Trivial Notes 2025

Hao Su

May 6, 2025

Contents

Welcome to the Era of Experience

Silver and Sutton 2025

In key domains such as mathematics, coding, and science, the knowledge extracted from human data is rapidly approaching a limit. And a sketch chronology of dominant AI paradigms is presented where era of *experience* is to arrive after era of *simulation* (RL) and *human data* (LLM).

The Era of Experience

Agents will inhabit streams of experience, rather than short snippets of interaction. This will allow them to take actions to achieve future goals, and to continuously adapt over time to new patterns of behaviour.

Their actions and observations will be richly grounded in the environment, rather than interacting via human dialogue alone.

Their rewards will be grounded in their experience of the environment, rather than coming from human prejudgement. Relying on human prejudgement usually leads to an impenetrable ceiling on the agent's performance: the agent cannot discover better strategies that are underappreciated by the human rater.

They will plan and/or reason about experience, rather than reasoning solely in human terms. It is highly unlikely that human language provides the optimal instance of a universal computer. More efficient mechanisms of thought surely exist, using non-human languages that may for example utilise symbolic, distributed, continuous, or differentiable computations. An agent trained to imitate human thoughts or even to match human expert answers may inherit fallacious methods of thought deeply embedded within that data, such as flawed assumptions or inherent biases ("an echo chamber of existing human knowledge"). Progressing beyond each method of thought required interaction with the real world: making hypotheses, running experiments, observing results, and updating principles accordingly.

RL

Early RL research yielded a suite of powerful concepts and algorithms. Concepts like options and inter/intra-option learning facilitated temporal abstraction, enabling agents to reason over longer timescales and break down complex tasks into manageable sub-goals (Sutton, Precup, and Singh 1999).

RLHF side-stepped the need for value functions by invoking human experts in place of machine-estimated values, strong priors from human data reduced the reliance on exploration, and reasoning in human-centric terms lessened the need for world models and temporal abstraction.

The era of experience presents an opportunity to revisit and improve classic RL concepts. This era will bring new ways to think about reward functions that are flexibly grounded in observational data. It will revisit value functions and methods to estimate them from long streams with as yet incomplete sequences. It will bring principled yet practical methods for real-world exploration that discover new behaviours that are radically different from human priors. Novel approaches to world models will be developed that capture the complexities of grounded interactions.New methods for temporal abstraction will allow agents to reason, in terms of experience, over ever-longer time horizons.

Consequences

Agents may even be able to exhibit capabilities previously considered the exclusive realm of humanity, such as longterm problem-solving, innovation, and a deep understanding of real world consequences. Future AI systems may be harder to interpret. This may also provide some safety benefits aside from risks.

Comment

The meaning of human being, should this era arrive, may be (1) consciousness and (2) intuitive reasoning (the latter may be related to vectors).

The agents should not do harm as long as they are asked not to.

What is reasoning? Is it the same as computing?

May 6

References

Silver, David and Richard S Sutton (2025). "Welcome to the Era of Experience". In: *Preprint of a chapter to appear in Designing an Intelligence, edited by George Konidaris, MIT Press (forthcoming).*

Sutton, Richard S, Doina Precup, and Satinder Singh (1999). "Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning". In: *Artificial intelligence* 112.1-2, pp. 181–211.