Non-Context Free Languages

Like Regular Languages, Context-Free languages have limits. The limits of regular languages are clearly defined. They can't track an infinite number of states. They can't count beyond the number of states they have. The limits of context-free gramars/languages are not as easy to find. In short they can't look at their entire input in the same order twice.

Examples

$\{ww \mid w \in \Sigma^*\}$

$\{a^n b^n c^n \mid n \geq 0\}$

$\{a^{n^2} \mid n \geq 0\}$

$\{a + b = c \mid a, b, c \in \{0,1\}^*$ and the sum of the binary numbers represented by a and b equals c $\}$

$\{a^i b^j c^k \mid 0 \leq i \leq j \leq k\}$

So how do we tell a language isn't context-free?

Context-Free Pumping Lemma.

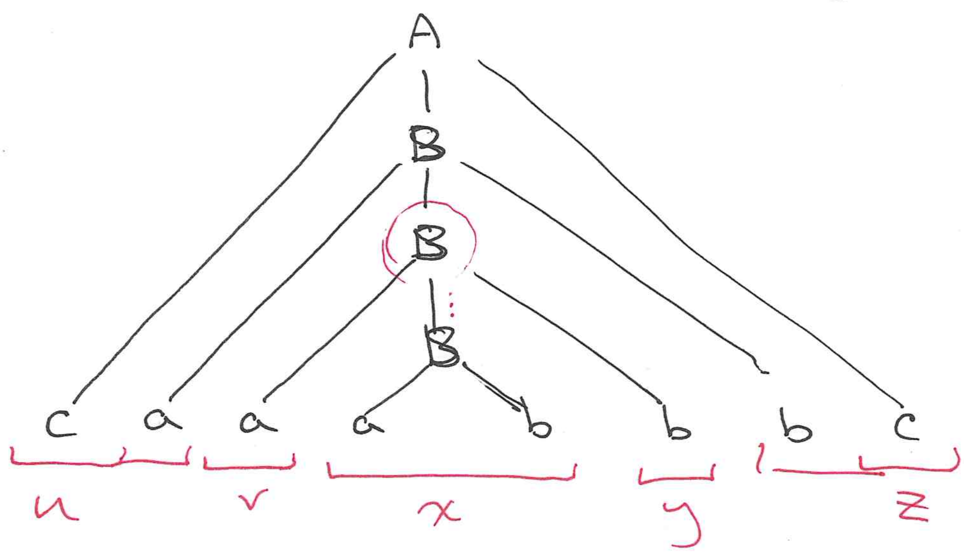
Context-Free Pumping Lemma

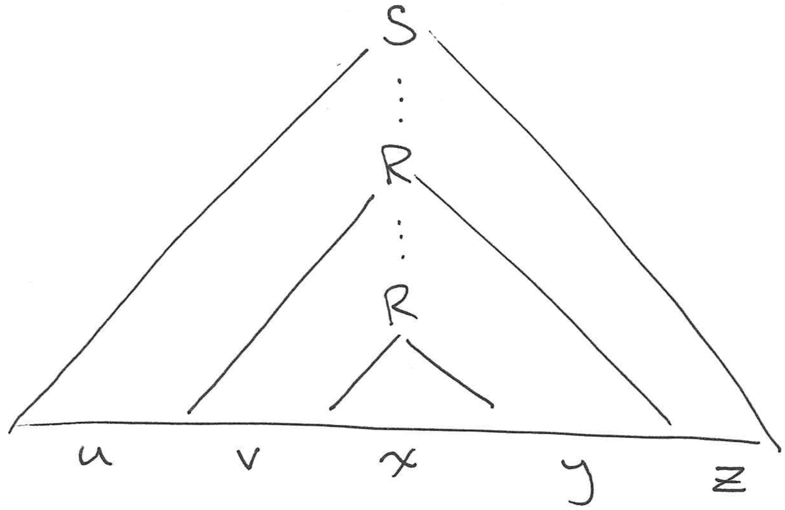
If A is a context-free language then there is a number p (the pumping length) where, if $s \in A$ and $|s| \geq p$, then s may be divided into 5 pieces $s = uvxyz$ satisfying the following conditions.

