

SIETTE: Sistema Inteligente de Evaluación mediante Test para TeleEducación



<http://www.lcc.uma.es/SIETTE>



Departamento de Lenguajes y Ciencias de la Computación.
Universidad de Málaga, SPAIN.

- ◆ 1. Introducción
- ◆ 2. Arquitectura del sistema SIETTE
- ◆ 3. Introducción a la Teoría de Respuesta al Item
- ◆ 4. Introducción a la teoría de Test Adaptativos Informatizados
- ◆ 5. Estudio simulado del comportamiento teórico
- ◆ 6. Estimación de las curvas características
- ◆ 7. Tipos de ítems
- ◆ 8. Integración con sistemas tutores inteligentes.
- ◆ 9. Conclusiones

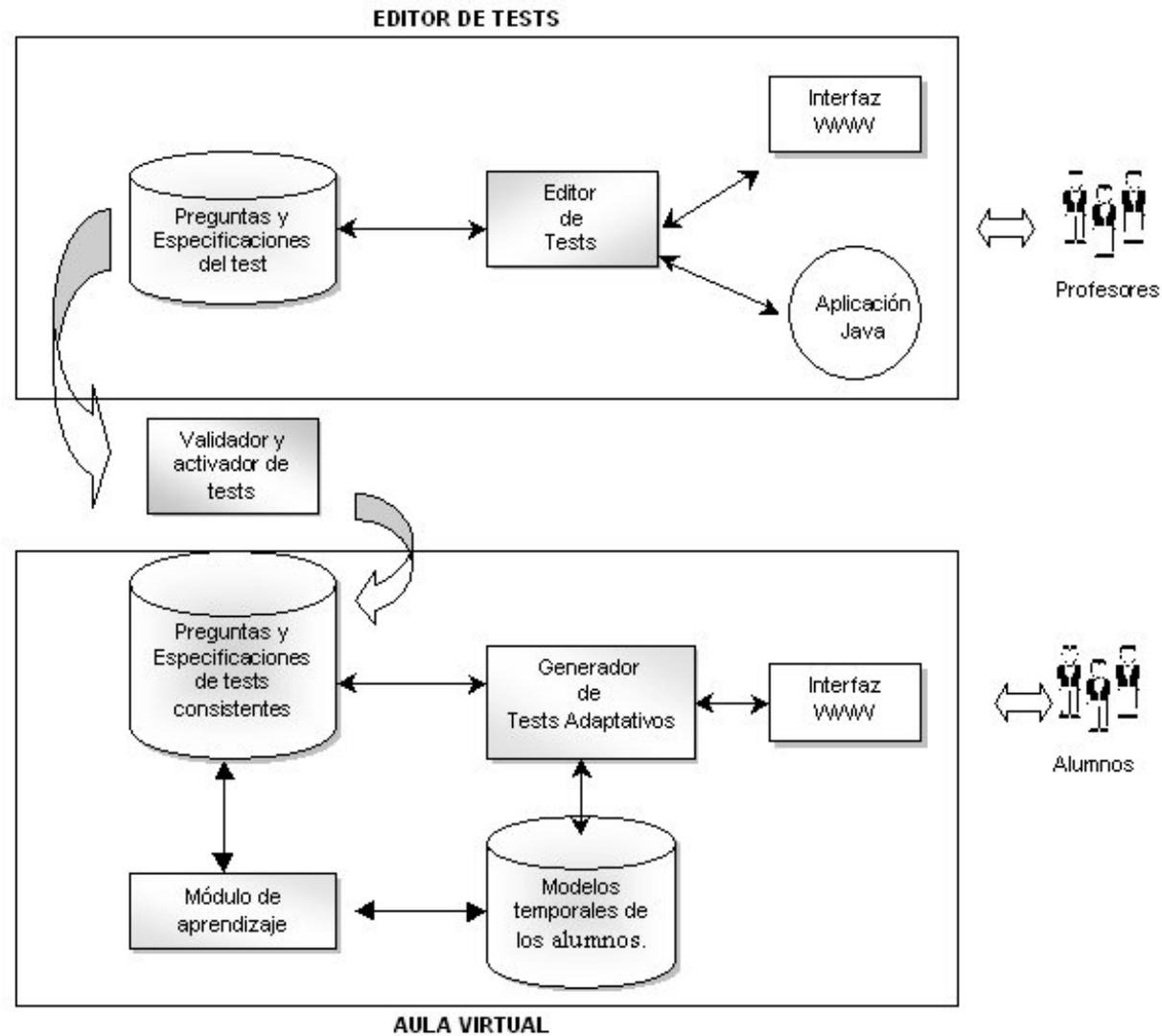
1. INTRODUCCIÓN

- ◆ Los tests son un mecanismo de evaluación muy usado en aplicaciones de enseñanza asistida por ordenador, y en sistemas tutores inteligentes.

- ◆ En el campo de la Psicometría se han estudiado los tests de forma rigurosa.
 - ☐ Toría Clasica de los Test
 - ☐ Teoría de Respuesta al Ítem
 - ☐ Teoría de Test Adaptativos Informatizados

- ◆ Tests sobre la WWW. (Brusilovsky 1999)
 - ☐ No requiere instalación
 - ☐ Actualización inmediata
 - ☐ En cualquier momento, en cualquier lugar
 - ☐ Mejora mediante aprendizaje automático

2. ARQUITECTURA DEL SISTEMA SIETTE



2.1. El generador de tests

MODOS DE USO DE SIETTE

- ◆ Como sistema autónomo
 - ☐ Identificación de alumno
 - * Usuario registrado
 - * Usuario anónimo
 - ☐ Selección y configuración del test del test
 - * Selección del test
 - * Idioma
 - * Mostrar evaluación: paso a paso /al final
 - * Ver respuestas: paso a paso /al final
 - ☐ Resolución de ítems.
 - ☐ Presentación de resultados.

- ◆ Como componente de un Sistema Tutor Inteligente

2.1. El generador de tests (Resolución del test)



Ítem del generador de tests



Corrección de un ítem

2.1. El generador de tests (Resultados)

Resultados del test - Microsoft Internet Explorer

Archivo Edición Ver Favoritos Herramientas Ayuda

Atrás Adelante Detener Actualizar Inicio Búsqueda Favoritos Historial Correo Imprimir

Dirección <http://mizar.lcc.una.es/cgi-bin/siette/generaTest.cgi> Ir a Vinculos >>

SIETTE  **TREE**
Sistema Inteligente de Evaluación mediante Test
para la TeleEducación.

Resultados del test:

Número de preguntas realizadas	Número de preguntas correctas	Nivel de conocimiento estimado	Intervalo de confianza
20	12	7	(6.236372, 8.099411)

APROBADO

[Mostrar todas las soluciones](#)

[\[Inicio\]](#)

Distribucion del conocimiento



X	p(X)
Nivel 0	0.000213
Nivel 1	0.000234
Nivel 2	0.000321
Nivel 3	0.000634
Nivel 4	0.002414
Nivel 5	0.017745
Nivel 6	0.143699
Nivel 7	0.496932
Nivel 8	0.314435
Nivel 9	0.023194
Nivel 10	0.000180
Media	Desv. típica
7.167891	0.808762

Listo Internet

2.2. El editor de tests

◆ INTERFAZ WEB

- ☐ Identificación del profesor
 - * Identificador por asignaturas
- ☐ Creación de la base de conocimientos
 - * Tests
 - * Temas
 - * Ítems.
- ☐ Contenidos multimedia

◆ INTERFAZ JAVA

- ☐ Creación del curriculum
- ☐ Ítems politómicos
- ☐ Editor mejorado
- ☐ Funcionamiento autónomo

2.2. El editor de tests (Interfaz web)

The screenshot shows a Netscape browser window titled "Acceso al Editor de Tests - Netscape". The browser's address bar is empty, and the menu bar includes "File", "Edit", "View", "Go", "Communicator", and "Help". The toolbar contains icons for Back, Forward, Reload, Home, Search, Netscape, Print, Security, Shop, and Stop.

The main content area is titled "SIETTE Sistema Inteligente de Evaluación mediante Test para la TeleEducación." and features a "Datos de la Cuestión" header. Below this is a "Datos generales" section with the following fields:

- Título:
- Esquema de generación de preguntas:
- Control de la corrección mediante un applet:
- Grado de dificultad:
- Factor de discriminación:
- Número de alternativas a mostrar:
- Temas a los que pertenecerá:
 - 1. Introducción
 - 2. Analisis lexico
 - 2.1. LEX
 - 3. Analisis sintactico
 - 4. Analisis semántico

Below the "Datos generales" section is a "Datos relativos a la pregunta" section with two text areas:

- Enunciado:
- Ayuda:

At the bottom of the main content area is a "Datos relativos a las posibles respuestas" section. Below this is a "Modo de Visualización" section with radio buttons for "Lineal" (selected) and "Ejemplo".

The left sidebar contains a "Tests" section with links for "Añadir", "Modificar", and "Eliminar". Below it are "Temas" and "Cuestiones" sections, each with "Añadir", "Modificar", and "Eliminar" links. The "Cuestiones" section also has a "(Parámetros)" link. At the bottom of the sidebar is a "Multimedia" section with an "Añadir/Eliminar" link.

The browser's status bar at the bottom shows "Document: Done" and a taskbar with various system icons.

2.2. El editor de tests (Interfaz JAVA)

Gestor de tests para SIETTE

Archivo Edición Insertar Gestión Ayuda

Curriculum

- Lógica
 - Temas
 - Prolegómenos
 - Verdad y validez
 - Verdad - falsedad
 - Verdad-falsedad 2
 - Verdad-falsedad 3
 - Escarabajo 1: El duende 4 no tiene ningún caracol
 - Escarabajo 2: El duende 3 tiene un caracol
 - Escarabajo 3: El duende 1 tiene dos caracoles
 - Escarabajo 4: El duende 2 tiene más caracoles que el 3
 - Escarabajo 5: El duende 1 tiene menos caracoles que el 4
 - Escarabajo 6: El duende 4 tiene cuatro caracoles
 - Verdad-falsedad 4
 - Gusano 1: El duende 2 está entre dos flores
 - Escarabajo 2: El duende 3 esta al lado de una flor
 - Gusano 2: El duende 5 está al lado de un caracol
 - Escarabajo 1: El duende 4 está al lado de un caracol
 - Caracol 1: El duende 1 está al lado de una flor
 - Caracol 2: El duende 5 está entre dos flores.
 - Verdad-falsedad 5
 - Gusano 1: La seta 1 es comestible
 - Caracol 2: La seta 3 no es comestible
 - Caracol 1: La seta 5 es venenosa
 - Escarabajo 1: Hay más setas venenosas que comestibles
 - Gusano 2: Todas las setas son comestibles
 - Escarabajo 2: Todas las setas son venenosas
 - Multirespuesta
 - Los lenguajes logicos
 - Proposiciones
 - Concepto de proposición
 - Proposiciones atómicas y moleculares
 - Proposiciones atómicas y moleculares
 - Atómicas y moleculares 2
 - El árbol tiene frutos
 - El duende está al lado de una seta
 - Hay un duende y un caracol debajo del árbol
 - Si hay nubes, entonces no brilla el sol

Pregunta

Título: Verdad-falsedad 3

Enunciado: ¿cuáles de los siguientes escarabajos dicen la verdad?

