ERIC's Special Participation A - ChatGPT-5.1 Pro on EECS 182 HW5

Testing Extended Thinking Capabilities on Written Problems

My Prompt to ChatGPT-5.1 Pro

Prompt: "Generate thorough solutions for HW5 problems 1-4. Show all work, be as accurate as possible, and explain your reasoning clearly."

Model Response:

- Entered extended thinking mode (approximately 30 minutes)
- Generated comprehensive PDF with complete solutions
- Single-shot output with no iteration required

Experimental Setup

LLM Tested: ChatGPT-5.1 Pro (with extended thinking)
Topic: EECS 182 Homework 5, Problems 1-4 (Written Problems)
Date: December 2025
Deconfliction: ChatGPT-5.1 Pro claimed in class spreadsheet

I tested ChatGPT-5.1 Pro's extended thinking capability by giving it a single prompt to solve all four written problems from HW5. The model activated its "thinking mode," spent approximately 30 minutes reasoning internally, then produced a comprehensive solution document as a PDF.

Key Findings

One-Shot Success Rate: 100%

All four problems were solved correctly on the first attempt:

- Problem 1 (Convolutional Networks): ✓ Correct
- Problem 2 (Batch Normalization): ✓ Correct
- Problem 3 (Depthwise Separable Convolutions): ✓ Correct

- Problem 4 (Regularization and Dropout): ✓ Correct

No human intervention, corrections, or follow-up prompts were needed.

Quality Observations:

- All mathematical derivations show complete step-by-step work
- Zero computational errors across ~10 pages of calculations
- All numerical answers match official solutions exactly
- Textbook-level explanations with proper mathematical formatting
- No misconceptions or hallucinations detected

Conclusion

ChatGPT-5.1 Pro with extended thinking completely solves graduate-level deep learning homework with perfect accuracy. The 30-minute thinking time produces solutions that match or exceed teaching staff quality in clarity, completeness, and correctness.

Key Observation: For well-defined mathematical problems, this model demonstrates superhuman accuracy and consistency. The extended thinking mode appears to provide internal verification and multi-perspective reasoning that eliminates typical LLM errors.

Bottom Line: Every homework problem was one-shot correctly. Zero errors, zero hallucinations, zero need for human guidance. This represents a fundamental shift in capability for LLMs on technical coursework.

Implications for EECS 182: Traditional "compute X" or "derive Y" homework is now fully automated. Any problem with a definite correct answer (derivations, calculations, parameter counting) can be perfectly solved by this model. Course design should shift toward conceptual "explain why" questions requiring insight, open-ended design problems, and creative synthesis of multiple concepts.