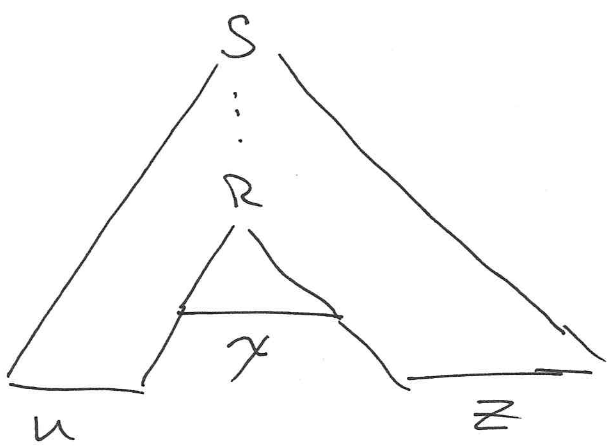
1. For each $i \geq 0$ $uv^i x y^i z \in A$
2. $|vy| > 0$
3. $|vxy| \leq p$

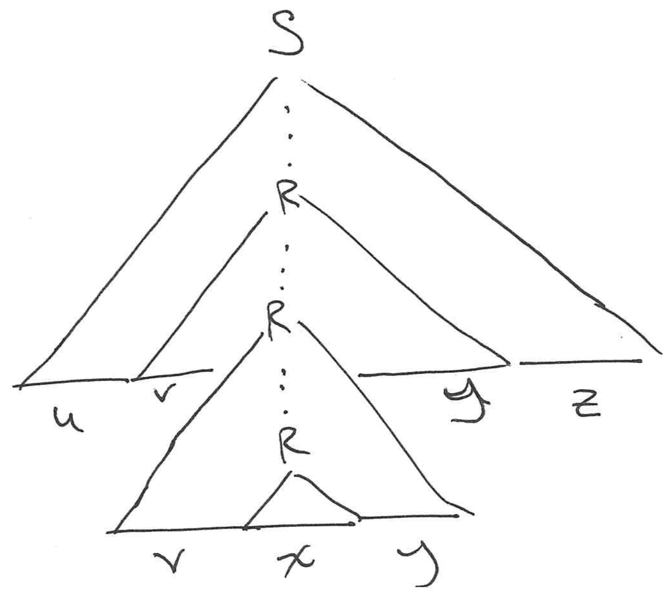
$A \rightarrow cBc$

$B \rightarrow aBb \mid ab$









So, what is p?

$b^h + 1$

b is the maximum branching factor

Theorem

The language $A = \{a^i b^j c^k \mid i \leq j \leq k\}$ is not context free.

Proof by Contradiction

- Assume A is a CFL. Let p be the pumping length given by the pumping lemma.
- Let $s = a^p b^p c^p$
- Since $|s| \geq p$ and $s \in A$, s can be divided into 5 pieces $s = uvxyz$ such that

$$|vy| > 0$$
$$|vxy| \leq p$$

- Consider 2 cases.
  1. either v or y contains more than one type of symbol. a & b or b & c
     - $uv^2xy^2z \notin A$ since it will contain some symbols out of order.
  2. Both v and y contain only one type of symbol.
     a) a's do not appear
        - $uv^0xy^0z = uxz \notin A$ because it contains the same number of a's and fewer b's or c's

     b) b's do not appear then either a's or c's must be in v, y
        - if a's appear $uv^2xy^2z \notin A$ (more a's than b's)
        - if c's appear $uv^0xy^0z \notin A$ (more b's than c's)

     c) c's do not appear
        - $uv^2xy^2z \notin A$ same number of c's but more a's or b's

Theorem

The language $B = \{ww \mid w \in \{0,1\}^*\}$ is not context free.

Proof by Contradiction

Assume B is a CFL. Let p be the pumping length.

$$s = 0^p 1 0^p 1$$

$$\overbrace{000\ldots0001}^{0^p1}\overbrace{0000\ldots0001}^{0^p1}$$

$u \quad v\,x\,y \quad z$

Let $s = 0^p 1^p 0^p 1^p$

$|vy| > 0$

$|vxy| \leq p$

1. all v and y contain a single character

2. v and y straddle the midpoint of one half.

3. vxy straddle the midpoint of s.