

Athénan Wins 11 Gold Medals at the 2024 Computer Olympiad

Quentin Cohen-Solal * Tristan Cazenave

LAMSADE, Université Paris Dauphine - PSL, CNRS, Paris, France

E-mail: Quentin.Cohen-Solal@dauphine.psl.eu

Athénan (Cohen-Solal, 2020; Cohen-Solal and Cazenave, 2021b) is a zero knowledge Deep Reinforcement Learning algorithm that has learned to play many games. It won eleven gold medals at the 2024 Computer Olympiad. Note that, it also won sixteen gold medals in 2023 (Cohen-Solal and Cazenave, 2023a), five gold medals in 2022, eleven in 2021, and again five in 2020 (Cohen-Solal and Cazenave, 2021a). Moreover, Athénan is now the defending champion on 17 games (11 new gold medals plus 6 previous uncontested gold medals).

Unlike Alpha Zero-like algorithms (Silver et al., 2018), Athénan is based on the Descent framework (Cohen-Solal, 2020). Thus, during the training process, it uses a variant of Unbounded Minimax (Korf and Chickering, 1996) called *Descent*, instead of Monte Carlo Tree Search, to construct the partial game tree used to determine the best action to play and to collect data for learning. With Descent, at each move, the best sequences of moves are iteratively extended until terminal states. During evaluations, another variant of Unbounded Minimax is used. This variant contains in particular a generic solver and it chooses the safest action to decide between actions. Moreover, contrary to Alpha Zero, Athénan does not use a policy network, only a value network. The actions therefore do not need to be encoded. In addition, unlike the Alpha Zero paradigm, with Athénan all data generated during the searches to determine the best actions to play is used for learning. As a result, much more data is generated per match (Cohen-Solal and Cazenave, 2023b), and thus the training is done more quickly and does not require a (massive) parallelization to give good results (contrary to Alpha Zero). Athénan can use end-of-game heuristic evaluations to improve its level of play, such as game score or game length (in order to win quickly and lose slowly). Further improvements are described in (Cohen-Solal, 2020).

Thus, in 2024, eleven gold medals were won by Athénan for the following games: Amazons, Breakthrough, Clobber, Havannah 8x8, Hex 11x11, Hex 13x13, Lines of Action, Surakarta, Shobu, Othello 16x16, and Ataxx.

Athénan's main opponent in this competition was the MiniZero program (Wu et al., 2024) which reimplements the Gumbel AlphaZero algorithm (Danihelka et al., 2022). Gumbel AlphaZero is a variant of AlphaZero which biases the action selection by sampling the k best values using a Gumbel distribution and the policy value (plus the Q-value during the final selection). In addition, on some games, MiniZero uses additional expert knowledge. For example, in the games Connect6 and Gomoku, they encode raw threat space as input features (the threatening positions are marked).

MiniZero has participated in 13 games (compared to 4 last year) and they have obtained 5 gold medals (compared to 1 last year). Athenan and MiniZero have faced each other on 12 games. Among them,

* Corresponding author. E-mail: Quentin.Cohen-Solal@dauphine.psl.eu.

MiniZero has won 4 gold medals: International draughts, Gomoku, Connect6 and Santorini. In addition, among them, Athenan has won 8 gold medals (namely: Amazons, Breakthrough, Clobber, Havannah 8x8, Hex 11x11, Hex 13x13, Lines of Action, Surakarta).

The other competitors for each game were :

- Amazons: MiniZero.
- Ataxx: Karin of Katsuki Ohto.
- Breakthrough: MiniZero ; Apply Alphazero of Hui-Hsian Weng, Jun-Yi Li, and I-Chen Wu ; SuperPudim of Jean-Noël Vittaut and Nicolas Jouandeau.
- Clobber: MiniZero.
- Connect6: MiniZero ; Clap_connect6 of Chi-Huang Lin, Shao-Xiong Zheng, Yuan-Hao Chen, and I-Chen Wu.
- Havannah (8x8): MiniZero.
- Hex (11x11): MiniZero.
- Hex (13x13): MiniZero.
- International Draughts: MiniZero.
- Lines of Action: MiniZero.
- Othello (16x16): Hollyhock of Katsuki Ohto.
- Outer-Open Gomoku: MiniZero ; Clap of Wei-Chen Liao, Po-Wei Huang, Shao-Xiong Zheng, Yuan-Hao Chen, and I-Chen Wu.
- Santorini: MiniZero.
- Shobu: Shizu of Katsuki Ohto.
- Surakarta: MiniZero.

where the Team of MiniZero is Yu-Hung Chang, Hung-Tse Lin, Jui-Yu Wang, Yan-Ru Ju, Hung Guei, Chung-Chin Shih, Pei-Chiun Peng, and Ti-Rong Wu.

REFERENCES

- Cohen-Solal, Q. (2020). Learning to Play Two-Player Perfect-Information Games without Knowledge. *arXiv preprint arXiv:2008.01188*.
- Cohen-Solal, Q. & Cazenave, T. (2021a). DESCENT wins five gold medals at the Computer Olympiad. *ICGA Journal*, 43(2), 132–134.
- Cohen-Solal, Q. & Cazenave, T. (2021b). Minimax Strikes Back. In *Reinforcement Learning in Games at AAAI*.
- Cohen-Solal, Q. & Cazenave, T. (2023a). Athénan wins sixteen gold medals at the Computer Olympiad. *ICGA Journal*, 1–2.
- Cohen-Solal, Q. & Cazenave, T. (2023b). Minimax Strikes Back. In *AAMAS* (pp. 1923–1931).
- Danihelka, I., Guez, A., Schrittwieser, J. & Silver, D. (2022). Policy improvement by planning with Gumbel. In *International Conference on Learning Representations*.
- Korf, R.E. & Chickering, D.M. (1996). Best-first minimax search. *Artificial intelligence*, 84(1-2), 299–337.
- Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., Lanctot, M., Sifre, L., Kumaran, D., Graepel, T., et al. (2018). A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science*, 362(6419), 1140–1144.

Wu, T.-R., Guei, H., Peng, P.-C., Huang, P.-W., Wei, T.H., Shih, C.-C. & Tsai, Y.-J. (2024). MiniZero: Comparative Analysis of AlphaZero and MuZero on Go, Othello, and Atari Games. *IEEE Transactions on Games*. <https://github.com/rlglab/minizero>.