MAST/PAST

Tuesday, November 2, 2021 12:28 PM

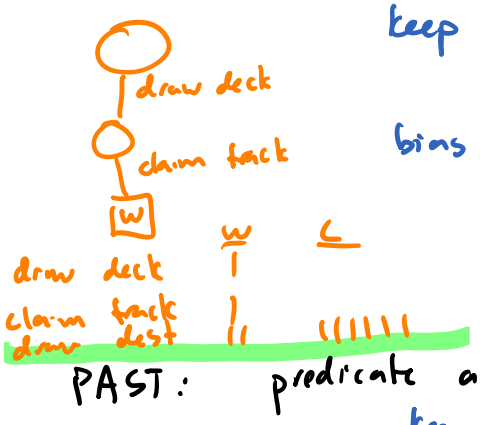
MAST : move averaging sampling technique

same move might be available at different points during play

keep global stats for each move

mean ↑ reward

bias default policy toward moves w/ better stats



PAST : predicate averaging ....

keep stats for each predicate

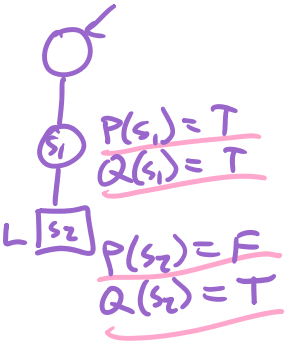
bias default policy toward moves that make predicates with good stats T

wins / traversals through states that make P(s) T

20 / ~~30~~ 31

Q(s) T

~~25~~ / ~~50~~ 52



when considering move to s_3

$P(s_3) = F$, $Q(s_3) = T$ $\frac{20}{31}$

vs

vs s_4

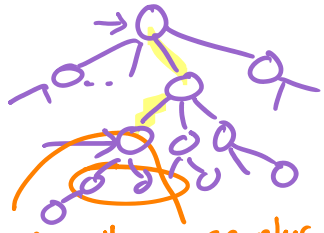
$P(s_4) = T$, $Q(s_4) = F$ $\frac{25}{52}$

bias toward s_3

Scout (principal variation search)

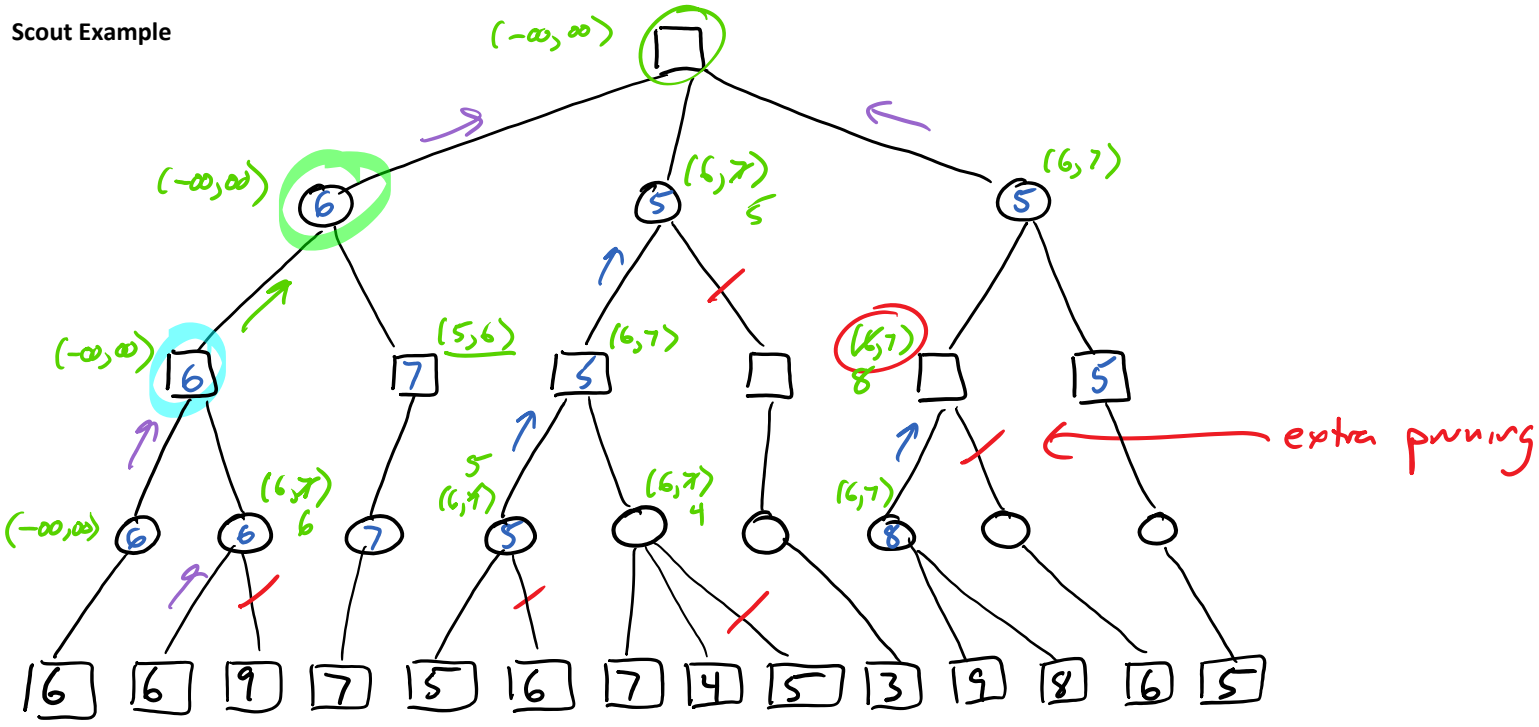
Alpha-beta wants to know and asks 1) is next child better than best so far, and if so, by how much?
(α, β)
result: $\leq \alpha \rightarrow$ NO
$> \alpha \rightarrow$ YES, and value tells you by how much

