

SISTEMAS TUTORES INTELIGENTES: LOS ESTILOS DEL ESTUDIANTE PARA SELECCIÓN DEL TUTORIZADO

Zulma Cataldi¹, Fernando Salgueiro^{1,2}, Fernando Javier Lage¹, Ramón García-Martínez^{3,2}

liema@fi.uba.ar

1. LIEMA. Laboratorio de Informática Educativa y Medios Audiovisuales y

2. LSI - Laboratorio de Sistemas Inteligentes.

Facultad de Ingeniería. Universidad de Buenos Aires. Av. Paseo Colón 850. 1063 Ciudad de Buenos Aires.

3. Centro de Ingeniería del Software e Ingeniería del Conocimiento. Escuela de Postgrado. ITBA.

Av. Madero 399. Anexo 3º Piso. Ciudad de Buenos Aires. Argentina.

Resumen

En este trabajo de investigación se presentan los pasos realizados en el Laboratorio de Informática Educativa y Medios Audiovisuales conjuntamente con el Laboratorio de Sistemas Inteligentes y el Centro de Ingeniería del Software e Ingeniería del Conocimiento, a fin de identificar las bases metodológicas para desarrollar un sistema tutor inteligente. El propósito es que el sistema de tutorizado exhiba un comportamiento similar al de un tutor humano, es decir, que se adapte al comportamiento del estudiante en lugar de ser un modelo rígido, además de contar con características tales como módulos reutilizables. El sistema debe poder proporcionar un trazado de la construcción de los pasos de la solución del problema como lo hace el tutor humano para que el alumno pueda continuar elaborando “su” solución, pero sin revelar cómo serán los siguientes pasos. La intención de la investigación es desarrollar un tutor que pueda adaptarse a diferentes estilos de estudiantes, con un fuerte base en psicología cognitiva. Para ello, se ha iniciado la elaboración de prototipos basado en redes bayesianas y de tipo SOM [Self Organizing Maps, también conocidas como mapas de Kohonen] entrenadas con datos provenientes de las planillas de Felder [2002]. De este modo se pretende obtener agrupamientos o en “familias” de individuos con estilos similares del conjunto original

Palabras clave: *modelado del estudiante, modelado del tutor, sistemas tutores inteligentes.*

INTRODUCCIÓN

Este tema de investigación surge motivado por la necesidad de encontrar diferentes formas alternativas para la enseñanza de la asignatura Algoritmos y Programación I de la Carrera Ingeniería Informática de la Facultad de Ingeniería de la Universidad de Buenos Aires.

Durante los últimos seis cuatrimestres se efectuó el seguimiento de los alumnos (a través de sus evaluaciones parciales y finales) a fin de saber por qué algunos no llegaban a aprobar la materia. Si bien la baja cantidad de alumnos que terminan la materia, en el orden del 30%, es alarmante, existen datos de Universidades se evidencia que la problemática es muy parecida, por lo que actualmente se encuentran elaborando estrategias tendientes a la paliación del problema. A lo largo de últimos ocho cuatrimestres se han aplicado diversas estrategias didácticas usando medios audiovisuales, foros de discusión, grupos de aprendizaje [Souto, 1995] y se ha observado que si bien se observan algunas mejoras, las mismas apuntan a los grupos de estudiantes que normalmente tienen menores dificultades.

Por este motivo, se pensó, en el desarrollo de un sistema que pudiera ser un “asesor inteligente” (utilizando tecnología de Sistemas Inteligentes) que pudiera llevar a cabo la tarea de tutorizado adaptando diferentes modalidades o estrategias de enseñanza, de acuerdo al estilo de aprendizaje que revelara el estudiante. Lo que se quiere obtener es un sistema de tutorizado que exhiba un comportamiento adaptativo, similar al de un tutor humano, es decir, que se adapte al comportamiento del estudiante en lugar de ser un modelo rígido.

