Playing Pig

2-players, turn-based or solitaire

On each turn
- roll
- if 1, then turn over (and no change in total)
- else add number to turn total
- decide: repeat
  - stop (and add turn total to score)

1st to 100 points wins minimize expected turns

Prof. Glenn		Class	
$\boxed{0}$	~~4+1~~	0	~~3+5+1~~
$\boxed{0}$	5+5+4+5+5	0	6+6+4
24	~~4+3+1~~	16	3+5+3+6+6
24		39	

not finite

⋮

Value iteration and policy iteration

Solitaire Pig

– expected turns to finish

$V(T,p,p',r)$ = value with target T, score p, turn total p', roll r

$\to V(100,24,0,1)$

$\to V(100,24,0,1)$

$$= \begin{cases} 0.0 & \text{if } p \ge T \\ -1 + \frac{1}{6}\sum_{r'=1}^{6} V(T,p,0,r') & \text{if } r=1 \text{ and } p<T \\ \max\left(\underbrace{\frac{1}{6}\sum_{r'=1}^{6} V(T,p,p'+r,r')}_{\text{value of rolling again}},\ \underbrace{\frac{1}{6}\sum_{r'=1}^{6}\left(V(T,p+p'+r,0,r')-1\right)}_{\text{value of stopping}}\right) & \text{otherwise } (r>1) \end{cases}$$

Add turn limit; game over when turn limit reached (game is now finite again)

$V'(T,p,p',r,t)$ = value with target T, score p, turn total p', roll r, t turns left

$$= \begin{cases} 0.0 & \text{if } t=0 \\ 0.0 & \text{if } p \ge T \\ -1 + \frac{1}{6}\sum_{r'=1}^{6} V'(T,p,0,r',t-1) & \text{if } r=1 \text{ and } p<T \\ \max\left(\underbrace{\frac{1}{6}\sum_{r'=1}^{6} V'(T,p,p'+r,r',t)}_{\text{value of rolling again}},\ \underbrace{\frac{1}{6}\sum_{r'=1}^{6}\left(V'(T,p+p'+r,0,r',t-1)-1\right)}_{\text{value of stopping}}\right) & \text{otherwise } (r>1) \end{cases}$$

as $t \to \infty$, $V'(s) \to V(s)$

$V(\ldots\ldots,0)$ ← initial v

$v(\ldots\ldots,1)$ — v after 1 iteration

$v(\ldots\ldots,2)$ — v after 2 iteration

Value iteration

$v(s) \leftarrow 0$ for all s

repeat

- for each s
  - $v'(s) \leftarrow \max_a \sum_{s'} P(s' \mid s,a)\cdot\left[R(s,a,s') + \gamma\cdot v(s')\right]$
- $\Delta \leftarrow \max_s |v(s) - v'(s)|$
- $v \leftarrow v'$

until Δ small enough (until converged to values under optimal policy)

for each s, $\pi_{OPT}(s) \leftarrow \arg\max_a \sum_{s'} P(s' \mid s,a)\cdot\left[R(s,a,s') + \gamma\cdot v(s')\right]$

Policy iteration

$v(s)$, $\pi'(s) \leftarrow$ whatever

repeat

- $\pi \leftarrow \pi'$
- repeat
  - for each s
    - $v'(s) \leftarrow \sum_{s'} P(s' \mid s,\pi(s))\cdot\left[R(s,\pi(s),s') + \gamma\cdot v(s')\right]$

for each s

$V'(s) \leftarrow \sum_{s'} P(s' \mid s, \pi(s)) \cdot \left[R(s, \pi(s), s') + \gamma \cdot V(s') \right]$

$\Delta \leftarrow \max_{s} |V(s) - V'(s)|$

$V \leftarrow V'$

until Δ small enough (until converged to values under optimal policy)

for all s, $\pi'(s) \leftarrow \arg\max_{a} \sum_{s'} P(s' \mid s, a) \cdot \left[R(s, a, s') + \gamma \cdot V(s') \right]$

until $\pi' = \pi$

now $\pi = \pi_{OPT}$!

Two-player Zero-sum, probabilistic finite games

↳ sum of reward for players is always 0

zero-sum
$1 + -1 = 0$
$-1 + 1 = 0$
$0 + 0 = 0$

constant sum
$1 + 0 = 1$
$0 + 1 = 1$
$\frac{1}{2} + \frac{1}{2} = 1$

non-constant sum
$2 + 0 = 2$
$0 + 2 = 2$
$2 + 1 = 3$
$1 + 2 = 3$

$3 + 0$ 3
$0 + 3$ 3
$1 + 1$ 2

For P1 choice position

$$V(s) = \max_{s'} \left(V(s') + R(s \rightarrow s') \right)$$

expected wins for P1 (P2)

reward for P1 for going from s to s'

For P2 choice position

$$V(s) = \min_{s'} V(s') + R(s \rightarrow s')$$

For non-final random event positions

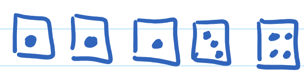
$$V(s) = \sum_{s'} P(s \rightarrow s') \left(V(s') + R(s \rightarrow s') \right)$$

2-player Yahtzee

optimal solitaire strategy is not optimal for 2-player

Aces	1	3	3K	15	20
Deuces	4	4	4K	0	
Treys	6	9	FH	25	0
Fours	16	12	SS	30	30
Fives	10	15	LS	0	40
Sixes	18	12	Chance		28
Bonus	0	0	Yahtzee		

Total



opt solitaire choice is
(0 + 21 vs 10 + 7)
0 in Y

0.0 vs small but > 0.0

anchors: Y 3^2
other cats 2^{24}
upper subtot 64^2
Δ score   ~3000

$\approx 2^{47}$ anchors   (2^{20} for solitaire)

(~3000 years of computation)