

RainbowZero: Combining Advancements in Search and Reinforcement Learning for Complex Environments

*Jonathan Lamontagne Kratz*¹

¹Department of Computer Science, McGill University, Montreal, QC, Canada

Keywords: Reinforcement learning, Monte Carlo Tree Search, MuZero, Stochastic optimization, Artificial intelligence

Corresponding author email: jonathan.lamontagnekratz@mail.mcgill.ca

Published: 05/05/2026

<https://doi.org/10.26443/msurj.v21i2.428>

While deep reinforcement learning has achieved superhuman performance in fully observable, zero-sum games, environments characterized by high stochasticity and complex state spaces remain a significant challenge. This ongoing research introduces "Rainbow Zero," a novel algorithmic framework that synergizes the forward-planning capabilities of search-based algorithms (such as MuZero and Muesli) with advanced value estimation techniques and world models (e.g., Dreamer and Rainbow). We utilize the board game *Settlers of Catan* as our primary testbed due to its complex mechanics, varying board layouts, and high variance. Currently, our framework is being evaluated on reduced-complexity, randomized Catan boards, alongside classic control environments and foundational games. By benchmarking against the Catantron library baselines, this project aims to isolate the impact of combining specific architectural components—such as search-based planning and distributional reinforcement learning—via planned ablation studies. Ultimately, the goal of this research extends beyond game-playing. By mastering the stochastic resource generation and long-term planning required in Catan, the Rainbow Zero architecture is being developed as a proxy for complex, real-world optimization problems, with future applications specifically targeting autonomous decision-making in mining operations and genomic sequence alignment.

© The Authors. This abstract is published under a CC-BY license:

<https://creativecommons.org/licenses/by/4.0/>