CSCI 383 Fall 2011	**Proof that a Language is Context-Free**

Main Steps

The main steps for proving a language A is context-free is as follows:

Step 1: Define a context-free grammar. Define a 4-tuple $G = (N, \Sigma, P, S)$ that represents your context-free grammar.

Step 2: Show that $L(G) \subseteq A$. To show that $L(G) \subseteq A$, you want to show that $x \in L(G) \Rightarrow x \in A$. This usually involves proving the following lemma, by induction on i (the number of productions in the derivation).

Lemma 1. *For all $x \in \Sigma^*$ and $i \geq 1$, $S \xrightarrow{i} x \Rightarrow x \in A$.*

In proving your lemma, your base case is usually $i = 1$.

Step 3: Show that $A \subseteq L(G)$. To show that $A \subseteq L(G)$, you want to show that $x \in A \Rightarrow x \in L(G)$. This usually involves proving the following lemma, by induction on the structure of your string x, or something like that...

Lemma 2. *For all $x \in \Sigma^*$, $x \in A \Rightarrow S \xrightarrow{*} x$.*

In proving your lemma, your base cases are usually the smallest strings in the language A.

Comments

- If your CFG has multiple non-terminals defined, you may need to do several induction proofs, one per non-terminal.

Example

Show that the language $P = \{x \in \{a,b\}^* : x = \mathsf{rev}(x)\} = \mathsf{palindromes}(\{a,b\})$ is context-free.

We start by defining a context-free grammar $G = (N, \Sigma, P, S)$ such that $L(G) = P$:

$$S \quad \to \quad aSa \mid bSb \mid a \mid b \mid \varepsilon \ .$$

Namely, $N = \{S\}$, $\Sigma = \{a,b\}$, $P = \{(S, aSa), (S, bSb), (S, a), (S, b), (S, \varepsilon)\}$ and $S = S$. The following two lemmas will combine to show that $L(G) = P$.

Lemma 3. *For all $x \in \Sigma^*$ and $i \geq 1$, $S \stackrel{i}{\to} x \Rightarrow x \in P$.*

Proof. We prove the lemma by induction on i. As a base case, consider $i = 1$. Then either

- $S \stackrel{1}{\to} a$. Then $x = a$ and $a = \mathsf{rev}(a)$ and $a \in P$.
- $S \stackrel{1}{\to} b$. Then $x = b$ and $b = \mathsf{rev}(b)$ and $b \in P$.
- $S \stackrel{1}{\to} \varepsilon$. Then $x = \varepsilon$ and $\varepsilon = \mathsf{rev}(\varepsilon)$ and $\varepsilon \in P$.

For the inductive step, suppose the claim is true for some $i \geq 1$. That is, suppose we know that for any $x \in \Sigma^*$, if $S \stackrel{i}{\to} x$ then $x \in P$. Now consider a $y \in \Sigma^*$ such that $S \stackrel{i+1}{\to} y$; we want to show that $y \in P$. There are two cases:

- $S \stackrel{1}{\to} aSa \stackrel{i}{\to} y$. Then $y = axa$ for some $x \in \Sigma^*$ with $S \stackrel{i}{\to} x$. By the induction hypothesis, $x \in P$ (that is, $x = \mathsf{rev}(x)$); therefore, $\mathsf{rev}(y) = \mathsf{rev}(axa) = a\mathsf{rev}(x)a = axa = y$, and $y \in P$.
- $S \stackrel{1}{\to} bSb \stackrel{i}{\to} y$. Then $y = bxb$ for some $x \in \Sigma^*$ with $S \stackrel{i}{\to} x$. By the induction hypothesis, $x \in P$ (that is, $x = \mathsf{rev}(x)$); therefore, $\mathsf{rev}(y) = \mathsf{rev}(bxb) = b\mathsf{rev}(x)b = bxb = y$, and $y \in P$.

This concludes the proof of the lemma. □

Lemma 4. *For all $x \in \Sigma^*$, $x \in P \Rightarrow S \stackrel{*}{\to} x$.*

Proof. We prove the lemma by induction on x. As a base case, consider x as ε, a, or b. Then $S \stackrel{1}{\to} \varepsilon$, $S \stackrel{1}{\to} a$, and $S \stackrel{1}{\to} b$ are derivations for x in each of these cases, respectively.

For the inductive step, suppose the claim is true for some $x \in P$. Now consider the following $y \in P$:

- $y = axa$. Then $S \stackrel{1}{\to} aSa \stackrel{*}{\to} axa$ is a derivation for $y = axa$, where the last step follows by the induction hypothesis.
- $y = bxb$. Then $S \stackrel{1}{\to} bSb \stackrel{*}{\to} bxb$ is a derivation for $y = bxb$, where the last step follows by the induction hypothesis.

□