

Mecanismos de ayuda inteligentes en test adaptativos

Ricardo Conejo, Eduardo Guzmán, José-Luis Pérez-De-La-Cruz y Eva Millán

Departamento de Lenguajes y Ciencias de la Computación
Universidad de Málaga, España.
{conejo, guzman, perez, eva}@lcc.uma.es

Resumen. En este artículo se presenta una propuesta para la introducción de mecanismos adaptativos de presentación de pistas en un entorno de test adaptativos. El artículo comienza con una discusión de aspectos relativos a mecanismos de selección adaptativos de pistas, que resultan en la definición de dos axiomas que tales pistas deben cumplir. Posteriormente se presenta una evaluación empírica preliminar con estudiantes reales, cuyo objetivo es evaluar un banco preliminar de preguntas con sus respectivas pistas para determinar si dichas pistas son realmente útiles para alumnos con diferentes niveles de conocimiento.

1. Introducción

La evaluación mediante pruebas tipo test se usa de modo habitual en diferentes contextos educativos y con diversos objetivos: calificación, auto-evaluación, diagnóstico cognitivo, etc. Para mejorar la eficiencia del proceso de diagnóstico, los sistemas de tests adaptativos seleccionan la mejor pregunta de acuerdo a la estimación actual de características relevantes del examinando. De este modo, se alcanza un grado de exactitud mayor en la evaluación a la vez que se reduce significativamente el número de preguntas. En este sentido, se pueden encontrar diferentes propuestas para test adaptativos [1], [2]. Una de las más utilizadas es la *Teoría de la Respuesta al Item (TRI)* [3], que tiene un fundamento teórico sólido que supone que la respuesta a una pregunta depende de un rasgo latente desconocido. Este rasgo latente θ representa un factor psicológico que se desea medir a través del test, y que no es directamente observable. En entornos educativos, se corresponde con el conocimiento del alumno que se examina.

En cualquier entorno educativo adaptativo, es necesario tener estimaciones fiables del nivel de conocimiento del alumno, con objeto de elegir qué tipo de acción instructiva es más apropiada en cada momento. En este sentido, los *Test Adaptativos Informatizados (TAI)* [4] basados

en la TRI proporcionan una herramienta de diagnóstico fiable y eficiente. SIETTE [5], [6] es un sistema de evaluación a través de Internet que implementa TAIs basados en una discretización de la IRT. Una de las características principales de SIETTE es que es una herramienta que puede ser integrada sin dificultad en cualquier entorno de aprendizaje Web (para un ejemplo de integración, véase [7]). Al realizar esta integración, SIETTE maneja todo lo relativo al modelado del alumno (básicamente, creación y mantenimiento del modelo del alumno). SIETTE está accesible en <http://www.lcc.uma.es/SIETTE>.

Por otro lado, actualmente no cabe duda de que una de las contribuciones principales a la psicología educativa en el siglo XX es la *Zona del Desarrollo Próximo* ZDP de Vigotsky [8]. Una definición corta y operativa para nuestros actuales objetivos es la dada en [9]: la zona definida mediante la diferencia entre el rendimiento que demuestra un niño (en nuestro caso, una persona) en un test en dos condiciones diferentes: con y sin ayuda.

Poco después de la definición de la ZDP, surgen los primeros intentos de aplicar este concepto en el contexto de la administración de tests, bajo las dos condiciones descritas (con o sin ayuda), normalmente con el objetivo de clasificar a los alumnos para poder asignarlos a los programas educativos más apropiados. Pero el mayor objetivo del trabajo que presentamos aquí es diferente: construir un modelo que permita la integración de mecanismos adaptativos de ayudas en el sistema SIETTE.

Las pistas constituyen una táctica adecuada y eficiente en el proceso de enseñanza. Muchos sistemas tutores inteligentes (STI) también dan pistas al alumno, como por ejemplo ANDES [10], que evalúa a los alumnos teniendo en cuenta no sólo el número de respuestas correctas, sino el número de pistas recibidas, o Animalwatch [11], que tiene diferentes tipos de pistas disponibles (muy/poco interactivas, concretas/simbólicas) y las selecciona adaptativamente en base a rasgos del usuario tales como nivel de desarrollo cognitivo y sexo. Asimismo, estudios psicológicos demuestran que un docente suele usar una estimación aproximada del rendimiento del alumno para seleccionar la pista más apropiada [12].

