

Bottom-up parsing: Ricostruzione di un albero di derivazione rightmost, per una data stringa di ingresso:

- Si parte dalla frontiera (foglie) dell'albero,
- Si procede verso la radice dell'albero (etichettata con lo start symbol della grammatica).

Tipologie:
 LALR(1):
 - La nostra
 - YACC, Bison
 LR(1)

$S \Rightarrow^* \alpha A w$ con un certo numero di passi
 per convenzione, indica tutti i simboli non-terminali

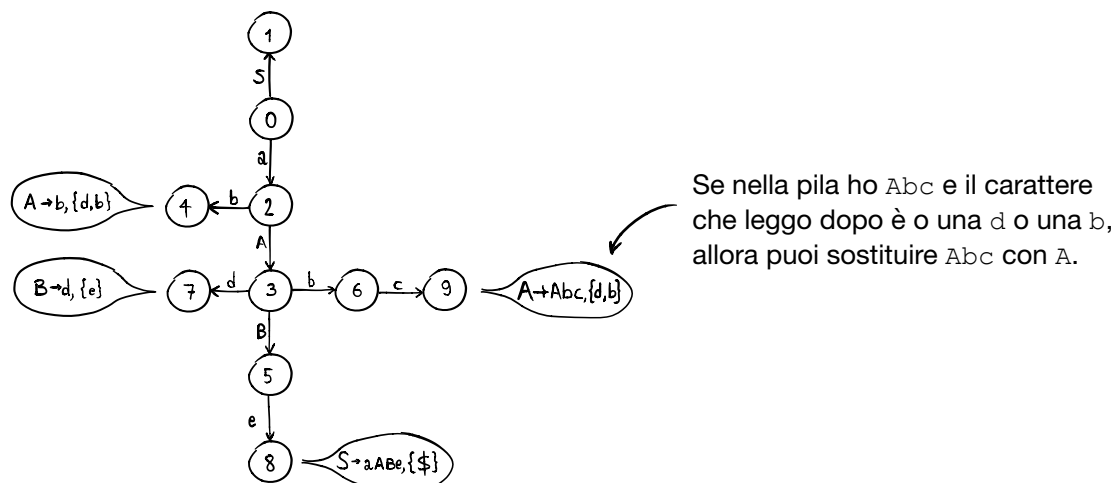
Supponendo che $A \rightarrow \beta \in \text{Produzioni}$, $S \Rightarrow^* \alpha A w \Rightarrow \alpha \beta w \Rightarrow^* w_1 \in \text{Linguaggio}$

Le decisioni fondamentali nel parsing bottom-up riguardano *quando* effettuare una riduzione e *quale* produzione applicare, cioè capire quando è il momento di dire che β è l'espansione di A .

Partiamo da piccoli passi, con la seguente grammatica:

$S \rightarrow aABe$
 $A \rightarrow Abc \mid b$
 $B \rightarrow d$

Da cui deriveremo in seguito l'automa, ma noi lo anticipiamo e lo disegniamo qua sotto:



I cammini dell'automa ripercorrono il driver delle produzioni.

Spiegazione del libro

Il parsing bottom-up è il processo di progressiva *riduzione* di una stringa di ingresso w fino al simbolo iniziale della grammatica.

Prima di vedere un esempio, diamo la definizione di **handle**, o maniglia:

- Un handle è una sottostringa corrispondente al corpo di una produzione.
- La sua derivazione rappresenta un passo nel processo di derivazione destra al rovescio.

Grammatica:

$E \rightarrow E + T \mid T$
 $T \rightarrow T * F \mid F$
 $F \rightarrow (E) \mid \text{Id}$

Stringa di ingresso:

$\text{Id1} * \text{Id2}$

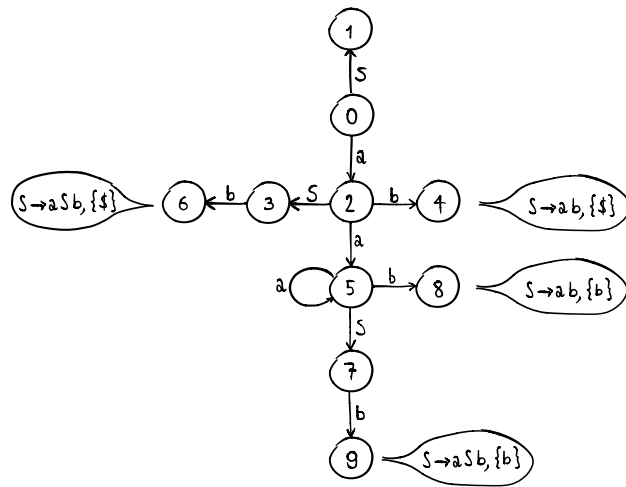
Un parsing corretto ed efficace è:

Forma sentenziale	Handle	Produzione riducente
$\text{Id1} * \text{Id2}$	Id1	$F \rightarrow \text{Id}$
$F * \text{Id2}$	F	$T \rightarrow F$
T $* \text{Id2}$	Id2	$F \rightarrow \text{Id}$
$T * F$	$T * F$	$T \rightarrow T * F$
T	T	$E \rightarrow T$
$E \quad // \text{È di fatto una produzione destra: } E \Rightarrow T \Rightarrow T * F \Rightarrow T * \text{id} \Rightarrow F * \text{id} \Rightarrow \text{id} * \text{id}$		

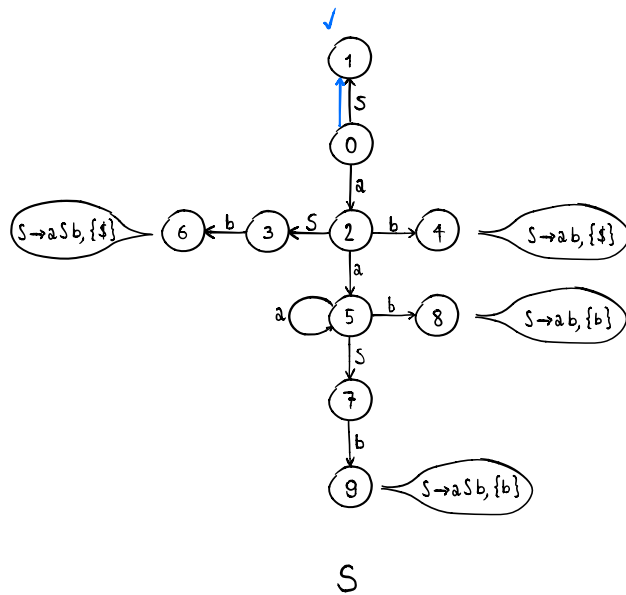
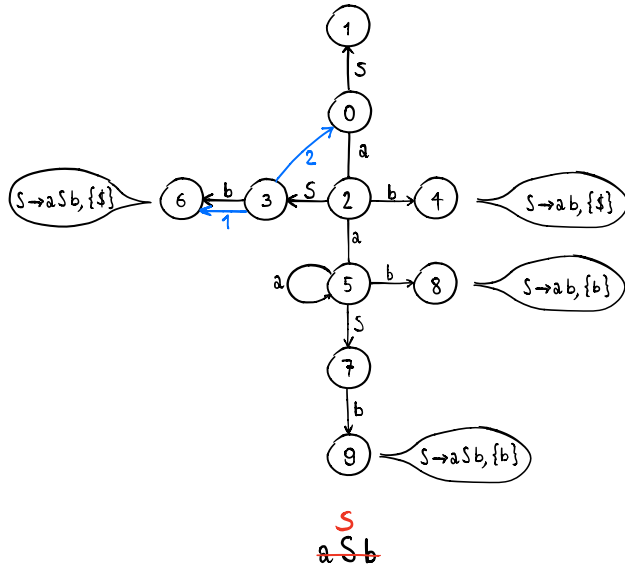
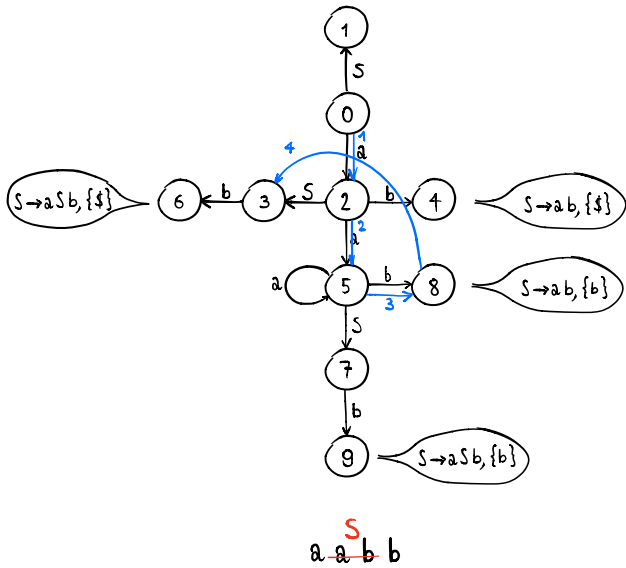
Attenzione che, nonostante la T in grassetto sia corpo della produzione $E \rightarrow T$, non può essere handle per la forma sentenziale $T * \text{Id2}$, altrimenti otterremmo la forma $E * \text{Id2}$ che non può essere derivata dal simbolo iniziale E .