Un tutor inteligente, por lo tanto “*es un sistema de software que utiliza sistemas inteligentes para asistir al estudiante que requiere de un tutorizado uno a uno y lo guía en su aprendizaje, adicionalmente posee una representación del conocimiento y una interface que permite la interacción con los estudiantes para que puedan acceder al mismo*” [VanLehn, 1988].

En un sistema como el descrito, *el modelo del tutor* es el encargado de definir y de aplicar una

estrategia pedagógica de enseñanza [socrática, orientador, magistral, etc.], de contener los objetivos a ser alcanzados y los planes utilizados para alcanzarlos. Es el responsable de seleccionar los problemas y el material de aprendizaje, de monitorear, y proveer asistencia al estudiante. Integrar el conocimiento acerca del método de enseñanza, las técnicas didácticas y del dominio a ser enseñado [con integración de planificación y curriculum] [Coll, 1994]. Esta selección y secuenciación del curriculum para los estudiantes requiere del uso de mecanismos de planificación bastante sofisticados que deben tener en cuenta la teoría de tutorizado empleada de acuerdo a las necesidades del aprendiz [Pozo 1999].

Aún hoy día, la mayoría de los desarrolladores de programas “didácticos” se basan en modelos instruccionales de neto corte conductista [Perkins, 1995]. En este sentido, la intención de la investigación es emular a un tutor humano aunque para la problemática planteada se piensa en un sistema para aprendizaje por refuerzo; ya que justamente el objetivo buscado es que el estudiante le encuentre significatividad a sus aprendizajes, que supere sus dificultades, incorporando lo nuevo de un modo significativo y permanente [Ausubel et al. 1983].

El aporte de este trabajo permitirá esclarecer algunos interrogantes respecto de cómo enseñan los tutores humanos, cuáles son sus estrategias en cada caso y cómo se puede capturar esa experticia, además de dar solución a los estudiantes de clases muy numerosas, que no pueden acceder al docente cuando lo requieren.

JUSTIFICACIÓN

Se ha observado que en los todos los cursos relevados que la cantidad de tutores humanos no es suficiente, la relación alumnos/tutores es demasiado alta y que existe un gran desnivel entre los conocimientos previos que traen los alumnos [Ausubel et al. 1983]. Un sistema de este tipo debería proveer algunas características en función de los propósitos por los que el estudiante recurre a él, tales como: la perspectiva desde la que debe impartir los conocimientos a los alumnos, la forma de adaptación del tutor a los conocimientos previos de los alumnos y la selección de la estrategia de enseñanza más adecuada para el alumno que lo consulta.

Y, cuando el mismo *guíe* al alumno deberá tener “reglas” almacenadas para saber qué hacer en casos cuando: el alumno no puede contestar una pregunta que le hace el tutor o contesta en forma incompleta una pregunta que le hace el tutor.

METODOLOGÍA

En trabajos previos [Šierra, et al. 2004] se ha establecido el marco referencial en el cual cobran valor los sistemas de tipo tutor inteligente en el ámbito universitario y se han determinado los pasos metodológicos para su construcción. Particularmente, en este caso, se trata de un diseño tecnológico orientado a la mejora de los aprendizajes de los estudiantes de grado que ingresan a la universidad. Las diferentes problemáticas de los alumnos deben ser resueltas con un modelado del tutor flexible, lo que es central para el desarrollo.

Gertner et al. [1998] en su trabajo sostiene que el 37% de las preguntas que se presentan cuando los alumnos intentan resolver una problemática corresponden al tipo “*estoy trabado, ¿cómo debo proseguir?*”. En la mayoría de los sistemas el camino a seguir esta pautado [es rígido, de tipo conductista], pero este no es el caso que mejor se adapta a la enseñanza de los Algoritmos, ya que un mismo problema puede admitir soluciones válidas a través de diferentes estructuras. En este caso el tutor debería poder determinar el modo de resolución que adoptó el alumno y proponer, un próximo paso o acción que sea consistente con la solución propuesta por el alumno.

