

Desarrollo de un guía virtual conversacional para museos en un entorno inmersivo de realidad virtual y aumentada

Isidro Hernán Martínez Juárez¹, María Eugenia Puga Nathal¹,
Juan Martínez-Miranda²

¹ Tecnológico Nacional de México, Instituto Tecnológico de Ciudad Guzmán,
Ciudad Guzmán, Jalisco, México

² CONACyT – Centro de Investigación Científica y Educación Superior de Ensenada,
Unidad de Transferencia Tecnológica Tepic (CICESE-UT3), Tepic, Nayarit, México
{isidroM18290017, mepuga}@itcg.edu.mx
jmiranda@cicese.mx

Resumen. El uso de agentes virtuales conversacionales, como un tipo de interfaces inteligentes de interacción, son cada vez más utilizados en diferentes áreas de aplicación. Una de estas áreas es el sector turístico, en el cual estos agentes actúan como guías turísticos para proveer información más detallada y personalizada acorde a los intereses de los turistas. En este artículo se describen las etapas iniciales del desarrollo de un agente virtual conversacional el cual actuará como guía de visitantes a un museo. En concreto, se presenta el diseño del mecanismo para generar el comportamiento emocional en el agente a partir de la interacción con el usuario. Una diferencia con trabajos similares es que el guía virtual propuesto estará ubicado en un entorno inmersivo de realidad virtual y aumentada. Esto permitirá evaluar si la interacción en este tipo de entornos mejora la experiencia de interacción con el usuario respecto a agentes virtuales ubicados en interfaces de usuario tradicionales.

Palabras clave: agente virtual, realidad virtual y aumentada, modelo cognitivo-emocional, sistemas conversacionales.

Development of a Conversational Virtual Guide for Museums in an Immersive Environment of Virtual and Augmented Reality

Abstract. The use of conversational virtual agents, as a type of intelligent interaction interfaces, is increasingly in different areas of application. One of these areas is the tourism sector, where these agents act as touristic guides able to provide more detailed and personalized information according to the interests of the tourists. This article describes the initial stages in the development of a conversational virtual agent that will act as a guide for the visitors in a museum. In particular, the design of the mechanism to generate the emotional behavior in the virtual agent during the interaction with the user is presented. A difference with similar works is that the proposed virtual guide will be located in an immersive environment of virtual and augmented reality. This will allow to evaluate if the interaction in this type of environments improves the user's

interaction experience regarding existent virtual agents located in traditional user interfaces.

Keywords: virtual agents, virtual and augmented reality, cognitive emotional model, conversational systems.

1. Introducción

El sector turístico es una de las áreas en las que el uso de las TIC está cambiando la manera de promocionar, ofrecer y proveer servicios por parte de todos los actores relevantes dentro del sector a nivel mundial. Además de las aplicaciones *clásicas* tales como la gestión de reservas en hoteles, vuelos y restaurantes a través de Internet y el uso de las guías de viaje electrónicas, se han desarrollado sistemas inteligentes e interactivos, así como aplicaciones móviles que han ayudado a aumentar la calidad de los servicios turísticos y mejorar las experiencias de los viajeros. El objetivo principal de estas aplicaciones es el atraer y retener a los visitantes de los diferentes destinos que se promocionan en cualquier parte del mundo, así como mejorar la competitividad de las empresas públicas y privadas proveedoras de servicios turísticos.

Una de las tecnologías que desde ya hace algunos años ha comenzado a aplicarse en el sector turístico es la realidad virtual y aumentada la cual ofrece el potencial de mejorar la experiencia de los turistas mediante el modelado de los alrededores del usuario en un espacio tridimensional y proveer información complementaria sobre un determinado contexto (ej. un punto de interés turístico) con apariencia de coexistir con el mundo real del turista. Además de la información que se pueda modelar como parte del ambiente en estos entornos inmersivos, un componente capaz de mantener una interacción personalizada y proporcionar información específica a petición expresa del usuario es el conocido como Agente Virtual Conversacional. Estos agentes virtuales toman la apariencia de una persona y son desarrollados como una herramienta avanzada de interacción hombre-máquina. Para comunicarse con el usuario estos agentes implementan un mecanismo para la generación de diálogos basados en lenguaje natural combinados con expresiones faciales y movimientos corporales para simular una conversación persona a persona [1].

