

ESCUELA TÉCNICA SUPERIOR DE

INGENIERÍA INFORMÁTICA

TRABAJO FIN DE MÁSTER EN  
INGENIERÍA DEL SOFTWARE E  
INTELIGENCIA ARTIFICIAL

**Título: Modelado del proceso de aprendizaje en un  
entorno de educación infantil**

Autor: Javier García-Herreros Castellero

Tutor: Dr. Ricardo Conejo Muñoz

Curso académico: 2009-2010



El Dr. D. Ricardo Conejo Muñoz, Catedrático de Universidad perteneciente al Área de Lenguajes y Sistemas Informáticos de la E.T.S.I. Informática de la Universidad de Málaga

certifica que,

D. Javier Garcia-Herreros Castellero, Ingeniero en Informática, ha realizado en el departamento de Lenguajes y Ciencias de la Computación de la Universidad de Málaga, bajo su dirección, el trabajo de investigación titulado:

*Modelado del proceso de aprendizaje en un entorno de educación infantil*

Revisado el presente trabajo, estima que puede ser presentado al tribunal que ha de juzgarlo, y autoriza la presentación de esta Tesina para la obtención del título de Máster en Ingeniería del Software e Inteligencia Artificial correspondiente al programa oficial de postgrado de Tecnologías Informáticas.

Fdo.: Ricardo Conejo Muñoz

Fdo.: Javier García-Herreros Castellero

Málaga, Septiembre 2010



# Índice

<b>1. Contexto: Proyecto PATIO</b>	<b>1</b>
<b>2. Motivación</b>	<b>1</b>
<b>3. Problemas que se presentan</b>	<b>2</b>
<b>4. Estado del arte</b>	<b>3</b>
4.1. Modelado de tareas . . . . .	4
4.2. Modelado de usuario . . . . .	5
4.3. IRT: Item response theory . . . . .	6
4.3.1. Modelado Unidimensional . . . . .	7
4.3.2. Modelado Multidimensional . . . . .	10
4.4. Estimación de parámetros . . . . .	15
4.4.1. Expectation-Maximization (EM) . . . . .	15
4.4.2. Suavizado de núcleo . . . . .	15
<b>5. Investigaciones iniciales</b>	<b>16</b>
5.1. Modelado de tareas . . . . .	16
5.2. Modelado de usuario . . . . .	20
5.2.1. Actualización del Modelo de Usuario . . . . .	21
<b>6. Trabajo en el aula</b>	<b>24</b>
<b>7. Conclusiones y plan de trabajo</b>	<b>26</b>



## 1. Contexto: Proyecto PATIO

El propósito de este proyecto es investigar en la modelización, implementación y evaluación de métodos y técnicas de representación para definir, realizar, apoyar y monitorizar tareas de aprendizaje individual y colaborativo en un marco multicultural, soportadas por ordenador y otros dispositivos tecnológicos, gracias al uso de técnicas derivadas de la Inteligencia Artificial, el Modelado Cognitivo y el diseño de recursos y materiales pedagógicos.

La finalidad del proyecto es investigar cómo las tecnologías de la información y las comunicaciones posibilitan la creación y gestión de entornos educativos innovadores que supongan una mejora en los procesos de aprendizaje individual, a la vez que favorezcan la integración de los niños en el entorno social al que pertenecen, tanto a un micro nivel (su aula y su colegio) como a macro nivel (la sociedad andaluza en la que viven).

Para abordar el proyecto, nos planteamos definir un conjunto de elementos tecnológicos que den soporte a la enseñanza, que ayuden al alumno y al profesor como una herramienta más en el aula. Como tema de investigación, se trabaja en modelizar los procesos de aprendizaje a partir de la Inteligencia Artificial de forma que permitan inferir automáticamente diferentes situaciones, problemas en el proceso educativo, etc., y de esta forma intervenir y mejorar el rendimiento de los niños en el aula y, en general, favorecer su desarrollo personal y social.

Para ello se necesita un modelo de usuario adaptado a las características de este tipo de enseñanza, es decir, debe poder representar todas las características necesarias para modelar el proceso de aprendizaje. Este modelo de usuario debe tener como premisa básica la multidimensionalidad, ya que los alumnos no aprenden un solo concepto a la vez. Esto implica un esfuerzo adicional en el diseño de los ejercicios y en la complejidad que se genera en forma de relaciones entre conceptos.

## 2. Motivación

En los últimos tiempos la tecnología ha avanzado notablemente y su introducción en diversas áreas ha tenido un gran impacto, trayendo consigo importantes beneficios. Sin embargo este avance no se ha visto reflejado en la educación infantil, o al menos no en todo su potencial. Creemos que el uso de estas nuevas tecnologías puede ayudar, por un lado al docente en la creación de tareas, monitorización del redimiendo del alumno, detección precoz de problemas, etc., y por otro lado al alumno, que podrá realizar los ejercicios de forma más atractiva, siendo más independientes. Además, hoy en día existen alumnos extranjeros, que a edades tempranas todavía no dominan correctamente el lenguaje español por lo que un sistema multilinguaje puede facilitar su integración.

Al diseñar un sistema con este propósito se deben tener en cuenta muchos aspectos, aunque uno de los más importantes es el aspecto del modelo del usuario,

que será el que posibilitará almacenar la estimación del conocimiento de cada alumno y adaptarse a él. Para comprender el problema al que nos enfrentamos, veamos el siguiente ejemplo. Se tiene un problema en el que se le pide al alumno que escoja las frutas rojas y las arrastre a un carrito de la compra, similar al mostrado en la figura 7. Para resolverlo correctamente el alumno debe tener conocimientos del color rojo y del concepto de fruta. Una vez completado el ejercicio, el sistema deberá comprobar la corrección del mismo y actualizar su modelo de usuario, pero en este caso, al no ser el problema unidimensional (ya que para resolverlo hace falta el conocimiento de más de un concepto), no se puede hacer individualmente por cada concepto sino que hay que actualizar el “concepto” fruta-rojo. En general, en caso de que el problema se resuelva correctamente, se puede considerar que se conocen todos los conceptos implicados, en mayor o menor medida, sin embargo, en caso de que el alumno no consiga resolverlo satisfactoriamente puede ser debido a que no conozca uno de los conceptos implicados. En determinados casos se podrá comprobar cuál es el concepto que falla (por ejemplo si el alumno hubiera seleccionado todos los objetos rojos pero no todos fueran una fruta), pero en general no se puede saber. Además, en general no sólo se puede actualizar el modelo de usuario para indicar que conoce un concepto en el dominio de las frutas y otro en el dominio de los colores, sino también en el dominio cognitivo, ya que en parte demuestra que distingue los colores y distingue las frutas.

El objetivo es tener una estimación de cada habilidad o conocimiento de cada concepto por separado, con lo que se deberán elaborar técnicas para tal fin.

### 3. Problemas que se presentan

Debido a las particularidades del entorno (niños en edad temprana en el aprendizaje, etc.) nos encontramos con determinados problemas específicos que se detallarán a continuación:

- Forma de aprendizaje: En general a los niños de 3 a 5 años no se les puede enseñar de la misma forma en la que se forma a un adulto. Primero porque aprenden conceptos mucho más generales que los adultos, donde la enseñanza es más específica. Y segundo porque dicha enseñanza se basa más en juegos que en teoría. Además, también se deben desarrollar habilidades psicomotrices, sociales, etc., y no sólo intelectuales con lo que, tanto el dominio como la forma de instrucción, es completamente diferente a lo que estamos acostumbrados en entornos como institutos o universidades. Los niños suelen trabajar en grupo, donde se pueden ayudar entre ellos. Esto también les enseña a socializar, etc. Normalmente el docente no puede atender a todos ellos a la vez, por lo que se debe centrar en cada alumno o grupo individualmente. Además, los conceptos aprendidos se deben reforzar cada poco tiempo para afianzarlos. Esto se consigue entrelazando los conceptos y relacionándolos entre sí (por ejemplo, si se está enseñando



seguridad vial se puede pedir al alumno que cuente el número de luces de un semáforo para que recuerde los números). Por tanto, el docente podría beneficiarse de un sistema de aprendizaje en el que los niños puedan ser autónomos.

- Diferencia de nivel según la edad: El cerebro de los niños pequeños tiene una plasticidad muy elevada. Esto significa que su tasa de aprendizaje es altísima, lo que implica que de un mes a otro puedan aprender una gran cantidad de información, conceptos, etc. Como en una determinada clase puede haber alumnos cuya diferencia de edad sea cercana al año, ocurre que unos niños tienen un nivel mucho más elevado que otros. Por tanto hay que adaptarse a esta circunstancia, donde hay una gran heterogeneidad en los niveles y las capacidades de los alumnos. Además, dependiendo de la edad, puede ocurrir que algo que esté bien resuelto para un alumno no lo esté para otro porque debiera tener un nivel más alto. Esto hace que haya que personalizar la corrección para cada alumno.
- Complejidad del modelo de usuario: Normalmente un modelo de usuario se suele definir para un dominio concreto. El problema al que nos enfrentamos aquí es que el currículum de los alumnos de infantil es extremadamente amplio ya que deben aprender una gran cantidad de conceptos básicos, desde el desarrollo de la grafomotricidad, formas y colores, pasando por objetos de la vida cotidiana, letras, formación de frases, etc. Además de las relaciones entre ellos, ya que raramente aprenden conceptos aislados. Por tanto es un modelo bastante heterogéneo en cuanto al dominio de los conceptos y además es multidimensional, lo que añade cierta dificultad.

