

1 Il problema

Problema (9 punti): Considera la seguente funzione da $\{0, 1\}^* \rightarrow \{0, 1\}^*$:

$$\text{stutter}(w) = \begin{cases} \varepsilon & \text{se } w = \varepsilon \\ aa \cdot \text{stutter}(x) & \text{se } w = ax \text{ per qualche simbolo } a \text{ e parola } x \end{cases} \quad (1)$$

Dimostra che se L è un linguaggio context-free sull'alfabeto $\{0, 1\}$, allora anche il seguente linguaggio è context-free:

$$\text{stutter}(L) = \{\text{stutter}(w) \mid w \in L\}. \quad (2)$$

2 Soluzione

Per dimostrare che la classe dei linguaggi context-free è chiusa rispetto all'operazione stutter, utilizzeremo una costruzione diretta basata sulla grammatica. L'idea principale è di modificare la grammatica originale in modo che, per ogni simbolo terminale generato, ne vengano prodotti due consecutivi.

Teorema 1. *Se L è un linguaggio context-free sull'alfabeto $\{0, 1\}$, allora $\text{stutter}(L) = \{\text{stutter}(w) \mid w \in L\}$ è anch'esso un linguaggio context-free.*

Proof. Sia L un linguaggio context-free sull'alfabeto $\{0, 1\}$. Quindi esiste una grammatica context-free $G = (V, \Sigma, P, S)$ che genera L , dove:

- V è l'insieme dei simboli non terminali
- $\Sigma = \{0, 1\}$ è l'alfabeto
- P è l'insieme delle produzioni
- $S \in V$ è il simbolo iniziale

Possiamo assumere, senza perdita di generalità, che G sia in forma normale di Chomsky (CNF), quindi ogni produzione ha una delle seguenti forme:

- $A \rightarrow BC$ dove $A, B, C \in V$
- $A \rightarrow a$ dove $A \in V$ e $a \in \Sigma$
- $S \rightarrow \varepsilon$ (solo se $\varepsilon \in L$)

Costruiamo ora una nuova grammatica $G' = (V, \Sigma, P', S)$ che genera $\text{stutter}(L)$. L'insieme P' di produzioni è ottenuto modificando P come segue:

- Per ogni produzione della forma $A \rightarrow BC$ in P , aggiungiamo la stessa produzione $A \rightarrow BC$ a P'
- Per ogni produzione della forma $A \rightarrow a$ in P dove $a \in \{0, 1\}$, aggiungiamo la produzione $A \rightarrow aa$ a P'

- Se $S \rightarrow \varepsilon$ è in P , aggiungiamo la stessa produzione $S \rightarrow \varepsilon$ a P'

Dimostriamo ora che $L(G') = \text{stutter}(L)$.

Parte 1: $L(G') \subseteq \text{stutter}(L)$

Vogliamo dimostrare che ogni stringa $y \in L(G')$ ha la forma $y = \text{stutter}(w)$ per qualche $w \in L$.

Per induzione sulla lunghezza della derivazione di y in G' , dimostriamo che se $A \Rightarrow_{G'}^* y$ per qualche $A \in V$, allora esiste una stringa x tale che $A \Rightarrow_G^* x$ e $y = \text{stutter}(x)$.

Caso base: La derivazione ha lunghezza 1. In questo caso, possiamo avere:

- $A \Rightarrow_{G'} \varepsilon$ mediante la produzione $A \rightarrow \varepsilon$. In questo caso, abbiamo anche $A \Rightarrow_G \varepsilon$ e $\text{stutter}(\varepsilon) = \varepsilon$.
- $A \Rightarrow_{G'} aa$ mediante la produzione $A \rightarrow aa$ (che deriva da $A \rightarrow a$ in P). In questo caso, abbiamo $A \Rightarrow_G a$ e $\text{stutter}(a) = aa$.

Passo induttivo: Supponiamo che la proprietà sia vera per tutte le derivazioni di lunghezza al più n . Consideriamo una derivazione di lunghezza $n + 1$.

Se la prima produzione applicata è $A \rightarrow BC$, allora abbiamo:

$$A \Rightarrow_{G'} BC \Rightarrow_{G'}^* y_1 y_2 = y \quad (3)$$

dove $B \Rightarrow_{G'}^* y_1$ e $C \Rightarrow_{G'}^* y_2$ sono derivazioni di lunghezza al più n .

