

Un Nuevo Enfoque para la Puntuación Automática de Problemas cuya Resolución se basa en Diagramas

Josep Soler, Imma Boada, Ferran Prados y Jordi Poch

Universidad de Girona, Dpto. de Informática, Matemática Aplicada y Estadística,
Avda. Lluís Santaló, s/n, 17071 Girona-España. (e-mail: josep.soler@udg.edu)

Recibido May. 17, 2016; Aceptado Jul. 19, 2016; Versión final Sep 16, 2016, Publicado Feb. 2017

Resumen

La corrección automática de diagramas a través de plataformas educativas es de gran importancia, especialmente en el ámbito de los estudios de Ingeniería. El objetivo de este artículo es presentar un nuevo enfoque para la puntuación automática de estos diagramas. Se presenta un nuevo método de forma genérica y su aplicación a la puntuación de diagramas entidad/relación utilizados en el diseño conceptual de bases de datos. Los resultados obtenidos nos indican que las diferencias entre la puntuación obtenida de forma automática y la obtenida a partir de la corrección manual por parte del profesor no exceden en ningún caso de 0.75 puntos sobre un total de diez. Como conclusión, se corrobora la validez del método propuesto para los diagramas entidad/relación facilitando un ahorro considerable de tiempo a los profesores al no tener que corregirlos, ni puntuarlos.

Palabras clave: aprendizaje electrónico; puntuación automática; evaluación automatizada; diagramas; solución de problemas

A New Approach to Automatically Evaluate Problems that are Solved using Diagrams

Abstract

Automatic correction of problems that are solved using diagrams through educational platforms is of great importance, especially in the field of engineering studies. In this paper, we present a new strategy to automatically assess diagrams. The proposed approach is described in detail as well as its application to assess entity/relationship diagrams used in the conceptual design of databases. The results indicate that the differences between manual and automatic assessment is less than 0.75 points over a total of ten which confirms the validity of the proposed approach. As a conclusion, the validity of the proposed method to assess entity/relationship diagrams reducing teacher correction time and unifying applied correction criteria is corroborated.

Keywords: e-learning; automatic marking; automated assessment; diagrams; problem solving

INTRODUCCIÓN

Hoy en día de todos es conocida la importancia de los entornos virtuales de e-learning en ámbitos universitarios. Estos entornos (Moodle, Sakai, Webct, etc.) permiten a los profesores, entre otras cosas, gestionar sus cursos, facilitar el acceso al material educativo, proponer ejercicios a los estudiantes, establecer canales de comunicación y realizar pruebas online y evaluarlas. Estos sistemas se pueden utilizar tanto en estudios on-line no presenciales (Open University en Inglaterra o la UNED y la UOC en España) como en estudios presenciales, donde estos recursos se usan como complemento a las actividades de aprendizaje (blended learning). Un valor añadido de estos entornos es que dispongan de funcionalidades que permitan la generación y asignación de actividades/ejercicios a los alumnos así como la corrección y puntuación totalmente automatizadas.

En este artículo nos centraremos en esta última funcionalidad, la puntuación automática de ejercicios que nos facilite la evaluación del estudiante. Se utilizan varios términos para describir el uso de una computadora para propósitos de evaluación, entre los más utilizados: *Computer-Based Assessment (CBA)*, *Computer-Assisted Assessment (CAA)* y *Online Assessment* (Conole y Warburton, 2005). Aunque estos términos se utilizan de manera intercambiable tienen matices distintos pero con el denominador común que el ordenador juega un papel decisivo en la evaluación del estudiante, usándose para una evaluación formativa, donde el objetivo principal es ayudar al estudiante en el proceso de aprendizaje, y también para una evaluación sumativa, con el objetivo de calificar el nivel del alumno al final de un tema, de una parte o de todo el curso (Thelwall, 2000). Sea de una forma o de otra, la evaluación, es y será una de las claves de los entornos de aprendizaje de este siglo (Scheuermann y Björnsson, 2009).

La tipología de ejercicios que pueden corregirse automáticamente es uno de los puntos clave en este contexto. La mayoría de plataformas de e-learning permiten la corrección de ejercicios de respuesta cerrada o fija como son los ejercicios tipo test, elección múltiple, relleno de espacios en blanco, etc. Ahora bien en un entorno universitario y más concretamente en el de la ingeniería, este tipo de ejercicios no es suficiente para evaluar el aprendizaje de actividades más complejas como por ejemplo el diseño de un circuito eléctrico, un programa de ordenador o un diagrama. En estos casos son necesarios entornos más especializados diseñados específicamente para el aprendizaje de un cierto tipo de materia (Uhomoi bhi y Ross, 2013). En ellos, el estudiante entrará la solución del problema utilizando un entorno determinado que la corregirá (programa de diseño de circuitos, de diagramación, etc.). La corrección automática de este tipo de ejercicios con respuestas totalmente abiertas y no prefijadas es una tarea muy compleja y un tema abierto de investigación (Poch et al., 2016).

Además de la corrección automática de los ejercicios será necesario que estos entornos puedan proporcionar feed-back a los alumnos con el fin de poderse utilizar para una evaluación formativa. De esta forma se usarán como entornos de aprendizaje virtual generalmente vinculados a las actividades presenciales que se desarrollan en el curso. El estudiante realizará los ejercicios propuestos por el profesor, entrará la solución a partir de los conceptos desarrollados en el curso y podrá enviar varias soluciones hasta obtener la solución correcta. En este proceso de aprendizaje, el sistema deberá facilitar pequeñas indicaciones que vayan guiando al usuario hacia la solución correcta. Los estudiantes de hoy en día están muy familiarizados con las tecnologías actuales y agradecen estos sistemas de aprendizaje y evaluación automatizada (Terzis y Economides, 2011).

Si queremos utilizar estos entornos para una evaluación sumativa, éstos deberán disponer de funcionalidades específicas para ello. La mayoría de estos entornos dispone de las funcionalidades necesarias para la puntuación de respuestas cerradas donde el resultado es binario, correcto o incorrecto y por eso resulta muy fácil tanto la corrección como la puntuación posterior. Ahora bien, la puntuación automática de respuestas abiertas de problemas complejos, por ejemplo un diagrama, es una tarea difícil en la que hay que tener en cuenta que la corrección no puede ser binaria (correcto o incorrecto, 0 o 10) ya que en ejercicios muy extensos es comprensible que el alumno cometa pequeños fallos. Por este motivo, no podemos calificar como incorrecto ejercicios con pequeños errores y habrá que calcular la nota en función de estos errores. Además, la nota obtenida por la plataforma de forma automática debe ser lo más parecida posible a la corrección manual realizada por un profesor experto en la materia.

Dada la dificultad de la puntuación automática de ejercicios complejos, en este artículo proponemos un nuevo método de puntuación automática de problemas cuya resolución se basa en un diagrama. La técnica propuesta se ha desarrollado en el marco de la plataforma ACME desarrollada en nuestra Universidad y que permite la corrección y puntuación automática de una amplia variedad de problemas complejos (Soler et al., 2012).

IMPORTANCIA DE LOS DIAGRAMAS

Los diagramas son representaciones gráficas formadas por distintos componentes y las relaciones entre ellos. En función del ámbito de aplicación variará la representación de los componentes y sus relaciones. Un tema importante en la enseñanza universitaria, especialmente en carreras tecnológicas, es la realización de diagramas y esquemas para representar un ámbito de conocimiento. Por ejemplo, en una ingeniería informática se utilizan diagramas de clase, diagramas entidad/relación, etc., en una ingeniería eléctrica se realizan diagramas que representan circuitos eléctricos, en una ingeniería industrial se realizan diagramas para representar los procesos industriales, la planificación de la producción, etc.

