Journal for Educators, Teachers and Trainers
The LabOSfor electronic, peer-reviewed, open-access Magazine

Journal for Educators, Teachers and Trainers
The PROFESIOLab electronic, peer-reviewed, open-access Magazine

ISSN 1989-9572
DOI:10.47750/jett.2020.11.01.20

AGENTIC AI AND REINFORCEMENT LEARNING: TOWARDS MORE AUTONOMOUS AND ADAPTIVE AI SYSTEMS

[1]Varun Bodepudi, Applab, [2] Niharika Katnapally, [3] Vasu Velaga, [4] Chethan Sriharsha Moore, [5] Purna Chandra Rao Chinta, [6] Laxmana Murthy Karaka

Journal for Educators, Teachers and Trainers, Vol. 11 (1)

https://jett.labosfor.com/

Date of Reception: 18 Oct 2020

Date of Revision: 15 Nov 2020

Date of Acceptance: 21 Dec 2020

[1]Varun Bodepudi, Applab, [2] Niharika Katnapally, [3] Vasu Velaga, [4] Chethan Sriharsha Moore, [5] Purna Chandra Rao Chinta, [6] Laxmana Murthy Karaka
(2020) AGENTIC AI AND REINFORCEMENT LEARNING: TOWARDS MORE AUTONOMOUS AND ADAPTIVE AI SYSTEMS
. *Journal for Educators, Teachers and Trainers*, Vol.11(1)177-193

Journal for Educators, Teachers and Trainers
The LabOSfor electronic, peer-reviewed, open-access Magazine

Journal for Educators, Teachers and Trainers
The PROFESIOLab electronic, peer-reviewed, open-access Magazine

Journal for Educators, Teachers and Trainers, Vol. 11 (1)
ISSN 1989 –9572

https://jett.labosfor.com/

AGENTIC AI AND REINFORCEMENT LEARNING: TOWARDS MORE AUTONOMOUS AND ADAPTIVE AI SYSTEMS

[1]Varun Bodepudi, Applab, [2] Niharika Katnapally, [3] Vasu Velaga, [4] Chethan Sriharsha Moore, [5] Purna Chandra Rao Chinta, [6] Laxmana Murthy Karaka

[1]Systems Inc, Computer Programmer
[2]Techsmart Global Inc, BI Lead
[3]Cintas Corporation, SAP Functional Analyst
[4]Impetus Technologies , Sr Software Engineer
[5]Microsoft , Support Escalation Engineer
[6]Microsoft, Senior Support Engineer

Received: 18/11/2020 **published:** 21/12/2020

Abstract

This essay presents agentic AI, its functioning in a complex dynamic environment, and the need for developing AI systems that are more autonomous and adaptive, working as an agent. The conclusion of requirements for agentic AI, in terms of agentic capabilities and axiomatic values, leads to the basic design philosophy of constructing AI systems that approximate such agentic perception and axioms, rather than trying to explicitly describe or entangle human meta-morals into the AI systems. Thus, we propose to focus on designing more autonomous and adaptive AI systems that are more similar to agentic entities whose utility functions are developed from the evolutionary functional requirements. Reinforcement learning methods permeate mainstream AI research since they often provide the best empirical results. Due to that, we focus our exploration on research into integrating agentic capabilities with RL methodologies. On the basis of software prototypes, we exemplify the problem of needing to wait for a print job and an RL-based solution for a delivery service. Further exploration into the research goals and full conclusions are covered in the debriefing of each use case.

We introduce an AI research trajectory for developing AI systems that are more agentic and show how these will act more autonomously and be more adaptable by design. By figuring out the agentic capabilities of agentic AI systems and further exploring their values, we could

develop AI systems that better adapt to society, instead of embedding society into AI systems. Such agentic AI systems will not be perfect, but they will progress towards being more agentic. We present two use cases to demonstrate our ideas, both of which require the AI systems to adapt: a way of needing to wait for a print job and a delivery service. We ground our second use case in today's hardware using the rapidly developing field of reinforcement learning. Our second use case is our first step forward to integrate agentic entities and one method of how researchers are working on the final design milestone presented on our agentic AI trajectory of future AI systems.

Keywords: Agentic AI, Autonomous Systems, Adaptive AI, Reinforcement Learning, AI Trajectory, Utility Functions, Evolutionary Requirements, Agentic Capabilities, Axiomatic Values, Dynamic Environments, AI Design Philosophy, AI Research, AI Adaptability, AI Autonomy, Software Prototypes, Use Cases, AI Integration, AI Ethics, AI Evolution, Intelligent Agents.

1. Introduction

In this research area, we opt for more narrow and specific interpretations of agentic AI systems. These more narrow interpretations include systems that exhibit autonomy or adaptability, especially in dynamic and uncertain environments. There is a growing demand for AI with such properties for domains such as smart living environments and autonomous vehicles. Current undertaking goals in AI, whether we speak of machine learning or intelligent agents characterized by rules and deliberative reasoning, need to be conducted within a constrained and preferably static environment. Yet autonomous and adaptive solutions are difficult not only from a complexity standpoint but also due to the need to exist without human intervention or at least under limited intervention. Agentic AI systems do not guarantee human-like agents, but they certainly contribute towards AI that is less dependent and interwoven with human activities and lifestyles, allowing for a more independent existence.

