Bridging AI and Cultural Heritage: Outcomes from the GameTable WG1 London Meeting

Dennis J.N.J. Soemers [a,*], Jakub Kowalski [b], Walter Crist [c], Summer Courts [d], Tim Penn [e] and Éric Piette [f]

[a] *Department of Advanced Computing Sciences, Maastricht University, the Netherlands*
E-mail: dennis.soemers@maastrichtuniversity.nl
[b] *Faculty of Mathematics and Computer Science, University of Wrocław, Poland*
E-mail: jko@cs.uni.wroc.pl
[c] *Centre for the Arts in Society, Leiden University, the Netherlands*
E-mail: wcrist@asu.edu
[d] *Department of Archaeology, University of Reading, United Kingdom*
E-mail: s.l.courts@pgr.reading.ac.uk
[e] *Department of Classics, University of Reading, United Kingdom*
E-mail: tim.penn@reading.ac.uk
[f] *ICTEAM, UCLouvain, Belgium*
E-mail: eric.piette@uclouvain.be

Abstract. On January 20 and 21, 2025, the GameTable COST Action's Working Group 1 convened in London for a meeting focused on game-playing AI, including search algorithms, knowledge representation, and reinforcement learning. The first day featured participant presentations, while the second day was dedicated to plenary and small group discussions. A significant outcome was the identified need for accessible resources to bridge AI and cultural heritage research. This report provides a summary of the discussions and talks that took place during the meeting.

Keywords: GameTable, Working Group Meeting, Tabletop Games, Artificial Intelligence

1. INTRODUCTION

On January 20 and 21, 2025, we held a Working Group (WG) 1 meeting for the *GameTable* COST Action CA22145 (Piette et al., 2024), in London. GameTable is an interdisciplinary network of researchers, studying tabletop games from perspectives such as artificial intelligence (AI), cultural heritage, archaeology, mathematics, and education. Within this network, the focus of WG1 is on AI techniques for automated playing of games (Soemers et al., 2024), such as search algorithms (Russell and Norvig, 2020) and reinforcement learning (Sutton and Barto, 2018). A total of 26 different people participated in the meeting, of which 20 in-person, with the remaining six joining remotely.

The first day of the meeting was primarily used for presentations contributed by some of the participants. These are summarised in Section 2. On the second day, our focus shifted to plenary discussions, as well as discussions in smaller groups. Some of the key topics of discussion are summarised in Section 3. Section 4 elaborates in further detail on one of the key conclusions from the meeting: the need

*Corresponding author. E-mail: dennis.soemers@maastrichtuniversity.nl

for easily accessible information that can help AI and cultural heritage researchers better understand and communicate with each other. Finally, Section 5 concludes this paper.

2. CONTRIBUTED TALKS

2.1. Imperfect Information games

The first session of the meeting focused on the challenges and advancements in imperfect-information games, a critical area in game AI research. This session provided a platform to discuss the complexities associated with reasoning under uncertainty and the development of general strategies for such games.

The first talk, titled "Belief Stochastic Game: A Model for Imperfect-Information Games with Known Position," was presented by Achille Morenville (Morenville and Piette, 2024a,b). Imperfect-information games present significant challenges for General Game Playing (GGP) (Genesereth et al., 2005) agents, as conventional models, such as Extensive Form Games (EFG) (von Neumann and Morgenstern, 1944), and more recent ones like Factored-Observation Stochastic Games (FOSG) (Kovařík et al., 2022), require agents to construct and maintain estimates of the game state. This often results in game-specific solutions and the unintended incorporation of domain-specific knowledge, limiting the generalizability of such approaches. To overcome these limitations, Achille introduced the Belief Stochastic Game model, a novel framework that externalizes state estimation by shifting it from the agent to the game model itself. This allows agents to concentrate exclusively on strategy development rather than on complex state inference. By exploiting common structural patterns in many imperfect-information games, this approach enhances the adaptability of AI agents, enabling them to generalize more effectively across diverse game environments.

