

Exámenes con corrección automática para la mejora de los sistemas de evaluación

Martínez-Castro, Alejandro E.

(1) Departamento de Mecánica de Estructuras e Ingeniería Hidráulica. Universidad de Granada, amcastro@ugr.es

Resumen

Se presenta un sistema de exámenes escritos basado en test de marcas con justificación de respuestas. La digitalización se basa en un escáner convencional, y en algoritmos de Visión por Computador. Los exámenes corregidos se envían a los alumnos de forma automática y personalizada. El sistema presenta importantes ventajas, tanto para el alumno como para el profesor.

Palabras clave: Exámenes, marcas, evaluación, visión por computador.

I. Introducción

En los sistemas docentes presenciales, la realización de pruebas escritas constituye un recurso fundamental que permite medir, de forma objetiva, los conocimientos y competencias adquiridos por el alumnado. Dentro de los sistemas de evaluación continua, y para garantizar el adecuado rendimiento del alumnado, el profesor debe, como tarea fundamental de la labor docente, elaborar un seguimiento individual y objetivo que permita al alumno conocer en todo momento su evolución. Esto permite incentivar el proceso de aprendizaje continuo, hasta llegar a superar los límites establecidos en la evaluación de competencias.

La elaboración de pruebas escritas presenta problemas cuando se realizan de forma tradicional exclusivamente: los tiempos requeridos para preparar las pruebas de evaluación, corregirlas, comunicar los resultados a los estudiantes, y revisar la corrección de las mismas, acaban haciendo inviable la posibilidad de realizar varias pruebas en un semestre. Es necesario por tanto un sistema que permita que el alumno tenga su corrección en un tiempo rápido, que reciba de forma individual la información detallada de la corrección de su prueba, y que pueda así modificar los conceptos mal adquiridos en aras a un aprendizaje continuo. Esto además permite generar una motivación positiva en el alcance de metas y logros (KOTSIANTIS, 2010).

Durante los últimos años se han desarrollado diferentes iniciativas que ayudan a reducir de forma significativa los tiempos de evaluación, corrección y comunicación. En general, los intentos llevados a cabo se han basado en exámenes con corrección automática basados en reconocimiento óptico de marcas (Optical Mark Recognition, OMR en lo que sigue). En estos casos, se requiere de un software y hardware específico, ajeno al profesor, para la corrección de las pruebas. El sistema es rígido, ya que sólo permite que el alumno escriba marcas a lápiz. Dentro de la Universidad de Granada, existe este servicio, el cual genera un coste adicional para los usuarios.

En los últimos años se han desarrollado alternativas a los sistemas clásicos de OMR, incorporando software de Visión por Computador, para la fase de reconocimiento de marcas impresas (QUEROL, 2014). La principal ventaja de estos sistemas es que no se necesitan dispositivos especializados en lectura de marcas. Estos sistemas pueden utilizarse con software libre, y con equipos disponibles en los departamentos (escáner multipáginas). Partiendo de exámenes de marcas escaneados, estos sistemas son capaces de reconocer las zonas marcadas y evaluar en base a unas reglas definidas por el usuario. El profesor puede diseñar directamente las pruebas, y plantear diferentes estrategias de verificación. Dentro de esta línea, aparece software libre, como SDAPS, QueXF, o Auto-Multiple-Choice.

En esta comunicación se presenta una experiencia llevada a cabo en diferentes asignaturas del departamento de Mecánica de Estructuras e Ingeniería Hidráulica, para las cuales, se ha utilizado la herramienta Auto-Multiple-Choice para generar un sistema de pruebas automáticas de evaluación. Se han desarrollado asimismo estrategias que permiten incorporar y evaluar partes escritas. El sistema se ha enriquecido con un conjunto de utilidades, basadas en Python, que permiten enviar el examen a cada estudiante, o proporcionarle su código de

identificación personal. La experiencia se ha llevado ya a cabo en dos cursos académicos, con gran satisfacción tanto para el profesorado como para el alumnado.