Ayuda:

Applet

Ruta:

Visualizar Configuración Examinar

Dificultad: 1 Modo Visualización: Secuencial

Adivinanza: 1

Discriminación: 1.2

Organización

Filas: 6

Columnas: 1

Respuestas

- Escarabajo 1: El duende 4 no tiene ningún cara...
- Escarabajo 2: El duende 3 tiene un caracol
- Escarabajo 3: El duende 1 tiene dos caracoles

Multirespuesta

Añadir Respuesta Añadir Multirespuesta Aceptar Cancelar

2.3. Validación y activación de tests

- ◆ Comprobar que los tests son factibles
 - ☐ Número de preguntas suficientes
 - ☐ Disponibilidad temporal

- ◆ Comprobar que los ítems son correctos
 - ☐ Número de alternativas por pregunta
 - ☐ Disponibilidad de componentes multimedia

- ◆ Integración de múltiples editores

- ◆ Separar la edición de la realización

3. TEORÍA DE RESPUESTA AL ÍTEM

- ◆ El nivel de conocimiento de un alumno sobre una materia puede ser medido mediante una variable real θ que toma valores entre $(-\infty, +\infty)$.
 - ☐ El nivel de conocimientos se mantiene constante a lo largo del test

- ◆ Para cada ítem i del test (cuestión) se conoce la probabilidad condicionada de responder dado el nivel de conocimientos: $P(u_i=1 / \theta) = P_i(\theta)$
 - ☐ función del intervalo $(-\infty, +\infty)$ en el intervalo real $[0,1]$,
 - ☐ monótona creciente
 - ☐ Curva Característica del Item

- ◆ Un test esta compuesto por n ítems: (u_1, \dots, u_n)
 - ☐ Se suponen ítems dicotómicos (solo dos posibles respuestas: acierto o fallo).
 - ☐ La respuesta a cada ítem es independiente.

3. TEORÍA DE RESPUESTA AL ÍTEM

- ◆ Curvas características según la distribución normal

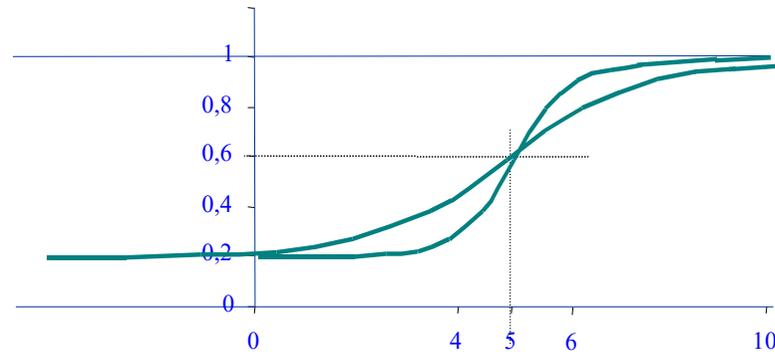
$$P_i(\theta) = P(u_i = 1 | \theta) = c_i + (1 - c_i) \frac{1}{a_i \sqrt{2\pi}} \int_{-\infty}^{\theta} e^{-\frac{(x-b_i)^2}{2a_i^2}} dx$$

- ◆ Curvas características según la distribución logística

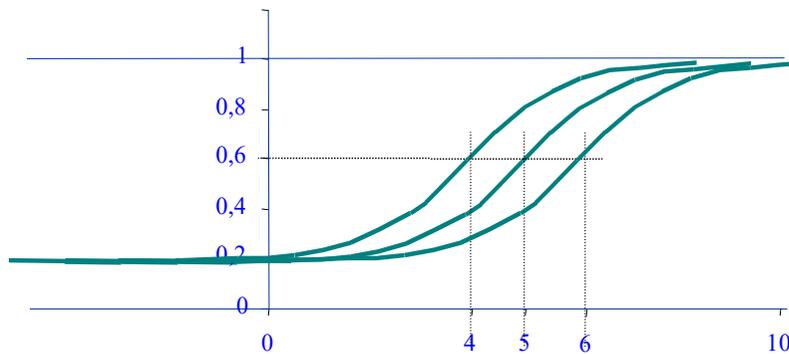
$$P_i(\theta) = P(u_i = 1 | \theta) = c_i + (1 - c_i) \frac{1}{1 + e^{-1.7a_i(\theta-b_i)}}$$

3. TEORÍA DE RESPUESTA AL ÍTEM

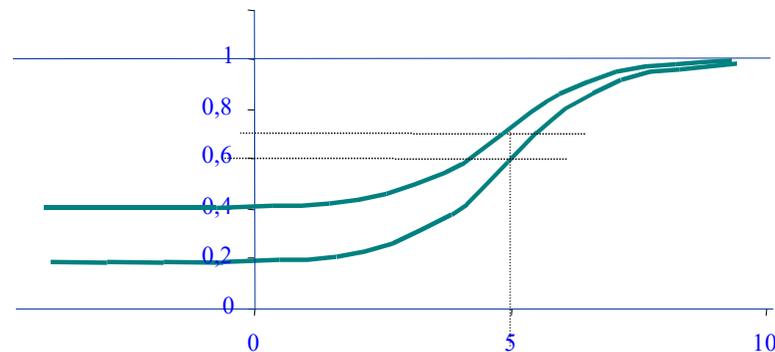
INTERPRETACIÓN DEL SIGNIFICADO DE LOS PARÁMETROS



$a = 0,6; a = 1,2$
 $b = 5$
 $c = 0,2$



$a = 1,2$
 $b = 4; b = 5; b = 6$
 $c = 0,2$



$a = 1,2$
 $b = 5;$
 $c = 0,2; c = 0,4$

3. TEORÍA DE RESPUESTA AL ÍTEM

ESTIMACIÓN DEL NIVEL DE CONOCIMIENTO DEL ALUMNO

- ◆ Método de la máxima verosimilitud

$$L(\mathbf{u} | \theta) = L(u_1 \dots u_n | \theta) = \prod_{i=1}^n P_i(\theta)^{u_i} (1 - P_i(\theta))^{(1-u_i)} \quad \left(\frac{\partial L(\mathbf{u} | \theta)}{\partial \theta} \right)_{\theta=\hat{\theta}} = 0$$

- ◆ Método bayesiano

$$P(\theta | \mathbf{u}) = P(\theta | u_1 \dots u_n) = \frac{\prod_{i=1}^n P_i(\theta)^{u_i} (1 - P_i(\theta))^{(1-u_i)}}{\prod_{i=1}^n P(u_i)^{u_i} (1 - P(u_i))^{(1-u_i)}} P(\theta)$$

3. Modelo de respuesta en SIETTE

DISCRETIZACION DE LA TEORIA DE RESPUESTA AL ÍTEM

- ◆ Probabilidad condicionada = Curvas características de los ítems

$$\mathbf{P}_i = \overline{P(u_i | \boldsymbol{\theta})} = (\Pr(u_i = 1 | \theta = 0), \Pr(u_i = 1 | \theta = 1), \dots, \Pr(u_i = 1 | \theta = K - 1))$$

$$\mathbf{Q}_i = \overline{Q(u_i | \boldsymbol{\theta})} = (1 - \Pr(u_i = 1 | \theta = 0), 1 - \Pr(u_i = 1 | \theta = 1), \dots, 1 - \Pr(u_i = 1 | \theta = K - 1))$$

- ◆ Probabilidad a priori = Estimación del nivel de conocimientos probable de un alumno

$$\mathbf{P} = \overline{P(\boldsymbol{\theta})} = (\Pr(\theta = 0), \Pr(\theta = 1), \dots, \Pr(\theta = K - 1))$$

- ◆ Probabilidad a posteriori (Aplicación de la regla de Bayes)

$$\overline{P(\boldsymbol{\theta} | \mathbf{u})} \propto \mathbf{P} \prod_{i=1}^n \mathbf{P}_i^{u_i} \mathbf{Q}_i^{1-u_i}$$

3. Modelo de respuesta en SIETTE

EJEMPLO 1: ESTIMACION DEL NIVEL DE CONOCIMIENTO

- Curvas características de los ítems

| | $\theta=0$ | $\theta=1$ | $\theta=2$ | $\theta=3$ |
|----------------|------------|------------|------------|------------|
| P ₁ | 0,1 | 0,3 | 0,7 | 0,9 |
| P ₂ | 0,5 | 0,6 | 0,9 | 1,0 |
| P ₃ | 0,3 | 0,6 | 0,8 | 0,9 |
| P ₄ | 0,3 | 0,4 | 0,7 | 0,9 |
| P ₅ | 0,1 | 0,2 | 0,3 | 0,9 |

- Vector de respuestas

$$(u_1=1, u_2=1, u_3=0, u_4=1, u_5=0)$$

- Distribución de la probabilidad a priori

$$(0,25 \quad 0,25 \quad 0,25 \quad 0,25)$$

- Cálculo de la probabilidad a posteriori

$$\begin{aligned}
 P(\theta | u_1 \dots u_5) &\propto \begin{pmatrix} 0,1 \\ 0,3 \\ 0,7 \\ 0,9 \end{pmatrix} \begin{pmatrix} 0,5 \\ 0,6 \\ 0,9 \\ 1,0 \end{pmatrix} \begin{pmatrix} 0,7 \\ 0,4 \\ 0,2 \\ 0,1 \end{pmatrix} \begin{pmatrix} 0,3 \\ 0,4 \\ 0,7 \\ 0,9 \end{pmatrix} \begin{pmatrix} 0,9 \\ 0,8 \\ 0,7 \\ 0,1 \end{pmatrix} (0,25 \quad 0,25 \quad 0,25 \quad 0,25) \propto \\
 &\propto (0,092 \quad 0,225 \quad 0,603 \quad 0,079)
 \end{aligned}$$

3. Modelo de respuesta en SIETTE

EJEMPLO 2: REDUCCION DE LA PRECISIÓN DEL CALCULO

- ◆ Curvas características de los ítems

| | $\theta=0$ | $\theta=1$ | $\theta=2$ | $\theta=3$ |
|----------------|------------|------------|------------|------------|
| P ₁ | 0,1 | 0,3 | 0,7 | 0,9 |
| P ₂ | 0,5 | 0,6 | 0,9 | 1,0 |
| P ₃ | 0,3 | 0,6 | 0,8 | 0,9 |
| P ₄ | 0,3 | 0,4 | 0,7 | 0,9 |
| P ₅ | 0,1 | 0,2 | 0,3 | 0,9 |

- ◆ Vector de respuestas

$$(u_1=1, u_2=1, u_3=0, u_4=1, u_5=0)$$

- ◆ Distribución de la probabilidad a priori

$$(0,5 \quad 0,5)$$

- ◆ Cálculo de la probabilidad a posteriori

$$\begin{aligned}
 P(\theta | u_1 \dots u_5) &\propto \begin{pmatrix} 0,20 \\ 0,80 \end{pmatrix} \begin{pmatrix} 0,55 \\ 0,95 \end{pmatrix} \begin{pmatrix} 0,65 \\ 0,15 \end{pmatrix} \begin{pmatrix} 0,35 \\ 0,80 \end{pmatrix} \begin{pmatrix} 0,85 \\ 0,40 \end{pmatrix} (0,5 \quad 0,5) \\
 &\propto (0,317 \quad 0,682)
 \end{aligned}$$

