REPORT 19

Mapping Possible Futures with Speculative Design and Deep Reinforcement Learning

[ANONYMIZED REPORT — ALL IDENTIFYING INFORMATION REMOVED]

Abstract

This paper explores the use of Deep Reinforcement Learning (DRL) and speculative design to project future scenarios for emerging technologies. Using speculative design, signals and trends that influence the advancement of DRL in different contexts were identified. Tools such as PESTEL Analysis, the Futures Wheel, and the Cone of Possibilities were applied to map possible and preferable scenarios, with a focus on sustainability and innovation. The analysis resulted in a future scenario that uses DRL to address complex challenges, highlighting both the contributions of speculative design and the ethical and social challenges of implementing these solutions at scale.

1. Introduction

1.1. Presentation of the topic

Deep Reinforcement Learning, represented in this paper as DRL (Deep Reinforcement Learning), is a technique that combines the concepts of deep learning (Deep Learning - DL) and reinforcement learning (Reinforcement Learning - RL). DRL brings together the ability of deep learning to process and interpret large volumes of unstructured data with the ability of reinforcement learning to make optimized sequential decisions based on continuous interactions with the environment. These are systems capable of making complex decisions and learning optimal behaviors through interaction with a dynamic environment.

The use of DRL has stood out in several applications, ranging from games to medical diagnostics.

1.2. Challenges in the field

Despite its potential, DRL faces several challenges. Generalization is a central concern, since many DRL algorithms tend to overfit training environments, making application to new contexts difficult. In addition, computational efficiency is a significant obstacle, given that training DRL agents can be extremely costly and time-consuming.

Scalability is also a challenge, especially in large-scale problems, such as the traveling salesman problem, where the number of possible solutions is vast and complex.

1.3. Literature exploration

The foundations of reinforcement learning (RL) were established between the 1950s and 1980s with the introduction of Markov Decision Processes (MDPs) by Richard Bellman and the development of the temporal difference (TD learning) method by Richard Sutton. During the 1990s, RL was formalized as a field of study, with the book *Reinforcement Learning: An Introduction* by Sutton and Barto becoming an essential reference.

In the 2000s, advances in computational capacity and training techniques highlighted the field of Deep Learning, with works such as those by Hinton and Salakhutdinov on Deep Belief Networks (DBNs). In the 2010s, advances involved the integration of RL with DL, such as DeepMind’s Deep Q-Network (DQN), which combined RL with convolutional neural networks to play Atari games from raw pixels.

Recent advances include the development of new algorithms to improve training efficiency and stability, such as a DRL-based hyper-heuristic framework to solve combinatorial optimization problems. Works such as [Kallestad et al. 2023] and [Li et al. 2021] demonstrate the flexibility and power of DRL in addressing complex optimization problems, expanding the scope of possible applications in various industrial and logistical areas.

1.4. Structure of the paper

This paper is structured beginning with the introduction, which presents the context of Deep Reinforcement Learning (DRL) and its challenges in the field of Information Systems. The theoretical background section provides a solid theoretical basis, discussing the main concepts and advances in the area.

Next, the speculative design section details the methodology used to explore future scenarios of DRL, incorporating techniques such as PESTEL Analysis, the Futures Wheel, and the Cone of Possibilities. Subsequently, a mapping of the current state of DRL is presented, with signals and trends for the future. The speculation of possible futures and the projection of desirable scenarios are explored based on emerging tools and technologies.

Finally, the paper presents an IT solution designed for a desirable scenario, culminating in the conclusion, where the main contributions and identified challenges are discussed.

2. Theoretical Background

The foundations of reinforcement learning (RL) were established between the 1950s and 1980s. Richard Bellman introduced Markov Decision Processes (MDPs) and the Bellman Equation in the 1950s, providing the mathematical basis for policy optimization in stochastic environments [Bellman 1957]. In the 1980s, Richard Sutton developed the temporal difference learning method, which improves value estimates based on the difference between expected and received rewards [Sutton 1988].

