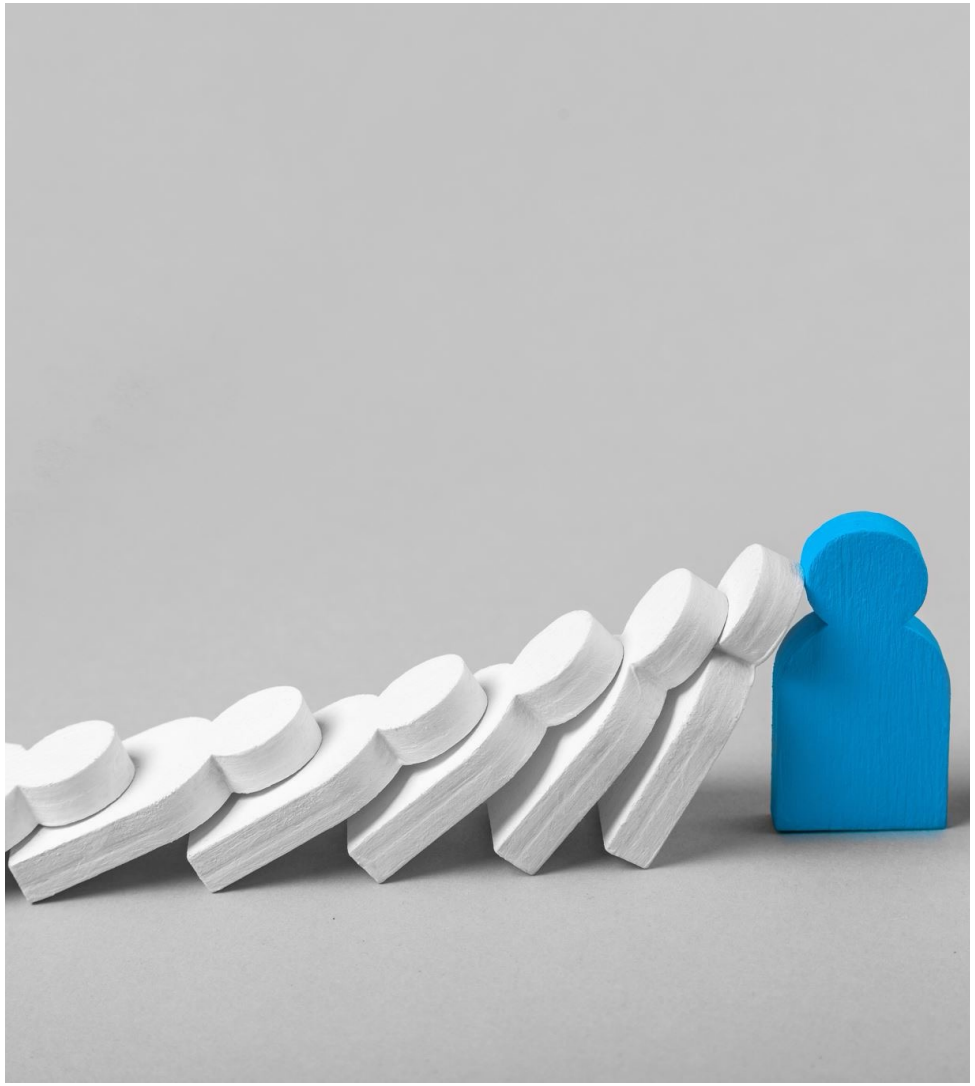


FUNDAMENTAL LIMITATIONS OF LARGE LANGUAGE MODELS IN REASONING

Temple University

Pei Wang

https://cis.temple.edu/~pwang/

TRiPS Talk, 4/1/2025

Can Large Language Models Reason?

- ✓A Survey of Reasoning with Foundation Models
- ✓Multimodal Chain-of-Thought Reasoning: A Comprehensive Survey
- ❖GSM-Symbolic: Understanding the Limitations of Mathematical Reasoning in Large Language Models
- ❖Can Large Language Models Reason and Plan?

Complexity: Success cases may fail after minor changes; a new model (or version) usually fixes some previous failures

Are there **fundamental** limitations of LLM in reasoning?

Reasoning in Logic and Psychology

Reasoning: to derive new knowledge from existing knowledge, step by step

Traditional models: normative (logic) vs. descriptive (psychology)

- Logical reasoning: each step follows an inference rule of a logic
  - Traditional logic: Aristotle's Syllogistic
  - Mathematical logic: Propositional Calculus, First-Order Predicate Calculus
  - "Anti-psychologism"
- Psychological theories of human reasoning
  - "Human thinking does not follow logic"
  - E.g., Wason Selection Task

