

Tipo de artículo: Artículo original
Temática: Inteligencia Artificial
Recibido: 28/02/2016 | Aceptado: 28/03/2016

Bots conversacionales para la corrección de errores ortográficos en idioma inglés

Bots conversational to correct spelling errors in English

Yeili Ibarra Monteagudo^{1*}, Ricardo Valle Priel², Rafael Aristides Quintanal Reytor¹, Fermín Lorenzo Carbajal³, Enelis Blanca Cuba Rondón¹, Karel Piorno Charchabal¹, Zoraida Fernández Guevara³

¹Vertex. Centro Entornos Interactivos 3d. Universidad de las Ciencias Informáticas, Carretera a San Antonio, km 2 ½, Torrens, Boyeros, La Habana, Cuba. CP: 19370. {yibarra, rreytor, ecuba, kpiorno}@uci.cu

² Centro de Soporte. Universidad de las Ciencias Informáticas, Carretera a San Antonio, km 2 ½, Torrens, Boyeros, La Habana, Cuba. CP: 19370. ricardovp@uci.cu

³ Departamento de Ciencias Sociales y Humanidades, Facultad 5. Universidad de las Ciencias Informáticas, Carretera a San Antonio, km 2 ½, Torrens, Boyeros, La Habana, Cuba. CP: 19370. {fcarvajal, zorlis}@uci.cu

Autor para correspondencia: yibarra@uci.cu

Resumen

En el mundo existen diferentes plataformas para la simulación de entornos virtuales como Second Life y OpenSim, donde la docencia puede realizarse en un espacio adecuado y accesible para un público que busca experiencias de aprendizaje práctico. Estos entornos, permiten simular espacios de la vida cotidiana de una manera fiel a la realidad, lo que causa mayor nivel de atracción para el público y que sean fuertemente usados para la formación en idiomas y la comunicación, a través de personajes virtuales. El presente trabajo propone realizar Bots conversacionales para interactuar y corregir errores en la producción escrita de los estudiantes, y alcanzar un nivel A3 del idioma inglés en un mundo virtual o metaverso. Se desarrollaron siete agentes inteligentes con la capacidad de corregir errores ortográficos y gramaticales distribuidos por los diferentes módulos de contenido de idioma inglés. La propuesta de solución contribuye a la retroalimentación sincrónica del contenido recibido durante la interacción en el mundo virtual, ventajas que el Entorno Virtual de Enseñanza Aprendizaje (EVEA) acoplado en Moodle no brinda, además de favorecer la colaboración del estudiantado en un entorno más ameno.

Palabras clave: aprendizaje; entornos virtuales; interactividad; OpenSim

Abstract

In the world there are different platforms for simulating virtual environments such as Second Life and OpenSim, where teaching can be done in an appropriate and accessible to an audience looking for hands-on learning experiences space. These environments, can simulate spaces of everyday life in a lifelike manner, causing greater level of attraction to the public and are heavily used for training in languages and communication through virtual characters. This paper intends to make Bots conversational to interact and correct errors in student writing, to reach a level of English A3 a metaverse or virtual world production. Seven Intelligent agents were developed with the ability to correct spelling and grammatical errors dispatched by the various modules of content English. The proposed solution contributes to the synchronous feedback received benefits content Virtual Learning Environment Moodle Learning does not provide and would enhance the collaboration of students in a more pleasant environment.

Keywords: learning; virtual environments; interactivity; OpenSim

Introducción

Hoy en día, los mundos virtuales o metaversos como *Second Life*, *The Sims*, *Habbo Hotel*, *EverQuest* y *Rune Scape* constituyen una forma de contribuir al aprendizaje, pues la comunicación entre el alumno y el sistema de educación online pueden ser de una forma síncrona, a diferencia de las plataformas en 2D como Moodle que ofrecen una comunicación asíncrona (MÁRQUEZ, 2011). Una comunicación síncrona puede incrementar la motivación del alumno y reducir los errores, pues se tiene una retroalimentación inmediata; mientras que una comunicación asíncrona incrementa el esfuerzo cognitivo del estudiante, ya que este tiene más tiempo para estudiar y pensar en el tema. La comunicación síncrona se puede utilizar de apoyo a la asíncrona para así conseguir una mayor participación personal del estudiante.

