

Integración de un agente virtual 3D en un entorno de inteligencia ambiental

Germán Montoro, Manuel García-Herranz,
Pablo A. Haya, Xavier Alamán, Daniel
Brande

Departamento de Ingeniería Informática
Universidad Autónoma de Madrid
German.montoro@uam.es

Sandra Baldassarri, Eva Cerezo,
Francisco J. Seron

Departamento de Ingeniería de Sistemas e
Informática
Instituto de Investigación en Ingeniería de Aragón
(I3A)
Universidad de Zaragoza
{sandra, ecerezo, seron}@unizar.es

Resumen

En esta comunicación presentamos el resultado de la integración de un agente animado virtual con un entorno de inteligencia ambiental. El agente animado proporciona información de contexto del entorno y se incorpora dentro de la interfaz de interacción oral del mismo. Para su integración se ha empleado un sistema de definición de reglas que permite la rápida y sencilla incorporación del agente dentro del entorno.

1. Introducción

La inteligencia ambiental proporciona “entornos altamente interactivos que usan computación embebida para observar y participar en los asuntos cotidianos del mundo que les rodea” [1].

Estos entornos inteligentes interactúan con el individuo y le ayudan de forma natural en la realización de las tareas cotidianas. Los ordenadores dentro de un entorno inteligente suelen quedar ocultos para el usuario, los servicios del sistema se obtienen mediante una interacción sensible al contexto. De este modo, se consigue que las habitaciones, oficinas y otros espacios habitables tengan una entidad propia, puedan tomar la iniciativa en la interacción y mejoren la calidad de vida de sus ocupantes.

Los usuarios a los que van dirigidos abarcan el más amplio abanico. Su aceptación pasa porque un usuario, en cualquier momento, pueda encontrarse dentro de un entorno inteligente sin tener que apreciar diferencias sustanciales, o al menos sin que suponga una desventaja, con respecto a entornos convencionales. Por ejemplo,

un campo donde su aplicación es de especial interés es el de las personas mayores, con discapacidades o enfermedades.

Estos entornos proporcionan nuevas posibilidades de interacción [2], ofreciendo nuevos retos a los diseñadores de las interfaces [3].

La aparición de estos entornos hace necesario construir nuevas interfaces que permitan interactuar con los usuarios o proporcionarles información de forma natural. Estas interfaces pueden ser capaces de entablar conversaciones o expresar datos relativos al entorno, sus elementos y los servicios que puede proporcionar a sus usuarios. Las interfaces se han de adaptar a estos sistemas [4], de modo que puedan hacer frente a los nuevos retos que proporcionan.

Dentro del proceso de interacción natural con el entorno, se ha demostrado que el uso de un agente animado virtual ayuda a seguir las instrucciones de forma más sencilla. El empleo de expresiones faciales aporta nueva información al proceso de comunicación y la interacción con un agente virtual que expresa emociones resulta más agradable [5].

En esta comunicación presentamos los resultados en el proceso de integración de un agente virtual animado con un entorno inteligente.

El entorno inteligente ha sido desarrollado dentro del departamento de ingeniería informática de la Universidad Autónoma de Madrid y el agente virtual animado en el departamento de ingeniería de sistemas e informática de la Universidad de Zaragoza.

El resto de la comunicación se distribuye de la siguiente manera. En la siguiente sección se presenta el agente animado virtual. A continuación, en el capítulo 3, se describe el

entorno inteligente empleado. El capítulo 4 presenta el sistema de reglas utilizado para controlar el agente. El apartado 5 muestra cómo se ha integrado este agente dentro del entorno. Por último, el capítulo 6 presenta las conclusiones y el trabajo futuro.

2. Maxine, el agente animado

Maxine [9], desarrollado en la Universidad de Zaragoza, es un motor gráfico para la creación y gestión de entornos virtuales tridimensionales. Dichos entornos pueden incluir actores sintéticos con la capacidad, entre otras, de síntesis de voz y de animación labial sincronizada, lo que los hace especialmente útiles para su uso como interfaces naturales.

Programado íntegramente en C++, está construido sobre un amplio conjunto de librerías de código abierto (o libre). La mayor parte de la aplicación es independiente de la plataforma, sin embargo, la síntesis de voz (basada en tecnología propietaria de Microsoft) hace que la versión actual de Maxine funcione en entorno Windows.

El control de la aplicación se lleva a cabo mediante scripting. Pueden ejecutarse comandos desde ficheros, scripts cargados al inicio de la aplicación o durante la misma, o bien ser introducidos a través de una consola de texto en pantalla.

