

# Sistema inteligente para evaluación de programas tutores

Luz Daiçy Alvarado<sup>1</sup>  
Nelson Becerra Correa<sup>2</sup>

## RESUMEN

El mecanismo evaluador que se presenta a continuación forma parte de un proyecto que busca el desarrollo de un Sistema de enseñanza asistida por ordenador (EAO) basado en tutores inteligentes.

Se pretende que este mecanismo, basado en herramientas de «aprendizaje automático», tenga la capacidad para capturar una «imagen» que indique el grado de comprensión de la materia por parte del alumno e identifique los aspectos en los que el estudiante tiene dificultades. Con esa información se retroalimentará a un sistema tutor de manera que este último tome una decisión acerca de la estrategia educativa por seguir.

### **PALABRAS CLAVE:**

EVALUACIÓN, TUTORES INTELIGENTES, ENSEÑANZA ASISTIDA POR ORDENADOR, APRENDIZAJE AUTOMÁTICO.

## ABSTRACT

The evaluating system we show in this paper is part of a project that ties to develop a computer assisted teaching system based on intelligent tutoring.

We pretend this mechanism based in automatic learning tools, have the capacity for capturing a image that reflect the subject comprehension degree by the student and identify the aspects where the student have problems. Whith this information the tutor will be feedbacking so he can to take a decision about the educational strategy to follow.

### **KEY WORDS:**

EVALUATION, INTELLIGENT TUTORING, COMPUTER ASSISTED TEACHING SYSTEM BASED, AUTOMATIC LEARNING.

## EL PROBLEMA

Como es bien sabido, la evaluación a nivel educativo tiene influencia decisiva y directa sobre la calidad de la educación. Sin embargo, tal y como se lleva a cabo actualmente, presenta algunos inconvenientes relacionados con el proceso mismo de evaluar, uno de ellos relacionado con la calificación. García Ramos (1989) afirma que “la validez de estos juicios depende en gran medida de la formación y experiencia del que juzga y por ello se presta a errores de origen diverso (“efecto halo”, tendencia a la severidad, a la benignidad o a evitar los extremos)”.

1 Profesora, Planta Universidad Distrital, adscrita a la Facultad de Ingeniería.

2 Profesor, Planta Universidad Distrital, adscrito a la Facultad Tecnológica.

Santos Guerra (1993) plantea también algunos cuestionamientos al respecto:

No se sabe si se comparan las capacidades de los sujetos, los esfuerzos realizados, los conocimientos adquiridos o la suerte que han gozado ... Partir de situaciones desiguales y pretender comparar los resultados utilizando los mismos raseros es una fórmula radicalmente injusta de ejercitar una aparente justicia ... la evaluación ha sido un instrumento de control, de amenaza, e incluso de venganza, respecto a algunos alumnos que se han permitido ejercitar el derecho a la crítica, a la discrepancia o a la indisciplina.

Según Cook (1983), “la asignación de números de una manera mecánica, como es común en los procedimientos cuantitativos, no garantiza la objetividad”.

Otro de los problemas tiene que ver con la continuidad y retroalimentación que tenga el estudiante respecto a los contenidos estudiados. Santos Guerra (1993) opina que el efecto retroalimentador se suele perder en las evaluaciones educativas. Y así, la evaluación que suponen las memorias finales de los centros, en nada iluminan el nuevo proyecto en el curso siguiente. El mundo educativo está lleno de situaciones en las que no se aprovecha esta riqueza potencial que lleva en su interior el proceso evaluador. Los profesores/alumnos repiten los mismos errores casi con obstinación; los cursos que finalizan una especialidad, en nada benefician con el análisis de su historia a los cursos siguientes.

Por otra parte, en los sistemas de enseñanza automatizados, el proceso de evaluación ha consistido en someter al alumno a una serie de preguntas sobre algún tema para obtener así una calificación. En los sistemas tutores dicha calificación es comunicada al tutor, y éste decide entre repetir la lección o avanzar al siguiente tema. Esta filosofía obliga a que el test se ajuste específica y exclusivamente al tema en estudio, sin posibilidades de capturar información acerca del conocimiento relacional de la materia. En estos casos, la eficacia de la evaluación está basada única y exclusivamente en el prerrequisito, dejando de lado la posibilidad de que otros temas puedan aportar información sobre las dificultades que el alumno tiene en el tema actual.

