

# Modelo de evaluación del conocimiento en un Sistema Tutorial Inteligente

Constanza Huapaya<sup>1</sup>, Francisco Lizarralde<sup>1</sup>, Jorge Vivas<sup>2</sup>,  
Graciela Arona<sup>1</sup>

<sup>1</sup>Facultad de Ingeniería, Universidad Nacional de Mar del Plata, Mar del Plata, Argentina

<sup>2</sup>Facultad de Psicología, Universidad Nacional de Mar del Plata, Mar del Plata, Argentina

## Resumen

El presente trabajo propone modificaciones de los modelos del instructor y del estudiante de un Sistema Tutorial Inteligente a fin de incorporar un nuevo método evaluativo del progreso del estudiante. Para tal fin, se integra un modelo basado en la distancia semántica (método DistSem) para mejorar el entendimiento del sistema computacional, sobre los logros del estudiante de ingeniería de los primeros años. El docente puede interactuar con el sistema computacional a fin de proponer pruebas de evaluación y visualizar resultados de exámenes diagnósticos, formativos o acumulativos tanto de grupos de estudiantes o de estudiantes en forma individual.

*Palabras clave:* evaluación del estudiante, distancia semántica, sistema tutorial, modelo del instructor, modelo del estudiante, ingeniería.

## 1. Introducción

La complejidad de la evaluación del conocimiento de los estudiantes es un tema ampliamente conocido. Los educadores 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 sean, parciales o finales). El resultado de estas evaluaciones es retornado al alumno luego de días o semanas. 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 adquirir niveles de logro de aptitudes intelectuales de alto nivel, como para guiar los procesos de enseñanza y de aprendizaje. La transposición de la metodología a un entorno informatizado complica el proceso toda vez que en este medio la evaluación sea continua y por lo tanto es necesario introducir mecanismos especiales que deben ser implementados en un medio informático.

## 1.1. Enseñanza de la Ingeniería

La necesidad de un cambio positivo hacia una educación activa, participativa y creativa en ingeniería ha sido señalada en la literatura [1], [2]. La nueva modalidad se concentra en modificar el contexto del aprendizaje a fin de mejorar el ámbito donde se desarrolla la enseñanza tradicional.

Generalmente, los cursos dados en la Universidad ponen un mayor énfasis en el entendimiento de grandes cantidades de información en lugar de enfatizar su aplicación a diversos problemas. Los conocimientos anteriores se controlan a través de cursos previos (pre-requisitos). La tendencia de los cursos tradicionales universitarios es entregar la misma información de la misma manera a todos los estudiantes. Además de estas limitaciones, se observa que los estudiantes difieren en sus características personales, en el esfuerzo que emplean en cada curso, en el tiempo que le dedican, etc. La comunidad que trabaja en Inteligencia Artificial en Educación busca ambientes de aprendizaje basados en computadora, que colaboren y mejoren el proceso de enseñanza/aprendizaje tradicional. Los Sistemas Tutoriales Inteligentes (STIs) son una de las herramientas más logradas.

En la enseñanza de la ingeniería, claramente se perfilan, el estilo de enseñanza de los profesores, el estilo de aprendizaje de los estudiantes y una interacción, a veces conflictiva, entre ellos. Felder y Brent sostienen que aquellos instructores que mejor entiendan las diferencias, tendrán una mejor oportunidad de conocer las diversas necesidades de aprendizaje de todos sus estudiantes [3].

En general, la falta de formación docente es un problema que sufren nuestras universidades. Asimismo, los docentes no tienen muchas oportunidades de seguir el progreso continuo de sus estudiantes, principalmente debido al carácter masivo de los cursos. La metodología propuesta, aún fuera de un STI, puede asistir al profesor para “ver de cerca” el progreso en el desenvolvimiento individual de cada alumno. Actualmente, nuestro grupo de investigación se encuentra diseñando un STI

especializado en Ingeniería y desarrollando dos de sus módulos [4].

## 2. Sistemas tutoriales

La comunidad científica que estudia y desarrolla los Sistemas Tutoriales Inteligentes [5] aspira a que estos sistemas informáticos muestren habilidades similares a las de un docente real en cuanto a enseñar un determinado tópico. Para lograrlo, un STI debería poseer la capacidad de entender, aprender, razonar y resolver problemas. Además, el sistema debe identificar las fortalezas y debilidades de un estudiante particular a fin de establecer un plan instruccional que debe ser consistente con los resultados obtenidos. Debe encontrar la información relevante sobre el proceso de aprendizaje de ese estudiante (como estilo de aprendizaje) y aplicar el medio de instrucción más adecuado a sus necesidades individuales.

