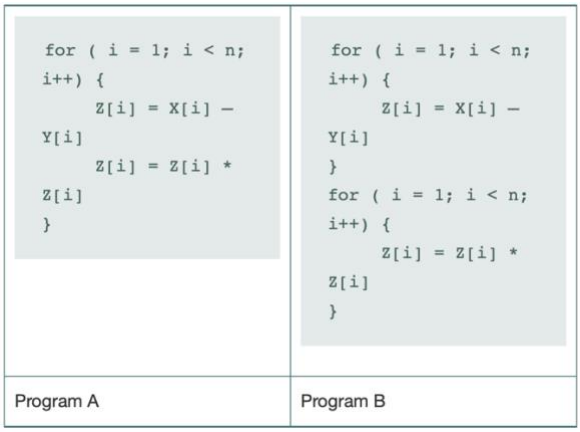
COSC 420: Computer Architecture

Assignment 3

Due date: February 23, 2024

Answer the following questions. The questions are numbered as they appear in the textbook.

4.2 (15 pts) Consider these two programs:



Program A

```
for ( i = 1; i < n;
i++) {
     Z[i] = X[i] –
Y[i]
     Z[i] = Z[i] *
Z[i]
}
```

Program B

```
for ( i = 1; i < n;
i++) {
     Z[i] = X[i] –
Y[i]
}
for ( i = 1; i < n;
i++) {
     Z[i] = Z[i] *
Z[i]
}
```

a. The two programs perform the same function. Describe it.

b. Which version performs better, and why?

Hint: think about the locality concept to answer (b)

4.3 (15 pts) Consider the following code:

```
for (i = 0; i < 20; i++)
   for (j = 0; j < 10; j++)
      a[i] = a[i] * j
```

a. Give one example of the spatial locality in the code.

b. Give one example of the temporal locality in the code.

4.5 (15 pts) Consider an L1 cache with an access time of 1 ns and a hit ratio of H = 0.95. Suppose that we can change the cache design (size of cache, cache organization) such that we increase H to 0.97, but increase access time to 1.5 ns. What conditions must be met for this change to result in improved performance?

Hint: You may review Section 4.4 and Equation 4.2 to answer this question.

4.6 (20 pts) Consider a single-level cache with an access time of 2.5 ns, a block size of 64 bytes, and a hit ratio of H = 0.95. Main memory uses a block transfer capability that has a first-word (4 bytes) access time of 50 ns and an access time of 5 ns for each word thereafter.

a. What is the access time when there is a cache miss? Assume that the cache waits until the line has been fetched from main memory and then re-executes for a hit.

Hint: When a cache miss occurs, the desired address is loaded onto the system bus, and the data (a block) is transferred from the main memory to the cache. So to answer the question, think about the time needed to decide that the address is not in the cache and the time required to transfer a block from the main memory to the cache.

b. Suppose that increasing the block size to 128 bytes increases the H to 0.97. Does this reduce the average memory access time?

Hint: First, review Example 4.1. You may then use Equation 4.2 to compare the average access time between the two cases (64 bytes vs. 128 block size). You should use your answer to (a) to answer (b).

5.3 (15 pts) A two-way set-associative cache has lines of 16 bytes and a total size of 8 kB. The 64-MB main memory is byte addressable. Show the format of main memory addresses.

5.9 (20 pts) Consider a machine with a byte addressable main memory of 2^16 bytes and block size of 8 bytes. Assume that a direct mapped cache consisting of 32 lines is used with this machine.

a. How is a 16-bit memory address divided into tag, line number, and byte number?

b. Into what line would bytes with each of the following addresses be stored?

0001	0001	0001	1011
1100	0011	0011	0100
1101	0000	0001	1101
1010	1010	1010	1010

c. Suppose the byte with address 0001 1010 0001 1010 is stored in the cache. What are the addresses of the other bytes stored along with it?

d. How many total bytes of memory can be stored in the cache?

e. Why is the tag also stored in the cache?