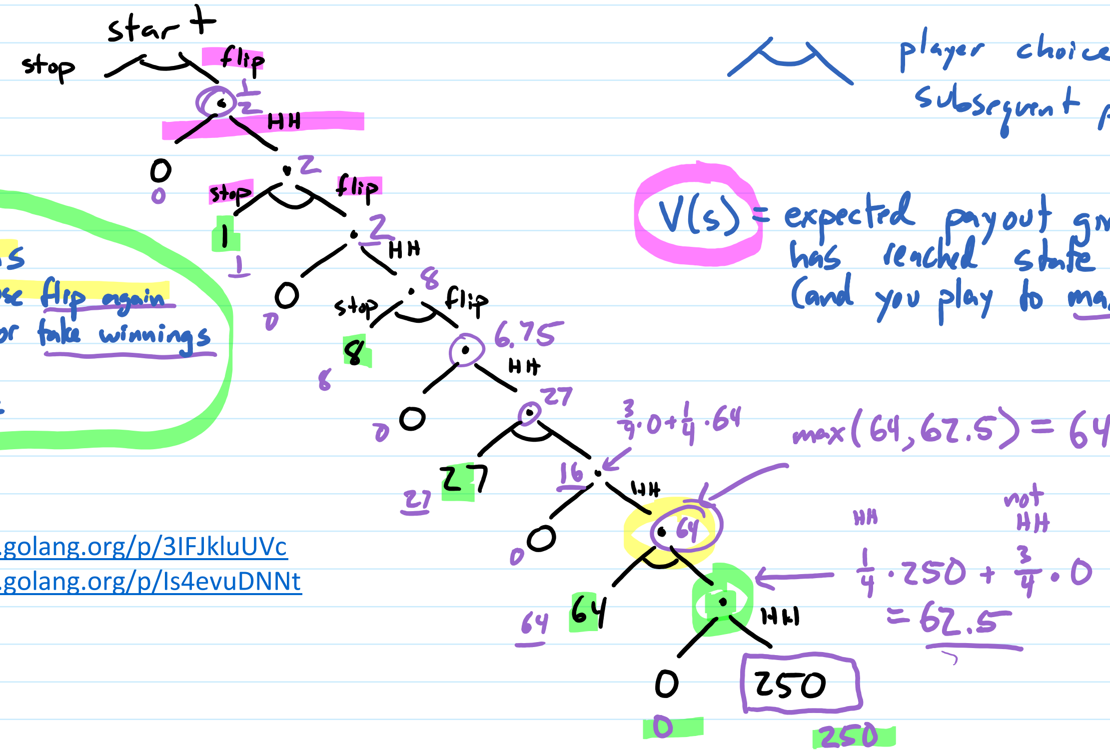
Coin Flipping Game



player choice of subsequent position

$V(s)$ = expected payout given game has reached state s
(and you play to maximize payout)

flip 2 coins
if HH choose flip again
or take winnings
otherwise
lose

https://play.golang.org/p/3IFJkluUVc
https://play.golang.org/p/Is4evuDNNt

Finite 1-Player Probabilistic Games

$V(s)$ = expected winnings having reached state s

game over

For terminal states s, $V(s)$ determined by rules

For nonterminal action states

$$V(s) = \max_{\text{action } a} V(\underbrace{next(s,a)}_{\text{the state that results from action } a \text{ in state } s})$$

For nonterminal random event states

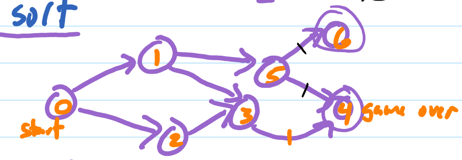
$$V(s) = \sum_{\text{outcome } \sigma} P(\sigma) \cdot V(\underbrace{next(s,\sigma)}_{\text{state resulting from outcome } \sigma \text{ in state } s})$$

```
for every terminal state s
    V[s] ← payout(s)
→ for every nonterminal state s  in order of topological sort
    if s is an action position
        max ← −∞
        argmax ← NIL
        for every action a
            v ← V[next(s,a)] + R(s,a)
            if v > max
                max ← v
                argmax ← a
        V[s] ← max
        OPT[s] ← argmax
    else
        v ← 0.0
        for every outcome σ
            v ← v + P(σ)·(V[next(s,σ)] + R(s,σ))
        V[s] ← v
```

R(s,a): reward for choosing a in state s

(for example, closest to end of game to furthest)

← 0 1 2 5 3 4 6



finite game = directed acyclic graph

topo sort = ordering of verts so edges all go →

Yahtzee

anchor : states at start of turn

component : states reachable from one position w/o going through another

number of anchors : score in each category

$2^{12} \cdot 64 \cdot 3 \approx \frac{3}{4}$ million TMI!