Para que el comportamiento de un STI sea realmente efectivo, éste debe poseer componentes que permitan modelar, el conocimiento a enseñar (dominio), el seguimiento de la actividad del estudiante (modelo del estudiante), el conocimiento sobre cómo enseñar los tópicos (modelo pedagógico) y la interfaz de comunicación (tanto con el estudiante como con el profesor). A continuación presentamos los cuatro primeros modelos, por considerarlos relevantes para nuestro estudio.

### 2.1. Modelo del dominio

Un sistema tutorial es creado para enseñar un tema en particular, por lo tanto debe tener suficiente conocimiento del tema y sobre la forma de enseñarlo. Este conocimiento debe estar estructurado de manera tal que el sistema construido pueda ser flexible. Los factores que hay que tener en cuenta son: cómo organizar los ítems; cómo relacionarlos con el avance del aprendizaje del alumno; cómo relacionarlos con el material de enseñanza.

La ontología propuesta a nivel conceptual se divide en dos estructuras, la descripción del dominio y la organización del material de enseñanza.

La metodología de representación mediante mapas conceptuales ha sido utilizada con éxito para el diseño instruccional [6]. El dominio es representado por nodos que representan los tópicos a enseñar y arcos que representan las relaciones de precedencia, inclusión o pertenencia entre los mismos. Esta estructura permite al sistema seleccionar el tópico a enseñar considerando el cumplimiento de los nodos que lo preceden y según la estrategia de enseñanza que se desee aplicar (de lo general a lo particular o viceversa). La enseñanza de distintos tipos de contenidos requiere distintas

metodologías, por lo que el sistema necesita reconocer estas diferencias. Se han establecido tres tipos de nodo: hecho, concepto y procedimiento. La granularidad con que se defina este mapa conceptual está directamente relacionado con el nivel de detalle (la granularidad) con que el sistema enseñará.

Relacionado con cada nodo del mapa de tópicos, el material de enseñanza se clasifica por el tipo de la instrucción que brinda (explicación, ejemplo, ejercicio, test, ayuda) y por las características propias (importancia, dificultad, tiempo). Dicho material es presentado al estudiante considerando su perfil y actuaciones anteriores, los cuales se hallan representados en el modelo del estudiante.

### 2.2. Modelo del estudiante

El modelo del estudiante es una descripción declarativa de las características de las actividades de aprendizaje de los estudiantes [7]. Nuestro STI usa para registrar su entendimiento sobre la evolución del aprendizaje del estudiante un modelo *overlay* básico (red de tópicos donde se compara el conocimiento del estudiante con el conocimiento experto).

El modelo se ha organizado en dos módulos principales: perfil del aprendiz y sesión individual del estudiante. El primero se refiere a las características personales y la evolución del conocimiento del estudiante a largo plazo, mientras que el segundo considera a la sesión globalmente para terminar analizando el “camino” recorrido y el nivel de rendimiento alcanzado por el alumno. Los resultados de ese análisis actualizan su perfil y su evolución personal.

El comportamiento del STI depende fuertemente del modelo del estudiante, sin embargo éste sólo puede representar una serie de fotografías instantáneas de la actividad del alumno en determinados momentos. La consecuencia inmediata es el registro incompleto del proceso de aprendizaje individual de cada estudiante y por lo tanto sólo disponer de modelos del estudiante inexactos.

### 2.3. Modelo pedagógico

Un currículo es una estructura que relaciona el tema a enseñar en términos de capacidades a adquirir por el estudiante, objetivos instruccionales, métodos didácticos y recursos pedagógicos. Específicamente, un currículo es una organización en red compuesta de nodos que representan capacidades, objetivos instruccionales definidos sobre esas capacidades y recursos pedagógicos que contribuyen al cumplimiento de los objetivos instruccionales. Los enlaces de esta red relacionan los objetivos instruccionales con las capacidades.

La red de transiciones de capacidades es la base para la representación y planificación de las estrategias tutoriales (adaptada de [8]). La red consta de componentes básicas denominadas transiciones pedagógicas (ver figura 1). Una capacidad de entrada  $C_E$  de esta transición es la capacidad pre-requisito para alcanzar el objetivo instruccional involucrado en la transición. Una capacidad de salida  $C_S$  de esta transición es la producida por el cumplimiento del objetivo instruccional involucrado. Las capacidades estarán construidas a partir de los nodos del mapa conceptual.

El enlace que une una capacidad de entrada y un objetivo instruccional se llama enlace pre-requisito

(●→) y el enlace que sale la transición y llega a la capacidad de salida se denomina enlace contribución

(→).

Un enlace pre-requisito desde la capacidad  $C_E$  al Objetivo Instruccional expresa el hecho que  $C_E$  es una pre-condición para la activación de OI. Este enlace se caracteriza por su naturaleza (obligatorio u opcional).

Un enlace de contribución califica la manera en la cual la realización de una transición contribuye a la adquisición de la capacidad  $C_S$ . El peso de esa contribución se asigna a este enlace y se lo califica como débil, moderado o fuerte. En la red de transiciones, varios nodos de transición pueden llegar a una capacidad que tiene ser adquirida por el estudiante a fin mejorar su aprendizaje.