En el caso en que el alumno contesta parcialmente una pregunta, el tutor no debe aceptar como válida solo una respuesta completa. En este caso de respuesta incompleta debe guiar al alumno para que la complete. Hume [1996a, 1996b] a partir de los estudios efectuados a través del comportamiento de los tutores humanos observó que éstos utilizan las “*pistas*” como un método

pedagógico válido, aunque esta táctica es bastante sutil y difícil de implementar en los sistemas tutores inteligentes.

Analizados los casos anteriores se tendrán en cuenta los tipos de respuestas de los alumnos [Yujian Zhou et al, 1999]: a) respuesta correcta, b) respuesta parcial (la respuesta es parte de la respuesta correcta), c) respuesta aproximada, la cual es pedagógicamente correcta pero no la respuesta deseada [Hume et. al, 1995] conceptualmente cerca, d) respuesta incorrecta, pero el alumno demuestra cierto entendimiento del tema y e) respuesta con error conceptual, una confusión de términos o un falso conocimiento del tema que se está explicando. [Cataldi, Lage, 2002]. A partir de cada una de estas respuestas el tutor debe tomar decisiones acerca de las pistas a dar y si el alumno no llegase a la solución deberá dar la determinada “*pista expositiva*” [Hume 1995].

En la literatura analizada se han encontrado dos posturas para la implementación de los conocimientos: una se basa en la estructura sintáctica de lo producido por los tutores humanos [Seu, Jai, 1991] y la otra en las metas pedagógicas que deben cumplir a fin de que el alumno pueda comprender el tema [Hume, et al. 1996, Evens et al, 1993]. Pero, reanalizando el problema y utilizando ambas teorías en forma conjunta se lograrían una serie de pasos que pueden resumir la forma de impartir los conocimientos [Freeva, et al, 1996]:

1. El tutor debe mantener una jerarquía de *metas* que debe cumplir mientras imparte los conocimientos al alumno.
2. El tutor debe poder explicar un mismo concepto de diferentes maneras, así si el alumno no entiende el concepto el tutor puede continuar efectuando otro acercamiento al mismo tema, explicando el concepto para luego continuar, utilizando un método iterativo para profundizar en el concepto cada vez más [paso a paso] o descartar este acercamiento al tema e intentándolo de otra manera.

En este contexto, surgen las posibilidades de aplicabilidad de los sistemas inteligentes a la resolución de problemas de modelado de este tipo. Dentro de los sistemas inteligentes se encuentran las redes neuronales (RN), las cuales son interconexiones masivas en paralelo de elementos simples, los cuales responden a una cierta jerarquía intentando interactuar con los objetos reales tal como lo haría un sistema neuronal psicológico [Kohonen, 1988]. Las redes neuronales poseen una característica que las hace muy interesantes, dado que pueden asimilar conocimiento en base a las experiencias mediante la generalización de casos. [García-Martínez et al., 2003]

GRADO DE AVANCE

Hasta el momento se ha determinado el estado actual del conocimiento en la temática, en este sentido se han analizado STI existentes, a fin de dar cuenta de los métodos de tutorizado utilizados y de obtener datos a fin de identificar los métodos de enseñanza que resultaron más efectivos en relación a la población estudiantil.

Para ello, se están estudiando los estilos de aprendizaje de los estudiantes de las carreras de ingeniería para obtener el perfil de los alumnos para relacionar los estilos de aprendizaje con los métodos o modos de enseñanza más adecuados [Figueroa, Cataldi et al, 2004a, 2004b]. En forma paralela, se están efectuando observaciones de las clases de aquellos docentes que en las encuestas a los estudiantes han sido puntuados con un alto porcentaje en el ítem: “*fomenta el interés por la materia*” a fin de saber como enseñan los *tutores* exitosos. [Denazis, Cataldi, et al. 2004].

