



Documento de trabajo: di 05-01-01
Enero 2005

Serie de innovación docente

Sistema distribuido de calificación de prácticas en web: una experiencia utilizando agentes que razonan con lógica difusa

Javier Carbó*, Jose M. Molina**

Dpto. Informática, Universidad Carlos III de Madrid
Escuela Politécnica Superior.

* Campus de Leganés. Avda. Universidad 30, 28911 Leganés

**Campus de Colmenarejo. Avda. Universidad Carlos III 22, 28270 Colmenarejo

Abstract

En este documento se describe un proyecto de innovación docente centrado en la evaluación de prácticas de forma distribuida. El proyecto se ha llevado a cabo en el marco de la asignatura 'Ingeniería de Conocimiento' de la titulación de Ingeniería Informática de la Universidad Carlos III de Madrid. Su propósito inicial consiste en la mejora del proceso de asimilación de conceptos mediante un mecanismo de evaluación participativo. La propia naturaleza de la asignatura, que consiste en la aplicación de una metodología a problemas de dominios extra-académicos escogidos por los propios alumnos, facilita la adopción de este mecanismo de evaluación, ya que las prácticas pueden ser defendidas en público sin resultar repetitivas. Este enfoque se ha reforzado con correcciones parciales puestas en común y sometidas a debate en las clases prácticas a lo largo del curso. En este trabajo se presenta una solución innovadora que facilita la participación directa de los alumnos en la evaluación de sus trabajos. A tenor de las calificaciones, de la actitud mostrada por los alumnos, y por los resultados de un test de evaluación de la experiencia de innovación docente, podemos concluir unos resultados muy positivos de su implantación.

Palabras clave: Evaluación definitiva, lógica difusa, agentes autónomos, sistemas distribuidos

Introducción

En este documento se describe un proyecto de innovación docente centrado en la evaluación de prácticas de forma distribuida. El proyecto se ha llevado a cabo en el marco de la asignatura 'Ingeniería de Conocimiento' de la titulación de Ingeniería

Informática de la Universidad Carlos III de Madrid. Se trata de una asignatura dedicada a cómo se construye un tipo especial de programas informáticos, denominados sistemas expertos, que emulan el comportamiento de expertos humanos [1][2].

Las prácticas de la asignatura consisten en la propuesta de un problema que requiera el saber-hacer de un experto en el dominio. Este problema ha de ser estudiado y analizado en profundidad para realizar un modelo del conocimiento necesario para resolverlo, con el que se podrá implementar un sistema informático que pueda sustituir al experto.

Las fases que han de seguirse, los pasos y técnicas que han de aplicarse constituyen una metodología. Aunque existen varias metodologías de construcción de sistemas expertos una de las más aceptadas, convertida en un estándar europeo de facto, es la metodología KADS [3][4] que se explica en la asignatura de ingeniería de conocimiento.

La primera fase de dicha metodología consiste en identificar si resolver el problema mediante ingeniería de conocimiento está justificado, es adecuado, plausible y tiene posibilidades de éxito. Este estudio de viabilidad del sistema experto es la primera tarea que han de abordar los alumnos, y para ello han de escoger a un experto (dispuesto a colaborar) y su correspondiente dominio de problemas. Así los alumnos escogieron problemas como: diagnóstico de averías eléctricas en automóviles, planificación de actividades en escuelas infantiles, clasificación setas, estimación de calidad del vino, identificación de toros de lidia, planificación de recetas de cocina, etc.

La segunda fase de la metodología es la adquisición de conocimientos: Para que un programa razone de forma semejante a un experto humano, es preciso adquirir el conocimiento adquirido por el experto humano en sus años de experiencia y que aplica, incluso de forma implícita en la resolución de los problemas de su dominio. En la adquisición de este conocimiento se suelen aplicar distintas técnicas como entrevistas, cuestionarios, análisis de protocolos, observación de tareas habituales, etc. Esta fase suele ser en la que más tiempo se invierte, y en ella los alumnos escogen un subconjunto de las técnicas explicadas en clase y las utilizan con su experto para adquirir el conocimiento relativo a cómo éste resuelve los casos que se le presentan.

