Exploring the synergies between transfer in reinforcement learning and procedural content generation

Müller-Brockhausen, M.F.T.

Chapter 1

Introduction

1.1 Background

Video games exemplify "perfect testbeds for artificial intelligence" [272], as they provide complex, interactive, and dynamic scenarios that challenge computational models. They consist of many different components (ranging from mathematics, and graphics engines, to approximations of the laws of physics [105]). The primary objective of these systems is to create an engaging and immersive experience for users. Achieving this requires the seamless integration of a multitude of technologies [30].

Designing Artificial Intelligence (AI) in video games presents a spectrum of challenges. When creating believable and immersive worlds the focus lies on replicating human-like behavior [121, 303] through Non-Player Characters (NPCs). Furthermore, in some highly competitive games, the goal may be either to achieve superhuman decision-making capabilities [43, 285], or to provide players with an interesting opponent [296].

Games of a competitive nature have seen significant advances in AI achieved through algorithms such as MiniMax [202, 205], Monte Carlo Tree Search (MCTS) [44, 56], Evolutionary Algorithms (EAs) [94, 152, 208], Reinforcement Learning (RL) [139, 298], and hybrids of RL & MCTS [168, 232, 279]. Particularly RL is well suited for learning interactive behavior and has made great strides in the field of robotics [249], where robots learn human-like movement [23]. Moreover, in video games, the promise of reinforcement learning has been well recognized [239].

The development of benchmarks is critical for advancing and validating any AI system. One notable example in the field of video games is the General Video Game Playing AI (GVGAI) Framework [196]. Recent advancements in competitive AI have shifted focus toward multi-agent systems, emphasizing coordinated decision-making through interactions in complex environments. Games such as Dota 2 have become prominent platforms for testing and refining these algorithmic frameworks [28]. Additionally, the emergence of Large Language Models (LLMs) has introduced novel capabilities for strategic planning, giving rise to novel benchmarks [184].

Human behavior is intrinsically complex, making its replication in video games a non-trivial challenge. Historically, game-level design relied on a static approach, where the content was manually crafted. Procedural Content Generation (PCG) encompasses any content for video games that is generated algorithmically. Examples include trees [118, 186], meshes [284], levels [129, 134, 209, 223, 261, 269, 282], textures [57], and dialogue [111, 165, 257]. This list is not exhaustive, for a more complete overview we refer the reader to [115]. To generate the required level of diversity in content for AI, developers have adopted algorithmic methods to procedurally generate content using a pseudorandom number generator (PRNG) [194]. In consequence, PCG has enabled video games to exhibit greater variability and unpredictability in their content [115], which in turn elevates the cognitive and strategic demands on players, allowing controllable levels of difficulty for gameplay [226, 251].

This trend towards more diversity can also be seen in machine learning [130]. Training new deep learning models for different data sets and environments is expensive [217] and wasteful [193]. It is much more efficient to transfer knowledge from a model that has been learned for one task to a model for a different task [265]. For example, in image recognition, the ImageNet model is trained on millions of images [67]. Subsequently, it is employed to fine-tune smaller models for specific recognition tasks [150]. Similarly, in natural language processing (NLP), LLMs are evolving into foundation models [18] capable of being adapted through fine-tuning for diverse applications [132]. The perfect example of this is the usage of a vision language action model for navigational tasks in robotics [237].

We wonder how the concept of transfer learning can contribute to video games. In turn, our thesis delves into the possible synergies between Transfer in Reinforcement Learning (TRL) and PCG by categorizing existing research and performing empirical studies.

Now that we have introduced key aspects of the relevant fields, we transition to defining the central research questions that will guide this thesis.

1.2 Research Questions

Throughout the research efforts of our thesis, we pose a variety of questions that we aim to answer by means of empirical analysis and experimentation in our publications. In the subsections, we first motivate the origin of our question based on related research.

Then we pose questions for which we list the chapters related to allowing us to answer the posed question. In Chapter 7 we present the answers to these individual questions.

Our main topic in this thesis is:

> Exploring synergies between TRL and PCG.

This general topic captures the research that is covered in this thesis at a high level. Combining TRL and PCG suggests the need for better benchmarks. To investigate this topic in greater depth, we have formulated the following research questions.

1.2.1 RQ1

The inherent difficulty in establishing robust benchmarks stems from the phenomenon known as Goodhart's law: "When a measure becomes a target, it ceases to be a good measure" [258]. This principle underscores the necessity for continuous benchmark evolution, which is why we ask:

> How can benchmarks for (transfer in) RL be improved?

This research question is studied by first benchmarking a variety of popular AI methods (Heuristic, MCTS, RL) on the game Tetris Link (Chapter 2). Both MCTS and RL exhibited disappointing results in comparison to a heuristic search. Additionally, our investigations into RL highlighted reproducibility challenges, a well-known issue in the field [201], prompting us to explore potential solutions (Chapter 4). Furthermore, we identified transfer in RL as a possible way to improve performance, guiding us to survey the field (Chapter 3) and conclude that PCG is a promising method to further improve benchmarks for (transfer in) RL. We answer this question in Section 7.1.1.

1.2.2 RQ2

The previously mentioned hybrids of RL & MCTS highlight that the combination of two methods can yield promising results, achieving superhuman performance in games like Go [232] or Starcraft II [279]. Our main topic is to explore such synergies between TRL and PCG, thus we ask more concretely:

> Can TRL improve PCG applications in video games?