A cada Objetivo Instruccional se le asocian recursos pedagógicos que contribuyen a su cumplimiento.

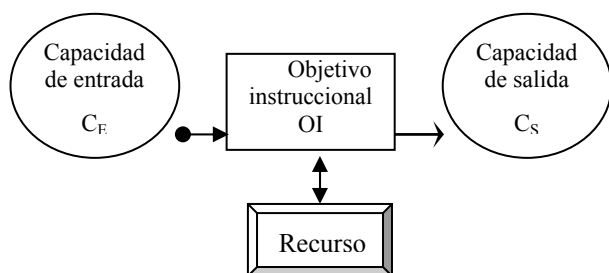


Figura 1: transición pedagógica

La red de transiciones de capacidades es la base para manipular el conocimiento estratégico del instructor y crear planes instruccionales altamente adaptados, dando lugar al Modelo Instruccional Global (MIG).

A partir de este modelo, se construyen las estrategias tutoriales como reglas de producción. Cada producción tiene como antecedente los vínculos entre los nodos del mapa conceptual y como consecuente las unidades de

aprendizaje (UA). Se consideran los enlaces del mapa conceptual (pre-requisito, es\_un y parte\_de) y los vínculos del MIG (pre-requisitos obligatorio y opcional; contribución fuerte y moderada).

## 2.4. Modelo del instructor

La enseñanza humano-a-humano raramente incorpora estrategias tutoriales ideales. Algunos de estos mecanismos pueden ser simulados en computadora, (como el método Socrático) mientras que otros, como aquellos donde intervienen el afecto o la emoción, son muy difíciles de lograr considerando el actual desarrollo de la tecnología computacional.

La actividad pedagógica de un STI, hasta ahora, se ha traducido en colecciones de reglas que sólo trabaja relativamente bien en la práctica. En parte, esto se debe a que no existe una enciclopedia científica donde consultar las mejores heurísticas tutoriales. En su lugar, se debe adoptar diferentes métodos de enseñanza porque diferentes estudiantes utilizan diferentes estilos de aprendizaje. Como afirma du Bulay [9] la riqueza del conocimiento del docente experimentado puede ser tomada como una de las fuentes más importantes donde extraer las líneas principales del conocimiento pedagógico de un STI.

En muchos STIs el *modelo del instructor* está implícitamente incluido y sólo recientemente ha sido desarrollado explícitamente [10]. La actividad de un instructor involucra tanto conocimiento sobre el tema del dominio como sobre la pedagogía propia de esa área del conocimiento. A partir de la observación de la actividad de los instructores puede deducirse que la funcionalidad aportada por este modelo comprende los siguientes aspectos:

- Mejoramiento en la adaptación y realismo de las estrategias tutoriales.
- Provisión de ayuda inteligente individualizada para el autor/instructor.
- Refinamiento en el diseño de STIs particulares.
- Aliento a instructores novicios para que se involucren en el diseño de sistemas tutoriales
- Mejoramiento en la coordinación y cooperación entre varios instructores durante la creación del mismo STI.
- Registro de la actividad de cada instructor.

Durante la etapa de creación del STI, los requerimientos pedidos por el sistema al autor (concretados a través de un modelador) pueden ser hechos siguiendo dos modalidades: a través de *preguntas directas* sobre sus preferencias personales y hábitos comunes (p.e. estilo

de enseñanza o periodicidad de exámenes) o a través de la *observación de la actividad del instructor* registrando los cambios sobre el perfil del instructor.

Un autor/instructor, cuyo nivel de experiencia puede ser muy variado, puede comenzar proponiendo un curso inicial, donde inicializa las estructuras básicas de los modelos del dominio y el material instruccional asociado, estudiante (grupo de estudiantes que tomarán el curso), parámetros del plan instruccional (pedagogía del tema del curso). El curso resultante puede ser revisado posteriormente, tanto por el mismo instructor como por otros.

Si el nuevo instructor posee un nivel superior de experticia, puede mejorar el diseño del curso. Si por el contrario, su nivel es inferior, puede aprender sobre desarrollo de cursos tutoriales inteligentes. I.e., el diseño e implementación de un STI queda sometido a un proceso de *refinamiento iterativo*, cuyo resultado será alcanzar sistemas tutoriales de mayor calidad, incentivando la cooperación de los autores.

El modelo del instructor examina la actividad de cada instructor así como la de los estudiantes a fin de verificar la eficacia de las estrategias tutoriales del plan instruccional. El resultado de ese análisis contribuye a mejorar el conocimiento pedagógico del sistema a través del modelo propuesto.