## PLANTEAMIENTOS QUE PUEDEN APORTAR SOLUCIONES

A la luz de los problemas relacionados con la evaluación educativa, planteados anteriormente, se propone un mecanismo evaluador que involucre herramientas de inteligencia artificial, con la idea de que proporcione alternativas de mejora en aspectos tales como imparcialidad y objetividad de la evaluación, globalidad, interacción, continuidad y retroalimentación acerca de los contenidos presentados al estudiante, pues el objetivo del mecanismo evaluador propuesto es identificar las áreas de conocimiento sobre el tema en cuestión en las cuales el alumno tiende a fallar, analizar dichas falencias y proponer estrategias al módulo tutor o informar al estudiante acerca de sus puntos débiles.

En este caso no se asigna valor alguno de comparación o calificación. A medida que el estudiante avanza el proceso evaluador se hace más y más personalizado porque dependerá de las áreas en las cuales el estudiante falla; de este modo se reducen tanto la comparación entre los estudiantes como las posibilidades de parcialidad y subjetividad por parte del evaluador.

Por otro lado, la incorporación de la hipermedia<sup>1</sup> a las redes de comunicaciones ha enriquecido el entorno de trabajo del usuario; sin embargo aún se tienen problemas debido a lo estático de la presentación y a la dificultad para acceder, rápida y directamente, a un tema en concreto. Este problema se evidencia con mayor énfasis en los sistemas de enseñanza cuando se desea que un estudiante, conectado desde un sitio remoto, tenga acceso a los principios que rigen una materia determinada.

Algunos autores plantean la creación de patrones para el desarrollo de aplicaciones hipermedia, aprovechando las posibilidades que ofrece la programación orientada a objetos<sup>2</sup> y ofreciendo la posibilidad de reutilizar el conocimiento en entornos diferentes. Es aquí donde se evidencia la necesidad de crear sistemas independientes de la materia por enseñar, de manera que se potencien los *shells* capaces de ser alimentados con un conocimiento

1 Utilización simultánea de diversos medios tales como video, sonido, animación, texto e hipertexto.

2 Según Meyer (1998), “la orientación a objetos es una técnica de ingeniería de software, la cual estudia los métodos y herramientas que se pueden utilizar para producir software práctico y de calidad”.

específico. Estos aspectos han sido tratados en artículos de Rossi, Schwabe y Garrido (1996), Gamma, Helm, Johnson y Vlissides (1995), Garzotto, Mainetti y Paolini (1995), Izakowitz, Stohr y Balasubramaniam (1995), Schwabe, Rossi y Barbosa (1996), entre otros.

## 1. Utilización de los mecanismos de aprendizaje automático

El término “mecanismo de aprendizaje” se ha utilizado para definir procedimientos que a partir de cierta información “aprendan” y generen elementos que pueden considerarse “solucionadores de problemas”. A este respecto Bahamonde (1997) define: “El problema del aprendizaje es sintetizar procedimientos computables para realizar tareas de las cuales sólo tenemos algunas descripciones particulares sobre el modo en que nos gustaría que trabajase el sistema”.

Existen varias clasificaciones de los algoritmos de aprendizaje automático. Para el trabajo que aquí se plantea, se utilizarán aquellos que realizan aprendizaje a partir de ejemplos. Estos algoritmos reciben un conjunto de ejemplos, cada uno de los cuales está compuesto por un grupo de atributos y una clasificación a la cual pertenece ese conjunto de atributos en particular. Generan como resultado un clasificador que puede recibir información similar y tiene una alta probabilidad de realizar la clasificación correcta.

Hasta ahora los mecanismos de aprendizaje clasificadores a partir de ejemplos han sido empleados en diversas áreas, tales como el diagnóstico del cáncer de mama, clasificación de robots, el precio de la vivienda en Boston, clasificación de lirios, pronóstico del tiempo, etc. Sin embargo, no se ha evidenciado una aplicación de estos mecanismos en el área educativa; se pensó entonces en la posibilidad de tomar las respuestas a un test, dadas por un estudiante, para presentarlas como ejemplos con el objeto de que el acierto o falla en dichas respuestas represente la clasificación de estos ejemplos. No obstante, la simple distinción entre acierto o falla no proporciona mucha información acerca del conocimiento que el alumno tiene sobre el tema, y por ello se hace necesario que cada una de las preguntas sea descrita como un conjunto de características relacionadas con el tema que se está tratando.

El mecanismo de aprendizaje clasifica los ejemplos en dos categorías: las preguntas en las que acertó y en las que falló, para luego producir unas reglas de clasifi-

cación que sean conclusión directa del conjunto de ejemplos. Dichas reglas agruparán los conceptos manejados por el estudiante (reglas de acierto) y aquellos en los que tenga dificultades (reglas de fallo).