Maxine es un motor para la creación, organización, manipulación y animación de diferentes elementos que constituyen un escenario virtual (ver figura 1), a través de un grafo de escena. Los siguientes elementos pueden ser representados en un grafo de escena:

- Imágenes y textos, que pueden ser mostrados y posicionados como rectángulos orientados en el espacio. Se permiten los principales formatos gráficos (bmp, gif, jpeg, pic, png, rgb, tga, tiff) y se soporta el canal alpha.
- Primitivas geométricas simples pero también modelos geométricos complejos. Maxine importa la mayoría de los formatos populares (3DS, flt, lwo, md2, obj, osg), permitiendo crear escenarios virtuales complejos con alto nivel de detalle.
- Luces simples.
- Sonido de ambiente y 3D.
- Personajes animados. Usan el formato de la librería de animación Cal3D, pero pueden ser

generados mediante herramientas de modelado y animación 3D comerciales. Se proporcionan diferentes tipos de animaciones, incluyendo animación secundaria para incrementar el realismo y la expresividad de los actores.

- Actores animados. Son una especialización de los anteriores y están provistos de síntesis de voz y animación facial con sincronización labial, siguiendo el estándar VHML y la especificación MPEG4 para la animación facial.
- Voces sintéticas. Las voces se pueden asociar a un elemento. Se pueden configurar y controlar ciertas características de audio como la selección de una voz concreta, el control de volumen, la velocidad del habla, inserción de pausas, el tono, la enfatización de las palabras, especificación de la pronunciación, etc.



Figura 1. Entorno y agente virtual 3D gestionado por Maxine

3. El entorno inteligente

El entorno inteligente, desarrollado en la Universidad Autónoma de Madrid, emplea una red domótica para el control de dispositivos del entorno y una red Ethernet para el control del flujo de información multimedia. El bus domótico utilizado es el bus EIB (European Installation Bus), actualmente denominado KNX. Tal y como considera necesario [6] el entorno posee cámaras, micrófonos y altavoces que transmiten y reciben la información que circula a través de estas redes.

Las entidades que forman parte del entorno se definen en un documento XML que sirve de base

para la creación automática de uno de sus elementos fundamentales, la pizarra, que sirve como un modelo del mundo. Las entidades se pueden añadir o eliminar del entorno de forma dinámica. Esta incorporación automática y dinámica de dispositivos al entorno y el empleo de un modelo del entorno coincide con las ideas expuestas por [7].

La pizarra actúa como un middleware centralizado que encapsula la interacción de las aplicaciones e interfaces de alto nivel con los elementos físicos del entorno, como se especifica en [8]. Además, sirve como un modelo del mundo que proporciona información de contexto donde cualquier aplicación puede acudir para conocer el estado actual del entorno.

Basándose en la información de contexto almacenada, el sistema es capaz de determinar cuál es el estado del entorno, quién se encuentra dentro del mismo y, en algunos casos, qué están haciendo sus ocupantes, tal y como propone [6].

Esta información se encuentra continuamente actualizada de acuerdo con la situación actual del entorno.

Por encima de la capa de middleware se han desarrollado un conjunto de aplicaciones que utilizan la información de contexto almacenada en la pizarra para realizar acciones dentro del entorno.



Figura 2. Imagen del entorno inteligente desarrollado

Para modificar el estado de alguno de los elementos del entorno bastará con enviar mensajes estándar de alto nivel a la pizarra. A su vez, para obtener el estado de algún elemento, basta con realizar una solicitud a la misma.

Los elementos físicos que se pueden controlar en un entorno inteligente se pueden clasificar en dispositivos de control y dispositivos multimedia. Entre los dispositivos de control se encuentran los sensores (de presencia, temperatura, etc.) y los actuadores (en luces, puertas, electrodomésticos, etc.). Los dispositivos multimedia engloban a radios, televisores, pantallas de información, micrófonos, altavoces, etc.

El entorno desarrollado se utiliza como sala de reuniones y banco de pruebas para el desarrollo de nuevas aplicaciones. Las aplicaciones desarrolladas se encuentran en continuo funcionamiento y están habilitadas para su uso generalizado, de modo que se puedan ejecutar y probar en ambientes reales. Una imagen de este entorno se puede encontrar en la figura 2.

El entorno ha sido provisto de un lenguaje basado en reglas que posibilita la adaptación y automatización del mismo a las necesidades del usuario para el contexto dado. A partir de este lenguaje se ha desarrollado un conjunto de agentes configurables capaz de adaptarse al entorno y a los hábitos de sus usuarios.

