



## AlphaZero++ with Neural Message Passing

AlphaZero, and its successor MuZero were breakthrough algorithms for teaching AI to perform tree-search. They have been successful in solving complex reasoning problems such as board games, demonstrating how AI can iteratively improve decision-making by simulating future moves and learning from self-play. More recently, Large Language Models such as the OpenAI o-series have incorporated “test-time compute”, a technique that enables additional reasoning, which is also speculated to leverage tree search.

The standard tree search framework in AlphaZero has a major limitation: information flow between search steps is very restricted. Each node in the search tree communicates with others only through a scalar value estimate (see the diagram on the right). In this project, we wish to elicit the type of reasoning shown in grey by incorporating neural message passing between the nodes. Using Graph Neural Networks in the search tree enables more expressive communication, leading to more structured reasoning and efficient search.

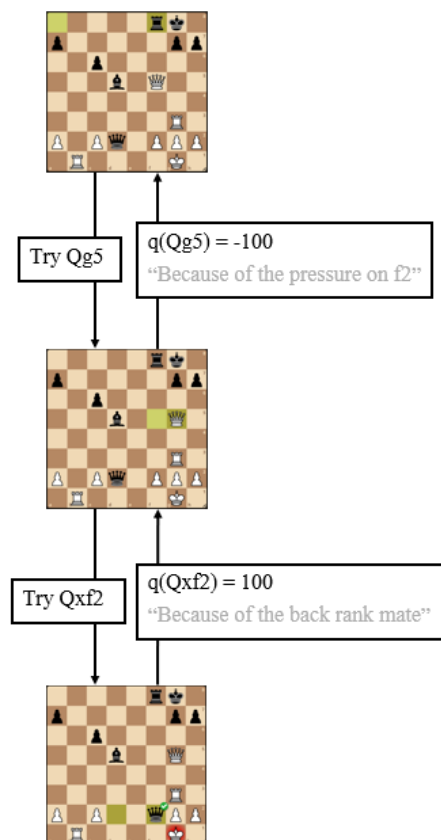
A potential extension of this work involves applying the developed techniques to open-source LLMs such as DeepSeek-R1.

### Requirements:

- Strong software engineering skills (ideally in the modern deep learning stack of Python & PyTorch/JAX) to quickly test & iterate on ideas
- Knowledge of Linear Algebra, Statistics, (ideally: Reinforcement Learning theory)

Interested? Please get in touch with us for more details!

- Saku Peltonen: [speltonen@ethz.ch](mailto:speltonen@ethz.ch), ETZ G63
- Sam Dauncey: [sdauncey@ethz.ch](mailto:sdauncey@ethz.ch), ETZ G61.1



Standard AlphaZero algorithm summarised in black, additional neural message reasoning in grey