COMS W3261: Computer Science Theory
Example of How to Prove that a Grammar Generates a Purported Language
by Sudip Das with some edits by Zeph Grunschlag

Consider the language of bitstrings:

$$L = \{(0 \cup 1)^m 0 (0 \cup 1)^m (0 \cup 1)^n 0 (0 \cup 1)^n | m, n \geq 0\}$$

Consider the grammar G:

$$\begin{aligned} S &\rightarrow ZZ \\ Z &\rightarrow 0|AZA \\ A &\rightarrow 0|1 \end{aligned}$$

Prove that G generates the language L:

To prove that G generates L, i.e. $L(G) = L$, you have to show two things:
1. Prove $L(G) \subset L$, i.e. all strings derivable from G are in L
2. Prove $L \subset L(G)$, i.e. every string in L has a derivation in G

Notice that the first production of G generates a concatenation of two Z's, each of which will generate a string. Also notice that strings in L are the concatenation of two strings in the language $L' = \{(0|1)^i 0 (0|1)^i\}$, i.e. $L = L'L'$. Thus it is enough to show that $L(Z) = L'$ where $L(Z)$ denotes the set of terminal strings generated from the variable Z.

Lemma 0.1. $L(Z) \subset L'$

Proof. For the base case, there is one string with derivation of length 1, that is the string 0. It is generated by the derivation $Z \Rightarrow 0$. The string 0 is in L' because it is of the form $(0|1)^k 0 (0|1)^k$ with $k = 0$.

Now make the inductive hypothesis: If $Z \Rightarrow^* w$ in less than n steps, then $w \in L'$.

Let w be a string such that $Z \Rightarrow^* w$ in n steps. Since the derivation has length greater than 1, it must use the production $Z \Rightarrow AZA$. The inner Z generates a string w' in less than n steps. By the inductive hypothesis, $w' \in L'$, i.e. $w' = (0|1)^k 0 (0|1)^k$ for some k. Since each A produces a 0 or 1, the result is that $Z \Rightarrow AZA \Rightarrow (0|1)w'(0|1) = (0|1)^{k+1} 0 (0|1)^{k+1}$. Therefore $w \in L'$. □

Lemma 0.2. $L' \subset L(Z)$

Proof. For the base case, the shortest string in L' is the string 0. This is generated by $Z \Rightarrow 0$.

Now make the inductive hypothesis: If $|w| \leq n$ and $w \in L'$, then $Z \Rightarrow w$.

Let w be any string of the form $(0|1)^{n/2} 0 (0|1)^{n/2}$ (here $n/2$ denotes integer division). This string can be broken up into $(0|1)(0|1)^{n/2-1} 0 (0|1)^{n/2-1} (0|1)$. The inner string w' of the form $(0|1)^{n/2-1} 0 (0|1)^{n/2-1}$ has length at most $n - 1$ and is

of the form $(0|1)^k 0 (0|1)^k$, i.e. $w' \in L'$. Thus we may apply the induction hypothesis and assume that w' is generated by Z, i.e. $Z \Rightarrow^* w'$. So $Z \Rightarrow AZA \Rightarrow^* (0|1)w'(0|1) = w$. Thus, we have extended the inductive hypothesis to the next higher length and proved the theorem.

□