

Una propuesta para el diagnóstico del estudiante en InfoSem

C. Huapaya¹, F. Lizarralde¹, J. Vivas², G. Arona¹

¹Laboratorio de Inteligencia Artificial, Facultad de Ingeniería, UNMDP
{huapaya,flizarra,grarona}@fi.mdp.edu.ar

²Grupo de investigación en Psicología Cognitiva y Educacional, Facultad de Psicología, UNMDP
jvivas@mdp.edu.ar

Resumen

El sistema informático InfoSem tiene por objetivo, entre otros, valorar el conocimiento conceptual de estudiantes universitarios. En el módulo del diagnóstico se encuentra el método DistSem, basado en la medición de las distancias semánticas entre conceptos. Considerando los resultados obtenidos y la naturaleza borrosa de la evaluación de los conocimientos, se ha integrado una técnica basada en lógica difusa. El conocimiento experto de los docentes-evaluadores fue capturado en conjuntos difusos a fin de obtener mayor flexibilidad en la evaluación y de este modo adaptar el aprendizaje de los alumnos de los primeros años de ingeniería.

Palabras claves: diagnóstico del estudiante de ingeniería, aprendizaje adaptativo, conjunto difuso, conocimiento conceptual.

Introducción

Los docentes utilizan diferentes medios para medir el nivel del aprendizaje de los temas que enseñan. Las evaluaciones propias de nuestras universidades comprenden, mayoritariamente, exámenes escritos (ya sea, parciales o finales) cuya evaluación es devuelta al alumno luego de varios días. El docente asigna una nota y de acuerdo a la escala elegida el alumno “aprueba” o “desaprueba” la asignatura, pero no necesariamente aprende los temas de la materia.

Con la finalidad de mejorar el proceso de evaluación proponemos una nueva metodología tanto para medir la adquisición

de aptitudes intelectuales de alto nivel, como para guiar el proceso de enseñanza/aprendizaje.

Llevar esta metodología a un entorno informatizado implica no solo replicar el proceso manual, sino agregarle las ventajas propias de los sistemas computacionales. En este trabajo se ha agregado al método DistSem un módulo basado en conjuntos difusos a fin de manipular la incertidumbre propia del proceso de evaluación.

Evaluación del conocimiento de un estudiante

“La valoración debería ser más que una mera prueba al finalizar la instrucción para conocer como los estudiantes actúan bajo condiciones especiales; en realidad, debería ser una parte integral de la instrucción, que informa y guía a los maestros cuando toman sus decisiones instruccionales. La valoración no debería ser hecha a los estudiantes; debería también ser hecha para los estudiantes, a fin de guiar y mejorar su aprendizaje” (The Assessment Principle, ¶ 1)[1].

La respuesta del estudiante, escrita, oral o mediada por computadora puede ser registrada por sistemas que contemplen los siguientes criterios:

1. Propósito e impacto: como usaremos la valoración y como impactará en la enseñanza y selección de temas?
2. Validez e imparcialidad: estoy midiendo lo que realmente quiero medir? Permite a los estudiantes demostrar lo que saben y lo que podrían hacer?

3. Confiabilidad: son los datos registrados confiables?

4. Significación: las habilidades y contenidos que están siendo evaluados, son considerados valiosos y se fundamentan en las corrientes de pensamiento actuales?

La valoración es el proceso de coleccionar información o evidencias (tomar muestras) desde un dominio de contenidos y habilidades, en un intervalo determinado del tiempo. La hipótesis subyacente de la valoración es que ésta provee una muestra representativa de las habilidades del estudiante. El tipo de muestra permite hacer inferencias sobre los logros, potenciales, aptitudes, inteligencia, actitudes y motivaciones. Esta tasación del conocimiento permite inferir el entendimiento del estudiante de una parte del dominio que está siendo explorado. La muestra puede incluir comportamientos, productos, conocimiento y rendimiento. La valoración es un proceso continuo que involucra examinar y observar el comportamiento de los alumnos, escuchar sus ideas y el desarrollo de preguntas para promocionar el entendimiento conceptual.

