Monte Carlo Tree Search

default policy: random

move-averaged sampling technique (MAST)
  extract possible moves from game definition
  maintain reward for each move over all playouts
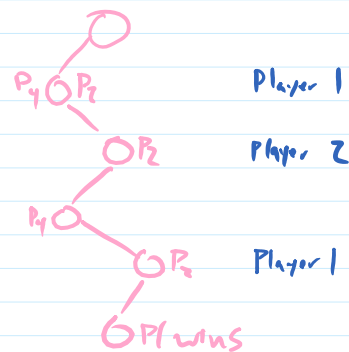 bias towards actions with higher observed reward

Kalah - pit to start from
$0, \ldots, 5, 7, \ldots, 12$
chess - all pairs of squares in same row/col/diag/left

.545
~~.1~~ .52 .53 .0 .49 .47
0   1   2   3   4   5
if we choose pit 0 and win
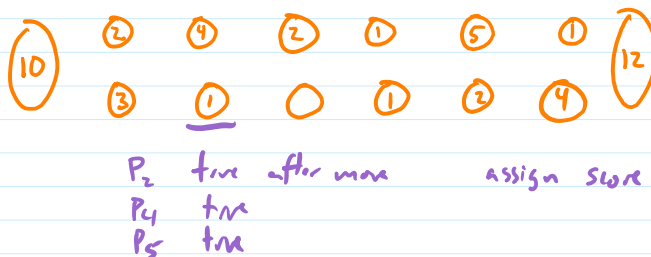
predicate average   (PAST)

P_1 : N has more in store
P_2 : S has more in store
P_3 : N has no pit with >2 seeds
P_4 : $\sim P_3$
P_5 : last pit is empty
P_6 : $\sim P_5$
    ⋮

	n	r
P_1 : N has more in store	100	0.52
P_2 : S has more in store	~~50~~ 53	~~0.4~~ 0.434
P_3 : N has no pit w/ >2 seeds	50	0.34
P_4 : $\sim P_3$	~~107~~ ~~100~~	(0.55) 0.56
P_5 : last pit is empty	50	0.7
P_6 : $\sim P_5$	100	0.3

Player 1
Player 2
Player 1
P1 wins

P_2 reward +1
P_2 n +3

P_1 move and capture
P_2 move and capture
P_1 move and free
P_2 move and free
    ⋮

10 ② ④ ② ① ⑤ ① 12
 ③ ① ○ ① ② ④

P_2 true after move   assign score
P_4 true
P_5 true

$$\frac{0.4 \cdot 50 + 0.55 \cdot 100 + 0.7 \cdot 50}{200}$$

bias towards moves with higher score

move was a capture
move resulted in free turn
    ⋮