Nuestro modelo del instructor consta de dos módulos principales: el módulo permanente y el módulo dinámico. El primero almacena los datos del instructor de largo plazo como sus hábitos y preferencias. El segundo guarda la información relevante de las sesiones del STI con cada estudiante.

### 3. Sistemas de evaluación

“La valoración debería ser más que una mera prueba al finalizar la instrucción para conocer cómo 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” [11].

La respuesta del estudiante, escrita, oral o mediada por computadora debe 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 [12]:

- 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 de juzgamiento de la calidad del trabajo del estudiante (juicio del maestro sobre los resultados de la valoración), estableciendo un criterio, y seguida por la asignación de un valor (i.e. nivel, nota numérica, etc.) que representa tal calidad. Este juicio se hace en relación al logro del currículo y objetivos, expectativas y resultados, usando la información recogida mediante una variedad de 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 otros contextos se utilizan tareas, proyectos, desarrollo de productos, muestras, etc. Pero, en todos los casos existe una secuencia finita de decisiones. Primeramente, clarificar cuales son los objetivos de aprendizaje a fin de medir determinadas aptitudes de los estudiantes; en segundo lugar, decidir posibles pruebas para medir el logro; y

finalmente, diseñar 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 podemos utilizar los objetivos instruccionales que requieren crecientes niveles de pensamiento, agrupados en las seis categorías de Bloom [13]: 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 sobre acciones y habilidades observables específicas que deben alcanzar los estudiantes.

Sin embargo, no todos los objetivos instruccionales pueden ser evaluados en forma directa. Si tenemos como objetivo “reproduzca el procedimiento de calibración de un tornó”, la evaluación de la prueba sólo consiste en observar directamente como el estudiante ejecuta el objetivo. En este caso es relativamente simple alinear la evaluación del objetivo con las acciones del estudiante.

Pero otras capacidades no son directamente observables. Las evaluaciones del logro de nuestros estudiantes operan sobre acciones, como emitir juicios, interpretar textos y gráficos y formar opinión, diseñar soluciones en el contexto de problemas abiertos, interpretar datos, etc. Muchas de las capacidades que contribuyen al logro de los objetivos instruccionales pueden describirse como completamente cognitivos. En estos casos el diseño de tareas de evaluación involucra el diseño de indicadores indirectos para las capacidades buscadas. Aunque no podamos observar directamente la capacidad, diseñamos una metodología que mida el logro del estudiante, y si la consideramos significativa, la aceptamos como indicador confiable de tal capacidad.

El punto inicial de la evaluación es diseñar pruebas, en un sentido general, para que sean ejecutadas. Y estas pruebas deben estar alineadas con los objetivos instruccionales. Nuestras tareas para evaluar deben diseñarse para que sean oportunidades en las cuales los estudiantes puedan demostrar sus logros en el aprendizaje de nuestra materia.

Tradicionalmente, hay dos usos primarios. 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 tutorial. 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, conocimiento y destrezas de cada estudiante. El conocimiento del perfil del estudiante permite al instructor (o al STI) 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 instruccional. 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, posee pruebas sin notas y sólo son usados para modificar y mejorar la enseñanza de acuerdo a las necesidades y progresos de los alumnos. La evaluación formativa es una parte integral de la interacción en curso entre docente y alumno. 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 sólo 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.

Nuestra propuesta considera la creación de una nueva prueba basada en la distancia semántica.

## 4. Distancia semántica

Si bien una distancia relativa puede entenderse como una medida entre objetos a través de la inspección de valores de determinados atributos, en este trabajo se propone el uso de la distancia semántica como la distancia entre conceptos, diferenciándola de la distancia física, lingüística o lexicográfica.

Existen diferentes modos para determinar la distancia semántica entre objetos. Por ejemplo, para datos numéricos puede usarse funciones de distancia entre ellos, como la euclidiana. Para datos no-numéricos, puede proponerse alguna forma de procedimiento de similitud a fin de correlacionar instancias no-numéricas con valores numéricos. Tales métodos producen un valor numérico indicando la cercanía de valores de acuerdo a alguna convención aceptada. La diferencia o

similitud entre dos objetos (o conceptos) pueden ser una función de varios atributos. Por ejemplo, la similitud entre las ciudades de Mar del Plata y Miramar puede ser estudiada analizando los recursos turísticos de cada una, clima y distancia física a la Capital Federal, por ejemplo.

Muchas aplicaciones utilizan modelos basados en grafos en lugar de árboles o jerarquías simples, tales como modelos de clasificación, rutas de transportes, etc. El modelo que vemos a continuación trata la distancia semántica entre conceptos representados en una red semántica.

#### 4.1. Método DistSem

En una red semántica dos conceptos se hallan semánticamente relacionados si se encuentran próximos en la red. 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 [14] 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 [15] 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. Además DistSem da lugar a la comparación entre dos redes semánticas, considerando a una de ellas un patrón.

