

Indice

1	Intro	2
1.1	Definizioni	4
1.1.1	Equivalenza tra le due nozioni di accettazione nel modello nondeterministico	9
1.2	Grammatiche di tipo 2	11
1.3	Equivalenza tra grammatiche di tipo 2 ad automi a pila	16
1.3.1	Da una grammatica che genera un linguaggio generiamo un automa che riconosce lo stesso	16
1.3.2	Da un automa a pila costruiamo una grammatica di tipo 2	19
1.4	Forme normali per le grammatiche di tipo 2	23
1.4.1	Forma normale di Greibach	23
1.4.2	Forma normale di Chomsky	24

Capitolo 1

Intro

Gli automi a pila sono automi che oltre ad avere un controllo a stati finiti hanno una memoria arbitrariamente grande, ma organizzata a pila; cioè una memoria a cui si può accedere solo all'elemento più in cima. Questo modello è one-way sul nastro di input (la versione two-way è più potente). Analizzeremo principalmente il modello nondeterministico, infatti come vedremo il modello deterministico è – a differenza degli FSA – meno potente di quello nondeterministico.

Un automa a pila è una tupla

$$M = \langle Q, \Sigma, \Gamma, \delta, q_0, Z_0, F \rangle$$

dove

- Γ è l'alfabeto della pila o alfabeto di lavoro
- δ è la funzione di transizione
- $q_0 \in Q$ è lo stato iniziale dell'automa
- $Z_0 \in \Gamma$ è lo stato iniziale della pila
- F è un insieme di stati finali

La funzione di transizione dipende da tre cose: dallo stato corrente, dal simbolo dell'input corrente e dal simbolo in cima alla

$$\delta : Q \times (\Sigma \cup \{\epsilon\}) \times \Gamma \rightarrow \text{PF}(Q \times \Gamma^*)$$

Quindi la funzione di transizione cambia lo stato dell'automa e rimpiazza il simbolo in cima alla pila con una stringa di stati della pila.

Nota 1. $PF(-)$ sta per le parti finite, infatti se utilizzassimo $2^{Q \times \Gamma^*}$ potremmo avere programmi infiniti, visto che Γ^* è un insieme infinito.

Nota 2. $\Sigma \cup \{\epsilon\}$ perché sono contemplate mosse in base allo stato dell'automa che modificano la pila senza leggere un simbolo in input.

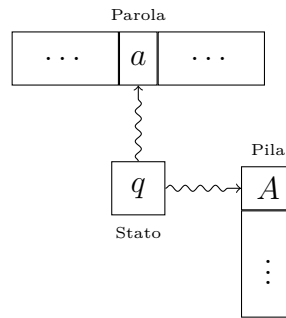
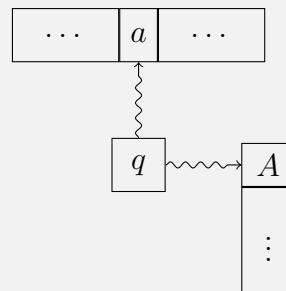


Figura 1.1: Rappresentazione delle varie parti di un PDA

Nota 3. Per convenzione la stringa di stati viene messa sulla pila di destra a sinistra, quindi il simbolo più a sinistra sarà in cima alla pila.

