

# Generación Automática de Reactivos en Evaluación Adaptable Computarizada

Joel Suárez Cansino, Emma Lucio Martínez

Universidad Autónoma del Estado de Hidalgo  
Instituto de Ciencias Básicas e Ingeniería  
Centro de Investigación en Tecnologías de Información y Sistemas  
Carretera Pachuca – Tulancingo Km. 4.5  
[jsuarez@uaeh.reduaeh.mx](mailto:jsuarez@uaeh.reduaeh.mx)

## Resumen

*El artículo describe el proceso de evaluación adaptable computarizada a través de una simulación. Se muestra cómo la media de las habilidades calculadas converge a la habilidad real. Asimismo, se plantea el problema de elaborar reactivos en forma automática considerando problemas de la lógica de proposiciones que son de la forma  $p \blacksquare q \blacksquare r$ , en donde  $p, q$  y  $r$  son proposiciones simples que pueden o no contener negación ( $\sim$ ) y el símbolo  $\blacksquare$  representa a cualquiera de los operadores binarios que simbolizan a la disyunción ( $\vee$ ), conjunción ( $\wedge$ ), implicación ( $\rightarrow$ ), reducción ( $\leftarrow$ ) y equivalencia ( $\leftrightarrow$ ). En el trabajo se muestra evidencia experimental de que las métricas basadas en co-ocurrencias tanto del modelo cognitivo propuesto como aquella de las muestras se encuentran correlacionadas.*

## Palabras clave

simulación, adaptable, co-ocurrencias, similaridad, psicométrico, cognitivo, outlier, interpretación, proposición.

## Introducción

Los resultados de evaluación de conocimientos en áreas de ciencias, español y matemáticas que México ha obtenido a nivel internacional, evaluaciones que han sido organizadas y realizadas por un comité localizado dentro de la OCDE (Organización para la Cooperación y el Desarrollo económicos), han mostrado que los estudiantes del último semestre del nivel secundaria, o que están próximos a ingresar al nivel medio superior, carecen de las habilidades mínimas necesarias dentro de las áreas mencionadas que les permitan enfrentar con éxito situaciones de la vida real.

Se han propuesto diversas acciones a nivel nacional con el objeto de resolver este problema. El enfoque de este artículo es el de mostrar una alternativa diferente a las soluciones expuestas y que probablemente sirva como complemento de las mismas. Esta alternativa ha sido y está actualmente siendo estudiada y aplicada en diferentes países de Europa y Estados Unidos.

Una evaluación que no sea diagnóstica y correctiva conduce a juicios erróneos acerca de los niveles de conocimiento que haya alcanzado un estudiante. El proceso de evaluación adaptable intenta cubrir esta dificultad haciendo que el examen se adapte al conocimiento del estudiante, más que éste al contenido del examen, como generalmente se hace. Este es un tema interesante que ocupa a los investigadores del Cuerpo Académico de Computación Inteligente del Centro de Investigación en Tecnologías de Información y Sistemas.

Actualmente, la investigación acerca de este relativamente novedoso enfoque de evaluación forma el sustento de proyectos de desarrollo tecnológico e investigación que tienen soporte financiero por parte del Consejo Nacional de Ciencia y Tecnología del Estado. Uno de los temas que interesa es el que se refiere a la construcción

automática de reactivos con la idea principal de hacer más eficiente y barata la construcción de éstos. Actualmente, algunos investigadores a nivel internacional han obtenido resultados alentadores en este sentido considerando reactivos con formato especial dentro de ciertas áreas del conocimiento.

La propuesta de un sistema para generación automática de reactivos no es nada sencillo ya que implica la aportación de un modelo psicométrico y un modelo cognitivo, con los cuales se pueda modelar la respuesta del examinado y la forma en que éste responde a cada uno de los reactivos propuestos dentro del área de conocimientos previamente establecida. La idea de proponer un sistema tal radica en el hecho de que en evaluación adaptable computarizada se requiere de la existencia de una base de reactivos a los que previamente, bajo un experimento bien diseñado, se les ha calculado su complejidad, además del hecho de que se han cuidado otros diferentes aspectos estadísticos, como por ejemplo el de su independencia estadística y su unidimensionalidad.

Este artículo se divide en tres secciones adicionales. La sección titulada **Evaluación adaptable** comenta algunas ideas esenciales de este concepto, además de que presenta algunos resultados que se han obtenido durante el desarrollo del proyecto y que sirven para la elaboración de tesis de maestría y doctorado de algunos de los colaboradores en el proyecto. La sección titulada **Modelo cognitivo en lógica de proposiciones** presenta también aspectos relacionados con la forma en que se construye el modelo cognitivo para especificar la validez de proposiciones de la forma  $p \blacksquare q \blacksquare r$ , en donde  $p, q$  y  $r$  son proposiciones simples que pueden o no contener negación ( $\sim$ ) y el símbolo  $\blacksquare$  representa a cualquiera de los operadores binarios que simbolizan a la disyunción ( $\vee$ ), conjunción ( $\wedge$ ), implicación ( $\rightarrow$ ), reducción ( $\leftarrow$ ) y equivalencia ( $\leftrightarrow$ ), como por ejemplo la proposición

$$\sim(\text{Cristobal Colón era italiano}) \wedge (\text{España dominó parte de América}) \rightarrow \\ \sim (\text{México está en el continente americano})$$