## 4. Estado del arte

A continuación describiremos diferentes técnicas que se utilizan hoy en día para conseguir los objetivos que nos proponemos.

En general, un sistema educacional debe tener diferentes módulos que realicen una tarea concreta. Los más comunes son:

- La interfaz gráfica, que se encarga de ser el intermediario entre el humano (ya sea profesor, alumno, etc.) y el sistema.
- El modelo del usuario, que proporciona información acerca de las habilidades de los alumnos representando todas las características que se deben tener en cuenta para una correcta instrucción del mismo.
- La base de conocimiento, donde se almacena toda la información que el alumno debe aprender y los recursos a disposición del profesor/sistema para crear los ejercicios.

- El razonador, el cual inferirá una estimación del conocimiento del alumno a partir de la resolución de un ejercicio. Además detectará determinados patrones que puedan sugerir algún problema en el proceso educacional de los alumnos para poder intervenir convenientemente.

Ejemplos de este tipo de sistemas se pueden ver en [10, 5].

#### 4.1. Modelado de tareas

El crear un entorno apto para la resolución de problemas por parte de alumnos no es un proceso trivial, especialmente con alumnos de educación infantil. Hoy en día existen determinados enfoques que intentan resolver esta problemática, tratados en el ámbito de HCI (*Human Computer Interaction*), que tratan de comprender el proceso que siguen los estudiantes a la hora de resolver una tarea conociendo el objetivo al que deben llegar, la representación del conocimiento del usuario en el ejercicio, etc. Uno de ellos se explica en [19], el cual está basado en análisis y modelado de tareas. El objetivo del análisis de tareas es comprender las metas y las tareas del usuario, describiendo de forma explícita qué involucra cada paso de las mismas. Al utilizar estas técnicas se consigue definir una tarea de una forma más formal a la par que puede ser evaluada en cuanto a usabilidad. También destaca la característica de que alguien que no sea el analista pueda comprender mejor el funcionamiento de la misma, sus objetivos, etc.

De acuerdo a la forma de trabajo propuesta en [19], primero se generaran unos prototipos del entorno educacional basados en las especificaciones iniciales y en suposiciones generales sobre resolución de problemas en los entornos educacionales, etc. A continuación estos prototipos se refinan según los datos recogidos en experimentos de campo (es decir, se evalúa el uso del sistema con usuarios reales). Este desarrollo es iterativo ya que el sistema se va refinando por fases, teniendo en cuenta el comportamiento de los usuarios con el sistema en los experimentos.

El propósito del marco de trabajo propuesto es identificar los problemas que puedan ocurrir en la interacción del estudiante con el sistema que dificulten la realización de la tarea. De esta forma, la resolución del problema debería depender única y exclusivamente del conocimiento del estudiante y no de factores externos tales como dificultades en el uso del sistema, etc. Para ello se consideran tres fases:

1. Análisis y desarrollo de un modelo preliminar de tareas realizado por el diseñador, basado en requisitos estimados a priori. El resultado se llama *Designer Task Model* (DTM).
2. Evaluación de campo, donde se recogen datos sobre el comportamiento de los alumnos frente al sistema y además se analizan y se interpretan los datos recolectados.

3. Desarrollo y refinamiento del *Student Task Model* (STM), basado en datos de campo.

El resultado de este proceso será un sistema basado en las necesidades reales de los alumnos y adaptado a ellos.

## 4.2. Modelado de usuario

El modelo de usuario es esencial en un sistema adaptativo. Este modelo representa el conocimiento del usuario, lo que es esencial para que el sistema pueda ofrecerle contenidos adaptados a él. Este modelo debe ser actualizado con los datos que se pueda recopilar de la interacción del usuario con el sistema, ya sea mediante la observación de sus acciones o bien solicitando directamente al usuario que introduzca determinada información. Este proceso se conoce como *modelado de usuario*. Dependiendo del tipo y la cantidad de adaptación que se quiera ofrecer, el modelo de usuario deberá ser más o menos complejo.

La información para realizar la adaptación puede ser de muy diferente índole como por ejemplo los intereses del usuario, sus metas, antecedentes, formación, experiencia, características individuales (personalidad, estilos de aprendizaje, etc.), el conocimiento sobre un tema, etc. En nuestro caso nos centraremos en este último punto ya que nuestro principal objetivo es que los alumnos aprendan los conceptos del currículum de infantil. Además es una de las características más importantes en cualquier sistema de enseñanza adaptativo, siendo en muchos casos la única modelada.

Hay que tener en cuenta que el conocimiento puede incrementarse o decrementarse de una sesión a otra, por lo que el sistema debe darse cuenta de este hecho y actualizar el modelo de usuario. Para diseñar el modelo de usuario existen diferentes técnicas, siendo la más simple la escalar, la cual asigna a cada alumno un valor de conocimiento. La precisión de este sistema es muy pequeña, ya que una persona puede tener diferentes conocimientos dentro de un mismo dominio, por lo que no se le podrá sacar todo el provecho a la adaptación. Por ello se utilizan diferentes tipos de *modelos estructurales*, que consideran que el dominio está a su vez dividido en subdominios. El grado de granularidad puede ser mayor o menor según las necesidades del sistema.

Uno de los modelos estructurales más populares es el *modelo de superposición* [2], que representa el conocimiento del individuo como un subconjunto del modelo del dominio (que a su vez representa el conocimiento del experto), el cual guarda una estimación del conocimiento por cada fragmento del mismo. Este enfoque ha recibido críticas por ser demasiado simple, ya que a veces el conocimiento del alumno no es exactamente un subconjunto del conocimiento del experto, por lo que se mejoró y se desarrolló el llamado *bug model*, que no sólo contiene información de lo que el alumno conoce correctamente sino también de las ideas equivocadas (*misconceptions*) que pueda tener. Uno de los problemas que tiene este modelo es la dificultad para su correcto desarrollo.

El modelo de superposición puede ser más o menos complejo. La forma más sencilla de implementación es que los conceptos sean independientes, sin ninguna relación entre ellos. Este enfoque se llama *set model* o *vector model*. Aunque en determinadas áreas pueda resultar de gran utilidad, el principal problema que presenta es que, al no existir relaciones entre los conceptos, el conocimiento sobre uno de ellos no ayuda a extraer información de los demás por lo que si el dominio es considerablemente amplio se necesitarían muchas observaciones y datos para tener una estimación del conocimiento de todos los conceptos. Existen dos tipos principales de modelos de superposición con conceptos relacionados: el primero consiste en un árbol de conceptos, donde la raíz del mismo es un curso completo y se va dividiendo en subtemas o subconceptos. En el segundo tipo, los conceptos están conectados entre sí por diferentes tipos de relaciones (prerrequisito de, parte de, etc.). Si los conceptos están relacionados entre sí, se puede inferir conocimiento de uno de ellos a partir del otro. Este tipo de extracción de información se llama *propagación del conocimiento* (*knowledge propagation*).

Para tratar con datos poco precisos o inciertos se suelen utilizar otras técnicas como redes bayesianas o lógica difusa. Las redes bayesianas proporcionan un modelo probabilístico basado en la causalidad, que se representa de forma gráfica con un esquema de nodos y conexiones entre los mismos. Cada nodo representa una variable y cada conexión representa una relación causal. Este tipo de modelado permite la diagnosis de manera más natural que otro tipo de enfoque. El modelo de superposición y el de redes bayesianas se pueden unir de tal forma que modelen el conocimiento del estudiante de forma más eficiente. Para ello primero se deben seleccionar las variables de interés para nuestro modelo, las cuales representan conocimiento en diferentes subdominios. Cada una de ellas será un nodo en la red bayesiana, de tal forma que se estructure el dominio en un esquema de relaciones causales. Además, se habrá de elegir correctamente la dirección de los arcos, de tal forma que se especifique correctamente qué habilidad/concepto es la causa de otro.

### 4.3. IRT: Item response theory

En el entorno educacional siempre existe alguna característica que uno quiere medir al realizar la evaluación. Ya sea la capacidad que tiene el alumno para resolver problemas matemáticos, si sabe cierta información de geografía, etc. El alumno debe tener ciertas habilidades para contestar correctamente a las preguntas que se le proponen.

La IRT [11, 15] es una teoría de test que abarca el diseño de los test, análisis de los resultados, etc. Se basa en dos principios: primero que el rendimiento de un estudiante en un test se puede explicar a partir del nivel de conocimiento/habilidad, y segundo que el rendimiento de un estudiante respondiendo a un ítem se puede predecir probabilísticamente y modelarse con funciones llamadas *curvas características*.