Existing AI systems require a lot of tuning, either offline using historical data or even by training processes that are really time-consuming. Additionally, current AI systems can only handle tasks within a restricted boundary, requiring external human intervention or recovery processes for real-world applications. These systems only partially fit the childhood dream of AI. When a robot is turned on, it should be able to continue to make decisions until the off button is pressed or until its power runs out. Just like humans or animals, the system should react and tolerate minor annoyances. This type of approach is now challenging research and will probably continue to exploit the attention of researchers in both the near and middle future: the immersive set of the RoboCup competition comprises adversarial real-time challenges. As portrayed in the trends of AI and robotics, research shows that the continual development of adaptive autonomous agents is becoming an important factor in the AI community. Besides the behavior of agents being more human-like, we also seek more autonomous capabilities and mechanisms in AI. For instance, in the fields of robotics, some directions are on the development of more autonomous robotic agents.

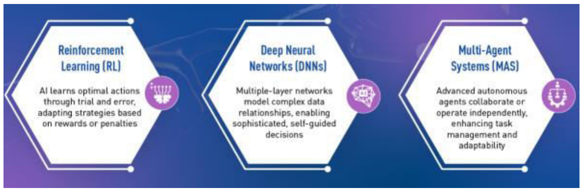

Fig 1 : Agentic AI

1.1. Background and Significance

Agentic AI and reinforcement learning have recently gained much attention. However, agentic concepts have a long history in ethology (proactivity, autonoetic consciousness, free will, and self-awareness) and across various psychological, cognitive, and neurological theories, human-computer interaction, as well as in robotics. In psychology, the concepts of proactivity, need for autonomy, and internalization have been central to self-determination theory since they are found to be central to motivation. The discussion of agentic concepts of AI has also become one of the last peak phases after the life cycle of constitutive properties. Before then, the emphasis of much early AI was on declarative properties. The role of AI was simply to create and manipulate knowledge representations. Maturana and Varela's circular cybernetic framework suggests two distinct phases: a generative and an operational phase. AI had mostly stayed in the representation generative phase throughout the symbolic AI term, but it has recently shifted toward more agentic models. Hence, an agentic capability is important to make AI principles reflect in AI systems in practice.

In many areas, the tasks are getting more complex and multi-dimensional, such as for education support (acquiring facts, understanding, enabling latent knowledge, sharpening practice); assistive technology (vehicle navigation, route planning, shopping recommendations); multimedia (natural language processing of texts, translation, and multimodal communication); and gaming (for story-building and automated planning as done in computer games). For example, the robot learned to carry out tasks in game worlds by emulating human players' strategies to learn and act autonomously, just as an embodied agentic agent would learn to "enact" - to modify its world through sensing toys in its environment and acting accordingly to its desires and motivations. So, the time is ripe for AI to formalize agentic capabilities of systems both theoretically and practically, beyond conventional symbolic methods. A review of much cybernetic AI or cyber AI research according to a theory of the philosophy of enaction, embodiment, autonomy, and cognition can be found in updated literature on this part. A conceptual review of other agentic AI uses distributed AI and concentrated on practical issues. Such old literature clearly takes up some aspects of agentic AI concerns at the time, but our focus is more general than these works. This review considers agentic AI more generally and primarily uses some of the extent to bring our point home with no specific theory reliance. For obvious concept-related ramifications of our AI review, interested readers are encouraged to read such criteria theories.

Equation 1 : Agentic AI Decision Function

$$A_t = \arg\max_{a \in \mathcal{A}} U(s_t, a)$$

A_t - Optimal action at time t,

$\mathcal{A}$ - Set of possible actions,

$U(s_t, a)$ - Utility function of action a in state s_t.

1.2. Research Objective

The goal of this work is to examine the convergence of two currently distinct domains: agentic AI and reinforcement learning. Our objective is to investigate an agentic AI system characterized by adaptation and autonomy, exploring aspects where the two domains interact. In particular, we aim to answer the following research questions: How can characteristics of agentic AI be linked to reinforcement learning to create AI that makes decisions autonomously? Furthermore, when designing autonomous and adaptive agents, we must consider both areas, meaning we also need to investigate possible conflicts between them. Consequently, we aim to answer the following: In what ways do agentic characteristics influence the application of reinforcement learning for agents?