Debido al creciente desarrollo tecnológico y el proceso de informatización, en Cuba ha sido posible la creación de “*Virtual English*” (Villavicencio y Fuentes, 2013), un entorno virtual 3D para la enseñanza de idiomas con el que se hace agradable el intercambio entre profesor y estudiante.

En la Universidad de las Ciencias Informáticas (UCI) existe un Entorno Virtual de Enseñanza-Aprendizaje (EVEA) donde estudiantes y profesores desarrollan sus tareas de aprendizaje a través del uso de sus recursos (foros, chat, multimedia, entre otros). El EVEA se desarrolló bajo la plataforma Moodle, la cual no fue concebida en su totalidad para el aprendizaje de idiomas y por lo tanto, no satisface el proceso de enseñanza - aprendizaje de profesores y estudiantes, pues dentro de las desventajas que presenta esta plataforma es que los textos son planos y carece de interactividad a la hora de realizar las tareas por parte de los estudiantes. Para suplir esta carencia se desarrolló “*Virtual English*” (Villavicencio y Fuentes, 2013), un entorno virtual 3D que propicia una interacción de forma amena a través de un agente inteligente y simula un mundo donde se produce el aprendizaje real. El agente inteligente que existe en el mundo no es capaz de brindar una retroalimentación sincrónica de los módulos del contenido de inglés, además de no poseer la característica de corregir los errores cometidos durante la producción escrita. A partir de las consultas realizadas a los expertos en la materia de inglés, se puede mencionar que este agente, ofrece el contenido de manera general y no por módulos que incluyan en cada uno las funciones comunicativas para obtener un nivel A3 del idioma inglés, lo que impide al estudiante desarrollar sus habilidades de escritura y comprensión lectora al ritmo que la asignatura lo requiere. El hecho de que exista un solo agente inteligente en el mundo virtual, puede causar desmotivación en el estudiante, además de no promover el aprendizaje mediante la actividad colaborativa en caso de no haber otros estudiantes conectados.

Para la solución de esta problemática se propone desarrollar siete agentes inteligentes distribuidos por los diferentes módulos de contenido de la asignatura y áreas del “*Virtual English*”. Se plantea además que los agentes inteligentes tengan la capacidad de corregir los errores ortográficos y de estructura gramatical que puedan cometer los estudiantes en la producción escrita, con el objetivo de promover el avance del estudiantado hacia un nivel A3 del sistema de conocimientos de inglés (Manuel, 2012).

Materiales y métodos o Metodología computacional

A continuación, se describen un conjunto de tecnologías utilizadas para el desarrollo de cada uno de los agentes inteligentes y sus funcionalidades. Se utiliza como metodología de desarrollo de software la metodología ágil XP, por ser una metodología flexible a los cambios que se puedan realizar en el proyecto, además permitir grupos de desarrollo relativamente pequeños.

Microsoft Visual Studio (MonoDevelop - Guía Ubuntu, 2014 y MonoDevelop 1.0 Released, 2014)

Microsoft Visual Studio es un Entorno Integrado de Desarrollo (IDE) que cuenta con varias características como (Mayo, 2014):

- Permite trabajar con múltiples versiones de .NET en un único entorno. El framework .Net hace que desarrollar aplicaciones sea una tarea mucho más sencilla que antes.
- Emplea menos tiempo depurando la aplicación cuando registra y realiza un seguimiento a través del historial de ejecución.
- Incluye herramientas para el desarrollo de aplicaciones dirigidas a las últimas versiones de Microsoft.

Con este entorno de desarrollo se ha programado la corrección de errores ortográficos de los agentes inteligentes utilizando el lenguaje de programación C#.

Biblioteca LibOpenMetaverse

La biblioteca LibOpenMetaverse fue usada para acceder al entorno virtual en 3D y a sus funcionalidades.

