Context-free languages

$$L = \{ 0^n 1^n \mid n \in \mathbb{N} \}$$

00001111

```
P ( ) {
  do { write 0;
       P();
       write 1; }
  or { ε        }   }
```

$(0(0(0(0(\epsilon)1)1)..)$

$$S \rightarrow 0S1 \mid \varepsilon$$

S is the <u>smallest</u> language s.t.

- $\epsilon \in S$
- If $w \in S$ then $0w1 \in S$.

$$S = \{ 0^n 1^n \mid n \in \mathbb{N} \}.$$

$$S \rightarrow 0S1 \mid \epsilon$$

Derivation of 000111

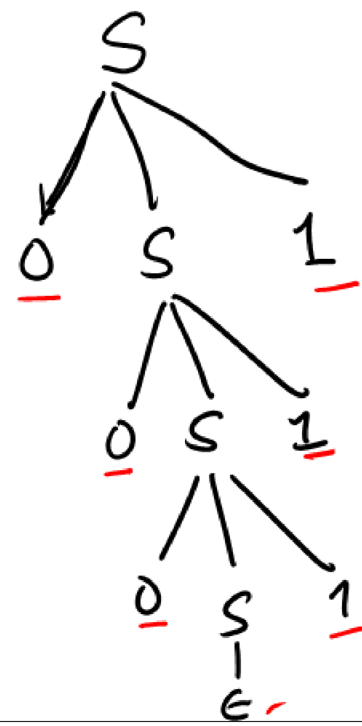
$$S \rightarrow \underline{0}S\underline{1} \rightarrow \underline{0}0S1\underline{1}$$

$$\rightarrow 000S111$$

$$\rightarrow 000\epsilon111 = 000111$$

$S \to 0S1 \mid \epsilon$

parse tree



$\Sigma = \{a, b\}$

$L =$ set of all palindromes over Σ

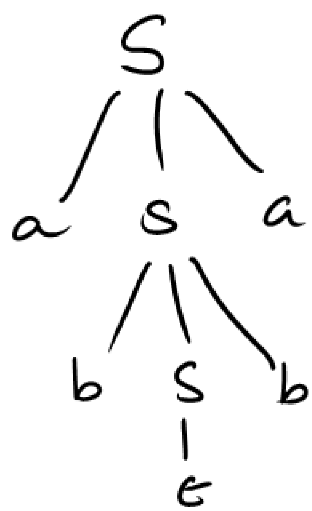
$= \{ w \mid w^R = w \}$

abbabbabba

If w is palindrome, awa & bwb are palindromes.

