the plot material of an ambient narrative.

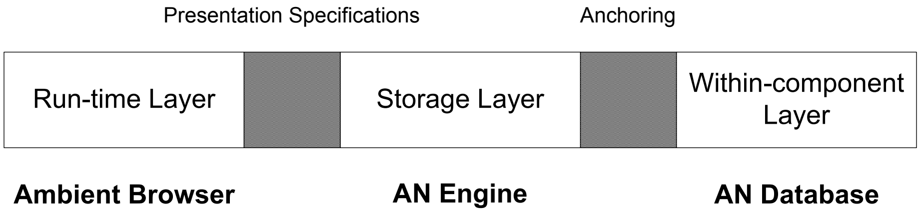

Figure 3: Mapping the high-level system architecture onto the Dexter reference model Run-time Layer

5.1. Run-time Layer: Ambient Browser

The Dexter model only provides a simple mechanism for presenting hypertext to the user for viewing and editing, it does not attempt to cover the details of user interaction with the hypertext. The interface with the storage layer is through presentation specifications. Presentation specifications are encoded into the hypertext network at the storage layer and contain information about how a component/network is to be presented to the user. This way, the presentation of a component to the user is a function of the specific hypertext tool that is doing the presentation (e.g. Microsoft Internet Explorer), the property of the component itself (e.g. the specific HTML document) and/or the link taken to that component. Presentation specifications provide a generic way for the storage layer to communicate with the run-time layer without violating the separation between these layers.

Although hypermedia can be added to the Dexter model as a data type in its storage layer, this approach cannot adequately support the complex temporal relationships among data items, high-level presentation attributes and link context that are important in supporting hypermedia. The Amsterdam Hypermedia Model (AHM) [22] extends the Dexter model by adding these hypermedia requirements. The AHM model inspired the definition of the Synchronized Multimedia Integration Language (SMIL) [40]. AHM and the SMIL language are designed for hypermedia presentations running on a single device and therefore do not mention the issue of timing and synchronization of media elements across multiple devices, characteristic for mixed reality hypermedia. To support timing and synchronization of media objects within and across devices, we use an in-house developed SMIL interpreter. This SMIL interpreter has the role of a networked service to which the other devices register themselves. The underlying reason for this choice is that we expect 'media' to include also lamps, fans and other output devices in the future. To use this functionality in SMIL, we have extended the toplayout element in the SMIL language with a proprietary 'target' attribute that specifies the rendering (or input) device. The author of the SMIL document can set the target attribute of the toplayout element in the SMIL head part to point to a specific rendering or input device. In the SMIL body part, the author can use (as he would normally do) the id of the toplayout element or the id of one of its region element children in the region attribute of a media element (e.g., an image, or a video fragment), to indicate the device/region on which the media element should be rendered. Figure 4 describes the outline of the SMIL document that belongs to the mirror TV example of figure 1. The advantage of this approach is that we can remain close to the AHM model and do not have to introduce spatial mapping functions outside the SMIL engine as described in [24] for example.