Scout wants to know and asks 1) is next child better than best so far?
if answer is N most of the time, can skip 2)
separate recursive calls
2) if answer to 1) is Y, how much better?

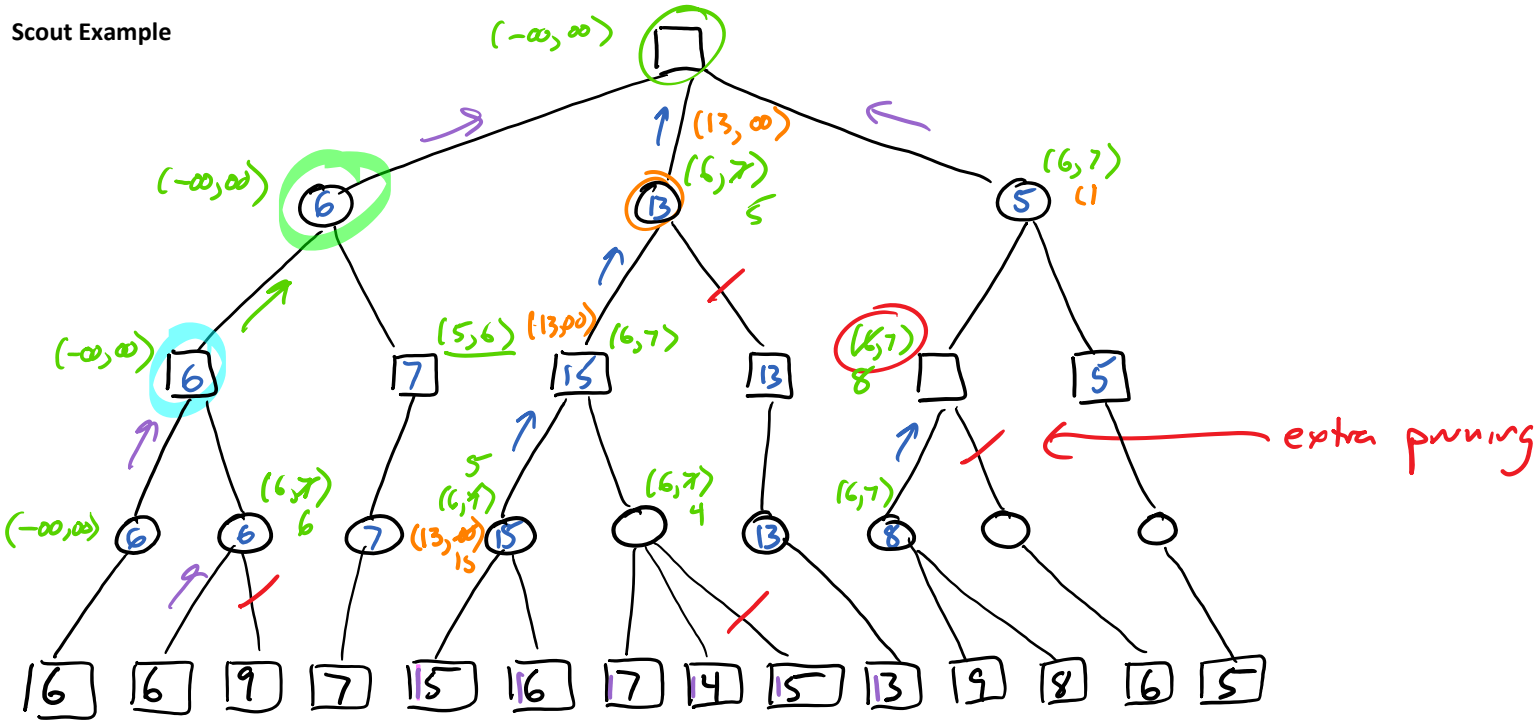

use results of shallower searches in transpo table
or heuristic
in rough order best... worst

```
Scout (p, α, β, h, depth)
    if p is terminal then return value(p)
    if depth = 0 then return h(p)
    if p is a max position
        for each reachable position p' and while α < β
            if p' is first child
                score ← Scout(p', α, β, h, depth-1)
            else
                score ← Scout(p', α, α+1, h, depth-1)      (9 means h(p') ≥ 9)
                if α < score < β                           (null window (for integer-valued h); faster b/c more pruning)
                    score ← Scout(p', score, β, h, depth-1)
            α ← max(α, score)
        return
    else
        ⋮
```

Scout Example



Scout Example



Scout Example