3. Modelo de respuesta en SIETTE

CURVAS CARACTERÍSTICAS PARAMÉTRICAS EN SIETTE

- ◆ A partir de los parámetros --> Hallar los vectores característicos

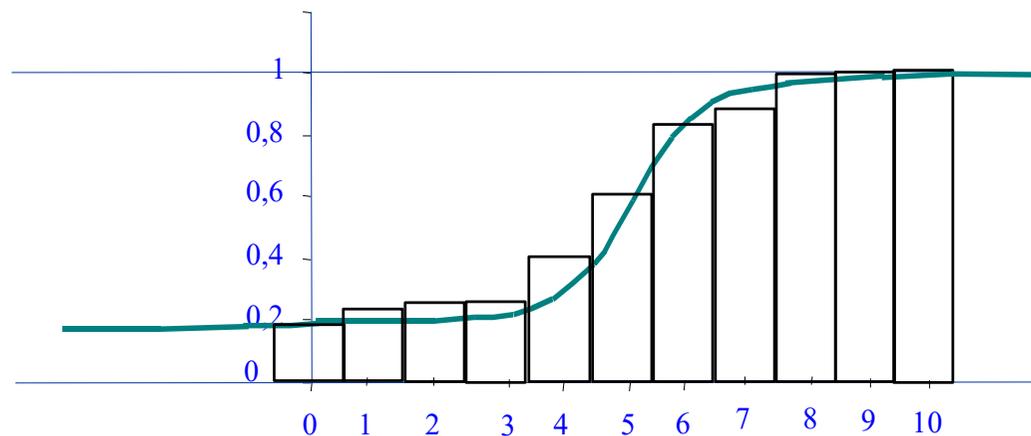
$$P(\theta) = (p_0, p_1, \dots, p_{K-1})$$

- ☐ Distribución normal
- ☐ Distribución logística

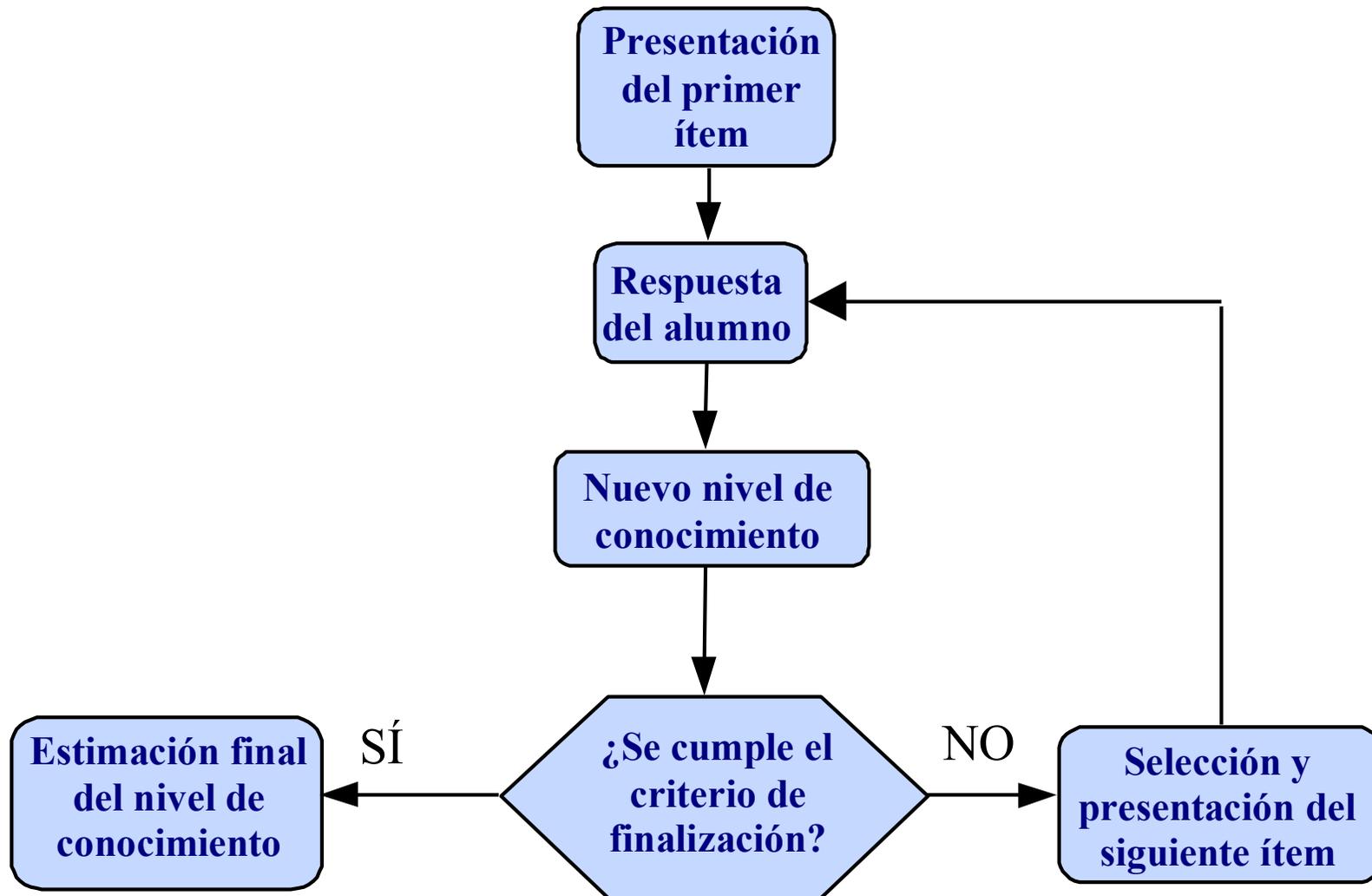
$$p_k = p_k(u = 1 | \theta = k) = c + (1 - c) \frac{1}{1 + e^{-1.7a \left(\left(k - \frac{K-1}{2} \right) - b \right)}}$$

- ◆ A partir de los vectores característicos --> Hallar los parámetros

- ☐ Aproximación por mínimos cuadrados



4. TEORÍA DE TEST ADAPTATIVOS



4. Modelo de adaptación en SIETTE

Los elementos básicos de un test adaptativo son:

- ◆ Banco de preguntas (Conjunto de preguntas que se pueden usar)
 - ☐ Debe de haber ítems de dificultad variada
 - ☐ Debe contener un gran número de ítems
- ◆ Modelo de respuesta (Predice cómo responde el alumno a cada pregunta)
 - ☐ Teoría de respuesta al Ítem
- ◆ Criterio de selección (Cual es la siguiente pregunta)
 - ☐ Métodos bayesianos --> minimizar la varianza esperada a posteriori
 - ☐ Métodos de la máxima información --> basado en la dificultad de las preguntas
 - ☐ Aleatorio (No adaptativo)
- ◆ Criterio de terminación (Cuando se han hecho suficientes preguntas)
 - ☐ Valor mas probable de la distribución superior a un umbral
 - ☐ Varianza de la distribución a posteriori menor que un cierto valor.
 - ☐ Número mínimo y máximo de preguntas
- ◆ Nivel de conocimiento de entrada (Estimación a priori)
 - ☐ Homogénea
 - ☐ Binomial
 - ☐ Según la estimación de la población

4. Modelo de adaptación en SIETTE

EJEMPLO 3: SELECCIÓN DE LA SIGUIENTE PREGUNTA (CRITERIO BAYESIANO)

- Curvas características de los posibles preguntas a realizar

| | $\theta=0$ | $\theta=1$ | $\theta=2$ | $\theta=3$ |
|----------------|------------|------------|------------|------------|
| P ₁ | 0,1 | 0,3 | 0,7 | 0,9 |
| P ₂ | 0,5 | 0,6 | 0,9 | 1,0 |
| P ₃ | 0,3 | 0,6 | 0,8 | 0,9 |
| P ₄ | 0,3 | 0,4 | 0,7 | 0,9 |
| P ₅ | 0,1 | 0,2 | 0,3 | 0,9 |

- Estimación actual del nivel de conocimientos

(0,1 0,2 0,6 0,1)

- Cálculo de la probabilidad a posteriori y la varianza esperada en cada caso

| | u=1 (acierto) | | | | media | var. | u=0 (fallo) | | | | media | var. | E[var] |
|----------------|---------------|-------|-------|--------|-------|-------|-------------|-------|-------|--------|-------|-------|--------|
| P ₁ | (0,017 | 0,103 | 0,724 | 0,155) | 2,017 | 0,327 | (0,214 | 0,333 | 0,429 | 0,024) | 1,262 | 1,240 | 0,784 |
| P ₂ | (0,062 | 0,148 | 0,667 | 0,123) | 1,852 | 0,497 | (0,263 | 0,421 | 0,316 | 0,001) | 1,053 | 1,215 | 0,856 |
| P ₃ | (0,042 | 0,167 | 0,667 | 0,125) | 1,875 | 0,443 | (0,250 | 0,286 | 0,429 | 0,036) | 1,250 | 1,150 | 0,796 |
| P ₄ | (0,048 | 0,129 | 0,677 | 0,145) | 1,919 | 0,461 | (0,184 | 0,316 | 0,474 | 0,026) | 1,342 | 0,979 | 0,720 |
| P ₅ | (0,031 | 0,125 | 0,563 | 0,281) | 2,094 | 0,522 | (0,132 | 0,235 | 0,618 | 0,015) | 1,515 | 0,879 | 0,701 |

5. EL SIMULADOR DE SIETTE

OBJETIVOS

- ◆ Evaluar y comparar el comportamiento de los algoritmos
- ◆ Analizar empíricamente la influencia de los distintos parámetros
- ◆ Estudiar en un entorno controlado nuevos algoritmos de aprendizaje
- ◆ Estudiar mecanismos para la identificación de ítems anómalos.

FACTORES QUE INFLUYEN EN LOS TEST

- Número de niveles de conocimiento.
- Tamaño de la base de preguntas del test.
- Criterio de selección de preguntas: aleatorio, basado en la dificultad o bayesiano.
- Criterio de finalización: fijo o variable en función de la precisión.
- Método de evaluación: modal, bayesiano o porcentual.
- Distribución del conocimiento real de los alumnos: homogénea o binomial.
- Número de alumnos (= número de test a realizar).
- Conjunto inicial de preguntas.
- Dificultad (real o estimada) de los ítems.
- Factor de discriminación (real o estimado).
- Factor de adivinanza (real o estimado).
- Factor de distracción (real o estimado).
- Modo de aprendizaje: incremental, por lotes, o no incremental

5. EL SIMULADOR DE SIETTE

FUNCIONAMIENTO DEL SIMULADOR

- ◆ Se generan N alumnos, cada uno de ellos de conocimiento k
- ◆ Para cada alumno se realiza un test de n preguntas. (según criterio terminación)
- ◆ Se simula la respuesta del alumno al ítem generando un número aleatorio entre 0 y 1 siguiendo una distribución homogénea (equiprobable). p
- ◆ Si p es mayor que $p(u=1|\theta=k)$ (componente k de la **curva característica real**) se considera que el alumno ha acertado la pregunta
- ◆ Se calcula la distribución de conocimiento a posteriori. (a partir de las **curvas características estimadas**)
- ◆ Al final del test se obtiene una estimación del conocimiento del alumno k'
- ◆ Se anotan los resultados de cada pregunta de acuerdo con las estimaciones. (**curva característica aprendida**)
- ◆ Aprendizaje: Se sustituye la **curva característica estimada** con la **curva característica aprendida**.

