308-420A Secondary Storage Algorithms and Data Structurs

Supplementary Questions and Exercises

Files and Secondary Storage

Section 1.2–1, Secondary Storage

1. What is the difference between algorithms which access data on secondary storage and algorithms which use only data in RAM?

   (a) Secondary storage algorithms are easier.

   (b) Secondary storage must only be used if there is too much data for RAM.

   (c) Transferring data from secondary storage is much more expensive than accessing it in RAM.

   (d) Finding data on secondary storage is much more expensive than finding it in RAM.

   (e) More than one of the above.

2. A secondary storage disk spins at 3000 r.p.m. and has a track capacity of 20,000 bytes. There are 60 tracks per cylinder. Average read head movement time between any one track and any other is 20 milliseconds (ms.). What is the access/transfer ratio?

   (a) 20,000

   (b) 30,000

   (c) 40,000

   (d) 1,210,000

   (e) None of the above.

3. What are *blocks, pages, tracks* and *cylinders*?

4. Define the *access-transfer ratio* and discuss its significance. Give some typical values.

5. Using some or all of the following quantities, give a formula for the access-transfer ratio, ρ : μ (average arm delay, sec.), σ (rotation speed, rev./sec.), β (no. bytes/track), γ (no. tracks/cylinder).

6. The significant difference between secondary storage and RAM can be measured by two quantities. Discuss the difference and the quantities.

7. What are the relationships among *rotation speed, average latency, transfer time per byte* and *bytes per track* for circulating memories?

8. Explain the relationship between the access-transfer ratio and blocks of data. What does it imply about block size?

9. How many tracks per cylinder are there on a head-per-track device, such as a magnetic drum?

10. How has the access-transfer ratio for disk changed since the time the textbook was published in 1984? Give an estimate of the magnitude of this change. Say precisely which fundamental characteristics of disk memory have changed, and in what way, to affect the access-transfer ratio.

11. A magnetic tape drive requires a distance of 1.5 centimeters to decelerate from writing (reading) speed to a stop, and to accelerate again for the next write (read). (That is, about 0.75 cm. for each operation.) Recording density is 4000 bits/cm., and the drive records eight data tracks and a parity track in parallel. Data cannot be written in this "gap".

    What blocksize (in bytes) would ensure that half the space on the tape is used for storing data and not lost to gaps? What must you do with this blocksize to use the tape more efficiently than this?

12. What are *blocks* and why are they used on secondary memory? Why can blocks not be very small or very large? Give an argument to suggest roughly how many bytes a block should hold.

13. Given the rotation speed of a disk, σ, in revolutions per minute,and the recording density, β, in bytes per track, give formulas for a) τ, the transfer time per byte, in microseconds, and b) λ, the average latency, in milliseconds.

14. What is the most important advantage of secondary storage? (Without this, everybody would use only RAM.)

15. What is the most significant *technical* characteristic of secondary storage that distinguishes it from RAM? (Without this, there would be no need to study secondary storage data structures and algorithms separately from RAM methods.)

16. A disc has rotation speed $\sigma = 50$ revs./sec., average arm delay $\mu = 10$msec, and track capacity $\beta = 20000$ bytes/track. From these three quantities, derive the three other quantities that are essential to the cost analysis of applications using this disk.

17. A disk drive holds 500,000 bytes per track, rotates at 6000 r.p.m., and requires 7ms. on average to move read/write heads from track to track. What is the access-transfer ratio?

18. Since the 1980s, the access-transfer ratio for disks has increased by a factor of about 50. What is the practical significance of this for the differences between secondary storage and RAM?

19. As of October 12, 1999, the population of the world was officially declared to be six billion. The genome of one human is a record of about a gigabyte. Imagine a disk with a capacity of 60 Gbytes, and a transfer time of 0.16 ns. (nanoseconds) per byte. How many such disks would be needed to store a genomic record for every person on earth, and how long would it take just to transfer the entire file from the disks?

    (Approximate magnitudes: an hour is 3.6 Ksec, a day is $\sim$ 86 Ksec, a month is $\sim$ 2.6 Msec, a year is $\sim$ 32 Msec.)

20. If the access-transfer ratio in the previous question were 1 million, and the data were divided (somehow) into pages of 1000 bytes, how much longer would it take to retrieve all the data if each page must be sought before transferring it, instead of just transferring it?

21. A disk drive has a rotation speed of 6000 rpm, track lengths of 10^6 bytes, and an average arm delay of 5 ms. What is its access/transfer ratio?

22. a) If the access-transfer ratio were the only consideration in determining block size, what size of block would ensure that at least half of the time used by a high-activity application is spent in transferring data to RAM? b) What are the other considerations?

23. Given a disk drive with access-transfer ratio of 1,500,000, transfer time per byte of 10 nsec., and average latency of 5 msec., what are the values of the three *basic* parameters that describe the disk hardware?

24. What are the three *basic* parameters for a magnetic disk unit that would give an access-transfer ratio of about 1 million? Show the calculation.

25. What is the largest blocksize possible, in principle and ignoring any other considerations, for a disk with access/transfer ratio of one million, a rotation speed of 6000 rpm, and an average arm delay of 5 ms.?

26. With the usual meanings for the symbols, what is α if N=900,000, n=10,000, and b=100? Say in words what your calculation means.