En sus inicios esta tecnología fue principalmente utilizada con objetivos de aprendizaje y/o entrenamiento en actividades concretas [2, 3], así como en el área de videojuegos actuando como personajes autónomos complementarios al jugador (“*non-player characters*”) [4, 5]. Sin embargo, en los últimos años estos agentes han mostrado su utilidad y beneficios en una gran área de aplicaciones, tales como sistemas interactivos dentro del sector salud [6–9], turístico [10, 11], e incluso como asistentes personales o acompañantes para adultos mayores [12]. Cuando el comportamiento y la apariencia de estos agentes son generados de la manera correcta, el agente virtual puede establecer un buen nivel de confianza con el usuario maximizando su utilidad y consecución de los objetivos para los cuales fueron creados [13].

Diferentes líneas de investigación son necesarias para el desarrollo de estos agentes virtuales para lograr el objetivo de generar una apariencia y comportamientos coherentes y acordes a los objetivos de la aplicación. Estas áreas de investigación incluyen el modelado y animación por computadora para hacer que el personaje gráfico se vea y mueva de una manera realista; técnicas de inteligencia artificial y computación

afectiva para crear modelos artificiales capaces de generar un comportamiento cognitivo y emocional que ayude a generar una mayor aceptabilidad por parte de los usuarios; o las áreas de psicología, sociología y ciencias cognitivas como bases teóricas para la construcción de estos modelos artificiales [14, 15].

El trabajo presentado en este artículo describe las primeras etapas en el diseño y desarrollo de un agente virtual conversacional con el objetivo de que proporcione información a los visitantes de un museo. Este agente virtual estará ubicado en un entorno virtual e inmersivo y adoptará el rol de un guía capaz de interactuar con los visitantes del museo y ofrecer información relevante sobre el material de exhibición de la sala en la que se encuentre visitando el turista. En particular, este trabajo presenta el diseño interno de este guía virtual mediante el cual se generarán los diálogos y el comportamiento emocional durante la interacción con el usuario. También se presenta el mecanismo de extracción de información para proporcionar datos relevantes al usuario sobre lo que se encuentra visitando dentro del museo.

El principal objetivo de este trabajo es valorar el impacto de esta tecnología en la experiencia turística de los visitantes a un museo. Uno de los objetivos particulares es medir el nivel de aceptabilidad y adherencia de los usuarios hacia un guía virtual inmerso en un entorno de realidad virtual y aumentada respecto a trabajos similares en los que el agente virtual se encuentra ubicado en una interfaz estática. El resto de las secciones del artículo presentan lo siguiente: la sección 2 describe algunos de los trabajos relacionados. La sección 3 presenta el trabajo inicial desarrollado para la construcción del modelo interno del agente, poniendo énfasis en el modelo cognitivo-emocional y en el mecanismo de generación de diálogos incluida la extracción de información sobre el área del museo que el turista se encuentra visitando y proporcionársela al usuario. Finalmente, la sección 4 presenta algunas conclusiones y el trabajo futuro para completar el desarrollo de este trabajo.

2. Trabajo relacionado

En años recientes, algunos museos han comenzado a utilizar agentes virtuales conversacionales como guías de los visitantes a los museos. Un ejemplo es el caso del Museum of Science, Boston (MoS) que en colaboración con el USC Institute for Creative Technologies han desarrollado *Ada y Grace*, un par de agentes mellizos que interactúan con los visitantes del MoS a través del procesamiento de lenguaje natural y usando una apariencia realística para atraer e incrementar el interés del público, principalmente estudiantes de secundaria, en áreas de ciencia y tecnología [16]. Una evaluación hecha a niños de edad entre 7 y 14 años demostró el impacto logrado por *Ada y Grace* en 14 de 15 indicadores en distintas áreas, entre las que se encontró un incremento en la atracción e interés hacia la informática y la tecnología; una actitud positiva hacia estas dos áreas, así como un aumento de su conciencia y conocimientos sobre dichas áreas. Las condiciones que permitieron esta evaluación fueron dos: interacción directa con el visitante y una combinación de la interacción con el personal del museo y el visitante. Los instrumentos de evaluación incluyeron la observación directa, entrevistas a los usuarios y la aplicación de cuestionarios de seguimiento [10].