used/unused 1-6, 3k, ..., C

uppertotal

unused, 0, or 50

$\cdot$ 1600 nodes/comp

$\approx$ 1 minute

7^6 unused, 0x, ..., 5x for 1-6

28^2 unused, 0, 5, ..., 30 for 3k, 4k

27 unused, 5, ..., 30 for C

3^3 unused, 0, nonzero for LS, SS, FH

15 unused, 0, 50, ..., 1250 Y

$\approx$ 1 trillion anchors (components)

$\cdot$ 1600 nodes/component

$\approx$ 1.6 quadrillion nodes

@ 1 billion/sec

1.6 billion seconds

$\approx$ 50 years

modification. $V(s) =$

For nonterminal action positions

$$V(s) = \max_{\text{action } a} V(\text{next}(s,a)) + \text{score}(s,a)$$

↳ 0 if s in middle of turn
determined by final roll + cat chosen at end

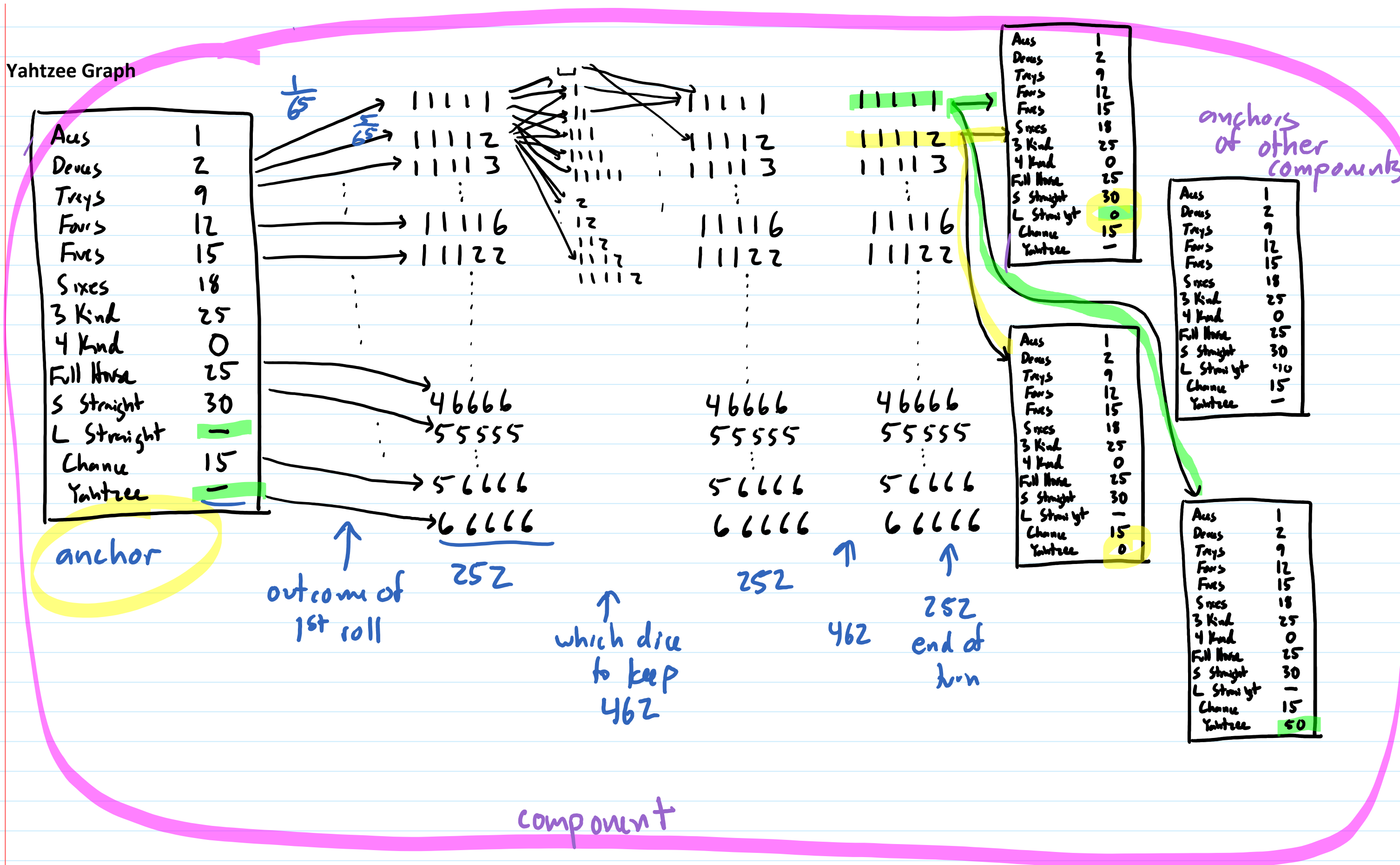
For nonterminal random event positions

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

↳ 0

anchors =

Yahtzee Graph



Hoot Owl Hoot

what captures state
how many states are there 1
how to simplify

what order to solve in

39 pos + nest
6 owls
14 sun pos

14 sun cards + 6 each of each color
6 colors

+ 2 rerolls

4K LS 1 1 1 2 2 ? FH or 1 1 1

LS 3 6 6 6 6 4K or 6 ?
end of turn