CS 124 Lecture 8

February 19, 2020

disjoint set

clever accounting!

MAKESET	$\log^*$
FIND	union by rank – smaller depth tree into larger depth tree
LINK	path compression
UNION	# operations × max time / operation

properties

1. if $v \neq p(v)$ then $\text{rank}(p(v)) > \text{rank}(v)$
2. whenever $p(v)$ is updated, $\text{rank}(p(v))$ increases
3. # of ECTs with rank $k \leq n/2^k$
4. # of ECTs with rank $\geq k$ is $\leq n/2^{k-1}$

MAX RANK $\rightarrow \log_2 n$

Groups

Groups	
0	$[0,1]$
1	$(1,2]$
2	$(2,4]$
3	$(4,16]$
4	$(16, 2^{16}]$

Group i → all elements w/ rank r w/ $\log^* r = i$ has form $(k, 2^k]$

Groups: $\log^*(\log n) + 1 = \log^* n$

of elements in group $(k, 2^k] \leq$?

$n/2^k$ by 4.

Type 1: if u, v are in different groups (or if v is root)

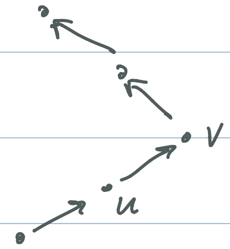
Easy, $\log^*(n)$ groups, m operations $\rightarrow O(m \log^* n)$

$u \rightarrow v$

Type 2: if u, v are in same group.

"charge" work to nodes

Node in group $[k, 2^k]$, give it 2^k tokens.



Work for group i:

- Why are tokens enough? Because ② says ranks increase on update and group $[k, 2^k]$ at most 2^k ranks.
- Total work: each set in $[k, 2^k]$ has 2^k tokens, $\le n/2^k$ sets $\rightarrow$ n tokens / group

$O((m+n) \cdot \alpha(n))$

Overview of Next Lectures

Greedy

Divide and Conquer

Dynamic Programming

Greedy

Satisfiability   2SAT!

$(x \vee \bar{y} \vee \bar{z} \vee \bar{w}) \wedge (\bar{x} \vee \bar{y} \vee \bar{w}) \wedge (\bar{y} \vee \bar{z} \vee w) \wedge (\bar{x} \vee y) \wedge (x) \wedge (\bar{z})$
$\wedge (\bar{x} \vee \bar{y} \vee w)$

Horn Formulae: each clause has ≤ 1 positive literal

Pure negative clauses

Implication: $y \wedge z \wedge w \rightarrow x = 1$ by 1st clause

$$X \wedge Z \to W$$
$$X \to y$$
$$\to X$$
$$X \wedge y \to W$$

Greedy Alg:

start w/ all false

while ∃ unsatisfied implication:

make implied variable true

check all pure negative are all still true.

Pf. Inductively set vars to true when I have to.

<u>Set Cover</u>

$X = \{X_1, \dots, X_n\}$ $\{X_1, X_3, X_5, X_7\}$

S = subsets of X $\{X_2, X_3, X_5, X_7\}$

$\bigcup_{S \in \mathcal{S}} S = X$ $\{X_4, X_6, X_{12}, \dots\}$

Find subcollection $\mathcal{I}$ of smallest size so $\bigcup_{T \in \mathcal{I}} T = X$

Greedy Alg: not always correct

pick set that covers largest # of uncovered elements

repeat

If the optimal set cover is k, this returns an answer of at most $k \log n$.

Y_i = # of still uncovered sets after i rounds

$|Y_0| = n$

$|Y_1| \le n - n/k = n(1-1/k) = |Y_0|(1-1/k)$ $\quad 1+x \le e^x$

$|Y_2| \le |Y_1|(1-1/k)$ $\quad 1-1/k \le e^{-1/k}$

$|Y_i| \le (1-1/k)^2$ $\quad |Y_0| = (1-1/k)^2 n < e^{-2k} \cdot n$

< 1

Huffman Coding

A - 00	70 million	0	prefix properting
C - 01	3 million	111	
G - 10	20 million	110	
T - 11	37 million	10	

↓ 260 million bits

↓ 213 million bits