In 1989, Watkins proposed the Q-learning algorithm, which allows the agent to learn the optimal policy through interaction with the environment, without requiring prior knowledge of the environment model [Watkins 1989]. This algorithm became one of the most widely used methods in reinforcement learning due to its simplicity and effectiveness.

During the 1990s, reinforcement learning was consolidated as a research field. The publication of the book *Reinforcement Learning: An Introduction* by Sutton and Barto in 1998 became a milestone, systematizing the main concepts, algorithms, and theoretical foundations of RL, including temporal difference methods, Q-learning, and policy gradient approaches [Sutton and Barto 1998].

In the 2000s, advances in computational power and optimization techniques enabled the emergence of Deep Learning as a prominent research area. Works such as those by Hinton et al. on Deep Belief Networks (DBNs) demonstrated the feasibility of training deep neural networks efficiently [Hinton et al. 2006]. Bengio et al. further explored the role of unsupervised pre-training in improving the performance of deep architectures [Bengio et al. 2007].

The integration of reinforcement learning and deep learning gained significant attention in the 2010s. In 2013, researchers at DeepMind introduced the Deep Q-Network (DQN), which combined Q-learning with deep convolutional neural networks to enable agents to learn control policies directly from high-dimensional sensory input, such as raw pixels from Atari games [Mnih et al. 2013]. This work marked a turning point in the field, demonstrating that deep neural networks could successfully approximate value functions in complex environments.

Subsequent advances extended these ideas through techniques such as experience replay, target networks, policy gradient methods, and actor–critic architectures, improving stability and convergence in deep reinforcement learning systems. In 2016, AlphaGo, developed by DeepMind, combined deep neural networks with reinforcement learning and Monte Carlo Tree Search to achieve superhuman performance in the game of Go, defeating world champions and further demonstrating the potential of DRL [Silver et al. 2016].

More recent research has focused on improving the scalability, sample efficiency, and generalization capabilities of DRL algorithms. Studies such as [Li et al. 2021] and [Kallestad et al. 2023] propose novel DRL-based frameworks and hyper-heuristics to solve large-scale combinatorial optimization problems, including variants of the traveling salesman problem. These works highlight the growing applicability of DRL beyond games, extending to industrial, logistical, and operational domains.

Despite these advances, challenges remain regarding computational cost, explainability, and ethical considerations. As DRL systems become increasingly autonomous and influential in decision-making processes, there is a growing need to examine their social, ethical, and environmental impacts, particularly when applied at scale within information systems.

3. Speculative Design

3.1. Technique

Speculative design is an approach that seeks to explore possible futures in order to question and investigate the implications of new technologies, policies, and behaviors. Rather than focusing on immediate practical solutions, speculative design uses the construction of scenarios and prototypes to open discussions about desirable futures and to prepare society for change and innovation.

In the context of this work, speculative design is used to explore possible and desirable futures for Deep Reinforcement Learning. The methodology applied in this study follows a structured process, composed of the following stages:

- Mapping the current state, through a literature review aimed at understanding the present scenario and identifying existing challenges and opportunities;
- Identification of signals and future trends that may influence the development and application of DRL;
- Speculation of possible futures based on the interaction of identified signals and trends;
- Projection of desirable futures through the creation of narratives and conceptual artifacts that represent preferred scenarios.

This process enables a systematic exploration of future possibilities while maintaining a critical perspective on technological development.

3.2. Tools Used

To support the speculative design process, several analytical and foresight tools were employed, each contributing to different stages of the exploration.

3.2.1. Literature Review

The literature review was conducted as a foundational step to map the current state of Deep Reinforcement Learning. Academic articles, conference proceedings, and technical reports were analyzed to identify recent advances, persistent challenges, and emerging application areas. Particular attention was given to issues such as generalization, computational efficiency, scalability, and ethical concerns.

The literature review provided the empirical and conceptual basis for identifying signals and trends and for grounding speculative scenarios in existing research.

3.2.2. PESTEL Analysis

PESTEL Analysis is a strategic tool used to identify and analyze external factors that may influence a technology or system. The acronym PESTEL refers to Political, Economic, Social, Technological, Environmental, and Legal factors.