Pertanto si conclude che la sottostringa sinistra corrispondente al corpo di una qualche produzione non deve necessariamente essere un handle.

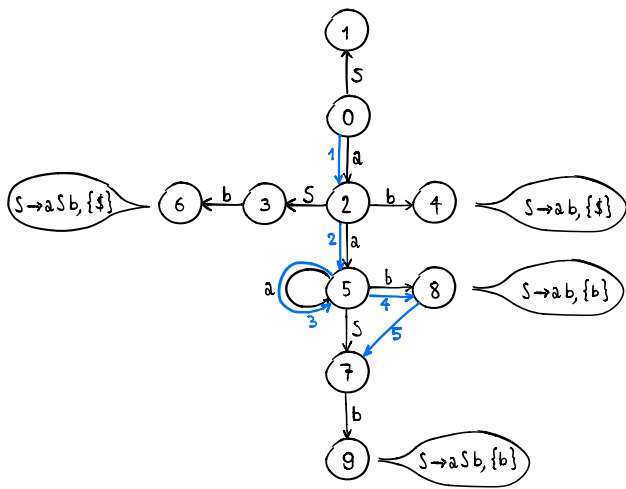
Vediamo la grammatica $S \rightarrow aSb \mid ab$ che ha il seguente automa:



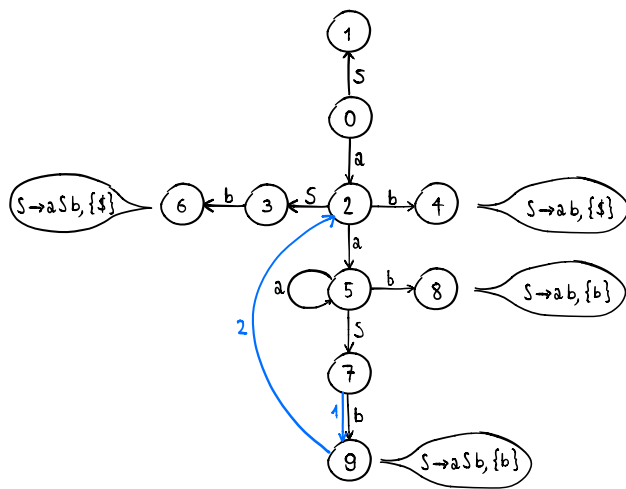
Come si deriva la stringa aabb? [$S \Rightarrow aSb \Rightarrow aabb$]



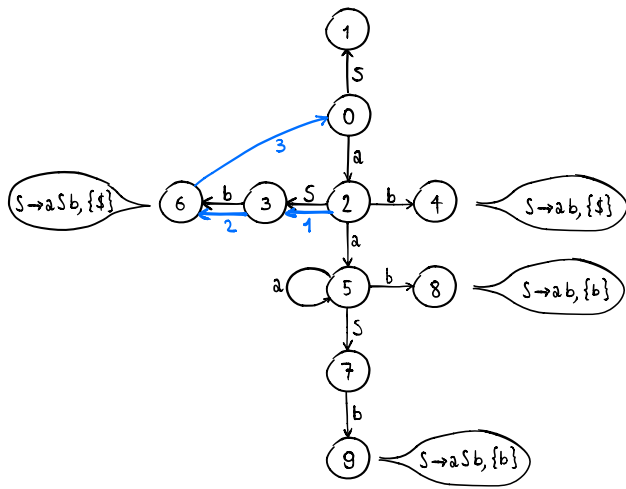
Come si deriva la stringa aaabbb? [$S \Rightarrow aSb \Rightarrow aaSbb \Rightarrow aaabbb$]



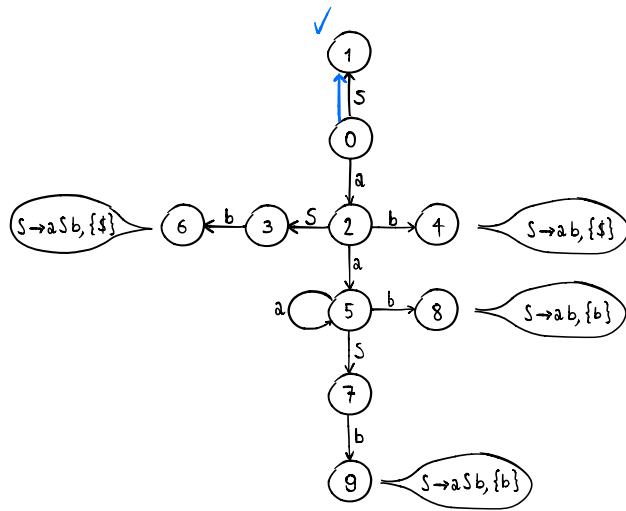
aaabbb



aaabbb



aaabbb



S