Consecuentemente, supondremos que la ayuda se representa mediante pistas que denotaremos h_1, \dots, h_n que proporcionan diferentes niveles de ayuda para cada pregunta tipo test (en adelante, ítem). Entenderemos que el mecanismo de selección de ayudas es *adaptativo* cuando la pista sea seleccionada automáticamente por el sistema. La

filosofía que subyace en esta adaptación es que se debe tener en cuenta la distancia a la que se encuentra la pista en la ZPD, de modo que se seleccionará siempre aquella pista que proporcione la información mínima que necesita el alumno para ser capaz de responder correctamente a la pregunta.

El trabajo que se presenta en este artículo extiende nuestras investigaciones previas sobre la introducción de pistas y mensajes de refuerzo en test adaptativos [13]. Ahora nuestros objetivos son:

- La definición de un marco teórico para selección adaptativa de pistas,
- Estudio empírico que permita validar la propuesta. Con este objetivo, hemos desarrollado un banco de preguntas real para una asignatura con sus respectivas pistas asociadas. Este banco ha sido usado en varios experimentos con examinandos reales, con el objetivo de validar el banco de pistas.

Una cuestión que merece la pena considerar y que planeamos realizar en un futuro próximo es la calibración de los pares ítem-pista, que permitirá la administración de tests en los cuales tanto los ítems como las pistas se seleccionen de modo adaptativo de acuerdo a la estimación actual del nivel de conocimiento del alumno.

La estructura del resto del artículo es la siguiente: en la sección 2, se introduce unos fundamentos teóricos básicos en TAIs e IRT. En la sección 3, discutimos diversos aspectos relativos a la introducción de pistas en entornos de test adaptativos, mientras que la sección 4 presenta los resultados de los experimentos realizados, en los que se ha desarrollado un banco preliminar de ítems con sus correspondientes pistas que posteriormente han sido presentados a dos grupos de alumnos reales, para determinar de modo empírico la validez de estas pistas. El artículo termina presentando las conclusiones que hemos alcanzado en este estudio junto con algunas líneas de trabajo futuro.

2. Los test adaptativos informatizados y la teoría de la respuesta al ítem

Los test adaptativos informatizados y la teoría de la respuesta al ítem constituyen hoy en día el marco teórico estándar para una evaluación fiable y eficiente del alumno. Un test adaptativo informatizado se define como un test en el cual la decisión de qué ítem presentar en cada

momento y la decisión de cuando finalizar el test se adoptan de un modo dinámico en función del rendimiento del alumno en las preguntas anteriores. Si se comparan dos examinandos, normalmente reciben preguntas diferentes y en distinto orden. Para lograr este objetivo, a cada pregunta i en un test se le asigna la llamada *Curva Característica del ítem* (CCI), que es una función que representa la probabilidad de que el alumno dé una respuesta correcta a ese ítem dado su nivel de conocimiento θ , que es el rasgo latente no observable que deseamos estimar a través de la realización del test. Una hipótesis importante es que este rasgo latente no cambia durante el test. La probabilidad P_i de contestar correctamente a un ítem u_i se puede calcular mediante la expresión $P_i = P(u_i = 1/\mathbf{q})$, y, en consecuencia, la probabilidad Q_i de fallar vendrá dada por la expresión $Q_i = P(u_i = 0/\mathbf{q}) = 1 - P(u_i = 1/\mathbf{q})$. Si el test se compone de n ítems, conocemos las CCIs, y suponemos la independencia local de los ítems, la función de verosimilitud L se calcula mediante la expresión:

$$L(u_1, u_2, \dots, u_n/\theta) = \prod_{i=1}^n P_i^{u_i} Q_i^{1-u_i}$$

El máximo de esta función da una estimación del valor más probable de \mathbf{q} . A La distribución de probabilidad de \mathbf{q} se puede obtener fácilmente aplicando n veces la regla de Bayes. Una hipótesis habitual es que las CCI pertenecen a una familia de funciones que dependen de uno, dos o tres parámetros y que se construyen basándose en la distribución normal o logística. Así por ejemplo, el modelo logístico de tres parámetros asume que la CCI se define mediante la expresión:

