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

- \( G \) is **sound** for \( E \) if \( L(G) \subseteq E \).
- \( G \) is **comprehensive** for \( E \) if \( L(G) \supseteq E \).

(logical term is "complete")

- \( G \) is **correct** if \( L(G) = E \), i.e., \( G \) is both sound and comprehensive.

**Example:** \( \Sigma = \{ (, ) \} \), \( E = \) the language of nonempty balanced parentheses strings.

\[ G = S \rightarrow (S)S \mid ( ) \]

**"Structural Induction"** (SI) for CFGs.

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

**Method**

- Assign to each variable \( A \) a property \( P_A \). \( P_S \equiv \) "Every string that I derive is balanced and nonempty?"

**Proof Script**

- \( P_S \) should imply membership in \( E \).
- Go through each rule \( A \rightarrow X \) and show that if every variable \( B \) on the L.H.S. derives a string \( u \) (or \( v \)) that obeys \( P_B \) then the whole resulting string obeys \( P_A \). Then we can deduce production (\( \sigma \)) for \( L(G) \subseteq E \) by SI.

\[ S \rightarrow ( ) : \]

Suppose \( S \Rightarrow \alpha X \) using this rule first (\( \sigma \)). Then \( X = ( ) \) which is balanced and nonempty. This **upholds** \( P_S \) on the L.H.S.

\[ S \rightarrow (S)S : \]

Suppose \( S \Rightarrow \alpha X \) \( \sigma \). Then \( X = (y)Z \) where \( S \Rightarrow \alpha y \) and \( S \Rightarrow \beta Z \). By IH \( P_S \) on the R.H.S. (twice), \( y \) and \( Z \) are both balanced and \( \neq \).

Then \( X = (y)Z \) is balanced because the 'i' shan makes the 'c'

\[ L( ) \Rightarrow ( ) \]

**yields** and \( y \) and \( Z \) are individually balanced. And \( X \neq \) clearly. \( \Rightarrow P_S \) on LHS since we uphold all rules, \( L(\gamma) \subseteq E \) by SI.
Is $G$ comprehensive? [S $\rightarrow$ (S)(S)] No. $G$ cannot derive $x = (x)$. 

One should not try to use S to prove comprehensiveness, but can use S to disprove it.

In fact, S obeys the stronger property $P_s' = \forall x \exists y : x \in E' \land (x = y \lor y \in E')$.

By Th $P_s$ on RHS, $x$ is also nonempty as well as balanced. Hence $x = (y)z$ has nesting. Thus we get $L(\alpha) \subseteq E' \cup \{''(\alpha)''\}$.

where $E' = \{\text{balanced } x : x = (x) \lor x \text{ has nesting (hence } x \notin E \text{ either way)}\}$.

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

Exercise: Show that $G$ cannot derive "(C)" (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:

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

$S \rightarrow (S)(S)$ Also sound, helps give $x$ but not $\exists x$. Generally, adding both rules makes the resulting $G'$ comprehensive, hard to prove.

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

$G_2: S \rightarrow SS | aB | bA | S | aS | bAA | bS | aBB | bB | aS | bS | aB | bB \theta$

$P_\theta$: Every $x \in E$ is in $E$.

$P_A = \exists y \forall x : a = y$ has 1 more.

$P_B = \exists x \forall y : B \neq x$ has 1 more by thumb.

Proof of soundness will be included Thursday and comprehensiveness sketched.

As for why $G' = S \rightarrow (S)(S) | (S)(S)(S)(S)$ is comprehensive for $E$, we will get that as a consequence of the Church-Ng conversion for $G_0 = S \rightarrow (S)(S) \in E$. 

$\equiv$