Currently, AI agent systems are developed almost exclusively in the domain of OSGC, while the parallel is true regarding agent characteristics—research has largely avoided the domain of AGCS. As a result, essential questions pertaining to how systems can improve their performance through learning and addressing gaps in the literature that focus on the development of agentic characteristics and autonomy in AI agents abound. Reinforcement learning, inter alia, is a domain of AI research that, when applied to an agent system, is expected to produce beneficial performance improvements. This class of AI systems with applied learning is recognized in the research domain of AI and is considered a form of 'adaptive agent.' Reinforcement learning specifically aims to make agents more autonomous concerning their decision-making capabilities. Thus, we argue that a correlation exists between the trained agents of reinforcement learning and their possession and exhibition of 'agentic' characteristics; as such, we posit that the inclusion of reinforcement learning in an AI system for an avatar character is conducive to inclusion in the research of AGCS. Our research will be implemented by experimental means, involving the application of reinforcement learning to a 2D AI agent system and observing the effect on its decision-making in a computer game.

2. Foundations of Agentic AI

In recent years, the concept of agentic AI began to take shape. According to established definitions of agentic systems, every agentic system is an autonomous agent, i.e., a system that can interact with the environment in an autonomous way in order to satisfy its own preferences. In this paper, we try to develop a detailed concept of an agentic system that might be an AI system, rather than a human.

In verification, just as in real life, one poses different verification questions that are meant for different kinds of systems. In the verification of agentic systems, these questions are very naturally cast in terms of the autonomy characteristics of the considered systems. For such kinds of questions, it is crucial to take into account the framework in which an agentic system operates, i.e., the assumptions made about the environment of an agentic system. In this paper, we introduce the conceptual framework that describes the relationships between an agentic system and its environment.

Following the above model, we say that an AI system is an agentic AI system when it is designed for the following properties: 1. The autonomy property states that the AI system can execute its intended purposes in interaction with its environment (the framework provides the set of assumptions made about that environment). It is the state of being self-governed or self-directed, effectively operating without control by humans. 2. The adaptivity property helps a system identify when its performance has not been successful, as well as pose a behavioral, functional, or operational adjustment to try to remedy the matter.

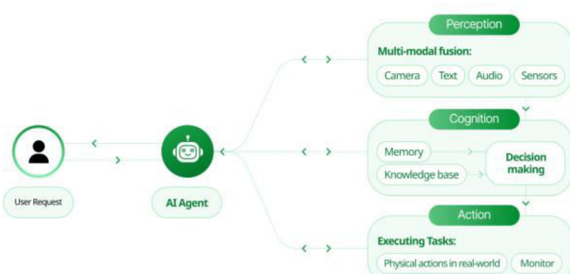

Fig 2 : Agentic AI Architecture

2.1. Definition and Conceptual Framework

Given the importance of shared definitions, particularly in fields of rapid and ongoing change, here we provide a robust definition and associated conceptual framework of agentic AI, intended to be well-grounded in existing philosophical and conceptual theory. The contemporary AI revolution is based on at least two essential phenomena: new technological possibilities to cope with big data and processing, and new conceptual and paradigmatic challenges that forced the development of new theoretical frameworks. Definitions and conceptual models will help the discussion to run homogeneously and pointedly, helping to define critical research issues in AI, including the kind of agency achievable in an agent-based AI, the perspective of a multi-tier or multi-strategy AI, interpretable and/or accountable AI, and the kind of hybridization among diverse AI strategies. After this tentative advance, the forms of agentic AI can evolve at different levels, finding ways and forms of hybridizing the machine learning paradigm, integrating, where possible, other paradigms of Adaptive AI, such as pre-symbolic/symbolic, logical and non-logical, already known in the history of AI.

In general, an agent is characterized by cognitive capacities of some kind. A more focused definition of what we call a "person-agent" can be as follows: A person-agent is an AI entity with cognitive capacities (we pose here a particular interest in the autonomous ability), but these cognitive capacities depend on the underlying AI techniques and paradigms. The term "cognitive" should be explained because different models of such capacities can be found in the AI domain. Generally, in AI, "cognitive" capacity may include, without any claim of universality, the functions of perception, abstraction, classification, simulation, prediction, anticipation, decision, motivation, attention, learning, planning, different kinds of language problem-solving, consciousness problems, abilities of navigation, adaptation, self-programming, and so forth. Using the kind of idiosyncratic definition we have chosen for “AI,” any combinatorial and intelligent capacities can also be presented by human operators with dedicated, specific, and sometimes signal-guided numerals, simplifying, or automating tools.

2.2. Key Components and Characteristics

Having outlined the concept of agentic AI, one can now describe its key components, or the characteristics of AI that may make it agentic. Important components that are characteristic of an agentic system include increased decision-making capabilities, adaptability, and learning due to their value for an AI system while deciding on its actions. To plan, an AI system needs to evaluate branching points in internal state space, including taking into account dynamic and evolving contexts. Finally, as uncertainty is an inherent characteristic of AI, agentic systems should have higher learning processes to refine their performance. The characteristics of agentic AI systems pertain to their effectiveness of operation in broader contexts or deployment. Agentic AI can be characterized as having or consisting of: - autonomy; - the ability to learn about internal and external states. Combining various decision-making processes can provide requisite decision-making capabilities. But in addition to decision-making, agentic systems should contain learning about external world states. Visioning is just one example of such a process, but essentially any decision-making system can be complemented with various learning processes. Agentic AI systems therefore differ from traditional or classical AI systems. Moreover, such systems provide a more natural form of human-computer intelligence because the AI will respond to how we interact with it; by adapting new strategies and learning curves. There are, of course, significant technical challenges associated with delivering such a technical solution.