27. A *magnetic drum* is a head-per-track circulating secondary storage device: read/write heads are mounted permanently over each track and so there is zero average arm delay; the access/transfer ratio is due solely to rotational latency. If there are a million bytes per track, what is the access/transfer ratio?

28. A drive spins CDs at from 300 revolutions per minute (when the read head is near the edge) to 600 r.p.m. (centre). Average arm delay is 100 ms (millisec.). The head reads 2 bytes per microsecond. a) How many bytes are there per track? b) What is the access-transfer ratio?

Section 1.2–2, Files

1. A telephone book of 3.2 million 100-byte records is organized into 100,000 blocks of 3200 bytes each, stored contiguously on a disk with a transfer time of 2 microseconds (μs.) per byte and an average seek time of 20 milliseconds (ms.). If the file is sorted by name and accessed sequentially, how much time is expected to find one entry?

   (a) 32,320 sec.

   (b) 2,640 sec.

   (c) 1,320 sec.

   (d) 640.02 sec.

   (e) 320.02 sec.

2. A very large file is stored in 10,000-byte blocks on a secondary storage device with access/transfer ratio = 90,000. How much longer would it take to read the whole file by direct access than to read the whole file by sequential access?

3. Explain the categories *activity, volatility* and *symmetry.*

4. A block of 100 records, 100 bytes each, is stored on a secondary storage device with access-transfer ratio 90,000. Two algorithms are used to retrieve the records: algorithm (A) retrieves all 100 in one access; algorithm (B) retrieves each in a separate access. What is the ratio of retrieval times, A:B?

5. What are the relationships among *load factor, number of records per block, number of blocks per file* and *number of records per file*?

6. A magnetic disc drive and a magnetic tape drive can each transfer data at the same rate. What is the advantage of using the disk for a sequential file?

7. Characterize the printed Montreal telephone book in terms of "symmetry" and "basic access speeds" (access complexity).

8. Eighty thousand and one (80,001) fixed-length records are to be stored in a file of 100-record blocks. The load factor should be not greater than 0.8, to keep room for future growth. What is the least number of blocks that will be needed? (Only data is to be stored: no pointers or other auxiliary structures.)

9. What are the relationships between N, the number of records in a file, n, the number of blocks in the file, b, the number of records per block, and α, the load factor? a) α in terms of the others. b) n in terms of the others.

10. Given that a file has N records, stored in blocks of capacity b, with an overall load factor, α, what is the expression that gives n, the number of blocks?

11. A file of a million records is stored, 1000 records per block, on 1250 blocks. What is the load factor?

12. a) Why are logarithmic files in RAM restricted to a base of 2? b) Why are they not so restricted on secondary storage, and what is the advantage?

13. The Thunder Bay Telephone Company prints, in its regular telephone directory, an *inverse directory*, which tells the name of the subscriber given the phone number, and a postal code listing, which tells the postal code given the street and street number. Comment on *symmetry* in this "data structure".

14. In a file, N= 1,000,000, b= 100, α= 0.8. What is n?

15. A file on a disk with access-transfer ratio of one million has a million records of 100 bytes each. An application needs to access 200 of these records. According to the breakeven activity criterion, is this a high- or a low-activity application? Why?

16. A file on a disk with access-transfer ratio of one million has 1000 records of 100 bytes each, and is stored in blocks of 10 records each. 100 of these records are needed by an application, and the file is organized such that all 100 are clustered together. Calculate the costs for a) retrieving each relevant block only once, and b) retrieving each record independently of the others by following a pointer to it.

17. A file with the capacity to hold 10 million records is 75% full. If it consists of 20,000 blocks, what is the blocksize?

18. An unordered sequential file with n blocks and a 70-30 usage distribution of records is searched separately for each of a set of records, none of which is in the file. What is the average number of blocks accessed per search?

19. List the following seven in order of size (together with their sizes).
`human genome record, 200-page novel (text only), 10,000-page encyclopedia (text only), small jpeg file, Landsat scan of Earth, RAM for a personal computer, magnetic hard disk for a personal computer`
Explain any assumptions that you make.

20. What is the "big-O" cost complexity for a file search that accesses $\frac{5}{6}n$ blocks? What access category is the file in?

21. On five separate days in one week, a file of 100 records is processed. The first day, 7 records are needed. Three records are accessed the second day. On subsequent days, 22, 6, and 13 records are needed, respectively. Calculate and comment on the *activity* of all of this processing.

22. a) How many records per block are stored, on average, in a file with N=1,000,000, b=100 and α=0.5? b) How many records in all are there in the file?

23. a) Ten separate query transactions on a file of 10,000 records read a total of 100 different records. What is the activity that the file organization must support? b) If the access/transfer ratio is 10,000 and each record is 100 bytes, is this high or low activity?

24. Which is more volatile, an update which deletes 2000 records of 100,000 and alters 1000 others, or an update which deletes 1000 records and alters 2000 others? Why?

25. A file has 200 million records in 1 million blocks at 300 records per block. Each record has 5 fields. What is the load factor?

26. 1000 people independently access a file of a million 20-byte records, each requesting 150 records. The access-transfer ratio for the device the file is stored on is 100,000. Should the file be designed for high or low activity?

The author gratefully acknowledges support from the taxpayers of Québec and of Canada who have paid his salary and research grants while this work was developed at McGill University, and from his students (who built the implementations and investigated the data structures and algorithms) and their funding agencies.