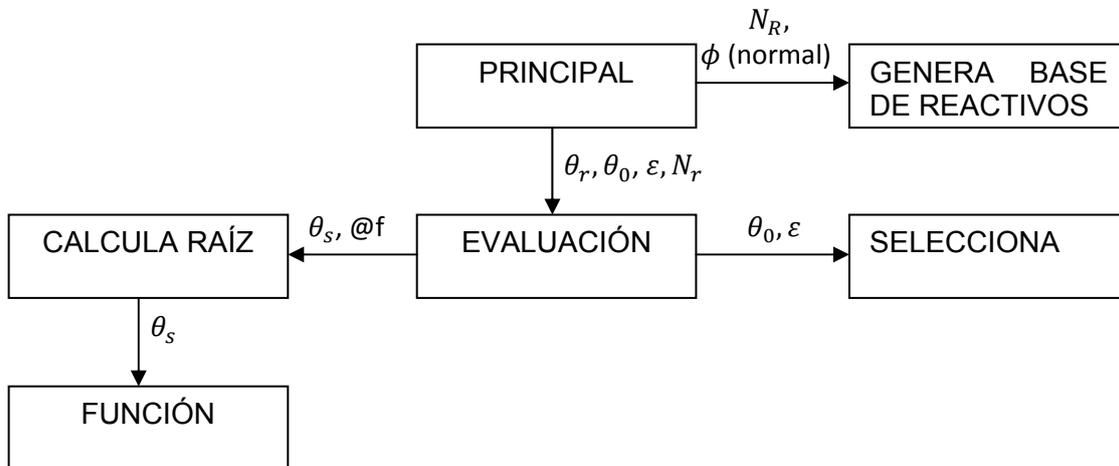
Esta parte del trabajo es también una componente primordial de las actividades de investigación que se realizan dentro del Cuerpo Académico de Computación Inteligente y que sirve como tema de tesis para algunos de sus colaboradores. Finalmente, la sección **Conclusiones** comenta los resultados obtenidos y plantea el trabajo futuro a realizar por los investigadores del Cuerpo Académico de Computación Inteligente en la dirección de generación automática de reactivos y evaluación adaptable en general.

## Evaluación adaptable

La forma en que se realiza el proceso de evaluación adaptable asume que inicialmente se cuenta con una base de reactivos que han sido calibrados; es decir, que se les ha asignado una complejidad. Esta complejidad juega el papel de parámetro dentro del modelo psicométrico que se haya seleccionado. Existen al menos tres modelos psicométricos, los cuales proporcionan la probabilidad de que el evaluado responda correctamente al reactivo de que se trate. Estos modelos son conocidos como logísticos de un solo parámetro, dos parámetros y tres parámetros (1PL, 2PL y 3PL, respectivamente).

La idea del proceso de evaluación consiste en seleccionar el siguiente reactivo con una complejidad que depende de la historia de las respuestas dadas a los reactivos que se han presentado previamente. Para ello se cuenta con un modelo Bayesiano con el cual se construye una función de verosimilitud cuyo máximo determina la complejidad del reactivo siguiente. El proceso continúa de esta forma hasta que se agote la batería de reactivos o hasta que no se note algún cambio en la complejidad del reactivo seleccionado (en realidad, la media de las habilidades o complejidades que se acumulan a lo largo del proceso de evaluación converge al valor de habilidad real, según el Teorema del Límite Central).

**Figura 1** bosqueja el proceso de evaluación adaptable en el que el módulo EVALUACIÓN se repite tantas veces como lo indiquen las condiciones ya especificadas. El módulo FUNCIÓN contiene la condición de maximización de la función de verosimilitud y el módulo SELECCIONA escoge el siguiente reactivo de acuerdo a una condición basada en la idea de información de Fisher. Los argumentos que se requieren en cada módulo también se indican.



**Figura 1.** Esquema del proceso de evaluación adaptable. Se supone que inicialmente se cuenta con una base de reactivos a los que se les ha determinado previamente su complejidad.

Estos argumentos son los siguientes: el número de reactivos en la base de reactivos ( $N_R$ ), la distribución de los mismos ( $\phi$ ), la habilidad real ( $\theta_r$ ) si se simula el proceso de evaluación y ésta se asume desconocida en una evaluación real, la habilidad inicial ( $\theta_0$ ), el radio de selección ( $\epsilon$ ), el número de reactivos a ser presentados máximo en una evaluación ( $N_r$ ), una habilidad semilla ( $\theta_s$ ) para determinar la raíz de la máxima verosimilitud, y la función a maximizar ( $@f$ ). Por ejemplo, en una simulación se obtienen los resultados que indica **Tabla 1**.

Se asume que se tiene que determinar una habilidad real igual a 0.75 e inicialmente se propone una habilidad igual a cero, la cual sirve como semilla para todo el proceso de evaluación. La columna referente a la media de la habilidad estimada señala la convergencia de ésta al valor de habilidad real.

La distribución de los reactivos en la base de reactivos es normal con media 0.5 y desviación estándar 0.2 y se especificó la existencia de 1000 reactivos. La escala de habilidad es de 0.0 a 1.0, aunque se permiten habilidades negativas y mayores que la unidad.

Actualmente se está investigando el impacto que tiene el radio de selección ( $\epsilon$ ) en la tasa de exposición de cada uno de los reactivos. Se piensa que existe un radio que minimiza a esta tasa, indicando con ello que en el diseño del proceso de evaluación este radio debe ser el más indicado.

