Intensive Computation

Prof. A. Massini

April 18, 2024

Midterm test

Student's Name

_____

***Matricola* number**

_____

Exercise 1 (4 points)	
Exercise 2 (5 points)	
Exercise 3 (4 points)	
Exercise 4 (3 points)	
Exercise 5 (4 points)	
Exercise 6 (4 points)	
Question 1 (4 points)	
Question 2 (4 points)	
Total (32 points)	

Exercise 1 (4 points) - GPU & CUDA

Technical specifications	Compute capability (version)									
	1.0	**1.1**	**1.2**	**1.3**	**2.x**	**3.0**	**3.5**	**3.7**	**5.0**	**5.2**
Maximum dimensionality of grid of thread blocks	2				3					
Maximum x-dimension of a grid of thread blocks	65535					2^{31}-1				
Maximum y-, or z-dimension of a grid of thread blocks	65535									
Maximum dimensionality of thread block	3									
Maximum x- or y-dimension of a block	512				1024					
Maximum z-dimension of a block	64									
Maximum number of threads per block	512				1024					
Warp size	32									
Maximum number of resident blocks per multiprocessor	8					16			32	
Maximum number of resident warps per multiprocessor	24		32		48	64				
Maximum number of resident threads per multiprocessor	768		1024		1536	2048				
Technical specifications	**1.0**	**1.1**	**1.2**	**1.3**	**2.x**	**3.0**	**3.5**	**3.7**	**5.0**	**5.2**
	Compute capability (version)									

Consider a matrix of size 1800x1400. You would like to assign one thread to each matrix element.

a) How would you select the **2D grid** dimensions and **2D square block** dimensions of your kernel to **minimize the number of idle threads** on a device having compute capability 3.7?

b) Is it possible to obtain fewer idle threads if the dimensions for **rectangular blocks** are appropriately selected?

Exercise 2 (5 points) - Number representation

- Represent the natural number range [0; 790] using the residue number system, considering:
  - the power-of-2 based choice consisting of 3 moduli S1={2^n -1; 2^n ; 2^{n+1} -1}
  - a moduli set consisting of 3 moduli at your choice S2.
- Give an estimation of the representational efficiency in both cases.
- Represent A= 35 and B=11 in **both residue systems** S1 and S2 defined
- Represent the product P=AxB using S1 or S2 of your choice and then obtain the **mixed radix representation** associated.

Exercise 3 (3 points) – Number representation

Consider the radix-10 digit set [-6, 6] and consider to use three digits.

a) How many values can you represent?
b) What is the redundancy?
c) What is the redundancy index?

Exercise 4 **(4 points) – Number representation**

Given the values A= 00 01 00 01 10 and B = 00 00 01 11 11 in the RB (Redundant Binary) representation, convert A and B in decimal and show the execution of operation A – B using the look-up table for addition. Verify the correctness of the result.

Exercise 5 (4 points) Performance equation

Suppose we have made the following measurements, where we are considering Arithmetic and Logic instructions AL and the **subset** of AL of only Logic instructions LOG:
Frequency of AL operations = 45%
Average CPI of AL operations = 5
Average CPI of other instructions = 2.8
Frequency of LOG = 25% (on the total)
CPI of LOG = 3

Assume that the two design alternatives are to decrease the CPI of AL to 3.5 or to decrease the average CPI of LOG operations to 2.
Compare these two design alternatives using the **processor performance equation** and compute the speedup in both cases, determining which alternative is more cost-effective.

Exercise 6 (4 points)

a) The analysis of a program has shown a speedup of 3 when running on 8 cores. What is the serial fraction according to Gustafson's law? And according to Amdahl's law?

b) Considering the serial fraction obtained in point a), compute the speedup when using 4 cores according to Gustafson's law and to Amdahl's law.

Question 1 (4 points)

Describe Flynn's taxonomy and explain why other classifications are needed and what they are based on.

Question 2 (4 points)

Describe the pipelined multiplier for unsigned and signed numbers and show their scheme.