Otro ejemplo de guía virtual es *Max*, quien interactúa con visitantes del Heinz Nixdorf Museums Forum. *Max* se encuentra dentro de una pantalla con una cámara que

detecta el paso de los visitantes de tal forma que él puede darse cuenta de la presencia de las personas y actuar como su guía con voz sintética en alemán, gestos, movimientos y expresiones faciales lo más naturales posibles. El agente está orientado principalmente a involucrar a los usuarios en conversaciones en las que él les proporciona información relevante del museo, así como de las exhibiciones dentro de él y otros temas de interés, de tal manera que sea comprensible e interesante. La interacción se da a través del procesamiento de lenguaje natural ya sea por voz o bien usando un teclado conectado al sistema [17].

Un tercer ejemplo de guía virtual es MINERVA el cual, a diferencia de los dos anteriores, es empleado en casos donde el usuario no está presente físicamente. MINERVA actúa como un agente “telepresente” que guía a los usuarios a través de diferentes espacios del Smithsonian’s National Museum of American History vía acceso remoto a través de Internet [11].

De manera complementaria, durante los últimos años múltiples proyectos de investigación han comenzado a crear aplicaciones de realidad aumentada (AR por sus siglas en inglés) que utilizan agentes de realidad aumentada, o AuRA (Augmented Reality Agents), como su principal paradigma de interacción y desarrollo [18]. La mayoría de las aplicaciones de realidad aumentada están relacionadas principalmente con permitir que un usuario explore un mundo virtual en 3D mezclado con el mundo real. Las interfaces AR más avanzadas permiten al usuario interactuar con el entorno mixto, pero la parte virtual suele ser bastante finita y determinista. Este tipo de aplicaciones pueden hacer uso de “agentes de realidad aumentada” los cuales son un caso particular de agentes virtuales. Su punto fuerte es que pueden apelar a múltiples sentidos e involucrar a un usuario en una conversación aún más realista y similar a la que se da de persona a persona [19].

3. Diseño del mecanismo interno del agente virtual

3.1. Generación del comportamiento

El trabajo llevado a cabo para el desarrollo del guía virtual conversacional se ha concentrado durante las primeras fases en obtener la arquitectura interna que generará el comportamiento del agente durante la interacción con el usuario. Basado en trabajos previos en los cuales se ha demostrado que un comportamiento consistente, coherente y empático en el agente virtual ayuda a promover una mayor credibilidad, aceptabilidad y nivel de confianza en los usuarios [14, 20], se ha trabajado en el mecanismo interno para generar un comportamiento de este estilo.

Actualmente existen diferentes arquitecturas computacionales que pueden ser integradas como el mecanismo interno para la generación de un comportamiento cognitivo-emocional tanto en agentes virtuales como en robots sociales [15]. Una de las teorías en la que la mayoría de estas arquitecturas se basan para implementar el mecanismo de generación de emociones es la conocida como teoría de valoración cognitiva [21, 22]. De manera general, esta teoría postula que los eventos que se producen en el entorno de una persona son constantemente (consciente o inconscientemente) evaluados acorde a las metas, creencias y preferencias del individuo. Esta evaluación *cognitiva* genera una respuesta emocional (de acuerdo con

la relevancia que el individuo asigna al evento ocurrido), lo cual a su vez genera un comportamiento específico para hacer frente a dicho evento.

Basado en esa teoría, el modelo interno de nuestro guía virtual (ver Fig. 1) implementa un mecanismo de valoración de los eventos ocurridos en su entorno. En nuestro caso, estos eventos son las acciones y diálogos generados por el usuario durante su recorrido por el museo y la interacción que mantiene con el agente. La valoración sobre el evento detectado puede producir una emoción (positiva o negativa) con una intensidad determinada acorde a la importancia que le asigna el agente al evento ocurrido. El conjunto de emociones que se generan en el agente está basado en el modelo OCC [23], el cual describe una organización jerárquica de 22 emociones a partir del proceso de valoración cognitiva respecto a las metas internas del agente virtual.