En el modelado unidimensional se supone que para cada pregunta hay una sola habilidad (*rasgo latente* en términos psicológicos) de la que depende la respuesta que el alumno proporcione. Por el contrario, en el modelado multidimensional se supone que para contestar correctamente a una pregunta se necesitan diferentes habilidades. En los siguientes puntos se explicará tanto el modelo unidimensional de la IRT como el multidimensional, las diferencias entre los dos, etc.

#### 4.3.1. Modelado Unidimensional

Como se ha comentado anteriormente, en el modelado unidimensional se supone que sólo se utiliza una determinada habilidad para responder a una pregunta. Esta habilidad debe ser medida y por tanto se necesita una escala. En la Teoría de Respuesta al Ítem esta escala tiene como punto medio el cero y un rango  $(-\infty, \infty)$ . En la práctica se utiliza el rango  $[-3, 3]$  ya que en dicho rango es donde se encuentra la información más relevante debido a que a partir de ahí, la función se va acercando muy poco a poco a la asíntota inferior si la habilidad tiende a  $-\infty$ , o a 1 si la habilidad tiende a  $\infty$ .

Para medir esta habilidad generalmente se crea un test que consiste en un número determinados de ítems (preguntas), para las cuales se necesita una determinada habilidad. En general la respuesta que se proporciona puede ser libre, donde la persona podría contestar cualquier respuesta que le parezca apropiada, y de esta forma minimizar el factor del azar. La persona que corrija el test deberá decidir si la pregunta es correcta o no. En función de la respuesta, se puntuará de una determinada forma. En la teoría clásica del test, el resultado del mismo sería la suma de cada una de las puntuaciones de los ítems, sin embargo la IRT se basa en que cada respuesta demuestra un cierto nivel de la habilidad, con lo que se traduce en una actualización continua del nivel de conocimiento del alumno con cada uno de los ítems.

Es razonable pensar que cada alumno tiene una determinada habilidad a la que se le puede asignar un valor numérico en la escala. Esta habilidad se denota por la letra  $\theta$ . Dependiendo de esta habilidad, la probabilidad de responder correctamente a una pregunta cambiará. Si la habilidad es alta, la probabilidad también lo será y viceversa. Esto define una curva llamada *curva característica del ítem*. Existen curvas con uno, dos o tres parámetros, dependiendo de la complejidad que se desee. En el caso del ejemplo de la Figura 1, existen tres parámetros:

- $a$ , que indica la "inclinación" de la curva, es decir, la discriminación que se hace entre los que tienen una alta o baja habilidad.
- $b$ , que indica la dificultad del ítem. En la práctica coincide con el punto de habilidad donde la curva tiene la mayor inclinación.
- $c$ , que es la asíntota inferior, indica la probabilidad de que alguien con el nivel de habilidad más bajo posible responda correctamente a la pregunta,

generalmente al azar. Esto suele ocurrir en los ítems donde hay un número de respuestas y hay que elegir una de ellas, ya de esta forma es más fácil responder correctamente.

La ecuación que define la curva es la siguiente:

$$P(\theta) = c + (1 - c) \frac{1}{1 + e^{-a(\theta - b)}}$$

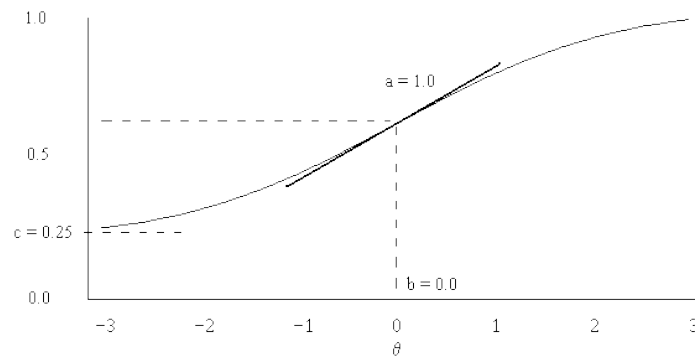


Figura 1: Curva característica del ítem de tres parámetros

La curva mostrada en la figura 1 es la herramienta básica sobre la que se sustenta la IRT, ya que por ejemplo, no es lo mismo que un alumno responda correctamente a un ítem muy difícil que a otro que sea relativamente fácil. En el primero se demuestra un nivel más alto de habilidad.

Los parámetros de esta curva deben ser ajustados a priori. Esto se llama "calibración de los ítems". La calibración consiste en la realización, por parte de alumnos, de determinados tests y, a partir de los resultados asignarles valores a los diferentes parámetros. Para ello existen diferentes técnicas, como por ejemplo la usada en [7].

Una vez conocida la curva característica del ítem se pueden calcular la estimación de las habilidades de los alumnos a partir de los tests. Cuando un alumno realiza un test se le asigna un 1 a cada ítem respondido correctamente y 0 a los que se ha respondido de forma incorrecta. El proceso del cálculo de la habilidad es un proceso iterativo que finaliza cuando de una iteración a otra la diferencia en la estimación del conocimiento del usuario se estabilice. Se supone una habilidad inicial (que puede ser aleatoria en caso de no tener información alguna) y se va actualizando según la siguiente ecuación:

$$\hat{\theta}_{s+1} = \hat{\theta}_s + \frac{\sum_{i=1}^N -a_i [u_i - P_i(\hat{\theta}_s)]}{\sum_{i=1}^N a_i^2 P_i(\hat{\theta}_s) Q_i(\hat{\theta}_s)}$$

donde:

- $\hat{\theta}_s$  es la estimación de la habilidad en la iteración  $s$ .
- $a_i$  es el parámetro de discriminación del ítem  $i$ .
- $u_i$  es la respuesta del estudiante al ítem  $i$  (1 si se ha respondido correctamente ó 0 en caso contrario).
- $P_i(\hat{\theta}_s)$  es la probabilidad de responder correctamente al ítem  $i$  con habilidad  $\hat{\theta}_s$ .
- $Q_i(\hat{\theta}_s)$  es  $1 - P_i(\hat{\theta}_s)$ , es decir, la probabilidad de fallar el ítem  $i$  con habilidad  $\hat{\theta}_s$ .

Esta es una de las formas de representar la habilidad de un alumno, pero también existen otras. Al estar basados el modelo en probabilidades, uno nunca tiene la certeza de que el alumno tenga ese nivel de habilidad. Lo que se está calculando es la habilidad más probable que se tiene. Pero ello también puede ser representado por una curva que indique, por cada nivel de habilidad, la probabilidad de que sea esa la real.

Otra función importante es la llamada *función de información del ítem*, que aporta datos sobre la estimación que puede realizar de la habilidad del alumno. Por ejemplo, si un ítem es muy fácil y un alumno con una habilidad muy alta lo responde correctamente, no nos aporta tanta información sobre el nivel específico del alumno como otro ítem con una dificultad cercana a la capacidad del mismo. En el caso del modelo de tres parámetros, esta función viene definida por la siguiente función:

$$I(\theta, a, b, c) = a^2 \frac{Q(\theta)}{P(\theta)} \left[ \frac{P(\theta) - c}{1 - c} \right]^2$$

donde:

- $a$  es el parámetro de discriminación del ítem.
- $P(\theta)$  es la probabilidad de responder correctamente al ítem teniendo una habilidad  $\theta$ .
- $Q(\theta)$  es la probabilidad de responder incorrectamente al ítem, es decir, es equivalente a  $1 - P(\theta)$ .

- $c$  es el parámetro que indica la asíntota inferior.

Por tanto, esta función nos muestra la idoneidad de presentar el ítem con parámetros  $a$ ,  $b$  y  $c$  a un alumno con una habilidad  $\theta$ . Mientras mayor sea el valor, mejor se adaptará al alumno y más información nos aportará en cuanto a la habilidad del mismo. En la Figura 2 se pueden observar dos funciones de información del ítem (en líneas sólidas) con respecto a sus curvas características (en líneas punteadas). Por tanto, a la hora de elegir un ítem para un alumno concreto no se deberá fijarse solo en la habilidad del mismo y en la dificultad del ítem.

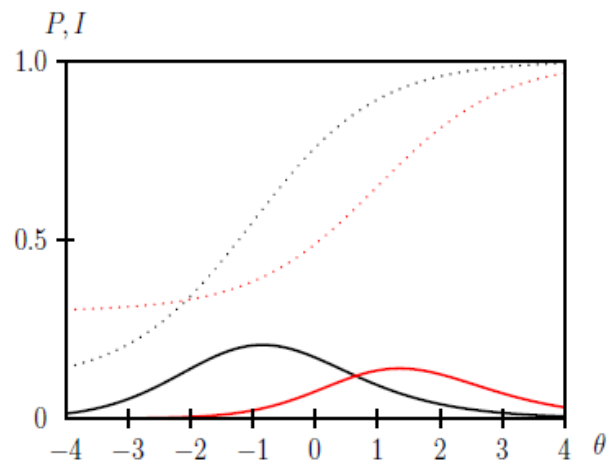


Figura 2: Ejemplos de dos funciones de información del ítem

#### 4.3.2. Modelado Multidimensional

Los procesos psicológicos en general son más complejos que lo que supone el modelado unidimensional. Raramente un razonamiento depende de una única habilidad, por lo que se han desarrollado los modelados multidimensionales. Este modelado debería llevar a mejorar la descripción de la interacción entre los alumnos y los ítems que realiza el modelado unidimensional.