3. Reinforcement Learning in AI

Reinforcement learning (RL) is a learning paradigm in AI, which is about agents that directly learn by interacting with an environment. The agent perceives states in the environment, selects actions, and receives reward signals based on those actions. This way, RL does away with the explicit signaling of correct and wrong answers in the training data that is necessary for learning in more classic machine learning (ML) paradigms. The fundamental principle of RL is to set up a learning problem in terms of an agent that interacts with an environment. The agent seeks to maximize a reward signal generated by the environment, typically selecting actions at each time step. The environment's dynamics are typically specified as a Markov decision process (MDP), providing a mathematical description of states in the environment, actions available to the agent, the reward signal, and the stochastic nature of the resulting state transitions.

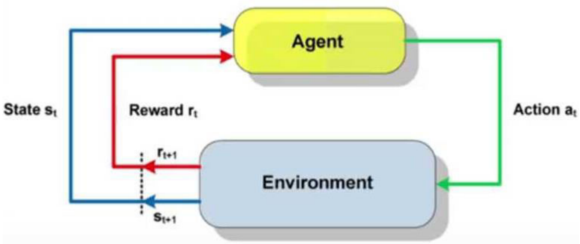

Fig 3 : Reinforcement Learning in AI

A key objective of RL research is to enable AI systems that are more adaptive and autonomous than existing systems. RL has naturally led to the design of a variety of learning algorithms, such as value-function and policy-based approaches, that stem from approaches in optimal control theory. These algorithms can be further combined in practice through strategy search and model-based control techniques that facilitate high-level decision-making over a large decision space. Some learning techniques model uncertainty to reach high rewards faster and more safely, often combining this modeling in parallel with learning frameworks. Researchers in neuroscience and related fields have been interested in a variety of RL algorithms, especially with respect to understanding computational learning mechanisms in the mammalian brain. RL is already having a significant impact in domains such as robotics, personalization software, and game playing. AI's capacity to learn within the RL framework will have a greater impact as adaptability becomes imperative in autonomous systems. The adaptability already visible in modern AI systems, for instance in perception, is due to supervised learning from larger and larger datasets. Without adaptability, the performance of reinforcement learning AI will degrade if executed on the same relatively stationary policy within the same environment. Currently, people have to handcraft and continually update reward functions to maintain adaptability in real systems, revealing the early promise of the field in robotics.

There are, however, multiple difficulties in training AI agents in real systems to date. First, there is the difficulty of training to perform optimally and safely in nondeterministic environments. Second, if the agent does not execute the same policy as was trained on, to achieve reliable behavior, it has to be robust to nonstationary environments. Other problems in real systems may result from difficulties in modeling the environment, modeling the reward function in the environment, or a small fraction of functionally different states that are hard to generalize across, for instance due to safety requirements. Another challenge is the ability to make learning decisions in natural environments without restarting after a failed task, where humans are able to store information about unsuccessful plans or actions for multiple steps in the future to increase the chances of a better outcome on the next sampling step. The rejection of agents that have been trained with state-of-the-art deep reinforcement learning to make human-lethal mistakes has already been presented. This result is so far the only demonstrated human safety impact in reinforcement learning.

Equation 2 : Reinforcement Learning Value Update (Q-Learning)

$$Q(s,a) \leftarrow Q(s,a) + \alpha[r + \gamma \max Q - Q]$$

$Q(s,a)$ - Q-value,

α - Learning rate,

r - Reward,

γ - Discount factor.

3.1. Overview and Principles

Reinforcement learning (RL) is an approach to gradual agent learning where AI agents maximize accumulated reward by selecting and tuning a policy (one that describes which actions to choose in a given state) using predictions that connect perceived world states and possible actions. Perceived consequences of an action are used to adapt this policy. Two mechanisms drive this agent learning. Intrinsic motivations push the agent to visit rarely visited state-action pairs, or novelties. The accumulation of reward, on the other hand, enables an agent to bias its policy towards actions that were found to be successful in the past. This trade-off between visitation and exploitation of behavior is commonly described informally in terms of exploration vs. exploitation. A mathematical framework can be used to find an optimal policy according to an optimal value function. Using this optimal value function as a baseline gives us the single global maximum of a given value function to provide direct feedback to steer learning. Specialized optimization techniques exist to find this optimal policy, like policy iteration and value iteration, that estimate good value functions through repeated sampling and updates.

