

## LA ENSEÑANZA DE LA MEMORIA CACHE: CONCEPTOS TEÓRICOS, PRÁCTICOS Y DIFICULTAD EN SU ESTUDIO

J. M. VALIENTE, J. SAHUQUILLO, X. MOLERO, A. PONT Y A. RODAS  
*Departament d'Informàtica de Sistemes i Computadors. Escola Universitaria d'Informàtica de Valencia. Universitat Politècnica. Camí de Vera s/n, 4602 València. España.*

*Las memorias cache son el mecanismo clásico e ineludible que incorporan los procesadores para reducir el tiempo medio de acceso a la información. Todos los planes de estudio universitarios de Informática dedican al menos un tema para abordar su estudio. Aunque los conceptos teóricos no suelen diferir notablemente de unas universidades a otras, la forma en que se plantea su estudio, tanto teórico como práctico difiere notoriamente. Este trabajo presenta cómo orientar su estudio y un análisis sobre el grado de asimilación de los conceptos relacionados con esta temática.*

### 1. Introducción

Las memorias cache han tenido una importancia crucial desde su aparición [1], de tal manera que todos los procesadores de altas prestaciones actuales incorporan varios niveles de memoria cache para reducir el tiempo medio de acceso a la información. Su estudio ha estado presente en todos los planes de estudio desde que se implantara la Escuela Universitaria en 1982, así como la Facultad de Informática en 1985, de la Universidad Politécnica de Valencia (UPV), en las cuales se enmarca el presente trabajo.

En los planes de estudio vigentes en la UPV, tanto para las titulaciones de Ingeniería Técnica en Informática como Ingeniería en Informática, el estudio de la memoria cache se encuentra recogido en la asignatura *Estructura de Computadores II* (ECII) que se incluye en las tres carreras. Esta asignatura, junto con *Estructura de Computadores I* (ECI), completa la visión de la organización interna del computador. En ECI el alumno estudia el procesador con sus elementos fundamentales: ruta de datos, unidad de control y unidad aritmético-lógica, mientras que ECII estudia el estudio de la memoria y de la entrada/salida (E/S). El bloque que estudia la memoria del computador está formado por 3 temas: (i) Introducción al sistema de memoria, (ii) Diseño de módulos de memoria, y (iii) Jerarquía de memoria.

### 2. Contenido temático y metodología

El bloque dedicado al sistema de memoria contempla las temáticas referidas a la memoria principal o central, la memoria cache y la memoria virtual. Previo al estudio de la memoria cache se estudian diferentes alternativas de diseño de la memoria central que permiten

soportar caches. Así, en el tema *Diseño de módulos de memoria*, se da una visión tecnológica de la memoria central, haciendo énfasis en las memorias comerciales que utilizan el modo de acceso paginado o por ráfagas como una de las soluciones actuales para transvases rápidos de bloques hacia la memoria cache. También se estudian memorias con gran ancho de banda y entrelazadas.

El estudio de la memoria cache empieza con la definición de principio de localidad de la información. A continuación, se hace hincapié en la organización interna en función de su capacidad, tamaño de bloque, algoritmo de correspondencia, etc. Se estudian las tres organizaciones clásicas de una memoria cache: correspondencia directa, asociativa por conjuntos y completamente asociativa, y su funcionamiento (algoritmos de ubicación y reemplazo). El estudio se realiza principalmente desde un enfoque cualitativo, que permita al alumno comprender su funcionamiento y las distintas opciones de diseño. Sin embargo, también se utilizan pequeñas trazas para estudios de prestaciones sencillos y se hacen pequeños cálculos del coste hardware en bits.

Las clases magistrales con transparencias sin apenas animaciones, se utilizan en las primeras clases para presentar al alumno el porqué de la aparición de las memorias cache en los procesadores [1], siguiendo el distintivo de "aprendizaje por evolución" (*learning by evolution*), que con frecuencia se utiliza en las asignaturas de arquitectura de computadores. Las diapositivas animadas se utilizan para explicar el funcionamiento de estas memorias. Seguimos un planteamiento paso a paso solucionando los aspectos básicos de funcionamiento de un sistema de cache. El objetivo más ambicioso que nos planteamos en este tema es que el alumno comprenda todos los pasos que transcurren desde que el procesador emite una petición de referencia a memoria a la cache hasta que le llega el dato (caso de instrucciones *load*) o hasta que el dato se almacena (caso de instrucciones *store*), ya que creemos que un elemento motiva más cuanto con más precisión describe la realidad. Por esta razón, se explica el TLB (*translation lookaside buffer*) y cómo se relaciona con la cache, siguiendo la línea metodológica "poniéndolo todo junto" [2].