A continuación de la adquisición del conocimiento del experto, se debe formular un modelo de estos conocimientos que refleje tanto la estructura de conceptos que maneja el experto como la secuencia de tareas/decisiones que sigue el experto para resolver el problema. Al primer tipo de conocimiento se le denomina estático y al segundo dinámico. Ambos modelos se describen de forma informal mediante una notación gráfica y a los alumnos se les alecciona para que busquen semejanzas con los modelos de problemas genéricos propuestos por la metodología KADS para su reutilización.

Finalmente los alumnos han de transformar estas representaciones intermedias del modelo conceptual a una notación formal que facilite al programador la implementación final del sistema experto. La notación que han de seguir es CML2 (Conceptual Modeling Language) [5]. Tanto la programación del código como la posterior verificación y validación solo se explican de forma teórica.

El objetivo de aprendizaje de la asignatura está centrado en la comprensión y desarrollo de la práctica por lo que el proyecto de innovación se centra en la misma. El ánimo del proyecto es mejorar el desarrollo de la práctica y fomentar la participación de los

alumnos en la crítica y valoración de las prácticas presentadas en la asignatura. El alumno debe comprender que el aprendizaje de la asignatura no está centrado únicamente en el desarrollo de su propia práctica, si no en la comprensión general de los problemas a los que puede aplicarse y de las distintas soluciones que pueden plantearse. Este aprendizaje se alcanza si se analizan y estudian en conjunto todas las prácticas de la asignatura y por lo tanto el proyecto pretende fomentar entre los alumnos interés no sólo por el desarrollo de su propia práctica si no también por las soluciones presentadas por sus compañeros a los diversos problemas planteados.

Planteamiento del problema

En este epígrafe trataremos de explicar la motivación del proyecto innovación docente. Las razones que motivan esta iniciativa están íntimamente relacionadas con la naturaleza concreta de la asignatura:

- Esta es de carácter eminentemente práctico y su docencia requiere la realización gradual de la práctica a lo largo de todo el cuatrimestre ya que el desarrollo de ésta debe seguir la metodología objeto de estudio en la asignatura.
- Además, requiere de una intensa coordinación con los profesores en la que se hagan visibles los errores cometidos en cada fase antes de reproducirlos en las fases siguientes.
- Así mismo la asignatura necesita de la participación activa de un experto en un dominio extra-académico con el que se deben realizar numerosas sesiones de adquisición de conocimiento.
- También la diversidad de temas y tipos de problemas abordados por los diferentes expertos escogidos por cada grupo de prácticas hace que cada una de ellas tenga muchas particularidades no compartidas con las demás prácticas. Esta circunstancia lleva a que los alumnos se encuentren un poco desorientados.
- A diferencia de gran parte de las asignaturas de la titulación, la ausencia de código ejecutable que muestre de forma clara la corrección y completitud del trabajo realizado genera un alto grado de incertidumbre en la evaluación del mismo por parte del alumnado. Así la evaluación se fundamenta en una valoración crítica de la forma de entender el problema y de conceptualizar la solución. Comprender esa valoración supone un valor añadido puesto que no sólo son capaces de construir una solución si no que además pueden entender y valorar otras soluciones posibles.
- Finalmente, por la observación de cursos pasados, las exposiciones orales del trabajo realizado no parecen interesar a los alumnos presentes en el aula, ni por su asistencia, ni por sus intervenciones en el ruego de preguntas y comentarios, ni por la reiteración mostrada en los errores cometidos en las prácticas.