La IRT multidimensional [1, 12, 16, 17, 18] es una metodología prometedora para tratar con este tipo de complejidades.

En general se pueden definir dos modelos de IRT multidimensional. Estos modelos vienen definidos por la forma en la que se combina la información a partir de un vector de  $\theta$ -coordenadas para especificar las probabilidades de respuesta al ítem. Un tipo de modelo está basado en una combinación lineal de las  $\theta$ -coordenadas, que se utiliza con una *normal ogive* o forma logística para especificar la probabilidad de una respuesta. El problema de este enfoque es que



puede haber diferentes combinaciones de las  $\theta$ -coordenadas que den el mismo resultado, es decir, que no se pueden saber los valores de las diferentes coordenadas que dan como resultado una determinada probabilidad. Estos modelos se suelen llamar *modelos compensatorios*.

El segundo tipo de modelo separa las diferentes tareas cognitivas de un ítem en partes y utiliza un modelo unidimensional para cada una de ellas. La probabilidad de una respuesta correcta es el producto de las probabilidades de cada parte. En general estos modelos se llaman *modelos no compensatorios* o *parcialmente compensatorios*.

En general se suele utilizar los modelos compensatorios debido a que puede resultar muy complicado conseguir dividir un ítem de forma efectiva. Por tanto tendremos un ítem cuya curva característica es función de dos o más coordenadas. En el caso de dos habilidades, la curva puede ser parecida a la mostrada en la Figura 3.

Como se puede comprobar, para una probabilidad dada es imposible saber qué valor tiene cada coordenada ya que si cortamos la curva en un nivel determinado, el resultado es una línea curva, es decir, hay infinitas combinaciones de valores de las dos variables para los cuales existe la misma probabilidad. Para estimar correctamente cada una de las coordenadas debería estimarse alguna de ellas mediante ítems unidimensionales.

Un ejemplo de modelo compensatorio es el propuesto por Reckase en [11]. Ese modelo es una generalización directa del modelo logístico de tres parámetros unidimensional al caso donde el conocimiento del alumno se representa por un vector de parámetros en vez de un solo escalar:

$$P(U_{ij} = 1|a_i, d_i, c_i, \theta_j) = c_i + (1 - c_i) \frac{e^{(a'_i \theta_j + d_i)}}{1 + e^{(a'_i \theta_j + d_i)}}$$

donde:

- $P(U_{ij} = 1|a_i, d_i, c_i, \theta_j)$  es la probabilidad de una respuesta correcta para la persona  $j$  en el ítem  $i$ .
- $U_{ij}$  representa la respuesta de la persona  $j$  al ítem  $i$ .
- $a_i$  es el vector de parámetros que define el poder de discriminación del ítem.
- $c_i$  es la probabilidad de respuesta correcta al azar.
- $\theta_j$  es el vector de habilidades de la persona  $j$ .

En el modelo unidimensional, la dificultad de un ítem es proporcional al valor de  $\theta$  donde la curva característica tiene mayor inclinación, sin embargo generalizar eso para el caso multidimensional no es trivial. La inclinación de una superficie

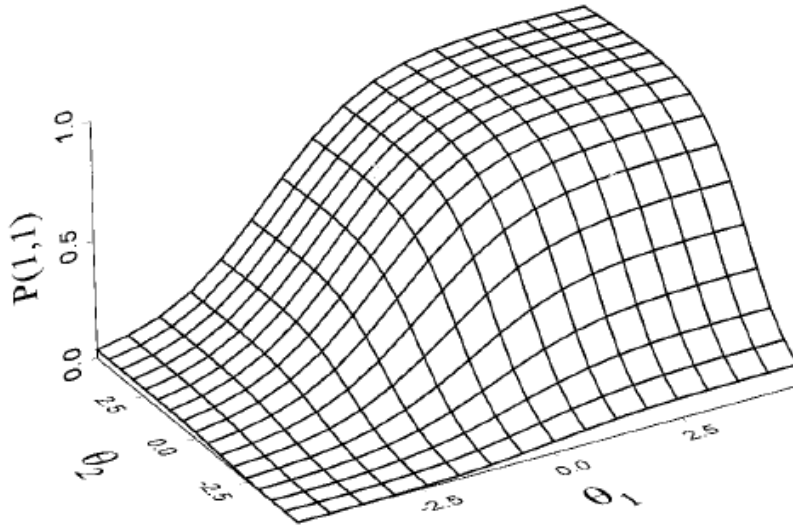


Figura 3: Curva característica de un ítem multidimensional

depende de la dirección del movimiento sobre ella con lo que el punto de mayor inclinación depende de la dirección que se considere. Sin embargo, en cada punto del  $\theta$ -espacio hay una dirección para la cual la inclinación es máxima. Si se considera todo el  $\theta$ -espacio y todas las direcciones, se puede obtener la máxima inclinación de toda la gráfica. Esto puede resultar útil para comprobar la capacidad de un ítem de distinguir las diferentes habilidades en dicho punto. También se puede utilizar para conocer la relación del vector de parámetros  $a_i$  con los valores de las pendientes de la curva en un punto determinado. Esto permite conocer la utilidad de un ítem para diferenciar entre las habilidades que tienen más peso.

El parámetro  $b$  en el caso unidimensional es un valor que indica si el punto de máxima inclinación está por debajo del 0 en el eje de habilidad (en este caso  $b$  sería negativo) o bien por encima ( $b$  positivo). Sin embargo, en el caso multidimensional, la proyección sobre el plano de  $\theta$ -coordenadas del punto de máxima pendiente de la curva no sólo viene definido por su distancia al origen de coordenadas (es decir, el punto donde las habilidades son 0), sino también por su dirección, que indicaría qué habilidad o característica es más importante en la respuesta del ítem. De esta forma, la dificultad viene determinada por un vector cuyo origen es el origen de coordenadas de las habilidades y su destino es la proyección del punto de máxima inclinación. Esto se puede representar en forma polar como un vector de ángulos (un valor para cada eje de coordenadas) y la distancia al origen. En el caso de un ítem bidimensional, la representación sería como la mostrada en la figura 4.

Para representar la dificultad genérica de un ítem multidimensional se utiliza la siguiente fórmula [17]:

$$B_i = \frac{-d_i}{\sqrt{\sum_{k=1}^m a_{ik}^2}}$$

donde:

- $m$  es el número de dimensiones.
- $a_{ik}$  es el valor del parámetro  $a$  en la dimensión  $k$ .
- $d_i$  es un valor escalar igual a  $a'_{ik}\xi$ , siendo  $\xi$  un vector de parámetros de dificultad del ítem.

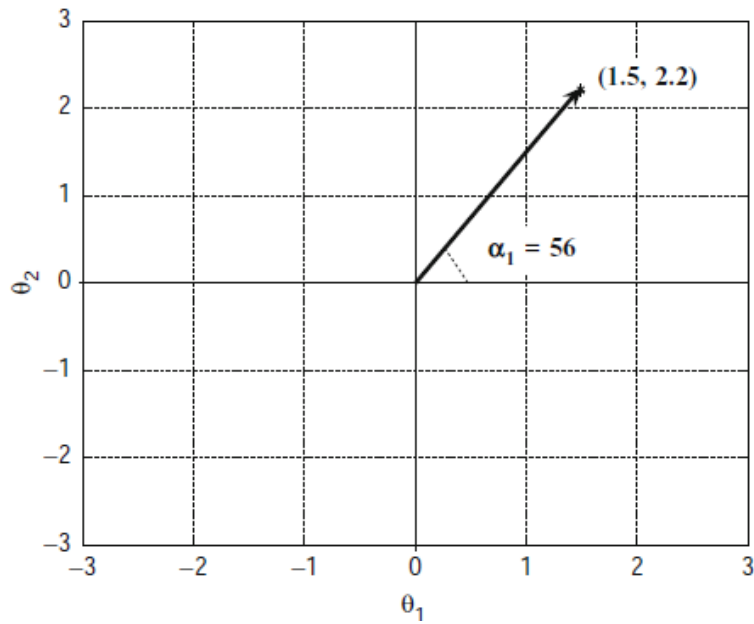


Figura 4: Representación de la dificultad de un ítem bidimensional

Análogamente también se puede definir el parámetro de discriminación en la IRT multidimensional. En el caso unidimensional, este parámetro indica la inclinación máxima de la curva característica del ítem. El equivalente en la multidimensional es la inclinación máxima de la superficie en la dirección desde el origen del  $\theta$ -espacio hasta el punto más cercano cuya probabilidad sea 0.5. Según [17], este parámetro viene definido por la siguiente expresión:

$$A_i = \sqrt{\sum_{k=1}^m a_{ik}^2}$$

Para poder comprender mejor estos conceptos, se presenta la Figura 5 donde se puede observar las líneas equiprobables de una superficie bidimensional. Si el ítem tiene una discriminación alta, significa que la inclinación de la curva es también alta, por lo que las líneas de probabilidad estarán más juntas entre sí, es decir, aunque se modifique muy poco la habilidad hay un cambio notable en la probabilidad de responder correctamente al ítem. Con respecto al parámetro de la dificultad, se puede observar que mientras más difícil sea el ítem, más separada estará la línea de probabilidad 0.5 con respecto al origen y con sentido hacia los valores positivos de las dos coordenadas. En caso de que el ítem sea muy fácil, esta distancia también será grande, pero el vector desde el origen hasta esta línea tendrá un sentido hacia los valores negativos de las características.