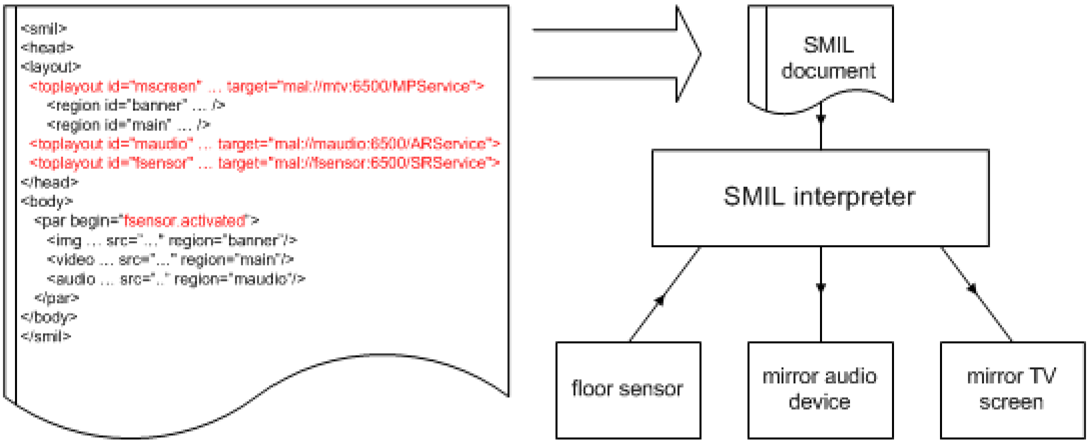

Figure 4: Extending SMIL for mixed reality media applications

From a narrative point of view, the presentation and interaction with SMIL belongs to the run-time layer in the Dexter model. This view differs to some extent from the role of media objects in the AHM, which addresses mostly the storage layer. We consider SMIL documents as basic building blocks, i.e., the components in the storage layer. The AHM model however views individual media objects as the principal components. The advantage of our approach is that we can abstract from the low-level positioning of media objects in space and time and focus on describing how such hypermedia presentations are selected based on context situations. This not mean that the AHM model is not present; it is just hidden. Inside the SMIL components, the AHM model is revealed.

5.2. Storage Layer: Ambient Narrative Navigation

The storage layer in the Dexter model describes the mechanisms by which nodes and links are connected to form networks. The storage layer itself is not specific about the internal structure of the components; components are treated as generic data containers. The components of the ambient narrative are called *beats*. The term beat originates from theater play and film script-writing and is defined [29] as a change in the behavior of a character. The beat is the smallest unit within a scene that represents a greater change in behavior. Sequences of scenes form acts which culminate into even greater change. This beat (sequencing) approach is also taken by e.g. Mateas [28], Magerko [27] and Brooks [7] in interactive drama (and since ambient narratives are also enacted, we choose the same terminology here). So, an ambient narrative is equivalent to a hypertext network consisting of beats (nodes) and links between these beats.

An individual Ambient Intelligence experience is represented by a sequence of beats and any specific parameters that may have been used in the instantiation of these beats. So like in Schank's script theory, specific memories (in our case Ambient Intelligence experiences) are represented as pointers to generalized episodes (beats) plus any unique events for a particular episode (represented by story values, as explained shortly).

5.2.1. Beat language

Now, we describe first a language for representing beats, followed by an explanation of how beats are sequenced by a beat sequencing engine or ambient narrative engine.

A beat is a component in the storage layer of the Dexter model. It consists of three parts: the *preconditions*, *action* and *post-conditions* part. Before we go on and explain these language parts in more detail, we introduce the notion of a *story memory* that contains *story values* that can affect the plot of the ambient narrative. The goal of this mechanism is to make the system less reactive. Writers can use story values in each part of the beat description as a mechanism to influence the selection of a new beat based on for example the current state of the story (e.g. the knowledge level of a particular user) or specific context information (ambient temperature or humidity). The story memory is simply a set of these dynamically changing story values or variables.

The preconditions part (component information) is checked during the selection or search for a new beat component by the beat sequencing engine or ambient narrative engine that controls the navigation through the ambient narrative. It specifies the conditions that must hold before the beat can be selected: it can contain information about the activity (performance), user(s) (actors), location (stage), tangible objects and devices present (props) and/or past user/system interaction (script).

The action part is executed after the beat has been selected by the engine for sequencing. It is subdivided in an initialization part (component information) and a main part (component content). In the initialization part, the author can specify which story values need to be modified in the story memory before the presentation is started. The main part specifies the interactive mixed reality presentation, encoded as a SMIL document in our case. We introduce an extra level of adaptation by allowing the multimedia presentation document to contain embedded queries with story values as parameters. These parameters are set at run-time by the ambient narrative engine. This way we allow for more variation in the media-enhanced performances and also reduce the amount of authoring effort needed. The SMIL document also contains links and/or queries with story values as parameters to the beat sequencing engine.

The post-conditions part (component information) is executed when the currently active beat is about to be replaced by a new beat, i.e. before the preconditions part of the next beat has been executed. The author can specify here which story values need to be modified in the story memory.

The design of this beat language is similar to the interactive drama approaches taken by Mateas and Magerko. Our preconditions part is similar to the selection knowledge stage in Mateas' Facade interactive drama architecture [28] that also uses tests on story memory values (and prior probability on beats). The action stage in Facade roughly corresponds to our action and post-conditions parts in the beat document. Magerko's IDA Architecture [27] represents plots at the scene level and consists of five stages: initial state, required events, background knowledge, content constraints and temporal constraints. The initial state sets up the scene and may be similar to the initialization part in our action stage. The required events and background knowledge are comparable with our preconditions stage, while the content constraints that limit the binding of the variables used in the required events are similar to the embedded queries in the main action description part of our ambient narrative beat documents.