Objetivos específicos

Todas estas circunstancias llevaron a plantearnos los siguientes objetivos respecto a la evaluación de las prácticas:

- Involucrar a los alumnos en la evaluación de las prácticas de la asignatura.
- Mejorar la comprensión de los criterios con los que se determina la nota de la práctica: justificación/adecuación del problema, valoración del trabajo realizado para la adquisición del conocimiento, claridad y corrección de la representación del modelo de conocimientos estático y dinámico.

- Animar a la participación y atención en clase durante la exposición oral de la práctica.
- Incentivar la mejora en la calidad de la memoria y la presentación de la práctica al saberse evaluados por sus propios compañeros.

Adicionalmente hemos re-estructurado la docencia de la asignatura en consonancia con la tendencia apuntada en la declaración de Bolonia [6] de forma que el peso dado a las clases prácticas en el esfuerzo y la calificación crezca en detrimento de la teoría, se distribuya el esfuerzo dedicado a la superación de la asignatura a lo largo de todo el cuatrimestre, se desplace la dedicación de los alumnos a actividades fuera de las aulas, y finalmente se sustituyan parte de las clases magistrales por debates y resolución colaborativa de casos prácticos.

Estrategia docente

La estrategia docente que se ha seguido incluye tipos de sesiones muy diferentes:

- Clases de teoría en las que el profesor de teoría explica cada una de las fases de la metodología en detalle: identificación del problema, adquisición, conceptualización y formalización.
- Las clases de teoría y prácticas correspondientes a la fase de formalización se imparten en un aula de ordenadores en la que se les explica el funcionamiento de la herramienta informática que guía y valida la corrección de la notación formal utilizada.
- Sesiones de resolución de casos en los que el profesor de prácticas expone problemas concretos de cada fase que resuelven entre todos. Y en las que los alumnos plantean la relación de estos casos con sus propios problemas.
- Sesiones prácticas fuera de las aulas en las que los alumnos aplican las distintas técnicas de adquisición de conocimientos con los expertos del dominio escogido por ellos mismos.
- Defensa oral del trabajo realizado mediante una presentación de 15m tras la cual se produce un debate sobre los puntos fuertes y débiles de la práctica en cuestión.

La participación activa en el mecanismo de calificación distribuido, en las sesiones de resolución de casos y en las defensas orales fue recompensada con una valoración subjetiva de los profesores de hasta el 10% de la calificación final. La práctica se valoró en un 65% de la calificación y el examen (claramente orientado a la práctica) en un 25%, siendo necesario aprobar tanto el examen como la práctica para superar la asignatura.

Medios

El medio utilizado para que cada grupo de prácticas evalúe a los demás grupos de prácticas es un sistema distribuido que se basa en la tecnología de agentes. Concretamente los cálculos matemáticos correspondientes a cómo se ponderan todas las evaluaciones se basan en trabajos de investigación de los profesores de la asignatura que han sido aplicados en la actualización de la reputación a través de recomendaciones en un entorno de comercio electrónico [10] [11]. Las recomendaciones son subjetivas y se representan mediante conjuntos borrosos y el algoritmo desarrollado permite la combinación de diversas opiniones representadas por conjuntos borrosos cada una de

ellas pesada por la confianza/reputación que merece el emisor de la misma, para detalles concretos del método véase [10].

De manera resumida podemos presentar el sistema como un conjunto de agentes, cada uno representa a un grupo de prácticas, que evalúan a los distintos grupos y a su vez son evaluados por ellos. Existe un tipo diferenciado de agentes que representa a los profesores de la asignatura, cada uno de ellos también distribuye las calificaciones a los destinatarios (la cual será tratada de forma especial como se verá en el método de cálculo), pero no ha de determinar la calificación propia.

Un agente es un programa informático que actúa en representación de una persona con un cierto grado de inteligencia artificial y autonomía [7]. Aunque existen diversas arquitecturas de agentes, la adoptada en este proyecto de innovación docente es la basada en el paradigma deliberativo [8]. Ésta representa tres niveles de conocimiento con el que el agente razona: creencias acerca de sí mismo y el exterior, intenciones que pretende ejecutar en el futuro próximo y deseos finales que persigue (véase figura 1).