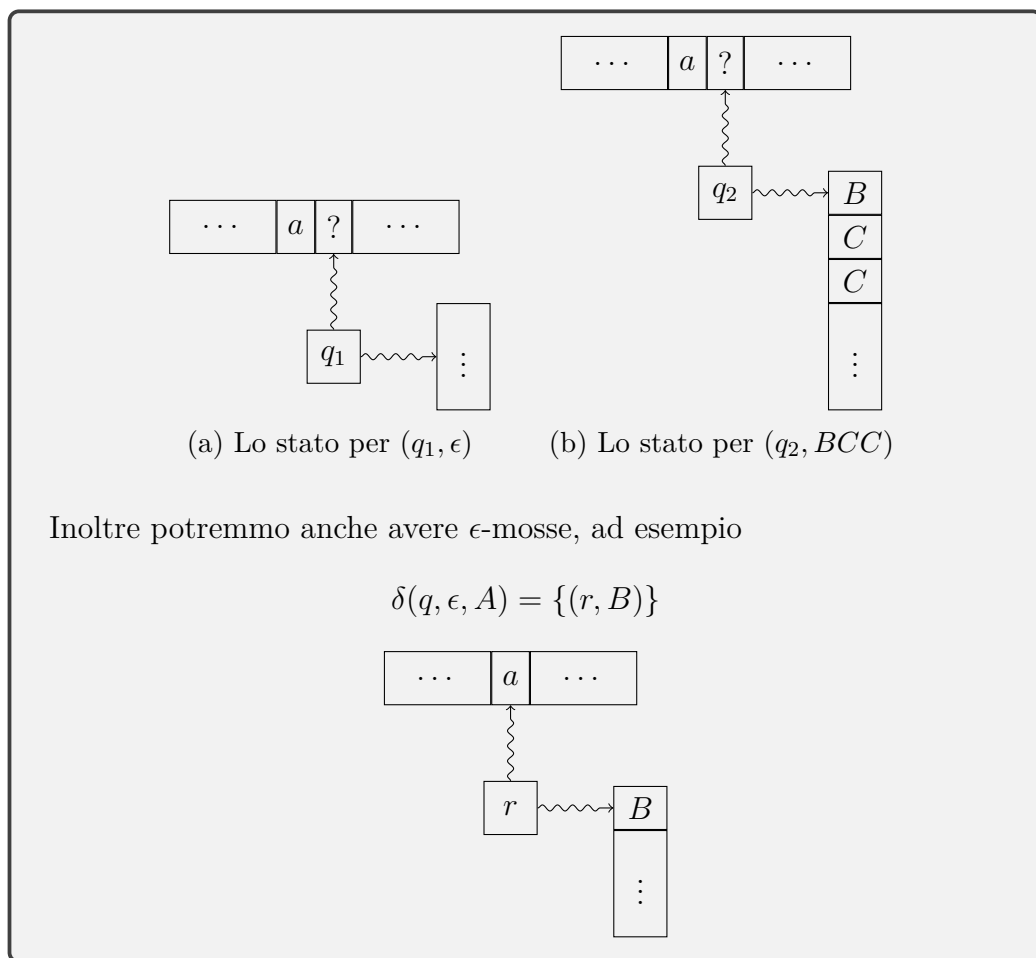
Supponiamo di essere nello stato



e che la funzione δ sia così definita

$$\delta(q, a, A) = \{(q_1, \epsilon), (q_2, BCC)\}$$

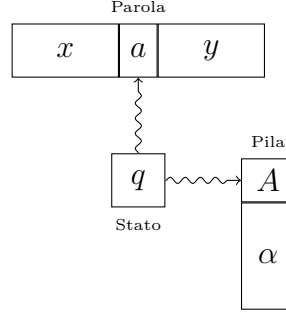
L'applicazione delle due alternative porterebbe l'automa nei seguenti stati



1.1 Definizioni

Ci riferiremo agli automi a pila come PDA (Push Down Automaton) e assumeremo che siano sempre nondeterministici, a meno che diversamente specificato.

Chiameremo lo stato complessivo dell'automa a pila la sua **configurazione**, questa verrà rappresentata compattamente come una tripla dello stato, la porzione di input ancora da leggere da quello corrente, e il contenuto della pila. Quindi



è rappresentato dalla configurazione

$$(q, ay, A\alpha)$$

con $q \in Q, a \in \Sigma \cup \{\epsilon\}, y \in \Sigma^*, A \in \Gamma, \alpha \in \Gamma^*$.

Una mossa, indica che da una configurazione q posso passare ad un'altra p , scritto

$$q \vdash p$$

Ad esempio sopra abbiamo che

$$\begin{aligned} (q, ay, A\alpha) &\vdash (q_1, y, \alpha) \\ (q, ay, A\alpha) &\vdash (q_2, y, BCC\alpha) \\ (q, ay, A\alpha) &\vdash (r, ay, B\alpha) \end{aligned}$$

Più rigorosamente, sia $(q, ay, Z\alpha)$ la configurazione corrente, con $Z \in \Gamma$ e M l'automa a pila, diciamo che

$$(q, ay, Z\alpha) \vdash_M (p, y, \beta\alpha)$$

sse $(p, \beta) \in \delta(q, a, Z)$ dove $q, p \in Q, y \in \Sigma^*, a \in \Sigma \cup \{\epsilon\}, Z \in \Gamma$ e $\alpha, \beta \in \Gamma^*$. Se l'automa è ovvio dal contesto possiamo ometterlo da \vdash e scrivere solo \vdash .

Da una configurazione C' arrivo ad una configurazione C'' in un certo numero di mosse – scritto

$$C' \vdash_M^* C''$$

sse esistono C_0, \dots, C_k con $C_0 = C'$ e $C_k = C''$ e $\forall i \in 1, \dots, k \ C_{i-1} \vdash_M C_i$.

La configurazione iniziale di un automa su input $w \in \Sigma^*$ è

$$(q_0, w, Z_0)$$

Per accettare possiamo dare alcune diverse definizioni di configurazione accettante:

- una volta finito l'input mi trovo in uno stato $q \in F$ e la pila può essere una stringa qualunque, questa è detta *accettazione per stati finali*, ed indichiamo il linguaggio accettato per stati finali dall'automa a pila M come

$$L(M) = \{w \in \Sigma^* \mid (q_0, w, Z_0) \stackrel{*}{\vdash} (q, \epsilon, \gamma), q \in F, \gamma \in \Gamma^*\}$$

Nota 4. *Siccome sono accettate le ϵ mosse può esserci il caso in arriviamo alla fine dell'input con uno stato non finale, e si può fare una ϵ -mossa ed arrivare ad uno stato finale.*

- è ragionevole pensare che tutto quello che viene messo sulla pila debba anche essere tolto, questa è detta *accettazione per pila vuota* per cui si deve arrivare alla fine dell'input ed aver svuotato l'intera pila, ingorandolo lo stato. Il linguaggio accettato per pila vuota dall'automa M lo indichiamo come

$$N(M) = \{w \in \Sigma^* \mid (q_0, w, Z_0) \stackrel{*}{\vdash} (q, \epsilon, \epsilon), q \in Q\}$$

In questo caso ovviamente si può omettere F dalla definizione dell'automa.

- si può pensare di richiedere entrambe le precedenti, come vedremo più avanti queste tre nozioni sono equivalenti (nel caso nondeterministico)

Nota 5. *Questa cosa la vedremo meglio, ma visto che la pila è la struttura fondamentale per la ricorsione, i linguaggi CF sono i linguaggi regolari a cui è stata aggiunta la ricorsione.*

Definiamo il linguaggio

$$\mathcal{L} = \{a^n b^n \mid n \geq 1\}$$

possiamo usare la pila per contare il numero di a .

$$\delta(q_0, a, Z_0) = \{(q_0, A)\}$$

$$\delta(q_0, a, A) = \{(q_0, AA)\}$$

$$\delta(q_0, b, A) = \{(q_1, \epsilon)\}$$

$$\delta(q_1, b, A) = \{(q_1, \epsilon)\}$$

E vale che data questa δ

$$\mathcal{L} = N(M)$$

Questo è un caso particolare di automa a pila in cui utilizziamo in simbolo solo (cioè A , oltre a Z_0) è detto *automa a contatore*.

