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

Müller-Brockhausen, M.F.T.

Chapter 7

Conclusions

This chapter will summarize the main findings of our work, and surmise the answers to our posed research questions (Section 1.2). Moreover, we highlight the current limitations of our research and present the next steps that could foster derivative works in the future (Section 7.2). We will start by providing answers to the research questions.

7.1 Answers to research questions

In this section, we provide answers to the research questions.

7.1.1 Improving Benchmarks

Our first research question sought to identify opportunities for improving the state of benchmarking for (transfer in) Reinforcement Learning (RL).

By benchmarking a variety of popular AI methods (Heuristic, Monte Carlo Tree Search (MCTS), RL) on the game Tetris Link (Chapter 2) we notice that RL experiments are hardly reproducible. Furthermore, we observed that training RL agents from scratch to perform effectively on complex tasks proves to be highly unreliable. Both MCTS and RL demonstrated underwhelming performance relative to a simple heuristic search baseline

To address reproducibility issues, we suggest the usage of *replay traces*. While this approach does not inherently guarantee reproducible training outcomes, it provides a mechanism for verifying experimental results by enabling the replay of every training and evaluation episode of an RL agent. Sadly, the manipulation of figures in scientific publications is a practice that to this day still prevails [170]. Replay traces enable reviewers to prove that displayed graphs are truly based on the collected data and have not been tampered with. Lastly, our survey of the Transfer in Reinforcement Learning (TRL) field (Section 3.5), highlights the underutilized potential of Procedural Content

Generation (PCG). It enables the creation of more varied task differences, which not only enhance generalization capabilities but also contribute to the development of improved future benchmarks for TRL research.

7.1.2 TRL & PCG

Our second research question (1.2.2) asked whether TRL can enhance PCG in video games.

Our findings demonstrate that TRL and PCG exhibit a bidirectional synergy. In addressing our first research question (7.1.1) we established that TRL benchmarks can profit by employing PCG methods for generating a broader spectrum of tasks. Furthermore, in our empirical experiments on 3D Linerider (Chapter 5), we highlight that his relationship is bidirectional. The first direction, as previously suggested, is to use PCG to vary the start and goal positions of each episode to train an RL policy with better generalization capability. The other direction is that TRL enables to tackle PCG tasks on a scale that would not be achievable if RL were trained from scratch. Specifically to our experiments that means instead of being limited to a world-size of 10 in 3D Linerider, we can achieve asymptotic performance, lower time to threshold, a higher total reward, and a jumpstart in training performance by incrementally transferring from smaller to larger world sizes of 30. These results satisfy all four performance criteria for TRL experiments outlined by [265].

7.1.3 LLMs & PCG

Our last research question examines the integration of Large Language Models (LLMs) into video games for PCG running locally on the machine of the user while being protected against adversarial attacks. Recent advancements in parameter-efficient LLMs such as LLAMA [274] consisting of 7B or 13B parameters (compared to 175B for, e.g., GPT3 [91]) enable running LLMs locally in parallel to a video game. To underline this feasibility, we show that 70% of consumer-grade gaming hardware would be able to load a 7B LLM next to a demanding game such as Cyberpunk running on Windows 11 (see Section 6.4).

To address adversarial attacks, we suggested disabling user input that could derail the generated content. Instead, developers should focus on producing only *Chatter* (see 6.3 for definition) that is generated by a pre-defined prompt. Empirical evaluation across five distinct scenarios revealed that this methodology yields desirable outputs with a reliability rate of 79.2% (see 6.3.3). The remaining 20% of cases do not produce toxic content but instead result in either refusal to generate specific outputs (e.g., due to graphic content or a sensitive subject matter) or harmless, albeit irrelevant, responses. These findings underline the practicality and safety of implementing LLM-driven PCG in video game environments.

7.2 Limitations and Future work

In this section we will first give a general outlook on the scientific field, and then in subsections talk about specific limitations and future work regarding our individual research questions.

The main limitation we see in our research is that TRL will not solve generalization in RL, even when combined with PCG. While TRL does enable scaling to tasks not achievable from scratch, as shown in Section 5.6.2, limitations regarding genuinely generalized behavior persist. Solutions based on Foundation Models seem better suited for this [216]. Specifically, Vision Action Models (VLA) show promise in developing agents that act in environments capable of generalizing across multiple domains [142], a feature that we have not observed in any TRL research yet. Furthermore, PCG remains a dynamic and rapidly evolving field. We already highlighted many of its applications for video games and research listed in our introduction and individual work. Especially the influx of generative AI algorithms that create Text [274], Images [57], and Videos [151] will only accelerate its adoption.

7.2.1 Improving Benchmarks

Our empirical experiments in Tetris Link have been conducted using a single RL algorithm (PPO [234]) and could be expanded by conducting comparative analyses across various RL algorithms.

Drawing insights from our work on 3D Linerider (7.1.2) and related work [300], future experiments could attempt to reduce the branching factor of the problem by starting to train on smaller fields that get incrementally bigger.

While we have initiated efforts to enhance reproducibility by establishing verifiability, a proper solution to achieve full reproducibility in RL experiments remains an open challenge.

Lastly, our TRL survey was performed five years ago, making a refresh of the findings with more current papers valuable.

7.2.2 TRL & PCG

Automated curriculum learning (ACL) is an actively evolving research domain. We are aware of one work [300] that combines ACL with PCG, but we observe significant potential for further exploration and innovation in this interdisciplinary space.

Concerning our Linerider experiment: It can be transformed into a Generative Adversarial challenge, akin to the game Super Monkey Ball [83], making it more interesting and enabling the incorporation of ideas from automated curriculum learning. Finally, the constrained scale of world sizes in our Linerider experiments limits their ability to capture the full spectrum of creativity and diversity typically observed in human-generated levels.

7.2.3 LLMs & PCG

The procedural generation of content through LLMs presents a promising area for future research and development that ranges beyond the generation of Non-Player Character (NPC) lines for video games. Examples include Mario [261] or Sokoban [269] level generation. These could also be generated using the locally runnable LLMs we focus on.

The ideal use case of local LLMs would be to have interactive dialogues that immerse the player more into the game world, achievable through dialogue systems [111, 135, 257].

On the topic of chatter generation, our experiment has been limited to a single model configuration (Koala [100], 7B parameter size). A greater variety of configurations should be explored. Especially newer models with even fewer parameters, such as Llama3.2 [78] and Gemma2 [99] would be even better suited as they only require between 1 to 2B parameters.

Bibliography

[1] Adesso, G.: Gpt4: The ultimate brain. Authorea Preprints (2022)

[2] Agarwal, R., Schuurmans, D., Norouzi, M.: An optimistic perspective on offline reinforcement learning. In: International Conference on Machine Learning (2020)

[3] Akiba, T., Sano, S., Yanase, T., Ohta, T., Koyama, M.: Optuna: A Next-generation Hyperparameter Optimization Framework. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. pp. 2623–2631. ACM (2019), https://doi.org/10.1145/3292500.3330701

[4] Al-Nassar, S., Schaap, A., Zwart, M.V.D., Preuss, M., Gómez-Maureira, M.A.: Questville: Procedural quest generation using nlp models. In: Proceedings of the 18th International Conference on the Foundations of Digital Games. pp. 1–4 (2023)

[5] Albrecht, J., Fetterman, A., Fogelman, B., Kitanidis, E., Wróblewski, B., Seo, N., Rosenthal, M., Knutins, M., Polizzi, Z., Simon, J., et al.: Avalon: A benchmark for rl generalization using procedurally generated worlds. Advances in Neural Information Processing Systems **35**, 12813–12825 (2022)

[6] Allain, R., Williams, R.: An analysis of a video game. The Physics Teacher **47** (2009)

[7] Ammar, H.B.: Automated Transfer in Reinforcement Learning. Maastricht University (2013), https://doi.org/10.26481/dis.20130613hb

[8] Ammar, H.B., Eaton, E., Ruvolo, P., Taylor, M.: Unsupervised cross-domain transfer in policy gradient reinforcement learning via manifold alignment. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 29 (2015), https://doi.org/10.1609/aaai.v29i1.9631

[9] Ammar, H.B., Eaton, E., Taylor, M.E., Mocanu, D.C., Driessens, K., Weiss, G., Tuyls, K.: An automated measure of mdp similarity for transfer in reinforcement learning. In: Workshops at the Twenty-Eighth AAAI Conference on Artificial Intelligence (2014), https://aaai.org/papers/aaaiw-ws1240-14-8824/

[10] Andersen, P., Goodwin, M., Granmo, O.: Cairl: A high-performance reinforcement learning environment toolkit. In: Conference on Games. pp. 361–368. IEEE (2022), https://doi.org/10.1109/CoG51982.2022.9893661

[11] Anwar, A., Raychowdhury, A.: Autonomous navigation via deep reinforcement learning for resource constraint edge nodes using transfer learning. IEEE Access **8**, 26549–26560 (2020), https://doi.org/10.1109/ACCESS.2020.2971172

[12] Arabasadi, Z., Didkar, N.: Learning transfer automatic through data mining in reinforcement learning. International Journal of Computer Applications **88**(13) (2014), `https://doi.org/10.5120/15411-3885`

[13] Arenius, P., Sasi, V., Gabrielsson, M.: Rapid internationalisation enabled by the internet: The case of a knowledge intensive company. Journal of International Entrepreneurship **3**(4), 279–290 (2005)

[14] Arnekvist, I., Kragic, D., Stork, J.A.: Vpe: Variational policy embedding for transfer reinforcement learning. In: 2019 International Conference on Robotics and Automation (ICRA). pp. 36–42. IEEE (2019), `https://doi.org/10.1109/ICRA.2019.8793556`

[15] Art_from_the_Machine: Chatgpt in skyrim vr with lip synced voice generation, `https://www.reddit.com/r/singularity/comments/12zveq0`, visited 15.04.2023

[16] Aumjaud, P., McAuliffe, D., Lera, F.J.R., Cardiff, P.: rl_reach: Reproducible reinforcement learning experiments for robotic reaching tasks. Software Impacts **8**, 100061 (2021)

[17] paper authors, A.: Code and trace repository (2022-02-27), `https://ipfs.io/ipfs/QmRDg98PaQEdvrj6tcDr5eQRLCPc2YovphMd7uHNJet1Ug`