El análisis de las reglas obtenidas dará luces acerca de los temas que el alumno maneja, aquellos en donde tiene dificultades o simplemente los temas puntuales en los cuales falló. Sin embargo, dicho análisis no resulta simple pues las reglas obtenidas por lo general combinan atributos o no tienen un 100% de aciertos. Por tanto, se hace necesario definir estrategias de evaluación para tales reglas, teniendo en cuenta el número de ejemplos cubiertos, el porcentaje de acierto y las probabilidades relacionadas con la clase a la que pertenece, dependiendo del valor tomado por cada atributo que participa como premisa de la regla. De acuerdo con esto, el evaluador analizará y generará un diagnóstico.

## 2. El evaluador

El papel de la evaluación no se reduce a obtener una calificación de las respuestas dadas en los tests o ejercicios presentados al estudiante con el objeto de permitirle o no continuar con el siguiente tema. Saber que un alumno acertó cierto número de preguntas en un test no aporta información acerca de los conceptos que tenga afianzados o aquellos que desconozca. Desde esta nueva perspectiva, los test se convierten en una nueva herramienta, con la cual se pretende establecer una representación cualitativa del grado de conocimiento que tiene el alumno sobre la materia en ese momento.

Así, cada vez que el alumno es sometido a un test, se extrae un paquete de características que describen parcialmente las áreas conocidas o aquellas en las cuales falla repetidamente, relacionadas con la materia evaluada. En la figura 1 se presenta el esquema general del módulo evaluador aquí planteado.

Como se mencionó antes, el proceso evaluador se inicia en el momento en que el estudiante es sometido a un test. La revisión de cada una de las respuestas, así como las características que se quieren evaluar con cada pregunta, se almacenan en una tabla estructurada de manera interpretable por el sistema de aprendizaje. Una vez realizado el proceso de aprendizaje automático, se obtienen las reglas que identifican los bloques de conocimiento o de fallas reiteradas por parte del alumno; el análisis de las reglas

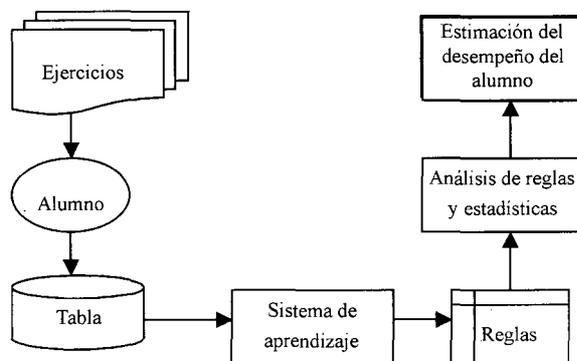


FIGURA 1. Arquitectura del módulo evaluador.

|        | Atrib. 1 | Atrib. 2 | Atrib. 3 | — | Atrib. 11 | Clase (acierto o fallo) |
|--------|----------|----------|----------|---|-----------|-------------------------|
| Preg 1 |          |          |          |   |           | Acierto                 |
| Preg 2 |          |          |          |   |           | Fallo                   |
| —      |          |          |          |   |           |                         |
|        |          |          |          |   |           |                         |

FIGURA 2. Estructura de la información obtenida de los tests.

producidas y otros datos estadísticos que proporciona el sistema de aprendizaje produce un acercamiento al nivel de desempeño del estudiante así como las áreas donde falla o acierta reiteradamente.

### 2.1. DESCRIPCIÓN GENERAL DEL PROCESO DE EVALUACIÓN

Se planteó entonces la posibilidad de aplicar mecanismos de aprendizaje automático a partir de ejemplos en el análisis de las respuestas dadas a un test, con el fin de comprobar que las reglas producidas tienen la posibilidad de describir el grado de conocimientos del alumno a través de las reglas concebidas como entidades conceptuales. Dicho test debía tener plenamente identificadas las características de los elementos que componían cada pregunta con el fin de saber exactamente qué parte del tema intentaba evaluar.

Una vez aplicado el test al alumno o alumnos seleccionados, sus respuestas –junto con las características

antes definidas– se organizan en archivos legibles por el mecanismo de aprendizaje elegido. La información obtenida de esta forma tiene una estructura similar a la presentada en la figura 2.

En este punto se aplica el mecanismo de aprendizaje que proporcione mayor información. Al obtener las reglas, se lleva a cabo una selección de las que mayor información aportan, y se analizan e interpretan para obtener de esta manera el nivel de desempeño del usuario.

### 2.2. INTERPRETACIÓN Y ANÁLISIS DE RESULTADOS

Como resultado de aplicar el mecanismo de aprendizaje, para cada uno de los alumnos se genera un archivo que contiene, además de las reglas, resultados adicionales relacionados con el nivel de errores, probabilidad de acierto de las reglas definidas, así como el número de ejemplos (o preguntas) en las cuales acertó o falló. En el análisis de los resultados se tienen en cuenta todos estos factores.