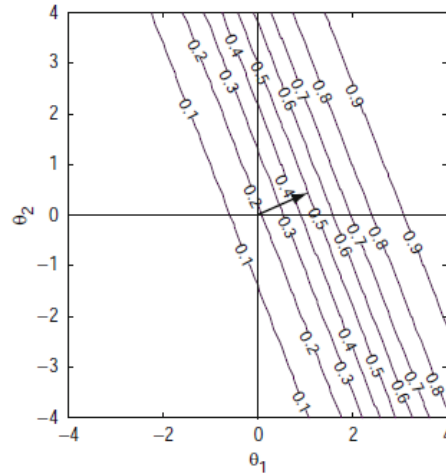


Figura 5: Representación de las características de un ítem multidimensional

En el caso de la *función de información del ítem* en el modelo multidimensional nos encontramos con el problema de que la inclinación en un punto concreto del espacio es variable dependiendo la dirección que se tenga en cuenta. En el modelo unidimensional, la función de información era dependiente del valor de la inclinación (parámetro  $a$ ). A esta dependencia hay que añadirle en este caso la dirección que se tome, por lo que la función en el caso multidimensional se generaliza de la siguiente forma:

$$I_{\alpha}(\theta) = \frac{[\nabla_{\alpha}P(\theta)]^2}{P(\theta)Q(\theta)}$$

donde:

- $\alpha$  es un vector de ángulos con respecto a los ejes de coordenadas que definen la dirección tomada desde el punto de características a considerar.
- $\nabla_{\alpha}$  es el gradiente en la dirección  $\alpha$

## 4.4. Estimación de parámetros

Existen diferentes métodos para la calibración de los ítems, es decir, la estimación de los parámetros de la curva característica del ítem. Hoy en día dos de los más utilizados son el de máxima verosimilitud (*Expectation-Maximization* o EM), que se centra en buscar aquellos parámetros de un determinado modelo que mejor se ajustan a los datos observados (es decir, con los que se comete menor error) y el suavizado de núcleo (*kernel smoothing*) que estima función a partir de datos observados cuando no se conoce ningún modelo para los mismos.

### 4.4.1. Expectation-Maximization (EM)

El algoritmo EM es un algoritmo iterativo donde cada iteración consiste en dos pasos: el paso de la expectación y el de la maximización. En el paso de la expectación se calculan los valores esperados de los datos que faltan (la probabilidad de acertar un ítem) a partir de la distribución condicional de los mismos dados los datos observados (las respuestas de los alumnos a los ítems) y el valor provisional de los parámetros. En el paso de maximización se calculan los parámetros que maximicen la función de máxima verosimilitud utilizando los datos estadísticos que se han obtenido en el paso de expectación, suponiendo conocidos los datos que queremos calcular. Estos pasos se repiten hasta que el algoritmo converja. Para más información sobre este algoritmo se puede consultar [9, 3].

### 4.4.2. Suavizado de núcleo

En general, para calibrar un ítem se suele tener en cuenta que la curva característica de un ítem es un modelo de dos o tres parámetros. Sin embargo, estos modelos paramétricos no siempre son capaces de estimar la función con suficiente precisión. Por ejemplo, el modelo de Rasch de tres parámetros explicado anteriormente es problemático cuando un ítem es muy sencillo ya que no hay suficientes datos para estimar el parámetro  $c$ . Esto hace que los cambios en el parámetro  $c$  sean compensados con cambios en la discriminación (parámetro  $a$ ), resultando en una mala estimación de este.

La idea básica del suavizado de núcleo [6, 20] es obtener una estimación no paramétrica de la curva característica del ítem. De esta forma, la función se puede adaptar a cualquier tipo de datos sin las restricciones que impone un modelo concreto. Para ello, el suavizado de núcleo se basa en medias locales de los datos existentes. La idea es la siguiente: si queremos calcular la imagen de un punto  $X_0$ , del cual no tenemos datos, tomamos una distancia  $\lambda$  alrededor de  $X_0$  y calculamos la media ponderada de las imágenes de los datos que existan en esa región. Mientras más lejos esté el punto de  $X_0$ , menor será su peso en la media. Para realizar la ponderación se pueden utilizar diferentes núcleos. Uno de ellos es el mostrado en la figura 6, donde tiene forma de campana.

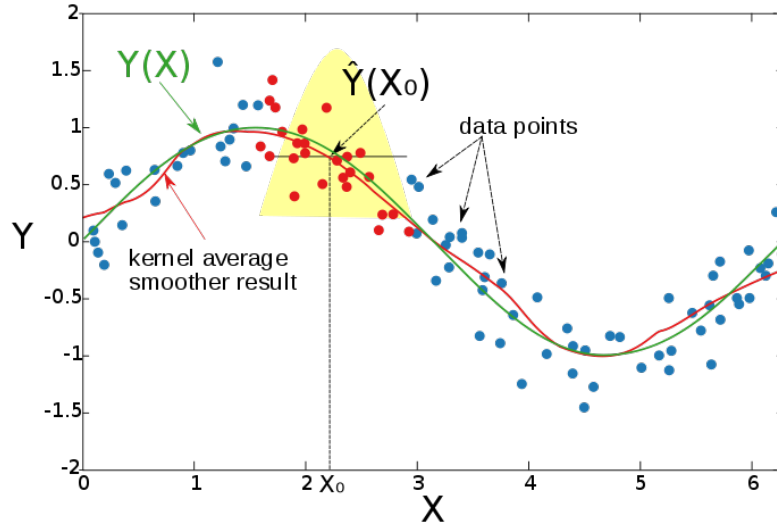


Figura 6: Ejemplo de suavizado de núcleo

## 5. Investigaciones iniciales

A continuación vamos a detallar el trabajo realizado hasta el momento, incluyendo tanto investigaciones teórica como prácticas. Los experimentos se llevaron a cabo en un colegio de infantil con alumnos de tres años.

### 5.1. Modelado de tareas

En el contexto del proyecto se tratan diferentes aspectos del currículum de infantil. Nosotros vamos a tratar el ámbito más genérico, donde se cubren temas como los colores, formas, series de elementos, conceptos como dentro-fuera, etc. Para cada uno de esos aspectos se han desarrollado una serie de actividades, las cuales están basada en una plantilla que define el tipo de ejercicio, el número de elementos, etc.

A partir de esa plantilla se genera el ejercicio concreto añadiéndole contenido. Este contenido es el que define los conceptos que se van a tratar, que normalmente será del mismo tema que se esté desarrollando en clase, como pudiera ser las estaciones del año, los números, las plantas, etc. Esto hace que se haya diseñado el sistema de tal forma que el profesor pueda generar los ejercicios de la forma más eficiente posible para poder adaptarse rápidamente a los cambios de temario que se producen en el currículum.

El proceso de creación de tareas se dividió en varias fases, equivalentes a las mencionadas en [19] y con características recomendadas en [14]. La primera consistió en el diseño de un primer prototipo con una cantidad limitada de ejercicios y tareas. Este prototipo fue llevado al aula y se recibieron consejos por parte de los profesores para cambiar determinada funcionalidad, además de ser probado por los alumnos. Se descubrieron las primeras deficiencias en cuanto a la interfaz que posteriormente se fueron subsanando. A continuación se fueron ampliando los ejercicios/tareas disponibles y refinándolos en sucesivas visitas al colegio y pruebas con los alumnos, de la misma forma que con los primeros prototipos. Hoy en día se está en contacto con psicólogos especializados en educación infantil para mejorar el proceso de aprendizaje.

La parte del proyecto que se va a desarrollar en esta investigación es la centrada en el ámbito más general. Existen dos formas de realizar los ejercicios. Por un lado tenemos el modo aprendizaje, donde el alumno recibe determinada información sobre si el ejercicio se está realizando de forma correcta o no. En este modo se tendrá en cuenta el *feedback* que reciba el alumno [4]. Por otro lado está el modo evaluación, donde el alumno realiza el ejercicio completo y no se muestra ningún tipo de información sobre la realización del mismo como por ejemplo pueda ser la puntuación, qué parte se ha realizado correctamente, etc.

