value
$+1-10$
maximise
minimise
maximiseminimum
two-player zero-sum
value
maximum
minimum
return$= 1 = -\infty$**each**$v3-v$**return**$= 2 = +\infty$**each**$v3-v$**return**
each$v3-v$
$-1+1$
$= 1$
$= -1$
two-player zero-sumperfect informationsequential moves
Nash equilibrium
perfectly
But...
$\approx 10^{13}$
$\approx 10^{47}$
terminal
Depth limiting
$+10-1$
d
$-1+1$
$d = 1$
stop early$+1-1$
alpha-beta pruning
best to worst
playing strengthheuristic
in-depth knowledge
heuristic evaluation function
Monte Carlo methodaveraging over random samples
averageexpected value
quickly approximatinglarge domains
converge in the limit
infiniteexact
number of samplesaccuracy
deterministic
pseudo-random number generators (PRNGs)
unpredictableseed
uniformly distributed
entropy
random rollouts
$smssm$
valuevalue$1-10$
expected value
$=$
Flat Monte Carlo search
$d > 1$
deterministic
rollouts
random
bias
plausible
tree
root node
Selection
Expansion
Simulation
Backpropagationselectionexpansion
Exploitation
Exploration
multi-armed bandit problem
do not know
Exploitation
Exploration
m
n_m
V_m
$n = \sum_m n_m$

$\frac{V_m}{n_m}$**exploitation**
$\sqrt{\frac{\log n}{n_m}}$**exploration**$n_m$
c
n_x
V_x
pq

game independent
enhance
anytime
any
$O(e^d)d$
horizon effect
$dd + 1$