### Modelo cognitivo en lógica de proposiciones

En evaluación adaptable existen diversos problemas que van desde la administración misma de la evaluación hasta el diseño de reactivos con los que sea posible analizar propiedades cognitivas y psicométricas del evaluado, sin la necesidad de someter estos reactivos a un proceso empírico de calibración previo (Real 1999).

La investigación realizada dentro de algunos campos del conocimiento, como por ejemplo las matemáticas, y específicamente con aritmética y análisis lógico (Real

1999)(Revuelta 1998), ha logrado establecer en un reactivo particular las características apropiadas que indican las facultades cognitivas que se requieren para su solución. Las propiedades cognitivas que se logran determinar se emplean como variables dentro de un modelo componencial que, al mismo tiempo, se incluye dentro de un modelo psicométrico del tipo de Rasch. Se asume que estas propiedades cognitivas contribuyen a la determinación de la complejidad del reactivo.

**Tabla 1.** Resultados de la simulación de un proceso de evaluación adaptable. Se muestra la habilidad real, supuestamente desconocida y a ser determinada por el proceso, así como la forma en que las habilidades calculadas convergen a esta habilidad real después de que se ha presentado un determinado número de reactivos.

Habilidad Real	Examinado	Habilidad Inicial	Radio de búsqueda	Evaluación	Reactivo	Habilidad Estimada	Media de Habilidad Estimada
0.75	2	0	0.1	1	1	0.4181	0.4181
0.75	2	0	0.1	1	2	0.8024	0.6102
0.75	2	0	0.1	1	3	0.5588	0.5931
0.75	2	0	0.1	1	4	1.2402	0.7549
0.75	2	0	0.1	1	5	0.5918	0.7223
0.75	2	0	0.1	1	6	0.5435	0.6925
0.75	2	0	0.1	1	7	-0.4101	0.5349
0.75	2	0	0.1	1	8	1.1033	0.606
0.75	2	0	0.1	1	9	1.2285	0.6752
0.75	2	0	0.1	1	10	1.2441	0.732
0.75	2	0	0.1	1	11	0.2744	0.6904
0.75	2	0	0.1	1	12	1.0011	0.7163
0.75	2	0	0.1	1	13	0.8904	0.7297
0.75	2	0	0.1	1	14	0.6603	0.7248
0.75	2	0	0.1	1	15	0.7926	0.7293
0.75	2	0	0.1	1	16	0.7962	0.7335
0.75	2	0	0.1	1	17	0.6772	0.7301
0.75	2	0	0.1	1	18	1.0764	0.7494
0.75	2	0	0.1	1	19	0.6905	0.7463
0.75	2	0	0.1	1	20	0.3239	0.7252
0.75	2	0	0.1	1	21	0.8638	0.7318
0.75	2	0	0.1	1	22	1.0887	0.748
0.75	2	0	0.1	1	23	1.1232	0.7643
0.75	2	0	0.1	1	24	0.6824	0.7609
0.75	2	0	0.1	1	25	0.3428	0.7442

Nuestro enfoque de estudio es verificar si es posible aplicar un modelo componencial en la generación automática de reactivos, dentro de un evaluador adaptable por computadora, que permita determinar la habilidad que un evaluado tiene para especificar el valor de verdad de una proposición de la forma  $p \blacksquare q \blacksquare r$ , en donde  $p, q$  y  $r$  son proposiciones simples que pueden o no contener negación ( $\sim$ ) y el

símbolo ■ representa a cualquiera de los operadores binarios que simbolizan a la disyunción ( $\vee$ ), conjunción ( $\wedge$ ), implicación ( $\rightarrow$ ), reducción ( $\leftarrow$ ) y equivalencia ( $\leftrightarrow$ ).

El modelo cognitivo de la interpretación de una proposición lógica necesariamente debe especificar los pasos que hipotéticamente debe seguir el sujeto cognitivo para determinar el valor de verdad de la proposición. Esta significación del modelo cognitivo en nuestro análisis particular está en correspondencia con el concepto general de modelo mental, con el cual se pretende explicar, a través del proceso de pensamiento, el comportamiento real del sujeto cognitivo. Bajo este esquema, se asume entonces que todos los sujetos siguen el mismo proceso de pensamiento para solucionar el problema de interpretación de una proposición lógica.

Nuestra propuesta de modelo cognitivo para la interpretación de una proposición lógica está fuertemente influenciada por el modelo cognitivo asociado a la solución de operaciones aritméticas básicas introducido por diferentes investigadores (Real 1999)(Revuelta 1998)(Jáuregui 2002). Resultan particularmente importantes en este modelo aritmético la presencia de operadores en la expresión y el manejo de los signos, aunque se soslaya el concepto de jerarquía de operadores debido a que las expresiones contienen únicamente un operador.