Cada diagrama está formado por una serie de componentes conectados entre sí. Por ejemplo, un diagrama entidad/relación, utilizado en el diseño conceptual de bases de datos, está formado por entidades y relaciones. Cada entidad viene determinada por un conjunto de atributos, donde uno o varios formarán la clave principal que identificará la entidad. Las relaciones vienen identificadas por su cardinalidad, las dos entidades relacionadas y posibles atributos de la relación. En un circuito eléctrico disponemos de resistencias, condensadores, transistores, etc. De cada uno de estos componentes solemos tener varias unidades con unas ciertas características.

En el aprendizaje de estas materias, es muy importante que el alumno realice bastantes problemas con la finalidad de poner en práctica los conceptos teóricos expuestos (Banday et al., 2014). La resolución de problemas es una característica clave en el contexto de la ingeniería (Hwang et al., 2014) ya que a menudo es difícil de entender conceptos sólo con clases teóricas. Ahora bien, la corrección de este tipo de ejercicios no es tarea fácil ya que un diagrama puede contener muchos elementos y su presentación totalmente variable. Además en algunos tipos de diagramas podemos tener varias soluciones correctas. Por todos estos factores la corrección automatizada de este tipo de ejercicios de respuesta abierta es compleja.

Así pues, dada la importancia de los diagramas, consideramos oportuno desarrollar módulos específicos para que la plataforma ACME pudiera realizar la corrección de distintos tipos de diagramas. Para cada uno de ellos se desarrolló el editor correspondiente que facilita su dibujo y no permite incoherencias en el diagrama. En una primera fase, y con el objetivo de que estos ejercicios sirvieran para la evaluación formativa del estudiante, se desarrollaron módulos correctores específicos con el objetivo de indicar si los estudiantes los habían realizado correctamente o no, facilitando el feed-back correspondiente en este último caso. Este sistema permite al alumno practicar este tipo de ejercicios y le deja enviar las soluciones que haga falta hasta llegar a la solución correcta (Soler et al., 2012).

Finalizada la primera fase y motivados por su buena acogida, hemos empezado la segunda fase consistente en la puntuación automática de los diagramas para su uso en la evaluación sumativa. Ahora el planteamiento es distinto, se trata de evaluar si el alumno ha adquirido el conocimiento necesario para llevar a cabo una determinada actividad mediante la realización del diagrama correspondiente que el sistema debe puntuar de forma automatizada con una nota. La puntuación automática de estos diagramas presenta varias ventajas. Por un lado el profesor no tendrá que realizar la tediosa labor de corrección, ya que el sistema calculará la nota y ya la registrará en el sistema. Además la corrección de este tipo de ejercicios no es tarea fácil, con lo que la automatización de este proceso supondrá un gran ahorro de tiempo, más si tenemos en cuenta que el número de estudiantes por curso es elevado. Otra ventaja a tener en cuenta cuando hay varios profesores que evalúan las pruebas, es que este sistema es totalmente independiente del profesor ya que aplica los mismos criterios de puntuación a todos los estudiantes. Por otro lado, el estudiante sabrá inmediatamente la nota obtenida.

REVISIÓN BIBLIOGRÁFICA SOBRE PUNTUACIÓN AUTOMÁTICA DE DIAGRAMAS

La mayoría de sistemas de e-learning soportan la corrección y puntuación automática de preguntas con respuesta fija. Este tipo de preguntas se pueden presentar de varias formas, como por ejemplo elección múltiple, relleno de espacios en blanco, escritura de una palabra o de un número, preguntas de tipo test, cualquier combinación de las anteriores, etc. Algunos entornos permiten la generación de distintas pruebas a partir de permutaciones de conjuntos de preguntas o preguntas variables a partir de diferentes valores con el objetivo de individualizar las pruebas, evitando la memorización y el plagio. Estos sistemas permiten calcular la nota obtenida por el estudiante a partir de las respuestas, valorando las correctas y también, en algunos casos, penalizando las incorrectas. El sistema facilita una nota muy precisa, aunque a veces no representa fielmente el nivel de conocimiento adquirido por el alumno ya que en algunos casos descarta por lógica o intuición algunas de las respuestas o bien puede acertarlas por pura casualidad. Las ventajas de este sistema de evaluar a partir de preguntas con respuesta fija son por una parte su sencillez y por otra parte que se pueden utilizar en muchos ámbitos (Jordan, 2013; Afzal et al., 2014).

Ahora bien, la principal limitación de este tipo de cuestionarios basados en respuestas cerradas es que sólo podemos evaluar los niveles cognitivos más bajos de la taxonomía de Bloom (como por ejemplo, conocimiento y comprensión) (Bloom, 1956) pero difícilmente podemos usarlos para evaluar los niveles de aplicación y análisis propios de una enseñanza universitaria (Birenbaum et al., 1992). Para poder evaluar estos niveles de forma automatizada se necesita entornos donde el estudiante pueda escribir, dibujar, expresar libremente su conocimiento de la materia y para ello necesitamos de ejercicios y actividades de respuesta abierta. Entre estos entornos citar los que permiten la corrección de un texto libre (Burrows et al., 2015), los que corrigen programas informáticos (Ihantola et al., 2010) (Caiza y Del Alamo, 2013) (Striwe y Goedicke, 2014) y los que corrigen diagramas.

En este artículo nos centramos en los entornos que puntúan un diagrama dibujado libremente por el estudiante como respuesta a la actividad planteada. No son muchos los trabajos realizados en este contexto. Stone et al. (2009) y Batmaz et al. (2010) proponen un sistema semi-automático que a partir de un diagrama correcto identifica las partes correctas de los diagramas de los estudiantes, marcándolas de color verde. El profesor sólo debe revisar las partes que no se han podido aparear dejando las incorrectas en color rojo y las aceptables en color naranja. Una vez revisados todos los elementos que forman el diagrama se obtiene la nota. Prados et al. (2011) van guardando en la base de datos del sistema los diagramas corregidos manualmente por el profesor junto con su nota. Para cada diagrama a puntuar, se busca en la base de datos si existe uno de igual. Si es así se facilita la nota asignada a este diagrama y en caso contrario se deja pendiente de corrección manual. Los trabajos más relevantes son los de (Thomas et al., 2012, 2013) y (Smith et al., 2013) y otros artículos de los mismos autores que proponen la interpretación y representación de diagramas imprecisos, disponiendo de una herramienta de evaluación basada en reconocer partes del diagrama, que llaman unidades mínimas significativas, que las comparan con las del modelo correcto que disponen. Aunque no especifican detalladamente como calculan la nota, los resultados que presentan en un estudio realizado para la puntuación de diagramas Entidad/Relación, afirman que en el 91,9% de los casos evaluados las diferencias entre la puntuación manual y la automática no exceden de 0,5 puntos sobre 7. Sousa y Leal (2015) proponen un método para evaluar ejercicios en que la solución se pueda transformar en un grafo, como es el caso de diagramas Entidad/Relación o diagramas UML. El método propuesto se basa en la comparación de dos grafos, por un lado el correcto y por el otro lado el del estudiante.

Si nos centramos en la corrección de tipos específicos de diagramas nos encontramos con diferentes entornos capaces de puntuarlos. La mayor parte de trabajos se basan en la puntuación de diagramas entidad/relación utilizados para el diseño conceptual de bases de datos y los diagramas asociados al Lenguaje Unificado de Modelado (UML, por sus siglas en inglés, Unified Modeling Language) como por ejemplo los diagramas de clase, diagramas de caso de uso, diagramas de secuencia, etc. En (Vachharajani y Pareek, 2014) se propone una arquitectura para la evaluación de diagramas de caso de uso. Outair et al., (2014) proponen la evaluación de diagramas de clase a partir de la transformación del diagrama en un grafo y comentan que el siguiente objetivo a realizar es la comparación de este grafo con el propuesto por el profesor y por similitud entre grafos obtendrán la puntuación final. Lackes y Siepermann (2009) utilizan diagramas para representar la planificación de la producción. Estos diagramas realizados por los estudiantes son transformados en un árbol que comparan con el árbol correcto. A partir de esta comparación y teniendo en cuenta distintos tipos de errores calculan la nota.