In this work, PESTEL Analysis was applied during the identification of signals and future trends to examine the broader contextual forces that may shape the evolution of Deep Reinforcement Learning. This analysis enabled a holistic understanding of how external dynamics, such as regulatory frameworks, economic incentives, social acceptance, environmental concerns, and technological infrastructure, may impact the adoption and development of DRL.

3.2.3. Futures Wheel

The Futures Wheel is a foresight tool originally developed by Jerome C. Glenn to explore the consequences of change. It enables the identification of first-order, second-order, and third-order consequences of a given trend or decision.

In this study, the Futures Wheel was used to map the potential consequences of advances in Deep Reinforcement Learning across different domains. By systematically identifying cascading effects, the tool helped uncover indirect impacts and unintended consequences that might arise from the widespread adoption of DRL-based systems.

3.2.4. Cone of Possibilities

The Cone of Possibilities, proposed by Joseph Voros, is a framework for categorizing future scenarios according to their likelihood and desirability. It distinguishes between possible, plausible, probable, and preferable futures.

This tool was employed to classify the speculative scenarios identified through the Futures Wheel. By positioning scenarios within the Cone of Possibilities, it became possible to differentiate between speculative extremes and more realistic or desirable outcomes, supporting informed discussion about future directions.

3.2.5. Innovation Map

The Innovation Map is an open innovation initiative that visualizes emerging technologies over a long-term horizon, typically up to the year 2035. Technologies are categorized according to their Technology Readiness Level (TRL) and aligned with the United Nations Sustainable Development Goals (SDGs).

In this work, the Innovation Map was used to identify technologies related to Deep Reinforcement Learning that could support sustainable and socially responsible futures. The tool helped ensure that speculative proposals were grounded in technologies with plausible development trajectories and positive societal impact.

3.2.6. Tarot Cards of Tech

The Tarot Cards of Tech are a speculative design tool designed to stimulate reflection on the social, ethical, and cultural implications of technology. The cards prompt discussion about issues such as power, trust, equity, scale, and unintended consequences.

In this study, the Tarot Cards of Tech were used to explore potential risks and ethical challenges associated with DRL-based solutions. They supported the identification of concerns related to accountability, inclusivity, environmental impact, and societal trust, informing the projection of desirable futures and mitigation strategies.

4. Mapping the Current State – Where Are We?

Recent advances in Deep Reinforcement Learning have focused on improving training efficiency, stability, and scalability. New algorithms and frameworks have been proposed to address limitations related to sample efficiency and convergence, enabling DRL systems to be applied to increasingly complex problems.

Works such as [Kallestad et al. 2023] and [Li et al. 2021] demonstrate the effectiveness of DRL in solving combinatorial optimization problems, including variants of the traveling salesman problem. These studies highlight the ability of DRL-based approaches to learn heuristics and decision policies that generalize across problem instances, outperforming traditional optimization methods in certain contexts.

The application of DRL has expanded beyond academic benchmarks and games, reaching industrial, logistical, and operational domains. Examples include scheduling, resource allocation, supply chain optimization, and autonomous systems. These applications benefit from DRL’s capacity to handle dynamic environments and sequential decision-making under uncertainty.

Despite these advances, significant challenges remain. Training DRL agents often requires substantial computational resources, leading to high energy consumption and environmental impact. Moreover, the complexity of DRL models raises concerns regarding explainability and transparency, making it difficult for users to understand and trust system decisions.

Another challenge involves the deployment of DRL systems in real-world environments. Issues such as safety, robustness, and adaptability to changing conditions must be addressed to ensure reliable performance outside controlled experimental settings.

4.1. Signals and Future Trends

The analysis of the current state reveals several signals and trends that are likely to influence the future of Deep Reinforcement Learning.

From a technological perspective, there is a trend toward the integration of DRL with other emerging technologies, such as the Internet of Things (IoT), edge computing, and cloud infrastructure. This integration enables real-time data collection and processing, expanding the range of applications for DRL-based systems.

