Grammars

[Linz (6 ed) p. 20]

Grammars are formalisms for **defining** or **generating** or **specifying** languages.

Definition: A grammar, G, is a **quadruple**:

$$G = (N, T, S, P)$$

N is a finite set of **non-terminals**.
(***Note***: Linz uses '***S***' for ***N*** and calls these **variables**)

T is a finite set of **terminals** in Σ.

S is the **start symbol**, $S \in V$.

P is a finite set of **productions**.

(Assume $N \neq 0$, $T \neq \emptyset$, $N \cap T = \emptyset$).

All **productions** are of the form

$$x \rightarrow y$$

Where $x \in (N \cup T)^+$ and $y \in (N \cup T)^*$.

Productions are applied as follows:

Given a **string** $w = uxv$ and a **production** $x \to y$,
we say the production $x \to y$ **applies** to w to get a **new string**

$$z = uyv.$$

Write: $w \Rightarrow z$ by production $x \to y$.
We say: z is **derived from** w by the production $x \to y$.

Suppose:

$$w_1 \Rightarrow w_2 \Rightarrow \ldots \Rightarrow w_n$$

by grammar G
(i.e. each step is **derived** by some production of G).

Then we write

$$w_1 \stackrel{\star}{\Longrightarrow} w_n,$$

i.e. w_n is **derived from** w_1 by G.

Note: Always $\quad w \stackrel{\star}{\Longrightarrow} w \quad$ (in zero steps)

Definition: For any $G = (N, T, S, P)$, the **language generated** by G is:

$$L(G) = \{w \in T^* \mid S \overset{\star}{\Longrightarrow} w\}$$

If $w \in L(g)$, then:

$$S \Longrightarrow w_1 \Longrightarrow w_2 \Longrightarrow w \ldots \Longrightarrow w_n = w$$

is a **derivation** of w by G.

Strings $w_1, w_2, \ldots$ which **contain non-terminals** are called **sentential forms** of G.

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

$$G = \{\ \{S\},\ \Sigma,\ S,\ P\}$$

where P is given by:

$$S \longrightarrow a\, S\, b$$

$$S \longrightarrow \lambda$$

Then:

$S \Longrightarrow \lambda$

$S \Longrightarrow a\,S\,b \Longrightarrow a\,b$

$S \Longrightarrow a\,S\,b \Longrightarrow a\,a\,S\,b\,b \Longrightarrow a^2b^2$

$S \Longrightarrow \underline{\boldsymbol{a\,S\,b \Longrightarrow a^2S\,b^2 \Longrightarrow a^3S\,b^3 \Longrightarrow a^3b^3}}$

So $L(G) = \underline{\{a^n\ S\ b \mid n \geq 0\}}$

[See Linz, Example 1.11]

Notation for P: $S \longrightarrow a\,S\,b \mid \lambda$

Example: [Linz. Example 1.12]

Find a grammar for $L = \{a^n\, b^{n+1} \mid n \geq 0\}$

$\Sigma = \{a, b\}, G = \{\ \{S, A\},\ \Sigma,\ S,\ P\}$

Where P consists of

$$S \longrightarrow A\,b$$
$$A \longrightarrow a\,A\,b \mid \lambda$$