4. El lenguaje de reglas

La regla es la unidad mínima de este lenguaje. Resulta una pieza fundamental del conocimiento que encapsula una preferencia del usuario o una inferencia contextual, esto es, una reacción deseada en respuesta a cierta situación emergente o bien información contextual de alto nivel inferida a partir de esa misma situación.

Este nuevo lenguaje basado en reglas facilita considerablemente la tarea de trasladar el sistema al terreno de interfaces más complejas como por ejemplo, la interacción mediante lenguaje natural, permitiendo una traducción mucho más sencilla de un dominio al otro.

Cada regla está dividida en tres partes: detonantes, conclusiones y acción. De esta manera una regla queda definida por la gramática mostrada en la figura 3.

```

<regla>:= <lista detonantes> :: <lista condiciones> -> <acción>

<lista detonantes> := <detonante> | <detonante> ; <lista detonantes>

<detonante> := <propiedad> | <relación>

<lista condiciones> := <condición> | <condición> ; <lista condiciones>

<condición> := <elemento> <operador> <elemento>

<elemento> := <entidad> | <propiedad> | <relación> | <valor>

<operador> := = | != | > | <

<acción> := <elemento> <operación> <elemento>

<operación> := := | => | += | -= | => | <=

```

Figura 3. Gramática de una regla del lenguaje del sistema

Tan importante como estas unidades mínimas o reglas es el agente. De esta manera un agente está constituido por un conjunto de reglas. En base a ellas, y a los cambios producidos en el contexto, reacciona produciendo una salida en forma órdenes o cambios que realizar sobre ese mismo contexto. La forma en que reacciona a esos cambios, las consecuencias internas que conllevan así como el proceso de inferencia que detonan puede representarse en forma de algoritmo.

El algoritmo de ejecución del agente consta de tres pasos básicos: educación, actualización y detonación que se encadenan en respuesta a un cambio en el entorno.

La educación o aprendizaje consiste en utilizar la información correspondiente al cambio producido en el contexto para potenciar o castigar las reglas que se ejecutaron de manera correcta o errónea respectivamente.

La actualización consiste en calcular el nuevo valor –verdadero o falso- de todas las condiciones que hagan referencia al elemento contextual objeto del cambio. Esta política permite tener la información que el agente necesita para la toma de decisiones siempre actualizada. De esta manera, al actualizar el valor de las condiciones a medida que se producen los cambios en el contexto se distribuye el esfuerzo computacional a lo largo del tiempo, agilizando el comportamiento del sistema

cuando es más necesario, esto es, a la hora de tomar una decisión.

Por último, la detonación consiste en, dadas las reglas que como detonante incluyen el elemento objeto de cambio en el contexto, comprobar que todas sus condiciones evalúen a verdadero y, en su caso, enviar la acción asociada a la pizarra.

Este sistema permite añadir comportamientos al entorno, configurar el comportamiento de dispositivos o describir contexto de alto nivel a partir de bajo nivel de forma altamente natural y en poco tiempo. Un ejemplo sencillo de aplicación de una regla se muestra en la figura 4, donde se gestionan las luces en función del encendido y apagado de la televisión. En este caso sólo hay un detonante que es la televisión. Las condiciones sirven para determinar si la televisión se enciende o se apaga y controlar a su vez las luces del entorno.

```

#Regla 1: cuando se enciende la TV se apaga la luz del techo
tv:eib6:status :: tv:eib6:status = 1 -> light:lamp_1:status := 0

# Regla 2: cuando se apaga la TV se enciende la luz del techo
tv:eib6:status :: tv:eib6:status = 0 -> light:lamp_1:status := 1

# Regla 3: cuando se enciende la TV, si la lámpara regulable esta alta se
baja a media intensidad
tv:eib6:status :: tv:eib6:status = 1 ; dimmablelight:lampv1:value > 20 ->
dimmablelight:lampv1:value := 20

# Regla 4: cuando se apaga la TV, si la lámpara regulable esta media se
sube un poco
tv:eib6:status :: tv:eib6:status = 0 ; dimmablelight:lampv1:value = 20 ->
dimmablelight:lampv1:value := 35

```

Figura 4. Ejemplo de reglas aplicadas a la televisión y las luces

5. Integración del agente 3D en el entorno

Basándose en la funcionalidad que proporcionan las reglas descritas se ha integrado el agente virtual con el entorno inteligente anteriormente explicado.

De esta forma, el entorno se beneficia de las características de interacción con un agente virtual 3D, proporcionando una nueva dimensión a los hechos que ocurren en el mismo.