Per ipotesi induttiva, esistono stringhe x_1 e x_2 tali che:

- $B \Rightarrow_G^* x_1$ e $y_1 = \text{stutter}(x_1)$
- $C \Rightarrow_G^* x_2$ e $y_2 = \text{stutter}(x_2)$

Quindi, nella grammatica originale G , abbiamo:

$$A \Rightarrow_G BC \Rightarrow_G^* x_1 x_2 = x \quad (4)$$

Ora, per la definizione di stutter , abbiamo:

$$\text{stutter}(x) = \text{stutter}(x_1 x_2) = \text{stutter}(x_1) \cdot \text{stutter}(x_2) = y_1 \cdot y_2 = y \quad (5)$$

Quindi, per ogni $y \in L(G')$, esiste $w \in L$ tale che $y = \text{stutter}(w)$, e quindi $L(G') \subseteq \text{stutter}(L)$.

Parte 2: $\text{stutter}(L) \subseteq L(G')$

Vogliamo dimostrare che per ogni $w \in L$, abbiamo $\text{stutter}(w) \in L(G')$.

Se $w \in L$, allora esiste una derivazione $S \Rightarrow_G^* w$ in G . Dimostriamo per induzione sulla lunghezza di questa derivazione che se $A \Rightarrow_G^* u$ per qualche $A \in V$ e $u \in \Sigma^*$, allora $A \Rightarrow_{G'}^* \text{stutter}(u)$.

Caso base: La derivazione ha lunghezza 1.

- Se $A \Rightarrow_G \varepsilon$ mediante la produzione $A \rightarrow \varepsilon$, allora abbiamo anche $A \Rightarrow_{G'} \varepsilon$ e $\text{stutter}(\varepsilon) = \varepsilon$.
- Se $A \Rightarrow_G a$ mediante la produzione $A \rightarrow a$, allora in G' abbiamo $A \Rightarrow_{G'} aa = \text{stutter}(a)$ mediante la produzione $A \rightarrow aa$.

Passo induttivo: Supponiamo che la proprietà sia vera per tutte le derivazioni di lunghezza al più n . Consideriamo una derivazione di lunghezza $n + 1$.

Se la prima produzione applicata è $A \rightarrow BC$, allora abbiamo:

$$A \Rightarrow_G BC \Rightarrow_G^* u_1 u_2 = u \quad (6)$$

dove $B \Rightarrow_G^* u_1$ e $C \Rightarrow_G^* u_2$ sono derivazioni di lunghezza al più n .

Per ipotesi induttiva, abbiamo:

- $B \Rightarrow_{G'}^* \text{stutter}(u_1)$
- $C \Rightarrow_{G'}^* \text{stutter}(u_2)$

Quindi, in G' , possiamo costruire la derivazione:

$$A \Rightarrow_{G'} BC \Rightarrow_{G'}^* \text{stutter}(u_1) \cdot \text{stutter}(u_2) = \text{stutter}(u_1 u_2) = \text{stutter}(u) \quad (7)$$

Quindi, per ogni $w \in L$, abbiamo $\text{stutter}(w) \in L(G')$, e quindi $\text{stutter}(L) \subseteq L(G')$.

Dalle parti 1 e 2, abbiamo dimostrato che $L(G') = \text{stutter}(L)$. Poiché G' è una grammatica context-free, $\text{stutter}(L)$ è un linguaggio context-free. \square

2.1 Osservazioni

È importante notare alcune proprietà specifiche dell'operazione stutter :

1. L'operazione stutter duplica ogni simbolo della stringa originale. Ad esempio, $\text{stutter}(010) = 001100$.
2. La lunghezza di $\text{stutter}(w)$ è esattamente il doppio della lunghezza di w , eccetto per il caso di $w = \varepsilon$.
3. La grammatica G' mantiene la struttura sintattica di G , modificando solo le produzioni che generano simboli terminali.

Questa dimostrazione mostra che l'operazione stutter preserva la proprietà di essere context-free. La chiave della dimostrazione è che possiamo modificare in modo sistematico le produzioni della grammatica originale per ottenere una nuova grammatica che genera esattamente le stringhe "stutterate" del linguaggio originale.