Definiamo alternativamente

$$\delta(q_0, a, Z_0) = \{(q_0, AZ_0)\}$$

$$\delta(q_0, a, A) = \{(q_0, AA)\}$$

$$\delta(q_0, b, A) = \{(q_1, \epsilon)\}$$

$$\delta(q_1, b, A) = \{(q_1, \epsilon)\}$$

$$\delta(q_1, \epsilon, Z_0) = \{(q_F, \epsilon)\}$$

con $F = \{q_F\}$, e vale che con questa δ

$$\mathcal{L} = L(M)$$

Vediamo ora il caso di sopra, ma in cui

$$\mathcal{L} = \{a^n b^n \mid n \geq 0\}$$

possiamo usare la pila per contare il numero di a .

$$\begin{aligned}\delta(q_0, \epsilon, Z_0) &= \{(q_0, \epsilon)\} \\ \delta(q_0, a, Z_0) &= \{(q_0, A)\} \\ \delta(q_0, a, A) &= \{(q_0, AA)\} \\ \delta(q_0, b, A) &= \{(q_1, \epsilon)\} \\ \delta(q_1, b, A) &= \{(q_1, \epsilon)\}\end{aligned}$$

E vale che data questa δ in cui si può direttamente accettare dallo stato q_0

$$\mathcal{L} = N(M)$$

L'introduzione della prima regola è problematica, perché con input non vuoto permette di svuotare la pila da Z_0 , bloccando la continuazione dell'automa, quindi abbiamo introdotto il nondeterminismo tra le due regole $\delta(q_0, \epsilon, Z_0)$ e $\delta(q_0, a, Z_0)$.

Definiamo similmente a sopra

$$\begin{aligned}\delta(q_0, \epsilon, Z_0) &= \{(q_F, \epsilon)\} \\ \delta(q_0, a, Z_0) &= \{(q_0, AZ_0)\} \\ \delta(q_0, a, A) &= \{(q_0, AA)\} \\ \delta(q_0, b, A) &= \{(q_1, \epsilon)\} \\ \delta(q_1, b, A) &= \{(q_1, \epsilon)\} \\ \delta(q_1, \epsilon, Z_0) &= \{(q_F, \epsilon)\}\end{aligned}$$

con $F = \{q_F\}$, e vale che con questa δ

$$\mathcal{L} = L(M)$$

sempre introducendo non determinismo.

Questa versione, a differenza di quello di sopra per pila vuota, può anche essere fatta senza non determinismo infatti definiamo q_I come nuovo stato iniziale che è anche finale, se la stringa è vuota possono direttamente accettare, mentre

$$\delta(q_I, a, Z_0) = \{(q_0, AZ_0)\}$$

ci riconduce all'automa di sopra.

Supponiamo invece di voler riconoscere

$$L = \{w \in \{a, b\}^* \mid \#_a(w) = \#_b(w)\}$$

bisogna contare sia le a che le b : finché si incontrano a si aggiungono alla pila, e appena si incontrano b le si tolgono. Se durante questa operazione si svuota la pila iniziamo a contare le b e facciamo la operazione inversa. Queste due operazioni continuano ad alternarsi finché la stringa non è svuotata.

Definiamo ora l'automa a pila deterministico. Questo in ogni configurazione permette una singola scelta:

- sono vietate configurazioni che ammettono una mossa e una ϵ -mossa, quindi $\forall q \in Q, z \in \Gamma$ se $\delta(q, \epsilon, z) \neq \emptyset$ allora $\forall a \in \Sigma \delta(q, a, Z) = \emptyset$
- $\forall q \in Q, z \in \Gamma, a \in \Sigma \cup \{\epsilon\} \mid |\delta(q, a, Z)| \leq 1$, quindi per ogni tripletta q, a, Z è ammessa al massimo una mossa.

A questo punto abbiamo definito quattro modelli: deterministico e nondeterministico che possono accettare per pila vuota o per stato finale. Vedremo che il caso nondeterminismo in questo caso è più potente del caso deterministico, e che nel modello nondeterministico automi che possono accettare per pila vuota o per stato finale sono equivalenti.

1.1.1 Equivalenza tra le due nozioni di accettazione nel modello nondeterministico

Da stati finali a pila vuota

Dato un automa $M = (Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ e supponiamo che $L = L(M)$ sia il linguaggio accettato per stati finali. Definiamo l'automa

$$M' = (Q \cup \{q'_0, q_e\}, \Sigma, \Gamma \cup \{X\}, \delta', q'_0, X, \emptyset)$$

con $q'_0, q_e \notin Q$ e $X \notin \Gamma$, vogliamo che $L = N(M')$.

Ad alto livello quando M arriva in uno stato finale, M' si sposta nello stato q_e in cui inizia a svuotare la pila. Infatti la e di q_e sta per “empty”.

Definiamo ora δ' :

1. prima di tutto

$$\delta'(q'_0, \epsilon, X) = \{(q_0, Z_0X)\}$$

questo serve solo ad infilare X in fondo alla pila. La X è necessaria per evitare che se l'automa iniziale M svuota la pila si accetti la stringa.

2. per ogni altra cosa M' si può comportare come M :

$$\forall q \in Q, a \in \Sigma \cup \{\epsilon\}, z \in \Gamma \mid \delta(q, a, Z) \subseteq \delta'(q, a, Z)$$

3. ogni qualvolta M entra in uno stato finale M' può – enfasi su può – iniziare a svuotare l'intera pila:

$$\forall q \in F, z \in \Gamma \cup \{X\} \mid (q_e, \epsilon) \in \delta'(q, \epsilon, Z)$$

4. una volta entrato nello stato di svuotamento, continua a svuotare:

$$\forall z \in \Gamma \cup \{X\} \mid \delta'(q_e, \epsilon, Z) = \{(q_e, \epsilon)\}$$

Questo necessariamente introduce nondeterminismo, infatti l'automa M potrebbe entrare in uno stato finale prima di essere arrivato alla fine della stringa. Ed anche se l'automa di partenza è deterministico il punto 3 potrebbe in ogni caso introdurre nondeterminismo.

