Intensive Computation

Prof. A. Massini

June 6, 2019

End-of-term test

Student's Name

_____

***Matricola* number**

_____

Exercise 1 (4 points)	
Exercise 2 (4 points)	
Exercise 3 (4 points)	
Question 1 (3 points)	
Question 2 (3 points)	
Exercise 4 (3 points)	
Exercise 5 (3 points)	
Exercise 6 (3 points)	
Exercise 7 (3 points)	
Exercise 8 (3 points)	
Total (33 points)	

Exercise 1 (4 points) - GPU & CUDA

You need to write a kernel that operates on an ultrasound color image represented by a matrix of size **640x480x24** (color images: have 24 bits for representing the three colors red, green, blue with 8 bits). You would like to assign one thread to each matrix element. You would like your thread blocks to use the maximum number of threads per block possible on your device, having compute capability 3.5.

a) How would you select the dimensions of a **2D grid** and **3D blocks** for your kernel? Consider the two cases of **rectangular** and **square** blocks for the x and y dimensions.

b) What is the best choice for grid and block dimensions with respect to the number of idle threads?

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)									

Exercise 2 (4 points) – Interconnection Networks – CLOS

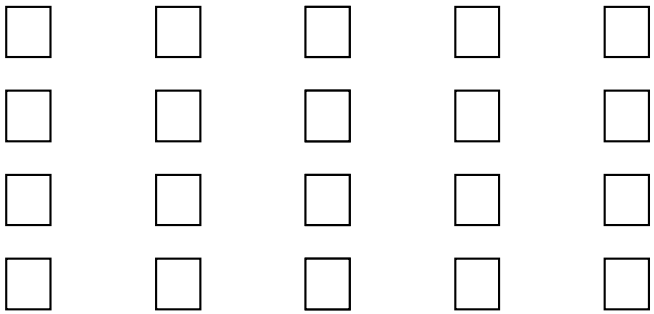
Design a Clos network of size 240 x 240, using in the first stage modules having 16 inputs. Consider both cases, strictly non-blocking and rearrangeable network.

Compare the cost of the crossbar 240 x 240 and the Clos network, strictly non-blocking and rearrangeable, designed in the previous point.

Exercise 3 (4 points) – Interconnection networks – (2 log N - 1) MIN

Briefly explain how the looping algorithm works.

Complete the scheme of the Benes network. Show the switch setting to realize permutation $P = \begin{pmatrix} 01\ 23\ 45\ 67 \\ 40\ 76\ 15\ 23 \end{pmatrix}$ according to the looping algorithm.



Question 1 (3 points)

Briefly explain the Amdhal's law and the performance equation

Question 2 (3 points)

Briefly describe the Flynn's taxonomy, highlighting its limitations.

Exercises 4 (3 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	6	3
Branch	6	5
Others	8	4

Assume that "*Arithmetic and logic*" instructions can be modified so that they take 4 cycle per instruction instead of 6 as in the table. Compute the speedup obtained by introducing this enhancement using the **Amdhal law**.

How many cycles should "*Branch*" instructions consist of to reach at least the same speedup obtained modifying the number of cycles of "*Arithmetic and logic*" instructions as above?

Exercises 5 (3 points) Performance equation

Suppose we have made the following measurements, where we are considering FP (Floating Point) instructions and FPM (Floating Point Multiplication) instructions:
Frequency of FP operations = 20%
Average CPI of FP operations = 3.2
Average CPI of other instructions = 1.8
Frequency of FPM = 10%
CPI of FPM = 8

Assume that the two design alternatives are to decrease the CPI of FPM to 4 or to decrease the average CPI of all FP operations to 2.4.
Compare these two design alternatives using the processor performance equation.

Exercise 6 (3 points) – Number representation

Given the values A= -12 give its RB (Redundant Binary) representation. Given B = 00 10 00 00 11 in the RB representation, convert it in decimal.

Show the execution of operation A+B. Verify the value of the results.

Exercise 7 (3 points) - N**umber representation**

- Determine two ways to choose the moduli set for using the **residue number system** to represent values in the number range [0; 479], considering:
  - the conventional choice consisting of **3 moduli** $\{2^n-1; 2^n; 2^n+1\}$,
  - a moduli set consisting of **4 moduli**.
- Compare the different choices with respect to the number of bits necessary for the representation, and consider also the number of bits needed for representing the range [0; 479] with the conventional binary system.

Represent A= 23 and B=36 using the considered different choices of the moduli sets, and show how to compute the sum A+B.

Exercise 8 (3 points) – Arithmetic circuit time and area

Draw a pipelined multiplier for **signed binary values** consisting of three bits. Compute the **time** (propagation delay) and **area** required by the **multiplier**.