La interpretación de la proposición dependerá de la interpretación que se dé a las proposiciones simples, en caso de que esta interpretación sea errónea, entonces la de la proposición seguramente lo será también. Así que la interpretación correcta de una proposición no solamente depende de los operadores presentes en ella, sino también de las interpretaciones que el examinado dé a las proposiciones simples. La aplicación repetitiva de los siguientes pasos es suficiente para interpretar correctamente la proposición, siempre y cuando la interpretación de las proposiciones simples sea la correcta (Genesereth 2007)

1. Si el valor de verdad de una proposición en una interpretación es verdadero, entonces la negación tiene una interpretación falsa.
2. El valor de verdad de una conjunción es verdadero bajo una interpretación si y solamente si el valor de verdad de sus operandos es también verdadero, de otra forma el valor de verdad es falso.
3. El valor de verdad de la disyunción es verdadero bajo una interpretación si y solamente si uno de sus operandos es verdadero, de otra manera es falso.
4. El valor de verdad de la implicación bajo una interpretación es falso si y solamente si el antecedente es verdadero y el consecuente es falso, de otra forma es verdadero.
5. El valor de verdad de la reducción bajo una interpretación es falso si y solamente si el antecedente es falso y el consecuente es verdadero, de otra forma es verdadero.
6. El valor de verdad de la equivalencia bajo una interpretación es verdadero si y solamente si los operandos son ambos falsos o ambos verdaderos, de otra forma es falso.

A la lista anterior debemos agregar también las prioridades de cada uno de los operadores, teniendo en cuenta que ellas también intervienen en la complejidad de la interpretación de la proposición, así

1. Si interviene una negación sobre una proposición, entonces esta operación tiene prioridad más alta.
2. Si interviene una conjunción, entonces la prioridad de esta operación es menor que aquella de la negación y mayor que la de los otros operadores.
3. Si interviene una disyunción, entonces la prioridad de esta operación es menor que la de la negación y disyunción, pero mayor que la de las otras operaciones.
4. Si interviene una operación de implicación, reducción o equivalencia, entonces estas operaciones tienen menor prioridad que las operaciones de negación,

conjunción y disyunción, pero ellas son igualmente prioritarias realizándose primero entonces el operador que se encuentre a la izquierda del operando que contiene a su izquierda y derecha un par de estos operadores.

Hay que agregar a la lista anterior el hecho de interpretar correctamente una proposición simple, ya que esta interpretación depende en gran medida de los conocimientos que el evaluado tenga acerca de lo que expresa la proposición. Adicionalmente, es necesario introducir paréntesis para considerar la situación en que se niega una proposición compuesta, en cuyo caso se aplica entonces de nuevo el criterio de prioridad máxima de la negación sobre la proposición dentro de los paréntesis. Solamente se emplean paréntesis en esta situación y de ninguna manera en otra donde sea redundante hacerlo.

Ahora bien, dejar abierta la posibilidad de escribir proposiciones de cualquier longitud haría intratable el problema en cierto momento. Por esto mismo, requerimos plantear el formato de los reactivos enfocados a evaluar la habilidad para determinar la interpretación de una proposición. Además de evaluar la forma en que el examinado interpreta a cada uno de los operadores y operandos, estamos interesados en determinar la manera en que comprende la prioridad de los operadores, por lo que es necesario pensar en una estructura que incluya la posibilidad de que al menos dos operadores se encuentren presentes en el reactivo. Al menos dos pero también no más para no hacer más complicado el problema que abordamos.

Si suponemos entonces que  $p$  es una proposición simple con interpretación dada a partir de hechos conocidos o familiares al examinado, entonces el formato de los reactivos tiene una estructura dada por un lenguaje libre de contexto cuyas reglas de producción son las siguientes:

$$\begin{aligned} \text{proposición\_compuesta} &\leftarrow \text{negación\_opcional}\{\text{operador\_binario negación\_opcional}\}_0^2 \\ \text{negación\_opcional} &\leftarrow \text{proposición\_simple}|\sim\text{proposición\_simple} \\ \text{proposición\_simple} &\leftarrow p \\ \text{operador\_binario} &\leftarrow \wedge|\vee|\rightarrow|\leftarrow|\leftrightarrow \end{aligned}$$

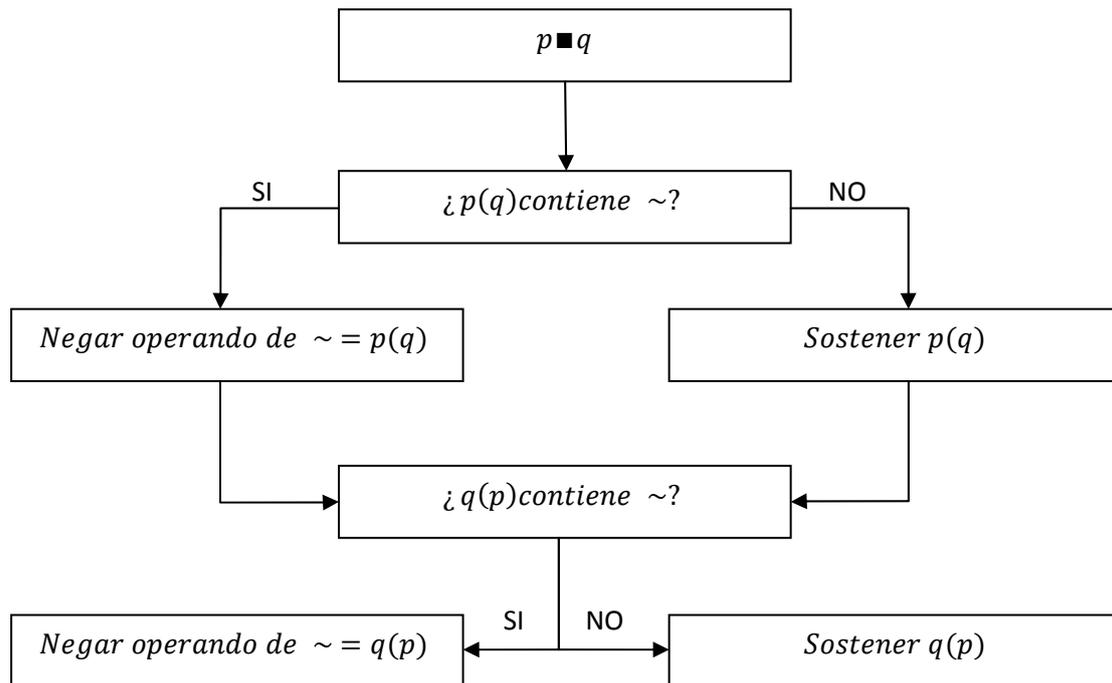
Por ejemplo, la proposición  $\sim p_1 \vee \sim p_2 \leftrightarrow \sim p_3$  es válida en el lenguaje definido por estas reglas de producción, al igual que la proposición  $p_1 \rightarrow \sim p_2 \wedge \sim p_3$ , como puede verificarse en forma relativamente simple. También se incluyen proposiciones más sencillas como  $p$  o bien  $\sim p_1 \leftarrow p_2$ . Pensamos que con esta simple estructura de lenguaje cubrimos el manejo de las prioridades de operadores, operandos e interpretaciones de proposiciones simples y compuestas.