METODOLOGÍAS PARA LA PUNTUACIÓN AUTOMÁTICA DE DIAGRAMAS

Aunque son pocos los trabajos existentes referentes a la puntuación automática de diagramas, los podemos clasificar según la metodología usada para determinar la nota. Básicamente consideramos tres metodologías: la manual-automática, semi-automática y automática.

La metodología manual-automática, se basa inicialmente en un proceso de corrección manual que a partir de la corrección de muchos diagramas se convierte en un sistema casi totalmente automatizado. La idea de estos sistemas es que van almacenando en su base de datos todos los diagramas junto con la nota que le ha asignado el profesor a partir de la corrección que ha realizado de forma manual. Los sistemas que utilizan esta metodología suelen funcionar de la siguiente manera: de entrada el sistema no guarda ninguna solución correcta del diagrama y es el profesor quien corrige manualmente cada diagrama, le asigna una nota y todo ello se guarda en la bases de datos. Cuando el sistema debe puntuar otro diagrama, compara si en la base de datos existe un diagrama igual. Si es así le asigna la nota guardada. Si en la base de datos no existe ningún diagrama igual pasa al estado de corrección manual. A medida que el profesor va corrigiendo diagramas se van guardando muchos diagramas con su nota y es muy probable que nuevos diagramas a corregir sean iguales a uno de la base de datos con lo que ya tenemos su cualificación. Las ventajas de esta metodología son: i) llega un momento que la puntuación es casi automática y sólo son unos pocos casos en

que la corrección debe ser realizada manualmente, ii) el profesor no tiene que entrar ningún diagrama de referencia en el sistema. Los inconvenientes principales de estas metodologías son: ii) el método utilizado para la comparación de diagramas debe estar muy optimizada para evitar el coste computacional, ya que la base de datos puede llegar a almacenar muchos diagramas de un problema dado; ii) para cada nuevo problema se tienen que volver a corregir muchos diagramas de forma manual. Como ejemplo de referencias que usan esta metodología, citar a Prados et al., 2011). Desde nuestro punto de vista sólo aporta beneficios en aquellos casos en que un mismo problema se utiliza durante varios cursos y lo tienen que resolver muchos estudiantes.

La segunda metodología es semi-automática y se basa en la comparación del diagrama a puntuar con uno o varios de referencia elaborados por el profesor y que son los diagramas correctos al problema propuesto. Para cada problema el sistema guarda el diagrama correcto que es utilizado en el proceso de comparación. Los métodos que utilizan esta metodología siguen dos fases: la primera fase es automática y el sistema identifica las partes coincidentes del diagrama a puntuar con el de referencia. En una segunda fase es el profesor quien revisa los diagramas. En esta revisión el sistema muestra de un determinado color las partes identificadas como correctas y es el profesor quien corrige las partes no identificadas en la fase anterior y que se muestran de distinto color. La nota final la determina el profesor. La ventaja de estos métodos es que el profesor sólo debe fijarse y corregir partes del diagrama. El inconveniente principal es que necesita la supervisión del profesor. Como ejemplo de sistema que usa esta metodología, citar a Stone et al. 2009 y Batmaz et al. 2010.

La tercera metodología es totalmente automática y se basa también en la comparación del diagrama a puntuar con uno o varios de referencia elaborados por el profesor y que son los diagramas correctos al problema propuesto. Esta es la metodología a la que tienden actualmente todos trabajos en puntuación automática de diagramas. Los sistemas que utilizan esta metodología utilizan métodos distintos para la puntuación del diagrama aunque todos ellos se basan en identificar las partes del diagrama que son iguales al de referencia y evalúan si las partes distintas son más o menos equivalentes al de referencia. En función de las partes identificadas como correctas o equivalentes el sistema asigna la nota. Los métodos existentes difieren en como realizan la comparación de los diagramas. Los métodos utilizados en (Thomas et al., 2012, 2013) y (Smith et al., 2013) descomponen el diagrama en unidades mínimas significativas, identificando las que son iguales e intentando comprobar si son equivalentes las demás. Los métodos utilizados por (Sousa y Leal, 2015) y (Outair et al., 2014) se basan en transformar los diagrama en grafos y a partir de algoritmos de similitud de grafos obtener la puntuación. El método que utilizan (Lackes y Siepermann, 2009) consiste en obtener un árbol que compara con el de referencia.

Como valoración final de estos trabajos, comentar que algunos de ellos son sólo propuestas, arquitecturas que presentan cómo tratar el problema pero que no están implementadas. La tendencia actual, tal como se refleja en los artículos más recientes, es seguir la metodología automática para puntuar y suelen basarse en ir sumando la calificación de las partes del diagrama que son exactamente iguales respecto al modelo de referencia y valorar las partes distintas mediante equivalencias con el diagrama de referencia. O bien, transformar el diagrama en un grafo o en un árbol y a partir de aquí compararlo con el de referencia. Generalmente no detallan los métodos o algoritmos utilizados sino que vagamente los describen y se centran más en el análisis de los resultados obtenidos. El método que presentamos se basa también en esta metodología automática pero partiendo de los errores detectados.

MÉTODO PROPUESTO

El método que proponemos en este artículo se basa en la tercera metodología, o sea la puntuación totalmente automatizada a partir de la comparación del diagrama a puntuar con el de referencia que ha introducido previamente el profesor cuando creó el problema. La justificación de este nuevo método se debe básicamente a dos motivos. El primer motivo se basa en la observación de como los profesores realizan la corrección manual. La mayoría de ellos marcan, normalmente con un bolígrafo rojo, los errores que encuentran en el diagrama y en función del tipo de errores detectados ponen la nota. Es lógico pensar que un sistema automatizado siga este mismo proceso. El segundo motivo se debe a que en el proceso de corrección automática de diagramas, los distintos sistemas ya detectan los errores del diagrama para facilitar el feed-back correspondiente al alumno. A partir de ahí, el método propuesto se basa en detectar experimentalmente si hay una correlación entre los errores del diagrama y la puntuación que manualmente realiza el profesor y si la hay obtener la puntuación del diagrama restando de la puntuación máxima un valor en función de los errores cometidos.

Para desarrollar esta idea y antes de poder aplicar el método propuesto se ha establecido el siguiente marco general de trabajo. En primer lugar es necesario que expertos en la materia asignen el valor de cada grupo de elementos del diagrama y determinen los errores que se pueden cometer, cuantificando su peso ya que algunos errores son más graves que otros. En segundo lugar, para que la nota que automáticamente

se obtenga se ajuste más a la realizada de forma manual se introduce el concepto del coeficiente penalizador. Estos coeficientes se obtiene de forma experimental y su función es minimizar los residuos. Para su cálculo proponemos que varios profesores corrijan los mismos diagramas y les asignen una nota. Para cada diagrama consideramos la nota media para atenuar los diferentes criterios de corrección, obteniendo la que llamamos nota final manual. A partir del entorno de corrección, en nuestro caso ACME, se corrigen automáticamente los mismos diagramas corregidos manualmente, obteniendo los elementos de cada grupo que hay en el diagrama y los errores de cada tipo. Con toda esta información y con un software estadístico calculamos mediante una regresión lineal múltiple si existe una correlación entre la nota final manual y los errores detectados, obteniendo los coeficientes penalizadores. Si los resultados de estos coeficientes son aceptables y producen diferencias mínimas podremos utilizar el método propuesto. A continuación detallamos cada uno de los pasos a realizar.