La biblioteca consta de varias dependencias (GAMOLAN, 2013):

- OpenMetaverse.dll: Es el núcleo de la biblioteca OpenMetaverse, que puede ser usado para crear aplicaciones clientes y servidor.
- OpenMetaverseTypes.dll: Usada para funciones específicas 3D y bibliotecas matemáticas.
- OpenMetaverse.StructuredData.dll: Conjunto de librerías para dar soporte a Linden Lab Structured Data (LLSD) y Java Script Object Notation (JSON).
- OpenJpeg: Un envoltorio .NET para codificar y decodificar los datos JPEG2000.

Biblioteca OpenNLP

OpenNLP es un conjunto de herramientas basadas aprendizaje de máquinas para el procesamiento de texto en lenguaje natural. Es compatible con las tareas más comunes de PNL, como tokenización, segmentación frase, de marcado parte de discurso, la extracción de entidades nombradas, fragmentación, el análisis y la resolución de la correferencia. Estas tareas suelen ser necesarios para crear servicios más avanzados de procesamiento de texto («Apache OpenNLP Developer Documentation», 2014 y Niveles - ESL cursos de idiomas, 2014). Esta biblioteca fue utilizada con el objetivo de realizar las correcciones de las estructuras gramaticales establecidas por parte de los agentes inteligentes.

Lenguaje AIML (A.L.I.C.E Artificial Intelligence Foundation, 2010)

El AIML (del inglés *Artificial Intelligence Mark-up Language*) es un lenguaje de programación basado en XML que utiliza la técnica de IA de los Sistemas Basados en Reglas. AIML cuenta con tres componentes fundamentales para establecer las conversaciones entre los usuarios y los agentes inteligentes (*category*, *pattern* y *template*) («AIML - The Artificial Intelligence Markup Language», 2014 y SÁNCHEZ, 2010).

Category: son las unidades fundamentales de conocimiento; agrupan dos componentes *pattern* y *template* los cuales relacionan un mensaje de entrada del usuario y uno de salida del bot respectivamente y que normalmente se codifican en ese orden.

Biblioteca AIMLbot.dll (LAPUENTE, 2014)

Se hace uso de la biblioteca AIMLbot.dll de código abierto que permite utilizar el lenguaje AIML desde código C#, con el fin de brindar las respuestas por parte de los agentes inteligentes a los usuarios, utilizando las clases con que cuenta (*Bot*, *User*, *Request*, *Result*).

La clase **Bot** contiene una referencia al conocimiento del bot conversacional (esto es, las rutas a los ficheros AIML) así como las conversaciones (almacenadas) que se han establecido con él.

La clase **User** identifica a un usuario concreto que manda mensajes a un bot específico. Estos mensajes del **User** para el bot se envían mediante la clase **Request**. Como respuesta, el bot genera un resultado a través de la clase **Result** que contiene la respuesta del bot.

Resultados y discusión

El presente trabajo incorpora al entorno virtual “*Virtual English*”, siete agentes inteligentes para mejorar el dinamismo y la interacción de los estudiantes con el contenido del curso de inglés que se encuentra en el servidor Moodle.

La arquitectura que se propone es una arquitectura por capas, con la que se logra establecer la comunicación estudiante-bot, la misma cuenta con las siguientes capas:

- Capa de usuario
- Capa del servidor OpenSim
- Capa de Acceso a datos
- Capa de comunicación