El método para la evaluación de distancias semánticas DistSem constituye un procedimiento amplio y flexible

que permite diferentes niveles de análisis tanto cualitativos como cuantitativos. La eficiencia de este método ha sido demostrada prácticamente con pruebas DistSem en el ambiente universitario [16], [17], [18].

En este trabajo se describe cómo el método Distsem ha sido incorporado en los módulos del estudiante y del instructor de un STI a fin mejorar el entendimiento del sistema sobre los progresos de un estudiante en temas de Ingeniería. El aspecto analizado y desarrollado en este trabajo es el enfoque cuantitativo.

#### 4.2. Evaluación basada en DistSem

La implementación de medios de evaluación como selección múltiple, verdadero/falso, etc., herederos de la tradición asociacionista en Psicología, involucran la programación de componentes específicos, así como la adquisición y procesamiento de la información producida por el estudiante. El STI debe disponer de algoritmos para analizar estos resultados y decidir la siguiente acción que el estudiante debe ejecutar.

Este tipo de mecanismos de evaluación son los más comúnmente usados por su simplicidad de implementación y procesamiento. Sin embargo presentan, algunas limitaciones basadas en la insuficiente representación de la actividad cognitiva. En un sentido amplio, el objetivo más importante de todo aprendizaje en el abordaje cognitivo es la comprensión, no los comportamientos observables que pueden ser medidos. La comprensión, como señalan DiVesta y Rieber [19], es un objetivo efectivo de la educación. Si el docente identifica las estructuras de conocimiento previo del aprendiz, puede evaluar y predecir la forma en que va a asimilar la novedad presentada.

Algunos desarrollos ([20], [21]) argumentan que los supuestos de la teoría asociacionista son inadecuados para prescribir estrategias instruccionales que tengan como objetivo la comprensión. Según Garrison [22] “existe poca posibilidad de negociar los objetivos del aprendizaje, de tomar decisiones en colaboración o, para los aprendices, asumir la responsabilidad de construir significados por sí mismos, basados en sus estructuras de conocimiento previo”. En una posición extrema, Winn [23] sugiere que el análisis de las condiciones y la selección de las estrategias, antes de la implementación, son adecuados para el logro de objetivos cognitivos de bajo nivel, la transferencia de información y la memorización mecánica. Empero resultan incompatibles con objetivos de alto nivel como la comprensión profunda en dominios complejos.

Es por este motivo que la evaluación del conocimiento de un alumno de Ingeniería del ciclo básico comienza con una nueva prueba, basada en DistSem. El docente propone nueve conceptos centrales al tópico que se

desea evaluar. Por ejemplo, en la tabla 1 vemos parte de una planilla sobre tópicos de la asignatura Análisis Numérico para Ingeniería. El estudiante debe estimar que tan relacionados están los 40 pares de conceptos, marcando con un círculo el valor correspondiente. El valor 1 significa conceptos poco relacionados y el valor 7 significa conceptos muy relacionados.

1	Punto Fijo - Converg. Lineal	1 2 3 4 5 6 7
2	Bisección - Punto Fijo	1 2 3 4 5 6 7
3	Lagrange - Bisección	1 2 3 4 5 6 7
...	.....	.....
39	Interpolación - Bisección	1 2 3 4 5 6 7
40	Converg. Cuad. - Newton	1 2 3 4 5 6 7

Tabla 1: vista parcial de la planilla de administración de Distsem

A partir de la planilla de administración, se construye una matriz de similitudes, la cual refleja la red semántica del estudiante. De la misma manera, construye una red semántica patrón propuesta por el instructor/experto. La comparación de las redes del experto y de un alumno particular, formarán parte de la evidencia que guiará al STI para decidir el marcado de su Modelo Instruccional Global. La adaptación del método DistSem a nuestro STI 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\_2” 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) [24], 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  $\rho$ . Luego se construye una correspondencia con nuestras tradicionales notas entre 0 y 10 puntos.

### 5. Nuevo modelo del instructor

Como ya mencionamos, los dos módulos principales del modelo del instructor constan de:

1. Módulo permanente: datos personales, nivel de experticia instruccional, nivel de experticia en herramienta, perfil, preferencias, hábitos, STI creados
2. Módulo dinámico: identificador STI, tipo, plan instruccional (implementación del Modelo Instruccional Global), frecuencia de uso por los estudiantes, test creados DistSem, resultados finales, informe para el estudiante, acciones inesperadas durante la creación STI

El nuevo modelo (ver figura 2) incluye la opción para que el autor/instructor diseñe y evalúe pruebas DistSem como un nuevo método a los ya creados.