El objetivo principal del método que presentamos es la obtención de una nota entre 0 y max para la evaluación de un diagrama enviado por el estudiante, siendo max la puntuación máxima. Cada diagrama está formado por varios grupos de componentes, por ejemplo entidades y relaciones en un diagrama entidad/relación, resistencias, condensadores, etc. en un circuito eléctrico. El sistema que proponemos se basa en asignar una puntuación P_i a cada grupo de componentes del diagrama de forma que con la suma total de ellos se obtenga la puntuación final máxima. Estas puntuaciones P_i las establece el profesor para cada tipología de problemas. Por ejemplo, si consideramos una nota máxima de 10 puntos, en un diagrama entidad/relación el profesor puede determinar asignar 4 puntos al conjunto de entidades y 6 puntos al conjunto de relaciones. Así pues, si consideramos n grupos de componentes, la suma de los pesos asignados a cada grupo de componentes será la puntuación máxima, tal como se refleja en (1)

$$\sum_{i=1}^n P_i = \max \quad (1)$$

Considerando que cada grupo i de componentes está formado por m_i elementos y considerando que a cada uno de ellos le damos el mismo valor, cada elemento de este grupo tendrá un valor v_i tal como se refleja en la ecuación (2)

$$v_i = P_i/m_i \quad (2)$$

Por ejemplo, si se asigna un peso de 2 puntos al grupo de las resistencias de un circuito eléctrico y la solución correcta tiene 10 resistencias, cada resistencia valdrá 0,2 puntos.

Con el método que proponemos la nota final expresada en la fórmula (3) la calculamos restando de la puntuación máxima un valor en función de los errores detectados. Este valor viene determinado por el producto del valor de cada elemento de grupo v_i , por el número de errores que ha cometido en cada grupo ne_i y por un coeficiente penalizador cp_i que experimentalmente determinaremos para ajustar la nota automatizada con la manual. El sistema obtiene los valores v_i y ne_i de forma automática en el proceso de corrección, a partir de la solución enviada por el estudiante con respecto a la solución correcta más parecida que tenemos guardada en el repositorio de soluciones.

$$Nota = \max - \sum_{i=1}^n (v_i * ne_i * cp_i) \quad (3)$$

En la tabla 1 se muestran, a modo de ejemplo, los parámetros necesarios para el cálculo de la nota de un diagrama formado por n grupos de componentes y como se obtienen.

Con la finalidad de que el método sea utilizable y fiable es necesario que las notas obtenidas de forma automática se ajusten el máximo posible a las obtenidas manualmente por el profesor. En el sistema planteado el profesor es quien decide los valores P_i y para cada ejercicio el sistema calcula el número de elementos m_i de este grupo con lo que ya sabemos el valor de v_i . El sistema también calcula de forma automática cuantos errores ne_i hay en cada grupo de elementos de la solución enviada por el estudiante. Para poder aplicar la fórmula planteada sólo nos falta disponer de los valores cp_i (coeficiente penalizador para cada grupo de componentes). El valor de estos coeficientes debe minimizar la diferencia de notas obtenidas de forma manual frente a las obtenidas de forma automática. Este cálculo de los coeficientes penalizadores se debe realizar una sola vez antes de utilizar el sistema de puntuación automática propuesto.

Tabla 1: Parámetros necesarios para el cálculo automático de la nota

GRUPOS	Valor Grupo	Número de elementos del grupo	Valor de cada elemento grupo	Número errores en el grupo	Coefficiente penalizador
Grupo 1	P_1	m_1	$v_1=P_1/m_1$	ne_1	cp_1
Grupo 2	P_2	m_2	$v_2=P_2/m_2$	ne_2	cp_2
...
...
Grupo n	P_n	m_n	$v_n=P_n/m_n$	ne_n	cp_n
	Fijados por el profesor	Calculados por el sistema	Calculados por el sistema	Calculados por el sistema	Obtenidos experimentalmente

El método propuesto para determinar inicialmente los coeficientes penalizadores es el siguiente: i) se obtiene una muestra de diagramas distintos realizados por los alumnos. Estos diagramas son las respuestas abiertas a diferentes problemas propuestos; ii) de estos diagramas se recoge una muestra considerable y significativa. Una copia de cada uno de estos diagramas es corregida y puntuada de forma manual e independiente por varios profesores expertos en la materia; iii) se considera como puntuación manual final de cada diagrama la nota media de las puntuaciones facilitadas por estos profesores. De esta forma se suavizan las diferencias de puntuación debidas a los criterios específicos de cada profesor; y iv) a partir de esta puntuación manual final de cada diagrama, de los valores de v_i y de ne_i específicos de cada diagrama y calculados por el sistema se realizan los cálculos mediante un sistema de regresión lineal para determinar los valores cp_i que consiguen que la nota automática de cada uno de los diagramas se ajuste al máximo a la puntuación manual final, o sea que las diferencias entre ellas sean mínimas. La figura 1 muestra las diferentes etapas del proceso propuesto para la obtención de los coeficientes penalizadores.

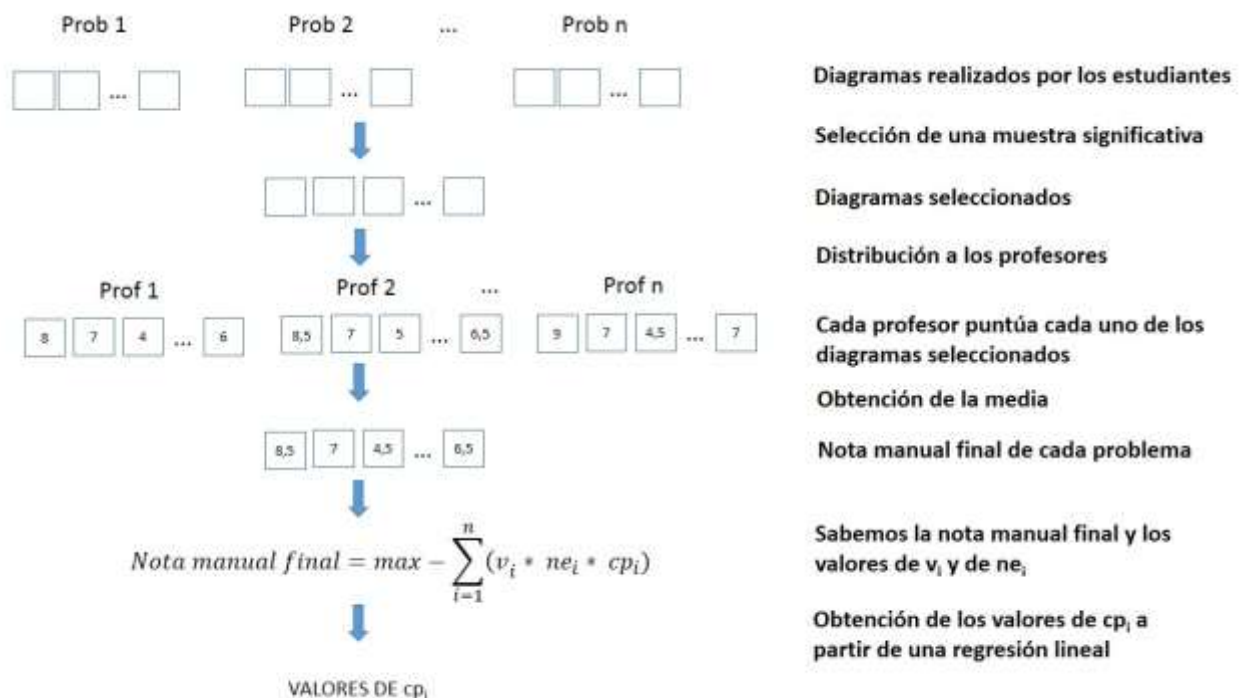


Fig. 1: Proceso seguido para la obtención de los coeficientes penalizadores

Una vez obtenidos los valores de cp_i y para validar el sistema hemos de verificar que la puntuación manual final de los ejercicios y la calculada de forma automática utilizando los coeficientes obtenidos no difieran en más de un punto sobre diez ya que consideramos que diferencias de un 10% son aceptables. Si las diferencias están dentro de este margen, podremos validar el sistema y utilizarlo como método de puntuación automática. A partir de este momento introduciremos los valores obtenidos de cp_i en el sistema que ya estará en disposición de calcular la nota de los diagramas enviados por los estudiantes.

