Then you may conclude $L(G) \subseteq E$. What you are actually doing is shorthand for a proof by induction on the <u>length n</u> of derivations in the grammar, with property $P(n)$ = "For each variable A, if A derives a terminal string x within n steps, then x satisfies P_A." The <u>basis</u>, with $n=1$, involves all productions $A \to X$ with $X \in \Sigma^*$. The induction step handles each other production $A \to RHS$ as a separate case. In longhand, if RHS = (say) BaC, it would go:

> Suppose $A \Rightarrow^* x$ by a derivation of <u>n</u> steps starting with the production $A \to BaC$. Then $x = yaz$ where y and z are terminal strings such that $B \Rightarrow^* y$ and $C \Rightarrow^* z$. The numbers m_1 of steps in the $B \Rightarrow^* y$ derivation and m_2 of steps in the $C \Rightarrow^* z$ derivation must sum to $n-1$, so both are $< n$. By IH $P(m_1)$, y satisfies P_B, while by IH $P(m_2)$, z satisfies P_C. Analytically, we can show that the string $x = yaz$ must then satisfy P_A [give details here], and this shows $P(n)$ in this case.

SI, however, takes away the "clunky" references to "n" and is much shorter:

> $A \to BaC$: Suppose $A \Rightarrow^* x$ using this production first (utpf). Then $x = yaz$ where $B \Rightarrow^* y$ and $C \Rightarrow^* z$. By IH P_B, P_C on RHS, y satisfies [text of P_B] and z satisfies [text of P_C]. Hence $x = yaz$ satisfies [you say it], which implies P_A on LHS.

This is AOK even for self-recursive productions. Consider $G = S \to aSb \mid \epsilon$ and $E = \{a^n b^n : n \geq 0\}$. Take P_S = "I derive strings x s.t. $x \in E$."

<u>Basis</u>: $S \to \epsilon$: $\epsilon \in E$ so P_S holds on LHS in this case.

<u>Induction</u>: $S \to aSb$: Suppose $S \Rightarrow^* x$ utpf. Then $x = ayb$ where $S \Rightarrow^* y$ on RHS. By IH P_S on RHS, $y \in E$, so y has the form $y = a^m b^m$ for some $m \geq 0$. Then $x = a \cdot a^m b^m \cdot b = a^{m+1} \cdot b^{m+1}$, which is in E, so P_S holds on LHS. No more productions, so <u>done</u>!

To be continued, "$E \subseteq L(G)$" parts next.

closer to text,