AI4media

ARTIFICIAL INTELLIGENCE FOR THE MEDIA AND SOCIETY

Roadmap on AI technologies & applications for the Media Industry

Section: "Reinforcement Learning"

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 951911

info@ai4media.eu www.ai4media.eu

Authors	David Melhart (modl.ai) Ahmed Khalifa (modl.ai) Daniele Gravina (modl.ai)

This report is part of the deliverable D2.3 - "*AI technologies and applications in media: State of Play, Foresight, and Research Directions*" of the AI4Media project.

You can site this report as follows:

F. Tsalakanidou et al., Deliverable 2.3 - AI technologies and applications in media: State of play, foresight, and research directions, AI4Media Project (Grant Agreement No 951911), 4 March 2022

This report was supported by European Union´s Horizon 2020 research and innovation programme under grant number 951911 - AI4Media (A European Excellence Centre for Media, Society and Democracy).

Reinforcement Learning

Current status

Reinforcement Learning (RL) is an area of machine learning that is concerned with learning how to solve a problem through trial and error. For example, an agent can learn to drive a car by allowing it to control the car and providing it with positive or negative rewards based on how well it is performing (Figure 1). Until the recent advancement of machine learning, reinforcement learning applications were limited to toy problems because it was very hard to learn to perform well from just interacting with the environment. Nowadays, reinforcement learning is being used in a lot of different fields and showcases super-human performance across complex problems that humans are traditionally good at solving: those include playing popular board games like Go[1] and video games such as arcade-type Atari[2] and StarCraft II[3].

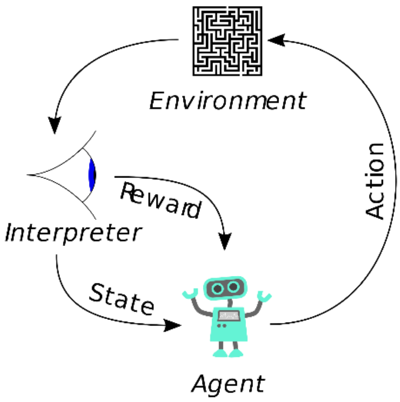

Figure 1: The components of the reinforcement learning framework.[4]

With the amazing performance of reinforcement learning algorithms over different games in the academic world, one might wonder if reinforcement learning is being used in other multimedia domains such as social media, online content production, movies or news outlets. Unfortunately, at the time of writing, RL is still not widespread in multimedia domains outside of video games. This is because although different types of multimedia provide a huge corpora of data where supervised and/or unsupervised learning can be used efficiently, traditionally RL is focused on controlling an agent in an interactive environment. However, looking at the recent advancements in reinforcement learning, one could see the advantage of using it in these fields, especially ***active learning***. Active learning algorithms involve agents that are continuously learning and adapting to an ever-changing situation. As such, active learning algorithms are well suited to address tasks involving human-computer interactions and domains where the focus is the preference or interaction style of humans using a system. For example, an RL bot could be

[1] Silver, David, et al. "Mastering the game of Go with deep neural networks and tree search." Nature (2016).
[2] Mnih, Volodymyr, et al. "Playing atari with deep reinforcement learning." NIPs (2013).
[3] Vinyals, Oriol, et al. "AlphaStar: Mastering the Real-Time Strategy Game StarCraft II." DeepMind Blog (2019).
[4] Image source: Wikipedia - https://en.wikipedia.org/wiki/Reinforcement_learning#/media/File:Reinforcement_learning_diagram.svg

used to assist directors during filming TV shows by learning their style of camera shooting over time and later propose camera positions that will fit the director's style and target aesthetics.

Research challenges

As mentioned before, RL is used more in games compared to other types of media. The lack of interactivity and the abundance of data usually prioritises different methods of machine learning over RL in other domains. Although RL is not a new field, its success is recent even within games. Consequently, there are only a handful of examples of successful industrial applications of RL. An indicative example is the *Black & White* (EA, 2002) strategy game[5] and "god simulator". The game uses RL to control a creature's behaviour during the game that acts as the player's companion throughout the levels. The player can affect their creature's behaviour by observing their actions and giving positive or negative reinforcement quickly after. During the game, the creature can be trained to perform different tasks, aiding the player. Even though *Black & White* enjoyed universal acclaim at the time, not a lot of companies followed suit in implementing RL algorithms in their games. The prototypical nature of the algorithm in game development did not help RL to advance further and become mainstream in games and their design.

