Lower bounds for Selection

A bound on the complexity of a problem is a bound on ***any*** algorithm that solves the problem

Technique:

Adversary Argument: Evil oracle dynamically decides outcomes of comparisons to thwart any algorithm.

See handout on Adversary arguments.

Lower bounds for Selection (cont)

Model of Computation **:** **Tournaments**

Keys (distinct values) ⟷ Players

Comparisons ⟷ Matches

Algorithms ⟷ Schedule of matches

Running time ⟷ Number of matches

Assumption**:** Outcomes of matches is consistent

Lower bounds for Selection (cont)

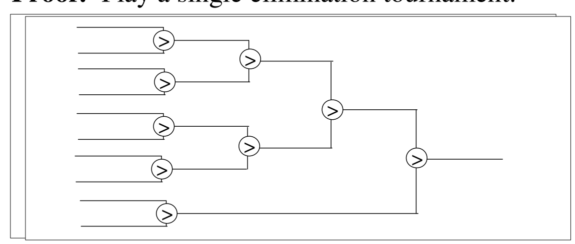
Problems	**Complexity**
Find the *kth* player out of *n*	$V_k(n)$
Find the top *k* players out of *n*	$U_k(n)$
Find the *1st, 2nd*,…, *kth* players out of *n*	$W_k(n)$

$$U_k(n) \le V_k(n) \le W_k(n)$$

Lower bounds for Selection (cont)

$V_1(n) = V_n(n) = n-1$

Proof: Play a single elimination tournament.



All players except champion lose once ⇒ *n-1* matches.

So $V_1(n) \le n-1$

Lower bounds for Selection (cont)

$V_1(n) = V_n(n) = n - 1$

Proof: (continued) To show that $V_1(n) \geq n - 1$ suppose you have played fewer than *n-1* matches. Then there are at least two players who have never lost and either could be the top player. So the top player can not be identified with fewer than *n-1* matches.

$V_1(n) = n - 1$ QED

Lower bounds for Selection (cont)

$V_2(n) = W_2(n)$

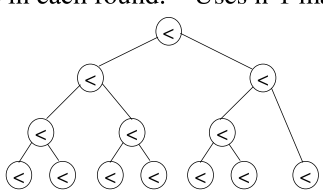
Proof: $V_2(n) \leq W_2(n)$ since if you know the 1st and 2nd players than you know the 2nd player. To show $V_2(n) \geq W_2(n)$ assume we know player B is 2nd. Since player B is not first it has lost a match, say to player A. If any player is better than A then B could not be second . So A must be the top player. QED

It is possible for $U_2(n) < V_2(n) = W_2(n)$.

Lower bounds for Selection (cont)

$W_2(n) = n + \lceil \lg n \rceil - 2$

Proof: Play a single elimination tournament among the n players with as many matches as possible in each round. Uses n-1 matches.

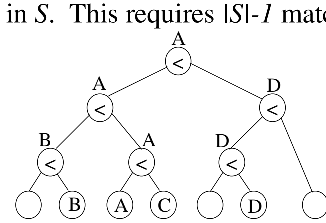


n=7

Lower bounds for Selection (cont)

$W_2(n) = n + \lceil \lg n \rceil - 2$

Proof: (continued) If the champion is A, collect all of the players that lost to A into a set *S*. Play a second single elimination tournament among the players in *S*. This requires |*S*|-*1* matches.



$S = \{B, C, D\}$

Lower bounds for Selection (cont)

$W_2(n) = n + \lceil \lg n \rceil - 2$

Proof: (continued) The champion A is the best player since everyone else has lost a match.

The players not in *S*, cannot be second since they have lost to a player who is not the best player.

Of the players in *S*, only the champion of the second tournament has not lost twice.

So the champion of the second tournament is the number 2 player.

Lower bounds for Selection (cont)

$W_2(n) = n + \lceil \lg n \rceil - 2$

Proof: (continued) The number of players in *S* is the number of matches played by A.

