Mach Learn (2012) 87:183–219
DOI 10.1007/s10994-012-5280-0

Temporal-difference search in computer Go

David Silver · Richard S. Sutton · Martin Müller

Received: 23 November 2010 / Accepted: 29 January 2012 / Published online: 21 February 2012

Abstract Temporal-difference learning is one of the most successful and broadly applied solutions to the reinforcement learning problem; it has been used to achieve master-level play in chess, checkers and backgammon. The key idea is to update a value function from episodes of real experience, by bootstrapping from future value estimates, and using value function approximation to generalise between related states. Monte-Carlo tree search is a recent algorithm for high-performance search, which has been used to achieve master-level play in Go. The key idea is to use the mean outcome of simulated episodes of experience to evaluate each state in a search tree. We introduce a new approach to high-performance search in Markov decision processes and two-player games. Our method, temporal-difference search, combines temporal-difference learning with simulation-based search. Like Monte-Carlo tree search, the value function is updated from simulated experience; but like temporal-difference learning, it uses value function approximation and bootstrapping to efficiently generalise between related states. We apply temporal-difference search to the game of 9×9 Go, using a million binary features matching simple patterns of stones. Without any explicit search tree, our approach outperformed an unenhanced Monte-Carlo tree search with the same number of simulations. When combined with a simple alpha-beta search, our program also outperformed all traditional (pre-Monte-Carlo) search and machine learning programs on the 9×9 Computer Go Server.

Keywords Reinforcement learning · Temporal-difference learning · Monte-Carlo search · Simulation based search · Computer Go

Editor: Csaba Szepesvari.

D. Silver (✉)
University College London, London, UK
e-mail: d.silver@cs.ucl.ac.uk

R.S. Sutton · M. Müller
University of Alberta, Edmonton, Canada