Con los datos relevados se pretende tomar el rediseño del módulo tutor planteado por Salgueiro et al. [2005] a fin de establecer las bases para la selección del protocolo pedagógico a utilizar en cada una de las sesiones pedagógicas.

Para ello se utilizarán las herramientas metodológicas que provee la Ingeniería del Conocimiento [García-Martínez y Britos, 2004] y las técnicas proporcionadas por la minería de datos inteligentes [Britos et al, 2005]. El estudio parte de una muestra de 121 estudiantes de una población de 800¹,

¹ Se debe señalar que para una población de 800 estudiantes, trabajando con fórmula que presenta Hernández Sampieri [2001] para el cálculo del tamaño muestral se tiene que: V es el error estándar, V^2 es el error estándar al cuadrado, σ^2 es la desviación estándar, S^2 es la varianza de la

tomando datos mediante el cuestionario de Felder [2002], se formarán grupos de estudiantes con características similares por medio de redes neuronales del tipo Self Organizing Maps. Estas redes, también conocidas con el nombre de mapas de Kohonen [1988] realizan una “clusterización” o agrupamiento a partir del conjunto de individuos original. Se recuerda que las redes SOM utilizan aprendizaje no supervisado, donde las salidas no poseen un valor esperado en la etapa de entrenamiento, por lo tanto los “clusters” resultantes de aplicar la red SOM a los datos de entrada si bien poseen características en común, estas no pueden verse a simple vista.

Las redes SOM, como la mayoría de las redes neuronales tienen dos modos de operación: el proceso de entrenamiento, en donde el mapa se construye y la red neuronal se organiza a si misma por medio de un proceso competitivo, las neuronas de la capa de salida compiten por activarse y sólo una de ellas permanece activa ante un conjunto de entradas determinada. Los pesos de las conexiones se ajustan en función de la neurona que haya resultado vencedora. Luego en la etapa de funcionamiento, el proceso de *mapeo* se realiza rápidamente entre los datos de entrada y la salida de la red, y el dato queda clasificado o categorizado por la red.

Una vez finalizado el proceso de formación del mapa topológico en cluster o familias con características comunes en la información de entrada a la red, se le agrega a cada uno de los individuos originales el identificador del grupo al que pertenece y se utiliza un algoritmo de inducción para hallar las reglas de que caracterizan a cada uno de estos grupos. En este caso se utilizarán algoritmos que pertenecen a la familia Top-Down Induction Trees [TDIT o en castellano Árboles inductivos de arriba hacia abajo], Si bien existen varios algoritmos que realizan estas funciones, uno de los mas completos es el C4.5 de Quinlan [1993], que es una extensión del algoritmo ID3 [Induction Decision Trees] propuesto también por Quinlan [1986]. Su objetivo es generar un árbol de decisión y luego las reglas de inferencia que caracterizan a dicho árbol.

