

Exercises

Problem 1 (review questions)

- 1) What are the differences among direct mapping, associative mapping, and set-associative mapping?
- 2) What is the distinction between spatial locality and temporal locality?

Problem 2 (mandatory for lab 1)

Consider a machine with a byte addressable main memory of 2^8 bytes and block size of 4 bytes. Assume that a direct mapped cache consisting of 8 lines is used with this machine.

- 1) How is an 8-bit memory address divided into tag, line number, and byte number?
- 2) Into what line would bytes with each of the following addresses be stored?

0001 1011

0011 0100

1101 0000

1010 1010

- 3) Suppose the byte with address 1010 0001 is stored in the cache. What are the addresses of the other bytes stored along with it?
- 4) How many total bytes of memory can be stored in the cache?
- 5) Why is the tag also stored in the cache?

Problem 3 (mandatory for lab 1)

Consider the following code:

```
cout << "Hello World";
cin >> a;
for(i = 0; i < 50; i++)
    cout<<i;
```

- 1) Give one example of the spatial locality in the code.
- 2) Give one example of the temporal locality in the code.

Problem 4 (additional)

Average memory-access time. A computer has a cache, main memory, and a disk used for virtual memory. If a referenced word is in the cache, 15 ns are required to access it. If it is in main memory but not in the cache, 70 ns are needed to load it into the cache, and then the reference is started again. If the word is not in main memory, 10 ms are required to fetch the word from disk, followed by 50 ns to copy it to the cache, and the reference is started again. The cache hit ratio is 0.95 and the main memory hit ratio is 0.8. what is the average time in nanoseconds required to access a referenced word on this system?

Problem 5 (additional)

Performance enhancement using cache. A computer system contains a main memory of 32K 16-bit words. It also has a 4K-word cache divided into four-line sets with 64 words per line. Assume that the cache is initially empty. The processor fetches words from locations 0, 1, 2, ..., 4351 in that order. It then repeats this fetch sequence nine more times. The cache is 10 times faster than main memory. Estimate the improvement resulting from the use of the cache. Assume an LRU policy for block replacement