A continuación mostraremos el tipo de problemas que se le presentan al alumno en este ámbito:

1) Ejercicios de dentro-fuera. En este tipo de ejercicios sólo se necesita especificar la imagen que va a tener el contenedor, los objetos de los que va a disponer el alumno y, para cada uno de ellos, si debe ir dentro del contenedor o fuera. Un ejemplo de este tipo de problema se puede observar en la Figura 7.

Cada objeto debe estar marcado con uno o más conceptos para poder actualizar el modelo de usuario correctamente. En este problema la IRT se aplica en dos partes: por un lado tenemos el concepto dentro-fuera que se evaluará dependiendo si el alumno es capaz de trasladar los objetos al contenedor (el carrito en el caso del ejemplo), independientemente de que los objetos sean los correctos. A continuación se habrá de evaluar si el objeto que se ha introducido es o no correcto. En este caso cada objeto se considerará un ítem, al cual se responderá correctamente si el objeto acaba en su posición objetivo (fuera o dentro del carrito) al finalizar el ejercicio.

2) Puzles, donde el alumno tendrá que formar una figura a partir de sus piezas. En este ejercicio simplemente se debe especificar el número de filas y columnas que debe tener el puzle y la imagen a partir de la cual se van a crear los trozos. El problema se puede ver en la Figura 8. En este caso, el problema entero se tratará como un ítem, ya que de otra forma los ítems no serían independientes entre sí.

3) Series. En este tipo de problema el alumno debe arrastrar diferentes figuras a sus correspondientes posiciones en la serie. Nosotros llamamos serie a un conjunto de elementos que están relacionados entre sí y con un orden determinado.



Figura 7: Ejemplo de un problema de tipo dentro-fuera

En este problema lo que se ha de especificar por parte del docente es el patrón de ejemplo, formado por dos objetos, que a su vez condicionará el orden que tengan que seguir el resto de piezas, y las piezas con las que puede interaccionar el alumno. Este tipo de problemas ofrece una amplia gama de posibilidades al profesor, ya que los objetos pueden ser de cualquier tipo y adaptarse al tema que en ese momento se esté impartiendo. En la Figura 9 se puede observar una serie en la que el alumno debe ir alternando las formas, en este caso círculo y cuadrado.

En este problema también se modelará cada movimiento que se tiene que hacer como un ítem de la IRT. En el ejemplo, lo primero que se debe hacer es seleccionar un círculo (con probabilidad al azar de  $4/12$ ) y colocarlo en la primera posición. A continuación se deberá escoger un cuadrado (con probabilidad al azar de  $4/11$ ) y ponerlo en la segunda, etc.

En este caso también se puede realizar el ejercicio en modo evaluación o aprendizaje, con lo que la actualización del modelo de usuario será diferente, al igual que en el ejercicio dentro-fuera.

4) Vocabulario. En este problema se presentan una serie de objetos y el sistema reproduce un sonido con el nombre de uno de ellos. El alumno deberá seleccionar el objeto que ha sido especificado por el sistema. Cuando lo seleccione, se reproducirá otro sonido y así sucesivamente hasta que el alumno haya seleccionado todos los objetos que se definieron en el problema. Lo único que se necesita



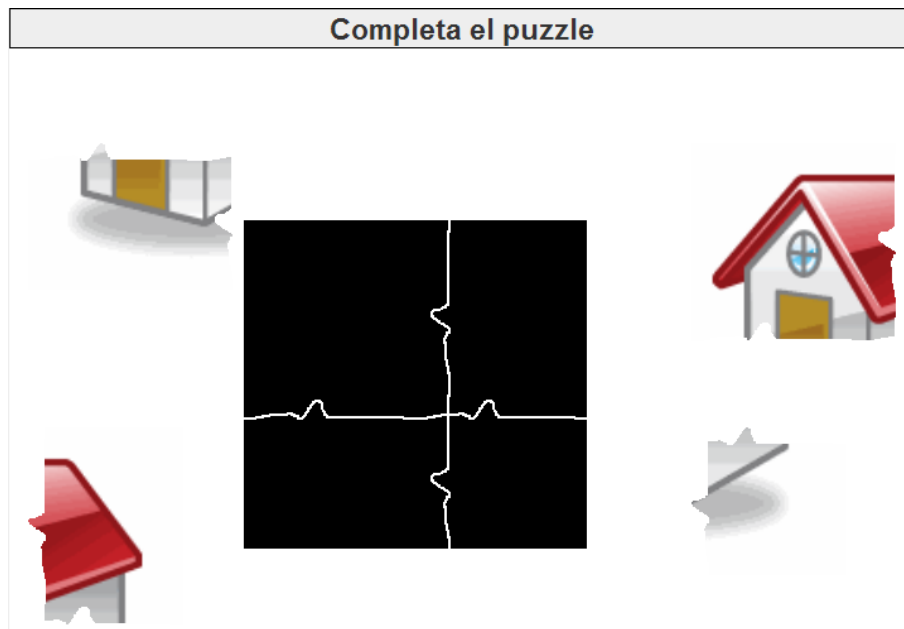


Figura 8: Ejemplo de puzle

para crear un problema de este tipo son los objetos que se quieren mostrar y el sonido asociado. El aspecto de este ejercicio es el mostrado en la Figura 10.

Cada uno de los objetos de cada ejercicio tiene asociado uno o más conceptos del currículum de educación infantil, tal y como se ha comentado anteriormente, para poder actualizarlo en el modelo de usuario cuando se complete el mismo. De esta forma podemos tener en todo momento una estimación del conocimiento del alumno para poder tomar las acciones necesarias.

En el cuadro se puede observar un resumen de los ejercicios implementados.

Tarea	Descripción	Objetivo
Dentro-Fuera	Se deben seleccionar los objetos de las características especificadas y arrastrarlos a un contenedor concreto	Trabajar el concepto dentro-fuera junto con los conceptos asociados a los objetos que seleccione el profesor para completar el ejercicio
Puzle	Se debe completar una imagen a partir de las piezas que se le muestran al alumno	Desarrollar la capacidad de generar una imagen mental completa a partir de piezas

Series	El alumno debe completar la serie teniendo en cuenta el patrón de ejemplo y los objetos que tiene disponibles	Comprender el concepto de serie y aquellos asociados a los objetos que se hayan utilizado de patrón
Vocabulario	El alumno debe seleccionar los dibujos que aparecen en pantalla según asociados al sonido que escuche	Ampliar el vocabulario del alumno con las palabras que decida el docente

## 5.2. Modelado de usuario

Para el modelado de usuario se ha tenido que contar con el apoyo de expertos en la materia que nos asesorasen sobre qué habilidades debe tener cada alumno a una determinada edad, cuáles son las características en las que hay que fijarse para comprobar si un ejercicio está bien resuelto o no, cuáles pueden demostrar que existen determinados problemas en el aprendizaje del niño para poder hacer hincapié en ellas, etc.

Para realizar el diseño del modelo de usuario se ha dividido el currículum de educación infantil en temas y conceptos. Cada uno de esos conceptos pertenecerán a un determinado tema y a su vez podrán tener subconceptos (por ejemplo, el concepto color podrá tener como subconceptos el rojo, verde, etc.). Estas relaciones se consideran como "parte de", es decir, que el conocimiento de un nodo será parte del conocimiento de su padre.

Cada vez que el alumno realiza un ejercicio, toda su interacción con el sistema se guarda en un archivo de log. Esto va a servir para realizar la corrección y actualizar el modelo de usuario.

Debido a que la mayoría de ejercicios que realizan los niños son multidimensionales dada la forma de aprendizaje de los mismos, se ha tenido que adaptar el modelo de usuario para que contemple esta posibilidad. En principio se va a utilizar la IRT multidimensional para modelar el conocimiento del alumno.

Un ejemplo simplificado de modelo de usuario se puede ver en la Figura 11. El currículum está dividido en dos conceptos, donde el primer concepto está formado por tres conceptos y el segundo por dos.