II. Metodología propuesta.

En esta experiencia llevada a cabo se han evaluado diferentes programas disponibles. Dentro de ellos, se ha elegido el software *Auto-Multiple-Choice* (AMC en lo que sigue). AMC ha sido desarrollado por Alexis Bienvenüe, y está disponible bajo licencia GPLv2+ para libre descarga y uso. Su instalación debe hacerse en sistema Linux (nativo o virtualizado desde Windows) Este código está disponible bajo sistema Linux. Sus características pueden resumirse en los siguientes puntos:

- Reconocimiento del alumnado en base a un código de identificación personal.
- Posibilidad de realizar el documento en LaTeX.
- Admite diferentes tipos de preguntas, incluidas las preguntas de evaluación directa de partes escritas (preguntas abiertas), donde es el profesor el que marca una calificación tras la lectura de lo escrito por el alumno en zonas habilitadas.
- Posibilidad de generar preguntas diferentes para cada alumno, seleccionadas dentro de un banco de preguntas, clasificadas por categorías.
- Posibilidad de modificar el orden de las preguntas en cada examen, así como de barajar las posiciones de las respuestas.
- Los ejercicios se imprimen con anotaciones, para facilitar la revisión individual de cada examen.
- Generación de hoja de cálculo con resultados individualizados y estadísticas del grupo de evaluación.

Para poder implementar el sistema completo, ha sido necesario además desarrollar dos códigos propios, implementados en el lenguaje Python.

- Código encargado de comunicar a cada estudiante su código CID personal.
- Código encargado de enviar a cada estudiante su examen corregido.

El código AMC está planteado para que la respuesta se establezca en base a marcas que el alumno rellena con bolígrafo. El código incorpora diversas opciones para puntuar cada respuesta, e incluso de penalizar. No obstante, para algunas asignaturas, se ha valorado como necesario que las respuestas marcadas sean justificadas por parte del alumno, acercando así el tipo de examen a uno tradicional, donde el planteamiento de los problemas sea o pueda ser también tenido en cuenta. Esto puede ser un requisito totalmente opcional, a disposición del profesorado. Para esto, se han generado ventanas donde el alumno debe escribir, como ocurre con un examen tradicional, el razonamiento que le lleva a determinar la solución de las casillas marcadas. Esto hace que el profesorado tenga que revisar algunas de las preguntas, como paso previo al escaneo final de los ejercicios y marcas.

La implantación de este sistema requiere que el estudiante conozca y utilice durante el curso este sistema, en pruebas cortas, de menor peso académico. De esta forma, en pruebas de mayor relevancia, el estudiante conocerá perfectamente cómo funciona el sistema, y cómo debe enfocar estos exámenes.

A continuación se describen las diferentes etapas.

II.1 Comunicación del código de identificación personal al alumnado.

El primer paso consiste en asignar a cada estudiante un código de identificación personal (CID en lo que sigue). Por cuestiones prácticas, este código se ha planteado a partir del propio DNI del estudiante. Al principio de curso, el profesor deberá disponer de una lista con el nombre, apellidos, DNI y email de sus estudiantes. Para estudiantes extranjeros es necesario generar un código CID basado en el número disponible de identificación para ellos.

Para comunicar a cada estudiante su código CID, se ha desarrollado una herramienta basada en Python. El lenguaje Python permite, de forma muy sencilla, realizar tareas cotidianas y repetitivas. En este caso, el script utiliza las librerías *smtplib* y *email*. De forma sencilla, el script genera un mensaje personalizado para cada alumno, informándole de sus datos personales y asignándole su código CID. La Figura 1 muestra un ejemplo impreso, aportado por un estudiante en un examen, de correo tipo que recibe el estudiante.