$$CCI_i(\mathbf{q}) = c_i + (1 - c_i) \frac{1}{1 + \exp[-1.7a_i(\mathbf{q} - b_i)]}$$

donde c_i es el factor de adivinanza, b_i es la dificultad del ítem y a_i es el *factor de discriminación*. El factor de adivinanza es la probabilidad de que un alumno sin conocimientos conteste la pregunta correctamente. La dificultad del ítem representa el nivel de conocimiento para el cual la probabilidad de que el alumno falle es igual a la de que acierte más el factor de adivinanza. El factor de discriminación es proporcional a la pendiente de la curva.

Basándonos en los TAIs y la IRT, en nuestro grupo se ha desarrollado la herramienta de test adaptativos a través de la Web SIETTE [5], [6]. A diferencia de la IRT tradicional, el nivel de conocimiento θ es en SIETTE una variable de $n + 1$ valores $v_0 < \dots < v_n$

(que en lo que sigue representaremos por $0, 1, \dots, n$). De este modo en SIETTE las CCIs son vectores de $n+1$ elementos y la aplicación de la regla de Bayes se reduce a un producto de $n+1$ valores seguido de un proceso de normalización.

3. Introducción de pistas en un modelo de test adaptativos informatizados

Para introducir las pistas en nuestro modelo, definamos en primer lugar algunos términos:

- *Ítem*. Utilizamos este término para representar de modo genérico una pregunta o ejercicio que el alumno debe resolver. La solución de esta pregunta o tarea se da respondiendo una pregunta de opción múltiple (un enunciado junto con un conjunto de posibles respuestas en el que una y sólo una es la correcta).
- Un *test* es una sucesión de ítems.
- *Pista*. Una pista es una información adicional que se presenta al alumno tras proponerle una pregunta y antes de que la conteste. Puede ser una explicación más detallada del enunciado, una ayuda que le permita descartar una o más respuestas, una indicación de cómo proseguir, etc. Las pistas se invocan de modo activo (el estudiante puede pedir las pistas sin más que pulsar un botón) o pasivo (el sistema selecciona y presenta la pista de acuerdo al comportamiento del alumno, cuando detecta que el alumno tiene dificultades, por ejemplo, si lleva mucho tiempo esperando la respuesta del alumno).

Veamos un ejemplo. Consideremos una pregunta y dos posibles pistas asociadas:

¿Cuál es el resultado de la expresión?: $1/8 + 1/4$?
 a) $3/4$ b) $2/4$ c) $3/8$ d) $2/8$

Algunas pistas posibles para esta pregunta podrían ser:

Pista 1. $1/4$ es equivalente a $2/8$

Pista 2. Primero encuentra fracciones equivalentes con el mismo denominador

Una hipótesis simplificadora que ha sido necesario realizar en esta primera etapa de nuestra investigación es que las pistas no modifican el

conocimiento del alumno (es decir, que el conocimiento del alumno permanece constante durante el test, incluidas las pistas). Esta es una hipótesis controvertida pero habitual en los test adaptativos, ya que hace que el modelo sea computacionalmente tratable. En nuestro caso, esta hipótesis puede interpretarse como que las pistas no cambian el conocimiento del alumno, sino que modifican la forma de la CCI. En este sentido, la pista lo que hace es acercar a la pregunta que actualmente está en la ZDP hasta el nivel de conocimiento del alumno. En este sentido, cabe considerar que la combinación ítem+pista podría considerarse como un nuevo ítem (virtual), que puede medirse y tratarse del mismo modo que otros ítems en el test, es decir, el nuevo ítem tiene asociada una nueva CCI cuyos parámetros pueden ser estimados utilizando las técnicas tradicionales de calibración de ítems. Sin embargo, ambas curvas no son independientes, sino que se deben satisfacer una serie de relaciones entre ellas. En primer lugar, la pista debe hacer que la pregunta sea más fácil. Esta condición se puede establecer en términos matemáticos mediante el siguiente axioma:

Axioma 1. Dada una pregunta q y una pista h , para todos los niveles de conocimiento k se debe tener que $CCI_q(k) \leq CCI_{q+h}(k)$, donde CCI_q representa la curva original y CCI_{q+h} representa la curva característica de la combinación ítem/pista.