Supponiamo di avere un automa deterministico che accetta la stringa w a pila vuota, allora ogni stringa che ha w come prefisso non può essere accettata, perché il prefisso w svuoterebbe la pila e un automa con pila vuota non può andare avanti. Quindi il nondeterminismo è in un certo senso necessario per automi a pila che accettano con pila vuota.

Da pila vuota a stati finali

Dato un automa $M = (Q, \Sigma, \Gamma, \delta, q_0, Z_0, \emptyset)$ che accetta per pila vuota il linguaggio $L = N(M)$, vogliamo creare un automa che accetti per stati finali. Sia questo

$$M' = (Q \cup \{q'_0, q_F\}, \Sigma, \Gamma \cup \{X\}, \delta', q'_0, X, F = \{q_F\})$$

con $q'_0, q_F \notin Q, X \notin \Gamma$.

Definiamo ora δ' :

- come prima inizialmente infiliamo X in fondo alla pila:

$$\delta'(q'_0, \epsilon, X) = \{(q_0, Z_0X)\}$$

X serve a riconoscere quando la pila è vuota.

- a questo punto copiamo tutte le mosse di M , per cui

$$\forall q \in Q, a \in \Sigma \cup \{\epsilon\}, Z \in \Gamma \mid \delta'(q, a, Z) = \delta(q, a, Z)$$

- nel momento in cui M svuota la prima, M' si trova X sulla pila, a questo punto può entrare in uno stato finale

$$\forall q \in Q \mid \delta'(q, \epsilon, X) = \{(q_F, \epsilon)\}$$

Supponendo che M sia deterministico, M' rimane deterministico – la trasformazione preserva il determinismo.

1.2 Grammatiche di tipo 2

Una grammatica è formata da quattro elementi:

$$G = \langle V, \Sigma, P, S \rangle$$

e nello specifico, in quelle di tipo 2 le produzioni hanno la forma

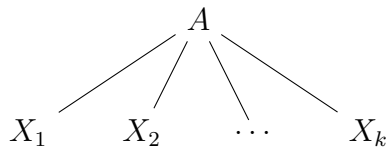
$$A \rightarrow \alpha \quad A \in V, \alpha \in (V \cup \Sigma)^*$$

Una rappresentazione utile per le derivazioni di linguaggi CF sono gli alberi, ad esempio data $w \in L(G)$, allora $S \xRightarrow{*} w$. Questa derivazione io la possono rappresentare come un albero di derivazione, o albero di parsing, o ancora parse tree. Questo è un albero

- con radice etichettata con il simbolo iniziale della grammatica
- le foglie da sinistra a destra sono w
- i nodi possono essere di tre tipi:
 - variabili, per i nodi interni

- terminali, per le foglie
- ϵ la parola vuota, in casi speciali per le foglie

Dato un nodo



rappresenta l'applicazione della regola di produzione

$$A \rightarrow X_1 X_2 \dots X_k \in P \quad A \in V, \forall i \in 1, \dots, k \mid X_i \in V \cup \Sigma$$

All'ultimo livello possiamo avere nodi



solo se $A \rightarrow \epsilon \in P$.

Abbiamo detto che gli automi a pila riconoscono linguaggi con ricorsione, dove questa nell'automa si esprime nella memoria a pila, nelle grammatiche si esprime nella struttura ad albero.

Definiamo la grammatica per le parentesi correttamente bilanciate

$$\begin{aligned} S &\rightarrow \epsilon \\ S &\rightarrow (S) \\ S &\rightarrow SS \end{aligned}$$

prendiamo ora la stringa $w = (()())()$ e scriviamone la derivazione

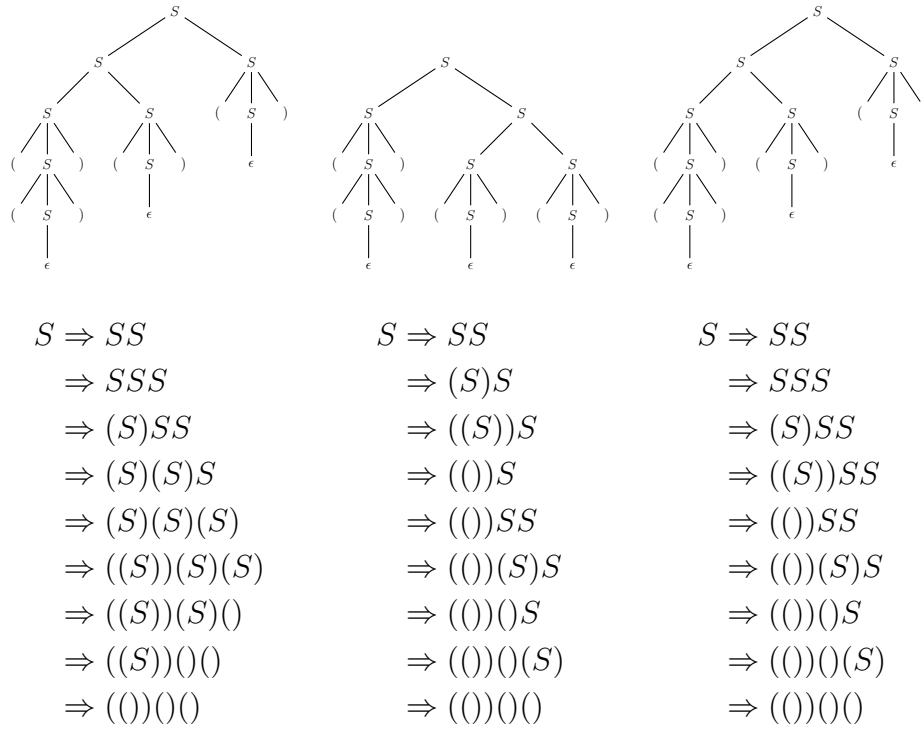
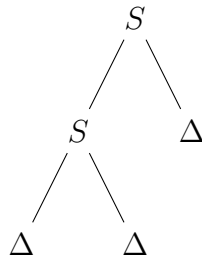
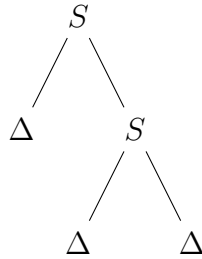