Quedan excluidas las operaciones de distribución de operadores, las cuales pertenecen más al ámbito de la argumentación y requieren además la introducción de símbolos adicionales del lenguaje, como por ejemplo los paréntesis. Un número de iteraciones menor que dos en la primera regla de producción dejaría fuera el manejo de prioridades de operadores, excepto la negación con cualquier otro operador, mientras que un número de iteraciones mayor que dos haría repetitivo el manejo de los conceptos arriba señalados, no aportando con ello mayor información acerca de la habilidad de interés.

La mayor jerarquía de la negación hace imprescindible determinar primero la existencia o no del operador de negación en una proposición en la que además solamente participa uno de los operadores binarios. Si  $p$  y  $q$  son las proposiciones iniciales, negadas o no, entonces la verificación de la presencia del operador de negación sigue como indica **Figura 2**.

Una vez que se ha determinado la interpretación asociada al operador de negación, se procede a interpretar las proposiciones asociadas a los operadores

binarios. **Figura 3** ilustra los pasos que se asume sigue un evaluado al tratar de interpretar correctamente la proposición.



**Figura 2.** Posibles procesos para interpretar la negación de proposiciones que son operandos de algunos de los operadores binarios.

La interpretación de una proposición, sea ésta simple o compuesta, asume la existencia de un modelo psicométrico. Nuestra propuesta de modelo supone que en el proceso intervienen cuatro tipos de conocimientos, suposición que también hacen otros autores cuando se trata de resolver problemas aritméticos con números enteros positivos (Real 1999), que son los siguientes

1. El conocimiento de rasgos distintivos de una situación nos permite diferenciar entre los diferentes operadores proposicionales; es decir, distinguir cuándo no interviene alguno de ellos, cuándo sí lo hace una negación, una conjunción, una disyunción, una implicación, una reducción y una equivalencia. Con ello, se considera que cada operación tiene asociada una dificultad diferente. Por esta razón, la definición del vector de características o rasgos del problema debe indicar la existencia o no de los operadores ya señalados.
2. El conocimiento de las restricciones asociadas a los rasgos distintivos del problema sustenta la verificación de inclusión de la negación en los operandos, y si la proposición que actúa como operando para la negación es verdadera o falsa. Asimismo, el procedimiento de una interpretación varía dependiendo del número de operandos que participan en la proposición, y lo mismo sería en el manejo de las prioridades de los operadores. Consideramos que estas son razones suficientes para que el vector de características incluya componentes que indiquen el grado de conocimiento requerido por el examinado para poder interpretar una proposición simple, cuando esto se requiera, el conocimiento de las prioridades de los operadores participantes, así como el número de operandos.