[18] Awais, M., Naseer, M., Khan, S., Anwer, R.M., Cholakkal, H., Shah, M., Yang, M.H., Khan, F.S.: Foundation models defining a new era in vision: a survey and outlook. IEEE Transactions on Pattern Analysis and Machine Intelligence pp. 1–20 (2025), `https://doi.org/10.1109/TPAMI.2024.3506283`

[19] Bahna, D.: Grand theft auto: A vehicle for public memory and prosthetic identity. The American Papers p. 81 (2017)

[20] Baker, B., Kanitscheider, I., Markov, T., Wu, Y., Powell, G., McGrew, B., Mordatch, I.: Emergent tool use from multi-agent autocurricula (2020), `https://openreview.net/forum?id=SkxpxJBKwS`

[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`

[22] Barekatain, M., Yonetani, R., Hamaya, M.: Multipolar: Multi-source policy aggregation for transfer reinforcement learning between diverse environmental dynamics. In: Bessiere, C. (ed.) Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20. pp. 3108–3116. International Joint Conferences on Artificial Intelligence Organization (7 2020). doi: 10.24963/ijcai.2020/430, `https://doi.org/10.24963/ijcai.2020/430`, main track

[23] Barrett, S., Taylor, M.E., Stone, P.: Transfer learning for reinforcement learning on a physical robot. In: Ninth International Conference on Autonomous Agents and Multiagent Systems-Adaptive Learning Agents Workshop (AAMAS-ALA) (2010), http://www.cs.utexas.edu/users/ai-lab?AAMASWS10-barrett

[24] Barros, G.A.B., Liapis, A., Togelius, J.: Murder mystery generation from open data. In: Pachet, F., Cardoso, A., Corruble, V., Ghedini, F. (eds.) International Conference on Computational Creativity /ICCC). pp. 197–204. Sony CSL Paris, France (2016), http://www.computationalcreativity.net/iccc2016/wp-content/uploads/2016/01/Murder-Mystery-Generation-from-Open-Data.pdf

[25] Barthet, M., Khalifa, A., Liapis, A., Yannakakis, G.: Generative personas that behave and experience like humans. In: Proceedings of the 17th International Conference on the Foundations of Digital Games. pp. 1–10 (2022)

[26] Barto, A.G., Sutton, R.S., Watkins, C.: Learning and sequential decision making. University of Massachusetts Amherst (1989), https://dl.acm.org/doi/10.5555/896920

[27] BELLMAN, R.: A markovian decision process. Journal of Mathematics and Mechanics **6**(5), 679–684 (1957), http://www.jstor.org/stable/24900506

[28] Berner, C., Brockman, G., Chan, B., Cheung, V., Debiak, P., Dennison, C., Farhi, D., Fischer, Q., Hashme, S., Hesse, C., Józefowicz, R., Gray, S., Olsson, C., Pachocki, J., Petrov, M., de Oliveira Pinto, H.P., Raiman, J., Salimans, T., Schlatter, J., Schneider, J., Sidor, S., Sutskever, I., Tang, J., Wolski, F., Zhang, S.: Dota 2 with large scale deep reinforcement learning. CoRR **abs/1912.06680** (2019), http://arxiv.org/abs/1912.06680

[29] Bevy: Game engine. https://github.com/bevyengine/bevy, accessed on 2024-09-08

[30] Bishop, L., Eberly, D., Whitted, T., Finch, M., Shantz, M.: Designing a PC game engine. IEEE Computer Graphics and Applications **18**(1), 46–53 (1998), https://doi.org/10.1109/38.637270

[31] Bjarnason, A., Reynisson, J.M.: Deeper: adventures in procedural game development in Rust. Bachelor's thesis, Tölvunarfræðideild / Department of Computer Science (2021), https://hdl.handle.net/1946/39502

[32] Blommaert, J., Varis, P.: Conviviality and collectives on social media: Virality, memes, and new social structures. Multilingual Margins: A journal of multilingualism from the periphery **2**(1), 31–31 (2015)

[33] Blondel, V.D., Guillaume, J.L., Lambiotte, R., Lefebvre, E.: Fast unfolding of communities in large networks. Journal of statistical mechanics: theory and experiment **2008**(10), P10008 (2008), https://doi.org/10.1088/1742-5468/2008/10/P10008

[34] Boettiger, C.: An introduction to docker for reproducible research. ACM SIGOPS Operating Systems Review **49**(1), 71–79 (2015)

[35] Bone, N.: A survey of transfer learning methods for reinforcement learning (2008), https://cedar.wwu.edu/computerscience_stupubs/5, computer Science Graduate and Undergraduate Student Scholarship

[36] Bontrager, P., Khalifa, A., Anderson, D., Stephenson, M., Salge, C., Togelius, J.: "superstition" in the network: Deep reinforcement learning plays deceptive games. In: Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE). vol. 15, pp. 10–16. AAAI Press (2019), https://doi.org/10.1609/aiide.v15i1.5218

[37] Bormann, C., Hoffman, P.: Concise binary object representation (cbor). Tech. rep., RFC 7049, DOI 10.17487/RFC7049, October 2013,< https://www. rfc-editor. org … (2013)

[38] Bradford, G.: The History and Analysis of the Supposed Automation Chess Player of M. de Kempelen: Now Exhibiting in this Country, by Mr. Maelzel. Hilliard, Gray & Company (1826)

[39] Brockman, G., Cheung, V., Pettersson, L., Schneider, J., Schulman, J., Tang, J., Zaremba, W.: Openai gym. arXiv preprint arXiv:1606.01540 (2016), https://arxiv.org/abs/1606.01540

[40] Brouwer, S., Akkermans, N., Hendriks, L., Van Uden, H., Wilms, V.: "lass frooby noo!" the interference of song lyrics and meaning on speech intelligibility. Journal of Experimental Psychology: Applied **28**(3), 576 (2022)

[41] Brown, J.A., de Araujo, L.J.P., Grichshenko, A.: Ai space invaders 2021 competition (2021)

[42] Brown, J.A., de Araujo, L.J.P., Grichshenko, A.: Snakes ai competition 2020 and 2021 report (2021)

[43] Brown, N., Sandholm, T.: Superhuman ai for multiplayer poker. Science **365**(6456), 885–890 (2019), https://doi.org/10.1126/science.aay2400

[44] Browne, C.B., Powley, E., Whitehouse, D., Lucas, S.M., Cowling, P.I., Rohlfshagen, P., Tavener, S., Perez, D., Samothrakis, S., Colton, S.: A survey of monte carlo tree search methods. IEEE Transactions on Computational Intelligence and AI in games **4**(1), 1–43 (2012), https://doi.org/10.1109/TCIAIG.2012.2186810

[45] Burgiel, H.: How to lose at Tetris. The Mathematical Gazette **81**(491), 194–200 (1997), https://doi.org/10.2307/3619195

[46] Bölling, L.: Librarian bully attack: Gaslighting state-of-the-art llms into obsolescence (2025-03-25), https://web.archive.org/web/20250331085213/https://humandataexperience.substack.com/p/librarian-bully-attack-gaslighting

[47] Campbell, J., Verbrugge, C.: Procedural generation of rollercoasters. In: Conference on Games. IEEE (2023), https://doi.org/10.1109/CoG57401.2023.10333156

[48] Cao, Y., Li, S., Liu, Y., Yan, Z., Dai, Y., Yu, P.S., Sun, L.: A comprehensive survey of ai-generated content (aigc): A history of generative ai from gan to chatgpt. arXiv preprint arXiv:2303.04226 (2023)

[49] Carr, D.: Applying reinforcement learning to Tetris. Department of Computer Science Rhodes University (2005)

[50] Carroll, J.L., Seppi, K.: Task similarity measures for transfer in reinforcement learning task libraries. In: Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005. vol. 2, pp. 803–808. IEEE (2005), https://doi.org/10.1109/IJCNN.2005.1555955

[51] Cavazza, M., Charles, F., Mead, S.J.: Characters in search of an author: Ai-based virtual storytelling. In: Virtual Storytelling Using Virtual Reality Technologies for Storytelling: International Conference ICVS 2001 Avignon, France, September 27–28, 2001 Proceedings. pp. 145–154. Springer (2001)

[52] Celiberto, L.A., Matsuura, J.P., López de Mántaras, R., Bianchi, R.: Using cases as heuristics in reinforcement learning: a transfer learning application pp. 1211–1217 (2011), https://doi.org/10.5591/978-1-57735-516-8/IJCAI11-206

[53] Center, C.C.: Chatgpt-4o contains security bypass vulnerability through time and search functions called time bandit (2025-03-25), https://web.archive.org/web/20250206190954/https://www.kb.cert.org/vuls/id/733789

[54] Ceron, J.S.O., Castro, P.S.: Revisiting rainbow: Promoting more insightful and inclusive deep reinforcement learning research. In: International Conference on Machine Learning. pp. 1373–1383. PMLR (2021)

[55] Chao, C.: Blokus Game Solver. In: Computational Engineering Commons. California Polytechnic State University (2018), https://digitalcommons.calpoly.edu/cpesp/290

[56] Chaslot, G.M.J.B.C.: Monte-Carlo Tree Search. Ph.D. thesis (2010), https://doi.org/10.26481/dis.20100930gc

[57] Chen, D.Z., Siddiqui, Y., Lee, H., Tulyakov, S., Nießner, M.: Text2tex: Text-driven texture synthesis via diffusion models. In: IEEE/CVF International Conference on Computer Vision, ICCV 2023, Paris, France, October 1-6, 2023.

pp. 18512–18522. IEEE (2023), https://doi.org/10.1109/ICCV51070.2023.01701

[58] Cobbe, K., Hesse, C., Hilton, J., Schulman, J.: Leveraging procedural generation to benchmark reinforcement learning. In: International conference on machine learning. Proceedings of Machine Learning Research, vol. 119, pp. 2048–2056. PMLR (2020), http://proceedings.mlr.press/v119/cobbe20a.html

[59] Cobbe, K., Hesse, C., Hilton, J., Schulman, J.: Leveraging procedural generation to benchmark reinforcement learning **119**, 2048–2056 (2020), http://proceedings.mlr.press/v119/cobbe20a.html

[60] Cohen, B., Pietrzak, K.: The chia network blockchain (2019)

[61] garage contributors, T.: Garage: A toolkit for reproducible reinforcement learning research. https://github.com/rlworkgroup/garage (2019)

