Topic 14: Reasoning and Thinking LLMs

1. Chain-of-Thought Prompting: https://arxiv.org/abs/2201.11903
2. Zero-shot CoT: https://arxiv.org/abs/2205.11916
3. Tree-of-Thoughts: exploring multiple reasoning hypotheses:
4. https://arxiv.org/abs/2305.10601
5. Self-Consistency: https://arxiv.org/abs/2203.11171
6. Graph of Thoughts graph-structured planning over intermediate “thoughts”: https://arxiv.org/abs/2308.09687
7. ReAct interleave reasoning with tool use and actions: https://arxiv.org/abs/2210.03629
8. Language Agent Tree Search combines planning, acting, and reasoning with MCTS: https://arxiv.org/abs/2310.04406
9. Program-of-Thoughts generates code as the reasoning trace, then executes it: https://arxiv.org/abs/2211.12588
10. Plan-and-Solve prompting plan first, then carry out the steps: https://arxiv.org/abs/2305.04091
11. Self-Ask prompting break questions into sub-questions with optional search: https://arxiv.org/pdf/2210.03350
12. CRITIC self-correct by using tools to critique and revise outputs: https://openreview.net/forum?id=Sx038qxjek
13. Self-Refine iterative self-feedback to improve an initial answer: https://arxiv.org/abs/2303.17651
14. Reflexion agents that write reflections to learn from mistakes across trials: https://arxiv.org/abs/2303.11366
15. STaR: self-taught reasoner that bootstraps rationales for fine-tuning: https://arxiv.org/abs/2203.14465
16. Quiet-STaR: learn internal “thoughts” during continued pretraining for better reasoning: https://arxiv.org/abs/2403.09629
17. OpenAI o1 methods overview on training models to “think before answering”: https://openai.com/index/learning-to-reason-with-llms/
18. Process supervision overview for improving mathematical reasoning with step-level signals: https://forum.openai.com/public/videos/improving-mathematical-reasoning-with-process-supervision-2023-07-26
19. Let’s Verify Step by Step train verifiers to check each reasoning step: https://openreview.net/forum?id=v8L0pN6EOi
20. Step-level reward models as navigators use PRMs to steer search at inference time: https://openreview.net/forum?id=RSQL6xvUYW
21. Multi-step Problem Solving Through a Verifier use a verifier to guide and validate multi-step solutions: https://aclanthology.org/2024.findings-emnlp.429.pdf
22. OpenAI reasoning best practices practical tips for using o-series reasoning models: https://platform.openai.com/docs/guides/reasoning-best-practices
23. OpenAI o3 and o4-mini overview of newest reasoning models and use cases: https://openai.com/index/introducing-o3-and-o4-mini/

24. Google Gemini “thinking” guide API usage for models with internal thinking: https://ai.google.dev/gemini-api/docs/thinking
25. DeekSeek-R1 paper: https://arxiv.org/abs/2501.12948
26. Scaling LLM Test-Time Compute Optimally can be More Effective than Scaling Model Parameters: https://arxiv.org/abs/2408.03314
27. Improve Mathematical Reasoning in Language Models by Automated Process Supervision: https://arxiv.org/abs/2406.06592v1
28. Training Language Models to Self-Correct via Reinforcement Learning: https://arxiv.org/abs/2409.12917
29. Towards System 2 Reasoning in LLMs: Learning How to Think With Meta Chain-of-Thought: https://arxiv.org/abs/2501.04682