5. EL SIMULADOR DE SIETTE

PARAMETROS DEL SIMULADOR

- ◆ **Generación del conjunto de ítems, (Curvas características)**
 - ☐ Definición a partir de un fichero
 - ☐ *Conjuntos de ítems correctos, equilibrados, aleatorio*
- ◆ **Generación del conjunto de alumnos**
 - ☐ Definición a partir de un fichero
 - ☐ Distribución normal
 - ☐ Distribución homogénea
- ◆ **Simulación del test**
 - ☐ Criterio de selección
 - ☐ Criterio de terminación
- ◆ **Aprendizaje de las distribuciones**
 - ☐ *Incremental*
 - ☐ *Por lotes*
 - ☐ *No incremental*
- ◆ **Análisis de resultados.**
 - ☐ *Medias, Varianzas, correlaciones, etc.*

5. EL SIMULADOR DE SIETTE

EJEMPLO 4: ESTUDIO DE LOS CRITERIOS DE SELECCIÓN

- ◆ 11 niveles de conocimiento: [0-10]
- ◆ 1000 alumnos, distribución uniforme de nivel de conocimiento.
- ◆ 100 cuestiones. Parámetros correctamente estimados
 - ☐ $a=1,2$;
 - ☐ b : distribución uniforme entre 0 y 10
 - ☐ $c = 0$;
- ◆ Criterio de finalización: factor de confianza del 90%.

| Número de clases K | Criterio de selección Aleatorio | | Criterio de selección Bayesiano | | Criterio de Selección Basado en Dificultad | |
|----------------------|--|--------------------------------------|--|--------------------------------------|--|--------------------------------------|
| | % de alumnos calificados correctamente | Numero medio de preguntas planteadas | % de alumnos calificados correctamente | Numero medio de preguntas planteadas | % de alumnos calificados correctamente | Numero medio de preguntas planteadas |
| 3 | 95.82 | 3.59 | 96.06 | 3.58 | 95.62 | 3.58 |
| 5 | 92.76 | 10.38 | 93.31 | 6.87 | 94.67 | 7.37 |
| 7 | 92.85 | 18.16 | 92.75 | 8.70 | 94.43 | 9.03 |
| 9 | 92.93 | 26.39 | 92.53 | 9.85 | 94.23 | 10.14 |
| 11 | 92.92 | 34.54 | 92.10 | 10.71 | 94.14 | 11.02 |

6. ESTIMACIÓN DE LOS PARÁMETROS

- ◆ *Métodos condicionales:* Construir la función de verosimilitud considerando como incógnitas tanto a los parámetros de las preguntas como los rasgos latentes.

$$L(u_{ij} | \theta_j, a_i, b_i, c_i) = \prod_{j=1}^m \prod_{i=1}^n P_i(\theta_j)^{u_{ij}} (1 - P_i(\theta_j))^{(1-u_{ij})}$$
$$\left(\frac{\partial \ln L(u_{ij} | \theta_j, a_i, b_i, c_i)}{\partial a_i} \right) = 0 \quad \left(\frac{\partial \ln L(u_{ij} | \theta_j, a_i, b_i, c_i)}{\partial b_i} \right) = 0$$
$$\left(\frac{\partial \ln L(u_{ij} | \theta_j, a_i, b_i, c_i)}{\partial c_i} \right) = 0 \quad \left(\frac{\partial \ln L(u_{ij} | \theta_j, a_i, b_i, c_i)}{\partial \theta_j} \right) = 0$$

- ◆ *Métodos incondicionales:* Asumir un conocimiento a priori sobre la distribución de la población de individuos que realiza el test, $h(\theta)$, Integrar en la función de verosimilitud, que solo dependerá de los parámetros de los ítems:

$$L(u_{ij} | a_i, b_i, c_i) = \prod_{j=1}^m \int P_i(\theta)^{u_{ij}} (1 - P_i(\theta))^{(1-u_{ij})} h(\theta) d\theta$$

6. Estimación de parámetros en SIETTE

¿QUE ES EL NIVEL DE CONOCIMIENTOS?

--> ¡ LO QUE MIDE EL TEST!

- ◆ *Definición porcentual:* El porcentaje de preguntas sobre el total que el alumno es capaz de contestar acertadamente.
 - ☐ Los resultados de un test de N preguntas aleatorias son un buen estimador del nivel de conocimientos del alumno
 - ☐ La TRI y la TAI sirven para reducir el número de preguntas y/o aumentar la fiabilidad.
- ◆ *Definición basada en la estimación del profesor.* El profesor decide cuales son las condiciones para que un alumno pertenezca a cada una de las clases.
 - ☐ El profesor define los parámetros de las preguntas.
 - ☐ Los alumnos se clasifican según estas estimaciones.

LA CLASIFICACION FINAL DEL TEST ES UN BUEN ESTIMADOR DEL NIVEL DE CONOCIMIENTOS --> ESTIMACION DIRECTA DE LA PROBABILIDAD

$$P_{ik} = p_{ik}(u_i = 1 | \theta = k) = \frac{m(u_i = 1 / \theta = k)}{m(\theta = k)}$$

6. Estimación de parámetros en SIETTE

- ◆ El profesor no puede estimar bien todos los parámetros.
 - ☐ Solo se le pide la estimación del parámetro de dificultad de cada pregunta
 - ☐ Solo se pide que la media de sus estimaciones se correcta.
 - ☐ Se estiman automáticamente los parámetros de discriminación y adivinanza

- ◆ HIPÓTESIS Conjunto equilibrado de ítems:

$$\forall k \in [0, K - 1] \quad \forall i \in I \quad d_r(i) = \frac{\sum_{d_e(j)=k} d_e(j)}{N_k}$$

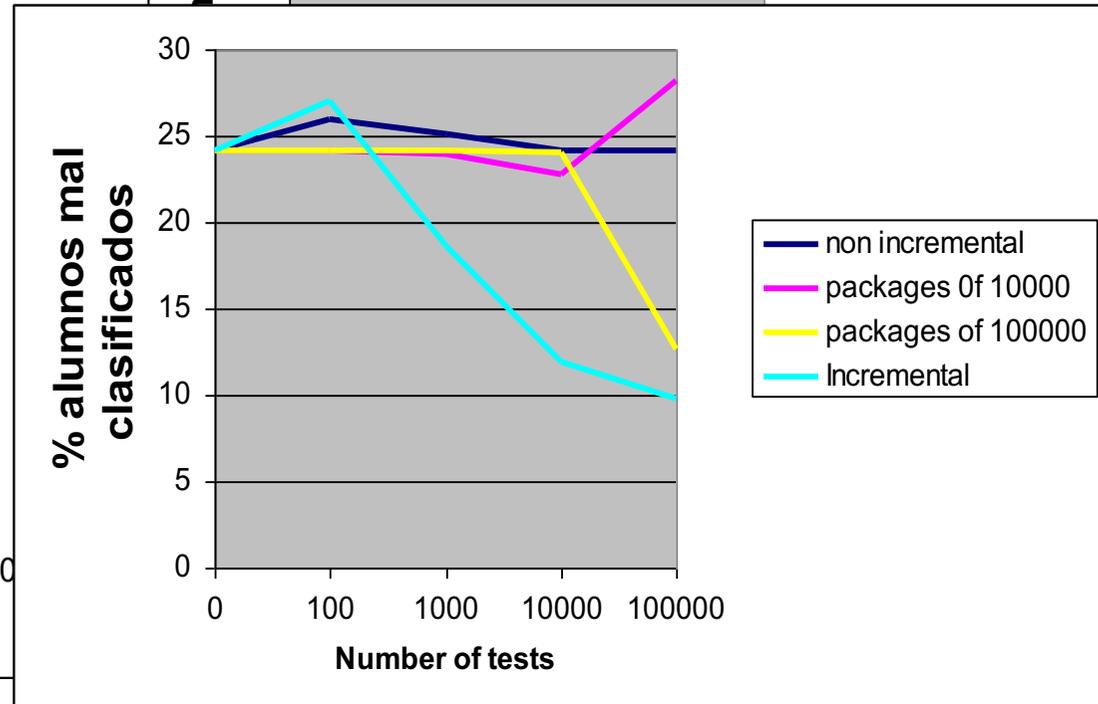
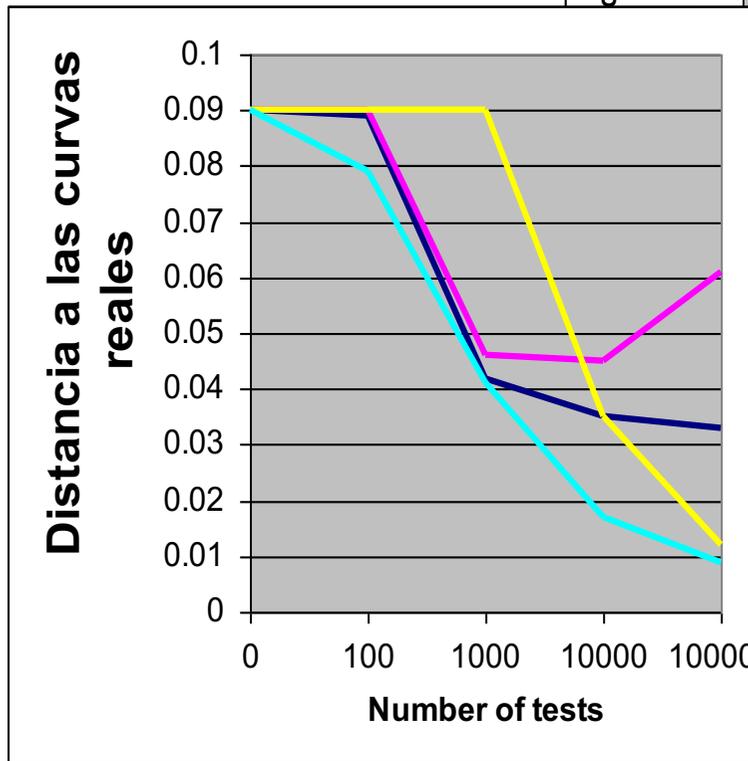
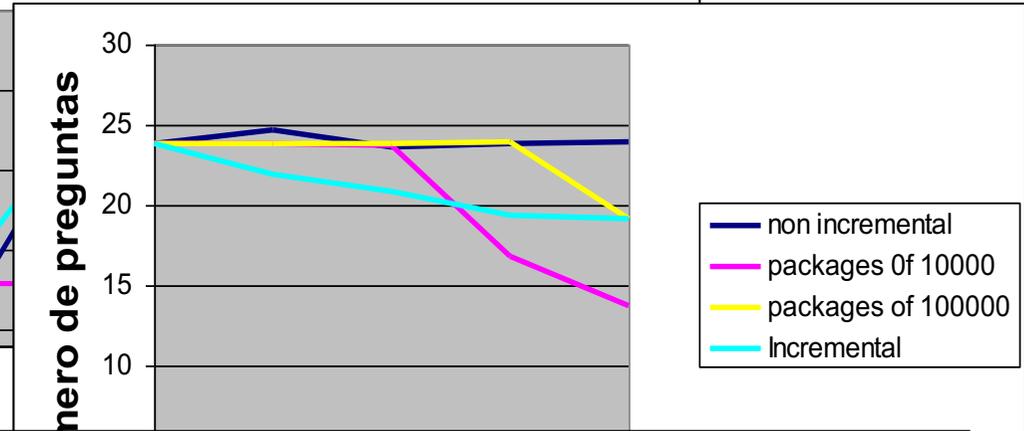
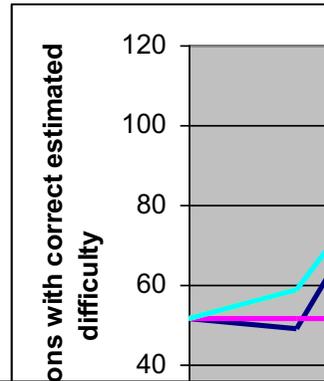
- ◆ Resultados del simulador:

| | Conjunto correcto de ítems | | Conjunto equilibrado de ítems | | | |
|---|---|---|---|---|---|---|
| | Criterio de Selección Aleatorio | | Criterio de Selección Aleatorio | | Criterio de Selección Basado en Dificultad | |
| <i>Factor de discriminación estimado a_e</i> | <i>% de alumnos calificados correctamente</i> | <i>Numero medio de preguntas planteadas</i> | <i>% de alumnos calificados correctamente</i> | <i>Numero medio de preguntas planteadas</i> | <i>% de alumnos calificados correctamente</i> | <i>Numero medio de preguntas planteadas</i> |
| 0.2 | 90.4 | 174.9 | 55.4 | 78.2 | 85.4 | 186.8 |
| 0.5 | 91.5 | 35.2 | 83.1 | 32.1 | 82.4 | 33.3 |
| 0.7 | 91.9 | 26.3 | 85.4 | 25.8 | 81.1 | 18.3 |
| 1.2 | 92.8 | 18.1 | 83.1 | 16.0 | 78.4 | 8.6 |
| 1.7 | 93.8 | 15.3 | 73.7 | 12.0 | 71.4 | 6.1 |