Los resultados relacionados con el nivel de errores, probabilidad de acierto y el número de ejemplos que clasifica cada regla, proporciona información acerca del grado de agrupación de las respuestas; por ejemplo, si se tiene una gran probabilidad de acierto en la clasificación con un número no muy elevado de reglas, se pueden identificar bloques de conceptos donde el estudiante acertó o falló reiteradamente, lo cual significaría que tiene afianzados ciertos conceptos o que comete errores sistemáticos. Igualmente puede determinarse el rango dentro del cual se considera que el alumno tuvo un rendimiento similar al que presentaría si hubiera respondido al azar.

La información relevante para el análisis, tomada de las reglas que el sistema de aprendizaje proporciona, corresponde al nivel de impureza de la regla (grado que indica la efectividad de la regla), el número de ejemplos a los cuales se aplica, los atributos involucrados, el valor o rango de valores, el peso (proporción que dentro de los aciertos o fallos tiene un valor de un atributo) relativo a un par atributo-valor en las clases de acierto y fallo, y la proporción o probabilidad de aciertos y fallos sobre el dominio restringido a un par atributo-valor.

De las reglas se seleccionan las que poseen un menor nivel de impureza; para cada una de ellas se analizan aisladamente uno a uno los atributos que posee, y la conjunción de atributos del antecedente de la regla.

De cada atributo se considera relevante la información del peso, la probabilidad y el valor del mismo. Con esta información se puede deducir qué grado de conocimiento posee el estudiante acerca de los conceptos relacionados con el atributo analizado cuando toma el valor definido en la regla.

Aunque dentro de una regla, la observación aislada de los atributos proporciona bastante información, es importante tener en cuenta que los valores correspondientes al nivel de impureza de la regla y número de ejemplos cubiertos por la misma, hace referencia a los ejemplos que satisfacen conjuntamente las condiciones impuestas por los antecedentes. Por tanto se deben analizar aquellas reglas relevantes que se correspondan con la probabilidad de acierto del alumno, con el fin de identificar los conceptos que mejor conoce y aquellos donde falla significativamente.

1 Si se tienen  $x$  opciones de respuesta, la probabilidad de acierto si se responde al azar será:  $1/x$ .

Dado que el nivel de impureza indica la calidad de las reglas, éste será el parámetro que se observe para elegir las mejores; son precisamente estas reglas las que indicarán el lugar donde se agrupan los aciertos o las fallas, es decir, identifican características muy relevantes de los conceptos que se manejan propiamente o de los adquiridos de manera errónea.

Teniendo la proporción de preguntas que el alumno acierta ( $P_a$ ) y falla ( $P_f$ ), la cantidad de preguntas del test ( $n$ ) y la probabilidad de acierto<sup>1</sup> en caso de responder al azar ( $P$ ) (que depende de las opciones de respuesta para cada pregunta), es posible obtener un intervalo de confianza  $[a, b]$  o rango de proporciones admisibles como consecuencia de respuestas al azar.

Una vez conocido el intervalo de confianza mencionado, es posible ir aún más allá en el análisis. Si se plantea que el conocimiento adquirido está dentro del rango  $[0, 1]$ , donde 1 indica el conocimiento total del tema, y en él se ubica el intervalo de confianza hallado antes, se considera que la ganancia en conocimientos estaría ubicada entre el límite superior del intervalo que nos señala las respuestas al azar ( $b$ ) y el 1 que nos indica el total de conocimientos.

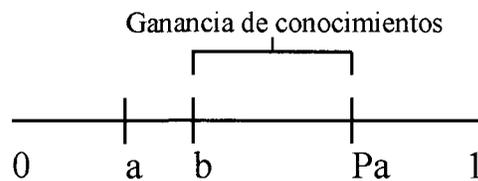


Figura 3. Representación de la ganancia de conocimientos

En la figura 3 puede observarse que si la proporción de aciertos es mayor que el extremo superior del intervalo de confianza  $[a, b]$ , el alumno tendrá un nivel de aprovechamiento positivo, cuya proporción, dentro del conocimiento que se puede adquirir, corresponde a la distancia desde el extremo superior del intervalo hasta el porcentaje de aciertos, dividida entre la distancia que hay desde el extremo superior del intervalo hasta 1:  $(P_a - b)/(1 - b)$ .