Another challenge faced by RL is that trained agents may reach a *super-human performance*[6,7,8]. While desirable in other domains, superhuman players are not an interesting feature in games because humans either won't be able to compete with these AI agents (if AI is an adversary) or they will be outshined by the agent which might make the game boring (if AI has a companion role). Besides that, designing these AI systems for any media sector is a complex task as there are no standard libraries that can be easily integrated and used out of the box. Finally, training these AI agents is often costly both in terms of money and time[9]. That cost, in turn, does not allow for fast iterations and quick testing of new ways of using RL in different media sectors.

Although reinforcement learning reached super-human performance in many fields, it still faces several challenges with *generalisation*. Generalisation is the problem of using the trained agent on a new problem other than the one it got trained on. For example, if we trained an agent to play *Super Mario Bros* (Nintendo, 1985)[10] optimally, we would like it to be able to do so in *Super Mario Bros 3* (Nintendo, 1988) too. Overcoming the generalisation challenge will help game and media productions reduce their overall cost as AI engineers won't need to retrain AI models across different games, similar to how computer vision models can operate directly without pre-training across dissimilar pattern recognition tasks[11]. We argue that transferable and generalisable world models and agents will be available in games and other media fields as soon as such AI models become easier to integrate, more accessible, and cheaper to train.

---

[5] Black and White Video games, EA (2022): https://www.ea.com/games/black-and-white
[6] Silver, David, et al. "Mastering the game of Go with deep neural networks and tree search." Nature (2016).
[7] Mnih, Volodymyr, et al. "Playing atari with deep reinforcement learning." NIPs (2013).
[8] Vinyals, Oriol, et al. "AlphaStar: Mastering the Real-Time Strategy Game StarCraft II." DeepMind Blog (2019).
[9] Wang, Ken. "DeepMind achieved StarCraft II GrandMaster Level, but at what cost?." Medium (2020)
[10] Super Mario Bros, Nintendo (1985): https://mario.nintendo.com/history/
[11] Shah, Hardik. "How Artificial Intelligence is Enhancing Mobile App Technology." RTInsights (2020).

Vignette: AI bots for game development & quality assurance

Gita and Rami are a programmer and a quality assurance (QA) manager working at a medium-size game company. Their company is working on a multiplayer action-adventure game "*Power of Legends*", which will be supported by continuous small content updates throughout the whole life-cycle of the game. Dependence on regular content updates means development never ends. Content must be added at speed while maintaining quality and creativity to keep players engaged and attract new ones, which means that the average cost of game development has increased ten folds each decade so far. To mitigate these costs, Gita and Rami's company uses AI-assisted testing tools that leverage reinforcement learning to quickly train agents that can perform the more menial tasks of game testing. This tool helps both the programmers and quality assurance staff to manage the rapidly increasing testing demand.

In the new content upgrade to "*Power of Legends*,'' the developers are introducing new maps and items to the game. The team has to make sure that the new map is bug-free and playable and that the newly added items do not break already existing levels. The work starts with Gita, who is responsible for the integration between the game and the testing software. She sets up the AI-powered bot in the game environment by defining its input space and the basic glitch or bug criteria. The goal of the setup is to prepare the work of QA staff, who will have a higher level control over the bot's interface. Once the bot is integrated into the game, the QA staff does not need to pull resources from the programmers; they can modify the bot's behaviour on their own to run a plethora of different tests. While not completely replacing traditional quality assurance, using AI-based game testing can deliver a faster, cheaper, and more consistent turn-around.