[62] Contributors, U.E.S.P.: Skyrim:generic dialogue - reactions to shouting, https://en.uesp.net/wiki/Skyrim:Generic_Dialogue#Reactions_to_Shouting, visited 04.05.2023

[63] Cook, M., Raad, A.: Hyperstate space graphs for automated game analysis. In: 2019 IEEE Conference on Games (CoG). IEEE (2019), https://doi.org/10.1109/CIG.2019.8848026

[64] Coulom, R.: Efficient selectivity and backup operators in Monte-Carlo tree search. In: International conference on computers and games. pp. 72–83. Springer (2006), https://doi.org/10.1007/978-3-540-75538-8_7

[65] Crataco: Tavernai: Atmospheric adventure chat. https://github.com/TavernAI/TavernAI (2023)

[66] Demaine, E.D., Demaine, M.L.: Jigsaw puzzles, edge matching, and polyomino packing: Connections and complexity. Graphs and Combinatorics **23**(1), 195–208 (2007), https://doi.org/10.1007/s00373-007-0713-4

[67] Deng, J., Dong, W., Socher, R., Li, L., Li, K., Fei-Fei, L.: Imagenet: A large-scale hierarchical image database. In: 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2009), 20-25 June 2009, Miami, Florida, USA. pp. 248–255. IEEE Computer Society (2009). doi: 10.1109/CVPR.2009.5206848, https://doi.org/10.1109/CVPR.2009.5206848

[68] Deshpande, A., Murahari, V., Rajpurohit, T., Kalyan, A., Narasimhan, K.: Toxicity in chatgpt: Analyzing persona-assigned language models. arXiv preprint arXiv:2304.05335 (2023)

[69] Deutsch, P., Gailly, J.L.: Zlib compressed data format specification version 3.3. Tech. rep., RFC 1950, May (1996)

[70] Dietterich, T.G.: Hierarchical reinforcement learning with the maxq value function decomposition. Journal of artificial intelligence research **13**, 227–303 (2000), https://doi.org/10.1613/jair.639

[71] Dimforge: Rapier physics engine. https://github.com/dimforge/rapier, accessed on 2024-09-08

[72] Donadigo, A.: Extracting inputs from replays (2021-12-15), https://donadigo.com/tminterface/input-extraction

[73] Donadigo, A.: Tmx replay investigation (2021-12-15), https://donadigo.com/tmx1

[74] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., et al.: An image is worth 16x16 words: Transformers for image recognition at scale. arXiv preprint arXiv:2010.11929 (2020)

[75] Driess, D., Xia, F., Sajjadi, M.S.M., Lynch, C., Chowdhery, A., Ichter, B., Wahid, A., Tompson, J., Vuong, Q., Yu, T., Huang, W., Chebotar, Y., Sermanet, P., Duckworth, D., Levine, S., Vanhoucke, V., Hausman, K., Toussaint, M., Greff, K., Zeng, A., Mordatch, I., Florence, P.: Palm-e: An embodied multimodal language model. In: arXiv preprint arXiv:2303.03378 (2023)

[76] Duan, Y., Chen, X., Houthooft, R., Schulman, J., Abbeel, P.: Benchmarking deep reinforcement learning for continuous control. In: International conference on machine learning. pp. 1329–1338. PMLR (2016), http://proceedings.mlr.press/v48/duan16.html

[77] Duan, Y., Schulman, J., Chen, X., Bartlett, P.L., Sutskever, I., Abbeel, P.: Rl^2: Fast reinforcement learning via slow reinforcement learning (2016), http://arxiv.org/abs/1611.02779

[78] Dubey, A., Jauhri, A., Pandey, A., Kadian, A., Al-Dahle, A., Letman, A., Mathur, A., Schelten, A., Yang, A., Fan, A., Goyal, A., Hartshorn, A., Yang, A., Mitra, A., Sravankumar, A., Korenev, A., Hinsvark, A., Rao, A., Zhang, A., Rodriguez, A., Gregerson, A., Spataru, A., Rozière, B., Biron, B., Tang, B., Chern, B., Caucheteux, C., Nayak, C., Bi, C., Marra, C., McConnell, C., Keller, C., Touret, C., Wu, C., Wong, C., Ferrer, C.C., Nikolaidis, C., Allonsius, D., Song, D., Pintz, D., Livshits, D., Esiobu, D., Choudhary, D., Mahajan, D., Garcia-Olano, D., Perino, D., Hupkes, D., Lakomkin, E., AlBadawy, E., Lobanova, E., Dinan, E., Smith, E.M., Radenovic, F., Zhang, F., Synnaeve, G., Lee, G., Anderson, G.L., Nail, G., Mialon, G., Pang, G., Cucurell, G., Nguyen, H., Korevaar, H., Xu, H., Touvron, H., Zarov, I., Ibarra, I.A., Kloumann, I.M., Misra, I., Evtimov, I., Copet, J., Lee, J., Geffert, J., Vranes, J., Park, J., Mahadeokar, J., Shah, J., van der Linde, J., Billock, J., Hong, J., Lee, J., Fu, J., Chi, J., Huang, J., Liu, J., Wang, J., Yu, J., Bitton, J.,

Spisak, J., Park, J., Rocca, J., Johnstun, J., Saxe, J., Jia, J., Alwala, K.V., Upasani, K., Plawiak, K., Li, K., Heafield, K., Stone, K., et al.: The llama 3 herd of models. CoRR **abs/2407.21783** (2024). doi: 10.48550/ARXIV.2407.21783, https://doi.org/10.48550/arXiv.2407.21783

[79] van Eck, N.J., van Wezel, M.: Application of reinforcement learning to the game of Othello. Computers & Operations Research **35**(6), 1999–2017 (2008), https://doi.org/10.1016/j.cor.2006.10.004

[80] Ecoffet, A., Huizinga, J., Lehman, J., Stanley, K.O., Clune, J.: First return, then explore. Nature **590**(7847), 580–586 (2021), https://doi.org/10.1038/s41586-020-03157-9

[81] Edelkamp, S., Korf, R.E.: The branching factor of regular search spaces. In: Fifteenth National Conference on Artificial Intelligence (AAAI) and Tenth Innovative Applications of Artificial Intelligence Conferenc (IAAI). pp. 299–304 (1998), https://dl.acm.org/doi/10.5555/295240.295615

[82] Edwards, B.: You can now run a gpt-3-level ai model on your laptop, phone, and raspberry pi, https://arstechnica.com/information-technology/2023/03/you-can-now-run-a-gpt-3-level-ai-model-on-your-laptop-phone-and-raspberry-pi/, visited 28.03.2023

[83] Egenfeldt-Nielsen, S.: Keep the monkey rolling: eye-hand coordination in super monkey ball. In: Proceedings of DiGRA 2003 Conference: Level Up. DiGRA (2003), https://dl.digra.org/index.php/dl/article/view/40/

[84] Eisner, D.: Reproducibility of science: Fraud, impact factors and carelessness. Journal of molecular and cellular cardiology **114**, 364–368 (2018)

[85] Enzenberger, M., Müller, M.: A lock-free multithreaded Monte-Carlo tree search algorithm. In: Advances in Computer Games. Lecture Notes in Computer Science, vol. 6048, pp. 14–20. Springer (2009), https://doi.org/10.1007/978-3-642-12993-3_2

[86] Ermi, L., Mäyrä, F.: Analyzing immersion. Worlds in play: International perspectives on digital games research **21**, 37 (2007)

[87] European Organization For Nuclear Research, OpenAIRE: Zenodo (2013). doi: 10.25495/7GXK-RD71, https://www.zenodo.org/

[88] Fandom.com: History of world generation - minecraft wiki, https://minecraft.fandom.com/wiki/History_of_world_generation, visited 11.05.2023

[89] Farooq, S.S., Oh, I.S., Kim, M.J., Kim, K.J.: Starcraft ai competition report. AI Magazine **37**(2), 102–107 (2016)

[90] Ferigo, D., Traversaro, S., Metta, G., Pucci, D.: Gym-ignition: Reproducible robotic simulations for reinforcement learning. In: 2020 IEEE/SICE International Symposium on System Integration (SII). pp. 885–890. IEEE (2020)

[91] Floridi, L., Chiriatti, M.: GPT-3: its nature, scope, limits, and consequences. Minds Mach. **30**(4), 681–694 (2020), https://doi.org/10.1007/s11023-020-09548-1

[92] Font, J.M., Mahlmann, T.: Dota 2 bot competition. IEEE Transactions on Games **11**(3), 285–289 (2018)

[93] Gaina, R.D., Couëtoux, A., Soemers, D.J., Winands, M.H., Vodopivec, T., Kirchgeßner, F., Liu, J., Lucas, S.M., Perez-Liebana, D.: The 2016 two-player gvgai competition. IEEE Transactions on Games **10**(2), 209–220 (2017)

[94] Gaina, R.D., Lucas, S.M., Liebana, D.P.: Rolling horizon evolution enhancements in general video game playing. In: IEEE Conference on Computational Intelligence and Games, CIG 2017, New York, NY, USA, August 22-25, 2017. pp. 88–95. IEEE (2017). doi: 10.1109/CIG.2017.8080420, https://doi.org/10.1109/CIG.2017.8080420

[95] Galván-López, E., Li, R., Patsakis, C., Clarke, S., Cahill, V.: Heuristic-Based Multi-Agent Monte Carlo Tree Search. In: IISA 2014, The 5th International Conference on Information, Intelligence, Systems and Applications. pp. 177–182. IEEE (2014), https://doi.org/10.1109/IISA.2014.6878747

[96] Games, E.: Replay system - unreal engine documentation (2021-12-15), https://docs.unrealengine.com/latest/INT/Engine/Replay/

[97] Gamrian, S., Goldberg, Y.: Transfer learning for related reinforcement learning tasks via image-to-image translation **97**, 2063–2072 (2019), http://proceedings.mlr.press/v97/gamrian19a.html

[98] Gao, C., Hayward, R., Müller, M.: Move prediction using deep convolutional neural networks in Hex. IEEE Transactions on Games **10**(4), 336–343 (2017), https://doi.org/10.1109/TG.2017.2785042

[99] Gemma Team: Gemma (2024), https://doi.org/10.34740/KAGGLE/M/3301

[100] Geng, X., Gudibande, A., Liu, H., Wallace, E., Abbeel, P., Levine, S., Song, D.: Koala: A dialogue model for academic research. Blog post (April 2023), https://bair.berkeley.edu/blog/2023/04/03/koala/

