**Hoot Owl Hoot**

State: owls, sun, cards (left/right), turn, ~~discards~~

$$\underbrace{\underbrace{\left(\binom{39}{6}+\binom{39}{5}+\cdots+\binom{39}{0}\right)}_{2.4\text{ million}} \cdot 14 \cdot 84 \cdot 84 \cdot 2}_{\approx 470 \text{ billion}} \quad \cancel{\text{millions}}$$

discards shuffled back in

$\approx 470$ billion $\approx$ 4TB

(1 million/sec $\rightarrow$ 5½ days)

pos of owls, sun, left cards, right cards, turn

$$\underbrace{V}_{\text{expected \# of wins}}[\Theta, s, C_0, C_1, t] = \begin{cases} 1.0 & \text{if } \Theta = \emptyset \\ 0.0 & \text{if } s = 13 \\ \max\limits_{\substack{c \in C_0 \\ o \in \Theta}} \sum\limits_{c'} P(c' \mid C_0 \cup C_1) \cdot V[\Theta', s, C_0 - \{c\} \cup \{c'\}, C_1, 1] & \text{if } t = 0 \text{ and SUN} \notin C_0 \end{cases}$$

$\Theta'$: set of owl pos resulting from moving owl at pos $o$ to next space of color $c$

$\sum_{c'}$: sum over all 6 colors and SUN

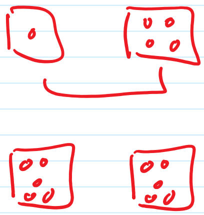
$$\sum_{c'} P(c' \mid C_0 \cup C_1) \cdot V[\Theta, s+1, C_0 - \{\text{SUN}\} \cup \{c'\}, C_1, 1] \quad \text{if } t=0 \text{ and SUN} \in C_0$$

(similar for $t=1$)

**Shut the Box**

~~1~~ ~~2~~ ~~3~~ ~~4~~ ~~5~~ 6 7 8 9

roll 2 dice if current sum > 6
1 die otherwise



shut open tiles to sum to 5

1 . . . 28 29 30 31 32

1.0 higher P(win)

0.5 0.5 lower expected score

$\boxed{3}$ $\boxed{5}$

4 2

1 5 over score 4+5+7 = 16

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

state $s$ is tuple $(S, r, t)$

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

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

for $P_1$ : $V(s)$ = expected wins

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

$$\sum_{r'=2}^{12} P(r') \cdot V(\{1, \ldots, 9\}, r')$$

$$\underbrace{\sum_{r'=2}^{12} P(r') \cdot V(\underbrace{\{1, \dots, 9\}, r'}\,)}_{\substack{\text{expected wins for P1} \\ \text{at start of game} \\ \text{(before 1st roll)}}}$$

Markov Decision Process

For nonterminal action positions

$$\underbrace{V(s)} = \max_{\text{action } a} \underbrace{V(\overbrace{next(s,a)}^{q})}_{\text{random event pos}} + \underline{score(s,a)}$$

For nonterminal random event positions

$$V(s) = \left(\sum_{\text{outcome } \sigma} P(\sigma) \cdot \left(V(next(s, \sigma)) + score(s,\sigma)\right)\right)$$

$$\max_a \sum_{\sigma} P(\sigma \mid s,a) \cdot \left[V(\overbrace{next(q, \sigma)}^{s'}) + \underbrace{score(q, \sigma) + score(s,a)}\right]$$

$$\max_a \sum_{s'} \underbrace{P(s' \mid s,a)} \cdot \left[V(s') + score(s, s', a)\right]$$

$\sum_{\substack{\sigma \text{ leads to} \\ s' \text{ from } s \\ \text{after } a}} P(\sigma \mid s,a)$

expected score for $s \to s'$ $= \sum_{\sigma} P(\sigma \mid s,a) \cdot score(s,a,\sigma)$

Markov Decision Process : $(S, A, P_a, R_a)$

states (positions) — actions — probability

$P_a : S \times S \to [0,1]$

$P_a(s, s') =$ prob $s \to s'$ under action $a$

$R_a(s,s') =$ expected reward for $s \to s'$ under action $a$ (points)

Policy : (strategy) $\pi : S \to A$ gives action $a$ to take in state $S$

$V_\pi(s)$ : value of $s$ under policy $\pi$ = expected total reward starting in $s$ following $\pi$

$$= \sum_{s'} P_{\pi(s)}(s, s') \cdot \left(R_{\pi(s)}(s, s') + \gamma V(s')\right)$$

discount factor $0 < \gamma < 1$ for a finite game $\gamma = 1$

$\pi_{opt}$ : policy that maximizes $V_\pi(s)$ for every $s$

$\pi_{opt}$ satisfies (for all $s$)

$$\pi_{opt}(s) = \arg\max_{a} \sum$$

$$\max_{a} \sum_{s'} P(s' \mid s, a) \cdot \left( V(s') + \text{score}(s, s', a) \right)$$