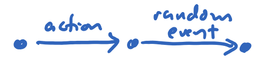
Shut The Box

1 ~~2~~ ~~3~~ ~~4~~ ~~5~~ 6 7 ~~8~~ ~~9~~ — P1 score = 14

~~1~~ ~~2~~ ~~3~~ 4 ~~5~~ ~~6~~ 7 ~~8~~ ~~9~~ — P2 score = 11

P2 wins!

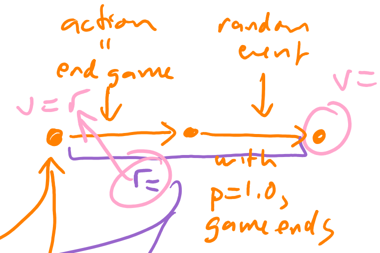
P1 MDP terminal state rewards = 1 − expected P2 reward for initial state w/ P1 score as target



for P_2 : $V_2(s)$ = expected wins for P2 (draw = $\frac{1}{2}$ win)

state s is tuple (S, r, t)

set of open tiles, roll, player 1's score



$$V_2(S, r, t) = \begin{cases} 0.0 & \text{if no subset } S' \subseteq S \text{ has } sum(S') = r \text{ and } sum(S) > t \\ 0.5 & \text{if no subset } S' \subseteq S \text{ has } sum(S') = r \text{ and } sum(S) = t \\ 1.0 & \text{if no subset } S' \subseteq S \text{ has } sum(S') = r \text{ and } sum(S) < t \\ \max\limits_{\substack{S' \subseteq S \\ sum(S') = r}} \sum\limits_{r'=2}^{12} P(\text{roll } r') \cdot V_2(S - S', r', t) & \text{if } sum(S - S') > 6 \\ \max\limits_{\substack{S' \subseteq S \\ sum(S') = r}} \sum\limits_{r'=1}^{6} P(\text{roll } r') \cdot V_2(S - S', r', t) & \text{if } sum(S - S') \le 6 \end{cases}$$

(P2 rolls to end even if already has winning score – can add cases below to avoid)

$g'((S, r, t), S')$

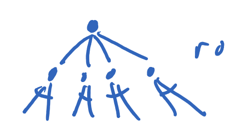
or 1.0 if $sum(S - S') < t$

for P1

$$V_1(S, r) = \begin{cases} 1.0 & \text{if } S = \emptyset \\ \max\limits_{\substack{S' \subseteq S \\ sum(S') = r}} \sum\limits_{r'=2}^{12} P(\text{roll } r') \cdot V_1(S - S', r') & \text{if } sum(S - S') > 6 \\ \max\limits_{\substack{S' \subseteq S \\ sum(S') = r}} \sum\limits_{r'=1}^{6} P(\text{roll } r') \cdot V_1(S - S', r') & \text{if } sum(S - S') \le 6 \\ 1 - \left[\sum\limits_{r'=2}^{12} P(\text{roll } r') \cdot V_2(\{1, \dots, 9\}, r', sum(S)) \right] & \text{if no } S' \subseteq S \text{ has } sum(S') = r \text{ (and } S \neq \emptyset) \end{cases}$$

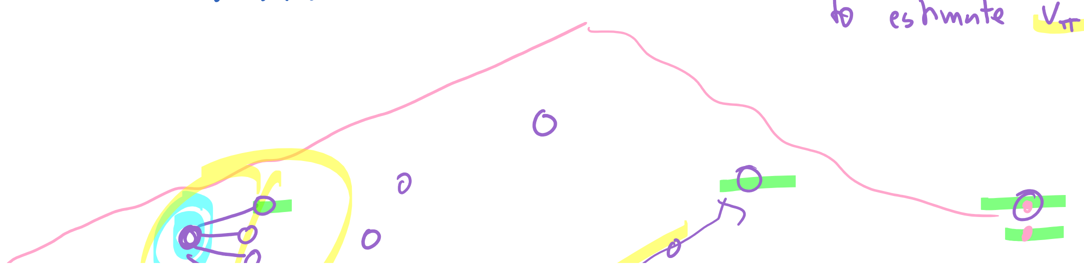
(add cases below to avoid rolls after box shut)

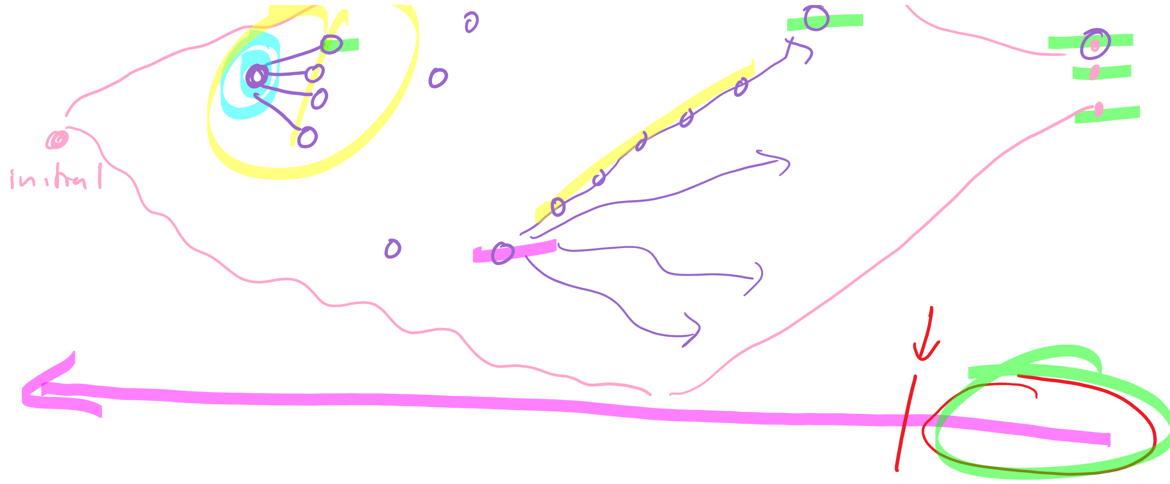
P_1's expected wins



compute $V^*(\text{initial})$ and π

simulate using π to estimate $V_\pi(\text{initial})$




initial