Building LLM Reasoners Midterm Instructions and Topics

The midterm is an **in-class exam**. You will have 2 hours to complete it, but it is designed to only take around 90 minutes or so. **You are allowed one 8.5x11 note sheet, double sided, typed is okay.** The questions will be more about synthesis and understanding of concepts than simple recall.

The exam will have the same broad types of questions (multiple choice, short answer, long answer) as past midterms available on my website. For long answer questions, you can show your work for partial credit, but for our sake and yours, please try to refrain from writing an essay as a response unless the question explicitly asks for it.

Topics   Below is the list of topics that may be covered on the midterm, which is the bulk of the course material thus far. The midterm will focus on these topics, but just because something isn't here doesn't mean that it couldn't come up in some fashion! While we won't quiz you about trivia regarding research papers, anything discussed in the lectures or assignments may show up.

- Transformers: architecture details

- Self-attention: mathematical definition, how it works

- Inference in decoders, nucleus sampling

- Transformer Language Modeling: everything from Assignment 1, including training, inference, and evaluation (perplexity)

- Positional encodings, RoPE

- Subword tokenization

- Optimizers, AdamW

- Resource accounting

- GPU architecture at a high level

- CUDA kernels: what they are

- Triton, especially the weighted sum example

- FlashAttention

- Scaling Laws: general forms of them, reasoning about model and data scaling tradeoffs (Chinchilla)

- SFT: training objective, reasoning about SFT datasets

- Direct Preference Optimization

- RL basics

- GRPO

Other content   You should expect to see examples of text and be comfortable reasoning about how these algorithms might work on such examples, as in the assignments so far.

Readings   We won't expect you to know content from the papers that hasn't been covered in lectures or assignments. You can focus on studying the content covered in the lecture slides.

Practice problems   The best source of practice problems is the prior midterms, although the topics differ a bit from this course.