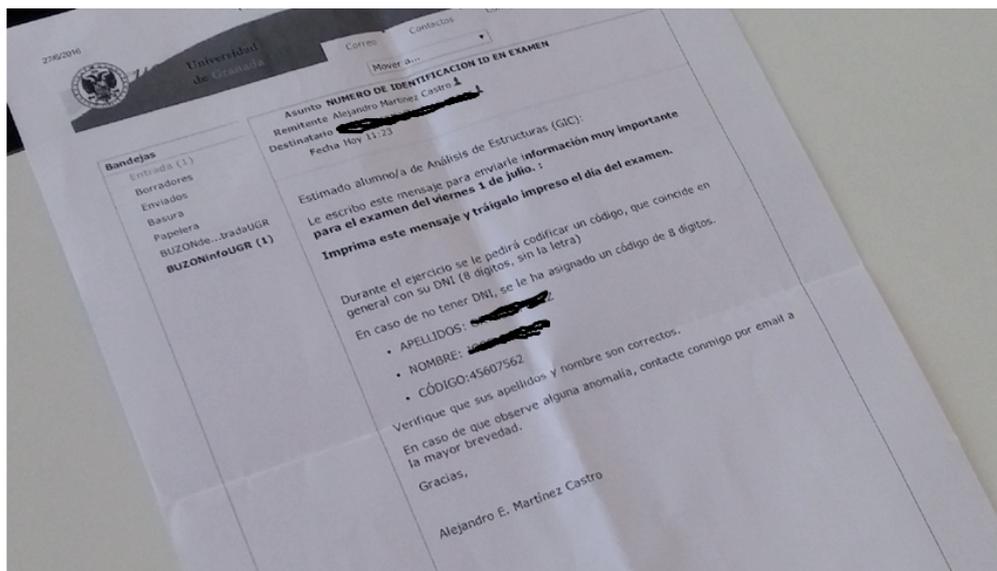


Figura 1: Ejemplo de correo personalizado enviado a estudiante informándole de su código de identificación personal

II.2 Preparación del examen.

La preparación del examen se realiza mediante un documento escrito en código LaTeX. Aunque AMC permite preparar los exámenes desde un fichero de texto, la preparación en código LaTeX permite un mayor control sobre las preguntas, así como activar otras ventajas conocidas de los documentos de LaTeX. Al instalar el código AMC, se instala un paquete de LaTeX, denominado *automultiplechoice*.

El programa principal requiere utilizar un entorno Linux. En general, se ha verificado sobre diferentes versiones de Ubuntu y Debian. Para Ubuntu, el paquete se instala de forma muy simple, ya que es oficialmente un paquete de la distribución (se instala con la utilidad *apt-get install*). La posibilidad de instalarlo en entorno Windows es posible, mediante una virtualización basada en Virtual Box de Oracle. Por tanto, no es una limitación que el entorno principal del equipo donde se esté trabajando esté en un sistema operativo diferente de Linux.

La documentación disponible en la página principal de AMC describe bien los diferentes tipos de preguntas. El propio programa aporta ejemplos donde se clarifican los sistemas de puntuación y las alternativas de tipos de preguntas. Dentro de los tipos de preguntas, cabe destacar algunos aspectos:

- El uso del paquete *fp* permite generar preguntas con respuesta numérica calculada, pero basada en números aleatorios. Este tipo de preguntas permiten presentar un problema para el cual, los datos de partida son diferentes para cada alumno, y la respuesta se obtiene mediante una fórmula que es programada.
- El uso de preguntas de respuesta simple o múltiple, con opciones de puntuación dependiente y diferenciada. Por ejemplo, es posible generar varias respuestas, y que la calificación de la pregunta dependa de qué respuestas se activen (llegándose en algún caso a la anulación total de la respuesta si el alumno marca alguna respuesta particularmente mal valorada).
- Preguntas abiertas. En este tipo de preguntas, se genera una escala de puntuaciones, y un espacio libre para que el alumno responda por escrito. En este caso, el alumno no marca sobre ninguna casilla, sino que es el profesor el que, tras leer lo que el alumno escribe, asigna la puntuación, marcando la casilla de la escala de puntuaciones.

Al elaborar el documento, se definen los bancos de preguntas, mediante etiquetas. En las últimas líneas, se define el examen-tipo, incluyéndose el número de copias que se generarán. El código genera copias individuales, todas ellas diferentes en cuanto a orden de preguntas, posiciones de respuestas, números aleatorios, etc. El control de cada examen es posible mediante un código binario escrito en la parte superior.