Si el examinando usa una combinación de pistas, la cuestión debería ser aún más fácil. Matemáticamente esta condición puede expresarse así:

Axioma 2. Dada una pregunta q , un conjunto de pistas H y una pista $h \notin H$, para todos los niveles de conocimiento k se tiene que $CCI_{q+H}(k) \leq CCI_{q+H+(h)}(k)$.

Para un conjunto de ítems y pistas, si tras la calibración de los parámetros de las CCIs de los ítems virtuales y reales las curvas resultantes no satisfacen los dos axiomas anteriores, significa que alguna de las informaciones suministradas como pistas no es tal, sino más bien un elemento que confunde al alumno en lugar de ayudarlo y por tanto debe ser descartado. Este enfoque es simple pero proporciona un método empírico útil que permite la validación de las pistas propuestas.

En entornos adaptativos, tiene sentido buscar criterios que permitan la selección de la mejor pista disponible. Bajo el marco de la ZDP, si el alumno no es capaz de contestar correctamente a la pregunta pero ésta se sitúa en su ZDP, la mejor pista disponible sería aquella que trae al ítem I desde la ZDP a la zona del conocimiento del alumno, y por supuesto

dependerá de cómo de “lejos” esté la pista en la ZDP, es decir, cuanto más lejos esté la pista del alcance del alumno según su conocimiento más detallada, completa o explicativa deberá ser la pista. De este modo, si por ejemplo un ítem I tiene asociada tres pistas h_1 , h_2 y h_3 a diferentes niveles de detalle, significa que cada una de ellas es apropiada para una zona diferente de la ZDP, tal como se representa en la Figura 1.

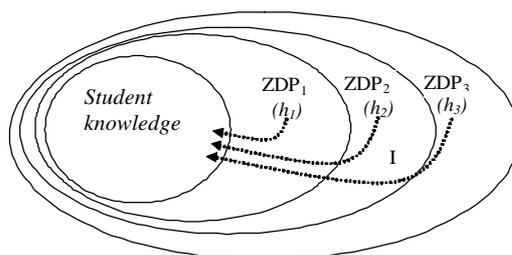


Fig.1. Conocimiento del alumno, ZDPs y pistas

De modo que la selección de h_i como la mejor pista a presentar significaría que el ítem I está en la zona ZPD_i para este alumno (en el ejemplo de la Figura 1, la mejor pista es h_2 ya que el ítem está en la zona ZPD_2). Una posibilidad de selección adaptativa de pistas es usar los mecanismos de adaptación de ítems clásicos: dada la estimación de conocimiento de un alumno $q(k)$ y dadas dos pistas h_1 , h_2 , y las respectivas curvas características $CCI_{q+\{h_1\}}(k)$ y $CCI_{q+\{h_2\}}(k)$, la mejor pista es aquella que minimice la varianza esperada de la distribución de probabilidad a posteriori. Este mecanismo es simple de implementar y no hace modificaciones sustanciales en el procedimiento adaptativo, porque el test se usa como mecanismo de evaluación y no como herramienta de aprendizaje. Sin embargo, el uso de mecanismos adaptativos de selección de pistas en este contexto puede servir de estímulo para el alumno y por tanto mejorar su grado de confianza en sí mismo.

4. Experimento

Un paso importante para la integración de este nuevo enfoque de pistas adaptativas en el sistema SIETTE es la calibración de ítems/pistas. Esta es un objetivo difícil, y para ello se debe seguir una metodología que consta de los siguientes pasos:

1. En primer lugar hay que acometer el costoso proceso de desarrollar un banco de ítems, que en este caso conlleva la dificultad añadida de tener que desarrollar también varias pistas para cada uno de estos ítems.
2. En segundo lugar se debe evaluar la calidad de este banco de ítems y pistas,
3. Finalmente, los parámetros necesarios para cada CCI correspondiente a una combinación ítem+pista) deben aprenderse de los datos lo que obliga a hacer pruebas con alumnos reales.