Podemos señalar como los propósitos de la valoración los siguientes puntos [2]:

- Asistir al aprendizaje del estudiante
- Identificar fortalezas y debilidades del estudiante
- Valorar la efectividad de una estrategia instruccional particular
- Valorar y mejorar la efectividad de la currícula.
- Valorar y mejorar la efectividad de la enseñanza.
- Proveer datos para apoyar la toma de decisiones.

La evaluación es el proceso por el cual se establece la discrepancia entre la calidad del trabajo del estudiante y los resultados deseables establecidos por el profesor. Con este juicio sobre los resultados se establece un

criterio seguido por la asignación de un valor (i.e. nivel, nota numérica, etc.) que representa la calidad del logro. Este juicio se hace en relación al logro deseable establecido por el currículo, las intenciones pedagógicas y los objetivos, usando la información recogida mediante distintas herramientas de valoración.

Existe una gran variedad de tipos de tarea que puede utilizarse para el diseño de evaluaciones. En el contexto de pruebas, encontramos la selección múltiple, la correspondencia, la respuesta corta, el ensayo corto, la solución computacional, etc. En todos los casos existe una secuencia finita de decisiones: primeramente, se especifica cuales son los objetivos de aprendizaje a fin de medir determinadas aptitudes de los estudiantes; en segundo lugar, se decide posibles pruebas para medir el logro; y finalmente, se diseña las tareas específicas ya sean preguntas en una prueba o actividades de una tarea que armonicen con los objetivos de aprendizaje.

Si nos proponemos como objetivo, p.e., evaluar habilidades cognitivas se puede utilizar los objetivos instruccionales que requieren crecientes niveles de pensamiento, agrupados en las seis categorías de Bloom [3]: conocimiento, comprensión, aplicación, análisis, síntesis y evaluación. Las últimas tres categorías se conocen como habilidades del pensamiento de alto nivel. Estos objetivos consisten en sentencias que afirman que acciones y habilidades observables específicas deben alcanzar los estudiantes.

Tradicionalmente, hay dos usos primarios de la evaluación. El primero es decidir una nota para seleccionar que estudiante aprobó la materia y quien no. En segundo lugar, la valoración del conocimiento del estudiante puede guiar al docente en su actividad de enseñanza. En este caso, el enfoque presenta tres tipos: diagnóstico, formativo y final (o acumulativo). Los tres casos son conceptualmente diferentes, implican roles educativos distintos y conviene manipularlos separadamente.

El objetivo de la valoración diagnóstica es establecer, antes de la instrucción, las fortalezas, debilidades, conocimientos y destrezas de cada estudiante. El conocimiento del perfil del estudiante permite al docente guiar al alumno cuando se inicia el aprendizaje según sus necesidades.

La valoración formativa es el monitoreo del progreso del estudiante durante la enseñanza y aprendizaje, no son juicios finales, debe transcurrir regularmente durante el proceso de enseñanza. Lo podemos visualizar como “diagnósticos en transición” a fin de decidir cual es el próximo paso en la enseñanza o cual es la respuesta a la necesidad creada en el. En su forma pura, esta valoración del progreso del estudiante, es usada para modificar y mejorar la enseñanza de acuerdo a las necesidades y progresos de los alumnos. En este contexto, se puede proveer ayuda parcial si el estudiante lo necesita o probablemente este puede finalizar su aprendizaje sin ayuda adicional.

La valoración final (o acumulativa) es una prueba cuya finalidad es determinar el nivel de logro alcanzado por un estudiante. Usualmente es tomado a la finalización de un período de tiempo (un trimestre, un semestre, etc.) o de una unidad de estudio (p.e. un capítulo). Esto es, luego de completar todas las clases y actividades de aprendizaje se desea ver, si finalmente y sin ayuda, un estudiante puede “resolver solo los problemas”. En este contexto, la autenticación del conocimiento del estudiante es importante. El propósito de esta valoración es llegar a un indicativo válido y confiable sobre las capacidades del estudiante.