II.3 Impresión de los ejercicios individuales.

Los ejercicios individuales se deben imprimir en impresora. Deben imprimirse tantas copias como exámenes individuales se deseen generar. Se han realizado pruebas con exámenes a doble cara o a simple cara. Los mejores resultados se obtienen con exámenes a simple cara, aunque con la debida preparación, es posible diseñar exámenes a doble cara, evitando solapamientos en las zonas de marcas. La elección del formato de impresión depende en gran medida del tipo de impresora y escáner en el que se vaya a trabajar.

II.4 Desarrollo de las pruebas presenciales.

En el día de la prueba oficial, el alumno va a encontrar diferencias con los sistemas tradicionales. Por este motivo, se ha planteado en esta experiencia una implantación paulatina, realizándose pruebas cortas en clase (de 15 minutos), que permiten que el alumno se familiarice con este nuevo tipo de examen. Gracias a esto, se ha podido implementar este sistema en exámenes oficiales.

La realización de la prueba (Figura 2), se asemeja a una prueba tradicional, con las siguientes particularidades.

- Cada estudiante recibe todo el examen completo, o parte del examen. Típicamente, el estudiante recibe impreso un examen único, diferente del de todos sus compañeros, pero parecido en su temática y problemas planteados.
- El alumno sólo escribe su nombre una vez, aunque reciba varios folios. El resto de folios quedan ligados al primero a través de los códigos binarios que aparecen en la parte superior.
- El estudiante debe venir provisto de bolígrafo negro (o azul oscuro), y cinta correctora.
- El estudiante conoce su código CID, el cual debe marcar en la primera hoja.
- Para contestar a las preguntas, el estudiante debe trabajar en borrador (en hojas aparte), y transcribir la justificación de las respuestas, si así se le requiere. En otros casos se podrá optar, si se desea, por recoger los folios de justificación (aparte), o por penalizar respuestas negativas.
- Puesto que cada estudiante tiene un documento único, se observa que cada estudiante está concentrado en su trabajo, observándose menores incidencias en intentos de copia o fraude.

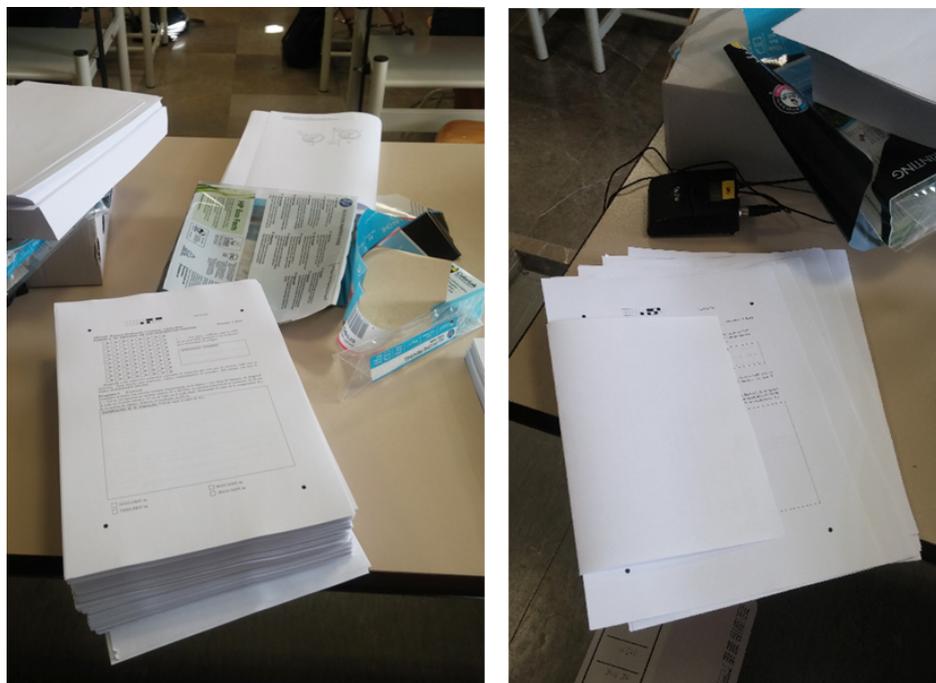


Figura 2: Grupo de exámenes impresos (izquierda) y examen individual para cada alumno

II.5 Corrección de las pruebas. Revisión del profesor.