Comparando el método propuesto con los existentes, comentar que nuestro método conlleva un trabajo adicional previo necesario para el cálculo experimental de los coeficientes penalizadores. Ahora bien, una vez obtenidos los coeficientes penalizadores, la gran ventaja que tiene es la rapidez y sencillez del método. La rapidez porque en la corrección automática de diagramas, los distintos entornos existentes ya detectan los errores para facilitar el feed-back al estudiante y simplemente se trata de aplicar la fórmula (3) para calcular la nota. Los otros métodos analizados han de realizar cálculos complejos para determinar la analogía/equivalencia de las partes del diagrama, especialmente de las partes distintas al diagrama de referencia y a partir de ahí estimar la puntuación que les corresponde. En otros métodos hay que transformar el diagrama a un grafo o a un árbol que luego es analizado para evaluar su semejanza con el de referencia.

APLICACIÓN DEL MÉTODO A LOS DIAGRAMAS ENTIDAD/RELACIÓN

Para probar el método hemos considerado la puntuación automática de diagramas entidad/relación utilizados en el aprendizaje del diseño conceptual de bases de datos. Un diagrama entidad/relación es una herramienta muy utilizada en el modelado de datos que permite representar a las entidades de un sistema de información así como las relaciones entre ellas. Los elementos más significativos de estos diagramas son las entidades con sus atributos en las que debemos remarcar el atributo identificador y las relaciones entre estas entidades anotando su cardinalidad (1:1, 1:n, n:m).

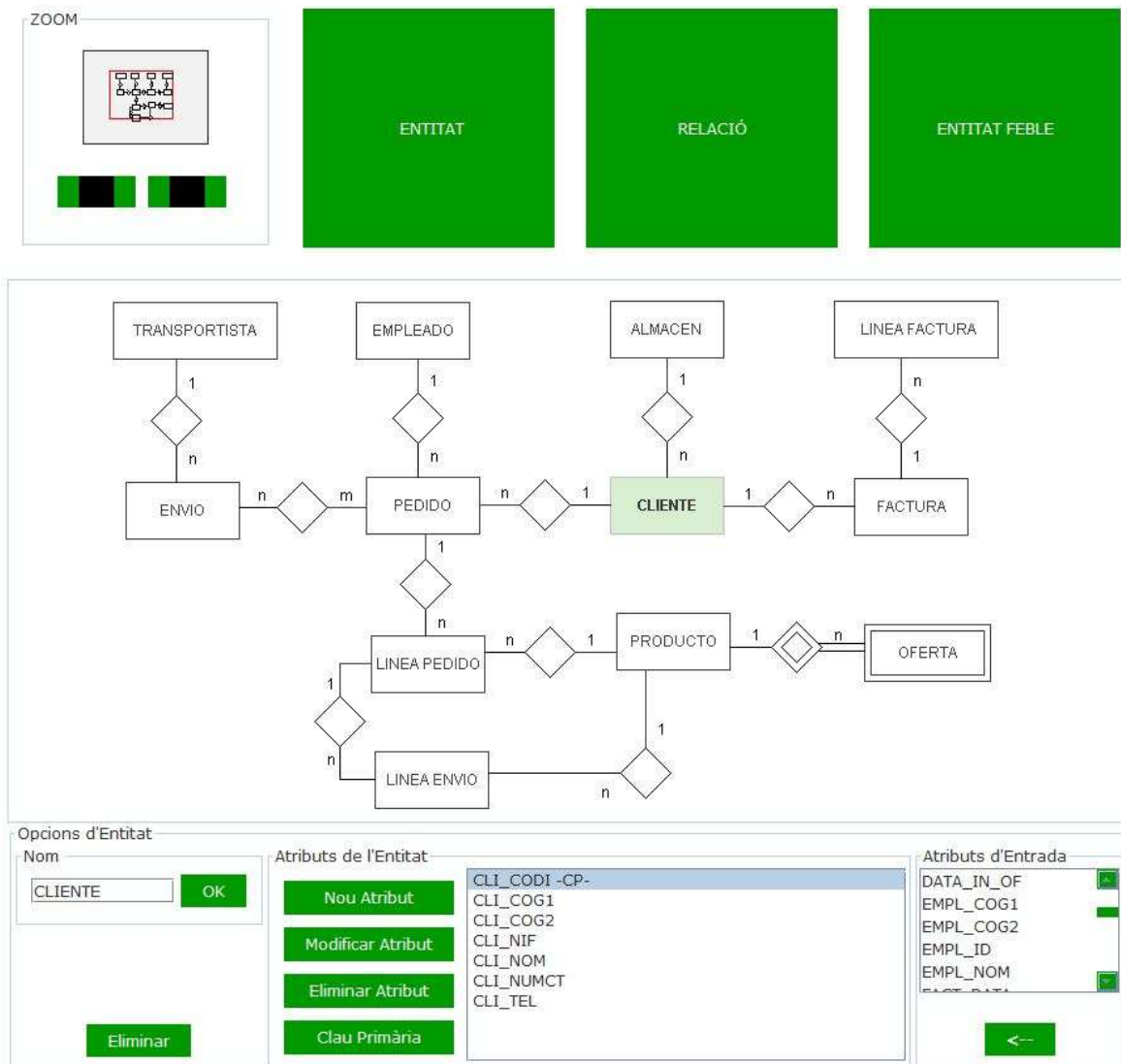
La plataforma ACME dispone de un entorno para la edición y corrección de diagramas entidad/relación en el que el estudiante debe realizar los diagramas correspondientes a los problemas asignados (Soler et al., 2011). Estos problemas y los correspondientes diagramas correctos están almacenados en el repositorio de problemas ACME y han sido creados y revisados por los cinco profesores que han ido impartiendo la asignatura de bases de datos en los últimos cursos. Tal como se muestra en la figura 2 el estudiante selecciona con los botones de la parte superior si quiera dibujar una entidad o una relación en la parte central de la pantalla tal como se muestra en la figura. A continuación, establece que atributos debe tener cada entidad y si es el caso los atributos de la relación, seleccionándolos a partir de la lista de la parte inferior derecha. Finalmente, establece la cardinalidad de cada relación y selecciona cual es el atributo identificador de cada entidad. Una vez completado el diagrama, el estudiante lo envía a corregir. Inmediatamente el sistema corrige el diagrama enviado y devuelve el resultado de la corrección, facilitándole las indicaciones oportunas en caso de que haya errores. La plataforma ACME también dispone del modo "prueba de evaluación", en este caso el sistema sólo admite la entrada de una respuesta, en nuestro caso un diagrama que posteriormente es puntuado de forma manual por un profesor.