Figura 1.3: Tre derivazioni diverse per la stringa $((()))()$ e gli alberi corrispondenti

Possiamo vedere che una stessa stringa ammette diverse derivazioni, ma non tutte queste portano allo stesso albero. Infatti la prima e la terza derivazione utilizzano le stesse sostituzioni, solo in ordine diverso, e quindi generano alberi uguali; mentre nel secondo albero applichiamo derivazioni diverse. Nella prima e nella terza derivazione abbiamo una struttura



mentre la seconda ha una struttura



Per evitare derivazioni multiple si utilizza un criterio detto di *derivazione leftmost*: una derivazione è leftmost se ogni volta che si fa una sostituzione sostituisco sempre la variabile più a sinistra della forma sentenziale.

Proposizione 1. *Esiste una corrispondenza uno a uno tra derivazioni leftmost e alberi di derivazione.*

La seconda e la terza derivazioni dell'esempio di sopra sono due derivazioni leftmost diverse.

Diciamo che una grammatica è *ambigua* se c'è una stringa che ammette almeno due alberi di derivazione – o derivazioni leftmost – diversi.

Nell'esempio di sopra si può vedere anche che ogni sottoalbero è una sequenza bilanciate di parentesi.

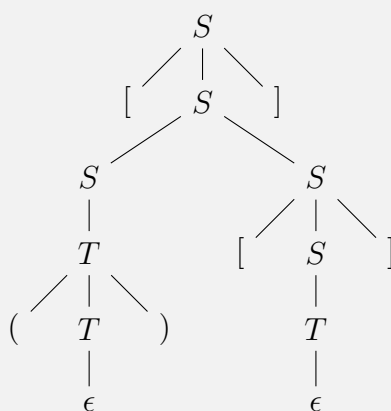
Se nella grammatica di sopra vorremmo anche le quadre, senza precedenze, questa è facilmente

$$\begin{aligned} S &\rightarrow \epsilon \\ S &\rightarrow (S) \\ S &\rightarrow [S] \\ S &\rightarrow SS \end{aligned}$$

Ma se si chiede che le quadre non possano stare all'interno delle tonde, diventa necessario suddividere le variabili in due livelli

$$\begin{aligned} S &\rightarrow T \\ S &\rightarrow [S] \\ S &\rightarrow SS \\ T &\rightarrow \epsilon \\ T &\rightarrow (T) \\ T &\rightarrow TT \end{aligned}$$

e vediamo un albero di derivazione di esempio



1.3 Equivalenza tra grammatiche di tipo 2 ad automi a pila

1.3.1 Da una grammatica che genera un linguaggio generiamo un automa che riconosce lo stesso

Data una grammatica

$$G = \langle V, \Sigma, P, S \rangle$$

di tipo 2, vogliamo costruire

$$M = \langle Q, \Sigma, \Gamma, \delta, q, Z_0, \emptyset \rangle$$

che accetta per pila vuota, con

- Q formato da un solo stato $\{q\}$
- $\Gamma = \Sigma \cup V$
- $Z_0 = S$

e δ definito come

- se $A \rightarrow \alpha \in P$ allora $(q, \alpha) \in \delta(q, \epsilon, A)$
- $\forall \sigma \in \Sigma, \delta(q, \sigma, \sigma) = \{(q, \epsilon)\}$, cioè si consuma il simbolo in cima alla pila

Si può dimostrare che il linguaggio generato dalla grammatica $L(G)$ è uguale al linguaggio accettato dall'automa per pila vuota $N(M)$.

Prendiamo

$$G = \langle \{S, T, U\}, \{a, b\}, P, S \rangle$$

con P definito

$$S \rightarrow TU$$

$$T \rightarrow aTb \mid \epsilon$$

$$U \rightarrow bUa \mid \epsilon$$

questo genera

$$L = \{a^n b^{n+m} a^m \mid n \geq 0, m \geq 0\}$$

Scriviamo le transizioni dell'automa corrispondente

$$M = \langle \{q\}, \{a, b\}, \{S, T, U, a, b\}, \delta, q, S, \emptyset \rangle$$

con δ definito come

$$\delta(q, \epsilon, S) = \{(q, TU)\}$$

$$\delta(q, \epsilon, T) = \{(q, aTb), (q, \epsilon)\}$$

$$\delta(q, \epsilon, U) = \{(q, bUa), (q, \epsilon)\}$$

$$\delta(q, a, a) = \{(q, \epsilon)\}$$

$$\delta(q, b, b) = \{(q, \epsilon)\}$$

Prendendo per esempio $w = abbbaa$, vediamo come viene accettata nondeterministicamente

$$\begin{aligned}
(q, abbbaa, S) &\vdash (q, abbbaa, TU) \\
&\vdash (q, abbbaa, TU) \\
&\vdash (q, abbbaa, aTbU) \\
&\vdash (q, bbbaa, TbU) \\
&\vdash (q, bbbaa, bU) \\
&\vdash (q, bbaa, U) \\
&\vdash (q, bbaa, bUa) \\
&\vdash (q, baa, Ua) \\
&\vdash (q, baa, bUaa) \\
&\vdash (q, aa, Uaa) \\
&\vdash (q, aa, aa) \\
&\vdash (q, a, a) \\
&\vdash (q, \epsilon, \epsilon)
\end{aligned}$$

Leggere la i terminali consumati fino a un certo punto e il contenuto della pila in quel punto restituisce la forma sentenziale durante la derivazione. Questo corrisponde a

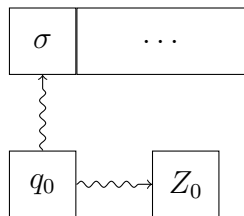
$$\begin{aligned}
S &\Rightarrow TU \\
&\Rightarrow aTbU \\
&\Rightarrow abU \\
&\Rightarrow abbUa \\
&\Rightarrow abbbUaa \\
&\Rightarrow abbbaa
\end{aligned}$$

L'automa a pila tenta di simulare il processo di derivazione leftmost della stringa.