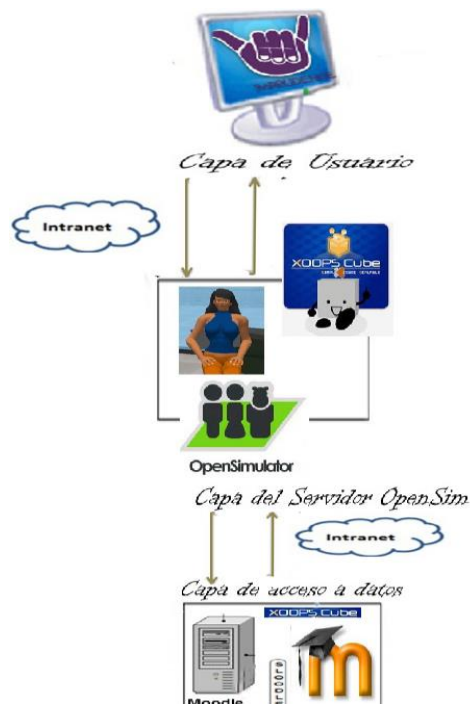


Figura 1: Descripción de la arquitectura del sistema.

En la figura 2 se muestran cada uno de los agentes inteligentes distribuidos por las diferentes áreas del entorno virtual.

Como se mencionaba, los agentes inteligentes independientemente del módulo de contenido que estos manejen, tienen la capacidad de corregir los errores ortográficos y gramaticales que se cometan durante la producción escrita.

Cuando se comete un error ortográfico se le muestra al usuario un menú con tres opciones, ver la **Figura 3**.

Si se selecciona la opción 1 (Figura 4) se le mostrará al usuario un conjunto de palabras como sugerencia a la palabra incorrecta, una vez seleccionada una palabra se le preguntará si desea establecerla como autocorrección para futuros errores.



Figura 2: Agentes Inteligentes.

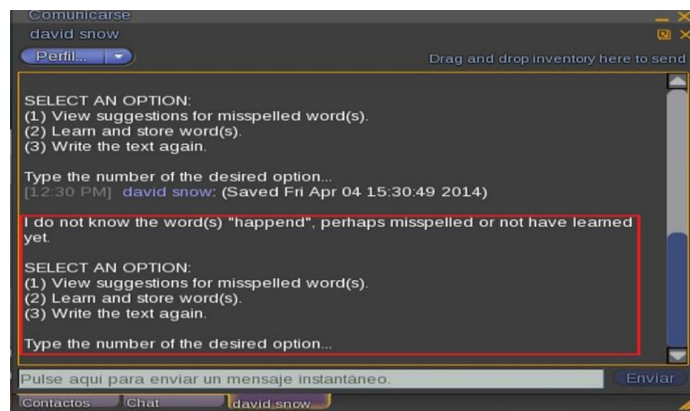


Figura 3: Menú que se muestra cuando se comete un error ortográfico.

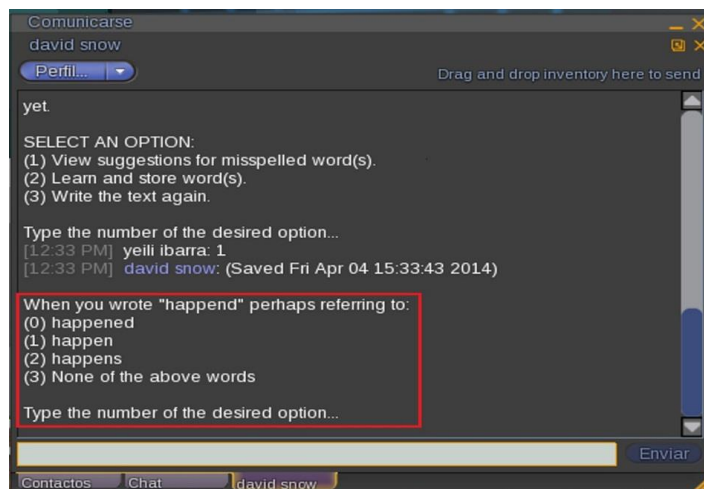


Figura 4: Sugerencias para la palabra incorrecta.

Si se selecciona la opción 2 el usuario podrá añadir la palabra a la base de datos, pues puede estar correcta y se reconozca como incorrecta por no aparecer en el diccionario de palabras.