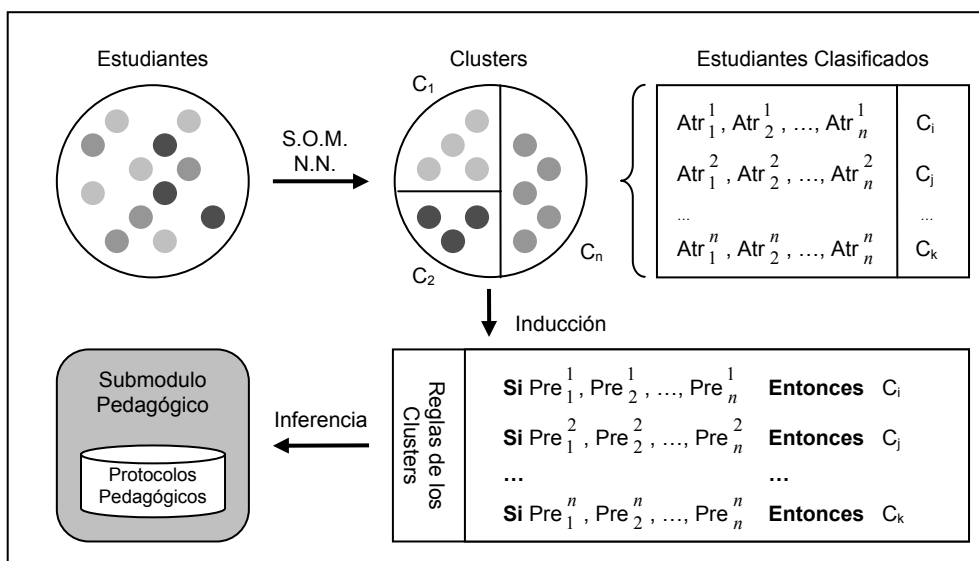


Figura 1: Esquema básico para la inferencia de los distintos patrones para categorizar los estilos de aprendizaje de los alumnos.

Una vez obtenida la menor cantidad posible de reglas por medio de *pruning* para evitar el *overfitting*, se las analiza y a través de un proceso de inferencia se selecciona cuales pertenecen a cada uno de los estilos pedagógicos que se encuentran disponibles en el sistema. En la Figura 1 se puede ver un esquema que representa el proceso completo.

muestra n la cual podrá determinarse en términos de la probabilidad p donde: $V=0.03$ y $V^2=[0.03]^2=0.0009$, por lo que el tamaño de la muestra sin ajustar a la población es: $n = [S^2 / \sigma^2] = 0.09/0.0009 = 100$, corrigiendo en función de la población N se tendrá el tamaño ajustado de la muestra: $n = [n' / (1 + n' / N)] = 100 / (1 + 100 / 800) = 89$ estudiantes. El error de generalización está por debajo de 0.03.

CONCLUSIONES

Se está completando la primera etapa del proyecto centrada en el modelado de las preferencias de los estudiantes, para obtener el método de tutorizado según su estilo de aprendizaje.

Actualmente, el trabajo de investigación se encuentra en la etapa de iteración para encontrar los parámetros de entrenamiento de las redes autoorganizadas de Kohonen. Los pesos de la red son modificados de diferentes maneras con respecto a la cantidad de iteraciones, los rangos aleatorios de los pesos de entradas, los parámetros de aprendizaje, etc.

Además, se está evaluando en qué medida los atributos relevados con las planillas de Felder se utilizarán como entrada para la clasificación a fin de generar los árboles de decisión de profundidad acotada por medio de los métodos de poda [o pruning] adecuados para obtener las reglas de las familias o cluster con la más alta confianza.