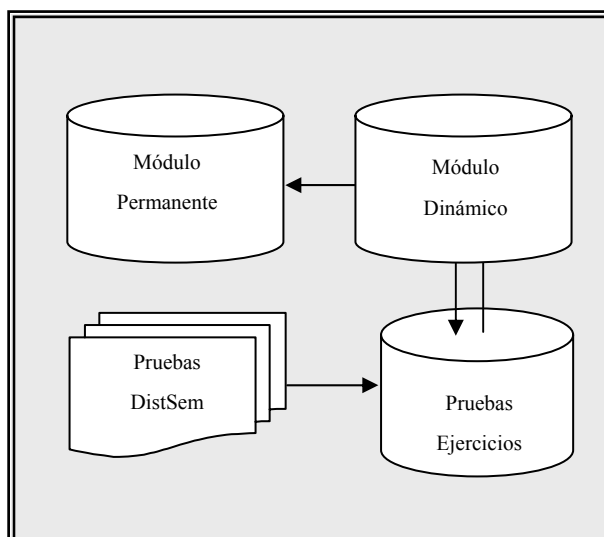


Figura 2: Nuevo modelo del instructor

Estas pruebas son construidas a través de un editor especializado (ver figura 3). El tema de las pruebas, actualmente en uso, pertenece a la asignatura Análisis Numérico. Posteriormente el sistema crea

aleatoriamente una planilla con los cuarenta pares de conceptos (ver figura 4).

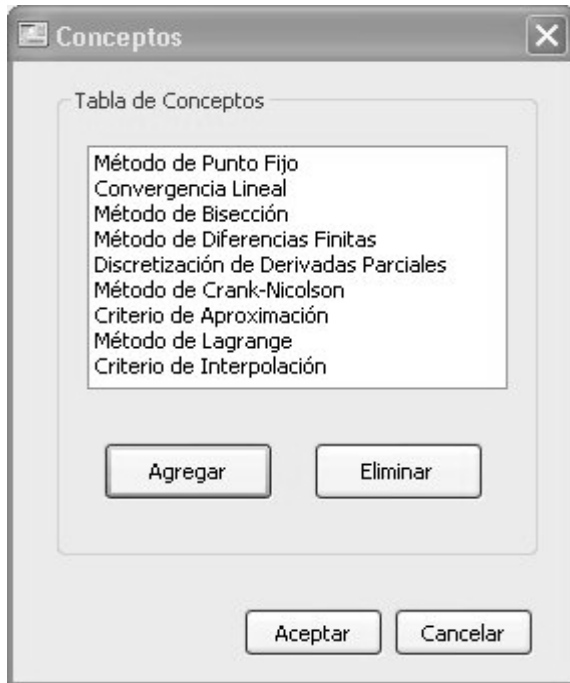


Figura 3: editor de conceptos



Figura 4: planilla de administración de la prueba DistSem

El autor puede modificar esta planilla las veces que lo desee. La eficiencia de las pruebas se comprueba durante el proceso iterativo de la creación de un STI.

Las pruebas DistSem son almacenadas en el módulo de “Pruebas y ejercicios” y luego son incluidas en el módulo dinámico, específicamente en el Plan instruccional. Luego de cada sesión estudiante/STI las pruebas DistSem son actualizadas consignando el éxito o fracaso de la prueba durante su ejecución. El resultado vuelve al módulo de Pruebas y ejercicios a fin de descartar las que fallaron. Igualmente los resultados, actualizan el perfil del instructor, incluido en el módulo permanente.

Este procedimiento mejora el conocimiento del sistema sobre el progreso del estudiante y en consecuencia las estrategias tutoriales alcanzan un mejor nivel de eficacia (en el modelo pedagógico).

## 6. Nuevo modelo del estudiante

El nuevo modelo del estudiante, agrega a los ya tradicionales “perfil del estudiante” y “sesión individual” una representación de las evaluaciones usando un “registro de juicio” donde se almacena la siguiente información:

- Identificador del estudiante
- Tópico/s evaluado/s
- Area (formativa o acumulativa)
- Tipo de evaluación (examen, examen preliminar, etc.)
- Significancia de la asociación entre las redes semánticas del estudiante y del experto
- Nivel de confianza del maestro en el juicio <opcional>
- Nota (juicio) del maestro <opcional>
- Comentarios (texto) del maestro respecto a su juicio <opcional>
- Inferencia final de evaluación dada por el sistema siguiendo un modelo de utilidad multi-atributo.

En el nuevo modelo del estudiante (ver figura 5) se aprecian sus componentes tradicionales y las figuras en gris presentan los componente del modelo que hemos tratado en este trabajo.