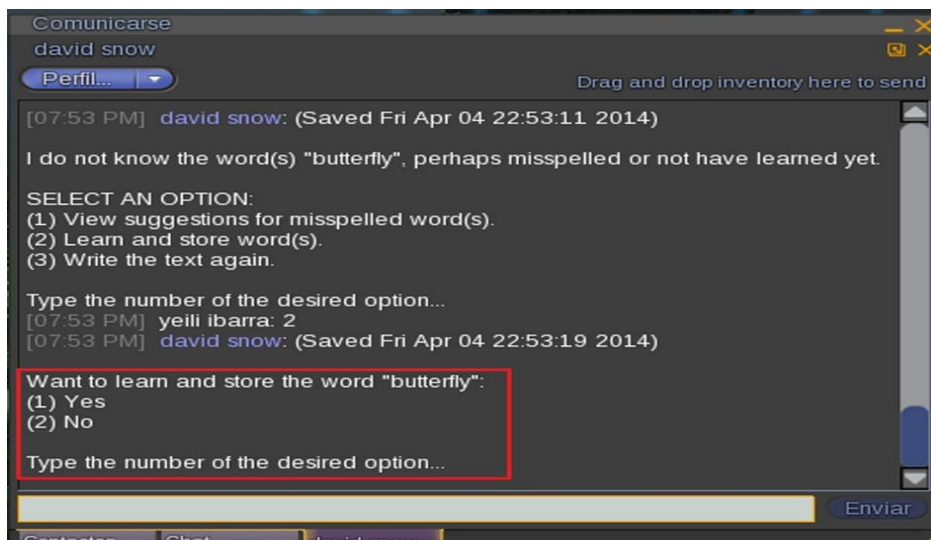


Figura 5: Añadir palabra a la base de datos.

La figura muestra un ejemplo para la corrección gramatical.

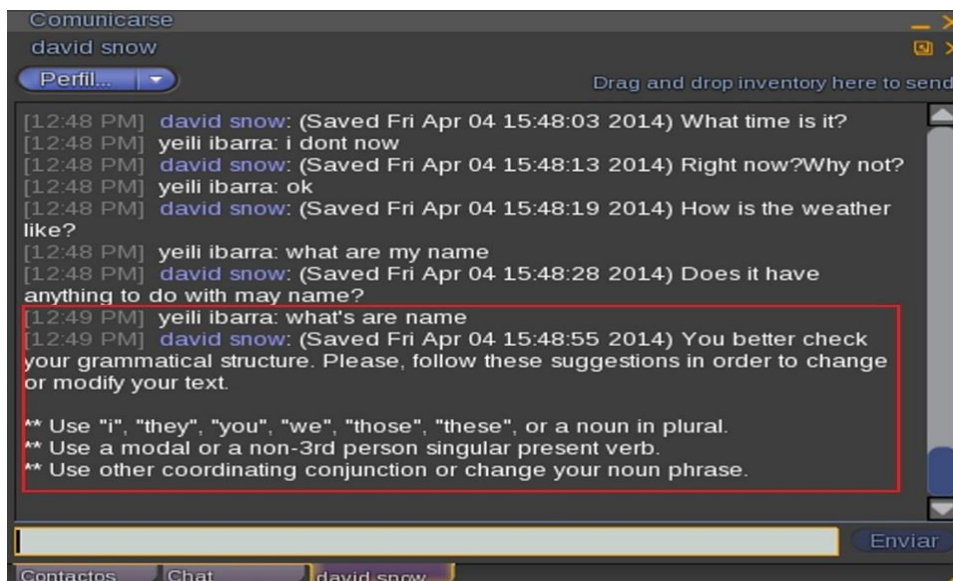


Figura 6: Sugerencias para un error gramatical.

Conclusiones

- Con la incorporación de los siete agentes inteligentes en el metaverso “*Virtual English*”, se promueve el aprendizaje a través de la actividad interactiva, se fortalece el proceso de enseñanza y el aprendizaje de los estudiantes, además de ser un elemento para lograr la inmersión dentro del entorno.
- La corrección ortográfica y gramatical por parte de los agentes inteligentes de forma inmediata, permite una retroalimentación sincrónica del conocimiento recibido.