Sin embargo, los evaluadores humanos podemos cometer errores cuando valoramos el progreso académico de un estudiante, errores provocados por la presencia de la incertidumbre en las mediciones del logro. La raíz de estos problemas la podemos encontrar en una cuantificación defectuosa de la naturaleza abstracta de la cognición humana

Además, para buscar un diagnóstico realista del estudiante, se debe considerar el constante cambio de sus estados mentales.

Vemos, entonces, que la valoración del progreso del aprendizaje del estudiante, involucrará alguna metodología que trate la incertidumbre. Por este motivo, usaremos la lógica difusa.

Método DistSem

En una red semántica dos conceptos se hallan próximos en la red si se encuentran semánticamente relacionados. Podemos medir la “proximidad” como la distancia entre ambos, esto es la distancia geodésica de los caminos que los unen. Esta técnica produce resultados poco satisfactorios en los modelos de redes semánticas jerárquicas, pues nodos que se hallan en dependencia jerárquica, por lo tanto más distantes, podrían resultar más próximos que nodos ubicados en la misma jerarquía. La Teoría Extendida de Propagación de la Activación [4] es un modelo reticular de búsqueda y comprensión en la memoria humana. La búsqueda es vista como una propagación de la activación desde dos o más nodos conceptuales hasta su intersección. El efecto de preparación (priming) se explica en términos de la propagación de la activación desde el nodo del primer concepto y constituye el proceso básico sobre el que se asienta la comprensión.

Estructuralmente, un concepto es representado como un nodo en una red. Sus propiedades son representadas como vínculos etiquetados con otros nodos conceptuales. Éstos tienen diferente peso según su relevancia para el significado del concepto.

El método DistSem [5] permite extraer la estructura de una red semántica en base a las distancias estimadas entre significados y representar esa información en una matriz de similitudes. Este método se aplica sobre conjuntos de conceptos seleccionados para su

evaluación y se desarrolla por etapas donde se realizan los siguientes tratamientos: conversión de las estimaciones a distancias geodésicas, escalamiento multidimensional, análisis de cluster, visualización y correlación de matrices. Los resultados de estos tratamientos dan lugar a la descripción, análisis y visualización de la red semántica.

Asimismo, DistSem da lugar a la comparación entre dos redes semánticas, considerando a una de ellas un patrón. En este trabajo enfatizamos este punto.

Consideramos la red semántica de un estudiante y la del docente-experto (considerada el patrón) siendo ambas representadas por las matrices de similitud. La adaptación del método DistSem para la evaluación del conocimiento del estudiante capta, además de la comparación del entendimiento del alumno y la del experto, la significancia de la asociación entre las matrices creadas por cada uno de ellos. Los pasos principales del algoritmo "Distsem" resultante son:

1. Toma de datos de la decisión del experto usando una planilla de administración
2. Construcción de la matriz de estimaciones de distancia del experto a partir de su planilla
3. Toma de datos del estudiante X con la planilla de administración
4. Construcción de la matriz de estimaciones de distancia del estudiante X a partir de su planilla
5. Cálculo de las permutaciones construidas sobre la matriz del estudiante X usando el método QAP
6. Cálculo del coeficiente de correlación de Pearson a fin de medir la significancia de la asociación entre el conocimiento del experto y el del estudiante

Considerando que los datos son diádicos e interdependientes (40 pares de conceptos), aplicamos el Procedimiento de Asignación Cuadrática (QAP, Quadratic Assignment Procedure) [6], a fin de medir la

significancia de la correlación observada entre las matrices del experto y la del estudiante. QAP es un procedimiento basado en permutaciones, no paramétrico que preserva la interdependencia entre las díadas. El procedimiento puro debería generar 362880 (9!) permutaciones sobre la matriz de los estudiantes y dejar fija la matriz del experto. A fin de mejorar los tiempos de cálculo, el sistema modifica el procedimiento original eligiendo al azar el cálculo de 2500 permutaciones. Luego se encuentra la distribución de las correlaciones. El nivel de significancia lo buscamos usando la correlación correspondiente a la matriz del alumno en la distribución encontrada. El porcentaje del número de permutaciones correspondiente a esa correlación nos da el nivel de significancia ρ .

