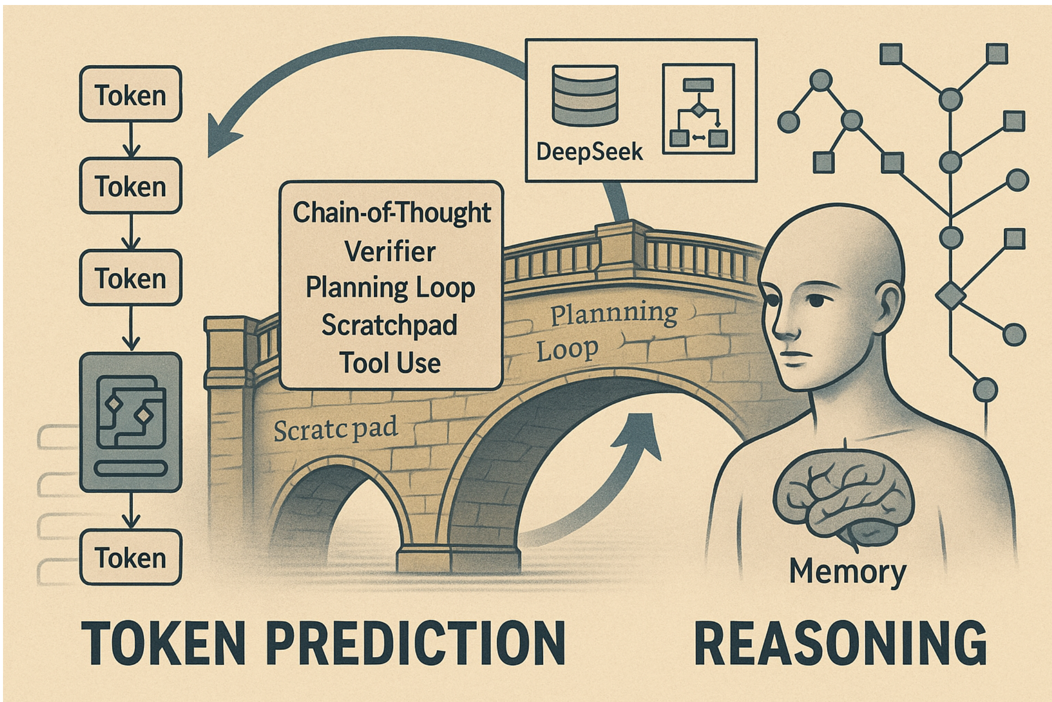
The Rise of Reasoning Models: A Revival of Planning in AI

Anjan Goswami



Subtitle: From Next-Token Prediction to Structured, Multi-Step Deliberation

1 The Gap Between Prediction and Inference

Early symbolic AI was grounded in monotonic logic, which lacked the ability to retract inferences when new information arose. This led to the development of non-monotonic reasoning systems. Modern LLMs, by contrast, do not reason through logic chains at all — their behavior is stochastic and context-sensitive, making them implicitly non-monotonic without formal inference states. This leads to inconsistencies across long reasoning chains and an inability to backtrack or revise inferences. Without external mechanisms for memory, verification, or constraint propagation, LLMs often lose coherence when asked to “think” over multiple logical steps.

2 Why Planning Matters: Lessons from Classical AI and Games

Classical AI systems were designed around explicit reasoning — from rule-based systems to STRIPS-style planning. In domains like chess or Go, systems like AlphaGo and AlphaZero demonstrated the strength of combining policy networks with explicit tree-based planning. These systems achieved superhuman performance by learning value functions and exploring action trajectories. The implication is clear: intelligent behavior often requires looking ahead, scoring options, and refining decisions based on structured representations. LLMs, which rely on autoregressive sampling, lack this deliberative apparatus.

3 Two Paths Forward: Reasoning by Training or Inference

To bridge this gap, researchers have developed two main strategies. Training-time reasoning, exemplified by DeepSeek-Reasoning, fine-tunes models with synthetic or curated multi-step tasks. The architecture remains unchanged, but the dataset structure induces the model to internalize reasoning behaviors. This is cost-effective and efficient at runtime but limited in flexibility. In contrast, inference-time methods like those used in O1 do not modify training but enforce reasoning behaviors via structured generation protocols. These systems generate intermediate steps and rationales at inference time, mimicking the process of planning without actual search trees or belief states.

4 Inference-Time Reasoning: Simulating Thought with Tokens

While language models lack an internal tree search, inference-time techniques simulate planning by treating intermediate generations as latent reasoning steps. Each token becomes a micro-action in a linearized plan. Models like O1 attempt to impose structure through prompt formats that encode subgoals, actions, and observations. This emulates a “reflect-act-observe” loop. However, the generation process remains strictly forward-moving and autoregressive, lacking any internal rollback or branching. Without memory or control flow, this planning approximation is brittle and compute-intensive.

5 From Prompting to Planning Agents

A growing direction treats LLMs as components in a larger agentic framework. These systems operate over a plan-act-observe loop, where the LLM proposes an action, observes the result (via a tool or environment), and revises future actions. Architectures like ReAct and recent open agent frameworks formalize this interaction pattern. Tool calling protocols like MCP standardize access interfaces, reduce sample complexity, and externalize planning structure. In such setups, reasoning shifts from being internal to being distributed across an orchestration layer.

6 What Reasoning Isn't: Why Token Sampling $\neq$ Search

Despite superficial similarities, token-level reasoning is not equivalent to symbolic planning. Classical planners operate over well-defined state spaces, allow backtracking, manage constraints, and propagate belief states. LLMs, even when guided by reasoning prompts, lack this fidelity. They cannot revise earlier steps without re-generating the entire sequence. Moreover, there’s no internal consistency check or verification step unless externally implemented. The implication is that while reasoning can be induced, it is neither stable nor grounded in model-internal representations.

7 Reasoning as the Next AI Substrate

The future of AI will likely hinge on hybrid systems that blend statistical fluency with explicit reasoning. This requires new primitives: memory modules for state persistence, symbolic scaffolds for constraint satisfaction, and interfaces that enable backtracking and revision. Whether realized through architecture (e.g., modular transformers) or orchestration (e.g., reasoning agents), reasoning will emerge not just as an emergent property, but as a core capability. As planning returns to prominence, the challenge is clear: to teach machines not just to generate, but to think.