La inferencia de las habilidades se realiza a nivel de hojas [8], que son los conceptos directamente medibles mediante ítems. El resto de conceptos deben ser calculados a partir de sus hijos. Por ejemplo, si se realiza un test de todo el currículum, los conceptos de los que se tendrá información real serán el Rojo, Verde, Azul, Triángulo y Cuadrado. Para obtener información sobre el color en conjunto habrá que utilizar los ítems relativos a los conceptos hijos. Para ello

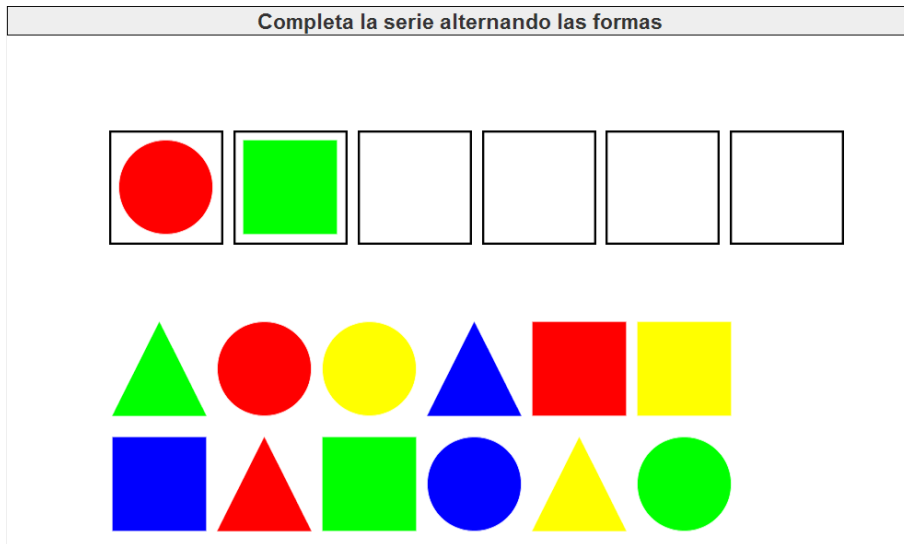


Figura 9: Ejemplo de serie

se administra un test que sólo contenga ítems que evalúen ese concepto, tanto directa como indirectamente. Las respuestas vendrán dadas por el nivel de conocimiento que haya demostrado el alumno en los test reales que haya realizado. Este proceso se repite hasta que se hayan calculado todos los nodos.

### 5.2.1. Actualización del Modelo de Usuario

En la práctica no se trabaja con un modelo de IRT continuo tal y como se ha expuesto anteriormente sino que se trabaja con un modelo discreto para facilitar y agilizar los cálculos. Nosotros dividiremos el intervalo  $[-3, 3]$  en 12 segmentos, tanto para la curva característica del ítem como para la estimación del conocimiento del usuario.

Para actualizar cada una de las habilidades, se deben tener en cuenta todos los ítems respondidos. El resultado del cálculo será la probabilidad de que el alumno pertenezca a cada una de las categorías. Para realizar este proceso, se calculan cada una de las probabilidades por separado según la siguiente ecuación:

$$C(\theta_x) = \frac{\prod_{i=1}^n (P_i(\theta_x) \cdot U_i + Q_i(\theta_x)(1 - U_i))}{\sum_{j=1}^m \left( \prod_{i=1}^n (P_i(\theta_j) \cdot U_i + Q_i(\theta_j)(1 - U_i)) \right)}$$

donde:

**Marca las palabras en el orden que indica el sonido**











Figura 10: Ejemplo de problema de vocabulario

- $C(\theta_x)$  es la probabilidad de que el alumno pertenezca a la categoría  $x$ .
- $P_i(\theta_x)$  es la probabilidad de que el alumno responda correctamente al ítem  $i$  perteneciendo a la categoría  $x$ .
- $Q_i(\theta_x)$  es la probabilidad de responder incorrectamente al ítem  $i$  perteneciendo a la categoría  $x$ .
- $U_i$  es 1 si el alumno ha respondido correctamente al ítem  $i$  y 0 en caso contrario.
- $m$  es el número total de categorías.
- $n$  es el número total de ítems.

Por tanto, por cada alumno y cada concepto se deberá guardar una serie de valores que corresponderán a las probabilidades de que el alumno tenga un conocimiento determinado de ese concepto. En nuestro caso, si se ha elegido dividir el intervalo  $[-3, 3]$  en 12 subintervalos, se necesitarán 12 valores por concepto.

En caso de que el ítem sea multidimensional el proceso a seguir es el mismo con la excepción de que la curva característica del ítem será una estructura de datos  $n$ -dimensional ( $n$  es el número de habilidades a tener en cuenta), donde los

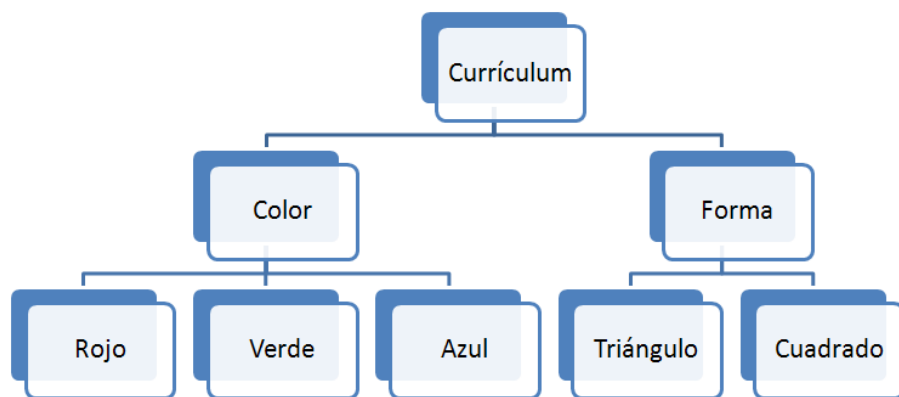


Figura 11: Ejemplo de currículum

valores de cada celda será la probabilidad de responder correctamente al ítem dados los niveles de habilidad concretos.

El problema radica en que al contestar correctamente una pregunta multidimensional no se puede actualizar individualmente cada uno de las habilidades, ya que no se puede conocer a priori cuál es la habilidad que más ha influido en la respuesta (por ejemplo, en el caso de que aparezcan dos colores, rojo y azul, y se le pida al alumno que seleccione el rojo, bastaría con que tenga un conocimiento del azul para, por descarte, saber cuál es el rojo, o simplemente podría conocer el rojo). En caso de que tengamos una estimación de alguna de ellas, se puede utilizar para calcular la otra en el caso bidimensional. Para realizar este proceso en el caso  $n$ -dimensional, habría que tener una estimación de  $n-1$  habilidades. Por tanto, si no se tienen ítems unidimensionales, no se podrá asignar a una habilidad o característica del modelo de usuario un valor concreto. En este caso habrá que guardar los datos necesarios que nos permitan conocer la probabilidad de responder correctamente a una pregunta dada cada una de las habilidades (por ejemplo, en el caso de que haya dos dimensiones sería una matriz). Esos datos irán asociados a todas las dimensiones implicadas, en vez de a una única habilidad, como se haría si el ítem fuese unidimensional. Conforme se vayan realizando diferentes ítem multidimensionales, al combinar sus diferentes curvas teniendo en cuenta las respuestas del alumno, se va formando una zona de mayor probabilidad en la que se encuentre el alumno, con lo que al final, teóricamente, podremos tener una aproximación de cada una de las habilidades individualmente, pero se cometería un error mayor que en el caso de que los ítem fueran unidimensionales.

## 6. Trabajo en el aula

Para trabajar con niños en una edad tan temprana se necesita un cierto conocimiento sobre su forma de trabajar (por ejemplo, no deben estar más de 20 minutos con la misma tarea), qué les resulta atractivo, su forma de aprendizaje, etc., como se ha comentado anteriormente. Para ello en [14] se explican ciertos conceptos básicos del comportamiento de los niños, se exponen diferentes ideas de cómo debe ser un sistema adaptado a ellos, su ciclo de desarrollo, el sistema de evaluación, cómo tratar los datos recopilados, métodos de observación, etc. Concretamente se describe el modelo PLU (*Players-Learners-Users*), que consiste en tratar a los alumnos como jugadores, estudiantes y usuarios. Como jugadores porque el niño se debe divertir al interactuar con el sistema. Como estudiantes porque el sistema debe, en cierta medida, sustituir al profesor, es decir, debe ser capaz de enseñar. Y como usuarios porque el sistema debe ser suficientemente intuitivo para que los niños no tengan dificultades al usarlo.

Hasta la actualidad se han ido realizando varios experimentos en el aula con alumnos de diferente edad y distintas aplicaciones. Los objetivos principales han sido conocer de primera mano cómo se trabaja en educación infantil, refinar la interfaz y hacerla más atractiva a los alumnos, refinar el software para que la evaluación se pareciera lo más posible a la que haría un profesor en esa situación y obtener datos de ejercicios para poder aplicar las técnicas aquí presentadas.

Los ejercicios se pueden realizar de forma individual en el tabletPC o colectiva con una pizarra táctil. La forma colectiva fomenta la comunicación entre los alumnos para resolver los problemas, lo que hace que también se desarrollen habilidades sociales. Sin embargo la forma individual está más indicada para el caso de la evaluación, donde el alumno debe demostrar por sus propios medios el conocimiento en unos determinados conceptos.

En el programa APRENDO, se realizaron los ejercicios expuestos con anterioridad: puzzles, series, vocabulario, dentro-fuera, etc. Para el experimento se utilizaron las tres clases de tres años que hay en el colegio, siendo una de ellas la clase de control. A los alumnos de las clases de control se les propuso realizar los ejercicios de aprendizaje con fichas plastificadas, de modo que simularan el mismo ejercicio que en el tabletPC. Para los ejercicios de evaluación del conocimiento se utilizó el tabletPC para todas las clases y que de esta forma no hubiera diferencia en la obtención de la estimación de las habilidades. Las series de ejercicios que se realizaron constaban de tres etapas:

- Ejercicios de evaluación, donde se comprobaba el conocimiento inicial del alumno. En estas pruebas el alumno no contaba con ningún tipo de ayuda y podía finalizar el ejercicio cuando considerase oportuno.
- Ejercicios de aprendizaje, donde los alumnos pueden recibir ayuda por parte de un adulto que le pueda indicar por qué ha cometido un error, o le dé pistas sobre la solución en caso de que éste no sepa continuar por sus propios medios.