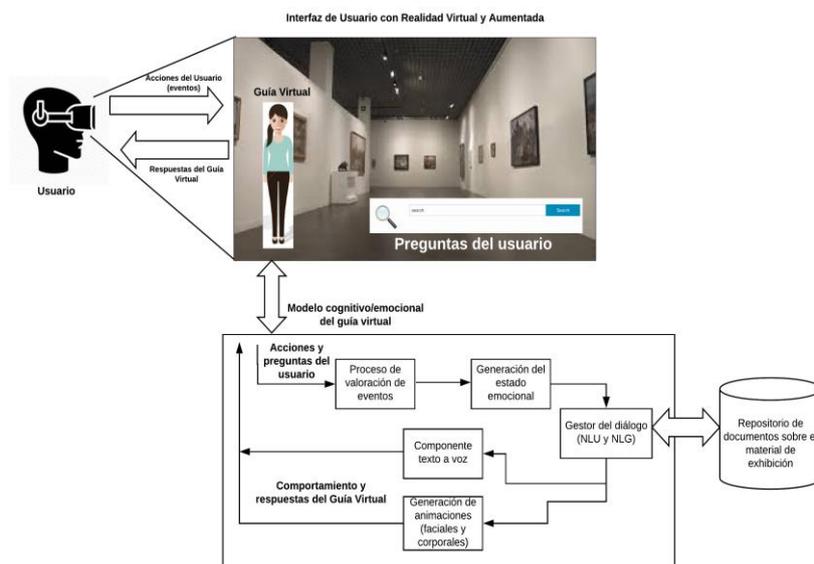


Fig. 1. Mecanismo Interno del Agente Virtual.

El estado emocional producido en el agente es utilizado por el componente gestor del diálogo para construir las oraciones –basadas en lenguaje natural– que se utilizarán como respuesta a las acciones del usuario. Una vez construido el diálogo, esta información y el estado emocional generado en el agente son procesados por el componente de texto a voz, y por el componente que genera las animaciones de expresiones faciales y movimientos corporales para reproducir las respuestas verbales y no verbales del agente hacia el usuario.

Para ejemplificar el mecanismo completo, considérese el siguiente escenario: el usuario visitante al museo recorre las piezas exhibidas en una sala y a través del componente de realidad aumentada se detecta que el usuario está observando un objeto determinado. Esta acción (evento para el agente que puede ser etiquetado como *objeto_atencion_usuario*) es identificada, la cual es evaluada por el componente de

valoración de eventos. Una de las metas internas pre-definidas del agente puede ser *ofrecer_informacion_objeto*. En consecuencia, la ocurrencia de dicho evento puede ser valorada positivamente generando una emoción positiva y activando dicha meta. La manera de llevar a cabo esta meta es a través de una determinada acción, por ejemplo, *confirmar_interes_objeto*.