REFERENCIAS

- Ausubel, D.; Novak, J. y Hanessian, H. [1983] *Psicología educativa: un punto de vista cognitivo*. México: Trillas.
- Britos, P., Hossian, A., García Martínez, R. Y Sierra, E [2005] *Minería de Datos basada en Sistemas Inteligentes*. Editorial Nueva Librería (en prensa).
- Cataldi, Z.; Lage, F. [2002] *los preconceptos de docentes y alumnos en el proceso de enseñanza y aprendizaje en carreras de grado y posgrado*. CBComp 2002. 26-30 Agosto. Itajaí, Sta. Catarina. Univali.
- Coll, C. [1994] *Psicología y Curriculum*. Paidós
- Evens, Martha W., John Spitkovsky, Patrick Boyle, Joel Michael, Allen A. Rovick. [1993]. *Synthesizing tutorial Dialogues*. Preceedings of the 15th Annual Conference of the Cognitive Science Society.
- Freeva, E. [1996] *Lecture Notes* in Computer Science. Pp 632-640.
- Denazis, J. M.; Cataldi, Z.; Alonso, A.; Ayam, V.; Lage, F. J. [2004]. *Las concepciones epistemológicas y didácticas en la enseñanza de la ingeniería*. IV CAEDI. 1-3 de setiembre. Instituto Tecnológico de Buenos Aires.
- Figuroa, N.; Cataldi, Z.; Costa, G.; Rendón, J.; Salgueiro, F.; Lage, F. y Perichinsky, G. [2004*]. *Nuevos enfoques para el estudio del desgranamiento universitario*. IV CAEDI. Cuarto Congreso de Enseñanza de la Ingeniería. 1-3 de setiembre. Instituto Tecnológico de Buenos Aires.
- Figuroa, N.; Cataldi, Z.; Costa, G.; Rendón, J.; Salgueiro, P. Méndez; F. y Lage, F. [2004b] *Los estilos de aprendizaje y el desgranamiento universitario en Ingeniería Informática*. Aceptado en X CACIC. X Congreso Argentino de Ciencias de la Computación. Universidad de la Matanza.
- Felder, R. y Silverman L. [2002] *Learning and teaching styles in engineering education*, Engineering Education Journal Vol. 78 Num. 7 p. 674-681
- García Martínez, R.; Servente; M. y Pasquini [2003] *Sistemas Inteligentes*. Editorial Nueva Librería.
- García Martínez, R. y Britos, P.[2004]. *Ingeniería de Sistemas Expertos*. Editorial Nueva Librería.
- Gertner, A. S; Conati; C y VanLehn, K. [1998]. *Learning Procedural help in Andes: Generating hints using a Bayesian network student model*. Research & Development. American Association for Artificial Intelligence.
- Hernandez Sampieri, R. [2001] *Metodología de la investigación*. Mc Graw Hil. México.
- Hume G., Michael, J; Rovick, A. y Evens, M. [1996a], *Hinting as a tactic in one-on-one tutoring*. Journal of Learning Sciences.
- Hume G., Michael, J; Rovick, A. y Evens, M. [1996b]. *The Use of Hints by Human and Computer Tutors: The Consequences of the Tutoring*. Illinois Institute of Technology.
- Hume, G. [1995]. *Using student modeling to determine when and how to hint in an intelligent tutoring system*. Illinois Institute of Technology.
- Kohonen, T. [1988] *Self-Organizing Maps* Springer Series in Information Sciences, Vol. 30, Springer, Berlin, Heidelberg, NY [pp 236]
- Perkins, D. [1995] *La escuela inteligente*. Gedisa
- Pozo Municio, I. [1999]. *Aprendices y Maestros*. Alianza.
- Quinlan, J. R. [1986]. Induction of decision trees. *Machine Learning*, 1, 81-106.
- Quinlan, J. R. [1993]. C4.5: Programs for Machine Learning. Morgan Kaufmann.
- Salgueiro, F. A, Costa, G., Cataldi, Z., García Martínez, R. y Lage, F. J. 2005. *Sistemas inteligentes para el modelado del tutor*. GCETE'2005, Global Congress on Engineering and Technology Education. may13-15.
- Seu, Jai, Ru-Cham Chang, Jun Li, Evens, M.; Michael, J. and Rovick, a. [1991]. *Language Differences in Face-to-Face and Keyboard-to-Keyboard tutoring Session*. Proceedings of the Cognitive Science Society.
- Sierra, E., García-Martínez, R.; Cataldi, Z. y Hossian, A. [2004] *Fundamentos para una metodología de diseño de sistemas tutoriales inteligentes centrada en la reparación de mecanismos*. X CACIC. X Congreso Argentino de Ciencias de la Computación. Universidad de la Matanza.
- Souto, M. [1995] *Hacia un Didáctica de lo grupal*. Miño y Dávila.
- VanLehn, K [1988]. *Student Modelling*. M. Polson. Foundations of Intelligent Tutoring systems. Hillsdale. N.J. Lawrence Erlbaum Associates, 55-78
- Yujian Zhou1, Freedman, R, Glass, M., Michael, J.; Rovick, A. y Evens, M. [1999]. *What Should the Tutor Do When the Student Cannot Answer a Question?*. Proceedings of the Twelfth Florida Artificial Intelligence Symposium.