6. Aprendizaje de las curvas características

APRENDIZAJE

- ◆ Incremental
- ◆ Por lotes
- ◆ No Incremental



7. TIPOS DE ÍTEMS

ITEMS DICOTÓMICOS

En un espacio euclídeo la línea mas corta entre dos puntos es una recta

Verdadero

Falso

- ◆ Respuesta correcta/incorrecta
- ◆ Estimación inicial de la probabilidad condicionada de acierto

$$\mathbf{P}_i = \overline{P(u_i | \boldsymbol{\theta})} = (p(u_i = 1 | \theta = 0), p(u_i = 1 | \theta = 1), \dots, p(u_i = 1 | \theta = K - 1))$$

$$\mathbf{Q}_i = \overline{Q(u_i | \boldsymbol{\theta})} = (p(u_i = 0 | \theta = 0), p(u_i = 0 | \theta = 1), \dots, p(u_i = 0 | \theta = K - 1))$$

$$\forall k \in [0, K - 1], p(u_i = 0 | \theta = k) = 1 - p(u_i = 1 | \theta = k)$$

7. TIPOS DE ÍTEMS

ITEMS POLITÓMICOS

| |
|--|
| ¿Quién descubrió América en el año 1492? |
| <input type="radio"/> Abraham Lincoln |
| <input type="radio"/> Cristobal Colón |
| <input type="radio"/> Americo Vespuccio |
| <input type="radio"/> James Cook |

- ◆ Respuestas: a_0 (en blanco) , a_1 , (inc.) a_2 , (inc.).... a_c (correcta)....., a_r (inc.)
- ◆ Estimación inicial de la curva característica de la respuesta correcta
 - ☐ Dificultad (*estimada por el profesor*)
 - ☐ Factor de adivinanza ($1/r$)
 - ☐ Factor discriminación ($0,5 - 1,0$)
- ◆ Se reparte inicialmente la probabilidad restante por igual entre las opciones

$$p_k(u = a_j)_{j \neq c} = \frac{1 - p_k(u = a_c)}{r}$$

7. TIPOS DE ÍTEMS

ITEMS MULTIOPCIÓN DE RESPUESTA INDEPENDIENTE

| | | |
|--|------------------------------------|-----------------------------------|
| ¿Cuales de las siguientes países pertenecen con pleno derecho a la Unión Europea en el año 2001? | | |
| <input type="checkbox"/> Francia | <input type="checkbox"/> Italia | <input type="checkbox"/> Alemania |
| <input type="checkbox"/> Japón | <input type="checkbox"/> Rusia | <input type="checkbox"/> Suiza |
| <input type="checkbox"/> Polonia | <input type="checkbox"/> Noruega | <input type="checkbox"/> Bélgica |
| <input type="checkbox"/> Holanda | <input type="checkbox"/> Finlandia | <input type="checkbox"/> España |

- ◆ Equivale a r ítems dicotómicos
- ◆ Estimación inicial de la curva característica de la respuesta correcta.
 - ☐ Se estima inicialmente que la dificultad de cada opción es igual.
 - ☐ Factor adivinanza de cada opción (1/2)
 - ☐ Se definen $2 \times r$ curvas
- ◆ Mecanismos de adaptación:
 - ☐ Bayesiano: Simular la aplicación conjunta del grupo de ítems dicotómicos
 - ☐ Por Dificultad: Estimar una dificultad media

7. TIPOS DE ÍTEMS

ITEMS MULTIOPCIÓN DE RESPUESTA DEPENDIENTE

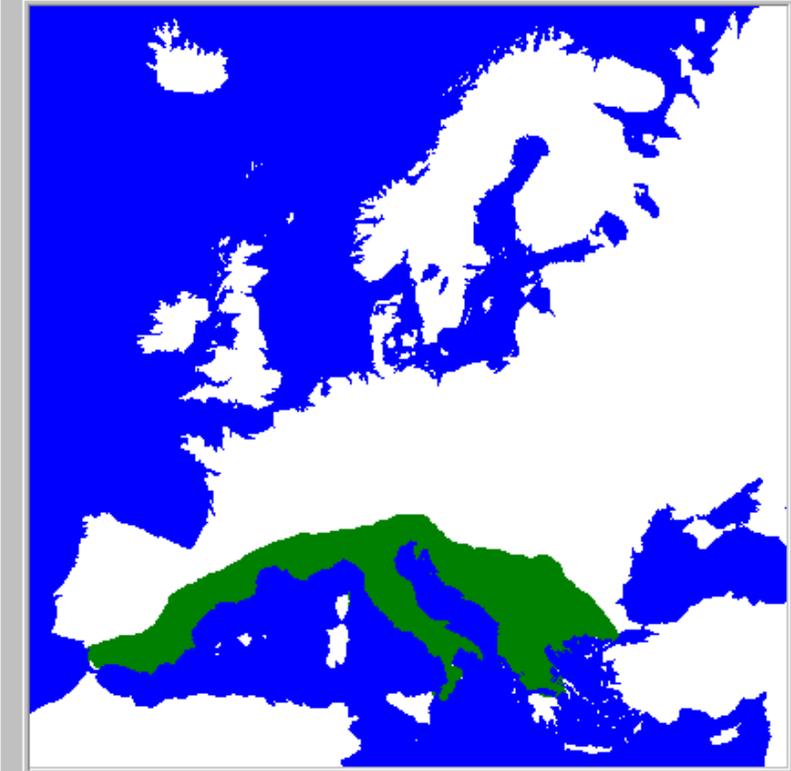
| ¿Cuáles son los componentes químicos de la pólvora? | |
|---|--|
| <input type="checkbox"/> Uranio | <input type="checkbox"/> Cloruro sódico |
| <input type="checkbox"/> Sulfuro ferrico | <input checked="" type="checkbox"/> Carbón |
| <input checked="" type="checkbox"/> Azufre | <input checked="" type="checkbox"/> Clorato potásico |
| <input type="checkbox"/> Nitrato potásico | <input type="checkbox"/> Acido sulfúrico |

- ◆ Equivale a *un* ítem politómico con 2^r opciones (sin posible respuesta en blanco)
- ◆ Estimación inicial de la curva característica de la respuesta correcta.
 - ☐ Podría haber mas de una combinación correcta.
 - * Se supone que tienen la misma dificultad
 - ☐ Factor adivinanza de cada opción ($1/2^r$)
 - ☐ Se definen 2^r curvas
- ◆ Mecanismos de adaptación:
 - ☐ Bayesiano: Sin cambios
 - ☐ Por Dificultad: Estimar una dificultad media de las respuestas correctas.

7.TIPOS DE ÍTEMS

ITEMS EVALUADOS MEDIANTE APPLETS

Señalar sobre el mapa cual es la distribución geográfica de la especie *Pinus nigra*



Area en verde 90.18 % Area en blanco 96.22 %

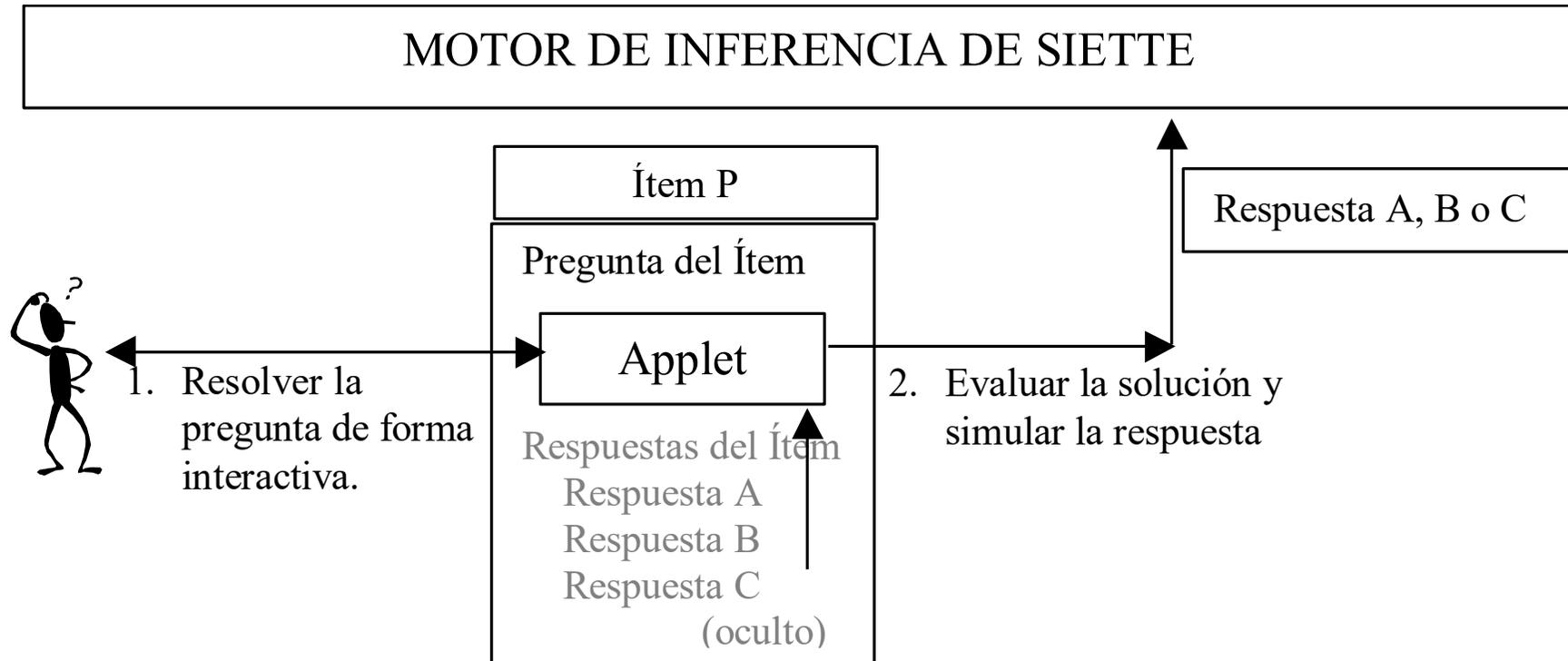
Area natural
Area no natural
Area en blanco

Borrar todas
Deshacer
Mostrar/Ocultar referencias
Corregir

The image shows a web-based applet interface for a geography quiz. At the top, the text asks the user to identify the geographical distribution of the species *Pinus nigra* on a map of Europe. The map shows a green area in the Iberian Peninsula. Below the map, statistics show 'Area en verde 90.18 %' and 'Area en blanco 96.22 %'. To the right is a control panel with a legend for 'Area natural' (green), 'Area no natural' (light green), and 'Area en blanco' (white), each with three size options. Below the legend are buttons for 'Borrar todas', 'Deshacer', 'Mostrar/Ocultar referencias', and 'Corregir'.

7. TIPOS DE ÍTEMS

ITEMS EVALUADOS MEDIANTE APPLETS



- ◆ Equivale a *un* ítem politómico con *r* opciones
- ◆ Factor de adivinanza pequeño.

7. TIPOS DE ÍTEMS

ITEMS DE RESPUESTA LIBRE

- ◆ Mediante expresiones regulares

¿Quien descubrió América en el año 1492?