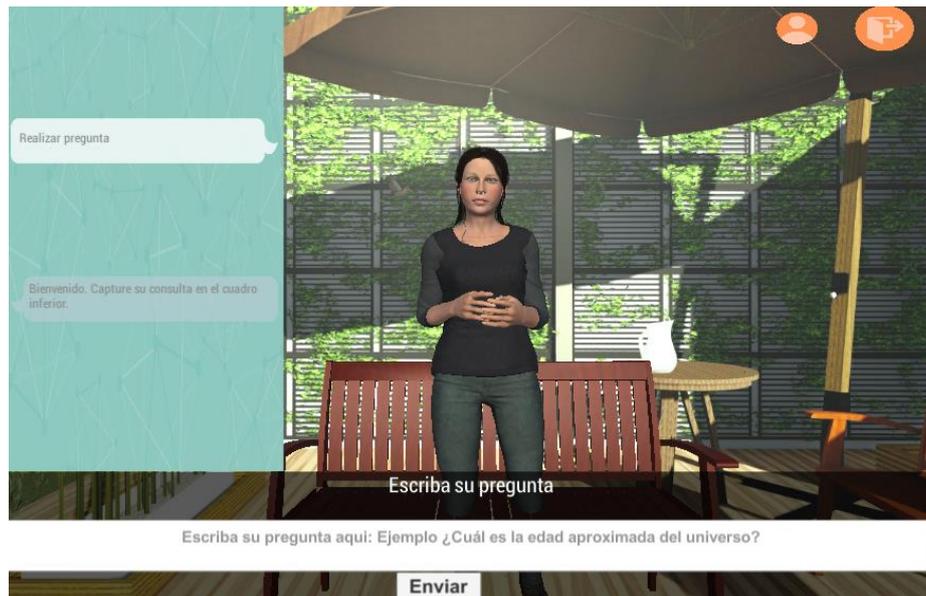


Fig. 2. Ejemplo de guía virtual dando la bienvenida al usuario.

Esta acción se puede considerar un acto de diálogo, el cual es transformado por el gestor del diálogo en una frase concreta construida a partir de la emoción generada en el agente, y un conjunto de frases pre-definidas para implementar los diferentes actos de diálogos que pueden ocurrir durante la interacción.

En este caso, el acto de diálogo *confirmar_interes_objeto* junto a la emoción positiva del agente puede generar tanto el comportamiento verbal (ej. “Me alegra que te interese ese objeto! –*parte emocional*– ¿Quieres saber más sobre él? –acción para confirmar el interés del usuario en el objeto–”); y no verbal (ej. el agente muestra una sonrisa) para dirigirse al usuario. La respuesta del usuario a la pregunta del agente se evaluará como un nuevo evento, iniciando el siguiente ciclo de interacción.

3.2. Extracción de información

Un componente importante adicional al mecanismo de generación del comportamiento del agente virtual es la manera de extraer y proveer la información sobre los objetos exhibidos en el museo. En nuestro caso, el guía virtual puede ofrecer información de un determinado objeto de dos formas: *i*) al detectar que el usuario está observando el objeto; o *ii*) a pregunta expresa del usuario. En cualquiera de las dos formas es necesario implementar el mecanismo de extracción de información sobre

documentos que contengan información sobre el dominio del objeto o pregunta hecha por el usuario. Estos documentos pueden ser los provistos por el propio museo, o incluso se puede extraer información de repositorios abiertos tales como Wikipedia.

Si la consulta de información es generada a partir de una pregunta realizada por el usuario, el sistema gestor de diálogo del agente tendrá integrado un componente de análisis y reconocimiento de lenguaje natural similar al que implementan los sistemas Question and Answering (QA) [24]. Este componente, el cual integrará algoritmos basados en ranking permitirá ir refinando y mejorando las respuestas ofrecidas a los usuarios a partir de recolectar la retroalimentación de los diferentes visitantes sobre la relevancia de la información proporcionada por el agente.

El desarrollo de este guía virtual está en sus primeras fases y de momento se tiene implementado únicamente el mecanismo interno de generación del comportamiento. Este mecanismo genera un conjunto de respuestas básicas para interactuar con el usuario (ver Fig. 2). El diseño y desarrollo del sistema QA para proveer información sobre un objeto y dominio determinado es el siguiente componente que se desarrollará y permitirá hacer una evaluación inicial completa del agente virtual antes de ubicarlo en el entorno inmersivo de realidad virtual y aumentada.

4. Conclusiones y trabajo futuro

El trabajo presentado en este documento describe las etapas iniciales en el desarrollo de un agente virtual que sirva como guía de los visitantes a un museo. Se espera que a través de este agente, los usuarios tengan acceso a una información más personalizada y contextualizada sobre las salas de exhibición que se encuentran visitando. La hipótesis de investigación de este trabajo es que este tipo de interfaces mejorarán la experiencia del usuario, ayudando a incrementar el interés y conocimiento sobre las temáticas expuestas en el museo. De manera concreta, se presentó el diseño y adaptación del componente cognitivo-emocional como el mecanismo interno de generación del comportamiento del agente virtual durante la interacción con el usuario.

