Intensive Computation

Prof. A. Massini

4 July 2022

Part 1

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

Student's Name

_____

***Matricola* number**

_____

Exercises 1 (5 points) Amdhal Law

The following measurements are recorded with respect to the different instruction classes for the instruction set running a given set of benchmark programs:

Instruction Type	Instruction Count (millions)	Cycles per Instruction
Arithmetic and logic	4	6
Load and store	8	3
Branch	5	7
Others	7	6

Assume that "Arithmetic and logic" instructions can be modified so that they take 5 cycles per instruction instead of 5, and "Load and store" instructions can be modified so that they take 3 cycle per instruction instead of 3 as in the table. Compute the speedup obtained by introducing **only one** enhancement and **both** enhancements using the **Amdhal law**.

How many cycles should the "Branch" instructions consist of if we wanted modify only this subset of instructions but get the same speedup obtained applying both of the above enhancements?

Exercise 2 (5 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)									

You need to write a kernel for a computational fluid dynamics problem that operates on a matrix of size **2500x4750.** You would like to assign one thread to each matrix element, and your thread blocks to use the maximum number of threads per block possible on your device.

a) How would you select the dimensions of a **1D grid and 2D rectangular blocks** for your kernel, minimizing the number of idle threads? Consider a device having compute capability 1.2.

b) How would you select the dimensions of a **2D grid** and **2D square blocks** for your kernel, minimizing the number of idle threads? Consider a device having compute capability 3.5.

Exercise 3 (6 points) - Number representation

- Represent the natural number range [0; 2599] using the residue number system, considering:
  - the conventional choice consisting of 3 moduli $\{2^n -1; 2^n ; 2^n +1\}$
  - a moduli set consisting of 3 moduli not power of 2.
- Give an estimation of the representational efficiency in both cases.
- Represent A= 45 in both residue systems defined and in the **mixed radix representation** for one of them at your choice.

Exercise 4 (5 points) – Number representation

Given the values A= 00 01 00 11 10 and B = 00 00 00 11 11 in the RB (Redundant Binary) representation, show the execution of operation A - B.

Question 1 (6 points)

Briefly describe the characteristics of a pipeline adder and a pipeline multiplier in case of signed (2's complement) operands.

Question 2 (5 points)

Illustrate the main characteristics of a GPU architecture.