Tal como hemos comentado, a partir de este sistema lo que se pretende ahora es establecer un sistema de puntuación automático que nos facilite una nota a partir de un diagrama de este tipo. En primera instancia se consideraron tres grupos de elementos: entidades, relaciones y atributos con pesos P_i de 4 puntos para el grupo de entidades, 4 puntos para las relaciones y 2 puntos para los atributos. Cada entidad, relación o atributo erróneo sería penalizado con un error. A partir de aquí se procedió a calibrar el modelo, así se calcularon los valores c_{pi} que debían minimizar la suma de los cuadrados de las diferencias con respecto a las correcciones manuales. Con esta elección de grupos y pesos los resultados que se obtuvieron no fueron muy buenos, con diferencias superiores a un punto entre las notas calculadas manualmente y las obtenidas por el sistema.

Entonces se pensó que el sistema debía contabilizar el número de errores cometidos de forma más parecida a la forma de actuar de un experto, en la que unos errores tienen más importancia que otros. Para ello establecimos dos grupos de elementos, uno para las entidades y el otro para las relaciones. El peso de cada grupo de elementos P_i es de cinco puntos (cinco puntos para las entidades y cinco puntos para las relaciones). Como ya hemos comentado, el sistema de corrección consiste en comparar la solución enviada por el alumno con las soluciones correctas guardadas en el repositorio. Para determinar los errores de la solución enviada se hace la comparación con la solución correcta más parecida a la enviada por el estudiante. La forma de contabilizar errores ha sido elaborada conjuntamente con los profesores de Bases de datos, expertos docentes y con una amplia experiencia en la corrección de diagramas. Consideran que hay distintos tipos de errores, unos poco importantes y otros más graves. A partir de sus comentarios la contabilización de errores en los dos grupos de componentes que forman un diagrama entidad/relación ha sido la siguiente:

Errores en el grupo de entidades: i) Un error por cada entidad incorrecta. Por entidad incorrecta consideramos entidades que sobran o que faltan y entidades con atributo identificador incorrecto; ii) 0,25 errores por cada entidad identificada correctamente pero en la que hay algún atributo, distinto al identificador, incorrecto. Al ser un error mucho menos grave que el anterior, penaliza mucho menos.

Errores en el grupo de relaciones: i) Un error por cada relación incorrecta. Por relación incorrecta entendemos relaciones que sobran, relaciones que faltan y relaciones entre entidades distintas a la solución correcta; ii) 0,5 errores por cada relación bien identificada pero con los atributos de la relación incorrectos (en caso de haberlos); iii) 0,75 errores por cada relación bien identificada pero con la cardinalidad incorrecta.



Corregir

Fig. 2: Entorno de edición de diagramas entidad/relación de ACME

Con esta nueva distribución de grupos y forma de contabilizar los errores se procedió a calibrar el modelo de nuevo. Para ello se seleccionaron 13 problemas de entre 5 y 10 entidades y entre 5 y 10 relaciones cada uno de ellos. Estos 13 problemas corresponden a los asignados con estas características en las pruebas de evaluación a los estudiantes del curso. En varias sesiones de laboratorio los estudiantes tenían un tiempo limitado para resolver el problema y entrar el diagrama obtenido a la plataforma ACME. Sólo se les permitía entrar un diagrama para cada problema que posteriormente era evaluado manualmente. Para evaluar el método propuesto, de cada uno de estos problemas se han seleccionado 3 diagramas de los enviados por los estudiantes. Los 39 diagramas han sido corregidos de forma manual por tres profesores de nuestro departamento expertos en bases de datos y la nota media obtenida es la que hemos considerado como nota manual final de cada diagrama. De cada problema se optó por escoger un diagrama con pocos errores, uno con bastantes errores y el otro con muchos errores para que el método contemplara todas las posibilidades. Con los valores de v_i , establecida la forma en que el sistema contabiliza los errores ne_i y a partir de estas 39 notas manuales finales, se han calculado los valores de los coeficientes penalizadores cp_i de forma que minimicen las diferencias entre las notas manuales y las calculadas de forma automática. Para el cálculo de los coeficientes se ha utilizado el modelo de regresión lineal comentado. Se ha utilizado la fórmula (3) que adaptada a este caso inicial resulta la fórmula (4).

$$\text{max} - \text{Nota_final_manual} = (\text{ve} * \text{nee} * \text{cpe}) + (\text{vr} * \text{ner} * \text{cpr}) \quad (4)$$

Donde *ve* representa el valor de cada entidad, *nee* representa el número de errores detectado en las entidades, *vr* representa el valor de cada relación, *ner* el número de errores en las relaciones. Todos estos valores nos los facilita el sistema en el proceso de corrección. En nuestro caso hemos asignado el valor 10 a la puntuación máxima. Conocidos todos estos valores sólo nos hace falta determinar el coeficiente de penalización para las entidades *cpe* y el coeficiente de penalización para las relaciones *cpr* mediante una regresión lineal.

Los valores obtenidos para *cpe* y *cpr* son respectivamente 1,45 y 1,62. Estos valores son significativos ya que el p-valor es 0,00. El coeficiente de determinación R^2 es de 0,993 con lo cual el modelo ofrece una calidad muy buena. En la figura 3 se puede ver en primer lugar (gráfico de probabilidad normal) que el modelo obtenido se ajusta a la normalidad. Si observamos los valores de los residuos en el histograma de frecuencias vemos que en ningún caso hay una diferencia superior/inferior a los 0,5 puntos respecto a las notas calculadas manualmente.