Una vez conocidas las proporciones de acierto y fallo del estudiante, se halla el intervalo de confianza de la probabilidad esperada de acierto o de fallo. Las reglas con una probabilidad mayor que el extremo superior del intervalo de confianza de acierto hallado identifican claramente los bloques conceptuales que se dominan o aquellos que provocan errores.

Para mayor claridad se presenta como ejemplo la regla siguiente, con el fin de mostrar el análisis que sobre ella se hace y las estrategias que, con base en la información analizada y recibida de la regla, se siguen.

Regla 1 (A \*-22,0147460308313\* (1) 8 8)

(Atrib1 | W:0,24 P:0,75)

(Atrib2 | W:0,6 P:0,61)

(Atrib3 | W:0,73 P:0,56)

En este caso se observa una regla que clasifica los aciertos (A), tiene un nivel de impureza muy bajo (-22.01474), una probabilidad de acierto de 1, ya que cubre ocho ejemplos y los acierta todos e incluye tres atributos. A su vez, de cada atributo se conoce el valor que toma, el peso (W) y la probabilidad (P).

En el caso del peso, para esta regla en el primer atributo ( $W = 0,24$ ), representa la probabilidad de que dicho atributo tome el valor 1 sabiendo que acertó. Por su parte, P corresponde a la probabilidad de acierto (0,75) sabiendo que el mismo atributo toma el valor 1.

De acuerdo con el propósito de este trabajo, los valores de P y W, para cada parámetro, tienen gran importancia ya que su análisis contribuirá a identificar las características conceptuales del alumno para, de esta manera, definir el nivel de conocimientos que posee.

De la regla anterior se tiene:

|    |  |      |
|----|--|------|
| P1 | $P(\text{acierto} / \text{Atrib 1})$     | 0.75 |
| P2 | $P(\text{acierto} / \text{Atrib 2})$     | 0.61 |
| P3 | $P(\text{acierto} / \text{Atrib 3})$     | 0.6  |
| W1 | $P(\text{Atrib 1} = 1 / \text{acierto})$ | 0.24 |
| W2 | $P(\text{Atrib 2} = 1 / \text{acierto})$ | 0.6  |
| W3 | $P(\text{Atrib 3} = 1 / \text{acierto})$ | 0.73 |

En esta regla se tienen en cuenta las preguntas donde el estudiante acertó, y adicionalmente los atributos 1, 2 y 3 toman el valor de 1.

El primer paso consiste en analizar para cada atributo el valor de W y P; por ejemplo, para la regla anterior y respecto al atributo 1, se concluiría que la probabilidad de que las preguntas con el valor 1 en este atributo estén dentro de los aciertos es de 0.75 (P1), mientras que la probabilidad de que se encuentren dentro de los fallos es solamente de 0.25. Gráficamente la figura 4, muestra la probabilidad P del atributo 1 dentro de la regla 1.

En este caso, el óvalo representa la totalidad (T) de los ejemplos con los que se aprendió, los cuales, dependiendo de las respuestas del alumno, tendrán una mayor

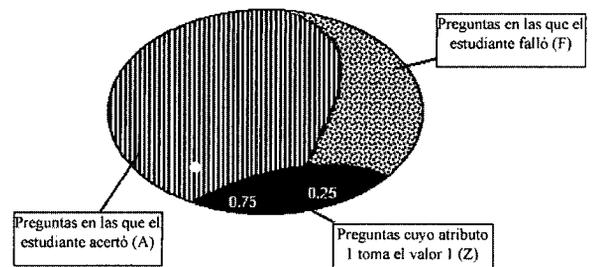


FIGURA 4. Representación de la probabilidad P del atributo1 dentro de la regla1.

o menor proporción de aciertos (A) o fallos (F). El área rayada identifica la proporción de ejemplos que acertó; el área punteada identifica la proporción de ejemplos que falló. La unión de estas dos áreas forma la totalidad de los ejemplos:  $A \cup F = T$ .

El área de color oscuro (Z) en la parte inferior del óvalo representa la probabilidad o peso que dentro del total de ejemplos, para el caso de la regla que se ha venido analizando, tienen aquellos cuyo atributo 1 toma el valor de 1. Como se observa, esta región tiene a su vez una parte dentro del área de aciertos y otra dentro del área de fallos, las cuales se podrían ver como  $A \cap Z$  y  $F \cap Z$ .

Finalmente, si se desea conocer la proporción de Z dentro de los aciertos sería  $(A \cap Z) / Z$ . Este valor, en el ejemplo que se está analizando, corresponde a  $0.75 = P1$ . Igual sucede con la proporción de Z dentro de F:  $(F \cap Z) / Z = 0.25$ . Esto daría indicios de que el alumno conoce el 75% de las respuestas a las preguntas cuyo atributo 1 toma el valor 1, lo cual en principio parece bastante significativo.