El trabajo por desarrollar en las siguientes etapas contempla la implementación del sistema QA para procesar las preguntas concretas del usuario, la extracción de información de documentos sobre un dominio determinado, así como el desarrollo de la interfaz de realidad virtual y aumentada en la cual estará ubicado el guía virtual. De manera paralela se identificarán las métricas de evaluación del sistema, el cual incluirá información objetiva (ej. logs del sistema durante la interacción del usuario con el sistema) y subjetiva (ej. cuestionarios, entrevistas) que permita establecer el nivel de adherencia, usabilidad y aceptabilidad por parte de los usuarios. Finalmente, una vez terminado el desarrollo, el sistema se validará en un museo con un conjunto de visitantes a los cuales se invitará a probar el sistema.

Se prevé que la versión con las que se llevarán a cabo las pruebas de evaluación sea a través de un dispositivo móvil y un equipo de realidad virtual. El equipo de realidad virtual se le proporcionará al usuario durante su recorrido por la(s) sala(s) del museo y consistirá en unos lentes mediante los cuales el usuario experimentará el entorno inmersivo en el cual se encuentre ubicado el guía virtual. Para la evaluación se establecerá un protocolo que permita identificar el tipo de interacción que se llevará a cabo con el usuario a partir de características tales como edad, sexo, preferencias sobre

la apariencia del agente virtual, etc., con el fin de ofrecer sesiones interactivas personalizadas.

A partir de este piloto de evaluación, se podrá valorar el impacto de esta tecnología sobre los usuarios objetivo y se identificarán las ventajas y desventajas de estos guías virtuales en el contexto propuesto y reconocer para qué tipo de usuarios, salas o exhibiciones sería más adecuado su uso.

Referencias

1. Cassell, J., Sullivan, J., Prevost, S., Churchill, E. (eds.): *Embodied conversational agents*. Cambridge, MA: MIT Press (2000)
2. Core, M., Traum, D., Lane, H.C., Swartout, W., Gratch, J., van Lent, M., Marsella, S.: *Teaching Negotiation Skills through Practice and Reflection with Virtual Humans*. *Simulation* 82(11), 685–701 (2006)
3. Martínez-Miranda, J., Jung, B., Payr, S., Petta, P.: *The Intermediary Agent's Brain: Supporting Learning to Collaborate at the Inter-Personal Level*. In: *Proceedings of the 7th Conference on Autonomous Agents and Multiagent Systems (AAMAS'08)* 3: pp. 1277–1280 (2008)
4. Gebhard, P., Schröder, M., Charfuelan, M., Endres, C., Kipp, M., Pammi, S., Rumpler, M., Türk, O.: *IDEAS4Games: Building Expressive Virtual Characters for Computer Games*. In: *Proceedings of the 8th International Conference on Intelligent Virtual Agents (IVA 2008)*, pp. 426–440 (2008)
5. Mateas, M., Stern, A.: *Façade: An experiment in building a fully-realized interactive drama*. In: *Game Developers Conference (GDC'03)* (2003)
6. Zhang, Z., Bickmore, T.W., Mainello, K., Mueller, M., Foley, M., Jenkins, L., Edwards, R.A.: *Maintaining Continuity in Longitudinal, Multi-method Health Interventions Using Virtual Agents: The Case of Breastfeeding Promotion*. In: *Proceedings of Intelligent Virtual Agents (IVA 2014)*, pp. 504–513 (2014)
7. Micoulaud-Franchi, J.A., Sagaspe, P., de Sevin, E., Bioulac, S., Sauteraud, A., Philip, P.: *Acceptability of Embodied Conversational Agent in a Health Care Context*. In: *Proceedings of Intelligent Virtual Agents (IVA 2016)*, pp. 416–419 (2016)
8. Bickmore, T.W., Utami, D., Matsuyama, R., Paasche-Orlow, M.: *Improving Access to Online Health Information with Conversational Agents: A Randomized Controlled Experiment*. *Journal of Medical Internet Research*. 18(1) (2016)
9. Martínez-Miranda, J.: *Embodied Conversational Agents for the Detection and Prevention of Suicidal Behaviour: Current Applications and Open Challenges*. *Journal of Medical Systems*, 41(9), 135 (2017)
10. Traum, D., Aggarwal, P., Artstein, R., Foutz, S., Gerten, J., Katsamanis, A., Leuski, A., Noren, D., Swartout, W.: *Ada and Grace: Direct Interaction with Museum Visitors*. *Intelligent Virtual Agents*, pp. 245–251 (2012)
11. Richards, D.: *Agent-based museum and tour guides: applying the state of the art*. In: *Proceedings of the 8th Australasian Conference on Interactive Entertainment: Playing the System (IE '12)*. ACM, New York, NY, USA, Article 15, 9 pages (2012)
12. Shaked N.A.: *Avatars and virtual agents – relationship interfaces for the elderly*. *Healthcare Technology Letters* 4(3), 83–87 (2017)
13. Lucas, G.M., Gratch, J., King, A., Morency, L.P.: *It's only a computer: Virtual humans increase willingness to disclose*. *Computers in Human Behavior* 37, pp. 94–100 (2014)
14. Scherer, K.R.: *Emotion and emotional competence: conceptual and theoretical issues for modelling agents*. In: Scherer, K.R., Bänziger, T., Roesch, E.B. (eds.): *Blueprint for Affective Computing: A Sourcebook*; Chap. 1.1. Oxford University Press; pp. 3–20 (2010)