$S \rightarrow aSa \mid bSb \mid a \mid b \mid \epsilon$

abba



$S \rightarrow aSa$
$\rightarrow abSba$
$\rightarrow abba$

$S \rightarrow NP \; VP$

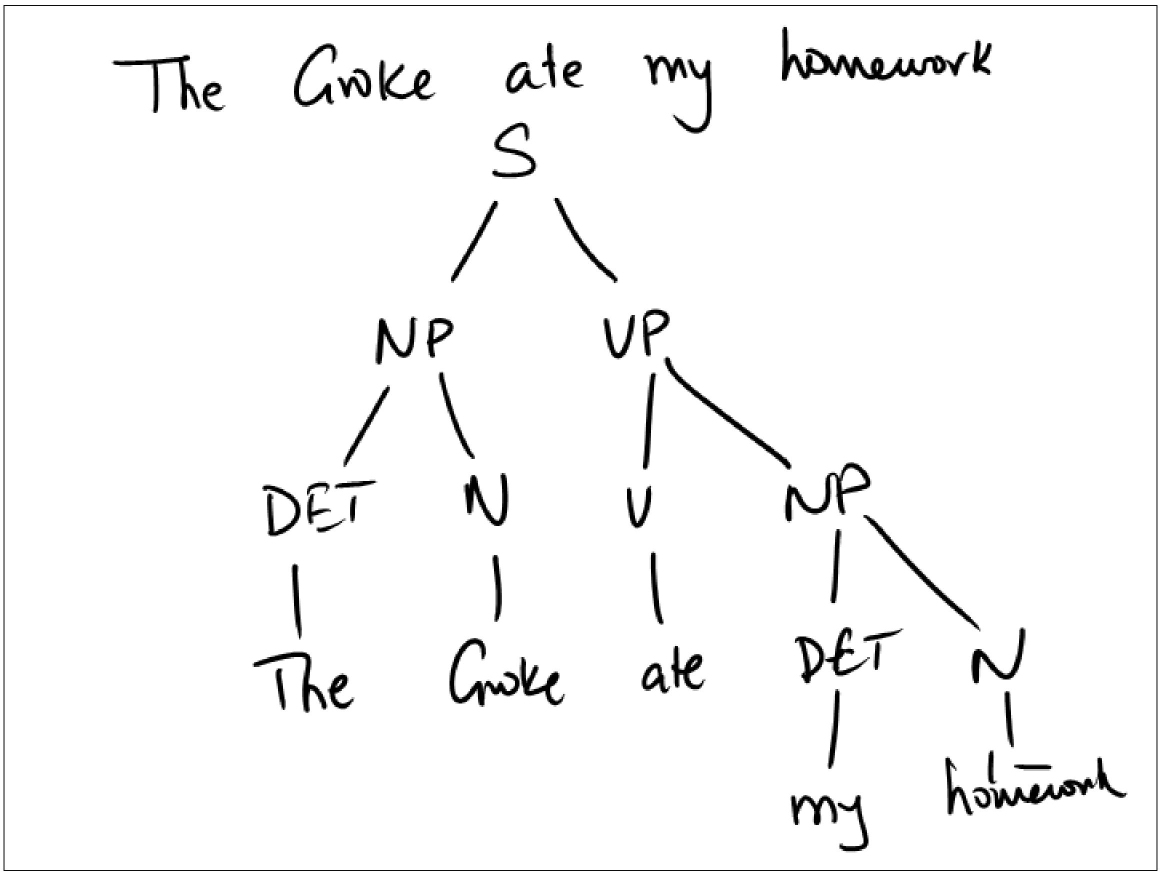
$NP \rightarrow DET \; N$

$VP \rightarrow V \; NP$

$N \rightarrow$ Groke | homework | lunch

$DET \rightarrow$ the | my

$V \rightarrow$ ate | corrected | washed ...


The Groke ate my homework
S
NP
VP
DET
N
V
NP
The
Groke
ate
DET
N
my
homework

Formal definition of a context-free grammar (CFG)

A CFG is a 4-tuple

$$(V, \Sigma, R, S)$$

where V is a finite set (called "variables")

Σ is a finite set ("the alphabet")

$S \in V$ - initial/start variable.

R is a finite set of rules of the form

$A \to w$ where $w \in (V \cup \Sigma)^*$ and $A \in V$

Eg. $S \to 0S1 \mid T$

$T \to aT \mid bT \mid \epsilon.$

$$\mathcal{L}(S) = \left\{ 0^n \, w \, 1^n \;\middle|\; \begin{array}{l} n \in \mathbb{N} \\ w \in \{a,b\}^* \end{array} \right\}$$

Formally $G = (V, \Sigma, R, S)$

$V = \{S, T\}$

$\Sigma = \{0, 1, a, b\}$

$R = \{S \to 0S1,\ S \to T,\ T \to aT,\ T \to bT,\ T \to \epsilon\}$

Let $x, y, w \in (V \cup \Sigma)^*$ and $A \in V$.
Defining "$\Rightarrow$" ("yields" relation)

$$xAy \Rightarrow xwy \quad \text{provided } A \to w \in R$$

If $x, y \in (V \cup \Sigma)^*$, then

$x \Rightarrow^* y$ if there is a sequence of words $z_1 \ldots z_n$

$$x = z_1 \Rightarrow z_2 \Rightarrow z_3 \Rightarrow \ldots \Rightarrow z_n = y$$

The language of a grammar:

$$G = (V, \Sigma, R, S)$$

is $\{ w \mid S \Rightarrow_G^* w \}$

$w \in \Sigma^*$

$$G \quad \begin{array}{l} S \rightarrow 0S1 \mid T \\ T \rightarrow aT \mid bT \mid \epsilon \end{array}$$

$$a\underline{S}b10T \Rightarrow_G a\underline{0S1}b10T$$

$$\Rightarrow a\underline{T}b10T$$

$$\Rightarrow aSb10\underline{aT}$$

$\Rightarrow^*$: $x \Rightarrow^* y$

iff $x \Rightarrow z_1 \Rightarrow z_2 \cdots \Rightarrow y$

$00a11 \in L(G)$

$S \Rightarrow^* 00a11$

$$
\begin{aligned}
S \Rightarrow 0S1 &\Rightarrow 00S11 \\
&\Rightarrow 00T11 \\
&\Rightarrow 00aT11 \\
&\Rightarrow 00a\epsilon 11 \\
&= 00a11
\end{aligned}
$$