`<<*Colón*|*Colombo*>>`

- ◆ Mediante expresiones aritméticas

Ernesto deja caer una maceta desde una terraza, que tarda 3 segundos en alcanzar el suelo. Suponiendo que no hay rozamiento por efecto del aire. ¿A qué altura en metros está la terraza?

`<<#(0.5*9.81*3**2)3%>>`

- ◆ Múltiples

Completar el siguiente texto:

"En un de la
de cuyo nombre no quiero acordarme...."

Fragmento de El

`<I>Completar el siguiente texto: </I>

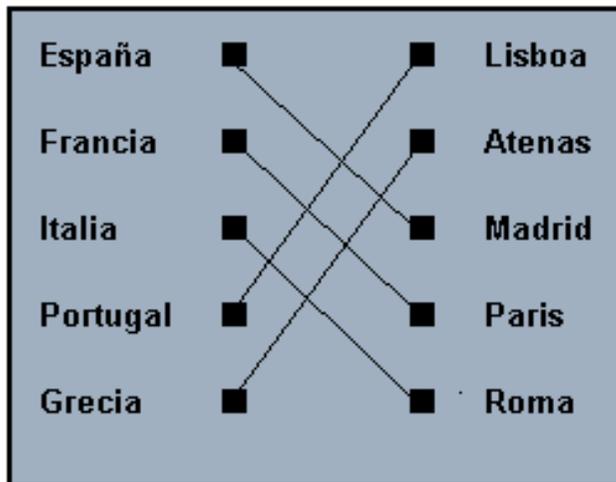
 "En un <<lugar>> de la
<<Mancha>>
 de cuyo nombre no
quiero acordarme..."
 Fragmento
de El <<Quijote>>`

7. TIPOS DE ÍTEMS

OTROS TIPOS DE ÍTEMS

- ◆ Bibliotecas de ítems genéricos o específicos de cada materia

Enlazar los países y sus capitales



Provincias de Andalucía



Buscar los números del 1 al 8

A 10x7 grid of letters. A path of red boxes highlights the sequence of letters: C, I, N, C, O, T, A, D, O, S, E, I, S.

| | | | | | | |
|---|---|---|---|---|---|---|
| D | I | C | I | N | C | O |
| O | S | N | T | A | D | O |
| O | I | I | R | S | O | R |
| C | J | S | E | I | S | T |
| H | L | U | S | T | Z | A |
| O | N | O | O | I | E | U |
| L | U | N | O | A | R | C |

7.TIPOS DE ÍTEMS

PLANTILLAS PARA LA GENERACIÓN DE ÍTEMS

| | | | |
|--|--|--|--|
| <pre>¿Cuál es el valor de x al final de este programa?

 <? srand(date("U")); \$randMax=getRandMax(); \$rand=Rand(); \$x =intval(doubleval(\$rand)*doubleval(10)/doubleval(\$randMax)); echo "<CODE><PRE>"; echo " x=\$x;
"; echo " x++;"; echo "</PRE></CODE>"; ></pre> | | | |
| <pre><? \$sol = \$x+\$x; echo \$sol; ></pre> | <pre><? \$sol = \$x+1; echo \$sol; ></pre> | <pre><? \$sol = \$x-1; echo \$sol; ></pre> | <pre><? \$sol = \$x; echo \$sol; ></pre> |

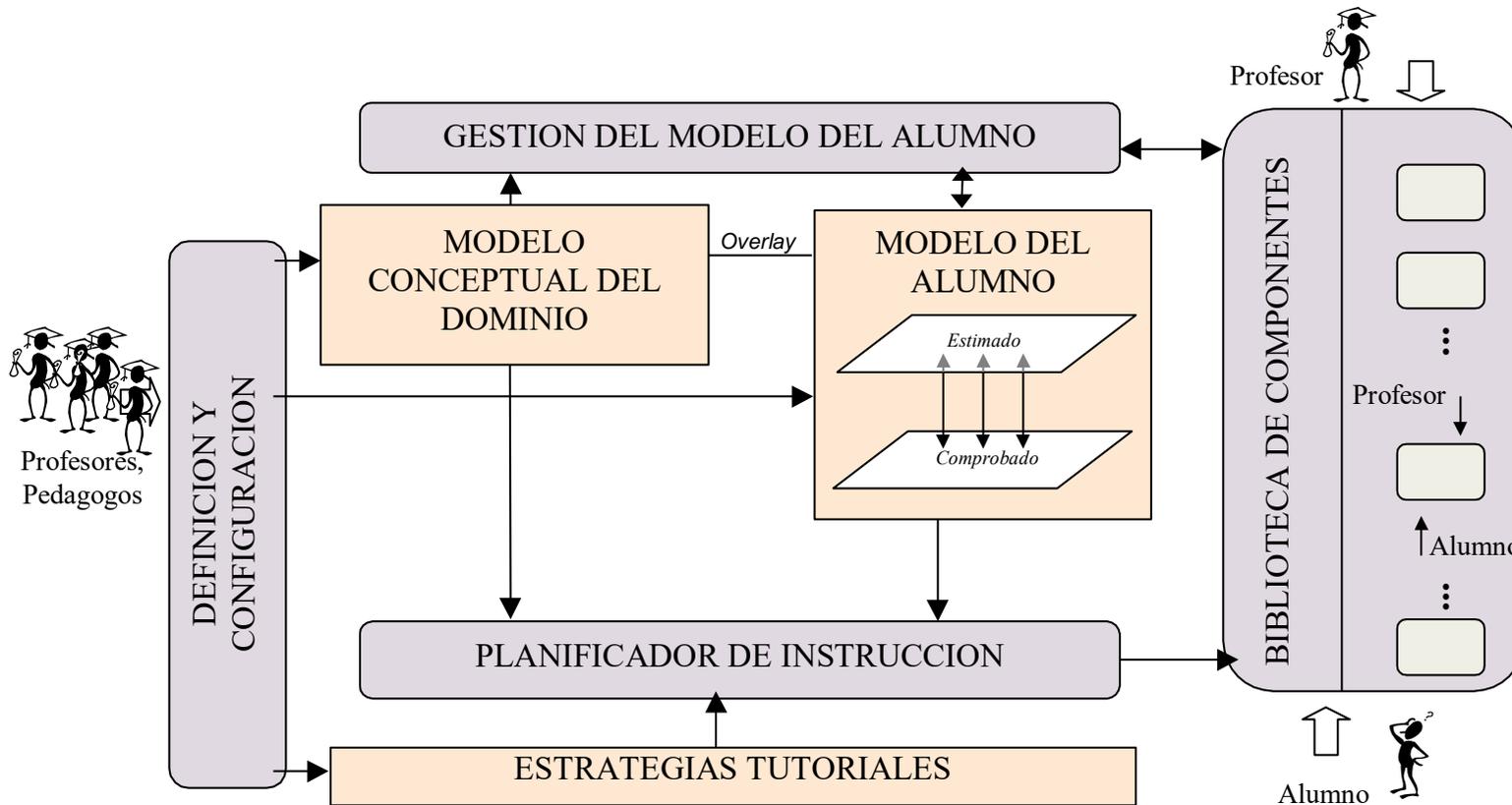
¿Cuál es el valor de x al final de este programa?

```
x=6;
x++;
```

12 7 5 6

8. INTEGRACIÓN EN S.T.I.

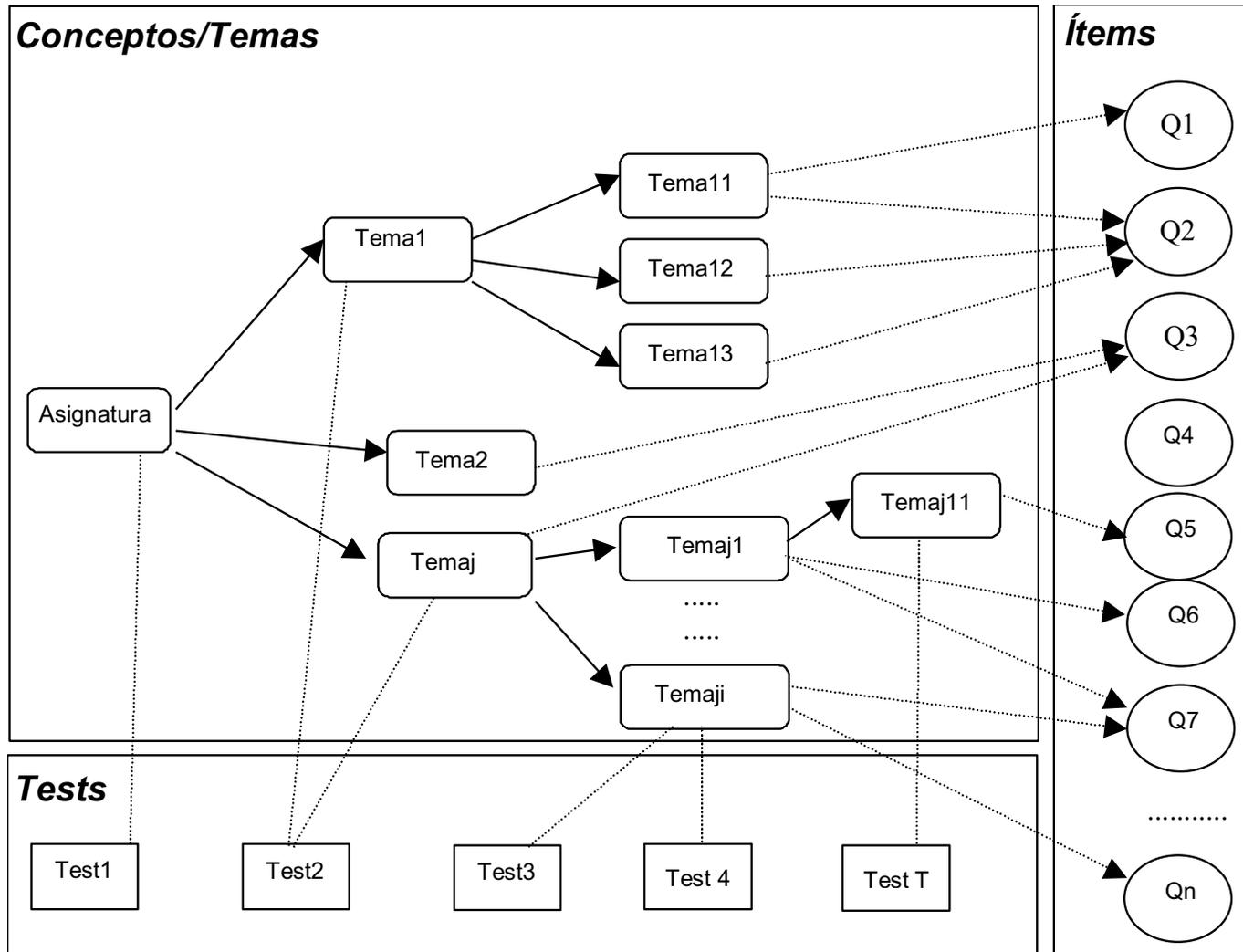
ESTRUCTURA DE UN SISTEMA TUTOR INTELIGENTE (MEDEA)



- ◆ El Conocimiento de la Materia está estructurado
- ◆ Es necesaria una evaluación precisa para el diagnóstico y la planificación