Indeed, this is mostly why RL models are so successful compared to classical optimization schemes. The system can incrementally learn better policies over time. An agent using these learning processes can improve its strategy and learn an optimal policy over time. In most contexts, estimated values smoothly converge within local optimal strategies and may require more dedicated exploration mechanisms. Reinforcement learning provides various ways to approximate value functions or policies based on the learning techniques. The exploration part that helps in adding variance to estimated value is the primary preference for criteria for this classification. It distinguishes methods that learn a model of the environment dynamics to use it explicitly in a search algorithm from those that directly model value functions or policies. It also identifies the major learning techniques used in each group. Additionally, agents utilizing model-based RL create an internal simulation of reality to make more advanced decisions to maximize their rewards. This is done by iteratively improving the internal dynamics model of the world by fitting internal parameters to the outcomes. While the explore vs. exploit trade-off remains core to the approach, this re-examination of the dynamics model means they do not need to rely on their action-value functions.

3.2. Applications in AI Systems

Reinforcement learning can be seen in various AI systems, where AI agents are able to interact with their dynamic environment. Taking a look at application domains, relevant case studies reveal successful utilization of RL techniques in problem solving within robotics, gaming, autonomous vehicles, and more. Despite their functional diversity, the systems have in common that they aim to learn and act robustly in highly dynamic, partially observable environments. Technical issues include designing agent-environment interactions, representing the state of a game or system, creating temporal-difference-based learning algorithms to evaluate and improve potential future strategies, methods to find good exploration policies, background knowledge integration through eligibility traces, as well as reapplying the learned policy in additional, comparable scenarios.

Technical problems in real applications include reasoning and planning under uncertainty, decision-making using available sensors, as well as the formalization and assessment of learning objectives. The underlying objective of RL in the context of AI systems is to develop capable, adaptive agents that are able to learn how to complete a given challenge or objective more and more efficiently. A crucial aspect in the case of practical AI applications is the continuous learning of the agent with a continuously changing or highly dynamic environment. Some of the successful applications of reinforcement learning also offer the flexibility of applying these methods to complex, AI-enabled systems to solve technically motivated problems. In the context of autonomous robots, these autonomous AI systems have shown high flexibility and capability in accomplishing specific and engaging tasks. In addition, reinforcement learning represents an important step towards making AI systems increasingly autonomous and adaptive since these rely on the capability for autonomous learning. Regarding the kinds of systems that are enabled by AGI, only a few arguments about blue sky ideas have been made so far.

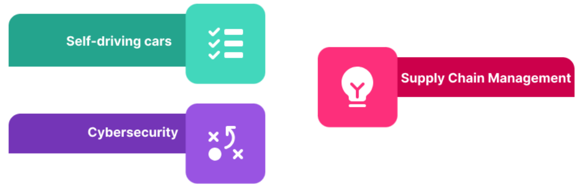

Fig 4 : Applications of Agentic AI in Real Life

4. Integration of Agentic AI and Reinforcement Learning

Despite their rooted conceptual basis in planned acting and situatedness, attempts to merge reinforcement learning and agentic AI have been few and preliminary. However, an integration of these two areas can give us novel insights into how agents develop the properties underlying the agentic AI principles of agency, proactivity, theoretical mind, and anticipatory modeling. It can help us experiment within specific theoretical frameworks in which these properties arise and could therefore enhance the autonomy and adaptability of AI systems in the wild.

A stronger synergy between both agentic characteristics and RL techniques will most likely foster better models of actual agent behavior, such as RL algorithms that employ reward shaping as a way of learning optimal policies that include the correct depth of affordance-footprint observations, the use of internal models of others to improve learning and performance in multi-agent systems, or the capacity of an AI agent to make increasingly accurate generative models through the active pursuit of novelty with the intention of maximizing the predictive information of the model. However, there are also unique insights to be gained by exploring the synergy between RL and the agentic properties: for example, an intelligent agent shows that the introduction of proactivity into a reinforcement learning agent can improve its performance by reducing the environmental search space of the system. Moreover, several of the tools and primitives found in architectural and functional AI and architectures can be nicely and easily modeled using RL and thus combined together in the AGI framework.

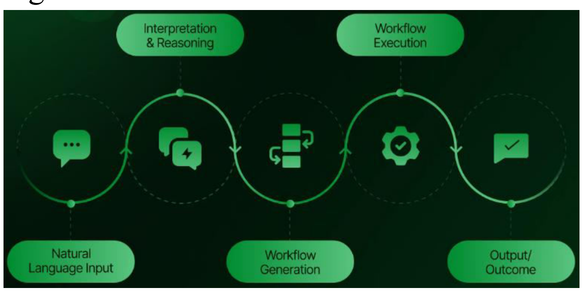

Fig 5 : Agentic AI Architecture

We accept that these are the first steps on a journey which, if fruitful, promises to bring us an AI — and a psychology — more in line with the agentic principle of RL, namely, acting for an expected future. Rather, we must take into account that the integration of agentic principles with mechanisms from a closely related domain still carries significant potential for breakthrough discoveries. For the Principles of Agentic AI, the combination with mechanisms from a domain where agentic characteristics also apply within the context of task accomplishment holds promising, novel insights and added value in its potential to establish a new theoretical framework. However, due to different scopes, research foci, and perspectives, the risks include integration that lacks adequate definition or the development of models that take into account interdisciplinary perspectives as well. There is often no formalized process to leverage the added value of an interdisciplinary perspective. The purpose of this paper is to experimentally fill this gap in the literature and exemplify, formally, the outcome of integration. Given this line of thought, we assume that by merging this approach into a domain connected to task planning and autonomous agent modeling, such as reinforcement learning, research could produce interesting insights to support possible breakthroughs. Although both internal and external mechanics refer to different levels of portrayal in AI, they do concentrate on the same phenomena. Socio-cognitive abilities demonstrated by an internal mechanism at the theoretical level represent an agent's path towards intelligent assisting behaviors. Moreover, we observe the ability of RL-based AI to "imagine" the future supported by the ability of reinforcement learning agents to develop internal models.