5.2.2. Beat sequencing

Beats are sequenced together to create a personalised story in mixed reality, the result of which is what we have called Ambient Intelligence. The beat descriptions and beat sequencing engine or ambient narrative engine can be een as an adaptive hypertext system. Brusilovsky [8] describes adaptive hypermedia as follows: "*By daptive hypermedia systems we mean all hypertext and hypermedia systems which reflect some features of the user in the user model and apply this model to adapt various visible aspects of the system to the user. In other words the system should satisfy three criteria: it should be a hypertext or hypermedia system, it should have a user model and it should be able to adapt the hypermedia model using this model.".* The Adaptive Hypermedia Application Model (AHAM) [6] builds further upon this definition and tries to fit adaptive hypertext and hypermedia in the Dexter model. It defines a hypermedia application as consisting of a domain model, user model, teaching model and adaptive engine. The domain model describes how the information is structured in nodes and links. The user model describes the knowledge level of the user and also keeps a record of the nodes visited by the user in the past. The teaching model consists of learning rules (called pedagogical rules in AHAM) which define how the domain model interacts with the user model. The adaptive engine performs the actual adaptation.

The beat descriptions and beat sequencing engine can be quite easily described in terminology of the AHAM model. The beats and their interrelationships form the domain model. The user model is implicit in the story memory of the ambient narrative engine: The story memory contains session knowledge and context knowledge which includes knowledge about the user and his context as we discussed before. The story memory dynamically evolves out of the continuous interaction between users and the ambient narrative. The teaching model is encoded in the beat descriptions. The action initialization part and post-conditions part allows the author to alter story values that can affect the selection of new beats and content items. The adaptive engine is the ambient narrative engine that sequences beats. The ambient narrative engine must implement an action selection mechanism as its main task is to find the next best beat. Instead of a rule-based AI approach, we prefer to implement the beat sequencing planner by an information retrieval (IR) approach. Queries for new beats are encoded in the action and post-condition parts of the beat description, that may contain both fixed parameters and story values. Beat queries are never explicitly entered by the user, they are selected and filled in by the beat sequencing engine based on user input and information present in the story memory.

The advantage of this approach is that it allows us to introduce adaptation at different levels of the narrative like the Facade [28] and Hyperdoc [30] systems. The notion of embedded content queries in beats allows us to describe adaptation at the subnode level, i.e., the media objects in the SMIL document that encodes the action part of the beat. The notion of embedded beat queries in beats allows us to describe adaptation at the link level: The choice of a beat can be made context-dependent by using story values in the embedded beat queries. The use of beat preconditions and embedded content queries allow us to easily add new beats and content without taking the system down. This way we can defer editing decisions by the narrative engine on both the node and subnode level until the moment they are played out. The same technique is used by the Automatist Storyteller system [14] for content items only. As a result authoring effort is lowered because the author does not have to explicitly sequence story elements into a finished story or rewrite the existing narrative when new material is added. This also provides the writer with the flexibility to add beats for specific devices that can be chosen if these devices are owned by the reader. Furthermore, this open approach creates possibilities for sharing and exchaing beats and content among narrative engines. Authoring tools may assist readers in creating their own beats and inserting them in their beat collection.

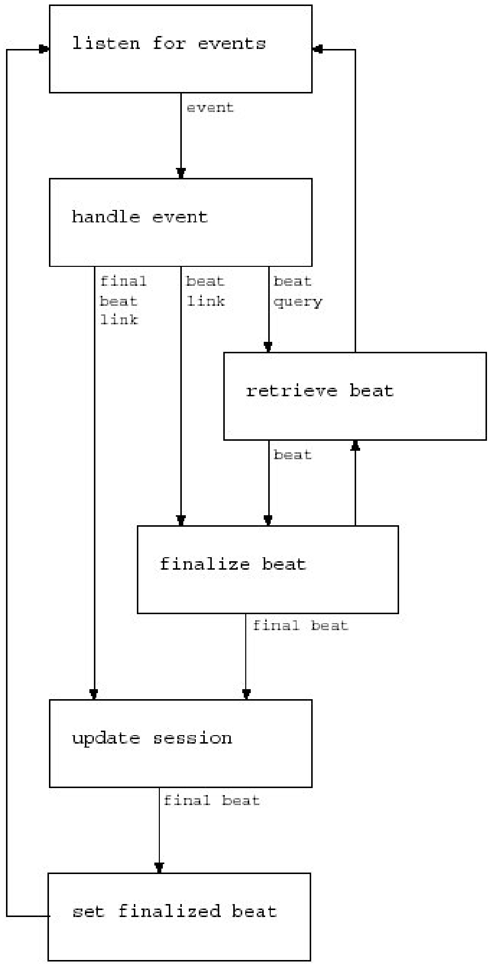

Figure 5: Process flow in ambient narrative engine.