La clase práctica que proponemos difiere de las clásicas prácticas de laboratorio donde se realiza la implementación de un programa o circuito similar al estudiado en clase. Creemos que es interesante que el alumno disponga de una herramienta de apoyo o simulador que valide si ha realizado correctamente o no el ejercicio. Por ejemplo, si ha aplicado correctamente el algoritmo de reemplazo a una traza. Más detalles sobre una herramienta con tales características se encuentra en [3].

### **3. Evaluación**

En esta sección nos planteamos medir el grado de dificultad y asimilación de la memoria cache en comparación con el resto de conceptos de la asignatura ECII. Actualmente el examen es de tipo test con un total de 20 preguntas con el mismo peso sobre la nota. Para evaluar el grado de dificultad del estudio de la memoria cache, agrupamos las preguntas de los exámenes en cinco bloques temáticos: memoria central, memoria cache, E/S, interrupciones y acceso directo a memoria (ADM). El peso que tienen las preguntas sobre la nota del examen para cada uno de los bloques es, respectivamente, de 30%, 20%, 13%, 27% y 10%.

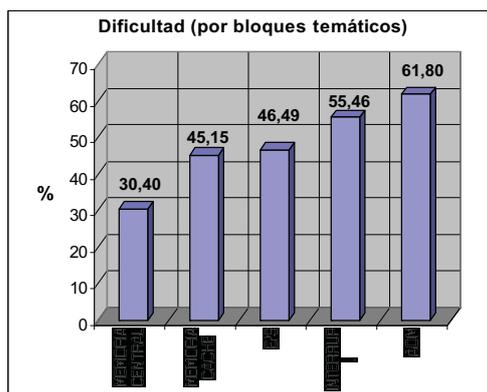
Para medir el grado de dificultad de las preguntas sobre la memoria cache con respecto al resto de bloques temáticos utilizamos los índices *dificultad* y *discriminación*. Esta información ha sido ofrecida por el Instituto de Ciencias de la Educación de la universidad. El índice dificultad se calcula como:

$$Dificultad = \left( 1 - \frac{n_{ACIERTOS}}{N_{TOTAL\_CASOS}} \right) \cdot 100 \quad (1)$$

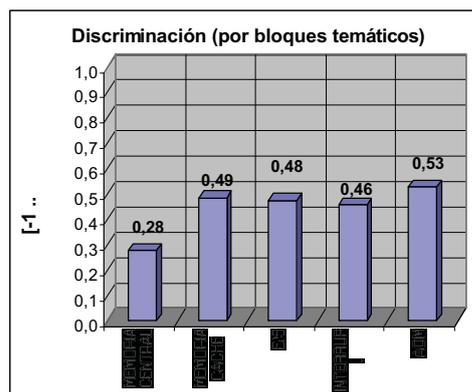
y varía desde 0 (dificultad mínima) hasta 100 (máxima dificultad). La métrica discriminación:

$$Discr = \left( \frac{n_{ACIERTOS\_GRUPO\_SUPERIOR} - n_{ACIERTOS\_GRUPO\_INFERIOR}}{N_{TOTAL\_CASOS} / 2} \right) \quad (2)$$

oscila entre  $-1$  y  $+1$ , e indica el grado de *discriminación* entre dos poblaciones. La dificultad se ha calculado para el total de los exámenes, pero para calcular la discriminación hemos extraído dos grupos (o poblaciones) del conjunto de alumnos evaluados: el *grupo superior* que incluye el 25% de los alumnos con mejores calificaciones, y el *grupo inferior* que incluye el 25% de los alumnos con las peores. Un valor cercano a  $+1$  en (2) indica que la pregunta distingue bien a los que “saben” de los que “no saben”. Las Figuras 1 y 2 muestran los resultados de ambos índices para los 5 bloques temáticos considerados.