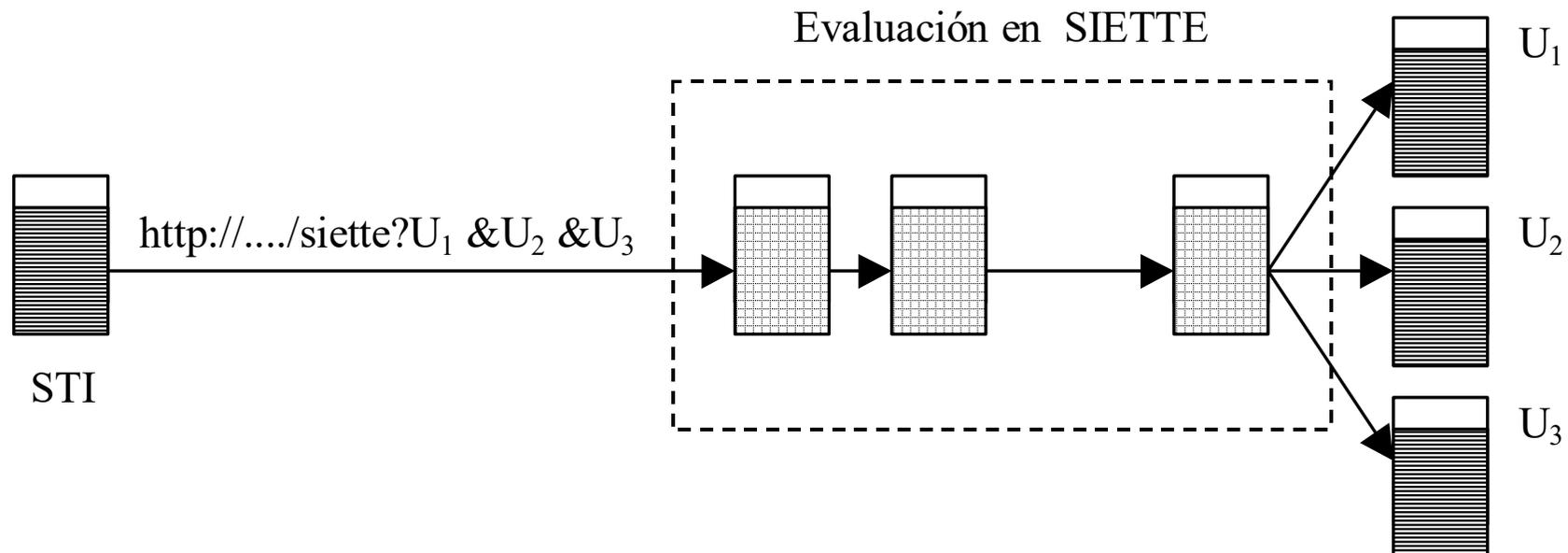
8. INTEGRACIÓN EN S.T.I.

ESTRUCTURA DE LA BASE DE CONOCIMIENTOS EN SIETTE



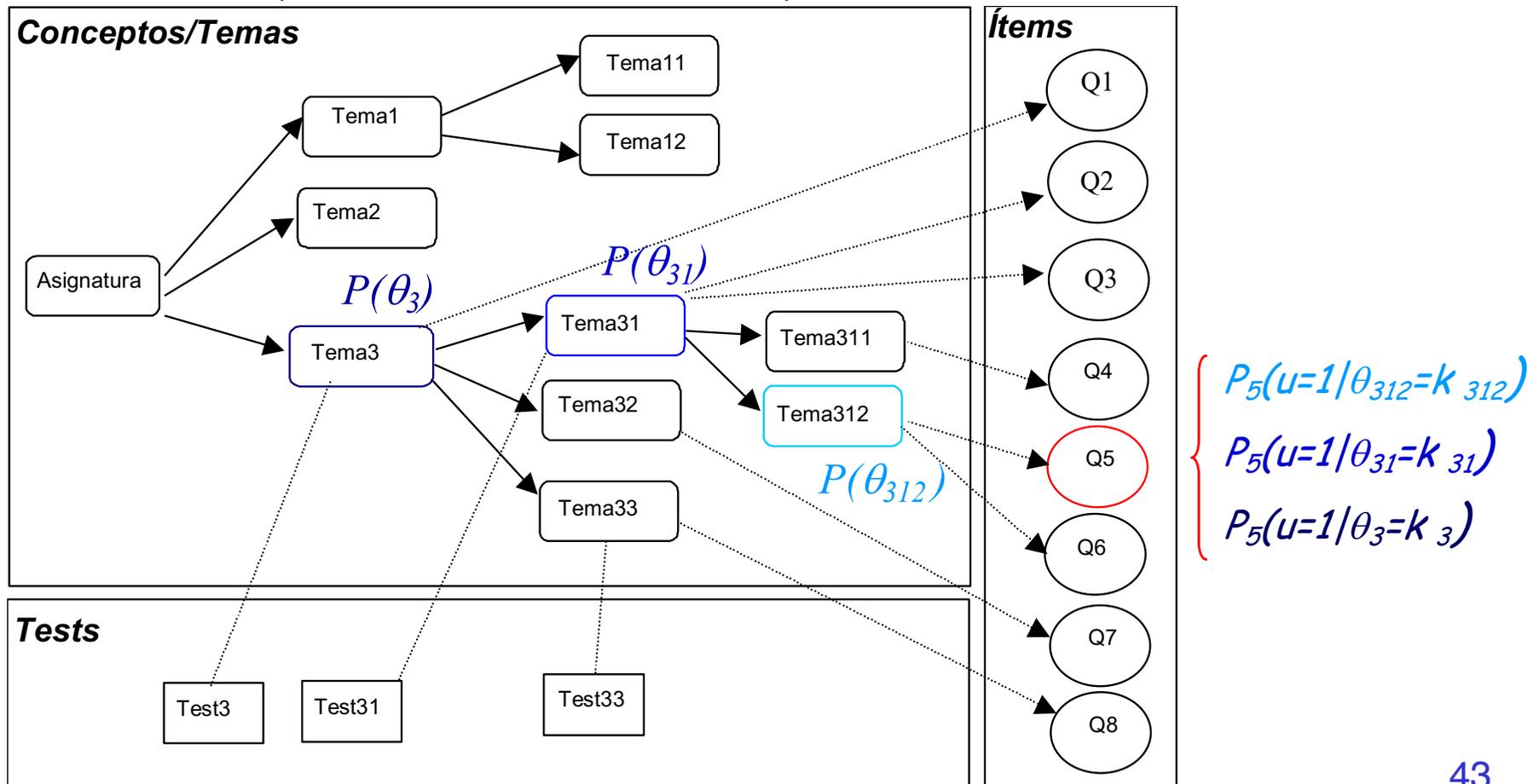
8. Evaluación con ítems unidimensionales.

- ◆ Un test independiente para cada tema/concepto.
 - ☐ Se requieren muchos tests para una evaluación múltiple.
 - ☐ Ítems asociados a un único tema.
 - ☐ Buena precisión
 - ☐ Fácil integración desde el punto de vista técnico.



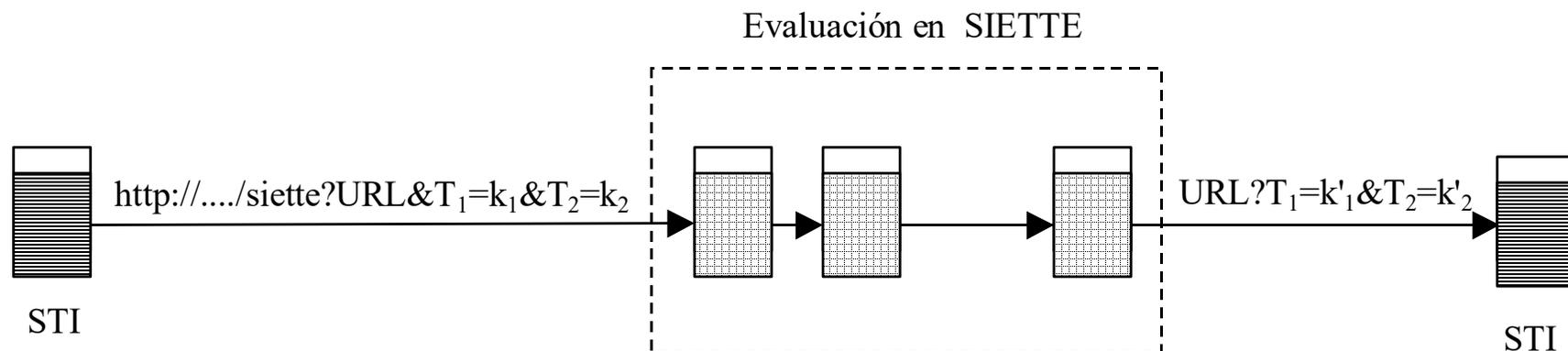
8. Evaluación con ítems unidimensionales. $P_i(\theta_k)$

- ◆ Un test único con múltiples evaluaciones simultáneas (1)
 - ☐ Ítems asociados a un único tema.
 - ☐ Los temas representan agrupación de otros conceptos/temas
 - ☐ Múltiples curvas características para cada ítem, según el curriculum
 - ☐ Evaluación para todos los nodos del árbol a partir de la raíz.



8. Evaluación con ítems unidimensionales.

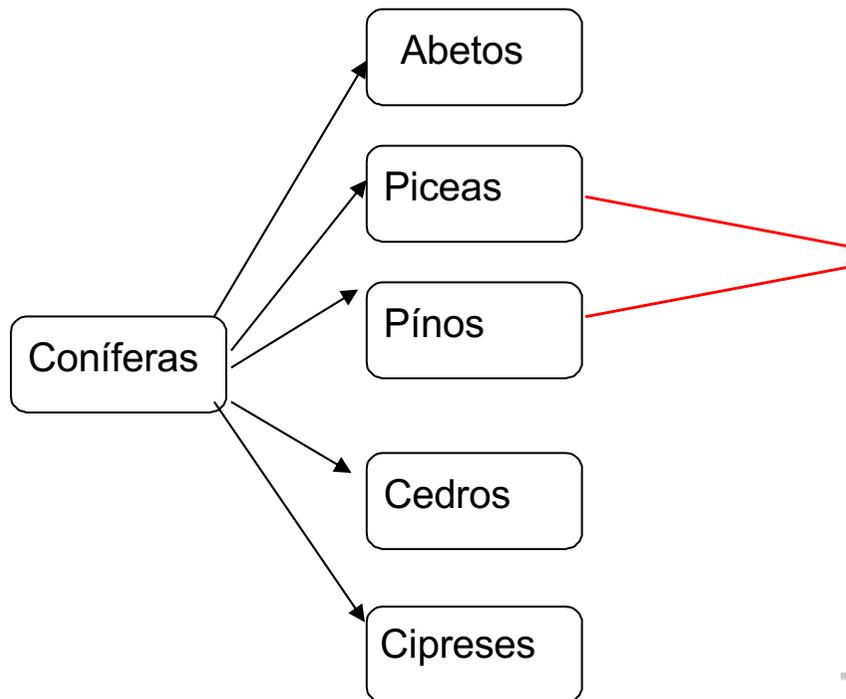
- ◆ Un test único con múltiples evaluaciones simultáneas (2)
 - ☐ Se requiere un único test
 - ☐ Precisión variable según la profundidad en la jerarquía
 - ☐ Número mínimo de preguntas de cada nodo
 - ☐ Mecanismo de adaptación
 - * Adaptación para el nivel de conocimientos del nodo raíz
 - * Adaptación conjunta para todos los nodos
 - ☐ La integración requiere un protocolo simple de paso de parámetros



8. Evaluación con ítems multidimensionales.

- ◆ Ítems asociados a varios temas.

- ☐ Son necesarios para representar adecuadamente ciertas preguntas
- ☐ Surgen a medida que el curriculum es mas detallado



¿Cuál de las siguientes fotografías corresponde a la especie Picea abies ?



8. Evaluación con ítems multidimensionales.

- ◆ Ítems asociados a varios temas.