15. Marsella, S., Gratch, J., Petta, P. Computational Models of Emotion. In: Scherer, K.R., Bänziger, T., Roesch, E.B. (eds.): *Blueprint for Affective Computing: A Sourcebook*. Chapter 1.2. pp. 21–46. Oxford University Press (2010)
16. Swartout, W., Traum, D., Artstein, R., Noren, D., Debevec, P., Bronnenkant, K., Williams, J., Leuski, A., Narayanan, S., Piepol, D., Lane, C., Morie, J., Aggarwal, P., Liewer, M., Chiang, J., Gerten, J., Chu, S., White, K.: *Ada and Grace: Toward Realistic and Engaging Virtual Museum Guides*. In: *International Conference on Intelligent Virtual Agents*. *Intelligent Virtual Agents*, pp. 286–300 (2010)
17. Kopp, S., Gesellensetter, L., Krämer, N., Wachsmuth, I.: *A Conversational Agent as Museum Guide – Design and Evaluation of a Real-World Application*. In: *International Workshop on Intelligent Virtual Agents*. *Intelligent Virtual Agents*, pp. 329–343 (2005)
18. Campbell, A.G., Stafford, J.W., Holz, T., O’Hare, G.M.P.: *Why, when and how to use augmented reality agents (AuRAs)*. *Virtual Reality* 18, pp. 139–159 (2014)
19. Barakonyi, I., Schmalstieg, D.: *Augmented reality agents for user interface adaptation*. *Computer Animation and Virtual Worlds* 19, pp. 23–35. Published online in Wiley InterScience (www.interscience.wiley.com) (2008)
20. Ortony, A.: *On making believable emotional agents believable*. *Emotions in Humans and Artifacts*, Cambridge, MA: MIT Press, pp. 189–211 (2001)
21. Scherer, K., Schorr, A., Johnstone, T.: *Appraisal Processes in Emotion: Theory, Methods, Research*. Oxford University Press (2001)
22. Lazarus, R.: *Emotion and Adaptation*. Oxford University Press (1991)
23. Ortony, A., Clore, G., Collins, A.: *The Cognitive Structure of Emotions*. Cambridge University Press (1998)
24. Ito, T., Iwama, Y.: *Implementation of Question Answering System Based on Reference-Based Ranking Algorithm*. In: Ali, M., Pan, J.S., Chen, S.M., Horng, M.F. (eds): *Modern Advances in Applied Intelligence*. IEA/AIE 2014. *Lecture Notes in Computer Science*, vol 8481. Springer, Cham (2014)