[101] Gerganov, G.: Llama.cpp: Port of facebook's llama model in c/c++, https://github.com/ggerganov/llama.cpp, visited 28.03.2023

[102] Gisslén, L., Eakins, A., Gordillo, C., Bergdahl, J., Tollmar, K.: Adversarial reinforcement learning for procedural content generation (2021), https://doi.org/10.1109/CoG52621.2021.9619053

[103] Glickman, M.E., Jones, A.C.: Rating the chess rating system. CHANCE-BERLIN THEN NEW YORK- **12**, 21–28 (1999)

[104] Green, M.C., Sergent, B., Shandilya, P., Kumar, V.: Evolutionarily-curated curriculum learning for deep reinforcement learning agents (2019), http://arxiv.org/abs/1901.05431

[105] Gregory, J.: Game engine architecture, Third Edition. AK Peters/CRC Press (2018), GameEngineArchitecture,ThirdEdition

[106] Grzes, M., Kudenko, D.: Plan-based reward shaping for reinforcement learning. In: 2008 4th International IEEE Conference Intelligent Systems. vol. 2, pp. 10–22. IEEE (2008), https://doi.org/10.1017/S0269888915000181

[107] Guez, A., Weber, T., Antonoglou, I., Simonyan, K., Vinyals, O., Wierstra, D., Munos, R., Silver, D.: Learning to search with mctsnets. In: International Conference on Machine Learning. pp. 1822–1831. PMLR (2018), http://proceedings.mlr.press/v80/guez18a.html

[108] Gupta, R., Kanade, A., Shevade, S.K.: Deep reinforcement learning for syntactic error repair in student programs. In: Conference on Artificial Intelligence. pp. 930–937. AAAI Press (2019), https://doi.org/10.1609/aaai.v33i01.3301930

[109] Guss, W.H., Castro, M.Y., Devlin, S., Houghton, B., Kuno, N.S., Loomis, C., Milani, S., Mohanty, S., Nakata, K., Salakhutdinov, R., Schulman, J., Shiroshita, S., Topin, N., Ummadisingu, A., Vinyals, O.: The minerl 2020 competition on sample efficient reinforcement learning using human priors (January 2021), https://www.microsoft.com/en-us/research/publication/the-minerl-2020-competition-on-sample-efficient-reinforcement-learning-using-human-priors/

[110] Guss, W.H., Codel, C., Hofmann, K., Houghton, B., Kuno, N.S., Milani, S., Mohanty, S., Liebana, D.P., Salakhutdinov, R., Topin, N., Veloso, M., Wang, P.: The minerl competition on sample efficient reinforcement learning using human priors. In: Thirty-third Conference on Neural Information Processing Systems (NeurIPS) Competition track (December 2019), https://www.microsoft.com/en-us/research/publication/the-minerl-competition-on-sample-efficient-reinforcement-learning-using-human-priors/

[111] Hämäläinen, M., Alnajjar, K.: Creative contextual dialog adaptation in an open world rpg. In: Proceedings of the 14th International Conference on the Foundations of Digital Games. pp. 1–7 (2019)

[112] Hamilton, N., Schlemmer, L., Menart, C., Waddington, C., Jenkins, T., Johnson, T.T.: Sonic to knuckles: evaluations on transfer reinforcement learning. In: Unmanned Systems Technology XXII. vol. 11425, p. 114250J. International

Society for Optics and Photonics (2020), https://ui.adsabs.harvard.edu/link_gateway/2020SPIE11425E..0JH/doi:10.1117/12.2559546

[113] Harsana, L.: How much ram does windows 11 use? [memory requirements], https://windowsreport.com/does-windows-11-consume-more-ram/, visited 28.03.2023

[114] Henderson, P., Islam, R., Bachman, P., Pineau, J., Precup, D., Meger, D.: Deep reinforcement learning that matters. In: Thirty-Second AAAI Conference on Artificial Intelligence. pp. 3207–3214. AAAI Press (2018), https://doi.org/10.1609/aaai.v32i1.11694

[115] Hendrikx, M., Meijer, S.A., Velden, J.V.D., Iosup, A.: Procedural content generation for games: A survey. Transactions on Multimedia Computing, Communications, and Applications **9** (2013), https://doi.org/10.1145/2422956.2422957

[116] Higgins, I., Pal, A., Rusu, A., Matthey, L., Burgess, C., Pritzel, A., Botvinick, M., Blundell, C., Lerchner, A.: Darla: Improving zero-shot transfer in reinforcement learning. In: International Conference on Machine Learning. vol. 70, pp. 1480–1490. PMLR (2017), http://proceedings.mlr.press/v70/higgins17a.html

[117] Hill, A., Raffin, A., Ernestus, M., Gleave, A., Kanervisto, A., Traore, R., Dhariwal, P., Hesse, C., Klimov, O., Nichol, A., Plappert, M., Radford, A., Schulman, J., Sidor, S., Wu, Y.: Stable Baselines. https://github.com/hill-a/stable-baselines (2018)

[118] Hoetzlein, R.: A procedural model for diverse tree species. In: Karpouzis, K., Gualeni, S., Pirker, J., Fowler, A. (eds.) FDG '22: Proceedings of the 17th International Conference on the Foundations of Digital Games, FDG22, Athens, Greece, September 5-8, 2022. pp. 72:1–72:8. ACM (2022). doi: 10.1145/3555858.3564251, https://doi.org/10.1145/3555858.3564251

[119] Hou, Y., Ong, Y.S., Feng, L., Zurada, J.M.: An evolutionary transfer reinforcement learning framework for multiagent systems. IEEE Transactions on Evolutionary Computation **21**(4), 601–615 (2017), https://doi.org/10.1109/TEVC.2017.2664665

[120] Hsu, F.h.: Computer chess, then and now: The Deep Blue saga. In: Proceedings of Technical Papers. International Symposium on VLSI Technology, Systems, and Applications. pp. 153–156 (1997), https://doi.org/10.1109/VTSA.1997.614748

[121] Hussain, T.S., Vidaver, G.: Flexible and purposeful NPC behaviors using real-time genetic control. In: International Conference on Evolutionary Computation(CEC). pp. 785–792. IEEE (2006), https://doi.org/10.1109/CEC.2006.1688391

[122] Inc., C.T.: Character.ai, https://character.ai, visited 28.03.2023

[123] Inc., S.I.E.: "race together," a new event with gran turismo sophy, available for a limited time only!, https://www.gran-turismo.com/us/gt7/news/00_5962264.html, visited 09.05.2023

[124] Islam, R., Henderson, P., Gomrokchi, M., Precup, D.: Reproducibility of benchmarked deep reinforcement learning tasks for continuous control (2017)

[125] Jacomy, M., Venturini, T., Heymann, S., Bastian, M.: Forceatlas2, a continuous graph layout algorithm for handy network visualization designed for the gephi software. Plos One **9**(6), e98679 (2014), https://doi.org/10.1371/journal.pone.0098679

[126] Jahanshahi, A., Taram, M.K., Eskandari, N.: Blokus Duo game on FPGA. In: The 17th CSI International Symposium on Computer Architecture & Digital Systems (CADS 2013). pp. 149–152. IEEE (2013), https://doi.org/10.1109/CADS.2013.6714256

[127] Jaschek, C., Beckmann, T., Garcia, J.A., Raffe, W.L.: Mysterious murder-mcts-driven murder mystery generation. In: 2019 IEEE Conference on Games (CoG). pp. 1–8. IEEE (2019)

[128] Jeurissen, D., Liebana, D.P., Gow, J., Çakmak, D., Kwan, J.: Playing nethack with llms: Potential & limitations as zero-shot agents. In: IEEE Conference on Games, CoG 2024, Milan, Italy, August 5-8, 2024. pp. 1–8. IEEE (2024), https://doi.org/10.1109/CoG60054.2024.10645630

[129] Jiang, Z., Earle, S., Green, M.C., Togelius, J.: Learning controllable 3d level generators. In: Foundations of Digital Games. ACM (2022), https://doi.org/10.1145/3555858.3563273

[130] Jordan, M.I., Mitchell, T.M.: Machine learning: Trends, perspectives, and prospects. Science **349**(6245), 255–260 (2015), https://doi.org/10.1126/science.aaa8415

[131] Joshi, G., Chowdhary, G.: Cross-domain transfer in reinforcement learning using target apprentice. In: 2018 IEEE International Conference on Robotics and Automation (ICRA). pp. 7525–7532. IEEE (2018), https://doi.org/10.1109/ICRA.2018.8462977

[132] Joshi, S., Khan, M.S., Dafe, A., Singh, K., Zope, V., Jhamtani, T.: Fine tuning llms for low resource languages. In: 2024 5th International Conference on Image Processing and Capsule Networks (ICIPCN). pp. 511–519 (2024), https://doi.org/10.1109/ICIPCN63822.2024.00090

[133] Justesen, N., Moore, P.D., Uth, L.M., Togelius, J., Jakobsen, C., Risi, S.: Blood bowl: A new board game challenge and competition for ai. In: 2019 IEEE Conference on Games (COG). IEEE (2019)

[134] Justesen, N., Torrado, R.R., Bontrager, P., Khalifa, A., Togelius, J., Risi, S.: Illuminating generalization in deep reinforcement learning through procedural level generation. arXiv preprint arXiv:1806.10729 https://arxiv.org/abs/1806.10729

[135] Kacmarcik, G.: Using natural language to manage npc dialog. In: Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment. vol. 2, pp. 115–117 (2006)

[136] Kaelbling, L.P., Littman, M.L., Moore, A.W.: Reinforcement learning: A survey. Journal of artificial intelligence research **4**, 237–285 (1996), https://doi.org/10.1613/jair.301

[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

[138] Kaspar, M., Osorio, J.D.M., Bock, J.: Sim2real transfer for reinforcement learning without dynamics randomization pp. 4383–4388 (2020), https://doi.org/10.1109/IROS45743.2020.9341260

[139] Kaufmann, E., Bauersfeld, L., Loquercio, A., Müller, M., Koltun, V., Scaramuzza, D.: Champion-level drone racing using deep reinforcement learning. Nature **620**(7976), 982–987 (2023), https://doi.org/10.1038/s41586-023-06419-4

[140] Khalifa, A., Bontrager, P., Earle, S., Togelius, J.: Pcgrl: Procedural content generation via reinforcement learning. In: Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment. vol. 16, pp. 95–101 (2020), https://doi.org/10.1609/aiide.v16i1.7416