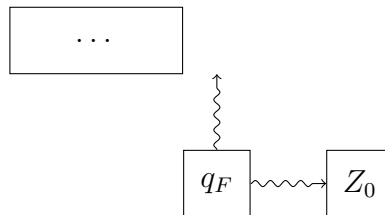
1.3.2 Da un automa a pila costruiamo una grammatica di tipo 2

Per la dimostrazione useremo una variazione degli automi a pila che non ne cambia la potenza computazionale. In questa forma normale

- all'inizio la pila contiene un simbolo speciale Z_0 che viene mai rimosso e non viene mai aggiunto



- alla fine l'input è stato letto completamente, la pila contiene solo Z_0 e lo stato è finale.



- le mosse sulla pila possono essere solo
 - push di un simbolo
 - pop di un simbolo
 - pila invariata

quindi il pop non è più implicito

- se una mossa legge un simbolo da input, allora non modifica la pila. Cioè le mosse che manipolano la pila sono separate da quelle che manipolano l'input.

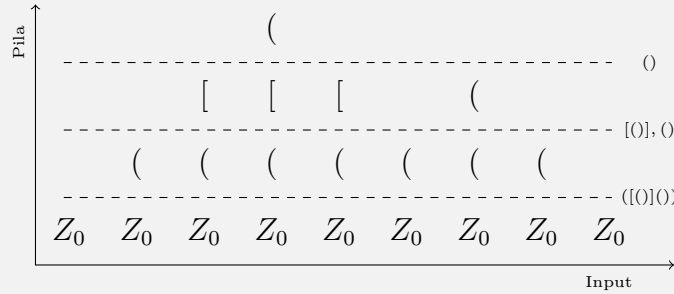
In questa forma

$$\delta : Q \times (\Sigma \cup \{\epsilon\}) \times \Gamma \rightarrow 2^{Q \times \{-, \text{pop}, a \in \Gamma | \text{push}(A)\}}$$

ed abbiamo che le mosse possono avere le seguenti forme

- mosse di lettura: $(p, -) \in \delta(q, a, A)$ $a \in \Sigma \cup \{\epsilon\}$
- pop: $(p, \text{pop}) \in \delta(q, \epsilon, A)$
- push: $(p, \text{push}(B)) \in \delta(q, \epsilon, A)$
- mosse che lasciano la pila invariata: $(p, -) \in \delta(q, \epsilon, A)$

Ad esempio se avessimo una sequenza di parentesi $([()])()$, la pila contiene inizialmente Z_0



questo disegno mostra la natura ricorsiva degli automi. Infatti visto che la pila di questo automa non può mai scendere sotto il suo livello iniziale, tutte le evoluzioni definite dalle linee tratteggiate definiscono parole valide del linguaggio.

Nota 6. *Gli automi che abbiamo visto fino ad ora possono essere simulati da questa versione normalizzata, scomponendo una mossa una pop ed una serie di push utilizzando degli stati ausiliari.*

Ripetiamo la versione di automa a pila semplificato vista a lezione scorsa, in questo per riuscire ad accettare dobbiamo arrivare in uno stato finale con solo Z_0 lo stato finale sulla pila.

Dobbiamo trovare un modo di trasformare un automa a pila come definito nella lezione scorsa, in una grammatica. Questa grammatica ha nonterminali della forma $[qAp]$ con $q, p \in Q$ e $A \in V$, e rappresenta:

- q è lo stato in cui si inizia
- p è lo stato in cui si finisce
- e A è il simbolo in cima alla pila all'inizio e alla fine della computazione.

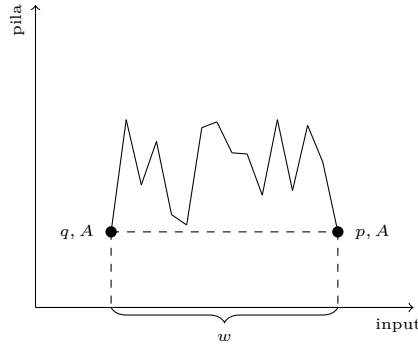


Figura 1.4: La computazione rappresentata dal simbolo $[qAp]$

Infatti negli automi come li abbiamo definiti, vale la proprietà per cui ???

Definiamo ora le regole di produzione della grammatica induttivamente come le stringhe riconosciute dalla computazione $[qAp]$:

- base: abbiamo due casi
 - caso 0: il caso più semplice è $[qAq]$, qui l'unica parola riconosciuta è ϵ , quindi è necessaria la regola

$$[qAq] \rightarrow \epsilon$$

Quindi creo tutte le produzioni della forma

$$\forall q \in Q, A \in \Gamma \mid [qAq] \rightarrow \epsilon$$

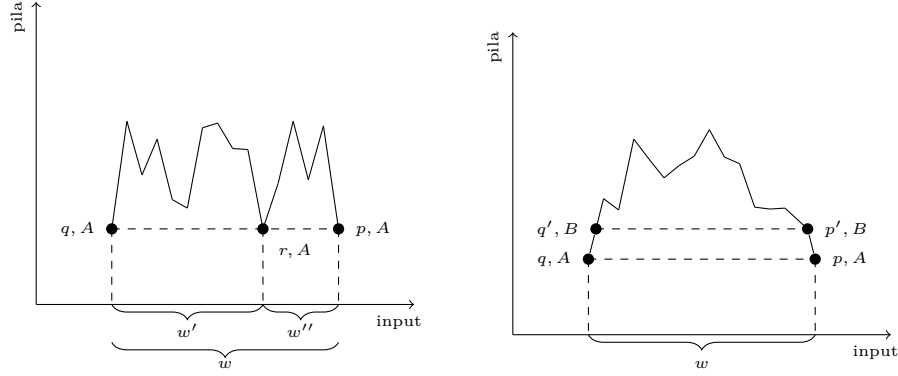
- caso 0': il secondo caso più semplice è quello in cui si è nello stato q , si consuma un carattere o nessuno, e questo ci porta nello stato p ; cioè $(p, -) \in \delta(q, a, A)$ con $a \in \Sigma \cup \{\epsilon\}$. Questo si traduce nella produzione