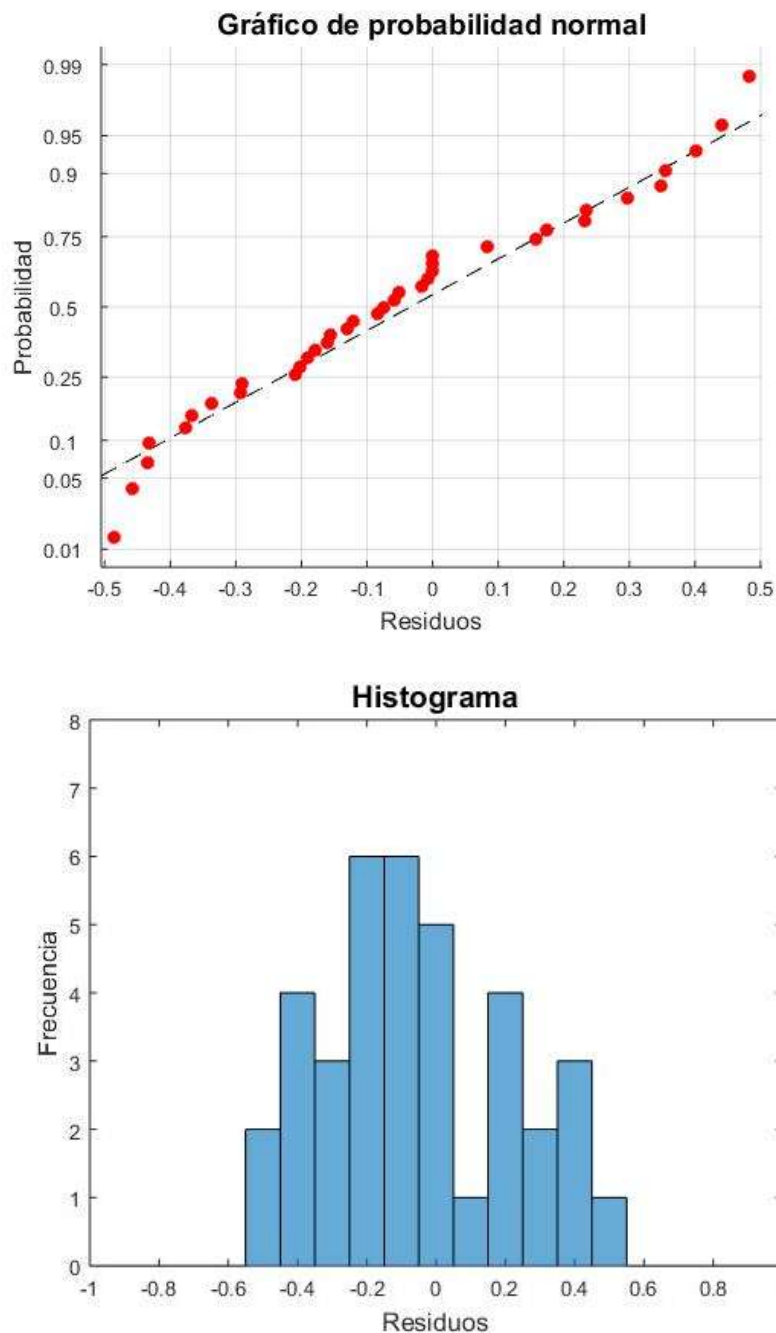


Fig. 3: Gráficas del modelo de puntuación obtenido. Resultados primera prueba

Comentar que para estos cálculos se han omitido las dos notas manuales finales con valores inferiores a 3. En estos casos en que hay muchos errores hemos observado que la nota que asigna el profesor suele estar entre 1 y 3 independientemente de si le correspondería una nota inferior. En estos pocos casos, las notas calculadas eran siempre considerablemente inferiores a la nota final manual y distorsionaban los valores obtenidos. No consideramos importante esta omisión ya que para notas inferiores a 3, el sistema responde simplemente "prueba no superada" sin mostrar la puntuación. Por otra parte vemos que no considerando estos casos, con los coeficientes penalizadores cpe y cpr obtenidos, el modelo se ajusta muy bien al resto de casos que son los que realmente nos interesa evaluar.

Para comprobar si los valores de cpe i cpr variaban al escoger otra muestra de diagramas a puntuar se ha llevado a cabo una segunda prueba. En este caso se han escogido 20 problemas resueltos por estudiantes de cursos anteriores. Estos problemas eran distintos de los de la prueba anterior pero de características y dificultad semejante. De cada uno de estos problemas se han seleccionado tres diagramas al azar. Estos 60 diagramas han sido corregidos también por otros tres profesores, dos de ellos distintos a los de la primera prueba. Igual que antes, hemos considerando la nota manual final como la media de las tres puntuaciones facilitadas. También, igual que la prueba anterior, se han descartado seis diagramas por obtener una nota manual final inferior a 3. Con los 54 restantes se ha realizado la regresión lineal múltiple obteniendo valores parecidos a los de la prueba anterior, concretamente 1,43 para cpe y 1,55 para cpr . Estos valores también son significativos (p -valor=0,0 y $R^2=0,991$). Igual que en la anterior prueba, en la figura 4 se muestran los resultados obtenidos. En este caso, los residuos (diferencia entre la nota calculada con estos coeficientes penalizadores y la nota final manual) son inferiores en 0,5 puntos en 48 casos y sólo en 6 las diferencias están entre 0,5 y 0,75 puntos.

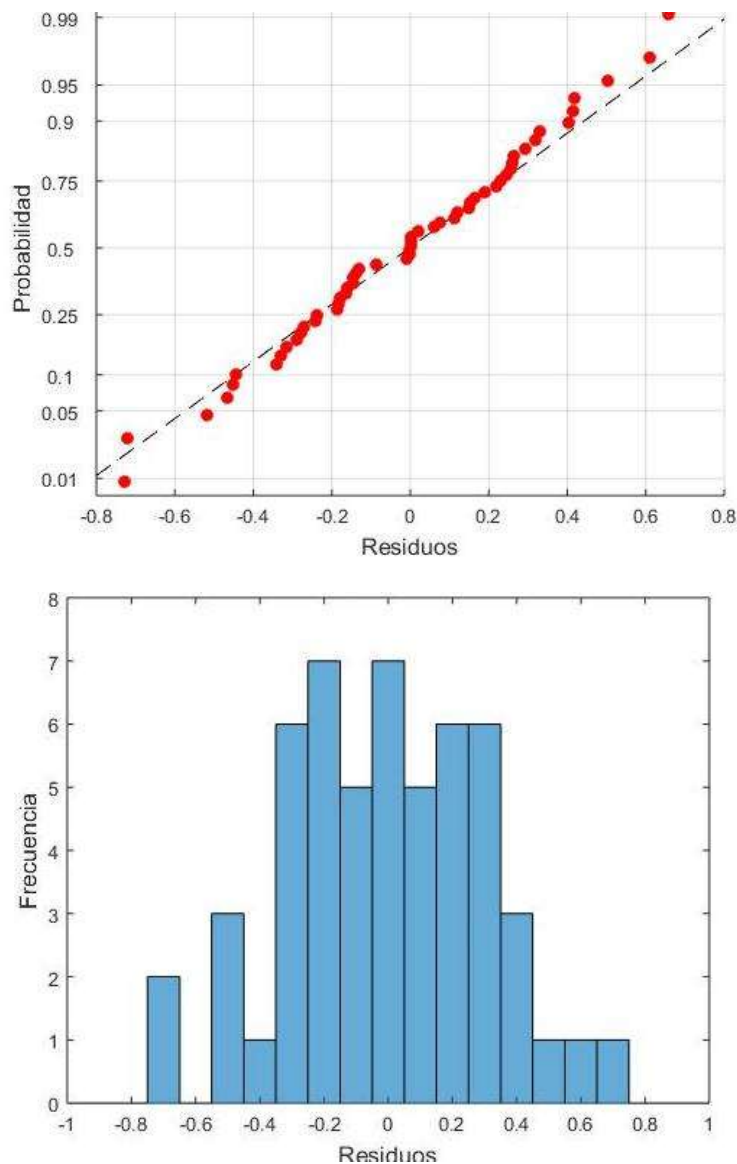


Fig. 4: Gráficas del modelo de puntuación obtenido. Resultados segunda prueba

Finalmente hemos querido comprobar si los distintos valores de cpe y cpr obtenidos en las dos pruebas pueden hacer variar considerablemente los residuos. Para ello hemos calculado cómo se comportarían los residuos de la primera prueba con los coeficientes penalizadores cpe y cpr de la segunda y viceversa. Si en la primera prueba utilizamos los valores de cpe y cpr de la segunda, los residuos continúan estando entre 0 y 0,5 puntos, excepto en dos casos que están entre 0,5 y 0,6 puntos. Si en la segunda prueba utilizamos los valores de cpe y cpr de la primera, los residuos son inferiores en 0,5 excepto en 8 casos con valores entre 0,5 y 0,75 puntos.

Así pues, las pruebas realizadas muestran que el método propuesto se aproxima con un error aceptable a la puntuación del profesor en el caso de los diagramas entidad/relación. Si comparamos estos resultados con los de (Thomas et al., 2012, 2013) que son los más relevantes en la puntuación de diagramas entidad/relación, observamos resultados parecidos ya que ellos sobre una muestra de 394 diagramas obtuvieron diferencias inferiores a 0,5 puntos sobre un total 7 en el 98,7% de los casos y en una segunda prueba de 108 diagramas obtuvieron diferencias inferiores a 0,5 puntos en el 92,3% por cien de los casos. Teniendo en cuenta las ventajas que proporciona la puntuación automática consideramos que el método propuesto se puede implementar en los entornos de corrección automática.

CONCLUSIONES

Hemos propuesto un nuevo método para la puntuación automática de diagramas. Este método se basa en la comparación del diagrama a puntuar con uno o varios de referencia almacenados en la base de datos del sistema. La idea de este método es realizar un proceso análogo al que realizan los profesores en la corrección manual y en función de los errores detectados van restando puntos a la puntuación máxima. En función del tipo de error se restan más o menos puntos ya que no todos los errores tienen la misma importancia.