[141] Khetarpal, K., Ahmed, Z., Cianflone, A., Islam, R., Pineau, J.: Re-evaluate: Reproducibility in evaluating reinforcement learning algorithms (2018)

[142] Kim, M.J., Pertsch, K., Karamcheti, S., Xiao, T., Balakrishna, A., Nair, S., Rafailov, R., Foster, E.P., Sanketi, P.R., Vuong, Q., Kollar, T., Burchfiel, B., Tedrake, R., Sadigh, D., Levine, S., Liang, P., Finn, C.: Openvla: An open-source vision-language-action model. In: Agrawal, P., Kroemer, O., Burgard, W. (eds.) Conference on Robot Learning, 6-9 November 2024, Munich, Germany. Proceedings of Machine Learning Research, vol. 270, pp. 2679–2713. PMLR (2024), https://proceedings.mlr.press/v270/kim25c.html

[143] Kocsis, L., Szepesvári, C.: Bandit based monte-carlo planning. In: European conference on machine learning. pp. 282–293. Springer (2006), https://doi.org/10.1007/11871842_29

[144] Konidaris, G.D., Barto, A.G.: Building portable options: Skill transfer in reinforcement learning. In: International Joint Conference on Artificial Intelligence (IJCAI). vol. 7, pp. 895–900 (2007), http://ijcai.org/Proceedings/07/Papers/144.pdf

[145] Kroon, J.: Nemesis Narratives: The relationship between embedded and emergent narrative in Middle Earth: Shadow of Mordor. Master's thesis, Utrecht University (2016), https://studenttheses.uu.nl/handle/20.500.12932/23201

[146] KURT, F., ÖZGÖVDE, B.: Edge computing for computer games by offloading physics computation. Journal of Science Part A: Engineering and Innovation **10** (2023), https://doi.org/10.54287/gujsa.1338594

[147] Langley, P.: Elements of Machine Learning. Morgan Kaufmann Publishers Inc., 340 Pine Street, Sixth Floor, San Francisco, CA, USA (1995), https://dl.acm.org/doi/10.5555/212543

[148] Lawrence, S., Giles, C.L., Tsoi, A.C.: Lessons in neural network training: Overfitting may be harder than expected. In: AAAI/IAAI. pp. 540–545. Citeseer (1997), http://www.aaai.org/Library/AAAI/1997/aaai97-084.php

[149] Lazaric, A., Hal, H.I., Lazaric, A.: Transfer in reinforcement learning: A framework and a survey. In: Wiering, M.A., van Otterlo, M. (eds.) Reinforcement Learning, Adaptation, Learning, and Optimization, vol. 12, pp. 143–173. Springer (2012), https://doi.org/10.1007/978-3-642-27645-3_5

[150] Li, X., Cen, M., Xu, J., Zhang, H., Xu, X.S.: Improving feature extraction from histopathological images through A fine-tuning imagenet model. CoRR **abs/2201.00636** (2022), https://arxiv.org/abs/2201.00636

[151] Li, Y., Min, M.R., Shen, D., Carlson, D.E., Carin, L.: Video generation from text. In: McIlraith, S.A., Weinberger, K.Q. (eds.) Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18), the 30th innovative Applications of Artificial Intelligence (IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-18), New Orleans, Louisiana, USA, February 2-7, 2018. pp. 7065–7072. AAAI Press (2018). doi: 10.1609/AAAI.V32I1.12233, https://doi.org/10.1609/aaai.v32i1.12233

[152] Liebana, D.P., Samothrakis, S., Lucas, S.M., Rohlfshagen, P.: Rolling horizon evolution versus tree search for navigation in single-player real-time games. In: Blum, C., Alba, E. (eds.) Genetic and Evolutionary Computation Conference, GECCO '13, Amsterdam, The Netherlands, July 6-10, 2013. pp. 351–358. ACM (2013). doi: 10.1145/2463372.2463413, https://doi.org/10.1145/2463372.2463413

[153] Lissa, P., Schukat, M., Barrett, E.: Transfer learning applied to reinforcement learning-based hvac control. SN Computer Science **1**(3), 1–12 (2020), https://doi.org/10.1007/s42979-020-00146-7

[154] Liu, C., Gao, C., Xia, X., Lo, D., Grundy, J., Yang, X.: On the replicability and reproducibility of deep learning in software engineering. arXiv preprint arXiv:2006.14244 (2020)

[155] López-Ibáñez, M., Branke, J., Paquete, L.: Reproducibility in evolutionary computation. ACM Transactions on Evolutionary Learning and Optimization **1**(4), 1–21 (2021)

[156] Lu, D.: I spent 11 years working on this line rider track. https://web.archive.org/web/20231020213756/https://delu.medium.com/i-spent-11-years-working-on-this-line-rider-track-96742fc0b709, accessed on 2024-09-08

[157] Maberry, K., Paustian, S., Bakir, S.: Using an artificial neural network to detect aim assistance in counter-strike: Global offensive https://www.cs.nmt.edu/~kmaberry/ann_fps_cheater.pdf

[158] Malisz, Z., Henter, G.E., Valentini-Botinhao, C., Watts, O., Beskow, J., Gustafson, J.: Modern speech synthesis for phonetic sciences: A discussion and an evaluation. In: International Congress of Phonetic Sciencesnces ICPhS 2019 5-9 August 2019, Melbourne, Australia Melbourne Convention and Exhibition Centre (2019)

[159] Markocki, M.: Reactive games as an example of extensive use of evocative narrative elements in digital games: cases of dwarf fortress and rimworld. Studia Humanistyczne AGH **20**(2), 71–83 (2021)

[160] Matsakis, N.D., Klock II, F.S.: The rust language. In: SIGAda annual conference on High integrity language technology (HILT). pp. 103–104. ACM (2014), https://doi.org/10.1145/2663171.2663188

[161] Matsubara, H., Iida, H., Grimbergen, R.: Chess, Shogi, Go, natural developments in game research. International Computer Games Associatoin (ICGA) **19**(2), 103–112 (1996), https://doi.org/10.3233/ICG-1996-19208

[162] Merriam-Webster: Chatter definition, https://www.merriam-webster.com/dictionary/chatter, visited 10.05.2023

[163] Michael Färber: The Microsoft Academic Knowledge Graph: A Linked Data Source with 8 Billion Triples of Scholarly Data. In: Proceedings of the 18th International Semantic Web Conference. pp. 113–129. ISWC'19 (2019). doi: 10.1007/978-3-030-30796-7_8, https://doi.org/10.1007/978-3-030-30796-7_8

[164] Miriti, K.: Integrating policy transfer, policy reuse and experience replay in speeding up reinforcement learning of the obstacle avoidance task. Ph.D. thesis, University of Nairobi (2014), http://erepository.uonbi.ac.ke/handle/11295/90177

[165] Mo, K., Li, S., Zhang, Y., Li, J., Yang, Q.: Personalizing a dialogue system with transfer reinforcement learning (2017)

[166] MobyGames: Line rider 2: Unbound. https://web.archive.org/web/20231203034915/https://www.mobygames.com/game/37725/line-rider-2-unbound/, accessed on 2024-09-08

[167] Mocanu, E., Nguyen, P.H., Kling, W.L., Gibescu, M.: Unsupervised energy prediction in a smart grid context using reinforcement cross-building transfer learning. Energy and Buildings **116**, 646–655 (2016), https://doi.org/10.1016/j.enbuild.2016.01.030

[168] Moerland, T.M.: The intersection of Planning and Learning. Ph.D. thesis, Technische Universiteit Delft (2021), http://hdl.handle.net/1842/879

[169] Moerland, T.M., Broekens, J., Jonker, C.M.: A framework for reinforcement learning and planning (2020), https://arxiv.org/abs/2006.15009

[170] Mole, B.: Top harvard cancer researchers accused of scientific fraud; 37 studies affected (2025-01-28), https://arstechnica.com/gadgets/2020/12/new-risc-v-cpu-claims-recordbreaking-performance-per-watt/

[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

[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

[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

[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

[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

[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

[177] Muralidharan, S., Ko, H.: An interplanetary file system (ipfs) based iot framework. In: 2019 IEEE International Conference on Consumer Electronics (ICCE). pp. 1–2 (2019). doi: 10.1109/ICCE.2019.8662002

[178] Muratore, F., Gienger, M., Peters, J.: Assessing transferability from simulation to reality for reinforcement learning. IEEE transactions on pattern analysis and machine intelligence **43**(4), 1172–1183 (2019), https://doi.org/10.1109/TPAMI.2019.2952353

[179] Murray, S.: High art/low life: The art of playing" grand theft auto". PAJ: A Journal of Performance and Art pp. 91–98 (2005)

[180] Müller-Brockhausen, M.: Interactive website to explore the data presented in this paper (2021), https://hizoul.github.io/trlsnap/

[181] Narasimhan, K., Barzilay, R., Jaakkola, T.: Grounding language for transfer in deep reinforcement learning. Journal of Artificial Intelligence Research **63**, 849–874 (2018), https://doi.org/10.1613/jair.1.11263

[182] Narvekar, S., Peng, B., Leonetti, M., Sinapov, J., Taylor, M.E., Stone, P.: Curriculum learning for reinforcement learning domains: A framework and survey. Journal of Machine Learning Research **21**(181), 1–50 (2020), https://jmlr.org/papers/v21/20-212.html

[183] Narvekar, S., Peng, B., Leonetti, M., Sinapov, J., Taylor, M.E., Stone, P.: Curriculum learning for reinforcement learning domains: A framework and survey. Journal of Machine Learning Research **21**(181), 1–50 (2020), https://jmlr.org/papers/v21/20-212.html

[184] Nasir, M.U., James, S., Togelius, J.: Gametraversalbenchmark: Evaluating planning abilities of large language models through traversing 2d game maps (2024), https://arxiv.org/abs/2410.07765

[185] National Academies of Sciences, Engineering, and Medicine and others: Reproducibility and replicability in science. National Academies Press (2019)

[186] Onishi, K., Murakami, N., Kitamura, Y., Kishino, F.: Modeling of trees with interactive l-system and 3d gestures. In: Ijspeert, A.J., Masuzawa, T., Kusumoto, S. (eds.) Biologically Inspired Approaches to Advanced Information Technology, Second International Workshop, BioADIT 2006, Osaka, Japan, January 26-27, 2006, Proceedings. Lecture Notes in Computer Science, vol. 3853, pp. 222–235. Springer (2006). doi: 10.1007/11613022_19, https://doi.org/10.1007/11613022_19