$$[qAp] \rightarrow a, \quad a \in \Sigma \cup \{\epsilon\}$$

Quindi creo tutte le produzioni della forma

$$\forall q, p \in Q, A \in \Gamma, a \in \Sigma \cup \{\epsilon\} \mid [qAp] \rightarrow a$$

- passo: si distinguono due casi



- caso 1: nei passi intermedi (tranne l'ultimo) la pila è sempre strettamente più alta di quando si è iniziato, questo si traduce in

$$[qAp] \rightarrow [q'Bp']$$

con $(q', \text{push}(B)) \in \delta(q, \epsilon, A)$ e $(p, \text{pop}) \in \delta(p', \epsilon, B)$. Quindi

$$\begin{aligned} \forall q, q', p, p' \in Q, A, B \in \Gamma \\ | (q', \text{push}(B)) \in \delta(q, \epsilon, A) \wedge (p, \text{pop}) \in \delta(p', \epsilon, B) \\ \Rightarrow [qAp] \rightarrow [q'Bp'] \end{aligned}$$

- caso 2: la computazione $[qAp]$ svuota la pila fino alla A inizia e poi continua, allora possiamo scomporre la computazione in due parti, quindi

$$\forall q, p, r \in Q, A \in \Gamma \mid [qAp] \rightarrow [qAr][rAp]$$

Si può dimostrare che

Lemma 1. $\forall q, p \in Q, A \in \Gamma, w \in \Sigma^* \mid [qAp] \xrightarrow{*} w$ sse l'automa M in una configurazione con A in cima alla pila, stato q , dopo aver letto w raggiunge una configurazione in cui il contenuto della pila è lo stesso dell'inizio, lo stato è p e nei passi intermedi la pila non scende mai sotto il livello iniziale.

Quindi durante una computazione quello che c'è sotto al simbolo in cima alla pila all'inizio della computazione non è rilevante.

Per finire di costruire la grammatica manca di definire l'assioma. Prima di tutto si può notare che visto che l'automa parte nello stato q_0 con Z_0 e basta sulla pila, una stringa w può essere generata solo dalle triple

$$[q_0 Z_0 q_F] \xRightarrow{*} w$$

con q_0 iniziale e q_F finale. Quindi definiamo l'insieme dei nonterminali della grammatica V come l'insieme di tutte le triple definite induttivamente sopra unito ad un nuovo nonterminale S tale che

$$\forall q_F \in F \mid S \rightarrow [q_0 Z_0 q_F] \in P$$

e questo S così definito è il simbolo iniziale.

1.4 Forme normali per le grammatiche di tipo 2

Si può vedere che tutte le produzioni generate dalla traduzione da automa a pila a grammatica sono di pochi tipi: variabile a terminale, variabile a variabile e variabile a coppia di variabili. Da questo fatto e dal fatto che i linguaggi riconosciuti dagli automi a pila sono esattamente quelli generati dalle grammatiche di tipo 2, ci rendiamo conto che possiamo restringere di molto il tipo di forma che il lato destro di una produzione di una grammatica CF può assumere. Nel caso di sopra appunto da variabile a terminale, da variabile a variabile e da variabile a coppia di variabili.

Vediamo ora due forme normali. Ogni grammatica può essere trasformata in una di queste forme normali a patto di sacrificare la parola vuota.

1.4.1 Forma normale di Greibach

In una grammatica in FNG (Forma Normale di Greibach) tutte le produzioni sono della forma

$$A \rightarrow aB_1 \dots B_k, \quad a \in \Sigma, A, B_1, \dots, B_k \in V, k \geq 0$$

Supponiamo di avere la gramamtica

$$\begin{aligned} A &\rightarrow aBB \\ A &\rightarrow b \\ B &\rightarrow bB \\ B &\rightarrow b \end{aligned}$$

e di aver fatto la trasformazione in automa a pila. In questa forma normale la pila avrà in cima sempre un terminale e quindi si può avere un simbolo di lookahead e scegliere più precisamente la prossima produzione da utilizzare, anche se non si toglie il nondeterminismo (v. $B \rightarrow bB$ e $B \rightarrow b$). Un altro vantaggio di avere sempre un terminale in cima alla pila è che in questo tipo di automa si possono eliminare le ϵ -mosse.

1.4.2 Forma normale di Chomsky

Nella FNC (Forma Normale di Chomsky) ci sono solo due tipi di regole

$$\begin{aligned} A &\rightarrow BC & A, B, C &\in V \\ A &\rightarrow a & A &\in V, a \in \Sigma \end{aligned}$$

Questa genera alberi di derivazione binari, salvo sulle foglie; ed è comoda per studiare alcune proprietà combinatorie.

Trasformazione in FNC

Data una grammatica genererica G eseguiamo i seguenti passi (l'ordine è importante) per trasformarla in FNC:

1. eliminazione delle ϵ -produzioni: diciamo che una variabile A è *cancellabile* sse $A \xrightarrow{*} \epsilon$. Induttivamente A è cancellabile se

- banalmente $A \rightarrow \epsilon$
- o se $A \rightarrow X_1 X_2 \dots X_k$ e X_1, X_2, \dots, X_k sono tutti cancellabili.