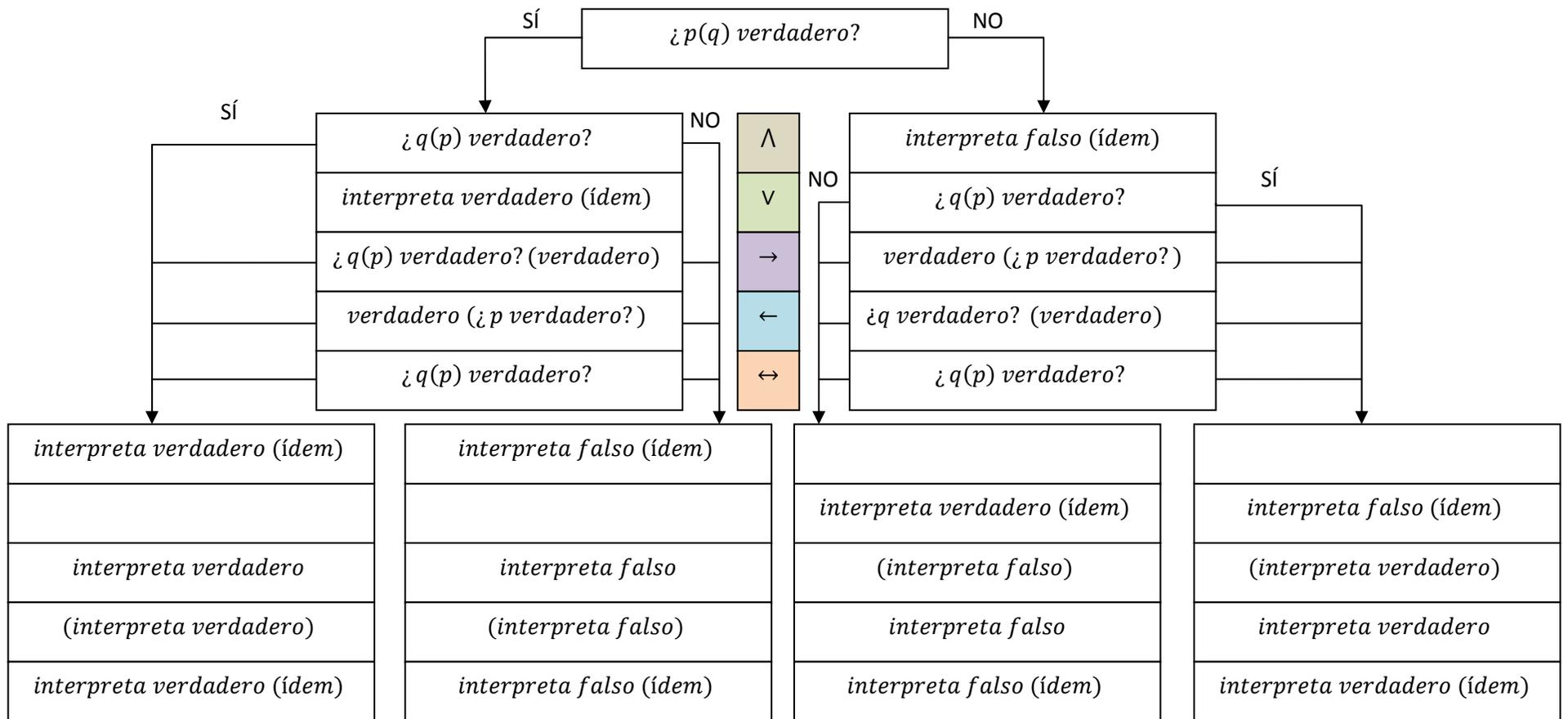
3. El conocimiento relativo a la planificación de la solución del problema sugiere que el estudiante debe ser capaz de identificar qué operadores tienen mayor prioridad y cuáles requieren, por lo tanto, ser negados en caso de que esto se requiera. Asimismo, este conocimiento indica al examinado que primeramente se deben interpretar las proposiciones simples. Obviamente, estos rasgos ya los hemos incluido como componentes del vector de características cuando analizamos los dos primeros tipos de conocimientos.
4. Finalmente, el conocimiento relativo a los algoritmos que definen los pasos que conducen a la solución del problema se encuentra relacionado con el conocimiento previo y tiene que ver con las diferentes etapas de los procesos señalados por **Figura 2** y **Figura 3**.

En resumen, el vector de rasgos o características para cada problema de interpretación de una proposición, tiene dieciocho componentes y ello se indica de la siguiente forma: (**operadores, prioridades, número de operandos, interpretaciones**), en donde cada componente está definida a su vez por componentes binarias, indicando la presencia o ausencia del rasgo distintivo. Así, **operadores** es el vector (**negación, conjunción, disyunción, implicación, reducción, equivalencia**) y un cero o uno en cada componente indica la ausencia o presencia de la operación en la proposición. Análogamente, **prioridades** consiste de cinco componentes, la primera señala la presencia de operandos simples, la segunda de operandos negados, la tercera la presencia de conjunciones, la cuarta la presencia de disyunciones y la quinta de implicaciones, reducciones y equivalencias.

El vector **número de operandos** consiste a su vez de tres componentes, cada una señalando si existen uno, dos o tres operandos, y finalmente el vector **interpretaciones** tiene cuatro componentes, la primera de ellas indicando el nivel de conocimiento lógico hasta el tema de interpretación de proposiciones, la segunda componente señala el nivel de cultura general, la tercera el conocimiento general tomando en cuenta su historial académico, y finalmente la cuarta componente determina el conocimiento específico de un área ajena a la del examinado. Cabe aclarar que, a excepción de **operadores, prioridades** e **interpretaciones** solo una de las componentes en **número de operandos** puede ser no nula e igual a uno.

Como un ejemplo, consideremos los siguientes problemas de interpretación y sus correspondientes vectores de características

1. La proposición  $p$  es dada por “El elemento neón en su estado natural es un gas”. Interprete la proposición  $\sim p$ . En este caso, el vector de características es  
 $(1\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 1\ 0)$ ,  
 indicando que existe un operador de negación cuya prioridad es necesario conocer cuando se aplica a un solo operando que es una proposición simple que requiere conocimiento específico.
2. Las proposiciones  $p$ ,  $q$  y  $r$  son respectivamente dadas por “El metalenguaje es un lenguaje que explica otro lenguaje”, “El elemento neón en su estado natural es un gas” y “Gran Bretaña es miembro de la comunidad europea”. Interprete la proposición  $p \leftarrow q \wedge \sim r$ . Aquí, el vector de características es  
 $(1\ 1\ 0\ 0\ 1\ 0\ 1\ 1\ 1\ 0\ 1\ 0\ 0\ 1\ 0\ 1\ 1\ 1)$ ,  
 el cual señala que en el problema aparecen operadores de negación, conjunción y reducción aplicados sobre tres proposiciones simples y no simples que se construyen con conocimientos sobre lógica, conocimiento de acuerdo al nivel académico y conocimiento específico. Se requiere además que el evaluado conozca acerca de las prioridades de operadores.