[187] OpenAI: Gpt-4 technical report (2023)

[188] Orenstein, P.: Does experiential learning improve learning outcomes in an undergraduate course in game theory–a preliminary analysis. Northeast Decision Sciences Institute (2014)

[189] Park, J.S., O'Brien, J.C., Cai, C.J., Morris, M.R., Liang, P., Bernstein, M.S.: Generative agents: Interactive simulacra of human behavior (2023)

[190] Partalas, I., Tsoumakas, G., Tzevanidis, K., Vlahavas, I.P.: Transferring experience in reinforcement learning through task decomposition. In: International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS). pp. 1193–1194. IFAAMAS (2009), https://dl.acm.org/doi/10.5555/1558109.1558208

[191] PassMark: Passmark software - hardware survey - amount of ram installed, https://www.memorybenchmark.net/amount-of-ram-installed.html, visited 28.03.2023

[192] Patel, D., Ahmad, A.: Google "we have no moat, and neither does openai", https://www.semianalysis.com/p/google-we-have-no-moat-and-neither, visited 05.05.2023

[193] Patterson, D., Gonzalez, J., Le, Q., Liang, C., Munguia, L.M., Rothchild, D., So, D., Texier, M., Dean, J.: Carbon emissions and large neural network training. arXiv preprint arXiv:2104.10350 (2021), https://arxiv.org/abs/2104.10350

[194] Pereira, A., Proenca, A.: Prng-broker: A high-performance broker to supply parallel streams of pseudorandom numbers for large-scale simulations. In: Advances in Parallel & Distributed Processing, and Applications. pp. 167–183. Springer (2021), https://doi.org/10.1007/978-3-030-69984-0_14

[195] Perez-Liebana, D., Liu, J., Khalifa, A., Gaina, R.D., Togelius, J., Lucas, S.M.: General video game ai: A multitrack framework for evaluating agents, games, and content generation algorithms. IEEE Transactions on Games **11**(3), 195–214 (2019), https://doi.org/10.1109/TG.2019.2901021

[196] Perez-Liebana, D., Lucas, S.M., Gaina, R.D., Togelius, J., Khalifa, A., Liu, J.: General Video Game Artificial Intelligence, vol. 3. Morgan & Claypool Publishers (2019), https://doi.org/10.2200/S00944ED1V01Y201908GCI005

[197] Pérez-Liébana, D., Samothrakis, S., Togelius, J., Schaul, T., Lucas, S.M.: Analyzing the robustness of general video game playing agents. In: 2016 IEEE Conference on Computational Intelligence and Games (CIG). pp. 1–8. IEEE (2016)

[198] Philidor, F.D.: Analysis of the Game of Chess. P. Elmsly (1790)

[199] Pineau, J.: Reproducibility, Reusability, and Robustness in Deep Reinforcement Learning (2020-01-18), https://www.youtube.com/watch?v=Vh4H0gOwdIg, international Conference on Learning Representations (ICLR)

[200] Pineau, J., Sinha, K., Fried, G., Ke, R.N., Larochelle, H.: Reproducibility Challenge (2020-01-18), https://www.cs.mcgill.ca/~jpineau/ICLR2018-ReproducibilityChallenge.html, international Conference on Learning Representations (ICLR)

[201] Pineau, J., Vincent-Lamarre, P., Sinha, K., Larivière, V., Beygelzimer, A., d'Alché Buc, F., Fox, E., Larochelle, H.: Improving reproducibility in machine learning research (a report from the neurips 2019 reproducibility program). Journal of Machine Learning Research **22** (2021)

[202] Plaat, A.: Research, Re: Search and Re-Search. Ph.D. thesis, Erasmus University Rotterdam (1996), https://arxiv.org/abs/2403.13705

[203] Plaat, A.: Learning to Play: Reinforcement Learning and Games. Springer Nature (2020)

[204] Plaat, A.: Deep Reinforcement Learning. Springer Nature (2021)

[205] Plaat, A., Schaeffer, J., Pijls, W., De Bruin, A.: Best-first fixed-depth minimax algorithms. Artificial Intelligence **87**(1-2), 255–293 (1996), https://doi.org/10.1016/0004-3702(95)00126-3

[206] Plisnier, H., Steckelmacher, D., Roijers, D.M., Nowé, A.: Transfer reinforcement learning across environment dynamics with multiple advisors. In: Proceedings of the 31st Benelux Conference on Artificial Intelligence (BNAIC 2019) and the 28th Belgian Dutch Conference on Machine Learning (Benelearn 2019), Brussels, Belgium, November 6-8, 2019. CEUR Workshop Proceedings, vol. 2491. CEUR-WS.org (2019), https://ceur-ws.org/Vol-2491/paper11.pdf

[207] Post, I., Ye, Y.: The simplex method is strongly polynomial for deterministic markov decision processes. Mathematics of Operations Research **40**(4), 859–868 (2015)

[208] Preuss, M.: Multimodal Optimization by Means of Evolutionary Algorithms. Natural Computing Series, Springer (2015), https://doi.org/10.1007/978-3-319-07407-8

[209] Prins, V.L.: Dungeons & Firearms: AI-Directing Action Intensity of Procedural Levels. Master's thesis, Leiden University, The Netherlands (2016), https://theses.liacs.nl/2822

[210] Prins, V.L., Prins, J., Preuss, M., Maureira, M.A.G.: Storyworld: Procedural quest generation rooted in variety & believability. In: Lopes, P., Luz, F., Liapis, A., Engström, H. (eds.) Proceedings of the 18th International Conference on the Foundations of Digital Games, FDG 2023, Lisbon, Portugal, April 12-14, 2023.

pp. 44:1–44:4. ACM (2023). doi: 10.1145/3582437.3587181, https://doi.org/10.1145/3582437.3587181

[211] Purdy, K.: Fallout 4 mod uses voice ai to add sensible reactions, more rpg-like choices, https://arstechnica.com/gaming/2023/03/ai-voice-generation-helps-mod-fallout-4s-narrow-dialogue-choices/, visited 28.03.2023

[212] Pyo3: Rust bindings for python. https://github.com/PyO3/pyo3, accessed on 2024-09-08

[213] PyTorch: Reproducibility - pytorch documentation (2021-12-15), https://pytorch.org/docs/stable/notes/randomness.html

[214] Qasemi, E., Samadi, A., Shadmehr, M.H., Azizian, B., Mozaffari, S., Shirian, A., Alizadeh, B.: Highly scalable, shared-memory, Monte-Carlo tree search based Blokus Duo Solver on FPGA. In: 2014 International Conference on Field-Programmable Technology (FPT). pp. 370–373. IEEE (2014), https://doi.org/10.1109/FPT.2014.7082823

[215] Raffin, A., Hill, A., Gleave, A., Kanervisto, A., Ernestus, M., Dormann, N.: Stable-baselines3: Reliable reinforcement learning implementations. Journal of Machine Learning Research **22**(268), 1–8 (2021), http://jmlr.org/papers/v22/20-1364.html

[216] Ramé, A., Ahuja, K., Zhang, J., Cord, M., Bottou, L., Lopez-Paz, D.: Model ratatouille: Recycling diverse models for out-of-distribution generalization. In: Krause, A., Brunskill, E., Cho, K., Engelhardt, B., Sabato, S., Scarlett, J. (eds.) International Conference on Machine Learning, ICML 2023, 23-29 July 2023, Honolulu, Hawaii, USA. Proceedings of Machine Learning Research, vol. 202, pp. 28656–28679. PMLR (2023), https://proceedings.mlr.press/v202/rame23a.html

[217] Ray, S.: A quick review of machine learning algorithms. In: 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon). pp. 35–39 (2019), https://doi.org/10.1109/COMITCon.2019.8862451

[218] Reed, S., Zolna, K., Parisotto, E., Colmenarejo, S.G., Novikov, A., Barth-Maron, G., Gimenez, M., Sulsky, Y., Kay, J., Springenberg, J.T., et al.: A generalist agent. arXiv preprint arXiv:2205.06175 (2022)

[219] Rimmel, A., Teytaud, F., Teytaud, O.: Biasing Monte-Carlo simulations through RAVE values. In: International Conference on Computers and Games. pp. 59–68. Springer (2010), https://doi.org/10.1007/978-3-642-17928-0_6

[220] Risi, S., Togelius, J.: Increasing generality in machine learning through procedural content generation. Nature Machine Intelligence **2**(8), 428–436 (2020), https://doi.org/10.1038/s42256-020-0208-z

[221] Romanycia, M.H., Pelletier, F.J.: What is a heuristic? Computational Intelligence **1**(1), 47–58 (1985), https://doi.org/10.1111/j.1467-8640.1985.tb00058.x

[222] Ruijl, B., Vermaseren, J., Plaat, A., Herik, J.v.d.: Combining simulated annealing and Monte Carlo tree search for expression simplification. arXiv preprint arXiv:1312.0841 (2013), https://arxiv.org/abs/1312.0841

[223] Rupp, F., Eberhardinger, M., Eckert, K.: Balancing of competitive two-player game levels with reinforcement learning. In: Conference on Games (COG). IEEE (2023), https://doi.org/10.1109/CoG57401.2023.10333248

[224] Russell, S.J., Norvig, P.: Artificial intelligence: a modern approach, 4th edition. Malaysia; Pearson Education Limited (2020)

[225] Salge, C., Green, M.C., Canaan, R., Togelius, J.: Generative design in minecraft (GDMC): settlement generation competition. In: Conference on the Foundations of Digital Games. ACM (2018)

[226] Sampaio, P., Baffa, A., Feijó, B., Lana, M.: A fast approach for automatic generation of populated maps with seed and difficulty control. In: 16th Brazilian Symposium on Computer Games and Digital Entertainment, SBGames 2017, Curitiba, Brazil, November 2-4, 2017. pp. 10–18. IEEE Computer Society (2017), https://doi.org/10.1109/SBGames.2017.00010

[227] Santiago III, J.M., Parayno, R.L., Deja, J.A., Samson, B.P.V.: Rolling the dice: Imagining generative ai as a dungeons & dragons storytelling companion. arXiv preprint arXiv:2304.01860 (2023)

