Emulating Human Learning and Evolution in Systems of AI Agents

Jiayu Chen

April 2024

1 Introduction

Ideally, AI agents should resemble digital human beings, with the capacity to discover, explore, and create new knowledge based on existing foundations. Large Language Models (LLMs) such as ChatGPT [4] have made significant progress in this direction by leveraging internet-scale human knowledge. However, the training of LLMs is resource-intensive, requiring substantial computational power and energy. For example, training a model like GPT-3 results in an estimated total energy consumption of 1,287 MWh [5]. Furthermore, LLMs primarily learn by imitating human knowledge, and due to their black-box nature, their reasoning and intelligence levels are often questionable. Therefore, a resource-efficient mechanism that enables AI agents to evolve continually towards acquiring human-level or superior intelligence is essential.

Research Gap: There is a pressing need for AI systems capable of emulating the stages of human cognitive development, learning and adapting autonomously from minimal interactions. Achieving this goal through human-engineered systems or predefined rules is notably challenging [7, 9]. Instead, it is imperative to develop an automated mechanism that can progressively improve within a given computational budget [10]. On the other hand, we notice that nature's ecosystem has enabled humans to evolve through a long history of interactions. By creating an ecosystem-like platform for AI agents and developing efficient self-evolution algorithms, we may facilitate the emergence of AI capabilities that are currently unforeseen, allowing agents to evolve in a process akin to human learning.

2 Methodology

Our primary objective is to create an ecosystem of AI agents, promoting unsupervised continual learning. The research content could specifically include the following aspects:

Ecosystem for AI Evolution: To enable human-like learning process, we aim to establish a digital playground for AI agents, simulating social environments crucial for advanced learning. Constructing such a platform involves creating an open world for AI agents, which demands the creativity and collaborative efforts of a massive open-source community. Our initial work seeks to lay the foundational efforts for this platform, encouraging further contributions from the global research community. Ideally, under a set of common criteria, each research institute could enhance the platform by integrating new environmental components, enriching the learning ecosystem for AI agents. Notably, each AI agent within the ecosystem can be based on Large Language Models. Through continuous evolution, these agents are expected to acquire abilities beyond the initial knowledge base and achieve this at a relatively lower computational cost compared to LLM training.

Reinforcement-Learning-based Evolution: Superhuman intelligence cannot rely solely on supervision from existing knowledge. Instead, learning should be objective-driven [3], where agents evolve by setting and achieving progressively complex goals within the ecosystem. Reinforcement Learning (RL) [11] is ideally suited for this purpose, but significant challenges remain:

- **Foundation Models for Deep RL:** Large-scale foundation models like Transformers [13] and Diffusion Models [2] have been successful in domains such as Natural Language Processing and Computer Vision. These models learn from vast amounts of data in a generalizable manner but are not inherently suited for the dynamic programming progress underlying RL. In this case, to facilitate the continual evolution of AI agents, it is essential to implement them with foundation models specifically designed for "Deep" Reinforcement Learning.

- **Efficient Online Planning:** Planning enables the agent to look ahead in a longer horizon to make better decisions and is an essential complement of RL [1]. So far, most successful applications of RL, such as AlphaZero [8], rely on planning. Thus, algorithms that integrate efficient planning and reinforcement learning for AI agent evolution are highly promising. The ability to plan ahead is also a symbol of intelligence.

- **Skill Discovery and Utilization:** Typically, in RL, agents are trained to generate decision/control sequences based on primitive actions. However, in a continual learning paradigm, humans develop and refine useful skills, subsequently forming strategies for complex tasks by combining these skills. Therefore, enabling agents with the capability for skill discovery and utilization is crucial. Moreover, the transfer of skills across similar tasks can enhance generalization; with acquired skills, agents are equipped to plan over longer horizons.

- **Fully Decentralized Multi-agent RL:** Concerning multi-agent RL, traditional centralized training paradigms, such as those referenced in [6], face scalability issues as the centralized state space can be vast, and learning a centralized function over all LLM-based agents is nearly impractical. Thus, the development of fully decentralized multi-agent RL methods is necessary.

- **Learning from Sparse Rewards:** Designing effective reward functions is typically crucial for RL success, but this can contradict the goal of self-evolution. Ideally, AI agents should learn under macro objectives, autonomously forming sub-goals based on naturally developed intrinsic motivations [12]. This requires the agent to learn its reward function while forming its policy, which is a challenging bi-level optimization problem.

3 Potential Impact

This research aims to bridge the gap between human and artificial intelligence, developing AI agents that learn continually, efficiently, and autonomously. Ideally, this approach will foster a new paradigm for AI development, complementing the current focus on Large Language Models and paving a resource-efficient path towards Artificial General Intelligence (AGI). There is no doubt that human-level or super-human AGIs, with reduced costs, would benefit nearly every aspect of our real life, through automation and intelligent augmentation of our daily work, decision-making, and entertainment processes.

References

[1] Dimitri Bertsekas. *A course in reinforcement learning.* Athena Scientific, 2023.

[2] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. In *Advances in Neural Information Processing Systems*, 2020.

[3] Yann LeCun. A path towards autonomous machine intelligence. *Open Review*, 62(1), 2022.

[4] OpenAI. Gpt-4 technical report, 2023.

[5] David Patterson, Joseph Gonzalez, Quoc Le, Chen Liang, Lluis-Miquel Munguia, Daniel Rothchild, David So, Maud Texier, and Jeff Dean. Carbon emissions and large neural network training, 2021.

[6] Tabish Rashid, Mikayel Samvelyan, Christian Schröder de Witt, Gregory Farquhar, Jakob N. Foerster, and Shimon Whiteson. QMIX: monotonic value function factorisation for deep multi-agent reinforcement learning. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 4292–4301. PMLR, 2018.

[7] Saber Sheybani, Himanshu Hansaria, Justin Wood, Linda B. Smith, and Zoran Tiganj. Curriculum learning with infant egocentric videos. In *Advances in Neural Information Processing Systems*, 2023.

[8] David Silver, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, Dharshan Kumaran, Thore Graepel, et al. A general reinforcement learning algorithm that masters chess, shogi, and go through self-play. *Science*, 362(6419):1140–1144, 2018.

[9] Linda B Smith. Can lessons from infants solve the problems of data-greedy ai? *Nature*, 2024.

[10] Richard Sutton. The bitter lesson. *Incomplete Ideas (blog)*, 13(1):38, 2019.

[11] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction.* MIT press, 2018.

[12] Richard S Sutton, Michael Bowling, and Patrick M Pilarski. The alberta plan for ai research. *arXiv preprint arXiv:2208.11173*, 2022.

[13] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems*, pages 5998–6008, 2017.