The second talk, titled "Finding Portfolios of Opponent Strategies in Large Imperfect Information Games," was presented by Karolina Drabent (Drabent et al., 2024). She explored the use of strategy portfolios to improve decision-making in complex games where computing Nash Equilibrium is infeasible. Instead of considering all possible strategies, agents can construct portfolios of opponent strategies to optimize their own play, either by minimizing worst-case losses (strategy optimization) or exploiting opponents (opponent exploitation). By restricting the game to a selected portfolio, computational efficiency is improved while maintaining strategic depth. She introduced methods like Gradient Clustering Transformations (GCT) and Regularized Nash Dynamics (RNaD) to refine portfolio selection, ensuring adaptability across large game spaces. Experimental results demonstrated the effectiveness of these approaches, but open questions remain on computing optimal mixed pessimistic portfolios. The findings highlight portfolio-based optimization as a promising alternative to exhaustive game-solving techniques.

The last talk of this session, titled "Adapting to opponents in large imperfect information games," was presented by David Milec (Milec et al., 2024, 2025). He explored methods for improving AI adaptability in complex game environments where computing an exact best response is infeasible. He discussed depth-limited approaches using value function approximation and heuristics like DeepStack (Moravčík et al., 2017) to estimate game values beyond computational limits. To enhance robustness, he introduced Worst-Case ModelMix, which blends opponent modeling with worst-case planning to mitigate errors in strategy adaptation. The talk also examined agent evaluation techniques, comparing traditional head-to-head testing with exploitability metrics, which provide a more general measure of an AI's adaptability. The talk concluded with open questions on adaptation in general-sum games and evaluating AI performance in Stratego and Dark Chess, highlighting ongoing challenges in imperfect-information game AI.

2.2. How humans and AI experience game playing

The second session was dedicated to how humans and AI experience game playing. GameTable members presented studies and projects on how AI models learn from experience, adapt to new challenges, and sometimes even exhibit behavior that mimics human intuition. Comparisons were drawn between human problem-solving approaches and the pattern recognition abilities of AI, highlighting both the strengths and limitations of each.

The first talk on this topic was titled “Quantifying tabletop games with AI - can we transfer anything to human experience?” and was given by James Goodman (Goodman et al., 2021a,b, 2024, 2025). This presentation focused on quantifying tabletop games using AI-driven metrics to analyze game characteristics such as difficulty, randomness, and skill depth. His research aims to develop game fingerprints, which are distributions of optimized AI parameters that help categorize games based on computational play. Using Monte Carlo Tree Search (MCTS) (Browne et al., 2012; Świechowski et al., 2022) optimization, he explored how games can be mapped into a structured landscape, allowing for comparative analysis across different titles. The talk covered several key topics, including game landscapes, measuring game difficulty using skill traces derived from AI performance, and assessing randomness by analyzing game outcomes under controlled conditions. While some findings aligned with human intuition—such as hidden information affecting strategy in Love Letter—others, like the reported high skill depth of Sushi Go, raised questions about the transferability of AI-based quantifications to human gameplay. This talk concluded with reflections on agent limitations, highlighting the need for human validation to ensure AI-generated metrics truly reflect player experience.

The second and last talk on this topic was titled “Winning is not everything - Towards human-like agents for tabletop games” and was given by Aloïs Rautureau. He explored the development of human-like AI agents for tabletop games, emphasizing that winning is not the only goal in game-playing AI. While traditional GGP agents optimize for victory using techniques like MCTS, this approach makes them unrealistic opponents for human players. The talk examined how human-likeness can be defined and measured, incorporating insights from cognitive science, psychology, and AI research in other fields (e.g., chatbots, NPC behavior in video games). A two-system thinking model was introduced, where System 1 represents intuitive pattern recognition, and System 2 involves deeper analytical reasoning. A proposed framework for human-like agents in GGP integrates these systems by filtering out intuitively bad moves and using MCTS only when needed. Initial implementations for Renju demonstrated promising results, with ongoing work aiming to refine the model using inverse reinforcement learning (IRL) (Russell, 1998; Ng and Russell, 2000) to infer human motivations. The talk concluded with future research directions, including integrating human-like AI into Ludii (Piette et al., 2020), exploring its role in ancient game reconstruction (e.g. (Browne et al., 2019, 2022; Crist et al., 2024)), and investigating AI models for cheating behavior in games.

2.3. Generalisation and Explainability

The last talk session was dedicated to Generalisation and Explainability. The session also discussed recent breakthroughs and ongoing debates on how best to balance model complexity with interpretability.

The first presentation, titled "Explainability of Board Game Agents," was delivered by Manuel Eberhardinger and focused on the findings of a Short-Term Scientific Mission (STSM) conducted during the first grant period of the GameTable COST Action. The talk focused on improving the explainability of AI agents in board games, particularly by using decision trees and genetic programming to generate human-interpretable explanations of AI decisions. The motivation behind this work is to address the black-box nature of game-playing AI, such as MCTS and reinforcement learning agents, which makes understanding their strategic choices difficult. The research aims to extract state-action features (e.g. (Soemers et al., 2023a,b)) without relying on expert policies or neural network logits, allowing for a more transparent decision-making process. The proposed method involves genetic programming to discover board game features and training decision trees that predict AI actions based on these features. Initial evaluations, using AlphaZero-trained agents, demonstrated that while decision trees can approximate AI strategies, they often fail to generalize correctly, leading to brittle decision-making. Future directions include testing the framework on simpler games like Tic-Tac-Toe, refining the feature selection process, and developing a learnable domain-specific language (DSL) to improve the explainability and robustness of AI-driven board game strategies.

The final talk, titled "Games and out-of-distribution generalisation" was given by Spyridon Samothrakis (Samothrakis et al., 2024; Soemers et al., 2025). It explored out-of-distribution (OOD) generalization in AI and game-playing agents, questioning whether current AI approaches can truly generalize beyond their training environments. The talk reviewed the evolution of game AI from rule-based systems to deep reinforcement learning, highlighting the "Bitter Lesson"—that AI progress tends to come from scaling computation rather than hand-crafted knowledge (Sutton, 2019). However, despite advances like AlphaZero, current AI models still struggle with sample efficiency and OOD generalization, requiring vast amounts of training data to adapt to unseen scenarios. The presentation introduced key OOD challenges, including systematicity, productivity, and substitutivity, which relate to how AI recognizes and applies learned patterns in novel situations. Samothrakis argued that procedural content generation (PCG) alone is insufficient for true generalization, and instead, new approaches combining neural networks with symbolic reasoning might be needed. He concluded with open research questions on integrating deep learning with first-order logic and developing more efficient algorithms that can generalize across a wide range of games and real-world tasks.

3. DISCUSSIONS

In addition to contributed talks, we reserved a substantial amount of time for more open-ended discussions. One of the main topics of discussion was how best to facilitate further communication and collaboration between researchers studying games from the AI perspective on the one hand, and the cultural heritage perspective on the other hand. We dedicate an entire separate section—Section 4—to this topic. The following other topics emerged as key topics for further consideration in the research community:

- Human-like AI: how can we implement AI algorithms that play tabletop games like humans do, such that any measures we collect from simulations accurately estimate the experience that humans would have had playing that ruleset? How can we make AI follow not only the explicit rules defined in games' rulesets—and maybe deliberately *not* follow them in plausible ways—but also follow social etiquette rules (e.g., avoid moving back and forth indefinitely). Another interesting factor is the thinking time used when playing: depending on culture and social context, certain amounts of thinking time may or may not be considered socially acceptable (or even allowed by tournament rules), but humans and AI players tend to be affected by time in different ways.

- Explainability in search and RL for game playing, and AI systems that can give advice or recommendations as to how to play to humans.
- How can we implement effective frameworks and benchmarks to facilitate research in the combination of general game playing with imperfect-information games?
- Which tabletop games, if any, remain as major challenges where AI cannot yet reach superhuman levels of playing strength?
- How can we improve methodologies used for benchmarking and evaluating different AI algorithms for game playing?
- Development of AI that can effectively collaborate with humans or other AI agents in collaborative games.
- What role can games play in benchmarking for Artificial General Intelligence?
- How can we use AI and games in education?
- How can we effectively share teaching materials, and in general collaborate in educational activities across universities, within the game AI research community?

4. COMBINING AI AND CULTURAL HERITAGE FOR TABLETOP GAMES RESEARCH

One of GameTable's overarching aims is to bring together experts in AI and the cultural heritage of games to identify and test new methodologies for approaching past ludic activity (Piette et al., 2024). To further this goal, four members (Walter Crist, Summer Courts, Tim Penn, and Ilaria Truzzi) of Working Group 2—"Cultural Heritage of Games"—attended the meeting to identify and discuss viable avenues for future research in this area. A key theme to emerge from conversations between the WG1 and WG2 members at this meeting is that experts in AI and experts in the cultural heritage of games work within highly divergent research traditions and frameworks. This divergence underscores the need for open dialogue to foster meaningful, collaborative research. Given that the application of AI to historical games remains a nascent field, participants agreed that developing well-defined case studies would be an effective capacity-building strategy to bridge these disciplinary gaps.

Very few concrete case studies that apply AI to answer questions about historical games have been published so far (Donkers et al., 2000; Browne, 2023; Crist et al., 2024). For the field to advance, continued collaboration between WG1 and WG2 members is essential in developing viable AI-based approaches to studying past games. This requires formulating specific research questions grounded in the distinct characteristics of specific traditional games. A key challenge is determining which metrics of traditional games can be reasonably calculated using AI and what types of research questions these methods can address to create useful new insights for scholars working on historical games. One promising area identified during the meeting is games that rely on chance—particularly those involving randomization devices such as dice or knucklebones (astragals). During discussions in the meeting, participants identified several avenues for future research, perhaps to be explored as part of a journal special issue on AI and historical games. **Possible case studies which might be meaningful for experts working on the cultural heritage of games** include:

- **Biased dice**: Ancient and historical dice were often asymmetrical, making certain outcomes more likely than others (Swift, 2017). How do AI playouts of ancient games that used such dice reveal the impact of these biased dice on gameplay?
- **Changing dice numbers or types**: The dice used to play some games changed over time or differ across space depending on the culture in which they are played. In some cases, the number of dice change, in others, the type of dice change (e.g., from binary to cubic dice). What do AI playouts suggest that this change would have had on play?

- **Changing board geometry**: The number or arrangement of lines or spaces of game boards have changed dramatically through time. This can affect the length and sequence of spaces in the track along which pieces move, as well as shrinking, expanding, or reconfiguring the field of spaces presented for games played on, for example, grids. What do AI playouts suggest that these changes would have had on play? Can these boards even be considered to have been used to play the same game?

Examples of techniques and methodologies from the field of AI research that may be applicable to such research questions include:

- Given a hypothesised ruleset for a game, AI-driven players can be used to simulate play at a large scale (e.g., running hundreds or thousands or more of simulated plays) and perform quantitative analysis at a level that would be infeasible to do with human playtesting. Essentially any quantity of interest that can be given a clear mathematical definition can be measured and analysed from such simulations. Examples include:
  - ∗ Duration (in number of moves or turn) per game. If games consistently and easily end in an extremely short time (e.g., the first player can win immediately), the evaluated ruleset is not plausible (Browne, 2023).
  - ∗ Various estimators of the "quality" of a game can be measured, following the intuition that rulesets that people enjoy playing are more likely to have been played than low-quality rulesets (Browne, 2018; Browne et al., 2019; Crist et al., 2024). Duration could again be one factor in this, if we assume that games are considered better or more fun if they take neither too long nor too short to complete. However, care should be taken to account for differences between cultures and social contexts within games were played, as these can affect how much time is considered too short or long. Other factors could include balance (does each player have a fair chance at winning), skill depth (is there room for different levels of skill expression) (Browne, 2022; Goodman et al., 2024), and more (Browne, 2009; Kowalski and Szykuła, 2016).
  - ∗ Usage of game equipment. If certain parts of a game's equipment (e.g., certain pieces or certain parts of the board) see substantially more use than others in simulated play, this could be correlated to signs of usage visible in the archaeological material.
- The impact of using different sources of randomness (e.g., different types of dice, as mentioned previously as a possible case study from cultural heritage research) on game outcomes and aspects such as the room for skill expression may be estimated from AI-driven playtesting (Goodman et al., 2025).
- Measures as described in the previous two points may be used to generate plausible explanations as to why certain changes in rules between closely related games may have been introduced. Differences in rules between games can be correlated to measures of quality, measures of balance between randomness and skill, and so on.
- When our knowledge of rulesets of an ancient game is incomplete (i.e., we know some parts of the rules, but not the complete rules), we may attempt to automatically fill in the missing parts. This may, for instance, be done by copying relevant rules from games that are closely related in terms of, e.g., cultures or social contexts in which the games were played. Such a process could procedurally generate a wide variety of hypothesised rulesets, each of which in turn could be evaluated for plausibility as described previously. If using solely AI-driven evaluations by themselves is not sufficiently reliable, an alternative approach can be to use a combination of AI-driven and human play-testing. AI-driven playtesting can filter a wide variety of hypothesised rulesets down to a smaller set, which may then be further tested by human players. This may require the development

of extensive tools that enable convenient, extensive playtesting of arbitrary (potentially procedurally generated) tabletop games, ideally in an online interface.
- Given an exhaustive database containing detailed information about games played throughout history, including extensive data on what is known about their rules, the social contexts in which they were placed, and any information concerning the geographical locations and periods of time in which they were played that may be derived from archaeological evidence (e.g., (Crist et al., 2022)), data science techniques may be used to generate plausible ways of filling in gaps. However, the sparsity of existing data remains a concern for this idea.

5. CONCLUSION

This paper summarised the key topics of discussion and outcomes from the January 2025 meeting of Working Group 1 of the GameTable COST Action (Piette et al., 2024). There were numerous discussions and talks surrounding various aspects of AI research for tabletop games, the use of tabletop games for AI research, and how to drive the field forwards. In this report, we placed particular emphasis on the matter of how best to facilitate interdisciplinary research between AI researchers and cultural heritage researchers. We identified a need to provide examples of (1) research questions that are of interest to researchers studying games from a cultural heritage perspective, and (2) techniques and methods that AI researchers could contribute to help answer such research questions. Such lists of examples should help researchers across the different disciplines to more easily and effectively communicate with each other. A first attempt at fulfulling this need is included in this report.

ACKNOWLEDGEMENTS

This article is based upon work from COST Action CA22145 - GameTable, supported by COST (European Cooperation in Science and Technology). We thank all the participants who attended and contributed to our meeting.

REFERENCES

Browne, C. (2022). Quickly Detecting Skill Trace in Games. In *Proceedings of the 2022 IEEE Conference on Games* (pp. 604–607).

Browne, C. (2023). Which Rules for Mu Torere? In *Computers and Games 2022*. To appear.

Browne, C.B. (2009). Automatic Generation and Evaluation of Recombination Games. PhD thesis, Faculty of Information Technology, Queensland University of Technology, Queensland, Australia.

Browne, C., Powley, E., Whitehouse, D., Lucas, S., Cowling, P.I., Rohlfshagen, P., Tavener, S., Perez, D., Samothrakis, S. & Colton, S. (2012). A Survey of Monte Carlo Tree Search Methods. *IEEE Transactions on Computational Intelligence and AI in Games*, *4*(1), 1–49.