Tras la realización del examen, el profesor debe revisar determinadas partes de los ejercicios, si así lo ha planteado. En el contexto de asignaturas de mecánica estructural, esta revisión afecta a las preguntas en las que se pide justificación expresa. En caso de encontrarse errores en la justificación, el profesor tiene dos mecanismos básicos para anular una pregunta: por un lado, puede marcar cualquier otra, en otro color (por

ejemplo, verde). De esta forma, el sistema lo va a interpretar como respuesta doble, y procederá a anularla. Otro recurso consiste en introducir una casilla especial, con un texto tipo “pregunta eliminada por el profesor”. Asimismo, en preguntas de tipo “abierto”, el profesor deberá marcar la nota asignada, en función de la evaluación de las partes escritas.

La intervención del profesor en fase de corrección es totalmente opcional, ya que el profesor puede optar por un examen de tipo test, en el cual no se revisan las preguntas, o no hay preguntas de tipo “abierto”. Por ejemplo, en una asignatura como *Puentes* del Máster de Ingeniería de Caminos, Canales y Puertos, se ha optado por realizar un test final de 20 preguntas (con penalización por respuestas incorrectas) que no requiere ningún tipo de revisión post-examen. En cambio, en *Análisis de Estructuras*, del Grado en Ingeniería Civil, sí se pide en ciertas partes que se obtenga la solución de un problema, y se justifique la respuesta.

II.6. Escaneo masivo de las pruebas.

La siguiente fase del sistema consiste en llevar los ejercicios a un escáner convencional, capaz de almacenar los exámenes en un documento de tipo *pdf*. El escaneo puede hacerse en bloques (varios ficheros) o en uno solo.

II.7. Corrección y preparación de pruebas corregidas.

La corrección de las pruebas se realiza dentro del software *Auto-Multiple-Choice*. El programa permite introducir uno o varios documentos *pdf*, y procede a corregir las pruebas. Una vez corregidas, se puede asignar un fichero de alumnos, para que el programa asigne cada examen con un alumno. El sistema es rápido y ágil. Para la asignación de estudiantes, el sistema permite utilizar el nombre escrito por el alumno para ayudar en la identificación de alumnos, en aquellos casos en los que se obtenga error de lectura del código CID.

Desde el programa AMC, se pueden exportar las calificaciones obtenidas en varios formatos: en un documento *pdf* único, o en una hoja de cálculo de Libre Office. En este último formato, el programa muestra de forma detallada qué respuestas dieron los estudiantes a las preguntas, destacando asimismo las que se dejaron en blanco, como aquellos alumnos que no asistieron a la prueba. Asimismo, se muestran estadísticas de respuestas, permitiendo así el profesor identificar qué preguntas fueron más complejas y cuáles tuvieron mayor porcentaje de respuestas correctas.

La hoja de cálculo resultante permite, a través de un identificador único, enlazar las calificaciones obtenidas con el resto de calificaciones del curso. En esta experiencia se ha utilizado el campo *email* para identificar a los estudiantes, ya que cuando estos resultados se quieren integrar en el sistema global de la asignatura, el código CID no necesariamente está presente en otros sistemas evaluadores (como los basados en Moodle).

El último paso de corrección lo establece también AMC. Consiste en la anotación detallada sobre el propio examen, generándose nuevos ficheros *pdf* que contienen el examen escaneado, y de forma detallada, el programa indica qué puntuación se ha obtenido en cada pregunta. Asimismo, el programa indica cuál era la respuesta correcta en cada caso.

II.8. Envío de pruebas corregidas a estudiantes.