Economically, increased investment in Artificial Intelligence research and development is accelerating innovation in DRL. Both public and private sectors are funding projects aimed at applying DRL to optimize complex systems and improve operational efficiency.

Socially, there is growing awareness of the ethical and environmental implications of AI technologies. Concerns about energy consumption, algorithmic bias, and accountability are driving demand for more sustainable and transparent AI solutions.

From a legal and regulatory perspective, emerging frameworks are beginning to address issues related to AI governance, data protection, and responsibility. These developments may shape how DRL systems are designed, deployed, and regulated in the future.

Environmental considerations are also becoming increasingly relevant. The high computational cost of training DRL models raises questions about sustainability, prompting research into more energy-efficient algorithms and hardware.

Together, these signals and trends provide the foundation for the speculative exploration of future scenarios presented in the following sections.

5. Speculation of Possible Futures – Where Are We Going?

Based on the signals and trends identified in the previous section, this stage of the speculative design process explores possible future scenarios for the development and application of Deep Reinforcement Learning. These scenarios are not intended as predictions, but as plausible futures that help illuminate opportunities, risks, and consequences associated with different technological trajectories.

5.1. Futures Wheel

Using the Futures Wheel tool, the potential impacts of advances in Deep Reinforcement Learning were mapped by identifying first-order, second-order, and third-order consequences. The analysis considered technological, social, economic, and environmental dimensions.

Among the first-order consequences identified are increased automation of complex decision-making processes, improved efficiency in resource allocation, and the expansion of autonomous systems across multiple sectors. These direct impacts may lead to second-order consequences such as changes in labor markets, increased reliance on algorithmic decision-making, and shifts in organizational structures.

Third-order consequences include broader societal transformations, such as reconfiguration of professional roles, new regulatory challenges, and long-term environmental effects related to energy consumption and infrastructure demands. The Futures Wheel enabled a structured examination of cascading effects that may arise from widespread adoption of DRL-based systems.

5.2. Cone of Possibilities

The scenarios identified through the Futures Wheel were classified using the Cone of Possibilities framework. This classification distinguished between possible, plausible, probable, and preferable futures, supporting critical reflection on the desirability and feasibility of different outcomes.

Possible futures include highly autonomous systems operating with minimal human oversight, while plausible futures involve collaborative human–AI systems designed to support decision-making. Probable futures reflect current trends toward increased integration of DRL in industrial and logistical applications, driven by economic incentives and technological feasibility.

Preferable futures emphasize sustainability, transparency, and ethical governance. These scenarios prioritize the development of DRL systems that align with social values, minimize environmental impact, and maintain human agency in decision-making processes.

5.3. Mapping the Future Scenario

Based on the classification within the Cone of Possibilities, several future scenarios were identified as particularly desirable. These include scenarios that promote sustainable computing practices, responsible innovation, and inclusive technological development.

Among these, sustainable computing practices were selected as the primary focus for further exploration. This choice reflects the high technological potential of DRL combined with the urgent need to address environmental and social challenges associated with large-scale computation.

The selected future scenario envisions DRL systems designed with energy efficiency, scalability, and social responsibility as core principles. This scenario serves as the basis for the projection of a desirable future and the design of an information technology solution presented in the next section.

6. Projection of a Desirable Future – Where Do We Want to Go?

Based on the speculative analysis of possible futures, this section projects a desirable future scenario for the application of Deep Reinforcement Learning, grounded in sustainability, ethical governance, and social responsibility.

6.1. Innovation Map

To support the projection of a desirable future, the Innovation Map was used to identify emerging technologies with potential alignment to sustainable and socially responsible development.

Through this tool, Artificial Superintelligence (ASI) was selected as the technological basis for the proposed future scenario, considering its high transformative potential and long-term speculative horizon.

The Innovation Map enabled the evaluation of ASI in relation to its Technology Readiness Level (TRL) and its alignment with the United Nations Sustainable Development Goals (SDGs), particularly those related to sustainable cities, innovation, and climate action.