El sistema informático InfoSem

InfoSem [7] tiene como objetivo la *integración* del método DistSem en un único proceso computacional. Asimismo se ha tenido como meta *automatizar* muchos de los sub-procesos tales como los cálculos algebraicos y almacenamiento de datos. Otra característica agregada al sistema es su *modularidad* con la finalidad de agilizar la inclusión de nuevos módulos con el mínimo ajuste al resto de ellos.

Esta última característica suministra a InfoSem flexibilidad para la inclusión de futuros módulos de procesamiento asociados al desarrollo de aplicaciones exploratorias de la potencialidad, restricciones y alcance del método de captura, análisis y representación de redes semánticas.

Actualmente, se han iniciado investigaciones en diversos ámbitos y campos de aplicación. Además de los desarrollos vinculados a la evaluación de los progresos escolares y el seguimiento y control de los aprendizajes, se investiga sobre aplicaciones que ponen a prueba la utilidad del método en la operativización cuali y cuantitativa de la Zona

de Desarrollo Próximo, la determinación de patrones de desarrollo y deterioro en enfermedades por trastornos mnésicos derivados de cuadros escleróticos y enfermedad de Alzheimer y, por otra parte, se explora la potencialidad de la captura de la configuración de redes semánticas diferenciales como posibles predictores para la detección temprana e intervención reparatoria antes de la irrupción clínica de patologías alimentarias.

Diagnóstico del estudiante en InfoSem

El proceso de evaluar el conocimiento de un estudiante sobre ciertos tópicos se facilita usando un criterio que concatene los objetivos y competencias que establezcan que aspectos del tópico debe saber el alumno. Siguiendo la evaluación por criterio basado en objetivos instruccionales /competencias, usaremos los niveles propuestos por el docente, a fin de razonar sobre los niveles de aprendizaje alcanzados por los estudiantes.

Una vez que el docente eligió los tópicos a evaluar, prepara las planillas DistSem considerando los objetivos instruccionales perseguidos. En InfoSem se puede definir un número finito de pruebas DistSem para un tópico específico, considerando tres objetivos instruccionales: *conocimiento*, *comprensión* y *aplicación*. Por ejemplo el tópico ‘Solución numérica de una Ecuación Diferencial Parcial’ puede ser evaluado desde el punto de vista de alguno de estos objetivos.

Particularmente, hemos elegido la *comprensión* para la presentación del diagnóstico en este artículo. La comprensión implica demostrar entendimiento sobre términos y conceptos, i.e., usar el conocimiento adquirido para comprender nueva información. Sabiendo que la comprensión es fundamental para la adquisición de nuevo conocimiento, alcanzar este objetivo es de gran importancia para que

el estudiante de ingeniería avance en su aprendizaje. Entre las acciones que el estudiante realiza encontramos explicar, describir, interpretar, clasificar, discutir, generalizar, relatar, seleccionar, resumir. Nosotros agregamos ‘relacionar conceptos’ según la distancia semántica como la acción principal propuesta por el método DistSem.

Siguiendo al algoritmo DistSem inferimos una estimación del nivel de conocimiento del estudiante sobre un tópico específico. Como deseamos que esta estimación sea tan cercana como sea posible a la estipulada por el docente, tratamos de seguir un proceso que involucre la manipulación de la ambigüedad propia de la evaluación. Usamos un modelo cualitativo que clasifica el nivel de aprendizaje sobre un tópico específico en uno de de los cuatro niveles de capacidad {*incompetente, mas o menos incompetente, mas o menos competente, competente*}.