- ▣ Curvas características s-dimensionales

- * Familias dependientes de varios parámetros

$$P(u_i = 1 | \theta_1 \dots \theta_s) = c_i + (1 - c_i) \frac{e^{\left(d_i + \sum_{j=1}^s a_{ij} \theta_j\right)}}{1 + e^{\left(d_i + \sum_{j=1}^s a_{ij} \theta_j\right)}}$$

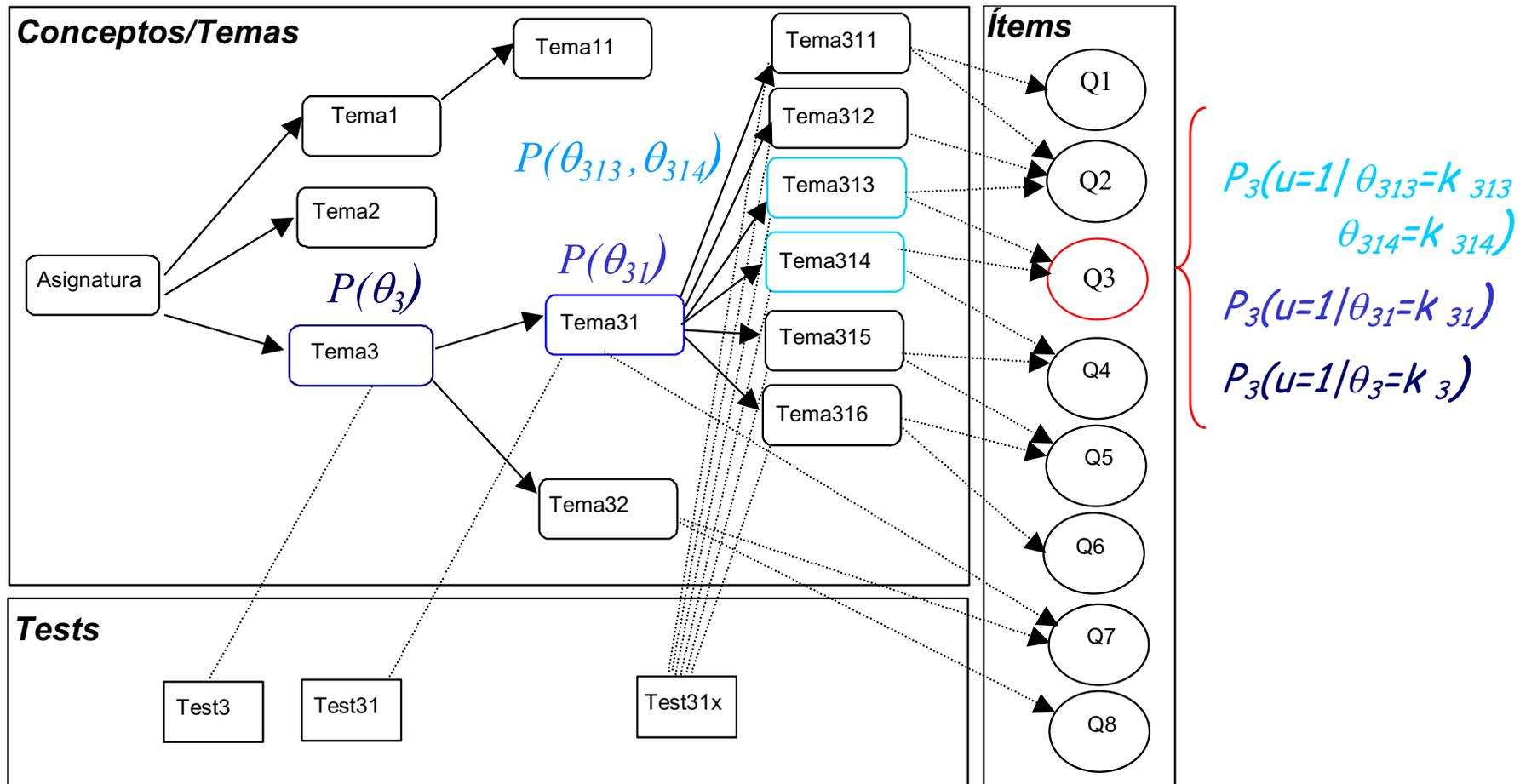
- ▣ Funciones de distribución s-dimensionales

$$P(\theta_1 \dots \theta_s | u_1 \dots u_n) \propto P(\theta_1 \dots \theta_s) \times \prod_{i=1}^n P_i(\theta_1 \dots \theta_s)^{u_i} (1 - P_i(\theta_1 \dots \theta_s))^{(1-u_i)}$$

- ▣ En SIETTE la discretización requiere k^s valores para cada curva característica
 - ▣ Problema de complejidad exponencial.

8. Evaluación con ítems multidimensionales.

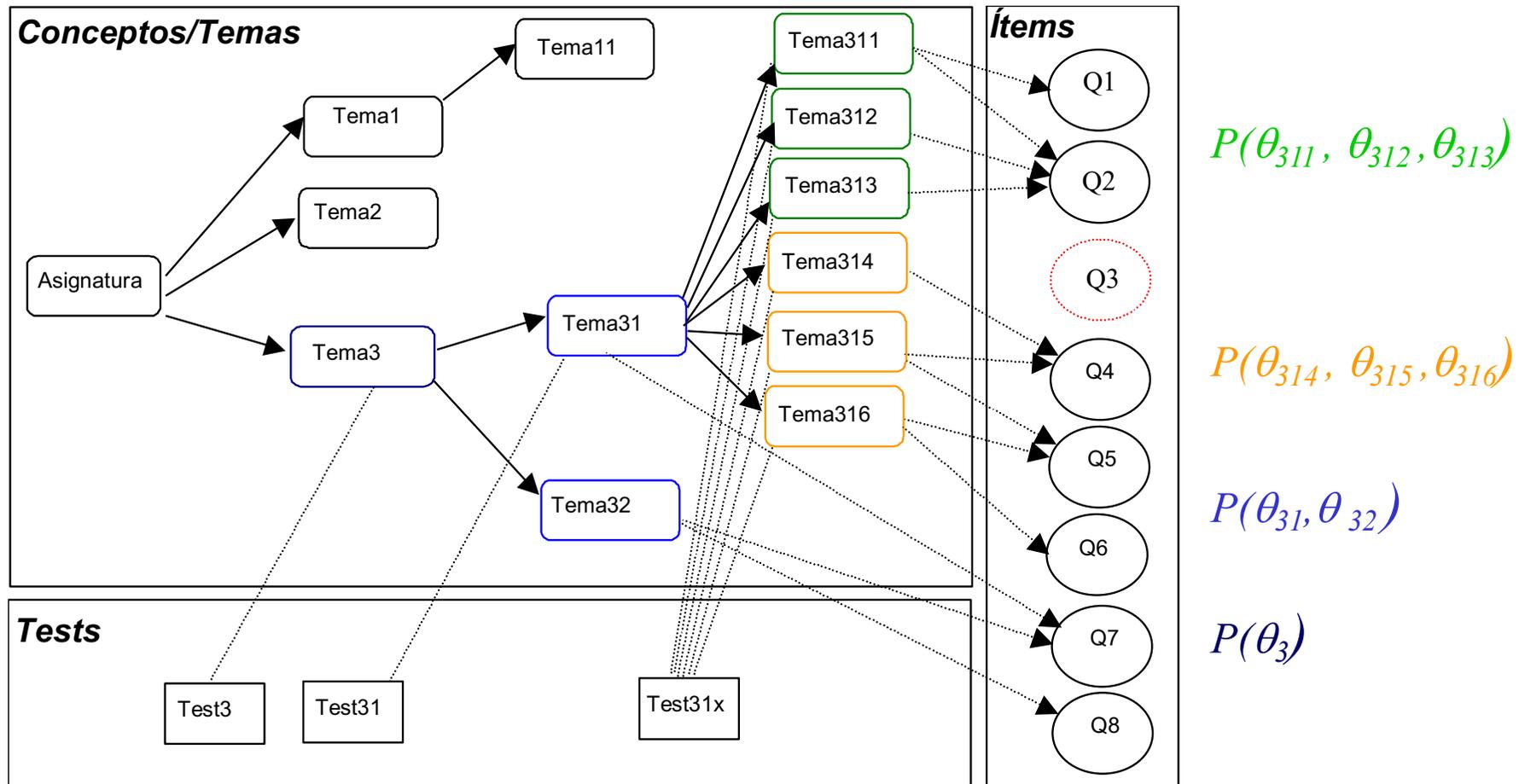
- ◆ Ítems asociados a varios temas.
 - ▣ Es compatible con la evaluación jerárquica múltiple



8. Evaluación con ítems multidimensionales.

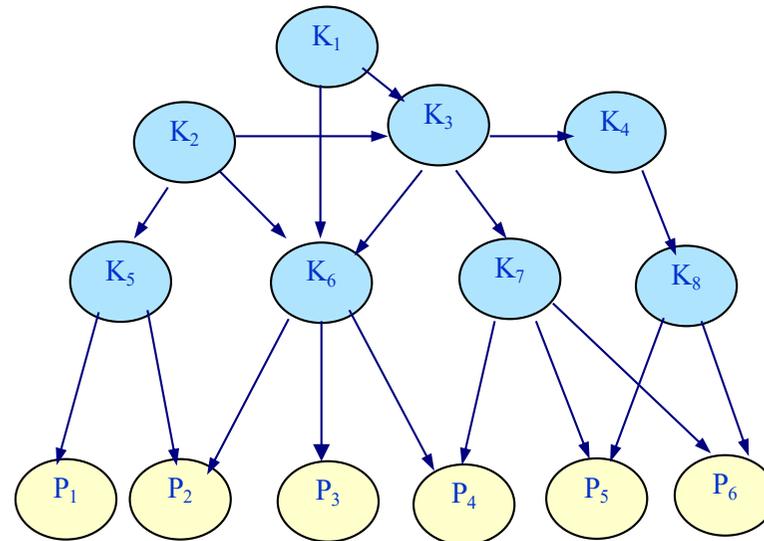
- ◆ Ítems asociados a varios temas.

- ▣ La dimensionalidad aumenta al combinar resultados de varios ítems
- ▣ La explosión multidimensional se puede controlar suprimiendo algunos ítems



8. Evaluación con ítems multidimensionales.

- ◆ Evaluación mediante redes bayesianas
 - ◆ Una red bayesiana es un grafo acíclico dirigido que consta de
 - ▢ Nodos → Variables $\{X_i, i=1,2,\dots,n\}$
 - ▢ Arcos → Relaciones de influencia causal
 - ▢ Parámetros → Probabilidades condicionadas
 $P(X_i / \text{pa}(X_i))$; $\text{pa}(X_i)$ = conjunto de padres del nodo X_i
 - ◆ Una vez definida la red se puede utilizar para hacer inferencias (algoritmos de propagación de probabilidades),



- ◆ Mayor complejidad.

8. CONCLUSIONES

- ◆ Los test adaptativos son mejores.
 - ☐ Reducen el número de preguntas
 - ☐ Aumentan la precisión

- ◆ Implementación eficiente.
- ◆ Multidimensionalidad.

- ◆ Valor añadido
 - ☐ Ventajas del uso de la WWW (instalación, ubicuidad, recogida datos)
 - ☐ Uso de ítems evaluados mediante programas
 - ☐ Uso de esquemas para generación dinámica de ítems
 - ☐ Mejora mediante aprendizaje automático
 - ☐ Integración en sistemas tutores inteligentes

SIETTE: Sistema Inteligente de Evaluación mediante Test para TeleEducación



<http://www.lcc.uma.es/SIETTE>



Departamento de Lenguajes y Ciencias de la Computación.
Universidad de Málaga, SPAIN.