6.1.1. Artificial Superintelligence

Artificial Superintelligence refers to a hypothetical form of Artificial Intelligence that surpasses human intelligence across a wide range of cognitive tasks. Unlike narrow AI systems designed for specific functions, ASI would possess general problem-solving capabilities, adaptive learning, and autonomous reasoning at a level beyond human performance.

While ASI presents significant opportunities for addressing complex global challenges, such as climate change and large-scale infrastructure optimization, it also raises profound ethical, social, and governance concerns. Issues related to control, alignment with human values, accountability, and societal impact are central to discussions about ASI.

In this speculative work, ASI is used as a conceptual element to explore the long-term implications of advanced DRL-based systems and to stimulate reflection on responsible technological trajectories.

6.2. Designing an IT Solution

Based on the desirable future scenario, an information technology solution was conceptualized to illustrate the application of DRL and ASI principles in a socially beneficial context. The proposed solution is an autonomous public transportation system designed to promote sustainability, efficiency, and accessibility.

The system leverages advanced DRL algorithms to optimize routing, scheduling, and resource allocation in real time. Autonomous vehicles are equipped with advanced sensors and communication capabilities, enabling coordinated operation and adaptive responses to changing conditions.

Sustainability is a core design principle. The system incorporates renewable energy sources, such as solar power, and employs energy-efficient computational strategies to reduce environmental impact. By optimizing transportation flows, the system aims to reduce traffic congestion, emissions, and operational costs.

Accessibility and inclusivity are also emphasized. The transportation system is designed to provide low-cost mobility options, ensuring access for diverse populations and supporting social equity.

6.3. Other Actions

To further examine the implications of the proposed solution, the Tarot Cards of Tech were used to identify potential risks and mitigation strategies. Key concerns include trust in autonomous systems, safety and reliability, job displacement, environmental impact, and inclusivity.

Mitigation strategies involve transparent communication about system behavior, robust safety mechanisms, workforce requalification programs, and continuous monitoring of environmental and social impacts. These actions aim to ensure that technological innovation contributes positively to society.

7. Conclusion

This work explored the use of Deep Reinforcement Learning within the field of Information Systems through the lens of speculative design. By mapping the current state, identifying signals and trends, and speculating on possible and desirable futures, it was possible to reflect on the potential trajectories of DRL and their implications for society.

Speculative design proved to be a valuable methodological approach for anticipating technological, social, and ethical challenges associated with advanced AI systems. Rather than predicting the future, the approach enabled structured reflection on the consequences of different technological choices and highlighted the importance of intentional design aligned with societal values.

The proposed desirable future emphasized sustainability, transparency, and social responsibility, illustrating how DRL and speculative projections of Artificial Superintelligence could be oriented toward positive societal impact. The conceptual IT solution demonstrated how advanced decision-making systems might contribute to sustainable urban mobility while addressing ethical and governance concerns.

Despite its contributions, this work is inherently speculative and subject to uncertainty. The realization of the proposed scenarios depends on factors such as technological progress, regulatory frameworks, political will, and societal acceptance. Future research could expand on these findings by exploring alternative scenarios, conducting empirical studies, and engaging stakeholders in participatory design processes.

In conclusion, the integration of Deep Reinforcement Learning and speculative design offers a powerful framework for exploring future possibilities and informing responsible innovation. By critically examining emerging technologies and their broader implications, it becomes possible to guide technological development toward futures that are not only technologically advanced but also socially desirable.

References

(The reference list is reproduced exactly as in the source document, without translation or normalization.)

[Analysis] Analysis, P. What is pestle analysis? https://pestleanalysis.com/

what-is-pestle-analysis/. Acesso em: 16 ago. 2024.

[Ashby 1956] Ashby, W. R. (1956). An Introduction to Cybernetics. Chapman & Hall,

London.

[Bellman 1957] Bellman, R. E. (1957). Dynamic Programming. Princeton University Press.