El perfil de los estudiantes, el organizador de la sesión individual y el repositorio de pruebas, problemas y ejercicios sólo fueron tratados someramente. Los dos componentes ampliados presentan los bordes resaltados: el “razonador de evaluación” donde está embebido el algoritmo DistSem2 y un analizador de resultados



donde se encuentra el “registro del juicio” para cada estudiante

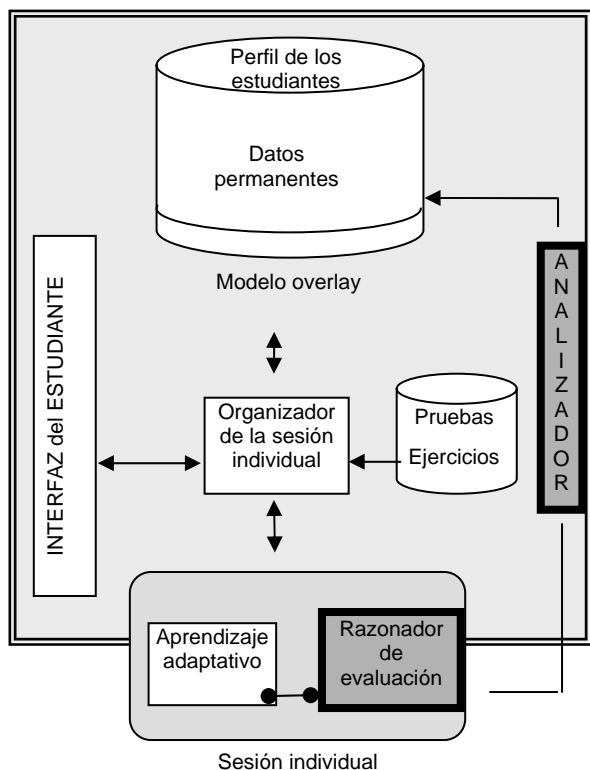


Figura 5: nuevo modelo del estudiante

Las reglas de producción devuelven los resultados finales de la evaluación al modelo overlay individual donde cada nodo correspondiente al tópico es marcado como parcialmente, totalmente o no aprendido.

En la figura 6 se aprecia la ventana de una de las preguntas que se le presenta a cada estudiante durante la toma de la prueba. De este modo puede navegar por los 40 pares de conceptos y cuando la prueba se da por terminada, es grabada para su posterior análisis.

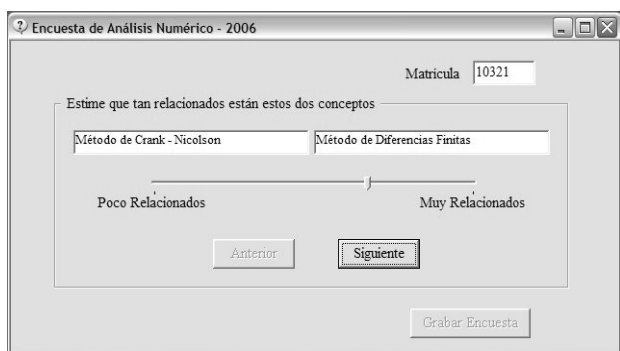


Figura 6: toma de la prueba DistSem

## Conclusiones

El nuevo diseño de nuestro Sistema Tutorial Inteligente involucra modelos del instructor y estudiante mejorados con el método DistSem. El objetivo perseguido fue evaluar la comprensión y las habilidades del pensamiento de alto nivel. Principalmente, se buscó mejorar la evaluación de los conocimientos con una técnica basada en la distancia semántica entre conceptos. Esta técnica ha demostrado tener éxito en las pruebas de campo efectuadas fuera del sistema tutorial.

Particularmente, el nuevo modelo del instructor, además de registrar el comportamiento del docente, le brinda un medio para identificar las fortalezas y debilidades del estudiante. La concepción de las pruebas basadas en DistSem respeta los tres roles educativos distintos de la evaluación.

Además, la inclusión del modelo ampliado del estudiante en el Sistema Tutorial Inteligente tiene como objetivo capturar información de mayor calidad de la actividad del estudiante a fin de alcanzar un modelo más exacto. De este modo, el sistema puede tomar mejores decisiones sobre el próximo paso tutorial.

Siendo DistSem el método que fundamenta nuestro estudio, tenemos como próximo paso en nuestra investigación el análisis cualitativo de las redes semánticas

## Referencias

- [1] R. Felder, L. Silverman, Learning and Teaching Styles in Engineering Education, Engineering Education, 78, pp 674-681,1988.
- [2] J. Turns, C. Atman, R. Adams, T. Barker, Research on Engineering Student Knowing: Trends and Opportunities, Journal of Engineering Education, 94(1), 27-40, 2005.
- [3] R. Felder, R. Brent, Understanding Student Differences, Journal of Engineering Education, 94(1), 57-72, 2005.
- [4] C. Huapaya, G. Arona, Sistemas de autoría especializado en STIs matemáticos, Revista Brasileira de Informática na Educação, 10(2), 37-49, 2002.
- [5] E. Wenger, Artificial Intelligence and Tutoring Systems. Computational and Cognitive Approaches to the Communication of Knowledge, Morgan Kauffmann, 1987.
- [6] D. Jonassen, T. Reeves, N. Hong, D. Harvey, K. Peters, Concept Mapping as Cognitive Learning and Assessment Tools, Journal of Interactive Learning Research, 8, 3/4, 289-308, 1997.
- [7] R. Sison, M. Shimura, Student Modeling and Machine Learning, International Journal of Artificial Intelligence in Education, 9, 128-158, 1998.