El diagnóstico nos dirá en cual de las cuatro caracterizaciones colocaremos a nuestro estudiante siguiendo los pasos que dados por un profesor. Para tratar el proceso de la inferencia involucrado hemos recurrido a los conjuntos difusos.

Lógica difusa y conjuntos difusos: breve introducción

El término lógica difusa [8,9,10] puede verse desde dos puntos de vista: el primero (más restringido) propone un sistema lógico cuyo objetivo es formalizar el razonamiento aproximado. Si bien la lógica difusa es una lógica multivaluada, no lo es en sentido tradicional de sistemas lógicos multivaluados; los conceptos centrales son variables lingüísticas, formas canónicas, reglas if-then difusas, cuantificadores difusos y modos de razonamiento tales como razonamiento silogístico y razonamiento interpolatorio.

El segundo enfoque (más amplio) es casi sinónimo de teoría de los conjuntos difusos. Esta es una teoría de clases con bordes blandos

y es mas amplia que la lógica difusa en su sentido estricto; contiene, además, aritmética difusa, programación matemática difusa, teoría de grafos difusos, etc.

Los conjuntos difusos son funciones, definidas a partir de un ‘universo de discurso’, X , en el intervalo real $[0,1]$, i.e., la función de pertenencia $\mu: X \rightarrow [0,1]$ especifica un conjunto difuso A y lo podemos representar como $A = \{(x, \mu_A(x)) \mid x \in X\}$. A la luz de la lógica tradicional el par $(x, 0)$ indica que $x \notin A$ y con el par $(x, 1)$ expresa que $x \in A$. Pero, con la lógica difusa, podemos expresar $\mu_A(x)$ con valores intermedios entre 0 y 1, indicando el grado de pertenencia al conjunto de manera gradual.

Como ejemplo veamos el conjunto difuso Agua = {(no potable, 0), (poco potable, 0.25), (medianamente potable, 0.5), (bastante potable, 0.75), (potable, 1) } donde se establece, en distintos grados, la potabilidad del agua. La variación parte de ‘no potable’ ($\mu_{\text{Agua}}(\text{no potable}) = 0$) hasta ‘potable’ ($\mu_{\text{Agua}}(\text{potable})=1$) pasando por 3 grados intermedios de potabilidad creciente. Los valores asignados a la potabilidad fueron dados según nuestra opinión (es una distribución personal). Esta distribución se la conoce como distribución de posibilidad.

Es interesante notar la diferencia entre probabilidad y posibilidad: suponga que le ofrecen dos botellas de agua y le dicen que una tiene probabilidad 0.75 de ser agua potable y la otra tiene una posibilidad de 0.75 de ser agua potable. Cuál elige? Le conviene inclinarse por una posibilidad de 0.75 (dada por el conjunto Agua) porque, si bien el agua no será totalmente potable, puede ser bebida; de otro modo, si bien es alta la probabilidad (0.75) de beber agua potable, corre el riesgo (1 en 4) de beber ácido sulfúrico!

Notemos que la especificación de la función de pertenencia es subjetiva, i.e., estas funciones pueden diferir para diferentes personas; esta subjetividad proviene de las

diferencias individuales en la percepción o expresión de conceptos abstractos. Este es el punto que diferencia la probabilidad (fenómenos aleatorios) de la posibilidad (subjetividad).

La teoría de la lógica difusa provee fortaleza matemática para capturar la incertidumbre asociada con los procesos cognitivos humanos.

Vemos que la lógica difusa es particularmente relevante para tratar fenómenos con incertidumbre y que involucran un cambio suave y progresivo. Por ejemplo, el razonamiento seguido durante un proceso de evaluación, se basa en un continuo cuyos extremos pueden denominarse, por ejemplo, “definitivamente competente” y “definitivamente incompetente” si se desea medir el nivel de pericia alcanzado.