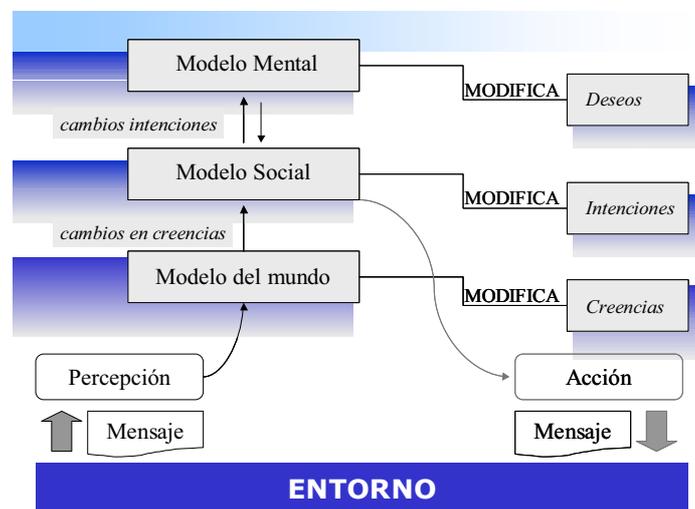


Figura 1. Modelos y tipos de conocimiento de un agente

Los deseos representan una abstracción de la motivación que da sentido al agente, y que se concretan en objetivos cuando se cumplen ciertas condiciones. Estos objetivos tienen asociados unos planes (secuencias de intenciones) que pretenden satisfacer el deseo perseguido. El algoritmo de ejecución típico de un agente deliberativo es:

```

Repetir
    intención=Selecciona(intenciones)
    Ejecuta(intención)
    evento=Percebe()
    Modifica(creencias,evento)
    Modifica(intenciones,creencias)
    Modifica(deseos,creencias)
Fin Repetir
    
```

En nuestro caso, cada agente representa a un grupo de prácticas y persigue dos deseos: manifestar las calificaciones introducidas por su representado con los destinatarios correspondientes junto con el peso inicial asociado a las mismas, y al tiempo determinar

la calificación final del grupo de prácticas representado por el agente. Las creencias que maneja se refieren a las calificaciones conocidas:

- las originadas por su representado y que provienen de una herramienta web diseñada ad-hoc.
- las que el representado ha recibido de los demás agentes,

Las calificaciones son introducidas en el agente de grupo a través de una página web. Por ejemplo en la figura 2 se observa que el Agente A recibe las calificaciones de los agentes B, C, etc.. a través de una página web y como el agente A se comunica con ellos para distribuir la información introducida por el grupo de prácticas.