Por otra parte, al observar W1, es posible deducir que dentro de los aciertos la proporción de las preguntas cuyo atributo 1 toma el valor 1 es 0.24, es decir que dentro de los aciertos –si se observa únicamente el atributo 1–, solamente un 24% corresponde a aquellos ejemplos que toman el valor 1 en dicho atributo. Para el ejemplo que se está analizando, la interpretación gráfica podría expresarse:

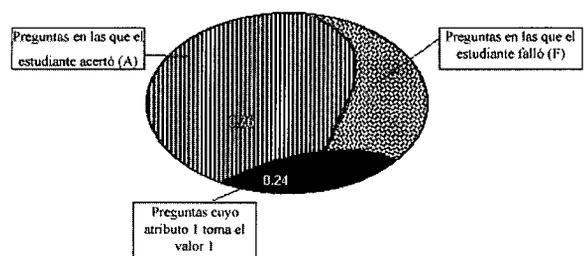


FIGURA 5. Representación del valor de W para el atributo1 en la regla1.

El óvalo representa todos los ejemplos con los cuales se aprendió. Al igual que en el caso anterior, esta región se divide en aciertos (A) y fallos (F). En la parte inferior, la región oscura representa los ejemplos cuyo atributo  $\ell$  toma el valor 1, la cual a su vez tiene una parte dentro del área de aciertos (ZA) y otra dentro del área de fallos (ZF). Como se observa, el área oscura rayada (ZA) muestra la probabilidad de que dentro de los ejemplos que pertenecen a la clase acierto su atributo  $\ell$  tome el valor 1. Esta proporción se puede expresar:  $(ZA \cap A)/A$  que en el ejemplo analizado corresponde a  $0.24 = W_1$ , lo cual implica que el resto de los ejemplos de acierto corresponden a  $(ZA \cap A)/A$ , cuyo valor en este caso sería 0.76.

De este análisis se puede deducir que aunque el alumno conoce bastante las respuestas a las preguntas cuyo atributo  $\ell$  toma el valor 1, hay pocas preguntas de este tipo.

Como puede verse, el análisis del significado correspondiente a los valores de P y W para cada uno de los atributos en una regla permite identificar aquellos atributos que tienen mayor incidencia en los aciertos y los que inciden especialmente en los fallos. En este caso se puede decir que P y W contribuyen a identificar conceptos que el estudiante conoce muy bien y aquellos en los cuales falla sistemáticamente; es decir, donde existe un concepto erróneo. Sin embargo, el análisis exclusivo de los atributos de las reglas puede conducir a conclusiones incompletas si no se tiene en cuenta el resto de la información que proporciona la regla que contiene dichos atributos junto con los datos adicionales relacionados con el número de respuestas acertadas y falladas, el número de preguntas y la probabilidad que se tiene de acertar al azar.

Adicionalmente, para llevar a cabo un análisis más detallado, se deberá tomar en cuenta datos como el número de preguntas, la probabilidad de acierto en ellas, así como la proporción de aciertos y fallos que presente el estudiante. Para complementar el ejemplo anterior, se asumirán 120 preguntas cuya probabilidad de acierto sea de 0.34, de las cuales se acertaron 62 (52%) y se fallaron 58 (48%).

Con la información anterior se puede hallar el intervalo de confianza, al 95%, dentro del cual se puede considerar que el alumno respondió al azar, teniendo como base una distribución de probabilidad normal:

$$0.34 \pm 1.96 \sqrt{\frac{0.34(1 - 0.34)}{120}} = (0.255, 0.425)$$

La proporción de aciertos (0.52 ó 52%) es mayor que el límite superior del intervalo (0.425), y por tanto se puede suponer que el alumno tiene un cierto conocimiento adquirido sobre el tema. Es posible evaluar el rendimiento (positivo en este caso) tomando como base el intervalo  $[0,1]$  donde un valor entre 0.255 y 0.425 corresponde al nivel de azar y 1 al acierto de todas las preguntas. En este intervalo se ubica el límite superior del intervalo de azar (dado que el rendimiento es positivo); de esta manera se halla el grado de rendimiento (proporción de conocimiento superior al azar) usando la siguiente expresión:

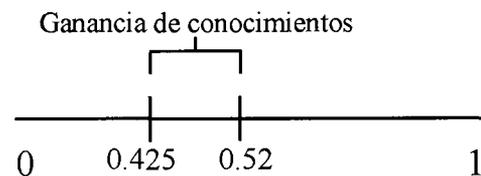


Figura 6. Ganancia de conocimientos en el ejemplo.

$$\frac{0.52 - 0.425}{1 - 0.425} = 0.16$$

En este caso, el alumno tiene un nivel de aprovechamiento de +16%. De igual forma es posible hallar el intervalo de confianza, al 95% (considerando una distribución de probabilidad normal), que representa el rango de probabilidad de acierto del alumno:

$$0.52 \pm 1.96 \sqrt{\frac{0.52 * 0.48}{120}} = (0.431, 0.609)$$

Se puede decir, entonces, que las áreas de conocimiento que presenten una probabilidad de acierto superior a este intervalo, son significativamente dominadas por el alumno.