Un alumno puede posicionarse, según algún criterio seguido de evaluación, en algún punto del continuo. El evaluador puede proponer distintas escalas para tomar decisiones, según el aspecto que desea medir. Por ejemplo, si se desea saber la “velocidad en la resolución de problemas” de un estudiante, puede pensarse en una escala cuyos extremos sean “rápido” y “lento”.

Inferencia del diagnóstico

Supongamos que un docente propuso cinco pruebas DistSem para evaluar la comprensión del tópico. Posteriormente a la actividad del estudiante, se calcula el porcentaje de las respuestas correctas y el porcentaje de las respuestas incorrectas. Considerando el coeficiente de Pearson y su significancia sobre cada prueba, sacamos dicho porcentaje.

A continuación veremos el proceso seguido para alcanzar una valoración formativa y diagnóstica del estudiante. Para tal fin usaremos los conjuntos difusos.

Por ejemplo los conjuntos difusos de “habilidad más o menos competente para la

comprensión del tópico x” y “habilidad competente para la comprensión del tópico x” están representados en la figura 1.

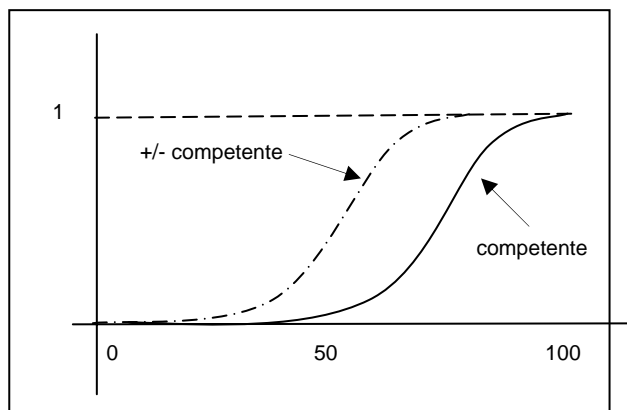


Figura 1: conjuntos difusos para el nivel competente

Notar que hay otros dos conjuntos difusos (para los niveles ‘más o menos incompetente’ e ‘incompetente’).

Cada conjunto difuso está representado por la función $f: A \rightarrow [0,1]$ siendo A el universo de discurso de la función y el intervalo $[0,1]$ nos da el grado de pertenencia al conjunto difuso. Para una entrada x (que representa el porcentaje de respuestas correctas), calculamos el grado de pertenencia de x a un conjunto difuso. P.e., $f(80) = 0.75$ (en la curva llena) significa que el conocimiento alcanzado por un estudiante que respondió correctamente el 80 % de las pruebas en el nivel de *comprensión*, puede ser considerado ‘competente’ con un grado de 0.75.

Para la creación de estas curvas se recurre a la opinión del docente-evaluador sobre el nivel considerado. Por ejemplo la respuesta a la pregunta: ¿si un estudiante ha respondido correctamente el 30 % de las pruebas, cuál considera Ud. que es el nivel de competencia alcanzado? Las respuestas sirven para construir el conjunto difuso.

Por razones prácticas la información solicitada es discreta obteniendo, por ejemplo, para $x = \{0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100\}$ la siguiente distribución $f(x) = \{0, 0, 0, 0, 0,$

$0.15, 0.2, 0.5, 0.75, 0.85, 1\}$ considerando el nivel de capacidad ‘competente’.

Una vez que el alumno ha alcanzado un determinado nivel sobre determinadas pruebas (en nuestro caso cinco tests para estimar el nivel de logro en un tópico específico) podemos continuar con la evaluación de otros tópicos, repitiendo el proceso.

El procedimiento presentado puede resultar de gran utilidad para la valoración formativa toda vez que los resultados de las pruebas guían al profesor sobre los niveles de pericia alcanzado por sus alumnos en la comprensión de las relaciones semánticas de los conceptos que componen algún tópico.