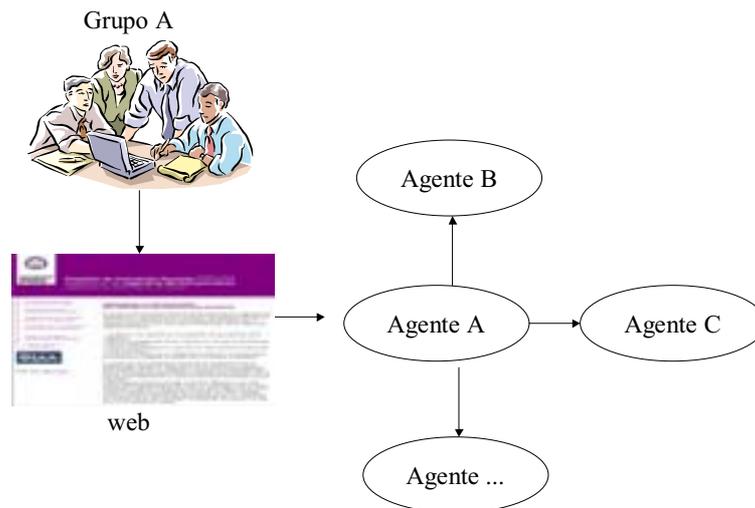


Figura 2. Comunicación entre agentes

Cada calificación tiene cuatro apartados: justificación/adecuación del problema, valoración del trabajo realizado para la adquisición del conocimiento, claridad y corrección de la representación del modelo de conocimientos estático y dinámico. Y se expresa en términos lingüísticos -no matemáticos- tales como: mucho, poco, etc., con la posibilidad de combinarlos con modificadores tales como: más o menos, extremadamente, etc. Esta valoración en lenguaje natural tiene por objeto facilitar la representación de la subjetividad inherente a la evaluación de estas prácticas. Estos términos tendrán asociada una representación matemática que permitirá operar con ellos. La representación matemática y los operadores son los correspondientes a la lógica borrosa. En 1965 Lotfi A. Zadeh [9] dio origen a esa disciplina al proponer un tipo especial de funciones para manejar formalmente la imprecisión y la vaguedad, conocidas como conjuntos difusos o borrosos. La utilización de términos vagos e inconcretos típica del lenguaje natural permite a los humanos razonar sobre fenómenos complejos y filtrar los detalles irrelevantes. En general por vaguedad se entiende la dificultad para hacer afirmaciones precisas en un cierto dominio, y así, la pertenencia de un objeto dado a un conjunto no puede determinarse de forma concluyente sino es mediante una escala de valores intermedios que da lugar a una lógica multivaluada donde las afirmaciones tienen asociado un grado de verdad.

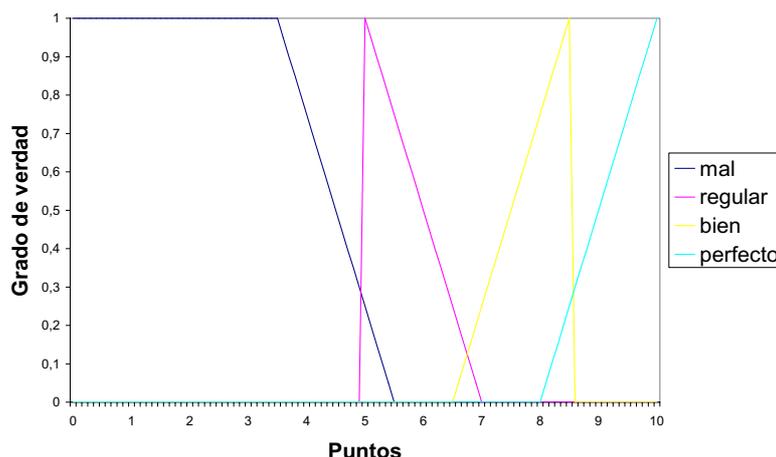


Figura 3. Definición matemática de los conjuntos difusos

Así en este proyecto de innovación se han utilizado cuatro términos difusos: perfecto, bien, mal y regular, junto con tres modificadores lingüísticos: extremadamente, bastante y más o menos. La definición matemática de los términos difusos puede observarse en la figura 3, y la de los modificadores lingüísticos en la tabla 1.

Modificador	Significado matemático
Extremadamente	$x=y^{0.5}$
Bastante	$x=y$
Más o menos	$x=y^2$

Tabla 1. Definición matemática de los modificadores lingüísticos utilizados

Para evitar que la influencia malintencionada de alumnos que evalúen prácticas de forma insincera, se han adoptado dos factores de reducción al peso aplicado a las calificaciones:

1- El peso inicial asociado a las calificaciones provenientes de un determinado grupo es proporcional a la semejanza global de todas ellas con las correspondientes que provienen de los profesores.

2- Para evitar la confabulación de grupos de prácticas diferentes, se aplica una modificación al peso inicial en el agente destinatario de forma que el peso de una valoración de un grupo A respecto de otro grupo B sufrirá una reducción inversamente proporcional a la valoración realizada en el sentido contrario (de B a A). Así si recíprocamente se evalúan con valoraciones muy positivas entre ellos, la aportación de estas valoraciones a la nota final será muy escasa.