De manera análoga, es posible realizar el mismo análisis para la proporción de fallos, con el fin de descubrir los bloques conceptuales cuyo dominio es significativamente inferior al esperado. En este caso, la probabilidad de fallo es de 0.48, con lo cual el intervalo de confianza al 95%, que representa el rango de probabilidad donde se espera que el alumno falle, es:

$$0.48 \pm 1.96 \sqrt{\frac{0.48 * 0.52}{120}} = (0.391, 0.569)$$

Por consiguiente, es posible concluir que cualquier grupo conceptual cuya probabilidad de fallo sea superior a 0.569 está adquirido con errores conceptuales que provocan una proporción de errores significativamente mayor de la esperada.

Después de ser analizada la información obtenida de las reglas y los resultados presentados por el estudiante al responder el test, así como los demás estadísticos relevantes a este análisis, es posible definir de manera genérica un grupo de estrategias que permitirán tomar decisiones relacionadas con las acciones por seguir en el proceso de enseñanza al estudiante.

### 2.3. ESTRATEGIAS

Una vez definida la forma como se analizarán los atributos (factores) que componen las reglas producidas por el sistema de aprendizaje, la estrategia por seguir se puede expresar de acuerdo con el algoritmo genérico siguiente:

*Si la proporción de respuestas corresponde al intervalo de confianza para el azar*

*Entonces: proponer el inicio del proceso de enseñanza*

*Si No*

*Si hay reglas de fallo con bajo nivel de impureza que se apliquen a varios ejemplos*

*Entonces: proponer el estudio de lecciones relativas a sus atributos*

*Sí No*

*Si hay reglas de acierto representativas cuyos atributos posean un valor W alto y P no sea muy alto*

*Entonces: extraer factores de acierto y proponer el enfocar el estudio en los aspectos restantes*

*Sí No*

*Si el número de fallos es pequeño*

*Entonces: mostrar soluciones a los fallos*

*Fin si.*

*Fin si.*

*Fin si.*

El análisis de las reglas producidas por el sistema de aprendizaje están basadas en la interpretación planteada en el ítem anterior, y dado que las reglas son mejores en tanto más ejemplos expliquen, la estrategia prioritaria es revisar primeramente las reglas que más ejemplos contemplen y que menor nivel de impureza tengan.

Inicialmente se centra el análisis sobre las reglas que clasifican los fallos. Como se mencionó antes, tienen prioridad aquellas que clasifican un mayor número

de ejemplos cuyo nivel de impureza sea menor, y realizando un análisis como el presentado anteriormente, se identifican grupos conceptuales donde el estudiante tiene mayor dificultad y se propone al módulo tutor el repaso de las áreas relacionadas. Igualmente se tendrán en cuenta los grupos conceptuales que el estudiante conoce para evitar su continua repetición, e identificar también aquellas áreas donde se falla pero no existen errores conceptuales; es decir, donde están representados los conceptos que el estudiante desconoce. En caso que de las reglas de fallo sean pocas y a su vez clasifiquen pocos ejemplos (es decir, el usuario cometió pocos errores que no están interrelacionados) simplemente se propondrá mostrarle al usuario directamente cuáles fueron los aspectos puntuales en los cuales falló.

Existe la posibilidad de que solamente se produzca la regla por defecto para clasificar los fallos. En ese caso se analizan a fondo las reglas de acierto, identificando aquellos atributos que intervienen con mayor preponderancia en los aciertos, para buscar aquellos que no aparecen o tienen un valor pequeño para P, y enfocar entonces la atención del estudiante sobre estos últimos.

## IV. CONCLUSIONES

Para la solución del problema, se utilizaron algoritmos que realizan aprendizaje a partir de ejemplo. Estos algoritmos reciben un conjunto de ejemplos, cada uno de los cuales está compuesto por un grupo de atributos y una clasificación a la cual pertenece ese conjunto de atributos en particular. Generan como resultado un clasificador que puede recibir información similar y tiene una alta probabilidad de realizar la clasificación correcta.