4.1. Theoretical Framework

Building on the discussion presented in section 3, this subsection aims to provide a more in-depth and theoretically grounded elaboration of the concept of agentic AI in combination with the field of reinforcement learning. To this end, a robust framework is introduced that outlines the distinct instrumental and conceptual paths towards converging agentic AI with RL. The framework distinguishes between different dimensions, including "how to model" approaches – i.e., a number of conceptual and practical models outlining the interaction between conduct and learning algorithms, together with an analysis of their resulting dynamics. Furthermore, we also outline "what is to be expected" based on these models – i.e., a series of developed heuristics showing which effect agentic features can be expected to have on their learning algorithms. In the remainder of this section, furnishing our theoretically grounded concept, we will outline the detailed framework for the integration of agentic and RL characteristics; the "how to model" agentic RL are the models of the interaction between learning and the agency, ambition, adaptive traits, and their dynamics and the "what is to be expected" regarding such interaction, shedding light on their potential.

Modeling Agentic Learning Integrating instrumental and conceptually motivated paths, the three models form the core of our theoretical framework, aiming to develop approaches that merge agentic behavior and adaptive learning. This integration is subject to large interest because not only do agents have to make and execute plans, but those plans might also have adaptive components: instead of (fully) searching an action plan space, one might choose to heuristically or greedily try out actions and adapt the plan as events unfold. Furthermore, the notion of "cost" of such inference and power - ambition - figures prominently. Model I is an abstract framework that can accommodate all ways in which simple behavior can be modeled, separate from the chosen learning. A general takeaway is that the more variance there is in the planned base behavior, likely the longer it takes to learn. In Ia we model learning based on abstract planned and current behavior, in Ib we model learning as an increase in realism of modeled "goals", and in Ic we model learning as being based on the difference between actual and preferred state. Each model section closes with mathematical instantiations for this abstract model: for Ia we review standard Q-learning in Cognitive Modeling and add adaptive planned intrinsic motivation, clarifying the dynamics and computational implications; in Ib we present another Q-learning variation, representing poverty-directed heuristics via adapted intrinsic motivation, again clarifying dynamics and computational implications. Finally, for Ic we represent these heuristics more simply with quick and slow learning with causally affecting a modifiable or non-modifiable variable. Model II explicitly describes a model for agentic behavior in the form of the Internal-External framing in this issue but does not incorporate a description of learning. Model III provides a simple model of instrumental learning in a dual-model way very familiar to both economics and psychological literature. All possible model combinations are possible: agentic and non-agentic learning towards agentic and non-agentic behavior. In sum, central implications of our framework concern the desirability and adaptation of ambition, along with the benefit of varying the realism of plan-based and plan-modification learning policies in a dynamically changing world. Importantly, this framework outlines a roadmap of foundational questions concerning the interrelation of agentic learning and octal control that critically rely on fluctuating task and other parameters, thus retaining high practical relevance.

4.2. Case Studies

In this section of the paper, we review the literature on scientific advances towards employing agentic AI with reinforcement learning (RL) based on case studies and implementations, spanning very different contexts, using methodological approaches such as interviews, case study development, or expert group discussions. The papers in this section look at how applying agentic AI in tandem with RL can work effectively, the challenges that are still present in getting agentic principles to work at all levels of an autonomous system, and how the combination of agentic and RL paradigms can improve adaptability and autonomy in practical implementations.

Case studies from industry show current RL algorithms integrated into game engines and virtual simulators in order to train game agents or autonomous agents are usually of a higher abstraction level, including Unity's ML-Agents and OpenAI's Gym. By contrast, the investigations in the context of Factorio, Pathologic 2, and The Turing Test, respectively, have looked at closer-to-prototype implementations, applicable to robotics or real-world scenarios, where a smaller, approximated environment of the system to control is extracted. Not all of these investigations led to an agentic handling of the system; the case of YuMi instead utilized multi-agent reinforcement learning (MARL) to train a second "wiki-captive" robotic arm that holds the gameplay arm. Following their investigation of agentic robotics in the screwdriver control chapter of this book, there have been parallel attempts to create an agentic control of "Engineer" in Factorio using frameworks generated around a library of implemented MoveIt controllers. There has also been a venture into the use of RRT in a project for safely navigating a forklift, which is an agentic AI approach for DM Planned control rather than RL, but is easily compatible with RL principles and can perform similarly to RL for some high-level motion planning cases. Evaluations of these case studies have shown that, while agentic behavior can be difficult to obtain at lower system levels in practice, combining RL with agentic AI at an embodiment level can result in systems that are more autonomous, adaptable, and advanced in related technical performance metrics. Interviews conducted at the lifting gear factory as well as the integration of end-users and industrial partners found that combining agentic AI with RL was viewed as necessary due to "imprecisions and non-exactness" or "dependency on changes of given environments and path feasibility" and that agentic "black-box protocols" bring several added benefits over Gaming AI or a traditional controller.