Figure 5 gives a blueprint of the ambient narrative engine, which is described in detail below. User input (implicitly derived from sensors or explicitly given by the user) and story memory together determine the direction through the ambient narrative and how the selected beats are customized and sequenced.

Step 1 (listen for events) involves the detection of events by sensors and input devices by the ambient browser and forwarding relevant events to the AN engine. In the mirror TV case (Figure 1) a sensor in the floor might send an event to the ambient browser that interprets the event and forwards it to the AN engine.

In step 2 (handle event) the AN engine checks if one of the triggers currently active is triggered by the event. If this is true, the AN engine uses the story memory and the user information to fill in the parameters of the trigger. This way, variation in navigation is introduced. If the trigger turns out to be a link to a parameterizable beat, the AN engine jumps to step 4, if the trigger is a link to a finalized beat, the AN engine jumps to step 5. We say a beat is finalized if all of its open parameters have been filled in. If the trigger turns out to be a query for a beat, the AN engine continues with step 3. If there is no such trigger, the database containing the beats is queried using only the story memory and user information present (step 3). Before the new beat is selected, the post-conditions part of the old beat is executed, any changes in story values are also communicated to the story memory. In our example, the floor sensor event is interpreted and triggers a query for a new beat which contains query attributes location ('bathroom') and the user ID of the person standing on the floor sensor (the child).

Step 3 (retrieve beat) involves the retrieval and ranking of beats. According to some similarity measure the query for a beat is compared with the top results. The best matching beat is selected and finalized in step 4. Story values used in the beat query are read from story memory and inserted as query terms. If the ranking is below a certain threshold, it is possible to return no beat, meaning that the currently running application, presentation in the ambient browser is still the most relevant one. In the mirror TV example, the best matching beat would be an interactive multimedia presentation describing a cartoon that makes the toothbrush activity more entertaining as show in Figure 1.

Step 4 (finalize beat) is concerned with finalizing the selected beat, i.e., filling in the open parameters. The embedded queries in the action part of the newly selected beat are executed. Story values used in the embedded content queries are read from story memory and inserted as query terms. The results of these queries (a list of URLs to content items) are checked for consistency/availability. If they are available, they are inserted in the beat description that is now finalized. If the beat cannot be finalized, a new one is chosen (step 3). The triggers and story values that the AN engine should react on when the ambient browser fires them are also identified at this stage (for step 5). The AN engine now reads the action part of the toothbrushing beat and retrieves the URL of today's cartoon from the server and inserts it into the SMIL template that is now a final form SMIL document.

In step 5 (update session) the story memory and user profiles active in this ambient narrative session are updated. Management of the session also takes place in this step.

The final step (step 6: set finalized beat) sets the triggers to which the AN engine will listen and sends the finalized beat to the ambient browser that updates the presentation and the triggers it should listen to. The engine is now again waiting in this session for events from the ambient browser (step 1). In the media-enhanced toothbrushing example, the final SMIL document describing the interactive cartoon presentation is now forwarded to the ambient browser and the initialization part of the beat description is executed, setting the 'toothbrush' story value in the story memory. The AN engine then sets the triggers that can respond to the beat queries and beat links specified in the SMIL document.

The design of the ambient narrative engine is inspired by the drama manager in the Facade system. This drama manager also uses user input and a story memory to sequence beats into an immersive interactive story. The difference with our approach is that we can not only adapt the link structure based on user interaction but also the nodes to some extent because we allow embedded queries in the beat description language. Similar concepts are also used by Magerko [27] and Szilas [41]: Margerko uses the word director while Szilas applies the term narrator to refer to a drama manager.

5.3. Within-component Layer: Plot Material

Depending on whether you look at the narrative level or at the hypermedia level, the SMIL document is either an atomic component or a composite component. In the first case the component content in the within-component layer are SMIL documents, in the second case the component content are media objects, text, images, audio, video and other modalities such as lightscripts. All the component information and content is indexed and stored in an ambient narrative database for fast retrieval by the ambient narrative engine.