Reasoning in AI

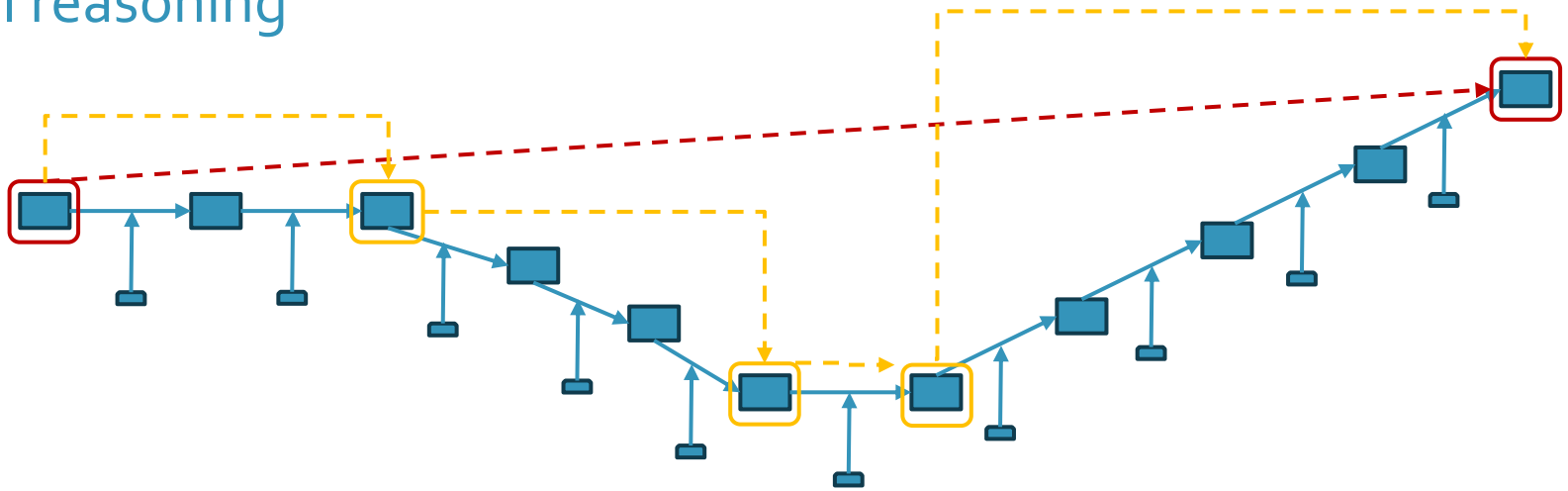
- Symbolic AI: Reasoning according to the inference rules of a logic
  - Classical logic (automated reasoning, theorem proving, …)
  - Non-classical logic (non-monotonic, probabilistic, fuzzy, …)
- LLM: Summarizing statistical patterns in human (linguistic behavior) data
  - "Next-token prediction is enough for AGI", with "emergent abilities"
  - ANNs learn cognitive tasks (including reasoning) as end-to-end mappings
  - LLMs "reason" by adding intermediate stops in these mapping processes (using "Chain-of-Thought", Reinforcement Learning, search, …)

Step and Process

Step-by-step learning

End-to-end learning

Logical reasoning



Two types of "Inference Rules"

"Rule": *derivation* (**{P, P → Q} |— Q**) vs. *implication* (**lighting → thunder**)

- Correspondences: "if-then", Deduction Theorem
- Differences:

DERIVATION	IMPLICATION
procedural	declarative
built-in	acquired
meta-level	object-level
formal	empirical
automatically triggered	deliberately applied

- Confusions between the two were denounced long ago but are still widespread

Evaluation of Reasoning in LLM

- Strengths: simplicity, efficiency, similarity (to human behaviors in many situations)
- Weaknesses: validity, reliability, justifiability, explainability
- Inference rules can be taught to LLM, but they will still be acquired as implications
- Why cannot ANN learn (meta-level, procedural) inference rules?
  - Variable binding in ANN (symbols with multiple interpretations)
  - Meta-learning: keeping coherence, rules/algorithms at the meta-meta-level
  - Choosing among logical models for a given problem

LLM can solve many "reasoning problems" without a "reasoning mechanism" (in the long-established sense)

Theoretical Issues

- Is logic acquired or innate?
- Two senses of "logic":
  - formal models (language, semantics, and inference rules)
  - regularities in thinking ("protologic", "laws of thought", …)
- Origin of the (innate) inference rules:
  - design (artificial systems)
  - evolution (natural systems)
- Nature vs. nurture: When creating an AGI, what should be built in, and what should be left for the system to learn?

NARS vs. LLM

Intelligence as adaptation in a realistic working environment (AIKR)

- Concept-Centered Knowledge Representation (CCKR): abstracting experience
- Reasoning as goal-guided concept substituting
- The inference rules of NAL are designed, but beliefs (including implications) are learned
- The system's behaviors depend on its experience, which may be different from human's

Recent developments in LLM:

- "Large Concept Model": Using "concepts" as "tokens" (but what is a "concept"?)
- "Agent AI": Taking goal-driven actions (but which "goals"?)

Current Works

- Extensions of NARS:
  - NAL-9: Self-monitoring and self-control
  - Summarizing derivation {T1, B} |— T2 as implication T1 → T2
- Using LLMs as tools:
  - NarsGPT
  - Natural Language Inference