CS 173 Lecture 15c: Induction on Grammar Trees

G:

$$S \to SS \mid Sac \mid ca \mid ba$$

Define for a parse tree T of G,

$a(T)$ = # of a's in T (leaf)

$b(T)$ = b's

$c(T)$ = c's

Claim: For all (grammar) parse trees T of G,

$$a(T) = b(T) + c(T)$$

Equivalently: in all strings generated by G, # a's = # b's + # c's.

Proof: by induction on height of the parse tree

every subtree of a parse tree is a parse tree

base case: $h = 1$: no parse trees of height 0

root: symbol
leaves: terminals
→ root not a leaf!

When $h = 1$, there are two possibilities:

$$\begin{array}{ccc} S & \quad & S \\ | & & | \\ ca & & ba \end{array}$$

$1 = 0 + 1$ ✓ → ca

ba ← $1 = 1 + 0$ ✓

in both cases, $a(T) = b(T) + c(T)$

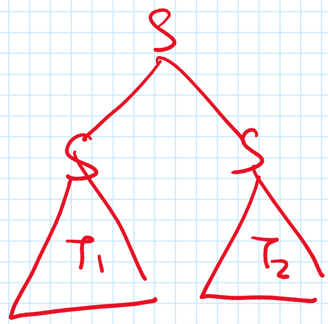
IH: assume $a(T) = b(T) + c(T)$ for all parse trees T of G w/ height $< h$.

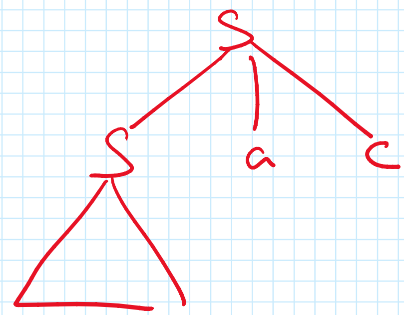
IS: Let T be a parse tree of G, of height $h \geq 2$.

The root is S.

There are two possibilities:

(i) $S \to SS$ or (ii) $S \to Sac$





In case i, let the two subtrees be called T_1 & T_2. Then

$$a(T) = a(T_1) + a(T_2)$$
$$b(T) = b(T_1) + b(T_2)$$
$$c(T) = c(T_1) + c(T_2).$$

heights of T_1 & T_2 are $< h$.

by I.H. $a(T_1) = b(T_1) + c(T_1)$

$a(T_2) = b(T_2) + c(T_2)$

So,
$$a(T) = a(T_1) + a(T_2)$$
$$= b(T_1) + c(T_1) + b(T_2) + c(T_2)$$
$$= b(T) + c(T).$$

In case ii, let the subtree be called T'.

$$a(T) = a(T') + 1$$
$$b(T) = b(T')$$
$$c(T) = c(T') + 1$$

height of T' $< h$.

By I.H. $a(T') = b(T') + c(T')$.

By I.H. $a(T') = b(T') + c(T')$.

So $a(T) = a(T') + 1$

$= b(T') + c(T') + 1$

$= b(T) + c(T)$

———— X ————

$\mathcal{H}$

$$A \to AaaA \mid \varepsilon.$$

Claim: $\mathcal{H}$ generates even length strings.

Proof: By induction on height of a parse tree of $\mathcal{H}$.

Base Case: $h = 1$.

The only possibility is $A \to \varepsilon$.

and $\text{length}(\varepsilon) = 0$, which is even.

IH. Assume for $h \geq 2$ that for all parse trees of H w/ height less than h, that the tree contains even number of terminals.

IS. Assume T is a parse tree of H of height $h \geq 2$.

Then the children of the root are labeled A, a, a, A.

The two subtrees have height $< h$, so each contain an even # of terminals, say $2p$ & $2q$

So. # of terminals in T is $2p + 2 + 2q$ which is even.

$\square$

which is even.

$\square$