[228] Sarkar, A., Guzdial, M., Snodgrass, S., Summerville, A., Machado, T., Smith, G.: Procedural content generation via knowledge transformation (pcg-kt). Transaction on Games **16**, 36–50 (2024), https://doi.org/10.1109/TG.2023.3270422

[229] Sarker, A., Gonzalez, G.: A corpus for mining drug-related knowledge from twitter chatter: Language models and their utilities. Data in brief **10**, 122–131 (2017)

[230] Satyanarayan, A., Russell, R., Hoffswell, J., Heer, J.: Reactive vega: A streaming dataflow architecture for declarative interactive visualization. IEEE Transactions on Visualization & Computer Graphics (Proc. InfoVis) (2016), https://doi.org/10.1109/TVCG.2015.2467091

[231] Schaeffer, J., van den Herik, J., Hsu, T.s.: Games, computers and artificial intelligence. Chips Challenging Champions: games, computer and Artificial Intelligence **134** (2002), https://doi.org/10.1016/S0004-3702(01)00165-5

[232] Schrittwieser, J., Antonoglou, I., Hubert, T., Simonyan, K., Sifre, L., Schmitt, S., Guez, A., Lockhart, E., Hassabis, D., Graepel, T., et al.: Mastering atari, go, chess and shogi by planning with a learned model. Nature **588**(7839), 604–609 (2020), https://doi.org/10.1038/s41586-020-03051-4

[233] Schrittwieser, J., Antonoglou, I., Hubert, T., Simonyan, K., Sifre, L., Schmitt, S., Guez, A., Lockhart, E., Hassabis, D., Graepel, T., et al.: Mastering atari, go, chess and shogi by planning with a learned model. Nature **588**(7839), 604–609 (2020), https://doi.org/10.1038/s41586-020-03051-4

[234] Schulman, J., Wolski, F., Dhariwal, P., Radford, A., Klimov, O.: Proximal policy optimization algorithms. arXiv preprint arXiv:1707.06347 (2017), http://arxiv.org/abs/1707.06347

[235] Schwartz, R., Dodge, J., Smith, N.A., Etzioni, O.: Green ai. Communications of the ACM **63**(12), 54–63 (2020)

[236] Serin, K.: The sims rival life by you reveals its open world, dialogue trees, and third person mode, https://www.rockpapershotgun.com/the-sims-rival-life-by-you-reveals-its-open-world-dialogue-trees-and-third-person-mode, visited 28.03.2023

[237] Shah, D., Osinski, B., Ichter, B., Levine, S.: Lm-nav: Robotic navigation with large pre-trained models of language, vision, and action. In: Liu, K., Kulic, D., Ichnowski, J. (eds.) Conference on Robot Learning, CoRL 2022, 14-18 December 2022, Auckland, New Zealand. Proceedings of Machine Learning Research, vol. 205, pp. 492–504. PMLR (2022), https://proceedings.mlr.press/v205/shah23b.html

[238] Shah, S., Dey, D., Lovett, C., Kapoor, A.: Airsim: High-fidelity visual and physical simulation for autonomous vehicles. In: Field and Service Robotics (2017), https://arxiv.org/abs/1705.05065

[239] Shao, K., Tang, Z., Zhu, Y., Li, N., Zhao, D.: A survey of deep reinforcement learning in video games. CoRR **abs/1912.10944** (2019), http://arxiv.org/abs/1912.10944

[240] Shao, K., Zhu, Y., Zhao, D.: Starcraft micromanagement with reinforcement learning and curriculum transfer learning. IEEE Transactions on Emerging Topics in Computational Intelligence **3**(1), 73–84 (2018), https://doi.org/10.1109/TETCI.2018.2823329

[241] Shi, F., Chen, X., Misra, K., Scales, N., Dohan, D., Chi, E., Schärli, N., Zhou, D.: Large language models can be easily distracted by irrelevant context. arXiv preprint arXiv:2302.00093 (2023)

[242] Shu, T., Liu, J., Yannakakis, G.N.: Experience-driven PCG via reinforcement learning: A super mario bros study. In: Conference on Games. IEEE (2021), https://doi.org/10.1109/CoG52621.2021.9619124

[243] Silva, F., Costa, A.: A survey on transfer learning for multiagent reinforcement learning systems. Journal of Artificial Intelligence Research **64**, 645–703 (2019), https://doi.org/10.1613/jair.1.11396

[244] Silva, F.L.d., Costa, A.H.R.: Object-oriented curriculum generation for reinforcement learning. In: Conference on Autonomous Agents and MultiAgent Systems. ACM (2018)

[245] Silva, R., Vasco, M., Melo, F.S., Paiva, A., Veloso, M.: Playing games in the dark: An approach for cross-modality transfer in reinforcement learning. In: Seghrouchni, A.E.F., Sukthankar, G., An, B., Yorke-Smith, N. (eds.) International Conference on Autonomous Agents and Multiagent Systems, AAMAS. International Foundation for Autonomous Agents and Multiagent Systems (2020), https://dl.acm.org/doi/10.5555/3398761.3398907

[246] Silver, D., Huang, A., Maddison, C.J., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., et al.: Mastering the game of Go with deep neural networks and tree search. Nature **529**(7587), 484–489 (2016), https://doi.org/10.1038/nature16961

[247] Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., Lanctot, M., Sifre, L., Kumaran, D., Graepel, T., et al.: A general reinforcement learning algorithm that masters chess, shogi, and go through self-play. Science **362**(6419), 1140–1144 (2018), https://doi.org/10.1126/science.aar6404

[248] Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., Hubert, T., Baker, L., Lai, M., Bolton, A., et al.: Mastering the game of go without human knowledge. Nature **550**(7676), 354–359 (2017), https://doi.org/10.1038/nature24270

[249] Singh, B., Kumar, R., Singh, V.P.: Reinforcement learning in robotic applications: a comprehensive survey. Artif. Intell. Rev. **55**(2), 945–990 (2022). doi: 10.1007/S10462-021-09997-9, https://doi.org/10.1007/s10462-021-09997-9

[250] Sinha, A., Shen, Z., Song, Y., Ma, H., Eide, D., Hsu, B.J.P., Wang, K.: An overview of microsoft academic service (mas) and applications. In: Proceedings of the 24th International Conference on World Wide Web. p. 243–246. WWW '15 Companion, Association for Computing Machinery, New York, NY, USA (2015). doi: 10.1145/2740908.2742839, https://doi.org/10.1145/2740908.2742839

[251] Smith, G.: Understanding procedural content generation: a design-centric analysis of the role of PCG in games. In: Jones, M., Palanque, P.A., Schmidt, A., Grossman, T. (eds.) CHI Conference on Human Factors in Computing Systems, CHI'14, Toronto, ON, Canada - April 26 - May 01, 2014. pp. 917–926. ACM (2014). doi: 10.1145/2556288.2557341, https://doi.org/10.1145/2556288.2557341

[252] Soni, V., Singh, S.: Using homomorphisms to transfer options across continuous reinforcement learning domains. In: AAAI Conference on Artificial Intelligence. vol. 6, pp. 494–499 (2006), http://www.aaai.org/Library/AAAI/2006/aaai06-079.php

[253] Sorokin, A.Y., Burtsev, M.S.: Episodic memory transfer for multi-task reinforcement learning. Biologically inspired cognitive architectures **26**, 91–95 (2018), https://doi.org/10.1016/j.bica.2018.09.003

[254] Sprague, T.: On the different possible non-linear arrangements of eight men on a Chess-board. Proceedings of the Edinburgh Mathematical Society **8**, 30–43 (1889), https://doi.org/10.1017/S0013091500030522

[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

[256] Stanley, K.O., D'Ambrosio, D.B., Gauci, J.: A hypercube-based encoding for evolving large-scale neural networks. Artificial life **15**(2), 185–212 (2009), https://doi.org/10.1162/artl.2009.15.2.15202

[257] van Stegeren, J., Myśliwiec, J.: Fine-tuning gpt-2 on annotated rpg quests for npc dialogue generation. In: Proceedings of the 16th International Conference on the Foundations of Digital Games. pp. 1–8 (2021)

[258] Strathern, M.: 'improving ratings': audit in the british university system. European Review **5**(3), 305–321 (1997)

[259] Strong, C., Mateas, M.: Talking with npcs: Towards dynamic generation of discourse structures. In: Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment. vol. 4, pp. 114–119 (2008)

[260] Strubell, E., Ganesh, A., McCallum, A.: Energy and policy considerations for deep learning in nlp. In: Association for Computational Linguistics (ACL). pp. 3645–3650 (2019), https://doi.org/10.18653/v1/p19-1355

[261] Sudhakaran, S., González-Duque, M., Freiberger, M., Glanois, C., Najarro, E., Risi, S.: Mariogpt: Open-ended text2level generation through large language models. In: Neural Information Processing Systems (NIPS). Neurips (2023), https://openreview.net/forum?id=aa8KsqfTPa

[262] Sutton, R.S.: Dyna, an integrated architecture for learning, planning, and reacting. ACM Sigart Bulletin **2**(4), 160–163 (1991), https://doi.org/10.1145/122344.122377

[263] Sutton, R.S., Barto, A.G.: Reinforcement learning: An introduction. Second Edition. MIT press (2018)

[264] Taori, R., Gulrajani, I., Zhang, T., Dubois, Y., Li, X., Guestrin, C., Liang, P., Hashimoto, T.B.: Stanford alpaca: An instruction-following llama model. https://github.com/tatsu-lab/stanford_alpaca (2023)

[265] Taylor, M.E., Stone, P.: Transfer learning for reinforcement learning domains: A survey. Journal of Machine Learning Research **10**(7), 1633–1685 (2009), https://dl.acm.org/doi/10.5555/1577069.1755839

[266] Taylor, M.E., Stone, P.: An introduction to intertask transfer for reinforcement learning. AI Magazine **32**(1), 15–15 (2011), https://doi.org/10.1609/aimag.v32i1.2329

[267] Tesauro, G.: Neurogammon: A neural network backgammon learning program. Heuristic Programming in Artificial Intelligence: The First Computer Olympiad, Chichester, England pp. 33–39 (1990), https://doi.org/10.1109/IJCNN.1990.137821

[268] Thrun, S., Schwartz, A.: Finding structure in reinforcement learning. Advances in neural information processing systems (NeurIPS) pp. 385–392 (1994), https://proceedings.neurips.cc/paper_files/paper/1994/hash/7ce3284b743aefde80ffd9aec500e085-Abstract.html