This is at most the height of the tree of the first tournament.

This tree has *n* leaves and height at most $\lceil \lg n \rceil$.

The total number of matches played is

$$n - 1 + |S| - 1 = n + \lceil \lg n \rceil - 2$$

So $W_2(n) \le n + \lceil \lg n \rceil - 2$

Lower bounds for Selection (cont)

$W_2(n) = n + \lceil \lg n \rceil - 2$

Proof: (continued) To show $W_2(n) \ge n + \lceil \lg n \rceil - 2$

we need to devise an oracle which will force this number of matches regardless of how the matches are organized.

Notation : $TOP = \{$currently undefeated players$\}$

$Win(x) =$ # of matches x has played and won against players that were (previously) undefeated

Lower bounds for Selection (cont)

$W_2(n) = n + \lceil \lg n \rceil - 2$

Proof: (continued)

$DOM(x, 0) = \{x\}$

$DOM(x, i) = \{y \mid y\text{'s first defeat was to a player in } DOM(x, i-1)\}$

$DOM(x) = \bigcup_{i=0}^{\infty} DOM(x, i)$

At start: $\forall x\ DOM(x) = \{x\}$

$TOP = \{$all players$\}$

$\forall x\ Win(x) = 0$

Lower bounds for Selection (cont)

$W_2(n) = n + \lceil \lg n \rceil - 2$ **Proof:** (continued)

Oracle decides matches as follows:

x wins if $x, y \in TOP$ and $Win(x) \ge Win(y)$
or if $x \in TOP$ and $y \notin TOP$
else whatever is consistent

Claim: After *m* matches are played, if *x* is still in *TOP*, then $|DOM(x)| \le 2^{Win(x)}$.

By induction on *m*.

Base: $m=0$ $|DOM(x)| = |\{x\}| = 1 = 2^0 = 2^{Win(x)}$

Lower bounds for Selection (cont)

$W_2(n) = n + \lceil \lg n \rceil - 2$ **Proof:** (continued)

Step: $m>0$. Assume claim holds after $m-1$ matches.

Case 1: *x* does not play in match *m*.

Then *DOM(x)* and *Win(x)* don't change.

Case 2: In match *m*, *x* plays $y \notin TOP$.

Then *DOM(x)* and *Win(x)* don't change.

Case 3: In match *m*, *x* plays $y \in TOP$.

$newDOM(x) \leftarrow DOM(x) \cup DOM(y)$

and $newWin(x) \leftarrow Win(x)+1$.

$|newDOM(x)| \le 2^{Win(x)} + 2^{Win(y)} \le 2^{Win(x)+1} = 2^{newWin(x)}$

Lower bounds for Selection (cont)

$W_2(n) = n + \lceil \lg n \rceil - 2$ **Proof:** (continued)

After m matches if $TOP=\{x\}$,

then $n = |DOM(x)| \le 2^{Win(x)} \Rightarrow \lceil \lg n \rceil \le Win(x)$

All but one of the losers to *x* must play a second match.

So *Win(x)-1* players lose at least twice.

$$\#\text{matches} \ge \underbrace{n-1}_{\text{1st losses}} + \underbrace{Win(x)-1}_{\text{2nd losses}} \ge n + \lceil \lg n \rceil - 2$$

QED

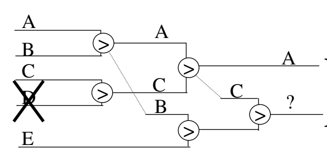
Lower bounds for Selection (cont)

Results:

$V_2(n) = W_2(n) = 5 + \lceil \lg 5 \rceil - 2 = 6$

$U_2(5) = 5$

It's possible to know the top 2 players out 5 without knowing the top player.



These two players are both better than 3 others and so are 1st or 2nd

Lower bounds for Selection (cont)

Finding first and last player requires at least $\left\lceil \frac{3n}{2} \right\rceil - 2$ matches.

See handout.