5. Challenges and Future Directions

One of the main challenges—which right now can be seen as an opportunity—in autonomous AI is whether people actually want to have completely autonomous systems, for example in customer service. We need more research to find out how to support and healthily balance autonomy and human oversight regarding the specific context and the social norms society will agree on. This will be different from context to context and between different cultures and countries, and AI should be able to adapt to this. One might also wonder what the benefits are of completely autonomous systems over systems that need human supervision and approval. In safety-critical domains, which are often put forward as interesting applications of agentic AI, where do the benefits of autonomy weigh up against the costs of mistakes an autonomous system could make? The public's acceptance of such systems is highly conditional and rightfully so. Due to the increasing autonomy of AI systems, the identification of accountability and response decision mechanisms is critical, also from an ethical perspective. To establish who is accountable when a decision does not meet expectations, AI should be provided with information to justify why a certain action has been chosen. However, without transparency, a condition must be met to ensure the ability of systems to explain their decisions or to provide the concerned users with knowledge about the AI process. Technological limitations also need to be tackled: now, AI is limited by the data available, the time and computational resources required to perform actions and learn. Long training time or very high requirements regarding the number of samples can be problematic in various scenarios. Improvements in prediction methods could be achieved if the networks were pushing more and more efficiently in learning algorithms. By learning to predict the future actions and events of others, the next action of a new, unfamiliar object could be learned by simulating the consequences of such an action with a physically plausible model. Additionally, unsupervised learning is required to handle uncertainties, such as those concerning the intentions of others.

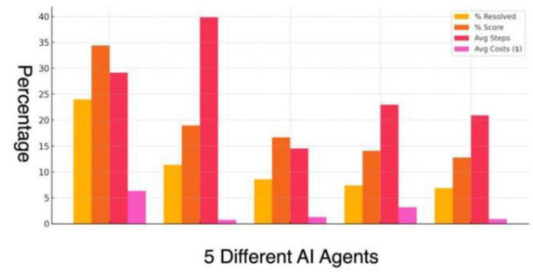

Fig 6 : Benchmarked Ai Agent Performances

Demanding unsupervised learning will lead to more general and robust representations thanks to data augmentation and the ability to cope with data anomalies. As explained, it is important to link together ethical reasoning and technological developments, needing thoughtful collaborations between experts in different fields. Initiatives already exist to establish such discussions and partnerships. Reminiscent of the concept of value sensitive design, some methods and tools have been suggested for intersectional collaboration between ethicists and technologists through values identification workshops. As an AI community, we will critically develop so-called value assessments that consider the predictability, autonomy, and intelligence of AI and its consequences. This perspective not only entails considering the consequences of technologies and including certain groups in the norms, but it also gives insight into self-organization, collaboration, and the overall development of technological innovation. This is more in contrast to technical ethics; we can highlight the following four unique contributions: 1. Technical. We limited our perspective to RL and AGT in the context of

AI systems being used in society generally, rather than being used for a specific sector or in a specific location. These AI systems are also software-only and, therefore, do not physically interact with the environment. It represents a class of AI systems used in society and is generally developed for use in society. In other words, they are not research-grade systems designed for demonstrations or competitions. We are looking at AI systems as they are designed to be used, rather than AI technologies. As such, we have prioritized practical considerations.

5.1. Ethical Considerations

There are a variety of ethical considerations that arise when discussing agentic AI and reinforcement learning-based systems. If an AI agent is increasingly granted autonomy in its decision-making processes, then the moral implications of these systems' designs become pertinent. At the very least, the architectures of these agents must be created with an eye toward future ethical norms and a consideration of how advances in AI may challenge existing moral intuitions. Tied directly to this consideration is the question of what the level of behavior regulation in these systems should be. These systems indeed require ethical guidelines for behavior that permit these AI agents to harness their knowledge in ways that can solve tasks or maximize value. These guidelines should encourage cooperation, understanding, and the respect of social and legal norms as much as expertise in a particular domain. Crucially, other interdisciplinary insights such as psychology and legal doctrine can help in developing ethical frameworks for advancements in AI, and legal scholars should be part of the necessary collaborations in addressing these questions. There are further discussions on the legal and regulatory impacts of agentic AI, arguing that they may be insufficient for application in critical systems. Similar concerns exist that as AI system complexity grows, the need for regulation becomes increasingly difficult to achieve. Thus, while regulatory frameworks will be necessary, they are not sufficient and ethical considerations must be intertwined with the development of these AI systems. There are also a variety of societal concerns around the marriage of agentic AI and reinforcement learning that should create further ethical hesitation in their deployment. A growing body of work has begun to grapple with the societal and ethical impacts of deploying these kinds of agentic AI technologies. Given the potential use of AI technology in domains such as health care and autonomous vehicles, it is crucial to focus on the primary concerns in these two fields: human safety and welfare. The autonomous deployment of these technologies in heavily regulated markets such as drugs or vehicle manufacturing could be further complicated by insurance market interests.

