**Mastering Atari, Go, chess and shogi by planning with a learned model**

* [Julian Schrittwieser](javascript:;),
* [Ioannis Antonoglou](javascript:;),
* [Thomas Hubert](javascript:;),
* [Karen Simonyan](javascript:;),
* [Laurent Sifre](javascript:;),
* [Simon Schmitt](javascript:;),
* [Arthur Guez](javascript:;),
* [Edward Lockhart](javascript:;),
* [Demis Hassabis](javascript:;),
* [Thore Graepel](javascript:;),
* [Timothy Lillicrap](javascript:;) &
* [David Silver](javascript:;)

[*Nature*](https://www.nature.com/nature) **volume 588**, pages604–609(2020)

## Abstract

Constructing agents with planning capabilities has long been one of the main challenges in the pursuit of artificial intelligence. Tree-based planning methods have enjoyed huge success in challenging domains, such as chess[1](https://www.nature.com/articles/s41586-020-03051-4#ref-CR1) and Go[2](https://www.nature.com/articles/s41586-020-03051-4#ref-CR2), where a perfect simulator is available. However, in real-world problems, the dynamics governing the environment are often complex and unknown. Here we present the MuZero algorithm, which, by combining a tree-based search with a learned model, achieves superhuman performance in a range of challenging and visually complex domains, without any knowledge of their underlying dynamics. The MuZero algorithm learns an iterable model that produces predictions relevant to planning: the action-selection policy, the value function and the reward. When evaluated on 57 different Atari games[3](https://www.nature.com/articles/s41586-020-03051-4#ref-CR3)—the canonical video game environment for testing artificial intelligence techniques, in which model-based planning approaches have historically struggled[4](https://www.nature.com/articles/s41586-020-03051-4#ref-CR4)—the MuZero algorithm achieved state-of-the-art performance. When evaluated on Go, chess and shogi—canonical environments for high-performance planning—the MuZero algorithm matched, without any knowledge of the game dynamics, the superhuman performance of the AlphaZero algorithm[5](https://www.nature.com/articles/s41586-020-03051-4#ref-CR5) that was supplied with the rules of the game.

References ( extract)

Silver, D. et al. Mastering the game of Go with deep neural networks and tree search. *Nature* **529**, 484–489 (2016).

Silver, D. et al. A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science* **362**, 1140–1144 (2018).

Silver, D. et al. Mastering the game of Go without human knowledge. *Nature* **550**, 354–359 (2017).

# Maîtriser les jeux Atari, le Go, les échecs et le shogi en utilisant la planification selon un modèle appris

## Abstract

Construire des agents avec des capacités de planification a longtemps été l'un des principaux défis poursuivi par l'intelligence artificielle. Les méthodes de planification basées sur les arbres ont connu un énorme succès dans des domaines difficiles, tels que les échecs et le Go, où un simulateur parfait est disponible. Cependant, dans les problèmes du monde réel, les dynamiques régissant l'environnement sont souvent complexes et inconnues. Nous présentons ici l'algorithme MuZero, qui, en combinant une recherche arborescente avec un modèle appris, atteint des performances surhumaines dans une gamme de domaines difficiles et visuellement complexes, sans aucune connaissance de leur dynamique sous-jacente. L'algorithme MuZero apprend un modèle itératif qui produit des prédictions pertinentes pour la planification: la politique de sélection d'action, la fonction d’évaluation et la récompense. Lorsqu'il est évalué sur 57 jeux Atari différents - l'environnement de jeu vidéo canonique pour tester les techniques d'intelligence artificielle, dans lequel les approches de planification basées sur des modèles ont toujours eu des difficultés, l'algorithme MuZero a atteint le meilleur niveau actuel. Évalué sur le Go, les échecs et le shogi - environnements canoniques pour la planification haute performance - l'algorithme MuZero correspondait, sans aucune connaissance de la dynamique du jeu, aux performances surhumaines de l'algorithme AlphaZero qui était fourni avec les règles du jeu.

References ( extrait)

Silver, D. et al. Mastering the game of Go with deep neural networks and tree search. *Nature* **529**, 484–489 (2016).

Silver, D. et al. A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science* **362**, 1140–1144 (2018).

Silver, D. et al. Mastering the game of Go without human knowledge. *Nature* **550**, 354–359 (2017).