El trabajo que presentamos aquí informa sobre avances en los pasos 1 y 2. En cuanto al primer paso, se ha desarrollado un banco de preguntas para la asignatura *Procesadores de Lenguajes* de la ETSI Informática de la Universidad de Málaga. Cada una de las preguntas tiene entre 2 y 4 pistas asociadas. Ejemplos de ítems (junto con sus pistas asociadas) que fueron propuestos en dichos tests son:

1. Indicar la salida que produce el siguiente programa LEX con la entrada `|*abc*|`

```
ab/c { printf("one "); }
c     { printf("two "); }
abc  { printf("three "); }
```

- a) tres b) uno dos c) uno d) uno dos tres

Pista 1. El contenido de `yytext` no incluye los caracteres correspondientes a la parte de la expresión que hay después del operador de lectura avanzada /

Pista 2. Cuando hay un operador de lectura avanzada /, la longitud del lexema que se forma es la correspondiente a la expresión situada a la izquierda del operador

2. Sea T el conjunto de todos los caracteres ASCII del 1 al 127. Para expresar en LEX el conjunto de todas las cadenas que pueden formarse con los símbolos de T, puede usarse la expresión:

a) `(.|\\n)*` b) `[a-zA-Z0-9]*` c) `.*` d) `[.]*`

Pista 1. El operador `.(punto)` representa a cualquier carácter ASCII, exceptuando el fin de línea

Pista 2. El operador `.(punto)` no tiene ningún significado

Pista 3. El alfabeto ASCII incluye no solo letras y números, sino también caracteres especiales, operadores, signos de puntuación, etc

En relación con el segundo paso, dos grupos diferentes de alumnos de la asignatura *Procesadores de Lenguajes* en la ETSI de Informática de la Universidad de Málaga (cursos 2003/2004 y 2004/2005) fueron calificados mediante un examen tipo test no adaptativo consistente en 20 preguntas que fue administrado a través de SIETTE, en el que se les

ofrecía la posibilidad de solicitar pistas para cada uno de los ítems. Los ítems eran de opción múltiple y cada uno de ellos tenía entre dos y cuatro pistas asociadas. Una vez que un alumno solicitaba una pista, ésta era seleccionada aleatoriamente dentro del conjunto de pistas disponibles. Sin embargo, el uso de una pista suponía una penalización en la valoración que recibía la respuesta correcta (1 punto por pregunta, con un mínimo para aprobar de diez puntos): en el primer experimento el alumno obtenía 0.5 puntos por respuesta correcta en la que hubiese utilizado una ayuda, mientras que en el segundo obtenía 0.75. Una vez que los alumnos habían acabado el examen, rellenaban un cuestionario sencillo en el cual informaban de cuántas pistas habían solicitado y cuántas de ellas les habían resultado útiles para seleccionar la respuesta correcta. Los resultados se presentan en la Figura 2, en la que se observa un uso escaso de las pistas que se explica por el factor de penalización. Podemos ver también que los alumnos no consideraron que las pistas fuesen demasiado útiles, probablemente debido a que la selección de pistas era aleatoria que es precisamente lo que intentamos mejorar con este trabajo.

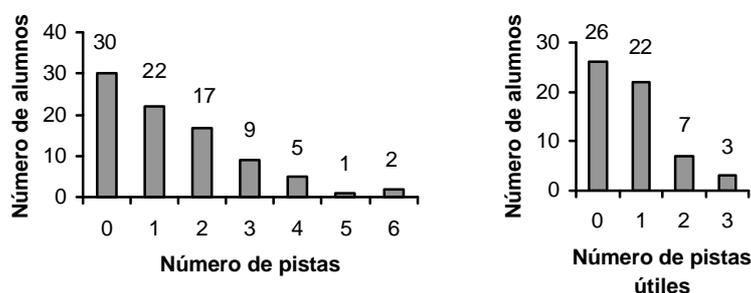


Fig. 2. Resultados de los cuestionarios de los alumnos en el curso 2003/2004

Hemos asimismo realizado un análisis de la traza de los dos experimentos¹. Con este objetivo, los alumnos se agruparon según su rendimiento en el test: los alumnos del Grupo 1 habían contestado correctamente menos de la mitad de las preguntas (84 alumnos en total), mientras que el Grupo 2 se formó con los 94 alumnos restantes. En las siguientes tabla y figura vemos los datos correspondientes a los dos ítems de ejemplo que se han presentado: el ítem 1, con dos pistas asociadas, y el

¹ En este análisis fue necesario unir los datos de ambos experimentos dado que en otro caso el conjunto de datos a analizar hubiese sido demasiado pequeño, debido al uso escaso de las pistas por parte de los alumnos provocado por la penalización asociada al mismo).