Esta aproximación se ha realizado de dos formas diferentes:

- Creando una entidad Maxine en el entorno, a la que pueden acceder todos los elementos del mismo, así como el sistema de reglas.
- Permitiendo asociar un agente Maxine personalizado a elementos específicos del entorno de forma exclusiva.

La comunicación con el agente Maxine se realiza de forma sencilla y estándar, empleando los mecanismos de comunicación del entorno proporcionados por la capa de middleware.

Esta capa está provista de un mecanismo para encapsular los procesos de comunicación entre los agentes del entorno. Así, cualquier elemento o aplicación del entorno que quiera utilizar Maxine sólo tiene que enviar un mensaje a la pizarra identificando al destinatario como la entidad "Maxine" del entorno. Esta entidad Maxine está suscrita a los mensajes de la pizarra que van dirigidos a ella, de modo que es capaz de leerlos y emitirlos para que puedan ser recibidos por los usuarios.

Además de poder enviar mensajes al agente Maxine, también es posible cambiar su estado de ánimo (Figura 5). Los estados de ánimo que puede adaptar Maxine son: "anger" (enfadada), "disgust" (asqueada), "fear" (temerosa), "joy" (feliz), "sadness" (triste), "surprise" (sorprendida) y "neutral" (neutral). Estos estados también se pueden modelar en intensidad, asignándoles un valor entre 0 y 1.0.



Figura 5. Expresiones del agente virtual 3D

Basándose en estas características y empleando el sistema de reglas descrito en la sección anterior, se han definido una serie de comportamientos específicos para Maxine que le permitirán actuar dependiendo del contexto actual del entorno.

Por ejemplo, Maxine saluda y se despide de forma personalizada a las personas que acceden o abandonan al entorno. Una transcripción de una de las reglas necesarias para realizar esta acción se encuentra en la figura 6.

```

person:Manuel:esta ::
person:Manuel:esta = 1 ->
person:maxine:say := Hola Manu

person:Manuel:esta ::
person:Manuel:esta = 0 ->
person:maxine:say := Adiós Manu

```

Figura 6. Reglas empleadas para que Maxine salude a un usuario del entorno

Como se puede observar, y se explicó en la sección anterior, cada una de las dos reglas representadas en este ejemplo se compone de tres partes (una por cada línea): el detonante, la conclusión y la acción.

Por ejemplo, para la primera regla, el detonante sería que cambiara el estado en el middleware de la propiedad “esta” de la entidad “Manuel”, que es de tipo “person”. Si se concluye que el estado de esta propiedad cambia para tomar el valor 1 (lo que significaría que Manuel ha entrado en la habitación) se realizaría la acción de enviar el mensaje “Hola Manu” a la propiedad “say” de la entidad “Maxine”, que es de tipo “person”.

Este sencillo mecanismo hace posible que el agente animado actúe de forma activa cuando se produce un cambio en el contexto del entorno, permitiendo establecer una nueva forma de comunicación con los usuarios del mismo.

A su vez, se están utilizando los diferentes estados de ánimo que puede presentar el agente animado para probar hasta qué punto incide esta información visual en los usuarios que la reciben. Como muestra, Maxine puede mostrar disgusto cuando ocurre en el entorno algo que no le parece adecuado. Ejemplos de estos casos serían que alguien subiera el volumen de la música si se está hablando por teléfono o que la temperatura del entorno superara un umbral determinado. Un ejemplo de las reglas necesarias para implementar este último caso se muestra en la figura 7.

Un caso específico de integración de Maxine con un módulo exclusivo de la pizarra es el de la

interfaz de interacción con el entorno mediante diálogos orales.

```

room:lab_b403:temperature ::
room:lab_b403:temperature > 25 ->
person:maxine:say := Uf qué calor

room:lab_b403:temperature ::
room:lab_b403:temperature > 25 ->
person:maxine:emotion := DISGUST

```

Figura 7. Ejemplo de reglas donde Maxine muestra un estado de disgusto

El entorno dispone de una interfaz en lenguaje natural para interactuar con los dispositivos que lo componen [10]. Esta interacción se adapta de forma automática a las características del entorno y permite comunicarse con las entidades del mismo y modificar su estado de forma natural. La interacción se realiza con las lámparas del entorno, entre las que se encuentran luces halógenas regulables, focos de lectura y fluorescentes, el sintonizador de radio, la puerta de entrada, etc.

