Automi e Linguaggi Formali - Esercizio

Gabriel Rovesti

Anno Accademico 2024-2025

Esercizio Shuffle Perfetto

Siano $L_A = (V_A, \Sigma_A, R_A, S_A)$ e $L_B = (V_B, \Sigma_B, R_B, S_B)$ due linguaggi context-free. Definiamo lo shuffle alternato di A e B come il linguaggio: $L = \{x_1y_1x_2y_2...x_ny_n \mid x_i \in A, y_i \in B, n \geq 0\}$.

Dimostrare che se A e B sono linguaggi context-free, allora anche L è un linguaggio context-free.

Soluzione

Teorema 1. *Se A e B sono linguaggi context-free, allora lo shuffle alternato $L = \{x_1y_1x_2y_2...x_ny_n \mid x_i \in A, y_i \in B, n \geq 0\}$ è anch'esso un linguaggio context-free.*

Proof. Dato che A e B sono linguaggi context-free, esistono delle grammatiche context-free $G_A = (V_A, \Sigma_A, R_A, S_A)$ e $G_B = (V_B, \Sigma_B, R_B, S_B)$ tali che $L(G_A) = A$ e $L(G_B) = B$.

Senza perdita di generalità, possiamo assumere che $V_A \cap V_B = \emptyset$ (se non lo sono, possiamo rinominare i simboli non terminali per renderli disgiunti).

Costruiamo una nuova grammatica context-free $G = (V, \Sigma, R, S)$ dove:

- $V = V_A \cup V_B \cup \{S, X, Y\}$, dove S, X, Y sono nuovi simboli non terminali non presenti in V_A o V_B
- $\Sigma = \Sigma_A \cup \Sigma_B$
- R contiene:
  - Tutte le regole di produzione in R_A e R_B
  - $S \rightarrow XYS \mid \varepsilon$
  - $X \rightarrow S_A$
  - $Y \rightarrow S_B$

Dimostriamo che $L(G) = L$:

1. Dimostriamo che $L(G) \subseteq L$: Consideriamo una stringa $w \in L(G)$. Se $w = \varepsilon$, allora $w \in L$ perché $n = 0$ è ammesso nella definizione di L. Altrimenti, w è derivabile da S utilizzando la regola $S \rightarrow XYS$ ripetutamente, seguita alla fine da $S \rightarrow \varepsilon$. Questo porta a una derivazione del tipo: $S \Rightarrow XYS \Rightarrow XY(XYS) \Rightarrow XY(XY(XYS)) \Rightarrow ... \Rightarrow XY...XY\varepsilon = (XY)^n$

Sostituendo X con S_A e Y con S_B, otteniamo: $S \Rightarrow S_AS_BS \Rightarrow S_AS_B(S_AS_BS) \Rightarrow ... \Rightarrow (S_AS_B)^n$

Poiché S_A può derivare qualsiasi stringa $x_i \in A$ e S_B può derivare qualsiasi stringa $y_i \in B$, la stringa w sarà della forma $x_1y_1x_2y_2...x_ny_n$ dove $x_i \in A$ e $y_i \in B$. Quindi, $w \in L$.

2. Dimostriamo che $L \subseteq L(G)$: Consideriamo una stringa $w = x_1y_1x_2y_2...x_ny_n \in L$ dove $x_i \in A$ e $y_i \in B$ per ogni i.

Se $n = 0$, allora $w = \varepsilon$ e possiamo derivare w in G usando la regola $S \to \varepsilon$.

Se $n > 0$, possiamo derivare w in G come segue:

$$\begin{aligned} S &\Rightarrow XYS \\ &\Rightarrow XY(XYS) \\ &\Rightarrow ... \\ &\Rightarrow (XY)^{n-1}XYS \\ &\Rightarrow (XY)^{n-1}XY\varepsilon \\ &\Rightarrow (XY)^n \end{aligned}$$

Sostituendo X con S_A e Y con S_B, otteniamo:

$$\begin{aligned} S &\Rightarrow S_AS_BS \\ &\Rightarrow S_AS_B(S_AS_BS) \\ &\Rightarrow ... \\ &\Rightarrow (S_AS_B)^{n-1}S_AS_BS \\ &\Rightarrow (S_AS_B)^{n-1}S_AS_B\varepsilon \\ &\Rightarrow (S_AS_B)^n \end{aligned}$$

Poiché S_A può derivare x_i e S_B può derivare y_i, possiamo continuare la derivazione per ottenere:

$$\begin{aligned} (S_AS_B)^n &\Rightarrow x_1S_B(S_AS_B)^{n-1} \\ &\Rightarrow x_1y_1(S_AS_B)^{n-1} \\ &\Rightarrow ... \\ &\Rightarrow x_1y_1...x_nS_B \\ &\Rightarrow x_1y_1...x_ny_n \end{aligned}$$

Quindi, w è derivabile in G, e $w \in L(G)$.

Poiché $L(G) \subseteq L$ e $L \subseteq L(G)$, concludiamo che $L(G) = L$. Dal momento che G è una grammatica context-free, L è un linguaggio context-free. □

Osservazione 1. La costruzione presentata mantiene una struttura semplice che sfrutta le grammatiche originali di A e B. È importante notare che, nonostante la possibile complessità delle grammatiche originali, la composizione tramite shuffle alternato preserva la natura context-free del linguaggio risultante.

Osservazione sulla conversione in Forma Normale di Chomsky: La grammatica G costruita nella dimostrazione non è direttamente in Forma Normale di Chomsky. Tuttavia, è ben noto che ogni grammatica context-free può essere convertita in una grammatica equivalente in Forma Normale di Chomsky. La conversione richiederebbe:

1. Eliminare la regola $S \to \varepsilon$ introducendo un nuovo simbolo iniziale 2. Sostituire le regole $X \to S_A$ e $Y \to S_B$ con appropriate regole in forma binaria 3. Scomporre la regola $S \to XYS$ in più regole binarie

Questo processo di conversione, sebbene non necessario per dimostrare che L è context-free, conferma ulteriormente che L appartiene alla classe dei linguaggi context-free.