El método desarrollado se basa en los siguientes puntos que han de realizarse antes de la puesta en funcionamiento: i) expertos en la materia determinan la puntuación a asignar a cada grupo de componentes del diagrama y determinan y cuantifican los errores que se pueden cometer; ii) se selecciona una muestra significativa de diagramas que son corregidos y puntuados manualmente por un grupo de expertos; iii) con las notas facilitadas se calcula la media obteniendo la nota final manual de cada diagrama; iv) en el proceso de corrección el sistema cuenta los componentes de cada grupo y calcula el valor de cada componente; v) en el mismo proceso anterior el sistema detecta los errores de cada grupo de componentes según los criterios fijados; vi) para que la nota que se calcula de forma automática se ajuste con la nota manual final se calculan, a partir de una regresión lineal múltiple, los coeficientes penalizadores de cada grupo; vii) si los valores obtenidos de los coeficientes penalizadores son estadísticamente aceptables ya se pueden introducir en el sistema. A partir de ahí el sistema ya podrá calcular la nota de forma totalmente automática mediante la fórmula (3).

Para probar método lo hemos aplicado al caso particular de los diagramas entidad/relación. Los profesores expertos en bases de datos nos han facilitado el valor de las relaciones y entidades del diagrama y el sistema de cuantificación de los errores. Hemos realizado dos pruebas experimentales y hemos obtenido valores muy parecidos de los coeficientes penalizadores. Con estos valores de los coeficientes penalizadores y aplicando el método propuesto, vemos que las puntuaciones calculadas de forma automática por el sistema y las manuales corregidas por el profesor son inferiores a 0,75 puntos sobre 10 en todos los casos, con lo que el sistema propuesto es aceptable para esta labor. Las ventajas que ofrecen estos sistemas de puntuación automática son: i) el estudiante sabe inmediatamente la nota, ii) el profesor se ahorra mucho tiempo en las tareas de corrección, iii) la corrección se realiza con los mismos criterios para todos los estudiantes.

Como trabajo futuro queremos plantear para el próximo curso la sustitución de las pruebas de diseño conceptual de bases de datos, de forma que la corrección ya sea automática. Por otra parte, al disponer también de un corrector de diagramas de clase, queremos evaluar el método propuesto en este artículo para este tipo de diagramas y evaluar si los resultados obtenidos son también aceptables.

AGRADECIMIENTOS

Este trabajo ha sido financiado en parte por subvenciones del Gobierno español (Nr. TIN2013-47276-C6-1-R) y del Gobierno catalán (Nr. 2014-SGR-1232).

REFERENCIAS

Afzal N. y R. Mitkov, *Automatic Generation of multiple choice questions using dependency-based semantic relations*, *Soft Computing* 18, 1269-1281 (2014)

- Banday M.T., M. Ahmed y T.R. Jan, *Applications of e-learning in engineering education: A case study*, Procedia - Social and Behavioral Sciences 123, 406-413 (2014)
- Batmaz F., R. Stone y C. Hinde, *Personalised feedback with semi-automated assessment tool for conceptual database model*, Innovation in Teaching and Learning in Information and Computer Sciences, 9(1), 105-109 (2010)
- Birenbaum M., K. Tatsuoka y Y. Gutvitz, *Effects of response format on diagnostic assessment of scholastic achievement*, Applied Psychological Measurement 16(4), 353-363 (1992)
- Bloom B.S., *Taxonomy of educational objectives*, David McKay Co Inc, New York, (1956)
- Burrows S., I. Gurevych y B. Stein, *The Eras and Trends of Automatic Short Answer Grading*. International Journal of Artificial Intelligence in Education 25, 60-117 (2015)
- Caiza J.C. y J.M. Del Alamo, *Programming assignments automatic grading: review of tools and implementations*, 7º International Technology, Education and Development Conference, Valencia, España, del 4 al 5 de Marzo, 5691-5700 (2013)
- Conole G. y B. Warburton, *A review of computer-assisted assessment*, Research in Learning Technology 13(1), 17-31 (2005)
- Hwang G., F. Kuo, N. Chen y H. Ho, *Effects of an integrated concept mapping and web-based problem-solving approach on students' learning achievements, perceptions and cognitive loads*, Computers & Education 71, 77-86 (2014)
- Ihantola P., T. Ahoniemi, V. Karavirta y O. Seppälä, *Review of recent systems for automatic assessment of programming assignments*, 10º Koli Calling International Conference on Computing Education Research. Koli, Finlandia, del 28 al 31 de Octubre, 86-93 (2010)
- Jordan, S., *E-assessment: past, present and future*, New Directions, 9(1), 87-106 (2013)
- Lackes R. y M. Siepermann, *Automatically Generating and Marking Calculation Exercises in e-learning*, International Conference on Multimedia Computing and Systems, Ouarzazate, Marruecos, del 2 al 4 de abril, 205-210 (2009)
- Outair A., A. Lyhyaoui y M. Tanana, *Towards an automatic evaluation of UML class diagrams by graph transformation*, International Journal of Computer Applications, 95 (21), 36-41 (2014)
- Poch J., I. Boada, J. Soler y F. Prados, *Automatic Creation and Correction of Mathematical Problems*, International Journal of Engineering Education, 32(1), 150-162 (2016)
- Prados F., J. Soler, I. Boada, J. Poch, *An Automatic Correction Tool That Can Learn*. IEEE Frontiers in Education Conference (FIE), Rapid City (USA), Octubre 2011
- Scheuermann F. y J. Björnsson, *JRC Scientific and Technical Reports: The Transition to Computer-Based Assessment - New Approaches to Skills Assessment and Implications for Large-scale Testing*, (2009)
- Smith N, P. Thomas y Waugh K, *Automatic Grading of Free-Form Diagrams with Label Hypernymy*, Learning and Teaching in Computing and Engineering, 136-142 (2013)
- Soler J., I. Boada, F. Prados, J. Poch y R. Fabregat, *A Web-Based E-learning Tool for Database Design Courses*, International Journal of Engineering Education, 27(1), 61-69 (2011)
- Soler J., F. Prados, J. Poch y I. Boada, *ACME: plataforma de aprendizaje electrónico (e-learning) con funcionalidades deseables en el ámbito de la ingeniería*, Formación Universitaria, 5(3), 3-16 (2012)
- Sousa R. y J.P. Leal, *A Structural Approach to Assess Graph-Based Exercises*. Languages, Applications and Technologies. Vol 563 of the series Communications in Computer and Information Science, 182-193, (2015)
- Stone R., F. Batmaz y C. Hinde, *Drawing and Marking Graph Diagrams*, Innovation in Teaching and Learning in Information and Computer Sciences 8(2), 45-52 (2009)

Striewe M. y M. Goedicke, *A review of static analysis approaches for programming exercises*, Computer Assisted Assessment. Research into E-Assessment, 439 of the series Communications in Computer and Information Science, 100-113, (2014)

Terzis V. y A. Economides, *The acceptance and use of computer based assessment*, Computers & Education 56, 1032-1044 (2011)

Theilwall M., *Compute-based assessment: a versatile educational tool*, Computers & Education, 34, 37-49 (2000)

Thomas P, N. Smith, Waugh K, *Automatically assessing free-form diagrams in E-assessment systems*. STEM Annual Conference, 1-8 (2012)

Thomas P. *Online Automatic marking of diagrams*, Systemic Practice and Action Research, 26(4), 349-359 (2013)

Uhomoibhi J. y M. Ross, *Trends in the development of e-learning in Engineering and Computing Education*, International Journal of Engineering Pedagogy, 3 (2), 26-29 (2013)

Vachharajani V. y J. Pareek, *A Proposed Architecture for Automated Assessment of Use Case Diagrams*, International Journal of Computer Applications, 108(4), 35-40 (2014)