Tutorial 7: Disk Storage, Basic File Structures, and Hashing

CS3402 Database Systems

Question 1

- If the database consists of 512 records and the blocking factor is 8. On average, how many blocks needed to be searched to if the records are in unordered format and ordered format?

Question 2

- Under which case the performance of hashing is better: (1) the values of hash key are uniformly distributed; and (2) the values of the hash key are in normal distribution.

Question 3 (1/2)

- Suppose that we use hashing to organize a *PRODUCT* file containing records with the following product# values: 2369, 3760, 4692, 4871, 5659, 1821, 1074, and 7115.

- (a) Let the hash function be **h(product#) = product# mod 5**, show the *static hash* structure for this file. Assuming that each bucket can hold at most three records as shown below, and records in each bucket is unordered.

Bucket i

	Pointer

Question 3 (2/2)

- (b) Some new records are inserted into the file with product# values: 1620, 2428, 3945, 4759, 6975, 4981, and 9206. Show the updated hash structure for this file when using *chaining* for *collision resolution*. That is, if collision occurs, new records are inserted in overflow buckets and pointers are set from the original buckets to the overflow buckets. Assuming that each overflow bucket can hold at most three records, as shown below.

Overflow Bucket

		Pointer
		Pointer
		Pointer

Question 4

- In extendible hashing, how many hash codes can you have in maximum if the global depth is 3?