Antes de integrar Maxine en el entorno, la comunicación con el usuario se realizaba mediante un módulo de síntesis de voz y no se proporcionaba ningún tipo de información visual. Ahora también es posible que sea Maxine quien se ocupe de comunicar las respuestas a los usuarios. De esta manera las personas que interactúan con el entorno no tienen la sensación de estar hablando a una entidad abstracta sino que establecen una comunicación con un agente que responde y se muestra de forma adecuada para cada caso.

6. Conclusiones y trabajo futuro

En esta comunicación hemos presentado los primeros resultados en la integración de un agente virtual animado en tres dimensiones, desarrollado en la Universidad de Zaragoza, con un entorno de inteligencia ambiental, implementado en la Universidad Autónoma de Madrid.

En ambos casos se partía de elementos ya definidos y utilizados de forma independiente en cada una de estas universidades, con el fin de comprobar la facilidad de integración de ambos componentes y la utilidad de dicha combinación.

El agente animado se ha utilizado en dos modalidades: para suministrar información de contexto del entorno y como interfaz de interacción con los dispositivos del mismo.

En el primer caso se ha empleado el sistema de reglas descrito para el entorno inteligente como mecanismo de actuación del agente ante los cambios de contexto producidos en el entorno. Se ha comprobado que su integración y utilización resulta rápida y sencilla.

En el segundo caso se ha añadido la interfaz proporcionada por el agente al módulo de interacción oral con el entorno. Esta nueva característica busca mejorar la comunicación con los usuarios proporcionando un agente a quien dirigir y de quien recibir los mensajes.

El trabajo de integración todavía se encuentra en sus primeras fases por lo que todavía hay un amplio campo de pruebas y experimentos que realizar.

Se ha de seguir aumentando el número de funcionalidades que presenta Maxine dentro del entorno, estudiando las nuevas posibilidades que ofrece a un espacio de estas características.

Resultaría de especial interés comprobar con usuarios reales la utilidad que presentan los mensajes contextuales proporcionados así como la utilización de diferentes estados de ánimo.

Además, se deberían hacer pruebas a la interfaz oral contrastando su uso original con el de la interacción con un agente animado, comprobando el nivel de mejora de su utilización y el de satisfacción de los usuarios.

Por último, se podría comprobar si se aprecian diferencias en el uso del agente animado entre diferentes tipos de usuarios, por ejemplo, entre personas mayores o poco habituadas a trabajar con tecnología.

Agradecimientos

Este trabajo ha sido parcialmente financiado por el Gobierno de Aragón gracias al convenio WALQA (ref. 2004/04/86) y al proyecto N° CTPP02/2006 así como por la Dirección General de Investigación a través de los proyectos TIN2004-07926 y TIN2004-03140.

Referencias

- [1] Coen, M.H. 1998. Design Principles for Intelligent Environments. In Proceedings of the AAAI Spring Symposium on Intelligent Environments (AAAI98).
- [2] Weiser, M. The world is not a desktop. *ACM Interactions*, 1, 1, 7-8, 1994.
- [3] Shafer, S., Brumitt, B., and Cadiz, J.J. Interaction Issues in Context-Aware Intelligent Environments. *Human-Computer Interaction*, 16, 363-378, 2001.
- [4] Paternò, F., and Santoro, C. One Model, Many Interfaces. In Proceedings of CADUI, 2002.
- [5] Ortiz, A.; Carretero, M.; Oyarzun, D.; Yanguas, J.J.; Buiza, C.; González, M.F.; Etxeberria, I.. Elderly Users in Ambient Intelligence: Does an Avatar Improve the Interaction? Proceedings of 9th ERCIM Workshop "User Interfaces For All", 2006
- [6] Pentland, A. 1996. Smart rooms. *Scientific American*, 274, 4, 68-76.
- [7] Shafer, S.A.N. 1999. Ten dimensions of ubiquitous computing. In Proceedings of MANSE'99, Dublin, Dec 1999.
- [8] Johanson, B.; Fox, A. and Winograd, T. 2002. The Interactive Workspaces Project: Experiences with Ubiquitous Computing Rooms. *IEEE Pervasive Computing Magazine*, 1(2).
- [9] Seron F., Baldassarri S., Cerezo. MaxinePPT: Using 3D Virtual Characters for Natural Interaction. E. Proc. 2nd International Workshop on Ubiquitous Computing & Ambient Intelligence, pp. 241-250, 2006
- [10] Montoro, G.; Haya, P.A. and Alamán, X. 2004. Context adaptive interaction with an automatically created spoken interface for intelligent environments. The 2004 IFIP International Conference on Intelligence in Communication Systems (INTELLCOMM 04), Bangkok, Thailand. November, 2004. Pag. 120-127.