[269] Todd, G., Earle, S., Nasir, M.U., Green, M.C., Togelius, J.: Level generation through large language models. In: Foundations of Digital Games. ACM (2023), https://doi.org/10.1145/3582437.3587211

[270] Todorov, E., Erez, T., Tassa, Y.: Mujoco: A physics engine for model-based control. In: 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems. pp. 5026–5033. IEEE (2012), https://doi.org/10.1109/IROS.2012.6386109

[271] Togelius, J., Yannakakis, G.N., Stanley, K.O., Browne, C.: Search-based procedural content generation: A taxonomy and survey. Transactions on Computational Intelligence and AI in Games **3** (2011)

[272] Torrado, R.R., Bontrager, P., Togelius, J., Liu, J., Pérez-Liébana, D.: Deep reinforcement learning for general video game AI. In: 2018 IEEE Conference on Computational Intelligence and Games, CIG 2018, Maastricht, The Netherlands, August 14-17, 2018. pp. 1–8. IEEE (2018). doi: 10.1109/CIG.2018.8490422, https://doi.org/10.1109/CIG.2018.8490422

[273] Torrey, L., Shavlik, J., Walker, T., Maclin, R.: Skill acquisition via transfer learning and advice taking. In: European Conference on Machine Learning. Lecture Notes in Computer Science, vol. 4212, pp. 425–436. Springer (2006), https://doi.org/10.1007/11871842_41

[274] Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F., Rodriguez, A., Joulin, A., Grave, E., Lample, G.: Llama: Open and efficient foundation language models (2023)

[275] Traoré, R., Caselles-Dupré, H., Lesort, T., Sun, T., Díaz-Rodríguez, N., Filliat, D.: Continual reinforcement learning deployed in real-life using policy distillation and sim2real transfer (2019), http://arxiv.org/abs/1906.04452

[276] Valve: Steam hardware and software survey february 2023, https://store.steampowered.com/hwsurvey, visited 28.03.2023

[277] Van Rijn, J.N., Bischl, B., Torgo, L., Gao, B., Umaashankar, V., Fischer, S., Winter, P., Wiswedel, B., Berthold, M.R., Vanschoren, J.: Openml: A collaborative science platform. In: Joint european conference on machine learning and knowledge discovery in databases. Lecture Notes in Computer Science, vol. 8190, pp. 645–649. Springer (2013), https://doi.org/10.1007/978-3-642-40994-3_46

[278] Van Rijswijck, J.: Search and evaluation in Hex. University of Alberta (2002), https://web.archive.org/web/20210623031417/https://www.cs.cornell.edu/~adith/docs/y_hex.pdf

[279] Vinyals, O., Babuschkin, I., Czarnecki, W.M., Mathieu, M., Dudzik, A., Chung, J., Choi, D.H., Powell, R., Ewalds, T., Georgiev, P., Oh, J., Horgan, D., Kroiss, M., Danihelka, I., Huang, A., Sifre, L., Cai, T., Agapiou, J.P., Jaderberg, M., Vezhnevets, A.S., Leblond, R., Pohlen, T., Dalibard, V., Budden, D., Sulsky, Y., Molloy, J., Paine, T.L., Gülçehre, Ç., Wang, Z., Pfaff, T., Wu, Y., Ring, R., Yogatama, D., Wünsch, D., McKinney, K., Smith, O., Schaul, T., Lillicrap, T.P., Kavukcuoglu, K., Hassabis, D., Apps, C., Silver, D.: Grandmaster level in StarCraft II using multi-agent reinforcement learning. Nature **575**(7782), 350–354 (2019), https://doi.org/10.1038/s41586-019-1724-z

[280] Vithayathil Varghese, N., Mahmoud, Q.H.: A survey of multi-task deep reinforcement learning. Electronics **9**(9), 1363 (2020), https://doi.org/10.3390/electronics9091363

[281] Volz, V., Naujoks, B.: Towards game-playing ai benchmarks via performance reporting standards. IEEE Conference on Games (CoG) pp. 764–771 (2020), https://doi.org/10.1109/CoG47356.2020.9231705

[282] Volz, V., Schrum, J., Liu, J., Lucas, S.M., Smith, A.M., Risi, S.: Evolving mario levels in the latent space of a deep convolutional generative adversarial network. In: Genetic and Evolutionary Computation Conference. ACM (2018)

[283] Wang, J.X., King, M., Porcel, N., Kurth-Nelson, Z., Zhu, T., Deck, C., Choy, P., Cassin, M., Reynolds, M., Song, F., Buttimore, G., Reichert, D.P., Rabinowitz, N., Matthey, L., Hassabis, D., Lerchner, A., Botvinick, M.: Alchemy: A structured task distribution for meta-reinforcement learning (2021), https://arxiv.org/abs/2102.02926

[284] Wang, N., Zhang, Y., Li, Z., Fu, Y., Yu, H., Liu, W., Xue, X., Jiang, Y.: Pixel2mesh: 3d mesh model generation via image guided deformation. IEEE

Trans. Pattern Anal. Mach. Intell. **43**(10), 3600–3613 (2021). doi: 10.1109/TPAMI.2020.2984232, https://doi.org/10.1109/TPAMI.2020.2984232

[285] Wang, T.T., Gleave, A., Tseng, T., Pelrine, K., Belrose, N., Miller, J., Dennis, M.D., Duan, Y., Pogrebniak, V., Levine, S., Russell, S.: Adversarial policies beat superhuman go ais. In: Krause, A., Brunskill, E., Cho, K., Engelhardt, B., Sabato, S., Scarlett, J. (eds.) International Conference on Machine Learning (ICML). Proceedings of Machine Learning Research, vol. 202, pp. 35655–35739. PMLR (2023), https://proceedings.mlr.press/v202/wang23g.html

[286] Weiss, K., Khoshgoftaar, T.M., Wang, D.: A survey of transfer learning. Journal of Big data **3** (2016), https://doi.org/10.1186/s40537-016-0043-6

[287] Weng, R.C., Lin, C.J.: A Bayesian Approximation Method for Online Ranking. Journal of Machine Learning Research **12**(Jan), 267–300 (2011), https://dl.acm.org/doi/10.5555/1953048.1953057

[288] Werning, M.: Predicting the past from minimal traces: Episodic memory and its distinction from imagination and preservation. Review of philosophy and psychology **11**(2), 301–333 (2020)

[289] Wikipedia: First-move advantage in chess (2019-08-18), https://web.archive.org/web/20190807121829/https://en.wikipedia.org/wiki/First-move_advantage_in_chess

[290] Wikipedia: Tetromino (2019-09-09), https://en.wikipedia.org/wiki/Tetromino

[291] Xenopoulos, P., Doraiswamy, H., Silva, C.: Valuing player actions in counter-strike: Global offensive. In: IEEE International Conference on Big Data (Big Data). pp. 1283–1292. IEEE (2020)

[292] Xiao, I.: A distributed reinforcement learning solution with knowledge transfer capability for a bike rebalancing problem (2018), http://arxiv.org/abs/1810.04058

[293] Xie, S., Galashov, A., Liu, S., Hou, S., Pascanu, R., Heess, N., Teh, Y.W.: Transferring task goals via hierarchical reinforcement learning (2018), https://openreview.net/forum?id=S1Y6TtJvG

[294] Yang, J., Petersen, B., Zha, H., Faissol, D.: Single episode policy transfer in reinforcement learning (2020), https://openreview.net/forum?id=rJeQoCNYDS

[295] Yang, T., Hao, J., Meng, Z., Zhang, Z., Hu, Y., Cheng, Y., Fan, C., Wang, W., Liu, W., Wang, Z., et al.: Efficient deep reinforcement learning via adaptive policy transfer pp. 3094–3100 (2020), https://doi.org/10.24963/ijcai.2020/428

[296] Yannakakis, G.N.: AI in Computer Games: Generating Interesting Interactive Opponents by the use of Evolutionary Computation. Ph.D. thesis, University of Edinburgh (2005), `http://hdl.handle.net/1842/879`

[297] Yin, H., Pan, S.: Knowledge transfer for deep reinforcement learning with hierarchical experience replay. In: AAAI Conference on Artificial Intelligence. vol. 31, pp. 1640–1646 (2017), `https://doi.org/10.1609/aaai.v31i1.10733`

[298] Yoon, I., Anwar, M.A., Joshi, R.V., Rakshit, T., Raychowdhury, A.: Hierarchical memory system with stt-mram and sram to support transfer and real-time reinforcement learning in autonomous drones. IEEE Journal on Emerging and Selected Topics in Circuits and Systems **9**(3), 485–497 (2019), `https://doi.org/10.1109/JETCAS.2019.2932285`

[299] Yu, T., Quillen, D., He, Z., Julian, R., Hausman, K., Finn, C., Levine, S.: Meta-world: A benchmark and evaluation for multi-task and meta reinforcement learning. In: Conference on Robot Learning. Proceedings of Machine Learning Research, vol. 100, pp. 1094–1100. PMLR (2020), `http://proceedings.mlr.press/v100/yu20a.html`

[300] Zakaria, Y., Fayek, M., Hadhoud, M.: Start small: Training controllable game level generators without training data by learning at multiple sizes. Alexandria Engineering Journal **72**, 479–494 (2023), `https://doi.org/10.1016/j.aej.2023.04.019`

[301] Zakaria, Y., Fayek, M.B., Hadhoud, M.: Procedural level generation for sokoban via deep learning: An experimental study. Transaction on Games **15**, 108–120 (2023), `https://doi.org/10.1109/TG.2022.3175795`

[302] Zhou, W., Zhang, S., Poon, H., Chen, M.: Context-faithful prompting for large language models. arXiv preprint arXiv:2303.11315 (2023)

[303] Zhu, X.: Behavior tree design of intelligent behavior of non-player character (npc) based on unity3d. Journal of Intelligent & Fuzzy Systems **37**(5), 6071–6079 (2019), `https://doi.org/10.3233/JIFS-179190`

[304] Zhu, Z., Lin, K., Zhou, J.: Transfer learning in deep reinforcement learning: A survey (2020), `https://arxiv.org/abs/2009.07888`

[305] Zittrain, J., Albert, K., Lessig, L.: Perma: Scoping and addressing the problem of link and reference rot in legal citations. Legal Information Management (LIM) **14**, 88 (2014)