Value Iteration

Initialize $v[s] \leftarrow 0$ for terminal s
$\leftarrow$ arbitrary for nonterminal s

repeat
$\Delta \leftarrow 0$
for each state s

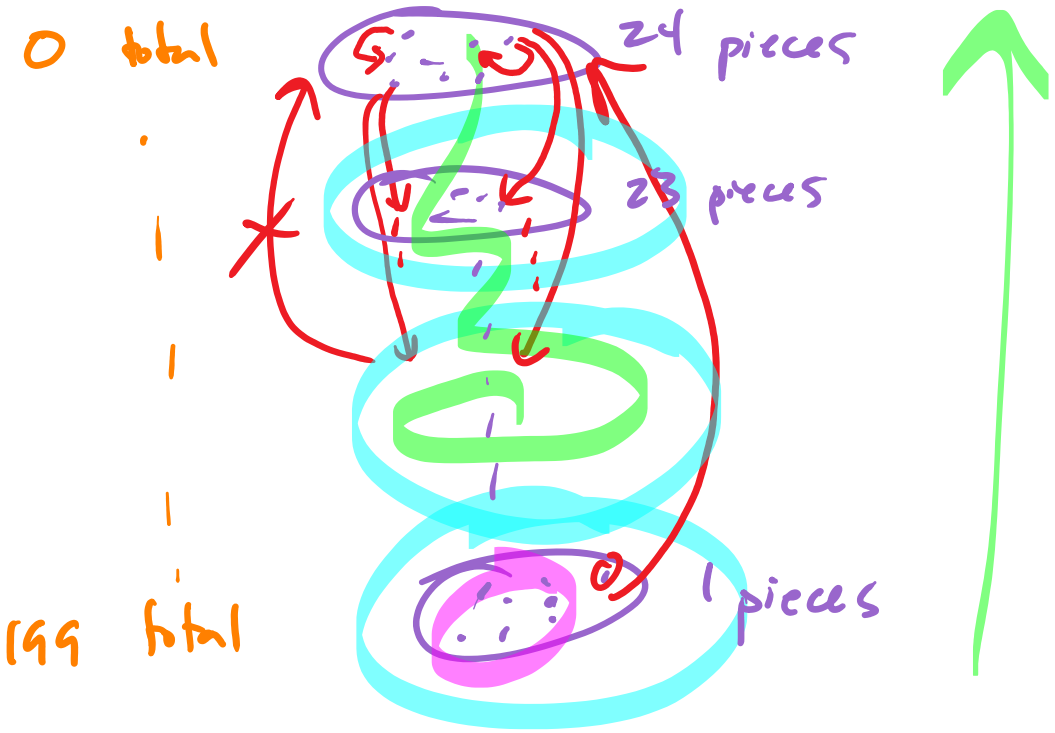
$v_{old} \leftarrow v[s]$

$v[s] \leftarrow \sum_{s'} \sum_{r} p(s', r | s, a) \cdot (r + \gamma v_{\pi}(s'))$

$\Delta \leftarrow \max(\Delta, |v[s] - v_{old}|)$

until Δ small enough

for each s
$\pi[s] \leftarrow$

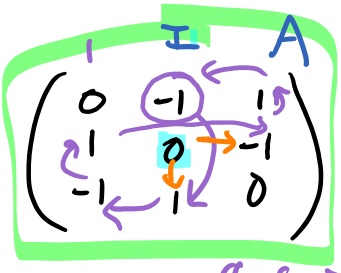

0 total
24 pieces
23 pieces
1 pieces
199 total

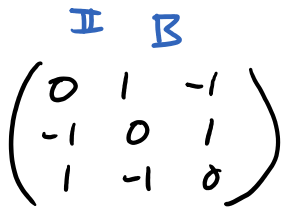
Simultaneous Play Games

I \ II	R	P	S
Rock	0,0	−1,1	1,−1
Paper	1,−1	0,0	−1,1
Scissors	−1,1	1,−1	0,0

$v^- = -1$, $v^+ = 1$

zero-sum: payoff for I + payoff for II





a_{ij} = payoff for I when I chooses action i, II chooses j

Penalty Kick

kicker \ goalkeeper	L	R
L	$\frac{1}{8}, -\frac{1}{8}$	$\frac{1}{4}, -\frac{1}{4}$
R	$\frac{1}{4}, -\frac{1}{4}$	$\frac{1}{6}, -\frac{1}{6}$

$v^- = \frac{1}{8}$, $v^+ = \frac{1}{4}$

$A = \begin{pmatrix} \frac{1}{8} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{6} \end{pmatrix}$

(zero-sum "saddle point")

equilibrium

constant-sum: $a_{ij} + b_{ij} = C$

zero-sum → 1) play game with $A - C$

no strategy → 2) award player I addl payoff C

I \ II	W	X	Y	Z
A	$\frac{1}{2}$	$\frac{3}{4}$	$\frac{1}{2}$	$\frac{3}{4}$
B	−1	$-\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{4}$	$-\frac{1}{2}$
C	−1	$\frac{1}{2}$	0	−1	$\frac{1}{2}$

max in col and min in row

v^- = amt I can guarantee = max of row mins

v^+ = amt II can guarantee not to exceed = min of col maxes

Stag Hunt

I \	Stag	Hare
Stag	2,2 (R)	0,1 (T)
Hare	1,0 (S)	1,1 (P)

$R > T = P > S$

Prisoners Dilemma

	Coop	Defect
Coop	3,3 (R)	0,5 (T)
Defect	5,0 (S)	1,1 (P)

equilibrium

R = reward for cooperating
P = punishment for defecting
S = sucker's reward
T = reward for temptation
$T > R > P > S$