Once a build of the game is prepared for testing, the game is uploaded into the testing platform, where a QA session can be initiated. At this point, Rami takes over and starts a test through a graphical user interface. The AI bots adapt to new content with ease (Figure 2) and quickly find bugs and glitches with different levels of severity. Some of these are game-breaking issues that have to be prioritised and addressed immediately by the programmers, while others are minor glitches or events that break the game's balance. Once the bot finishes, Rami receives a detailed test report which includes a smart summary of the encountered issues and a structured telemetry log of these issues. The created bug reports are sorted based on severity and sent back to Gita, who can immediately start working on an update, giving Rami time to go through the telemetry logs and - if needed - replay or reproduce more complex issues. This faster iteration cycle between Gita and Rami aided by the AI bots frees up resources in both the development and testing process. As the AI bots work faster and more consistently than human testers, there is less idle time during development. Gita can deliver faster updates to the game and Rami can focus on parts of the QA process that require more insight.

Recently, large-scale game developer studios started looking into using reinforcement learning in the game development process instead of using it to control characters in the game. The superhuman performance[9,11,12] and fast exploration power allowed these companies to utilise

[12] Koster, Ralph. "The cost of games" Venture Beat (2018).

these agents to test their games. For example, EA uses reinforcement learning to detect navigation problems such as unreachable areas or missing collision boxes[13,14]. Companies like modl.ai are working on creating plug-and-play AI tools, which can integrate into different games, which could help bring these innovations to small and mid-sized companies as well, making Gita and Rami's story a reality.

Future trends for the media sector

The use of RL will increase in the game industry and other media sectors at large over the next decade. In the game industry, the use of reinforcement learning in the area of ***automatic game testing*** is going to grow more over the next decade and receive much research and innovation attention (Figure 2). This increase of interest in RL-based game testing is due to games gradually becoming online services (for instance *Fortnite*, *Dota 2*, *Candy Crush*, etc.) that require continuous updates and new content to keep players engaged. To keep that supply running, any new updates and content need to be tested periodically, resulting in a substantial increase in the production cost of a game. This cost increase is far from being sustainable. Thus, if game companies wish to scale their productions and expand their online services, they will need to rely on reliable and explainable AI methods that are deeply integrated into the testing process.

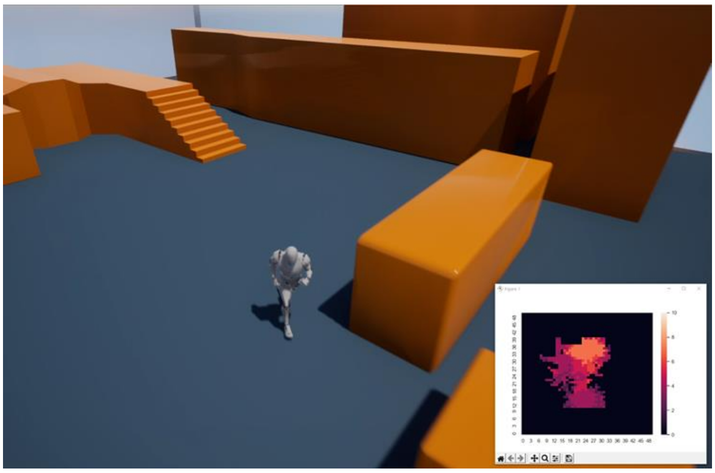

Figure 2: Modl.ai AI bot exploring a previously unseen game level.

Besides games, one might think about the various ways RL can be used in other domains throughout other media sectors. For instance, it could be used to create ***new types of media content*** such as (parts of) games, levels, environments, visuals, cloth simulations for fashion design, camera positions during filming movies and TV shows, etc. These avenues are not explored yet in any sector of the media industry but we expect that recent advancements in machine learning and RL will push innovation in these fields as well. But to see this happening, content creators should always keep their eyes and minds open to embrace these super-human performance agents and find new ways to incorporate RL in different media sectors. For example, the EA research team has been using RL recently to generate new playable levels for

[13] Thompson, Tommy. “Training the Shadow AI of Killer Instinct.” AI and Games (2017).
[14] Yannakakis, Georgios N., Antonios Liapis, and Constantine Alexopoulos. "Mixed-initiative co-creativity." Foundations of Digital Games (2014).

different games with different levels of difficulty[15] . Such generative AI methods could be used in the other media sectors as well for creating/curating content that will most likely be enjoyed by their users.