ítem 2, con tres pistas. Las tablas muestran los resultados por cada ítem, es decir, el número de alumnos que resolvieron correctamente el ítem en cada uno de los grupos, así por ejemplo el 26/75 de la primera posición de la primera fila significa que, en el grupo 1, 75 alumnos contestaron el ítem 1 sin pistas y 26 de ellos fueron capaces de resolverlo correctamente; 2/3 significa que a 3 de los estudiantes que pidieron pistas se les presentó la pista 1, y 2 de ellos contestaron correctamente, y así sucesivamente. Estos mismos datos se representan en porcentajes en la Figura 3.

Ítem 1	<i>Sin pista</i>	<i>Pista 1</i>	<i>Pista 2</i>	Ítem 2	<i>Sin pista</i>	<i>Pista 1</i>	<i>Pista 2</i>	<i>Pista 3</i>
<i>Grupo 1</i>	26/75	2/3	2/6	<i>Grupo 1</i>	48/71	1/4	2/5	3/4
<i>Grupo 2</i>	56/86	4/5	3/3	<i>Grupo 2</i>	71/79	2/2	5/8	4/5

Tabla 1. Proporción de alumnos que contestaron correctamente en cada situación.

Somos conscientes de que, desafortunadamente, el uso de la penalización ha hecho que el número de casos de pistas utilizadas sea demasiado bajo para permitirnos establecer conclusiones que sean estadísticamente significativas, pero pensamos que nos pueden servir de primera aproximación para determinar el tipo de información que puede obtenerse mediante el uso de estos análisis. Así por ejemplo en la Figura 3 se observa que para el ítem 1 la pista 1 ayuda tanto a los alumnos más aventajados como a los otros, mientras que la pista 2 sólo es útil para los alumnos más avanzados. Para el ítem 2, la pista 1 confunde a los alumnos de nivel más bajo pero ayuda a los más avanzados, la pista 2 confunde a ambos tipos de alumnos y la pista 3 sólo es informativa para los menos avanzados. Esto sugiere que este tipo de análisis permitiría eliminar las pistas “malas” (las que confunden a ambos grupos, las que no tienen efectos positivos, las que tienen un efecto demasiado positivo) del banco de pistas, mejorando así su calidad. Del mismo modo, las pistas pueden seleccionarse de modo adaptativo dependiendo de la estimación actual del nivel de conocimiento del alumno para satisfacer mejor sus necesidades.

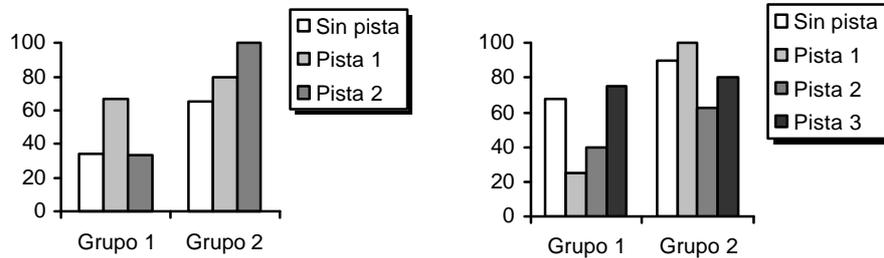


Fig. 3. Porcentajes de respuestas correctas según el nivel de los alumnos.

Estos mismos datos pueden utilizarse para hacer aproximaciones a las curvas CCI, que se presentan en la Figura 4.

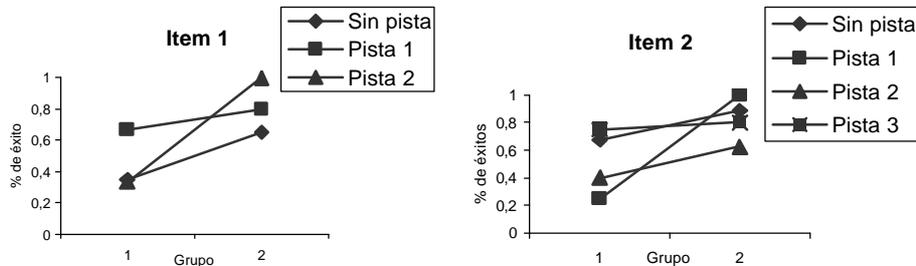


Fig. 4. Aproximaciones para las CCIs.