En la creación del evaluador, el estudiante se somete a un test; los resultados, junto con las características que se requieren evaluar con cada pregunta, se almacenan en una tabla estructurada de manera interpretable por el sistema de aprendizaje. Una vez realizado el proceso de aprendizaje automático, se obtienen las reglas que identifican los bloques de conocimiento o de fallas reiteradas por parte del alumno; el análisis de las reglas producidas y otros datos estadísticos que proporciona el sistema de aprendi-

zaje produce un acercamiento al nivel de desempeño del estudiante, así como las áreas donde falla o acierta reiteradamente.

La eficacia de los resultados obtenidos con los mecanismos de aprendizaje depende especialmente de los tests presentados al estudiante, las características o atributos que definen lo que se quiere evaluar con cada una de las preguntas y la manera de representar la información que se obtiene de dichos tests; es aquí

donde el papel del experto humano desempeña un papel preponderante.

El desarrollo de mecanismos de evaluación como el que se plantea aquí contribuyen a disminuir el impacto que tienen, en el sistema educativo tradicional, los problemas relacionados con la evaluación, tales como los relativos a la continuidad en los contenidos, la retroalimentación, la asignación de una calificación y la subjetividad de quien evalúa.

## BIBLIOGRAFÍA

- [Anderson *et al.*, 90] Anderson, J., Boyle, F.; Corbett, A.; & Lewis, M. *Cognitive Modeling and Intelligent Tutoring in Artificial Intelligence & Learning environments*, Ed W. J. Clancey & E. Soloway, MIT 1990.
- [Bahamonde, 97] Bahamonde, A. Aprendizaje automático: tercer ciclo en ciencias de la computación. Centro de Inteligencia Artificial, Universidad de Oviedo en Gijón. Curso académico 1996-1997.
- [Ferrero, Urreatavizcaya, Fernández, 97] Ferrero, B.; Urreatavizcaya, M.; & Fernández-Castro, I. Un sistema de diagnóstico y evaluación basado en detectives. Actas de la VII Conferencia de la Asociación Española para la Inteligencia Artificial. Málaga, 1997.
- [García, 89] García, J. M. *Bases pedagógicas de la evaluación: guía práctica*. Ed. Síntesis. Barcelona, 1989.
- [Gutiérrez, 94] Gutiérrez, J. INTZA: Un sistema tutor inteligente para entrenamiento en entornos industriales. PhD Dissertation, University of the Basque Country, San Sebastian. Spain. 1994.
- [Jerinic, Devedzic, Radovic, 97] Jerinic, L. J., Devedzic, V., Radovic, D.: The GET-BITS model of control knowledge in intelligent tutoring. Actas de la VII Conferencia de la Asociación Española para la Inteligencia Artificial. Málaga, 1997.
- [Milech *et al.*, 93] Milech, D., Krisner, K., Roy, G., & Waters, B. "Applications of Psychology to Computer-Based Tutoring Systems". In: *International Journal of Human-Computer Interaction*, University of Western Australia, Nedlands, Australia. 1993.
- [Ranilla, 98] Ranilla, J. Abanico: Aprendizaje basado en agrupación numérica en intervalos continuos. Memoria de Investigación del tercer ciclo en ciencias de la computación e inteligencia artificial, Universidad de Oviedo, 1998.
- [Ranilla, Mones, Bahamonde, 97] Ranilla, J., Mones R., Bahamonde, A. Segmentación de valores numéricos para el aprendizaje a partir de ejemplos. Actas de la VI Conferencia de la Asociación Española para la Inteligencia Artificial. Málaga, 1997.
- [Rossi, Schwabe, Garrido, 97] Rossi, G., Schwabe, D., & Garrido, A. Design Reuse in Hypermedia Applications Development. Proceedings of ACM International Conference on Hypertext (Hypertext'97), Southampton, ACM Press. 1997.
- [Santos, 93] Santos, M. *La evaluación: un proceso de diálogo, comprensión y mejora*. Ediciones Aljibe. Granada. 1993.
- [Vadillo, Díaz de Ilarraza, 97] Vadillo, J. & Díaz de Ilarraza, A. GenEx: Un sistema para la generación dinámica de explicaciones en sistemas tutores inteligentes de entrenamiento. Actas de la VII Conferencia de la Asociación Española para la Inteligencia Artificial, Málaga, 1997.
- [Weber, Specht] Weber, G., Specht, M. Episodic Learner Model: The Adaptive Remote Tutor. <http://www.psychologie.unitrier.de:8000/projects/ELM/elmart.html>