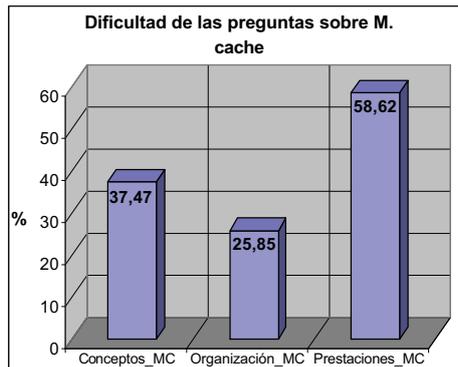
**Figura 1:** Grado de discriminación de los bloques temáticos de ECII.



**Figura 2:** Grado de dificultad de los bloques temáticos de ECII.

Los resultados muestran que la memoria central es el bloque temático más fácil de asimilar, la memoria cache y los conceptos sobre la E/S tienen un grado de dificultad medio, mientras que las interrupciones y el ADM constituyen los apartados más difíciles. El resultado del ADM no se considera significativo por ser este tema el último de la asignatura, que se imparte generalmente en dos clases. Comparando ambas figuras, se observa que el grado de discriminación de los distintos bloques no es proporcional a su grado de dificultad. Así, las preguntas acerca de la memoria central son poco discriminatorias, probablemente por ser más fáciles, y son contestadas correctamente en similares proporciones por los grupos superior e inferior. Sin embargo, el resto de grupos temáticos tienen grados de discriminación más altos. Por ejemplo, para el caso de la memoria cache, los alumnos del grupo superior que contestan correctamente a estas preguntas superan en un 49% a los del grupo inferior. Esto significa que aquellos alumnos que realmente no han comprendido los conceptos de memoria cache suelen fallar en el examen.

Una causa probable del resultado puede ser que las preguntas sobre memoria cache requieren el manejo de varios conceptos al mismo tiempo, su interrelación y, a menudo, la realización de sencillos cálculos matemáticos que ligan dichos conceptos. El alumno que realmente no lo ha comprendido bien suele llegar a conclusiones erróneas. Esto mismo no ocurre en otros casos, como por ejemplo en las interrupciones. En este caso, aun siendo un concepto más difícil, su discriminación es más baja, pues requiere menos cálculos o menor interrelación de ideas, con lo cual se producen muchos aciertos en el grupo inferior, probablemente más por tanteo o por eliminación que por conocimientos.



**Figura 3:** *Dificultad de las preguntas acerca de la memoria*

Como se puede observar, las preguntas acerca de análisis de prestaciones en casos concretos resultan notablemente más difíciles para los alumnos que el resto. La necesidad de hacer cálculos complejos y de mezclar en un mismo ejercicio varios conceptos provoca en un menor número de respuestas correctas. En cambio, las preguntas sobre la organización de la cache son las más fáciles, seguramente porque las ideas son más directas y sistemáticas, sin necesidad de cálculos complicados.

#### 4. Conclusiones

En este trabajo hemos presentado una propuesta sobre el contenido y la metodología a seguir para el estudio de la memoria cache, enmarcada dentro de la asignatura de Estructuras del Computador. Se ha realizado un análisis detallado sobre la dificultad que entraña el tema en comparación con otros bloques temáticos de la asignatura.

El resultado es que, aunque el tema en sí tiene una dificultad media, no todos los aspectos relacionados tienen la misma dificultad. Por ejemplo, las preguntas relacionadas con las prestaciones entrañan más de doble de dificultad que las relacionadas con la organización. Este análisis es especialmente interesante, ya que nos indica en qué puntos debemos profundizar para ayudar al alumno en su estudio.

#### Referencias

- [1] A. J. Smith, *Cache Memories*, ACM Computing Surveys, pp. 473-530, (1982).
- [2] D. A. Patterson and J. Hennessy, *Computer Organization and Design: the hardware/software Interface*, Morgan-Kauffman, (1997).
- [3] X. Molero, y otros. *WebCache: a graphic simulation tool of cache memories using the internet*. Int. Conf. on Simulation and Multimedia in Engineering Education (2000).