Pero, si ese docente desea brindar ayuda (feedback positivo) sobre un tópico, conviene recurrir a los conjuntos difusos que representan los niveles ‘incompetente’ y ‘mas o menos incompetente’ (ver figura 2). Por ejemplo, el docente puede decidir que si el grado de ‘incompetencia es mayor que 0.7’ se debe brindar ayuda inmediata a ese alumno.

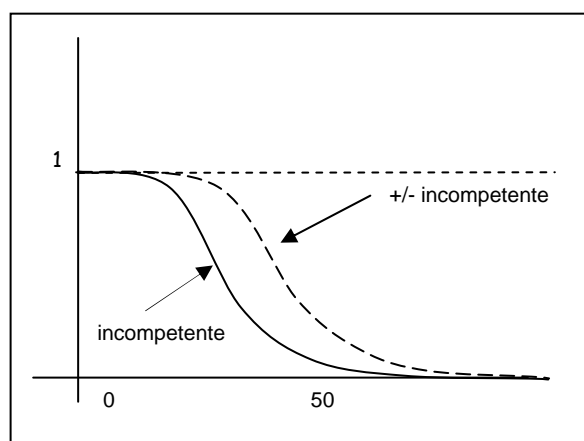


Figura 2: conjuntos difusos para el nivel incompetente

Del mismo modo, las pruebas DistSem pueden ser usadas para una valoración diagnóstica del estudiante. En este caso se puede recurrir al conjunto difuso ‘mas o menos incompetente’ para medir el nivel de desconocimiento que posee el alumnos sobre el tópico que interesa enseñar.

Conclusión

Infosem, como herramienta computacional que asiste al docente e investigador educativo, ha incorporado un procedimiento diagnóstico adaptativo del alumno basado en conjuntos difusos. El punto central del trabajo fue unir el método DistSem, ya probado previamente, con una técnica que trata la incertidumbre propia de la evaluación a fin de incorporar el conocimiento experto del docente en conjuntos difusos y mejorar la flexibilidad en la asistencia al estudiante.

El uso de conjuntos difusos dio lugar a una evaluación mejorada con respecto al tradicional aprueba/desaprueba, considerando básicamente la evaluación diagnóstica y formativa como guía para mejorar la enseñanza.

Referencias

- [1] National Council of Teachers of Mathematics, Principles and Standards for School Mathematics, 2001.
- [2] R. Kellough, N.Kellough, Secondary School Teaching: A Guide To Methods And Resources; Planning For Competence, Prentice Hill, 1999.
- [3] B. Bloom, Taxonomy of Educational Objectives, The Classification of Educational Goals: Handbook I, Cognitive Domain, Longmans, 1956.
- [4] A. Collins , E. Loftus, A spreading-activation theory of semantic processing, Psychological Review, 82, 407-428, 1975.
- [5] J. Vivas, Método Distsem: procedimiento para la evaluación de distancias semánticas. Revista Perspectivas en Psicología, 1, 56-61, 2004.
- [6] D. Krackhardt, QAP partialling as a test of spuriousness, Social Networks 9, 171-186, 1987.
- [7] C. Huapaya, G. Arona, F. Lizarralde, J. Vivas. InfoSem: Distancia Semántica entre conceptos como base de la Evaluación Cognitiva. En el libro "Las ciencias del Comportamiento en los albores del Siglo XXI". UNMDP-Conicet. , 405-411, 2005.
- [8] L. Zadeh. Fuzzy sets. Inf. and control. Vol 8, 338-353.1965.
- [9] L. Zadeh. Fuzzy logic. IEEE Computer. Vol 21(4), 83-93. 1988.
- [10] L. Zadeh. Fuzzy logic, neural networks and soft computing. Communications of the ACM. Vol 37(3), 77-84, 1994.