6. Current Implementation Status

The work described above is still very much work in progress. The ambient browser is based on a SMIL engine we have developed in-house that implements the toplayout element. This allows us to describe simple interactive multimedia presentations that involve multiple devices. Examples are hypervideo documentaries where additional information is presented on a separate web tablet or a music playlist navigation where the audio is streamed over a networked audio device and the playlist is shown on a remote control display. The Ambient Narrative engine is still in the design stage. We are looking at AI and IR techniques to refine the beat language model and beat sequencing algorithm.

7. Conclusions

Ambient Intelligence is a vision on the future of consumer electronics that refers to the presence of a digital environment that is sensitive, adaptive and responsive to the presence of people. Since it is technologically not possible to mass produce Ambient Intelligence with the current state of the art in artificial intelligence and economically not feasible to manually develop tailor-made Ambient Intelligence products or services for each customer individually, we believe a different approach is needed to move Ambient Intelligence out of research laboratories into the real world. In this paper we described a mass customization strategy for Ambient Intelligence services offered over a collection of networked devices to customize Ambient Intelligence on a mass basis. In this approach modular parts of Ambient Intelligence descriptions are assembled based on user feedback and interaction history into a personalized flow of personalized, interactive media presentations in which multiple devices may participate simultaneously. The Ambient Intelligence service thus allows people to create their own Ambient Intelligence within the scope of possibilities set down by the designer or writer of an ambient narrative, an interactive narrative in mixed reality that is designed to support people at home or in a particular service encounter in performing their everyday activities. We also explained that this mass customization strategy for ubiquitous hypermedia applications can be implemented in the existing Amsterdam Hypermedia Model (AHM) and Adaptive Hypermedia Application Model (AHAM). Although we believe this approach of mass customization in ambient narratives is conceptually sound, and supported by economical and social-cultural drivers, it will need to be implemented and tested in practice. To this end we are working on an example ambient narrative situated in HomeLab, the usability and feasibility lab at Philips Research.

Acknowledgments

We would like to thank in particular Ramon Clout (Philips) and Richard Doornbos (Philips) for developing the SMIL engine. We are grateful to Frank Nack (CWI) for providing useful feedback on an earlier draft of this document.

References

[1] E. Aarts and S. Marzano, editors. The New Everyday: Views on Ambient Intelligence. 010 Publishers, 2003.

[2] B. Arons. Hyperspeech: Navigating in Speech-Only Hypermedia. In UK Conference on Hypertext, pages 133–146, 1991.

[3] J. Austin. How to Do Things With Words. Harvard University Press, 2nd edition, 1975.

[4] J. Bates, A. Loyall, and W. Reilly. Broad Agents. Sigart Bulletin, 2(4):38–40, 1991.

[5] J. Bitner. Servicescapes: The Impact of Physical Surroundings on Customers and Employees. Journal of Marketing, 54(4):69–82, 1992.

[6] P. D. Bra, G.-J. Houben, and H. Wu. AHAM: A Dexter-Based Reference Model for Adaptive Hypermedia. In UK Conference on Hypertext, pages 147–156, 1999.

[7] K. Brooks. Metalinear Cinematic Narrative: Theory, Process and Tool. PhD thesis, MIT Media Lab, 1999.

[8] P. Brusilovsky. Methods and Techniques of Adaptive Hypermedia. User Modeling and User-Adapted Interaction, 6(2-3):87–129, 1996.

[9] M. Castells. The Information Age: Economy, Society and Culture, Vol.I: The Rise of the Network Society. Cambridge MA: Blackwell Publishers, 1996.

[10] M. Cavazza, F. Charles, and S. Mead. Interactive Storytelling: From AI experiment to New Media. In Proceedings of the Second International Conference on Entertainment Computing (ICEC '03), pages 1–8, 2003.

[11] R. Clarke and A. Schmidt. Beyond the Servicescape: The Experience of Place. Journal of Retailing and Consumer Services, 2(3):149–162, 1995.

[12] M. Csikszentmihalyi. Flow: The Psychology of Optimal Experience . Perennial, 1991.

[13] G. DaSilveira, D. Borenstein, and F. Fogliatto. Mass Customization, Literature Review and Research Directions. International Journal of Production Economics, 72(1):1–13, 2001.

[14] G. Davenport and M. Murtaugh. Automatist Storyteller Systems and the Shifting Sands of Story . IBM Systems Journal, 36(3):446–456, 1997.