De esta forma se les explicará en clase que si todos actúan de forma no sincera, la única valoración que influirá de forma significativa en la nota final será la realizada por los profesores de la asignatura. Así mismo solo se permitirá calificar aquellas prácticas a cuya presentación han asistido.

Implementación

A lo largo del cuatrimestre se han dedicado 12h a la defensa oral de las prácticas, 4h a la utilización de la herramienta informática para editar la notación formal CML2, 6h a

debates, 16h a teoría, 12h a resolución de casos prácticos, y se ha estimado en 6h el tiempo dedicado a las sesiones con el experto (véase figura 4). Así mismo se han realizado 4 entregas parciales de la práctica (una por cada fase) en las que se ha procedido a una corrección sin influencia en la calificación final.

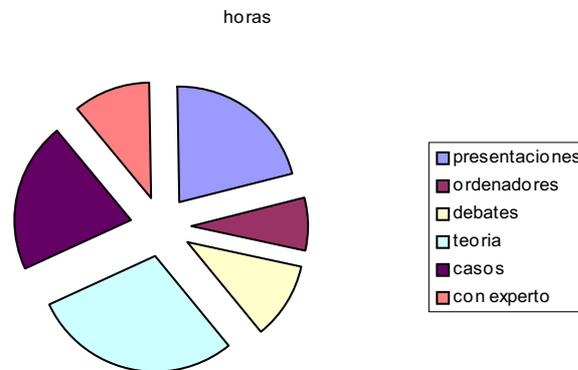


Figura 4. Distribución por sectores de las horas dedicadas a la asignatura

Para implementar todo el sistema de calificaciones distribuidas se ha utilizado el código correspondiente a los agentes escrito ad-hoc en java, que se comunican entre ellos mediante comunicación por sockets. La ejecución local de dichos agentes tiene lugar en la máquina donde reside la página web [12] habilitada para que los alumnos introduzcan sus calificaciones (véase figura 5). La ejecución de el código java asociado a la página se traduce en una comunicación vía sockets con el agente representante del grupo.



Figura 5. Web del proyecto de innovación docente

Discusión y conclusiones

Así mismo se ha estudiado el nivel de satisfacción de los objetivos específicos planteados mediante una encuesta al alumnado sobre su opinión acerca del proyecto de innovación docente. Los resultados de dicha encuesta son:

- Al 81% le pareció adecuado el equilibrio entre teoría y práctica
- Al 95% le pareció adecuado realizar entregas parciales
- Al 77% le pareció adecuado realizar presentaciones orales
- Al 27% le pareció adecuado evaluar a sus compañeros
- El 90% cree que ha asimilado mejor los conceptos gracias a la orientación innovadora de la asignatura
- El 85% dice que ha evaluado con justicia a sus compañeros
- El 86% dice que ha aprendido bastante observando las presentaciones de los otros grupos de prácticas
- El 71% ha cambiado la orientación de su presentación por tener que ser evaluado por sus compañeros
- El 54% se ha sentido más comprometido a escuchar las presentaciones de los demás por tener que evaluarlas
- El 71% cree que la calificación que ha recibido es justa
- El 95% ha dedicado más esfuerzo a esta asignatura que a las demás
- El 81% querría que un esquema similar se aplicara en otras asignaturas

De estos resultados es preciso destacar la baja aceptación de la evaluación de sus propios compañeros que contrasta con las opiniones extendidas de que la nota recibida y dada ha sido justa y de que han asimilado mejor los conceptos, han aprendido de la observación de las presentaciones de los demás y de que la mayoría ha modificado su presentación para convencer a sus compañeros y no solo al profesor. En consecuencia consideramos que esta opinión se debe a una resistencia psicológica a asumir el papel de evaluador que los alumnos no están acostumbrados a ejercer, ya que creen en la justicia del mecanismo evaluador y consideran que ha afectado positivamente a sus presentaciones.