Referencias

AIML - The Artificial Intelligence Markup Language. *AIML - The Artificial Intelligence Markup Language* [en línea] 2014. [Consulta: 10 octubre 2014]. Disponible en: <http://www.alicebot.org/aiml.html>.

A.L.I.C.E Artificial Intelligence Fundation. A.L.I.C.E Artificial Intelligence Fundation [online] 2010. 12/2/2014.
<http://www.alicebot.org/aiml.html>.

ANA MARÍA VILLAVICENCIO y YASMANY ALFONSO FUENTES 2013. *Entorno Submersivo Virtual en OpenSim en Función Del Aprendizaje de Idioma Extranjero I.pdf*. Tesis de Grado. La Habana: Universidad de las Ciencias Informáticas.

Apache OpenNLP Developer Documentation. *Apache OpenNLP Developer Documentation* [en línea] 2014. [Consulta: 10 octubre 2014]. Disponible en: <http://opennlp.apache.org/documentation/manual/opennlp.html>.

ARTURO ALEJANDRO CASTRO, NADIA ALEJANDRO CASTRO, GRACIELA TERESA GARCÍA VIZCAYA, MARIA DE LOS ANGELES HERNÁNDEZ y MARIA LUISA TORRES LARA, 2005. *Agentes_Inteligentes.pdf*. 22 agosto 2005. S.l.: Instituto Tecnológico de Nuevo Laredo.

EDUARDO ROJO SÁNCHEZ 2010. *Aplicación de la Herramienta Open Source Sloodle y las Tecnologías del Procesamiento del lenguaje Natural para el Desarrollo de una Plataforma de Virtual Learning en la Universidad Carlos III de Madrid.pdf*. Tesis de Grado. Madrid: Carlos III.

GAMOLAN ADAMSKI. Ayuda para Second Life. [en línea] 2013. [Consulta: 10 octubre 2014]. Disponible en: <http://ayuda-secondlife.blogspot.com/imprudence-viewer.html>.

(INTECO), INSTITUTO NACIONAL DE TECNOLOGÍAS DE LA COMUNICACIÓN. Bots Conversacionales. In.: INTECO-CERT, 2010.

ISRAEL V. MÁRQUEZ 2011. *METAVERSOS Y EDUCACIÓN Second Life como plataforma educativa.pdf*. *Revista Icono*, vol. 2, no. 14, pp. 16. ISSN 1697-8293. G - 84075977

JOE MAYO 2014. *Visual_Studio_2010.pdf*. Primera Edición. New York: McGraw Hill. ISBN 978-0-07-166896-5. 0-07-166896-9

LAPUENTE, M.J.L. 2014. Hipertexto: el nuevo concepto de documento en la cultura de la imagen. En: Tesis doctoral sobre hipertexto, el nuevo concepto de documento en la cultura de la imagen. Autora María Jesús Lamarca Lapuente. Tema Biblioteconomía y Documentación [en línea]. [Consulta: 10 octubre 2014]. Disponible en: <http://www.hipertexto.info/>.

Manuel, J. Hablando de clase. 2012 [En Línea] [cited 2014 21 Enero]. Disponibles en: <http://hablandodeclase.blogspot.com/>

MonoDevelop - Guía Ubuntu. *MonoDevelop - Guía Ubuntu* [en línea] 2014. [Consulta: 10 octubre 2014]. Disponible en: <http://www.guia-ubuntu.com/index.php?title=MonoDevelop>.

MonoDevelop 1.0 Released - MonoDevelop. *MonoDevelop 1.0 Released - MonoDevelop* [en línea] 2014. [Consulta: 10 octubre 2014]. Disponible en: http://monodevelop.com/download/monodevelop_1.0_released.

Niveles - ESL cursos de idiomas. *Niveles - ESL cursos de idiomas* [en línea] 2014. [Consulta: 10 octubre 2014]. Disponible en: <http://www.esl-idiomas.com/es/idiomas-en-el-extranjero-niveles.htm>.