CA HW5 Handwritten part

Problem 1

Block size (B) can affect both miss rate & miss latency. Assuming a machine base CPI = 1, average 1.35 references per instruction. find the block size that minimizes the total miss latency

block size	8	16	32	64	128
miss rate	4%	3%	2%	1.5%	1%

(1) What is the optimal block size for a miss latency of 20 * B cycles?

(2) What is the optimal block size for a miss latency of 24 + B cycles?

(3) For constant miss latency, what is the optimal block size?

Problem 2

In this exercise we look at memory locality properties of matrix computation. The following code is written in C, where elements within the same row are stored contiguously. Assume each word is a 64-bit integer.

```
for (I=0; I<8; I++)
    for (J=0; J<8000; J++)
        A[I][J]=B[I][0]+A[J][I];
```

(1) How many 64-bit integers can be stored in a 16-byte cache block?

(2) Which variable references exhibit temporal locality?

(3) Which variable references exhibit spatial locality?

(4) How many 16-byte cache blocks are needed to store all 64-bit matrix elements **being referenced**? (Assume each row contains more than one element.)

Problem 3

In this exercise, we will examine how replacement policies affect miss rate. Assume a two-way set associative cache with four one-word blocks. Consider the following word address sequence:

`0, 1, 2, 3, 4, 2, 3, 4, 5, 6, 7, 0, 1, 2, 3, 4, 5, 6, 7, 0`

(1) Assuming an LRU replacement policy, which accesses are hits? (List only the addresses that are hits.)

(2) Assuming an MRU (most recently used) replacement policy, which accesses are hits? (List only the addresses that are hits.)

(3) Describe an optimal replacement policy for this sequence. Which accesses are hits using this policy? (List only the addresses that are hits.)

Problem 4

Consider the following portions of two different programs running at the samw time on four processors in a *symmetric multicore processor*(SMP). Assume that before this code is run, both x & y are 0.

```
Core 1: x = 2;
Core 2: y = 2;
Core 3: w = x + y + 1;
Core 4: z = x + y;
```

(1) list all the possible resulting values of w, x, y and z. For each possible outcome explain how to achieve this result.

(2) How to make the outcome more stable?

Problem 5

In this exercise, we compare the design of CPU and GPU.

(1) In terms of hardware architecture, what are the key differences between CPU and GPU.

(2) Consider the following tasks. For each of them, determine which kinds of processor we should use to get better performance, please also explain why.

- sorting 10000 integers by quick sort
- sorting 10000 integers by counting sort
- training a large neural network
- summing 1073741824 numbers
- running physics simulation
- running a virtual machine