[Bengio et al. 2007] Bengio, Y., Lamblin, P., Popovici, D., and Larochelle, H. (2007). Greedy

layer-wise training of deep networks. In Advances in Neural Information Processing

Systems, volume 19, pages 153–160.

[Betz 2021] Betz, C. T. (2021). Managing Digital: Concepts and Practices. Elsevier.

[Bytheway 2014] Bytheway, A. (2014). Investing in Information: The Information Management

Body of Knowledge (IMBOK). Springer.

[Delve] Delve. Speculative design and a cone of possibilities.

https://www.delve.com/insights/

speculative-design-and-a-cone-of-possibilities. Acesso

em: 16 ago. 2024.

[Group] Group, A. The tarot cards of tech. https://www.artefactgroup.com/

resources/the-tarot-cards-of-tech/. Acesso em: 16 ago. 2024.

[Hinton et al. 2006] Hinton, G. E., Osindero, S., and Teh, Y.-W. (2006). A fast learning

algorithm for deep belief nets. Neural Computation, 18(7):1527–1554.

[Kallestad et al. 2023] Kallestad, J., Hasibi, R., Hemmati, A., and S¨orensen, K. (2023). A

general deep reinforcement learning hyperheuristic framework for solving combinatorial optimization problems. European Journal of Operational Research, 309(2):446-

468.

[Krizhevsky et al. 2012] Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet

classification with deep convolutional neural networks. Advances in Neural Information

Processing Systems, pages 1097–1105.

[Laudon and Laudon 2021] Laudon, K. C. and Laudon, J. P. (2021). Management Information

Systems: Managing the Digital Firm. Pearson, 16th edition.

[Li et al. 2021] Li, K., Zhang, T., Wang, R., Wang, Y., and Han, Y. (2021). Deep reinforcement

learning for combinatorial optimization: Covering salesman problems. IEEE

Transactions on Cybernetics, 14(8):1–15.

[Map] Map, I. Innovation map. https://innovationmap.at/. Acesso em: 16 ago.

2024.

[Maturana and Varela 1987] Maturana, H. R. and Varela, F. J. (1987). The Tree of Knowledge:

The Biological Roots of Human Understanding. Shambhala, Boston, MA.

[Mnih et al. 2013] Mnih, V. et al. (2013). Playing atari with deep reinforcement learning.

arXiv preprint arXiv:1312.5602.

[Mnih et al. 2015] Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare,

M. G., and Hassabis, D. (2015). Human-level control through deep reinforcement

learning. Nature, 518(7540):529–533.

[Ramage and Shipp 2009] Ramage, M. and Shipp, K. (2009). Systems Thinkers. Springer.

[Scouting] Scouting, F. Futures wheel. https://futurescouting.com.au/

futures-wheel/. Acesso em: 16 ago. 2024.

[Silver et al. 2016] Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den

Driessche, G., and Hassabis, D. (2016). Mastering the game of go with deep neural

networks and tree search. Nature, 529(7587):484–489.

[Sutton 1988] Sutton, R. S. (1988). Learning to predict by the methods of temporal differences.

Machine Learning, 3(1):9–44.

[Sutton and Barto 1998] Sutton, R. S. and Barto, A. G. (1998). Reinforcement Learning:

An Introduction. MIT Press.

[Tesauro 1995] Tesauro, G. (1995). Temporal difference learning and td-gammon. Communications

of the ACM, 38(3):58–68.

[Tools] Tools, M. Pest analysis: Identifying the external factors that influence your

organization. https://www.mindtools.com/aqa3q37/pest-analysis.

Acesso em: 16 ago. 2024.

[Wang et al. 2020] Wang, H., Liu, N., Zhang, Y., et al. (2020). Deep reinforcement learning:

a survey. Frontiers of Information Technology & Electronic Engineering, 21:1726–

1744.

[Watkins 1989] Watkins, C. J. (1989). Learning from Delayed Rewards. PhD thesis, University of Cambridge.

[Watkins and Dayan 1992] Watkins, C. J. and Dayan, P. (1992). Q-learning. Machine Learning, 8(3-4):279–292.