CSE 396 Lecture Tue 3/26 Spring 2019

Top Hat #2233

Suppose G is a context-free grammar that is trying to meet some external specification $E \subseteq \Sigma^*$.

- G is <u>sound</u> for E if $L(G) \subseteq E$.
- G is <u>comprehensive</u> for E if $L(G) \supseteq E$.

(Logical term is "complete.")

- G is <u>correct</u> if $L(G) = E$, i.e., G is both sound and comprehensive.

> The term "sound" comes from logic, in which <u>proofs</u> are a more general kind of <u>derivation</u> and a proof system is sound if every theorem it derives is actually true.

<u>E</u>xample: $\Sigma = \{(,)\}$, E = the language of <u>nonempty</u> balanced-paren strings.

$G = S \rightarrow (S)S \mid ()$ "<u>Structural</u> <u>Induction</u>" (SI) for CFGs

Is G sound? A particular technique for seeing this kind of thing.

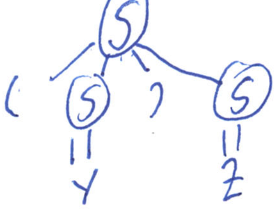
<u>Method</u> (Proof Script)

- Assign to each variable A a <u>property</u> P_A. $P_S \equiv$ "Every string x that I derive is balanced and nonempty."
- P_S should imply membership in E.
- Go through each rule $A \rightarrow X$ and show that if every variable B on the R.H.S. derives a string u (or v...) that obeys P_B then the whole resulting string obeys P_A. Then we can deduce $L(G) \subseteq E$ <u>by SI</u>.

(B can be A itself or S too.)

<u>$S \rightarrow ()$</u>: Suppose $S \Rightarrow^* x$ using this rule first (<u>utrf</u>.) (utrf = production). Then $x = ()$ which is balanced and nonempty. This <u>upholds</u> P_S on the L.H.S.

<u>$S \rightarrow (S)S$</u>: Suppose $S \Rightarrow^* x$ <u>utrf</u>. Then <u>$x =: (y)z$</u> <u>where</u> $S \Rightarrow^* y$ and $S \Rightarrow^* z$. By IH P_S on the R.H.S. (twice), <u>y and z are both balanced and $\neq \varepsilon$</u>. Then <u>$x = (y)z$ is balanced</u> because <u>the ')' shown matches the '('</u> <u>and y and z are individually balanced</u>. And $x \neq \varepsilon$ clearly. $\therefore P_S$ on LHS.



yields $(y)z$. 0 1 1 0 0

Since we upheld all rules, $L(G) \subseteq E$ by SI. ☒

Is G comprehensive? $\boxed{S \to (S)S \mid ()}$ No: G cannot derive $x = ()()$.

☆ One {should not try / cannot} to use SI to prove comprehensiveness, but can use SI to disprove it.

In fact, S obeys the stronger property $P'_S \equiv$ Every string x I derive is in E and either equals $()$ or has nesting.

By IH P_S on RHS, y is also nonempty as well as balanced. Hence $x = (y)z$ has nesting. Thus we get $L(G) \subseteq E' \cup \{\text{"()"}\}$ where $E' = \{$balanced x: $x = \text{"()"}$ or x has nesting (hence $x \neq \varepsilon$ either way)$\}$.

Since clearly "()()" $\notin E'$, $E' \subsetneq E$, so $L(G) \subsetneq E$. Hence G is not correct either.

Study exercise: Show that G cannot derive $y =$ "(())" (this "fully-nested" string) either. This leads to the question: Can we expand G, keeping it sound, by adding rules so it becomes comprehensive? Try adding:

$S \to (S)$ Clearly sound but does not help us derive x.

$S \to ()S$ Also sound, helps give x but not y.

Adding both rules makes the resulting G' comprehensive. Generally hard to prove.

Second example of Soundness: $E = \{x \in \{a,b\}^* : \#a(x) = \#b(x)\}$.

G_2: $S \to SS \mid aB \mid bA \mid \varepsilon$ P_S: Every x I derive is in E.

$A \to aS \mid bAA$ $P_A = \{$every y st $A \Rightarrow^* y$ has 1 more a than $b\}$

$B \to bS \mid aBB$ $P_B = \{$every z st $B \Rightarrow^* z$ has 1 more b than $a\}$

Proof of soundness will be included Thursday, and comprehensiveness sketched.

As for why $G' = S \to (S)S \mid (S) \mid ()S \mid ()$ is comprehensive for E, we will get that as a consequence of the Chomsky NF conversion for $G_0 = S \to (S)S \mid \varepsilon$