Questo può essere definito come una chiusura dove

$$C_0 = \{A \mid A \rightarrow \epsilon\}$$

e

$$C_i = C_{i-1} \cup \{A \mid \exists A \rightarrow X_1 X_2 \dots X_k \text{ con } \forall i \in 1, \dots, k \mid X_i \in C_{i-1}\}$$

Visto che

$$C_0 \subseteq C_1 \subseteq \dots \subseteq V$$

e V è finito, allora esiste un i tale che $C_i = C_{i-1}$.

Ora sia C l'insieme delle variabili cancellabili, costruiamo una grammatica $G' = \langle V, \Sigma, P', S \rangle$ con P' costituito da tutte le produzioni di P eccetto le ϵ produzioni e per ogni produzione $A \rightarrow X_1 X_2 \dots X_k$ con $X_1, X_2, \dots, X_k \in V \cup \Sigma$ aggiungo a P' le produzioni $A \rightarrow X_{i_1} X_{i_2} \dots X_{i_j}$ tali che $1 \leq i_1 < i_2 < \dots < i_j \leq k$ e per $\forall X_l \notin X_{i_1}, \dots, X_{i_j} \mid X_l \in C$ e $j \geq 1$.

Supponiamo di avere nella grammatica G che vogliamo trasformare la produzione

$$A \rightarrow BCaD$$

e che l'insieme delle variabili cancellabili è $C = \{C, D\}$.

Nella mia grammatica G' simulo la cancellazione di C e D aggiungendo le produzioni

$$A \rightarrow BaD$$

$$A \rightarrow BCa$$

$$A \rightarrow Ba$$

Supponiamo di avere nella gramantica G che vogliamo trasformare la produzione

$$A \rightarrow CDE$$

e che l'insieme delle variabili cancellabili è $C = \{C, D, E\}$.
Nella mia gramantica G' simulo la cancellazione di C e D , quindi aggiungo le produzioni

$$A \rightarrow CD$$

$$A \rightarrow CE$$

$$A \rightarrow DE$$

$$A \rightarrow C$$

$$A \rightarrow D$$

$$A \rightarrow E$$

Visto che le produzioni da aggiungere sotto tutti i sottoinsiemi delle variabili cancellabili di un lato destro meno l'insieme vuoto, vengono aggiunte nel caso peggiore un numero esponenziale di produzioni.

2. eliminazione delle produzioni unitarie: una produzione unitaria è una produzione della forma

$$A \rightarrow B, \quad A, B \in V$$

Costruiamo similmente a prima l'insieme di tutte le coppie di variabili X, Y tali per cui $X \xRightarrow{+} Y$, cioè per cui vale Abbiamo quindi

$$X \rightarrow A_1 \rightarrow \cdots \rightarrow Y$$

questo processo infatti le catene sono di lunghezza al più $|V|$ senza contenere cicli.

Nella nuova grammatica tolgo tutte le produzioni unitarie e se $X \rightarrow \cdots \rightarrow Y \rightarrow \alpha$ e $\alpha \in \Sigma$ oppure $|\alpha| > 1$, allora aggiungo la produzione $X \rightarrow \alpha$.

3. eliminazione simboli inutili: $X \in V \cup \Sigma$ è utile sse $\exists S \xRightarrow{*} \alpha X \beta \xRightarrow{*} w \in \Sigma^*$. Questi sono eliminati utilizzando algoritmi sui grafi (chiusura bottom up, chiusura top down, non lo ha spiegato ma ci sono negli appunti del Santini).
4. eliminazione dei terminali: in tutte le produzioni $A \rightarrow \alpha$ con $|\alpha| > 1$ si introducono nonterminali per ogni terminale.

Supponiamo di avere le produzioni

$$A \rightarrow Aaabc$$

$$A \rightarrow bC$$

$$A \rightarrow bb$$

introduciamo i nonterminali X_a e X_b e le regole

$$A \rightarrow AX_aX_aX_bC$$

$$A \rightarrow X_bC$$

$$A \rightarrow X_bX_b$$

$$X_a \rightarrow a$$

$$X_b \rightarrow b$$

5. binarizzazione delle produzioni: per ogni produzione $A \rightarrow B_1B_2 \dots B_k$ con $k > 2$, si introducono delle produzioni intermedie

$$A \rightarrow B_1Z_1$$

$$Z_1 \rightarrow B_2Z_2$$

$$\vdots$$

$$Z_{k-2} \rightarrow B_{k-1}B_k$$

Date le produzioni

$$S \rightarrow aB$$

$$S \rightarrow bA$$

$$A \rightarrow a$$

$$A \rightarrow aS$$

$$A \rightarrow bAA$$

$$B \rightarrow b$$

$$B \rightarrow bS$$

$$B \rightarrow aBB$$

questa è già priva di ϵ -produzioni, produzioni unitarie e tutti i simboli sono utili.

Ora eliminiamo i terminali e otteniamo

$$S \rightarrow X_a B$$

$$S \rightarrow X_b A$$

$$A \rightarrow a$$

$$A \rightarrow X_a S$$

$$A \rightarrow X_b AA$$

$$B \rightarrow b$$

$$B \rightarrow X_b S$$

$$B \rightarrow X_a BB$$

$$X_a \rightarrow a$$

$$X_b \rightarrow b$$

ed ora binarizziamo le produzioni

$$S \rightarrow X_a B$$

$$S \rightarrow X_b A$$

$$A \rightarrow a$$

$$A \rightarrow X_a S$$

$$A \rightarrow X_b E_1$$

$$E \rightarrow AA$$

$$B \rightarrow b$$

$$B \rightarrow X_b S$$

$$B \rightarrow X_a E_2$$

$$E_2 \rightarrow BB$$

$$X_a \rightarrow a$$

$$X_b \rightarrow b$$