This research question is studied by identifying limitations in TRL experiments of existing research (Chapter 3) and empirically showing that combining TRL and PCG is promising (Chapter 5). We answer this question in Section 7.1.2.

1.2.3 RQ3

The use of LLMs in video games holds significant promise as it can be used for AI [128] or for PCG to generate levels [261, 269], and interactive dialogues [257]. Moreover, the existence of mods for existing games that incorporate LLMs [15, 211] highlights a clear demand by the user base. However, as a developer, controlling the output is an open problem. That is why we wonder:

> Can LLMs procedurally generate content locally for video games while preventing adversarial attacks?

This research question studies the computational requirements and safety of deploying LLMs that procedurally generate content in video games. We run empirical experiments (Chapter 6) to help answer this question (Section 7.1.3).

1.3 Contributions

This doctoral thesis comprises independent research publications, each constituting a distinct chapter. We provide a concise overview of the key findings and arguments presented in each chapter. We will now list our main contributions:

Chapter 2 studies the game of Tetris Link using three different popular AI algorithms: Heuristic Search, MCTS, and RL. To our surprise, the classical method, heuristic search, performs best. We show that this is due to a combination of a large branching factor and restrictive game rules. This chapter is based on:

[175] Müller-Brockhausen, M., Preuss, M., Plaat, A.: A new challenge: Approaching tetris link with AI. In: 2021 IEEE Conference on Games (CoG), Copenhagen, Denmark, August 17-20, 2021. IEEE (2021), `https://doi.org/10.1109/CoG52621.2021.9619044`.

This surprising result has caused us to investigate transfer methods in RL in Chapter 3. Here we find that transfer learning is studied widely in reinforcement learning and is successful when levels are simple and pre-programmed. However, we find that few works use transfer in combination with PCG, identifying a field for further study. This chapter is based on:

[176] Müller-Brockhausen, M., Preuss, M., Plaat, A.: Procedural content generation: Better benchmarks for transfer reinforcement learning. In: 2021 IEEE Conference on Games (CoG), Copenhagen, Denmark, August 17-20, 2021. IEEE (2021), `https://doi.org/10.1109/CoG52621.2021.9619000`.

Reproducibility is a desirable trait, especially for TRL experiments. However, it remains an unsolved and often ignored problem in the field of RL [201]. As we cannot fix reproducbility, we propose focusing on verifiability (Chapter 4). This enables reviewers to validate empirically reported results, such as tables or graphs. This chapter is based on:

[173] Müller-Brockhausen, M., Plaat, A., Preuss, M.: Towards verifiable benchmarks for reinforcement learning. In: IEEE Conference on Games, CoG 2022, Beijing, China, August 21-24, 2022. pp. 159–166. IEEE (2022), `https://doi.org/10.1109/CoG51982.2022.9893715`.

Next in Chapter 5 we empirically show the synergies of TRL and PCG at the example of Linerider 3D (see 7.1.2). Employing PCG broadens the explored state space during training, yielding scalable transfer unachievable when learning from scratch (see 5.6.2). Complementarily, the task of the RL environment is to procedurally generate a track for the video game Linerider (see 5.4). This chapter is based on:

[172] Müller-Brockhausen, M., Khalifa, A., Preuss, M.: Scalable procedural content generation via transfer reinforcement learning. In: Data Science and Artificial Intelligence (DSAI). Springer (2024), `https://doi.org/10.1007/978-981-97-9793-6`.

In our final chapter 6, we study the viability of local LLMs to generate content for video games, while preventing derailment through the user. To that end we suggest the usage of Chatter, which disables user input, relying solely on pre-defined prompts. This prevents users from breaking the model out of its safety guidelines, for which researchers keep finding new creative ways [46, 53]. This chapter is based on:

[171] Müller-Brockhausen, M., Barbero, G., Preuss, M.: Chatter generation through language models. In: IEEE Conference on Games, CoG 2023, Boston, MA, USA, August 21-24, 2023. IEEE (2023), `https://doi.org/10.1109/CoG57401.2023.10333244`.

1.3.1 Published Collaborations of the Author

In addition to these published works, we have collaborated on a variety of topics which has led to the following publications:

[137] Kampert, D., Varbanescu, A.L., Müller-Brockhausen, M., Plaat, A.: Mimicking the human approach in the game of hive. In: IEEE Symposium Series on Computational Intelligence, SSCI 2021, Orlando, FL, USA, December 5-7, 2021. IEEE (2021), `https://doi.org/10.1109/SSCI50451.2021.9659999`.

[174] Müller-Brockhausen, M., Plisnier, H.: Transferring while playing the rl agent. In: Demo at BNAIC/BeNeLearn 2022: Joint International Scientific Conferences on AI and Machine Learning (2022), https://bnaic2022.uantwerpen.be/wp-content/uploads/BNAICBeNeLearn_2022_submission_6564.pdf.

[255] van der Staaij, A., Prins, J., Prins, V.L., Poelsma, J., Smit, T., Müller-Brockhausen, M., Preuss, M.: Believable minecraft settlements by means of decentralised iterative planning. In: IEEE Conference on Games, CoG 2023, Boston, MA, USA, August 21-24, 2023. IEEE (2023), https://doi.org/10.1109/CoG57401.2023.10333146.

[21] Barbero, G., Müller-Brockhausen, M., Preuss, M.: Challenges of open world games for ai: Insights from human gameplay. In: Data Science and Artificial Intelligence (DSAI). Springer Nature (2024), https://doi.org/10.1007/978-981-97-9793-6.