**Figura 3.** Proceso seguido en la interpretación de una proposición compuesta en la que participa un operador binario. La proposición en paréntesis es un proceso alternativo que puede ser seguido por el examinado. Se asume que previamente se han interpretado las negaciones en las proposiciones, tal y como lo señala .

La especificación de la matriz de las características de cada reactivo que se emplea en la evaluación se denota como

$$W = \begin{pmatrix} w_{1,1} & w_{1,2} & \cdots & w_{1,j} & \cdots & w_{1,k-1} & w_{1,k} \\ w_{2,1} & w_{2,2} & \cdots & w_{2,j} & \cdots & w_{2,k-1} & w_{2,k} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ w_{i,1} & w_{i,2} & \cdots & w_{i,j} & \cdots & w_{i,k-1} & w_{i,k} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ w_{n-1,1} & w_{n-1,2} & \cdots & w_{n-1,j} & \cdots & w_{n-1,k-1} & w_{n-1,k} \\ w_{n,1} & w_{n,2} & \cdots & w_{n,j} & \cdots & w_{n,k-1} & w_{n,k} \end{pmatrix}$$

donde  $n$  es el número de reactivos y  $k$  el número de características por reactivo. Los valores de  $w_{i,j}$  son binarios, indicando que si en la solución del reactivo  $i$  se emplea la característica  $j$ , entonces  $w_{i,j} = 1$  y en caso contrario  $w_{i,j} = 0$ . No puede ocurrir que alguna fila fuera nula ya que esto implicaría que en la solución del problema no se emplea al menos una de las características que definen la estructura cognitiva.

Cada fila de la matriz  $W$  representa un punto en el espacio  $k$ -dimensional, haciendo entonces posible la definición de clases de reactivos en base a sus características. La especificación de clases se hace en este trabajo a través de la métrica de co-ocurrencias que se define como sigue

$$C = WW^T$$

la cual simplemente mide el número de coincidencias de entradas unitarias que tiene cada reactivo con algún otro en particular, inclusive con él mismo.

En lo que respecta a la similaridad entre la métrica de co-ocurrencia definida por el modelo y la métrica de co-ocurrencias definida por el experimento, la contradicción de la hipótesis nula “ambas matrices no son similares” radica en lo siguiente. Similaridad entre ambas métricas indica que no importando cuál de ellas se emplea para agrupar, el agrupamiento resultante es el mismo. Esto implica la existencia de una alta dependencia entre ambas matrices en lo que respecta a los valores que definen las filas de cada una de ellas; es decir, una alta correlación, la cual puede ser medida en términos del producto de las matrices que definen a las métricas.

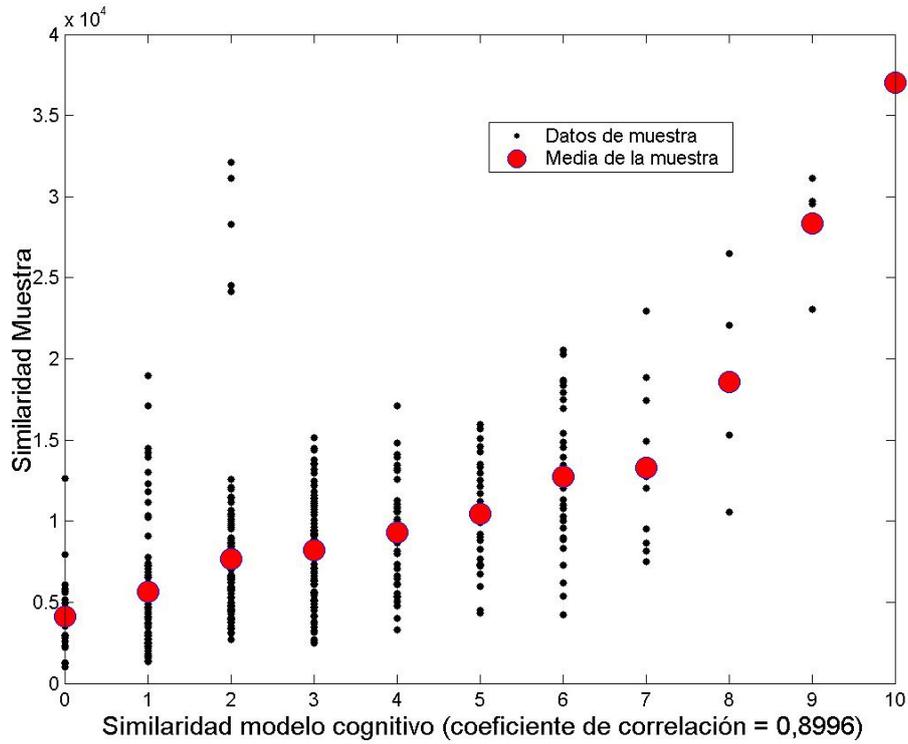
Los resultados de este trabajo muestran que efectivamente existe una correlación aceptable entre ambas métricas. Se realizaron varias muestras en experimentos que involucraron a estudiantes de nivel Técnico Superior Universitario de la Universidad Tecnológica de Xicotepec de Juárez, localizada en el Estado de Puebla, validando con ello en forma parcial el modelo cognitivo propuesto para el tipo de reactivo indicado. **Figura 4** muestra la manera en que se correlacionan ambas métricas.

