Some CFL's that really require Proof
By William Gasarch

1 Introduction

We give CFG's for $\{w \mid n_a(w) = n_b(w)\}$ and $\{w \mid n_a(w) \neq n_b(w)\}$.

2 A CFG for $L_{a's=b's} = \{w \mid n_a(w) = n_b(w)\}$

Let G be the CFG:
$S \rightarrow aSb$
$S \rightarrow bSa$
$S \rightarrow SS$
$S \rightarrow e$

Theorem 2.1 $L_{a's=b's} = L(G)$.

Proof:
1) $L(G) \subseteq L_{a's=b's}$.
It is easy to show that if $S \Rightarrow \alpha$ (and α may have S's in it) then $n_a(\alpha) = n_b(\alpha)$ by an induction on the number of steps in the derivation. Hence any string in $L(G)$ (so all terminals) is in $L_{a's=b's}$.

2) $L_{a's=b's} \subseteq L(G)$.
We proof this by induction on $|w|$. If $|w| = 0$ then use $S \rightarrow e$.
Assume that any string in $L_{a's=b's}$ of length $< n$ is in $L(G)$. Let $w \in L_{a's=b's}$ and $|w| = n$. We show $w \in L(G)$. There are two cases depending on if w begins with a or b. We do the w begins with a case. The other case is similar. Let $w = a\sigma_2 \cdots \sigma_n$.
Find the least i such that $a\sigma_2 \cdots \sigma_i \in L_{a's=b's}$.
If $i = n$ then $\sigma_n = b$ so $w = aub$ where $u \in L_{a's=b's}$. Clearly $|u| < |w|$ so $u \in L(G)$ inductively. Use $S \rightarrow aSb$ and then $S \Rightarrow u$ to get $w = aub \in L(G)$.
If $i < n$ then let $u_1 = a\sigma_2 \cdots \sigma_i$ and $u_2 = \sigma_{i+1} \cdots \sigma_n$. Since $u_1, u_2 \in L$, $|u_1| < |w|$, and $|u_2| < w$, both u_1, u_2 are in $L(G)$ inductively. Use $S \rightarrow SS$ and $S \Rightarrow u_1$ and $S \Rightarrow u_2$ to get $S \Rightarrow u_1u_2 = w$. ∎

3 A CFG for $L_{a's\neq b's} = \{w \mid n_a(w) \neq n_b(w)\}$

Note that

$$L_{a's\neq b's} = \{w \mid n_a(w) < n_b(w)\} \cup \{w \mid n_a(w) > n_b(w)\}$$

We show that

$$L_{a's<b's} = \{w \mid n_b(w) < n_a(w)\}$$

is a CFL. The proof that

$$L_{a's>b's} = \{w \mid n_b(w) > n_a(w)\}$$

is a CFL s similar. Then we just take the union.

Let G be:

$S \to aT$

$S \to aS$

$S \to bSS$

$T \to aTb$

$T \to bTa$

$T \to TT$

$T \to e$

Note that the last four rules are our grammar for $\{w \mid n_a(w) = n_b(w)\}$ from the last section. Hence we have the following:

Lemma 3.1 $\{w \mid T \Rightarrow w\} = L_{a's=b's}$.

Theorem 3.2 $L(G) = L_{a's<b's}$.

Proof:

1) $L(G) \subseteq L_{a's<b's}$.

We do this by the length of the derivation. If $S \Rightarrow w$ with a derivation of length 1 then the lemma is true vacously. If $S \Rightarrow w$ with a derivation of length 2 then the only one possible is $S \to aT$ and then $T \to e$ which yields $S \Rightarrow a$.

Assume that if $S \Rightarrow w'$ with a derivation of length $\leq n-1$ then $w' \in L_{a's<b's}$. Let $S \Rightarrow w$ via a derivation of length n. We look at what the first rule could be

If $S \to aT$ is the first rule then we don't need the ind hyp. We know that T yields strings $u \in L_{a's=b's}$ so $S \Rightarrow au \in L_{a's<b's}$.

If $S \to aS$ then note that $S \Rightarrow u$ by a derivation of length $\leq n-1$. Hence inductively $u \in L_{a's<b's}$. Clearly $au \in L_{a's<b's}$.

If $S \to bSS$ then the two S's both have derivations of length $\leq n-1$ to strings. Let those strings be u_1 and u_2. Inductively $u_1, u_2 \in L_{a's<b's}$. Hence $bu_1u_2 \in LL$.

2) $L_{a's<b's} \subseteq L(G)$.

We prove this by induction on $|w|$. Since $w \in L_{a's<b's}$, $|w| \geq 1$. If $|w| = 1$ then $w = a$ which is easily derived.

Assume all strings in $L_{a's<b's}$ of length $\leq n-1$ can be generated. Let $w \in L_{a's<b's}$ be of length n.

1. If $w = au$ where $u \in L_{a's=b's}$ then use $S \to aT$ and then $T \Rightarrow u$.
2. If $w = au$ where $u \in L_{a's<b's}$ then use $S \to aS$ and then $S \Rightarrow u$.
3. If $w = bu$ then note that $n_a(u) \geq n_b(u) + 2$. One can easily show that $u = u_1u_2$ where $u_1, u_2 \in L_{a's<b's}$. Use $S \to bSS$, $S \Rightarrow u_1$ and $S \Rightarrow u_2$.

∎