Equation 3 : Adaptive AI Policy Optimization (Policy Gradient)

$$\nabla J(\theta) = \mathbb{E}\left[\sum_{t=0}^{T} \nabla_\theta \log \pi_\theta(a_t|s_t) R_t\right]$$

$J(\theta)$ - Policy objective function,

$\pi_\theta(a_t|s_t)$ - Policy function parameterized by θ,

R_t - Expected reward,

T - Time horizon.

5.2. Technological Limitations

Currently, we face several technological limitations that hamper the further development of agentic AI and reinforcement learning systems. Computational resources continue to place stark limits on the training of large, complex networks over protracted periods. Limited data: Experimental data is expensive, and problems are easy to describe but hard to solve. User engagement in interactive systems is low. For some tasks, company data might be available, but access is restricted due to intellectual property and privacy concerns. Moreover, the quality of the data at our disposal might leave much to be desired, seriously complicating the training of AI agents. When dealing with augmented environments or applications, the latency between the AI agent and its environment grows with the number of layers through which the requests and responses need to commute. These communication times represent an inflexible computational bottleneck, which tightens even further due to the rapid deployment of geo-distributed services. Finally, although popular reinforcement learning algorithms are based on a rich body of theory, their guarantees do not apply to complex problems or evolve over the course of the learning process, making the optimization field empirically oriented. While there are techniques for making existing testing procedures for reinforcement learning more efficient, estimating the key metric, finding relevant actions to test, re-prioritizing actions to re-test, and balancing between exploration and exploitation to maximize the quality of a policy or test a distribution rather than a function or black-box policy, this part has not been explored. Moreover, to get a sample set of the environment, a user needs to collect an amount of data that demonstrates effective coverage for the entire distribution of the environment, as well as the underlying loss function in order to select objects in the environment. This has not been considered and would warrant its own research. Altogether, the ability of a system to function as a self-improving, agentic agent in a real-world environment remains an open question.

6. References

[1] Ganesan, P. (2020). DevOps Automation for Cloud Native Distributed Applications. Journal of Scientific and Engineering Research, 7(2), 342-347.

[2] Chintale, P., Korada, L., Ranjan, P., & Malviya, R. K. (2019). Adopting Infrastructure as Code (IaC) for Efficient Financial Cloud Management. ISSN: 2096-3246, 51(04).

[3] Gollangi, H. K., Bauskar, S. R., Madhavaram, C. R., Galla, E. P., Sunkara, J. R., & Reddy, M. S. (2020).Unveiling the Hidden Patterns: AI-Driven Innovations in Image Processing and Acoustic Signal Detection. (2020). JOURNAL OF RECENT TRENDS IN COMPUTER SCIENCE AND ENGINEERING (JRTCSE), 8(1), 25-45.

https://doi.org/10.70589/JRTCSE.2020.1.3.

[4] Ganesan, P. (2020). Balancing Ethics in AI: Overcoming Bias, Enhancing Transparency, and Ensuring Accountability. North American Journal of Engineering Research, 1(1).

[5] Gollangi, H. K., Bauskar, S. R., Madhavaram, C. R., Galla, E. P., Sunkara, J. R., & Reddy, M. S. (2020). Exploring AI Algorithms for Cancer Classification and Prediction Using Electronic Health Records. Journal of Artificial Intelligence and Big Data, 1(1), 65–74. Retrieved from https://www.scipublications.com/journal/index.php/jaibd/article/view/1109

[6] Sikha, V. K. Ease of Building Omni-Channel Customer Care Services with Cloud-Based Telephony Services & AI.

[7] Sarisa, M., Boddapati, V. N., Patra, G. K., Kuraku, C., Konkimalla, S., & Rajaram, S. K. (2020). An Effective Predicting E-Commerce Sales & Management System Based on Machine Learning Methods. Journal of Artificial Intelligence and Big Data, 1(1), 75–85. Retrieved from https://www.scipublications.com/journal/index.php/jaibd/article/view/1110

[8] Ganesan, P., Sikha, V. K., & Siramgari, D. R. TRANSFORMING HUMAN SERVICES: LEVERAGING AI TO ADDRESS WORKFORCE CHALLENGES AND ENHANCE SERVICE DELIVERY.

[9] Manikanth Sarisa, Venkata Nagesh Boddapati, Gagan Kumar Patra, Chandrababu Kuraku, Siddharth Konkimalla and Shravan Kumar Rajaram. “The power of sentiment: big data analytics meets machine learning for emotional insights”, International Journal of Development Research, 10, (10), 41565-41573.

[10] Sikha, V. K. INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING.