Doctoral School of Information and Biomedical Technologies Polish Academy of Sciences (TIB PAN)

SUBJECT: Deep Reinforcement Learning and Game-Theoretic Frameworks for Solving Complex Multi-Agent Environments

SUPERVISOR: Ph. D., DSc, Eng. Agnieszka Pregowska, Institute of Fundamental Technological Research, Pawinskiego 5B, 02-106 Warsaw, PL, e-mail: aprego@ippt.pan.pl

Name of the institute in which the topic will be realized: **Institute of Fundamental Technological Research**
Scientific discipline: **Information and communication technology**

PROJECT DESCRIPTION

Decision-making in massive adversarial environments remains a frontier AI challenge. Traditional Deep Reinforcement Learning (DRL) and search algorithms systematically fail in complex games characterized by hidden information, chance events, and enormous state spaces ("fog of war" conditions). In these environments, algorithms lacking rigorous game-theoretic foundations are highly exploitable by rational adversaries.

This project explores state-of-the-art, scalable game-theoretic search and learning frameworks. Based on recent breakthroughs, research will focus on real-time subgame solving without common knowledge, imperfect recall [3], and advanced equilibrium computation (e.g., proper and sequential equilibria) [5]. By integrating recursive belief-based learning with highly efficient optimizers like Regret Matching (RM), this thesis aims to achieve robust policy convergence in large-scale imperfect-information games where traditional methods become computationally intractable.

Research Objectives and Approach

This thesis pushes the boundaries of computational game theory and reinforcement learning through three core objectives:

1. Advanced Subgame Solving and Equilibrium Refinements: Computing a full-game strategy is intractable in massive imperfect-information games. Real-time subgame solving techniques will be researched to dynamically refine precomputed blueprint strategies. Because this must account for hidden states and uncertainty about opponents' past actions, "gadget games" will be utilized. By adapting sequence-form linear programming and Counterfactual Regret Minimization (CFR), convergence to *sequential equilibria* within these gadget games will be studied, guaranteeing robust local strategy refinements that minimize exploitability [2].

2. Search Without Common Knowledge and Imperfect Recall: Standard search algorithms assume a shared understanding of the game state, which collapses under severe fog-of-war constraints or decentralized team dynamics. Search paradigms operating without common knowledge will be investigated for depth-limited, real-time planning. Furthermore, limited communication and hidden states will be modeled as *imperfect-recall decision problems*. Frameworks will be explored to maintain strategic cohesion even when agents lack access to previously held information [3].

3. Scalable Regret Matching and Belief-Based Learning: Traditional first-order optimizers (like gradient descent) frequently stagnate in massive game trees. The Regret Matching (RM) family of algorithms, which have demonstrated orders-of-magnitude improvements in solving large zero-sum games, will be further analyzed [4]. Combined with self-play belief learning, expanding the "state" to include probabilistic beliefs about hidden information, this approach will be studied to efficiently navigate massive decision-tree branching.

Simulation Environments & Synthetic Data

The proposed algorithms will be rigorously evaluated against highly complex, asymmetric, and stochastic imperfect-information games. Evaluation metrics will move beyond simple win-rates to focus on theoretical soundness, specifically measuring the minimization of exploitability (Nash/KKT gaps) and robustness in the face of worst-case adversarial behavior. To overcome computational bottlenecks, the project will leverage accelerator-oriented engines and massive vectorization, ensuring rapid iteration and scalable, real-time decision-making in environments with enormous state-action spaces.

BIBLIOGRAPHY

[1] Z. Rudnicka, J. Szczepanski, A. Pregowska, "Impact of Neuron Models on Spiking Neural Network Performance," *Neuroinformatics*, 24(1), 2026.
[2] O. Kubíček et al., "Equilibrium Refinements Improve Subgame Solving," *arXiv:2601.17131* (2026).
[3] E. Tewolde et al., "Decision Making under Imperfect Recall," *arXiv:2602.15252* (2026).
[4] I. Anagnostides et al., "Convergence of Regret Matching in Potential Games," *arXiv:2510.17067* (2025).
[5] B. H. Zhang et al., "The Complexity of Proper Equilibrium," *arXiv:2602.10096* (2026).