Por último se ha estudiado el efecto producido sobre las calificaciones. De media las calificaciones introducidas por los alumnos han tenido una semejanza del 30% con respecto a las de los profesores. Así mismo la modificación media reflejada en la nota final respecto la que hubieran recibido de los profesores es de un 5% de incremento (0,5 puntos), con una marcada tendencia a hacia los valores intermedios (bajaban las notas elevadas y subían las inferiores). Es preciso destacar especialmente dos excepciones en las que la modificación sobre la nota final ha alcanzado el 15% en un sentido descendente lo que provocó la queja airada a posteriori de los alumnos afectados.

De dicho análisis de los comentarios y sugerencias incluidos en las respuestas al test se concluye que los alumnos no han comprendido suficientemente el proceso de adquisición, no les ha agradado la formalización en CML2 (sugiriendo la utilización de un editor gráfico más amigable que el utilizado) y les interesa profundizar en la fase de conceptualización. Así mismo para mejorar el proceso de calificación distribuida han solicitado la disponibilidad de la memoria de los trabajos durante el día de la presentación.

Por último comentar que la explicación de un sistema distribuido de calificaciones que utiliza conjuntos borrosos para representar las calificaciones nos permite introducir los conceptos básicos de la lógica difusa, la cual está muy relacionada con la ingeniería de conocimiento. Este tema se impartía otros años en la misma asignatura correspondiente a la titulación de 2º ciclo de informática que ahora se extingue, y, actualmente se imparte en la asignatura optativa “Razonamiento con Incertidumbre” correspondiente a la nueva titulación de ingeniería de informática.

Bibliografía

- [1] Gómez A., Juristo, N., Montes C., Pazos, J., “Ingeniería del conocimiento”. Editorial Ramón Areces, 1997.
- [2] Alonso A., Guijarro B., Lozano, A., “Ingeniería del conocimiento”. Pearson Prentice Hall, 2004.
- [3] Schreiber A.T., Wielinga B.J., Breuker J.A. (eds) "KADS: A Principled Approach to Knowledge Engineering". Academic Press, 1993.
- [4] Breuker, J.A. Van de Velde, W. (eds) "The Common Kads Library for Expertise Modelling". IOS Press, 1994.
- [5] CML2 syntax (2.2.1) <http://www.swi.psy.uva.nl/projects/kads22/cml2doc.html>
- [6] Joint Declaration of the European Ministers of Education, Bologna 1999.
- [7] Wooldridge M., Jennings N., “Intelligent Agents: Theory and Practice”, Knowledge Engineering Review 10, No. 2, pp. 115-152 (1995).
- [8] Rao A.S., Georgeff M.P., “Modeling Rational Agents within a BDI Architecture”, Actas del 2nd Int. Conf. on Principles of Knowledge Representation and Reasoning, pp. 473-484, (Cambridge, EEUU, 1991).
- [9] Zadeh, L.A. “Fuzzy Sets”. Information and Control 8, pp.338-353. 1965.
- [10] Carbo J., Molina J.M., Davila J., “Trust management through Fuzzy Reputation”, International Journal of Cooperative Information Systems 12(1), pp. 135-155, 2003.
- [11] Carbo J., Molina J.M., Davila J., “A fuzzy model of reputation in Agent-mediated Electronic Commerce”. Actas del 5th International Conference on Autonomous Agents. (Montreal, Canada, 2001).
- [12] Web del proyecto de innovación docente <http://www.giaa.inf.uc3m.es/id/index.htm>

Agradecimientos

Al servicio de gestión académica de la universidad, el apoyo prestado y la ayuda concedida a esta iniciativa. A Fernando Fernández por su participación en este proyecto de innovación docente. A los alumnos de la asignatura por su colaboración con esta experiencia innovadora. A Elena Cobos por el diseño profesional de la web. A David García por su ayuda con la puesta en marcha del servidor web.