

# TDDD56, Mock exam of GPU part

Important! This is not the actual exam questions. The purpose of this document is to give you an idea of what kind of questions you are expected to answer!

This will be about 1/3 of the whole exam.

## 1. Algorithms

*(This kind of questions may to varying extent demand that you write code or pseudo code. Partial score will be common for vague or partially correct answers. We may get even closer to actual code than in this example.)*

Describe in pseudo code and figures how an optimized matrix multiplication works on the GPU. (5p)

## 2. Question that calls for a longer explanation

*(I expect a fairly long answer, possibly figures too. Rarely more than half a page, though. Note that you can often get partial score.)*

- a) List three different kinds of GPU memory and describe for each their characteristics in terms of performance, usage and accessibility. CUDA terminology is assumed, please note if you use OpenCL terminology. (3p)
- b) Describe the principle for *reduction* algorithms. Give two examples of problems handled by reduction. Demonstrate with a simple figure. (2p)

## 3. Quickies

*(I expect short answers here, one or two lines of text. A simple figure can be called for in some cases. Partial score is usually not relevant for separate questions. These questions are similar to the "lecture questions".)*

- a) What geometry is usually used for shader-based GPU computing?
- b) What concept in CUDA corresponds to a streaming multiprocessor in the GPU architecture?
- c) How can pinned (page-locked) CPU memory improve performance?
- d) List three different kinds of hardware that OpenCL runs on. (Similar systems by different vendors count as one.)
- e) Which parts of the FFT algorithm are hardest to accelerate on a GPU?