- Ejercicios de evaluación, donde se comprobaba el progreso que había tenido el alumno en la fase de aprendizaje.

Al realizar todos estos ejercicios se guarda automáticamente un fichero de *log* que se utiliza para comprobar la interacción del alumno con el sistema, de tal modo que se puede conocer paso a paso qué acción ha realizado el alumno y en qué momento. A partir de estos *logs* se debe actualizar el modelo de usuario acorde a la estimación realizada.

En cuanto a recopilación de datos para estimar el conocimiento de los alumnos, sólo se ha utilizado el ejercicio de las series para el experimento realizado debido a restricciones de horarios del colegio, cantidad de alumnos, etc., no siendo así para la recopilación de datos con el objetivo de mejorar el sistema, donde se han utilizado todos los tipos. En la figura 12 pueden verse a dos alumnos resolviendo un puzle. Los ejercicios propuestos fueron los siguientes:

1. Elegir todas las figuras azules (evaluación)
2. Elegir todos los triángulos (evaluación)
3. Elegir todas las figuras rojas (aprendizaje)
4. Elegir todos los cuadrados (aprendizaje)
5. Series de dos elementos: primero cuadrado y luego un triángulo (evaluación)
6. Series de dos elementos: primero verde y después azul (evaluación)
7. Series de dos elementos: primero círculo y luego cuadrado (aprendizaje)
8. Series de dos elementos: primero rojo y después amarillo (aprendizaje)
9. Series de dos elementos: primero círculo y luego cuadrado (evaluación)
10. Series de dos elementos: primero verde y después azul (evaluación)

Entre el ejercicio 8 y 9 hubo dos semanas en los que los alumnos no trabajaron con el sistema, con lo que se trataba de observar si recordaban lo que habían aprendido. Hay que destacar que de los ejercicios de aprendizaje también se puede sacar cierta información ya que se puede comprobar el tiempo que tarda un alumno en finalizar la tarea, lo que puede demostrar un alto grado de conocimiento en caso de que se realice rápidamente.

Los resultados todavía están siendo estudiados.



Figura 12: Alumnos realizando el ejercicio del puzle

## 7. Conclusiones y plan de trabajo

Se ha podido comprobar que es factible introducir las tecnologías de la información en el aula como una herramienta adicional a las tradicionales. No sólo no interrumpen el correcto funcionamiento de la clase sino que lo hace más dinámico y ameno, característica que es bastante importante y apreciada por los alumnos.

El proceso de modelado del alumno y el de tareas es complejo, especialmente en educación infantil, que requiere la ayuda de especialistas (profesores de educación infantil, psicólogos, etc.) ya que el proceso de aprendizaje de estos alumnos es muy particular a la par que importante.

Existen tecnologías disponibles para conseguir implementar con éxito un sistema de este tipo, aunque habrá que introducir mejoras para la aplicación a las nuevas necesidades.

A continuación describimos las mejoras que se van a realizar sobre el sistema que se tiene en la actualidad.

- Incorporar un nuevo sistema de interacción con el alumno basado en objetos reales, donde cada uno de ellos tiene unas marcas específicas para



realizar su seguimiento mediante procesamiento de imágenes. Este sistema estará basado en el propuesto por [13].

- Aumentar la cantidad de tipos de ejercicios según las necesidades del currículum infantil.
- Implementar el modelo teórico de la IRT multidimensional para actualizar el modelo de usuario. En esta parte hay que investigar sobre la estimación de parámetros o habilidades individuales a partir de ítems multidimensionales. Además habrá que estudiar la mejora que aporta el enfoque multidimensional con respecto al unidimensional.
- Creación de un entorno para la edición de tareas de forma cómoda para los profesores. De esta forma se conseguirán crear los ejercicios más rápidamente y por tanto los alumnos dispondrán de una mayor cantidad de material. Esto es esencial ya que mientras más material se cree, mejor se podrá seleccionar el adecuado al nivel de cada alumno.
- Explorar diferentes modelos y modos de diagnóstico de dificultades en los alumnos. Muchos de estos problemas pueden venir dados por causas que se pueden corregir si se detectan a tiempo, por lo que es esencial este aspecto para el correcto desarrollo del alumno.
- Crear un módulo de asignación de ejercicios que sea capaz de seleccionar aquellos que mejor se adaptan al nivel de cada uno de los alumnos y, de esta forma, agilizar el proceso de aprendizaje. Para ello se utilizará la función de información de cada uno de los ítems o ejercicios que se encuentren disponibles.
- Elaborar un sistema de generación automática de ejercicios que sea capaz de crearlos de forma eficiente a partir de los objetos.
- Estudio de la aportación de las redes bayesianas al modelo de usuario.

## Referencias

- [1] Tamás Antal. On multidimensional item response theory – a coordinate free approach. *Electron. J. Stat.*, 1:290–306, Jul 2007.
- [2] Peter Brusilovsky and Eva Millán. User models for adaptive hypermedia and adaptive educational systems. In Peter Brusilovsky, Alfred Kobsa, and Wolfgang Nejdl, editors, *The Adaptive Web*, volume 4321 of *Lecture Notes in Computer Science*, chapter 1, pages 3–53. Springer Berlin Heidelberg, Berlin, Heidelberg, 2007.
- [3] Michael Collins. The em algorithm. In *In fulfillment of Written Preliminary Exam II requirement*, 1997.
- [4] E.; Pérez-de-la-Cruz J.L. Conejo, R.; Guzmán. Towards a computational theory of learning in an adaptive testing environment. In *Proceedings of AIED'03*, 2003.
- [5] Aaron D'Souza, Jeff Rickel, Bruno Herreros, and W. Lewis Johnson. An automated lab instructor for simulated science experiments. In *In: Proceedings of AIED 2001, 10th International Conference on Artificial Intelligence in Education*, pages 65–76. IOS Press, 2001.
- [6] Pere J. Ferrando. Kernel-smoothing estimation of item characteristic functions for continuous personality items: An empirical comparison with the linear and the continuous-response models. *Applied Psychological Measurement*, 28(2):95–109, 2004.
- [7] Eduardo Guzmán and Ricardo Conejo. Towards efficient item calibration in adaptive testing. In *User Modeling*, pages 402–406, 2005.
- [8] Eduardo Guzmán, Ricardo Conejo, and José-Luis Pérez-de-la Cruz. Adaptive testing for hierarchical student models. *User Modeling and User-Adapted Interaction*, 17(1-2):119–157, 2007.
- [9] B Hanson. Irt parameter estimation using the em algorithm. 2000.
- [10] Essam M. Kosba, Vania Dimitrova, and Roger D. Boyle. The evaluation of an intelligent teacher advisor for web distance environments. In *AIED*, pages 370–377, 2005.
- [11] Ronald K. Linden, Wim J. van der; Hambleton. *Handbook of Modern Item Response Theory*, chapter 16, pages 271–286. Springer, 1997.
- [12] Jingyu Liu. *Comparing multidimensional and unidimensional computer adaptive strategies in psychological and health assessment*. PhD thesis, Graduate Faculty of School of education, 2007.

- [13] Javier Marco, Sandra Baldassarri, and Eva Cerezo. Bridging the gap between children and tabletop designers. In *IDC '10: Proceedings of the 9th International Conference on Interaction Design and Children*, pages 98–107, New York, NY, USA, 2010. ACM.
- [14] Panos Markopoulos, Janet C Read, Stuart MacFarlane, and Johanna Hoysiemi. *Evaluating Children's Interactive Products: Principles and Practices for Interaction Designers (Interactive Technologies)*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2008.
- [15] Ivailo Partchev. A visual guide to item response theory. 2004.
- [16] Mark D. Reckase. The past and future of multidimensional item response theory. *Applied Psychological Measurement*, 21(1):25–36, 1997.
- [17] Mark D. Reckase. *Multidimensional item response theory*. Springer, 2009.
- [18] Steven G. Spencer. *The strength of multidimensional item response theory in exploring construct space that is multidimensional and correlated*. PhD thesis, Brigham Young University, 2004.
- [19] Nikolaos K. Tselios, Nikolaos M. Avouris, and Maria Kordaki. Student task modeling in design and evaluation of open problem-solving environments. *Education and Information Technologies*, 7(1):17–40, 2002.
- [20] Yinggan Zheng, Mark J. Gierl, and Ying Cui. Using cochrans z statistic to test the kernel-smoothed item response function differences between focal and reference groups. *Educational and Psychological Measurement*, 70(4):541–556, 2010.