

ATHÉNAN Wins Sixteen Gold Medals at the Computer Olympiad

Quentin Cohen-Solal^a, Tristan Cazenave^a

^a LAMSADE, University Paris-Dauphine, PSL, CNRS, France

E-mail: quentin.cohen-solal@dauphine.psl.eu

Athénan (Cohen-Solal, 2020; Cohen-Solal and Cazenave, 2021) is a zero knowledge Deep Reinforcement Learning algorithm that has learned to play many games. It won sixteen gold medals at the 2023 Computer Olympiad. By contrast, it won five gold medals in 2020, eleven in 2021, and again five in 2022.

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, 2023), 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 2023, sixteen gold medals were won by Athénan for the following games: Amazons, Arimaa, Ataxx, Breakthrough, Canadian Draught, Chinese Chess, Clobber, Havannah 8x8, Havannah 10x10, Hex 11x11, Hex 13x13, Hex 19x19, Lines of Action, Othello 10x10, Santorini, and Surakarta. Silver medals were won by our programs at Brazilian Draughts, Connect6, International Draught. Bronze medals were also won by our programs at Othello 8x8 and Outer-Open-Gomoku.

The other competitors for each game were:

- Amazons: program Asura of Ohto Katsuki (it is based on Policy Gradient and it uses a unique neural network for all the games).
- Arimaa: Asura of Ohto Katsuki.
- Ataxx: Asura of Ohto Katsuki.
- Brazilian Draughts: Asura of Ohto Katsuki and Tdk88 of Ton Tillemans.
- Breakthrough: Apply Alphazero of Hui-Hsian Weng, Jun-Yi Li, and I-Chen Wu ; Asura of Katsuki Ohto ; SuperPudim of Jean-Noël Vittaut and Nicolas Jouandeau.
- Canadian Draughts: Asura of Katsuki Ohto ; TdkCan of Ton Tillemans.
- Chinese Chess: Asura of Ohto Katsuki.

- Clobber: Asura of Ohto Katsuki.
- Connect6: Asura of Katsuki Ohto ; Clap_connect6: of Chi-Huang Lin, Shao-Xiong Zheng, Yuan-Hao Chen, and I-Chen Wu.
- Havannah (8×8): Asura of Katsuki Ohto ; Clap_havannah of Hong-Jun Chiu, Nian-Rong Wu, Yu-Chun Chen, Yen-Chi Chen, Yuan-Hao Chen, and I-Chen Wu.
- Havannah (10×10): Asura of Ohto Katsuki.
- Hex (11×11): Asura of Katsuki Ohto ; Minizero of Markus Fjellheim, Ti-Rong Wu, and I-Chen Wu.
- Hex (13×13): Asura of Ohto Katsuki.
- Hex (19×19): Asura of Ohto Katsuki.
- International Draughts: Asura of Ohto Katsuki and Tdking of Ton Tillemans..
- Lines of Action: Asura of Ohto Katsuki.
- Othello (8×8): Aiiia_othello of Hsuan-Chia Kang, Tin-En Hsu, and Hsin-Hung Chou ; Asura of Katsuki Ohto ; Maverick of Yen-Chi Chen and Shun-Shii Lin ; Minizero_othello of Pei-Chiun Peng, Po-Wei Huang, Hung Guei, Ti-Rong Wu, and I-Chen Wu.
- Othello (10×10): Asura of Ohto Katsuki.
- Outer-Open Gomoku: Asura of Katsuki Ohto ; Clap_oog of Wei-Chen Liao, Po-Wei Huang, Shao-Xiong Zheng, Yuan-Hao Chen, and I-Chen Wu ; Corking of Bing-Lung Tsai and Shun-Shii Lin ; Minizero_oog of Yu-Chen Pai, Ti Wu, Hung Guei, and Ti Rong Wu ; Minizero_tssoog of Chi-Feng Liu, Po-Wei Huang, and Ti Rong Wu ; Peace_oog of Yu-Chang Lin and Shun-Shii Lin ; Stone_oog of De-Wei Huang and Shun-Shii Lin.
- Santorini: Asura of Ohto Katsuki.
- Surakarta: Asura of Ohto Katsuki.

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. (2021). Minimax Strikes Back. In *Reinforcement Learning in Games at AAAI*.
- Cohen-Solal, Q. & Cazenave, T. (2023). Minimax Strikes Back. In *AAMAS* (pp. 1923–1931).
- 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.