Un aspecto que también se cuida en la obtención de los resultados es que éstos se distribuyan en forma normal. Para ello, se aplican pruebas de normalidad para dada una de las nubes de puntos experimentales asociados a cada una de las distancias entre reactivos cuyas características son determinadas por medio del modelo cognitivo. **Figura 5** muestra el comportamiento para algunas de estas distancias o proximidades.

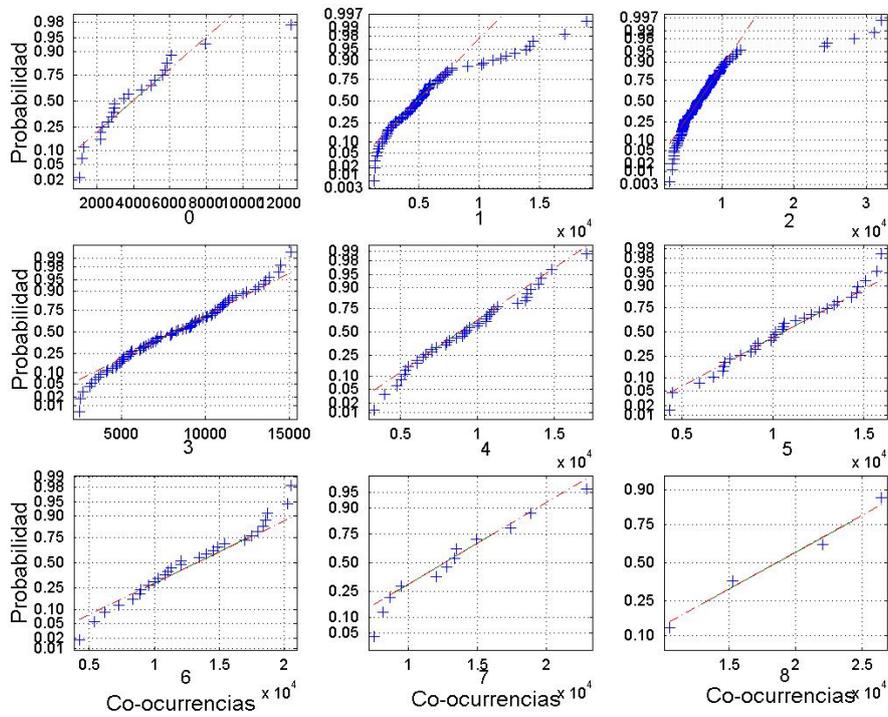
## Conclusiones

Se hace hincapié en la demostración parcial de que el modelo cognitivo propuesto es útil para especificar, en base a componentes, las complejidades de reactivos con los que se miden las habilidades interpretativas de proposiciones lógicas de la forma  $p \blacksquare q \blacksquare r$ . Para completar la prueba es necesario que al menos se determine que existe una correlación entre las complejidades predichas por el modelo y aquéllas obtenidas de manera usual; es decir, calibrando los reactivos como tradicionalmente se hace en evaluación adaptable computarizada (Muñiz 1997). Otro tema a considerar es también el relacionado con la existencia de outliers. Pensamos que

aumentando el tamaño de las muestras es suficiente para eliminar la existencia de estos puntos.



**Figura 4.** Correlación entre proximidades de muestra y modelo cognitivo.



**Figura 5.** Evidencia gráfica de normalidad de co – ocurrencias determinadas por sujetos de muestreo.

Finalmente, queda la consideración de que lo que aquí se muestra se basa principalmente en un análisis estadístico global y que, por ello, se han perdido ciertas especificidades que implicarían más un análisis estadístico con distribución de Bernoulli en lugar de Gaussiana. Existe trabajo en curso con el que se intenta instrumentar en una situación práctica parte de lo que aquí se ha planteado (Suárez 2003)(Suárez Cansino 2008).

## Referencias

**Genesereth, Michael.** «Chapter 2: Propositional Logic.» En *Lecture Notes on Computational Logic at Stanford Logic Group*, editado por Stanford Computer Science Department. Stanford, 2007.

**Jáuregui, L. et al.** «Formulación y validación de un modelo logístico lineal para la tarea de adición y sustracción de fracciones y números mixtos.» *Psicothema* 14, nº 4 (2002): 802 - 809.

**Muñiz, José.** *Introducción a la Teoría de Respuesta a los Ítems*. Madrid: PSICOLOGÍA PIRÁMIDE, 1997.

**Real, Eulogio et al.** «Análisis de la dificultad de un test de matemáticas mediante un modelo componencial.» *Psicológica*, nº 20 (1999): 121- 134.

**Revuelta, Javier y Ponsoda, Vicente.** «Un test adaptativo informatizado de análisis lógico basado en la generación automática de ítems.» *Psicothema* 10, nº 3 (1998): 709-716.

**Suárez Cansino, Joel and Hernández, Roberto.** «Adaptive testing system modeled through fuzzy logic.» En *Modern Topics of Computer Science*, de Alexander and Grebennikov, Alexander Zemliak, 85-89. Acapulco, Mexico: WSEAS Press, 2008.

**Suárez, Joel.** «Student evaluation through membership functions in CAT systems.» *Revista Mexicana de Física* 49, nº 4 (AGOSTO 2003): 371-378.