[8] R. Nkambou, M. Frasson, C. Frasson, Generating courses in an intelligent tutoring system, Proceedings of the 9<sup>th</sup> International Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert systems, 261-267, 1996.

[9] B. du Boulay, A. Luckin. Modelling human teaching tactics and strategies for tutoring systems, International Journal of Artificial Intelligence in Education, 12, 235-256, 2001.

[10] Virvou, M., M. Moundridou, Adding an instructor modelling component to the architecture of ITS authoring tools, International Journal of Artificial Intelligence in Education, 12, 185-21, 2001.

[11] National Council of Teachers of Mathematics, Principles and Standards for School Mathematics, 2001.

[12] R. Kellough, N. Kellough, Secondary School Teaching: A Guide To Methods And Resources; Planning For Competence, Prentice Hill, 1999.

[13] B. Bloom, Taxonomy of Educational Objectives, The Classification of Educational Goals: Handbook I, Cognitive Domain, Longmans, 1956.

[14] A. Collins & E. Loftus, A spreading-activation theory of semantic processing, Psychological Review, 82, 407-428, 1975.

[15] J. Vivas, Método Distsem: procedimiento para la evaluación de distancias semánticas. Revista Perspectivas en Psicología, 1, 56-61, 2004.

[16] J. Vivas Análisis de redes semánticas aplicado a contenidos académicos. Métodos e instrumentos. En Kronmüller, E. y Cornejo, C. (Comp.). Ciencias de la Mente: Aproximaciones desde Latinoamérica. Santiago, Chile: JCSáez Editor. (En prensa).

[17] J. Vivas, A. Comesaña, L. Vivas. Evaluación de las redes semánticas de conceptos académicos en estudiantes universitarios. PsicoUSF, 11,2. 2006.

[18] A. Comesaña, M. Gonzalez, J. Vivas. El aprendizaje de contenidos académicos y el conocimiento previo. Un abordaje desde las redes semánticas de los alumnos. Las Ciencias del Comportamiento en los Albores del Siglo XXI. XRAAC. Mar del Plata. UNMDP. 386-393. 2005.

[19] DiVesta, F. y Rieber, L., Characteristics of cognitive engineering. Educational Communication and Technology Journal, 35,4,213-230, 1987.

[20] Shuell, T., Cognitive conceptions of learning. Review of Educational Research, 56,4, 411-436, 1987.

[21] Resnick, L. B., Shared cognition en Perspectives on socially shared cognition. Washington:A.P.A., 1991.

[22] Garrison, D., A cognitive Constructivist view of Distance Education. Distance Education, 14,2,199-211, 1993.

[23] Winn, W., Some implications of cognitive theory for instructional design. Instructional Science, 19, 1,53-69, 1990.

[24] D. Krackhardt, QAP partialling as a test of spuriousness, Social Networks 9, 171-186, 1987.

*Dirección de Contacto del Autor/es:*

**Constanza Huapaya**

Juan B. Justo 4302

7600- Mar del Plata

Argentina

e-mail: huapaya@fi.mdp.edu.ar

**Francisco Lizarralde**

Juan B. Justo 4302

7600-Mar del Plata

Argentina

e-mail: flizarra@fi.mdp.edu.ar

**Jorge Vivas**

Funes 3250

7600-Mar del Plata

Argentina

e-mail: jvivas@mdp.edu.ar

**Graciela Arona**

Juan B. Justo 4302

7600- Mar del Plata

Argentina

e-mail: grarona@fi.mdp.edu.ar

---

**Constanza Huapaya.** Profesora titular y directora del GI en Inteligencia Artificial de la Facultad de Ingeniería de la UNMDP. Cursa la Maestría en Tecnología Informática Aplicada a Educación de la Facultad de Informática de la UNLP.

---



---

**Francisco Lizarralde.** Ingeniero Electrónico. Docente e investigador en la Facultad de Ingeniería de la UNMDP. Actualmente cursa la Maestría en Tecnología Informática Aplicada a Educación de la Facultad de Informática de la UNLP.

---



---

**Jorge Vivas (Mg).** Director del GI en Psicología Cognitiva y Educacional de la Facultad de Psicología de la UNMDP. Profesor titular de Psicología Cognitiva y Teorías del Aprendizaje en la UNMDP.

---



---

**Graciela Arona.** Docente e investigadora en la Facultad de Ingeniería de la UNMDP. Profesora titular en la Facultad de Ingeniería de la Universidad Fasta.

---