[15] J. Derrida. Differance (trans. by Alan Brass: Writing and Di_erence). University of Chicago Press, 1980.

[16] J. Falk. Funology: From Usability to Enjoyment, chapter Interfacing the Narrative Experience. Kluwer Academic Publishers, 2003.

[17] R. Fisk and S. Grove. The Service Experience as Theater. Advances in

Consumer Research, 19:455–461, 1992.

[18] T. Galyean. Narrative Guidance of Interactivity. PhD thesis, MIT Media Lab, 1995.

[19] E. Goffman. The Presentation of Self in Everyday Life. Doubleday: Garden City, 1959.

[20] K. Gronbaek, P. Vestergaard, and P. Orbaek. Towards Geo-Spatial Hypermedia: Concepts and Prototype Implementation. In Proceedings of the 13th Conference on Hypertext and Hypermedia, Maryland, USA, 2002.

[21] F. Halasz and M. Schwartz. The Dexter Hypertext Reference Model. Communications of the ACM, 37(2):30–39, February 1994.

[22] L. Hardman, D. Bulterman, and G. V. Rossum. The Amsterdam Hypermedia Model: Adding Time and Context to the Dexter Model. Communications of the ACM, 37(2):50–64, February 1994.

[23] R. Hightower, M. Brady, and T. Baker. Investigating the Role of the Physical Environment in Hedonic Service Consumption. Journal of Business Research, 55(9):697–707, 2002.

[24] C. Kray, A. Kruger, and C. Endres. Some Issues on Presentations in Intelligent Environments. In First European Symposium on Ambient Intelligence (EUSAI), 2003.

[25] B. Laurel. Toward the Design of a Computer-Based Interactive Fantasy System. PhD thesis, Ohio State University, 1986.

[26] B. Laurel. Computers as Theatre. Addison-Wesley, 1993.

[27] B. Magerko. A Proposal for an Interactive Drama Architecture. In AAAI 2002 Spring Symposium Series: Artificial Intelligence and Interactive Entertainment, 2002.

[28] M. Mateas. An Oz-Centric Review of Interactive Drama and Believable Agents. Technical report, School of Computer Science, Carnegie Mellon University, Pittsburgh, 1999.

[29] R. McKee. Story: Substance, Structure, Style and The Principles of Screenwriting . Regan Books, 1997.

[30] D. Millard and e.a. Hyperdoc: An Adaptive Narrative System for Dynamic Multimedia Presentations. In Proceedings of the 14th Conference on Hypertext and Hypermedia, 2003.

[31] J. Murray. Hamlet on the Holodeck: The Future of Narrative in Cyberspace. MIT Press, 1998.

[32] J. Pine and J. Gillmore. The Experience Economy. Harvard Business School Press, 1999.

[33] J. Rifkin. The Age of Access: The New Culture of Hypercapitalism, Where all of Life is a Paid-For Experience. J.P. Tarcher, 2001.

[34] L. Romero and N. Correia. HyperReal: A Hypermedia Model for Mixed

Reality. In Proceedings of the 14th ACM Conference on Hypertext and Hypermedia, Nottingham, UK, 2003. ACM.

[35] N. Sawhney, D. Balcom, and I. E. Smith. HyperCafe: Narrative and Aesthetic Properties of Hypervideo. In UK Conference on Hypertext, pages 1–10, 1996.

[36] R. Schank and R. Abelson. Scripts, Plans, Goals, and Understanding. Hillsdale, NJ: Earlbaum Assoc., 1977.

[37] R. Schechner. Performance Studies: An Introduction. Routledge: New York, 2002.

[38] J. Searle. Speech Acts : An Essay in the Philosophy of Language. Cambridge University Press, 1969.

[39] P. Sinclair, K. Martinez, D. Millard, and M.Weal. Links in the Palm of your Hand: Tangible Hypermedia using Augmented Reality. In Proceedings of the 13th Conference on Hypertext and Hypermedia, Maryland, USA, 2002.

[40] Synchronized Multimedia Integration Language (SMIL) W3C standard. http://www.w3.org/AudioVideo/.

[41] N. Szilas. Interactive Drama on the Computer: Beyond Linear Narrative. In AAAI 1999 Fall Symposium on Narrative Intelligence, 1999.

[42] M. Weal, D. Michaelides, M. Tompson, and D. D. Roure. The Ambient Wood Journals - Replaying the Experience . In Proceedings of Hypertext ‘03, Nottingham, UK, 2003.

