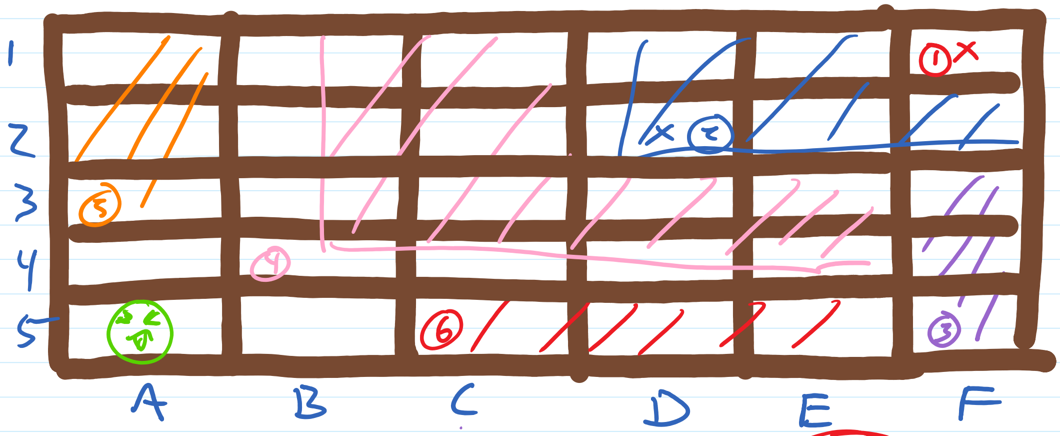
Chomp

Play on m×n grid. Take turns selecting remaining cell, remove all above and to right

Last move loses.

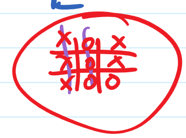
misère

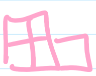


# of browns in each column

# of browns

000 N
100 P D
110 N
200 N
111 N
210
211
220
221
222



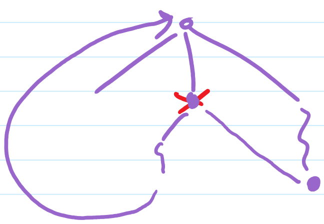


outcome_class(p)

if p is end of game (terminal)
  return value according to rules — normal: P (lose); misère: N (win)
else
  S ← positions reachable in 1 move from p
  if S contains a P position
    return N
  else
    return P

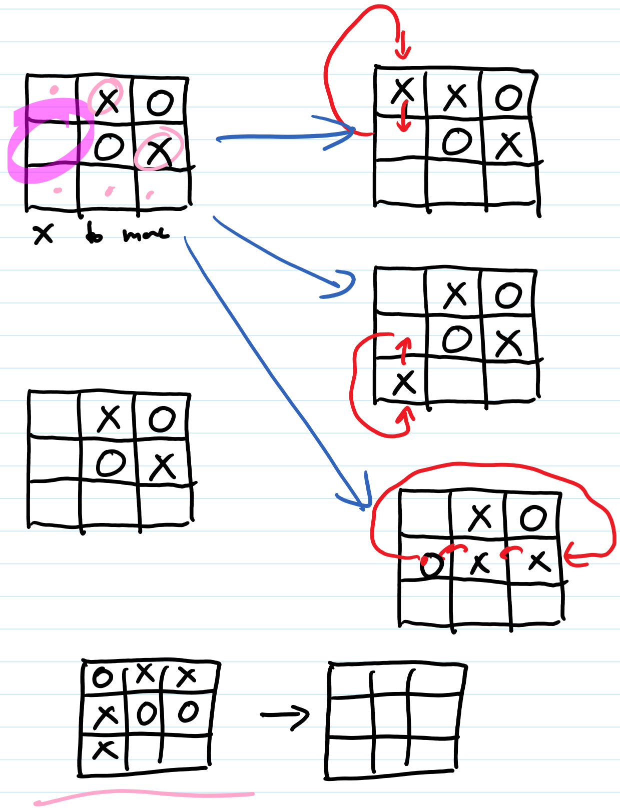
solve(s)
solve(s')
solve(s'')
solve(t)

10^5



Quixo

[Quixo | Board Game | BoardGameGeek]



Value Iteration

mark all positions as draw
mark all terminal positions mark according to rules
repeat
  for all positions p marked as draws   player to move at p
    if some successor is win for current player
      mark as win for current player
    else if all successors losses for current player
      mark as loss for current player
until no new positions marked

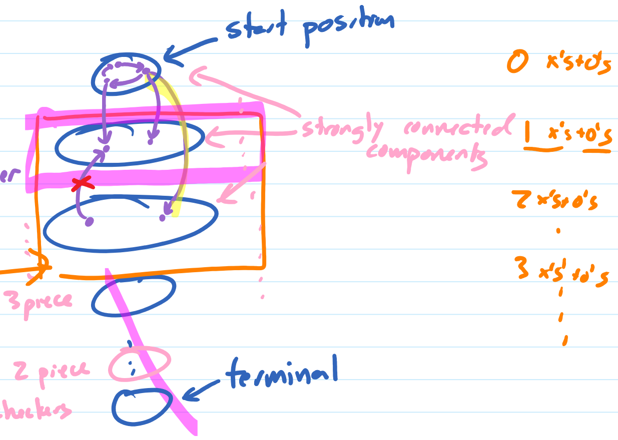
33666

33366 initial
1 2 4 5 6 FH
1 1 1 2 6 FH end of turn

break states into layers

for each layer terminal → start
  apply value iteration in layer

current storage



Markov Decision Process

Scoresheet, roll, rerolls ← states (positions)

Markov Decision Process : (S, A, p, R) — actions (moves) $(0, A(s))$; rewards (scores)

1 2 4 5 6 → 6 6 6 6 6

$p : S \times A \times S \times R \to [0, 1]$

(current state, action, next state, reward)

$p((\square, 66666, 1), 0 \mid (\square, 12456, 2), 6) = \frac{1}{6^4}$

$p((\square, 55555, 0), 100 \mid (\square, 12456, 2), 6)$

$= 0$

$P((\boxed{\text{50}}, 12566, 2), 50 \mid (\square, 66666, 2), 4)$

$= P(\text{rolling } 12566)$

$p(s', r \mid s, a)$ = prob. of reaching s' w/ reward r given current state s, action a

$$p(s' \mid s, a) = \sum_r p(s', r \mid s, a)$$

$$r(s, a) = \sum_r \sum_{s'} r \cdot p(s', r \mid s, a)$$

Episode : $(s_0, a_0, s_1, R_1), (s_1, a_1, s_2, R_2), \ldots, [(s_{T-1}, a_{T-1}, s_T, R_T)]$

$G = R_1 + R_2 + R_3 + \cdots$ Goal : maximize sum of rewards