Properties of Context-Free Languages (Section 11.7)

Properties of Context-Free Languages

- This section has lots of information about how grammars can be transformed into particular normal forms. This is not relevant for anything that we are doing in this class.

C-F Pumping Lemma

- Again, there is one topic from this section that we will be covering. That topic is how we can use an extension of the *Pumping Lemma* to show that certain languages are not context-free.
- Pumping Lemma: If L is an infinite context-free language, then any grammar for L must be recursive, so there must be derivations of the following form where u, v, w, x, and y are terminal strings.
  - $S \Rightarrow^{+} uNy$
  - $N \Rightarrow^{+} vNx$ (where v and x are not both Λ)
  - $N \Rightarrow^{+} w$

Pumping Lemma continued

- These derivations lead to derivations like:
  - $S \Rightarrow^+ uNy \Rightarrow^+ uvNxy \Rightarrow^+ uv^2Nx^2y \Rightarrow^+ uv^kNx^ky \Rightarrow^+ uv^kwx^ky \in L$ for all $k \in \mathbb{N}$.
- This is the basis for the Pumping Lemma:
- There is an integer $m > 0$ such that if $z \in L$ and $|z| \geq m$, then z has the form $z = uvwxy$ where $1 \leq |vx| \leq |vwx| \leq m$ and $uv^kwx^ky \in L$ for all $k \in \mathbb{N}$.

C-F Pumping Lemma Example

- *Example:* The language $L = \{a^n b^n c^{n+k} | k, n \in \mathbb{N}\}$ is not context-free.
- *Proof:* Assume, BWOC (by way of contradiction), that L is context-free. L is infinite, so the pumping lemma applies. Choose $z = a^m b^m c^m = uvwxy$ where $1 \leq |vx| \leq |vwx| \leq m$ and $uv^k wx^k y \in L$ for all $k \in \mathbb{N}$. Observe neither v nor x can contain distinct letters. For example, if $v = \ldots a \ldots b \ldots$, then $v^2 = \ldots a \ldots b \ldots a \ldots b \ldots$, which can't appear as a substring of any string in L. So v and x must be strings of repeated occurrences of a single letter.

Example continued

- Now since $|vwx| \leq m$, there are two possible places in $a^m b^m c^m$ where v and x must occur:
  1. v and x occur in $a^m b^m$
  2. v and x occur in $b^m c^m$

  But we obtain the follow contractions because v and x are not both Λ:
  1. Let $k = 2$ to obtain $uv^2wx^2y = a^{m+i} b^{m+j} c^m$, where $i > 0$ or $j > 0$. So $uv^2wx^2y \notin L$
  2. Let $k = 0$ to obtain $uwy = a^m b^{m-i} c^{m-j}$, where $i > 0$ or $j > 0$. So $uwy \notin L$

  These contradictions imply that L is not context-free. QED

Another Proof

- Prove that the language $L = \{ss | s \in \{a, b\}^*\}$ is not context-free.
- *Proof:* Assume, BWOC, that L is context-free. L is infinite, so the pumping lemma applies. Choose $z = a^m b^m a^m b^m$ where m is the positive integer from the lemma. Then $z = a^m b^m a^m b^m = uvwxy$ where $1 \leq |vx| \leq |vwx| \leq m$ and $uv^k wx^k y \in L$ for all $k \in \ltimes$.
- Now, since $|vwx| \leq m$, there are three possible places in $a^m b^m a^m b^m$ where v and x must occur:
  1. v and x occur in $a^m b^m$ (on the left of z)
  2. v and x occur in $b^m a^m$ (in the center of z)
  3. v and x occur in $a^m b^m$ (on the right of z)
- Notice that v and x can consist only of repetitions of a single letter, for similar reasons to what we saw in the previous proof.

Proof continued

- We need to find a contradiction in each of the three cases. We'll do it by using $k = 0$. This tells us that $uwy \in L$. But we obtain the following contradictions because v and x are not both Λ.
  1. $uwy = a^{m-i}b^{m-j}a^m b^m$ where either $i > 0$ or $j > 0$. So $uwy \notin L$.
  2. $uwy = a^m b^{m-i}a^{m-j}b^m$ where either $i > 0$ or $j > 0$. So $uwy \notin L$.
  3. $uwy = a^m b^m a^{m-i}b^{m-j}$ where either $i > 0$ or $j > 0$. So $uwy \notin L$.
- Therefore L is not context-free. QED