Grammatiche libere dal contesto

Una grammatica può essere intesa come un insieme di regole che permettono di generare un linguaggio. Le grammatiche libere dal contesto permettono di descrivere la sintassi dei linguaggi di programmazione.

Una **grammatica** può essere definita come una quadrupla $< V, T, P, S >$ (per la spiegazione vedere linguaggi).

Una **produzione di un linguaggio context free** è del tipo:

$$A \rightarrow \alpha$$

ovvero in ogni stringa posso sostituire A con α indipendentemente da ciò che ho a destra e a sinistra[1].

Esempio:

$$\gamma A \delta A w \rightarrow^1 \gamma \alpha \delta A \rightarrow^2 \gamma \alpha \delta \alpha w$$

Linguaggio libero dal contesto

Un linguaggio generato da una grammatica G si definisce come segue:

$$L(G) = \{x \in T^* | S \rightarrow^* x\}$$

ovvero consiste di tutte quelle stringhe che, a partire dal simbolo iniziale e dopo un numero finito di riduzioni ($\rightarrow^*$), arrivano ad x.

Un linguaggio si dice libero dal contesto se è generato da una grammatica libera dal contesto. Un linguaggio può essere generato da infinite grammatiche.

Definizione Una grammatica si dice ambigua se esiste una stringa per la quale esistono due alberi di derivazione diversi.

Esempio: dimostrare che il linguaggio è context free

$$\mathfrak{L} = \{a^n b^n | n \geq 0\}$$

Per dimostrare che il linguaggio è CE devo costruire la grammatica CE che lo genera.

$$\begin{aligned} T =& \{a, b\} \\ V =& \{S\} \\ S =& iniziale \end{aligned}$$

Ora devo costruire le produzioni:

- Poichè posso avere anche $n = 0$, il linguaggio contiene la string avuote ($\varepsilon = a^0 b^0$);
- Se $a^n b^n$ ($S \rightarrow a^n b^n$) è una stringa del linguaggio, allora anche aSb appartiene al linguaggio ($== a a^n b^n b$).

Quindi, schematizzando:

$$P = \begin{cases} S \rightarrow \varepsilon \\ S \rightarrow aSb \end{cases} \qquad oppure \qquad P = \varepsilon | aSb$$

Ora devo dimostrare che il linguaggio generato dalla grammatica che ho costruito è uguale a quello di partenza ($\mathfrak{L}(G) = \mathfrak{L}$):

Dimostrazione $\mathfrak{L}(G) \leftarrow \mathfrak{L}$

Caso $n = 0 \rightarrow \varepsilon \in \mathfrak{L}$	$S \rightarrow^1 \varepsilon \in \mathfrak{L}(G)$
Caso $n > 0 \rightarrow a^n b^n \in \mathfrak{L}$ e $a^n b^n \in \mathfrak{L}(G)$ suppongo $S \rightarrow^* a^n b^n$	$S \rightarrow aSb \rightarrow^* a a^n b^n b = a^{n+1} b^{n+1}$

[1] Da questo deriva il nome "libere dal contesto"

Dimostrazione $\mathfrak{L}(G) \to \mathfrak{L}$

Caso $n = 1$ $\varepsilon \in \mathfrak{L}$

Caso $n > 1$ suppongo $(S \to^n a^k b^k)$ $\underline{S \to aSb \to^n}_{n+1} \ aa^k b^k b$

Esempio: costruire la grammatica

$$\mathfrak{L} = \{a^m b^{n+2} c^n a^{m+2} | m, n \geq 0\}$$

Scompongo la stringa principale in due sottostringhe $b^{n+2}c^n$ e $a^m a^{m+2}$:

$S \to aAaaa|Aaa$ (caso stringa vuota)
$A \to bbbAc|bb$ (caso stringa vuota)

Questa soluzione presenta però un errore: ?. Quindi Riscrivo la stringa principale come $a^m[bb(b^n c^n)aa]a^n$:

$S \to aSa|bbBaa$
$B \to \varepsilon|bBc$ $(B == b^n c^n)$