In addition to RL, ***imitation learning (IL)*** has been gaining popularity in recent years. Imitation learning is an area of machine learning concerned with the problem of learning how to imitate humans through their behavioural data (also called human demonstrations). The reason for their popularity is that IL algorithms may yield agents that perform and behave in a human-like fashion. This feature is important in competitive multiplayer games such as *Fortnite* (Epic, 2017)[16] and *Dota 2* (Valve, 2013)[17] as in these games players usually need believable and human-standard AI to play with or against. Most scripted (not machine-learned) AI agents do not perform as well as humans and they tend to follow specific strategies which can be exploited easily by players. Through imitation learning, we make multiplayer games engaging and plausible even when human players are not always there. By using the collected player data we can train AI agents that capture a player's style and performance which we can, in turn, use to create NPCs in the game for humans to play against (or with). *Forza* (Microsoft, 2018)[18] and *Killer Instinct* (Rare, 2013)[19], for instance, use the player's data to create a shadow player (driver and fighter, respectively) that other players can play against and test their skills[13,20]. Most notably the *Drivatar AI* system of the Forza Series (2005-present) is the longest living imitation learning algorithm existent in video games with more than 15 years of continuous development and operation.

Goals for next 10 or 20 years

We expect more media and game companies to start adopting RL in all stages of media production from early design to shipping. In the next 10 years, RL algorithms will not be used as a fully autonomous system but it will be working in tandem with designers, developers, content creators, and testers in a mixed-initiative fashion[21]. These tools will provide designers with new ways of exploring new rules and mechanics, help them design levels, test the designed content, and provide statistics on the user experience. They will elevate the tester's job to focus more on high-level challenges instead of tedious repetitive work. They will also help developers to optimise their code and find better solutions toward simulating certain physics or world history. They will help content creators in the content creation process by suggesting improvements and edits and sometimes whole topics that consumers are looking for. Because of all the aforementioned reasons, content creation cycles will start to be shorter which will not only cut the cost of production but also it will allow the industry to be more experimental and take more risks.

[15] Gisslén, Linus, et al. "Adversarial reinforcement learning for procedural content generation." IEEE Conference on Games (CoG) (2021).
[16] Fortnite, Epic: https://www.epicgames.com/fortnite/en-US/home
[17] Dota 2, Valve: https://www.dota2.com/home
[18] Forza, Microsoft: https://www.microsoft.com/en-us/store/collections/forzacollection
[19] Killer Instinct, Rare: http://www.raregamer.co.uk/games/killer-instinct/
[20] T. Thompson, "How Forza's Drivatar Actually Works." AI and Games (2021).
[21] Yannakakis, Georgios N., Antonios Liapis, and Constantine Alexopoulos. "Mixed-initiative co-creativity." Foundations of Digital Games (2014).

To achieve the long-term goal of wide AI (RL) adoption in media and game productions certain decisive steps need to be made towards rapid deployment and wide accessibility. Most notably, the Unity game engine has been addressing the accessibility issue since 2018: it released an easy to integrate library called ML-Agents[22]. This library allows Unity developers to train agents using reinforcement learning in their own games. Although this is an important first step towards making reinforcement learning/machine learning more accessible, the algorithms are still computationally heavy and hard to debug. Besides, ML-Agents are currently operational only within Unity which is far from being ideal for game developers that use other engines.

We think that in the next 20 years, easy-to-use RL libraries will be released (similar to ML-Agents). These libraries will allow non-technical users to easily interact with them and train them on more specific problems by a press of a button. Besides that, the advancements in hardware and software will enable RL algorithms to run efficiently on small devices like smartphones, which will allow users to have the power of these models at the tip of their fingers.

[22] A. Juliani, et al. "Unity: A general platform for intelligent agents." arXiv preprint arXiv:1809.02627 (2018).

AI4media

ARTIFICIAL INTELLIGENCE FOR THE MEDIA AND SOCIETY

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 951911

info@ai4media.eu www.ai4media.eu