En la Figura 4 podemos ver que, a excepción de la pista 1 del ítem 1, todas las pistas violan el Axioma 1 que establece que para cualquier nivel de conocimiento k se debe tener que $CCI_q \leq CCI_q + \{h_k\}$. Por tanto, si tal hipótesis se considera razonable las pistas que no la cumplen deberían ser descartadas, a menos que la adaptación de las pistas al nivel de conocimiento del alumno sea el objetivo deseado.

5. Conclusiones y trabajo futuro

En este artículo se han presentado algunas ideas sobre la introducción adaptativa de pistas en un entorno de aprendizaje adaptativo, basado en conceptos de la TRI.

Las pistas no se consideran como modificadores del conocimiento del alumno, sino como modificadores del CCI de una pregunta. Algunos axiomas formales que todo modelo de pistas deben satisfacer han sido presentados y justificados de un modo informal. Un estudio preliminar ha

sugerido que el uso de las pistas en estos entornos es adecuado y factible. Como se hemos indicado anteriormente, el paso siguiente es la calibración de las CCI para cada par ítem-pista utilizando datos empíricos. Las curvas obtenidas permitirán la validación de tales pistas y servirán de base para la integración e implementación de este modelo en SIETTE, para permitir la selección de adaptativa tanto de ítems como de pistas en nuestro sistema de tests.

Referencias

- [1]. Chua-Abdullah, S. (2003) Student Modelling by Adaptive Testing - A Knowledge-based Approach. PhD Thesis. University of Kent. Canterbury.
- [2]. Rudner, L. M. An (2002) Examination of Decision-Theory Adaptive Testing Procedures. Paper presented at the annual meeting of the American Educational Research Association; New Orleans.
- [3]. Hambleton, R. K.; Swaminathan, J., and Rogers, H. J.(1991) Fundamentals of Item Response Theory. California, USA: Sage publications.
- [4]. H. Wainer (ed.), (1990). Computerized adaptive testing: a primer, Lawrence Erlbaum Associates, Hillsdale, NJ.
- [5]. R. Conejo, E. Guzmán, E. Millán, M. Trella, J. L. Pérez-De-La-Cruz and A. Ríos, (2004). *SIETTE: A Web-Based Tool for Adaptive Testing*. International Journal of Artificial Intelligence in Education, 14, pp. 29-61.
- [6]. Guzmán, E. and Conejo, R. (2004) A brief introduction to the new architecture of SIETTE. LNCS 3137 . New York: Springer-Verlag; pp. 405--408.
- [7]. Millán, E., García_Hervás, E. et al (2004). TAPLI: An Adaptive Web-Based Learning Environment for Linear Programming. LNCS 3040, 676-685.
- [8]. Vygotskii, L. S., (1994). The Vygotskii Reader, Blackwell, Cambridge, Massachusetts.
- [9]. G. Wells, (1999). Dialogic inquiry: Towards a socio-cultural practice and theory of Education, chapter 10. New York: Cambridge University Press.
- [10]. Gertner, A. S., Conati, C., VanLehn, K. (1998): Procedural Help in ANDES: Generating Hints Using a Bayesian Network Student Model. Proc. 15th Nat. Conf. on Artificial Intelligence, pp. 106-111.
- [11]. Arroyo, I. Animalwatch: an arithmetic ITS for elementary and middle school students. "Learning Algebra with the Computer" Workshop. Fifth International Conference on Intelligent Tutoring Systems. Montreal, Canada. June 2000.
- [12]. Hume, G. D., Michael, J., Rovick, A., Evens, M. W. (1996): Hinting as a tactic in one-on-one tutoring. Journal of Learning Sciences 5(1) pp. 23-47.
- [13]. Conejo, R, Guzmán, E. and Pérez-de-la-Cruz, J.L. Towards a computational theory of learning in an adaptive testing environment. Proceedings of the 11th International Conference on Artificial Intelligence in Education, IOS Press.