Browne, C., Soemers, D.J.N.J., Piette, É., Stephenson, M., Conrad, M., Crist, W., Depaulis, T., Duggan, E., Horn, F., Kelk, S., Lucas, S.M., Neto, J.P., Parlett, D., Saffidine, A., Schädler, U., Silva, J.N., de Voogt, A. & Winands, M.H.M. (2019). Foundations of Digital Archæoludology. Technical report, Schloss Dagstuhl Research Meeting, Germany.

Browne, C., Piette, É., Crist, W., Stephenson, M. & Soemers, D.J.N.J. (2022). Report on the 2nd Digital Ludeme Project Workshop. *ICGA Journal*, *44*(2), 56–66.

Browne, C. (2018). Modern Techniques for Ancient Games. In *IEEE Conference on Computational Intelligence and Games* (pp. 490–497). Maastricht: IEEE Press.

Crist, C., Stephenson, M. & Browne, C. (2022). Ludii Games Database. https://dataverse.nl/dataset.xhtml?persistentId=doi:10.34894/BP8G8U.

Crist, W., Piette, É., Soemers, D.J.N.J., Stephenson, M. & Browne, C. (2024). Computational Approaches for Recognising and Reconstructing Ancient Games: The Case of Ludus Latrunculorum. In A. Pace, T. Penn and U. Schädler (Eds.), *The Archaeology of Play: Material Approaches to Games and Gaming in the Ancient World*. Monographies Instrumentum (pp. 63–80). Dremil-Lafage: Mergoil.

Donkers, J., de Voogt, A. & Uiterwijk, J. (2000). Human versus Machine Problem-Solving: Winning Openings in Dakon. *Internationa Journal for the Study of Board Games*, *3*, 79–88.

Drabent, K., Milec, D., Kubicek, O. & Lisý, V. (2024). In the Search for Optimal Portfolios of Counterstrategies in the Large Imperfect Information Games. In *OPT 2024: Optimization for Machine Learning*.

Genesereth, M.R., Love, N. & Pell, B. (2005). General Game Playing: Overview of the AAAI Competition. *AI Magazine*, *26*(2), 62–72. http://www.aaai.org/ojs/index.php/aimagazine/article/view/1813.

Goodman, J., Perez-Liebana, D. & Lucas, S. (2021a). Fingerprinting Tabletop Games. In *Proceedings of the 2021 IEEE Conference on Games* (pp. 860–863).

Goodman, J., Perez-Liebana, D. & Lucas, S. (2021b). Visualizing Multiplayer Game Spaces. *IEEE Transactions on Games*, *14*(4), 663–675.

Goodman, J., Perez-Liebana, D. & Lucas, S. (2024). Skill Depth in Tabletop Board Games. In *Proceedings of the 2024 IEEE Conference on Games* (pp. 1–8).

Goodman, J., Perez-Liebana, D. & Lucas, S. (2025). Seeding for Success: Skill and Stochasticity in Tabletop Games. *IEEE Transactions on Games*. Accepted.

Kovařík, V., Schmid, M., Burch, N., Bowling, M. & Lisỳ, V. (2022). Rethinking formal models of partially observable multiagent decision making. *Artificial Intelligence*, *303*(103645).

Kowalski, J. & Szykuła, M. (2016). Evolving Chesslike Games Using Relative Algorithm Performance Profiles. In *EvoApplications 2017: Applications of Evolutionary Computation*. LNCS (Vol. 9597, pp. 574–589).

Milec, D., Kovařík, V. & Lisý, V. (2025). Adapting Beyond the Depth Limit: Counter Strategies in Large Imperfect Information Games. https://arxiv.org/abs/2501.10464.

Milec, D., Kubíček, O. & Lisý, V. (2024). Continual Depth-limited Responses for Computing Counter-strategies in Sequential Games. In *Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems*. AAMAS '24 (pp. 2393–2395–). International Foundation for Autonomous Agents and Multiagent Systems.