Tras el paso anterior, se generan un conjunto de ficheros que pueden enviarse a los estudiantes. Mediante un script en Python, el sistema se encarga, de forma automática, de enviar a cada estudiante su examen corregido. El estudiante recibe un correo personal, con su prueba. De esta forma, el estudiante no sólo obtienen en un tiempo rápido el resultado de su prueba, sino que puede revisar de forma detallada cómo se ha corregido la misma. En este proceso, hay estudiantes que pueden encontrar algún error que deba subsanarse. Esto permite agilizar la revisión de los exámenes. En estas experiencias se ha observado que se reduce notoriamente el número de estudiantes que acude a revisión de examen.

IV. Conclusiones y aportaciones.

Se presenta un sistema de exámenes con corrección automática basado en software libre y el uso de un escáner de uso habitual en los departamentos. El sistema se ha implantado desde el curso 2015-2016 en diferentes asignaturas del área de Mecánica de Estructuras (Análisis de Estructuras, Teoría de Estructuras,

Mecánica, Análisis Avanzado de Estructuras, Puentes). Las principales ventajas del sistema se resumen en los siguientes puntos:

- Gran rapidez en las etapas de evaluación y difusión de los resultados. Prácticamente al día siguiente de realizarse un examen, pueden recibir los estudiantes las correcciones en su correo electrónico.
- Mayor confort para el profesorado en la etapa de corrección.
- Mayor objetividad de las pruebas. Las calificaciones se reparten entre más ítems que se evalúan de forma objetiva.
- El estudiante recibe en su correo electrónico su examen corregido. Esto es opcional, si el profesor así lo desea.
- Las pruebas quedan escaneadas en formato digital, pudiéndose almacenar durante décadas siempre que se disponga de espacio de almacenamiento. No es necesario conservarlas en papel.
- El sistema permite introducir varias pruebas intermedias, que pueden realizarse en clase, que pueden ayudar a evaluar el aprendizaje del alumno. Estas pruebas sirven también para que el alumno conozca el sistema de examen.
- Gran flexibilidad para el profesor en la elaboración de las pruebas, ya que, al ser un documento LaTeX, se tiene gran libertad para incluir figuras, efectos tipográficos, fórmulas matemáticas, etc. Este aspecto es muy importante en enseñanzas técnicas.
- Puesto que el examen de cada estudiante es único, se observa una mayor concentración del alumnado en el trabajo individual, evitándose intentos de fraude o copia en exámenes.
- La introducción (totalmente opcional) de ventanas de verificación permiten que el examen no pierda la necesidad de una lectura por parte del profesor, en determinadas preguntas. Se abandona así ciertas dificultades asociadas a los exámenes de tipo test, abriendo las puertas a un examen semi-automático, asistido por computador, pero que no pierde la esencia de exámenes con partes escritas.
- El software *Auto-Multiple-Choicek* es totalmente libre y gratuito.

Agradecimientos

Este sistema se ha implantado durante dos cursos académicos en asignaturas del área de Mecánica de Estructuras del departamento de Mecánica de Estructuras e Ingeniería Hidráulica de la Universidad de Granada. Esto ha sido posible porque un equipo docente ha apostado por incorporar este tipo de exámenes en sus pruebas de evaluación. Las contribuciones de los profesores compañeros han permitido explorar nuevas vías de aplicación y particularización a las diferentes asignaturas. En este sentido son muy notorias las contribuciones de mis compañeros, Rafael Gallego Sevilla, María Esther Puertas García, Juan José Granados Romera, y Germán Rodríguez Salido. Ningún sistema de este tipo puede implantarse sin el apoyo del equipo docente de las asignaturas. Sirvan estas líneas para agradecerles su entusiasmo y apoyo desde el primer momento, y sus brillantes contribuciones.

Referencias

- KOTSIANTIS, S., PATRIARCHEAS, K. (2010). A combinational incremental ensemble of classifiers as a technique for predicting students' performance in distance education. *Knowledge-Based Systems*, 23, 529-535.
- QUEROL, E., PEREZ-BENEDITO, J.L., MEDIC, L.(2014). Aplicación del reconocimiento óptico de marcas en el proceso de evaluación continua. *IEEE-VAEP-RITA*, 2, 29-35.