Moravčík, M., Schmid, M., Burch, N., Lisỳ, V., Morrill, D., Bard, N., Davis, T., Waugh, K., Johanson, M. & Bowling, M. (2017). DeepStack: Expert-level artificial intelligence in Heads-up No-limit Poker. *Science*, *356*(6337), 508–513.

Morenville, A. & Piette, E. (2024a). Belief Stochastic Game: A Model for Imperfect-Information Games with Known Positions. In *Computer and Games (CG)*.

Morenville, A. & Piette, E. (2024b). Vers une Approche Polyvalente pour les Jeux à Information Imparfaite sans Connaissance de Domaine. In *Rencontres des Jeunes Chercheurs en Intelligence Artificielle (RJCIA)* (pp. 44–46). In French.

Ng, A.Y. & Russell, S. (2000). Algorithms for inverse reinforcement learning. In *Proceedings of the Seventeenth International Conference on Machine Learning* (Vol. 1, pp. 663–670).

Piette, É., Soemers, D.J.N.J., Stephenson, M., Sironi, C.F., Winands, M.H.M. & Browne, C. (2020). Ludii – The Ludemic General Game System. In G.D. Giacomo, A. Catala, B. Dilkina, M. Milano, S. Barro, A. Bugarín and J. Lang (Eds.), *Proceedings of the 24th European Conference on Artificial Intelligence (ECAI 2020)*. Frontiers in Artificial Intelligence and Applications (Vol. 325, pp. 411–418). IOS Press.

Piette, É., Crist, W., Soemers, D.J.N.J., Rougetet, L., Courts, S., Penn, T. & Morenville, A. (2024). GameTable COST Action: Kickoff Report. *ICGA Journal*. To appear.

Russell, S. (1998). Learning agents for uncertain environments (extended abstract). In *Proceedings of the Eleventh Annual Conference on Computational Learning Theory*. COLT' 98 (pp. 101–103–). Association for Computing Machinery.

Russell, S. & Norvig, P. (2020). *Artificial Intelligence: A Modern Approach* (4th ed.). Pearson.

Samothrakis, S., Soemers, D.J.N.J. & Machlanski, D. (2024). Games of Knightian Uncertainty as AGI Testbeds. In *Proceedings of the 2024 IEEE Conference on Games* (pp. 1–4). IEEE.

Soemers, D.J.N.J., Samothrakis, S., Piette, É. & Stephenson, M. (2023a). Extracting Tactics Learned from Self-Play in General Games. *Information Sciences*, *624*, 277–298.

Soemers, D.J.N.J., Piette, É., Stephenson, M. & Browne, C. (2023b). Spatial State-Action Features for General Games. *Artificial Intelligence*, *321*.

Soemers, D.J.N.J., Kowalski, J., Piette, É., Morenville, A. & Crist, W. (2024). GameTable Working Group 1 meeting report on search, planning, learning, and explainability. *ICGA Journal*, *46*(1), 28–35.

Soemers, D.J.N.J., Samothrakis, S., Driessens, K. & Winands, M.H.M. (2025). Environment Descriptions for Usability and Generalisation in Reinforcement Learning. In *Proceedings of the 17th International Conference on Agents and Artificial Intelligence* (Vol. 3, pp. 983–992).

Sutton, R. (2019). The Bitter Lesson. http://www.incompleteideas.net/IncIdeas/BitterLesson.html.

Sutton, R.S. & Barto, A.G. (2018). *Reinforcement Learning: An Introduction* (2nd ed.). Cambridge, MA: MIT Press.

Świechowski, M., Godlewski, K., Sawicki, B. & Mańdziuk, J. (2022). Monte Carlo Tree Search: A Review of Recent Modifications and Applications. *Artificial Intelligence Review*, *56*, 2497–2562.

Swift, E. (2017). *Roman artefacts and society: design, behaviour, and experience* (1st ed.). Oxford: Oxford University Press.

von Neumann, J. & Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, NJ, USA.