

Closure Properties of Context-Free Languages

Let L_1, L_2 be context-free languages.

1. Closure under union

Theorem. *If L_1 and L_2 are context-free, then $L_1 \cup L_2$ is context-free.*

Let

$$G_1 = (V_1, \Sigma, R_1, S_1), \quad G_2 = (V_2, \Sigma, R_2, S_2)$$

be CFGs generating L_1 and L_2 . Rename variables if necessary so that

$$V_1 \cap V_2 = \emptyset.$$

Construct

$$G = (V_1 \cup V_2 \cup \{S\}, \Sigma, R, S),$$

where S is fresh and

$$R = R_1 \cup R_2 \cup \{S \rightarrow S_1, S \rightarrow S_2\}.$$

We prove

$$L(G) = L_1 \cup L_2.$$

First containment: $L(G) \subseteq L_1 \cup L_2$. Take $w \in L(G)$. Then

$$S \Rightarrow^* w.$$

The first step must be either $S \Rightarrow S_1$ or $S \Rightarrow S_2$. If $S \Rightarrow S_1$, then all subsequent productions are from R_1 , so

$$S_1 \Rightarrow^* w,$$

and hence $w \in L_1$. Similarly, if $S \Rightarrow S_2$, then $w \in L_2$. Thus $w \in L_1 \cup L_2$, and therefore

$$L(G) \subseteq L_1 \cup L_2.$$

Second containment: $L_1 \cup L_2 \subseteq L(G)$. Take $w \in L_1 \cup L_2$. If $w \in L_1$, then

$$S_1 \Rightarrow^* w,$$

so

$$S \Rightarrow S_1 \Rightarrow^* w.$$

Hence $w \in L(G)$. The case $w \in L_2$ is identical. Therefore

$$L_1 \cup L_2 \subseteq L(G).$$

Combining the two containments,

$$\boxed{L(G) = L_1 \cup L_2.}$$

2. Closure under concatenation

If L_1 and L_2 are context-free, then L_1L_2 is context-free. Using a fresh start variable S , add

$$S \rightarrow S_1S_2.$$

Then every generated word is a word of L_1 followed by a word of L_2 , and every such concatenation can be generated. Hence

$$L(G) = L_1L_2.$$

3. Closure under Kleene star

If L_1 is context-free, then L_1^* is context-free. Let S_1 be the start variable of a CFG for L_1 . Introduce a fresh start variable S and add

$$S \rightarrow S_1S \mid \varepsilon.$$

This generates any finite concatenation of words from L_1 , including zero copies. Thus

$$L(G) = L_1^*.$$

4. Mixing CFLs with regular languages

Every regular language is context-free. Therefore, if L is context-free and R is regular, then

$$L \cup R, \quad LR, \quad RL$$

are context-free.

The more interesting mixed operation is intersection.

Theorem. *If L is context-free and R is regular, then $L \cap R$ is context-free.*

Proof. Let P be a PDA recognizing L and let D be a DFA recognizing R . Construct a PDA P' that simulates P and D simultaneously. Its states are pairs

$$(q, r),$$

where q is a state of P and r is a state of D .

Write the transition relation of P in Sipser's notation. For an input-consuming move, if

$$(q', B) \in \delta_P(q, a, A) \quad \text{and} \quad \delta_D(r, a) = r',$$

then define

$$((q', r'), B) \in \delta'((q, r), a, A).$$

Thus, whenever the PDA consumes the symbol a , the DFA component also makes its a -transition, while the stack operation is exactly the one performed by P .

For an ε -transition of P , if

$$(q', B) \in \delta_P(q, \varepsilon, A),$$

then define

$$((q', r), B) \in \delta'((q, r), \varepsilon, A).$$

Thus an ε -move changes only the PDA component; the DFA state remains r .

The start state of P' is (q_0, r_0) , and the accepting states are

$$F' = F_P \times F_D.$$

Hence P' accepts exactly the strings accepted by both P and D , so

$$L(P') = L(P) \cap L(D) = L \cap R.$$

Therefore $L \cap R$ is context-free. □

5. CFLs are not closed under intersection

Consider

$$L_1 = \{a^i b^i c^j : i, j \geq 0\}$$

and

$$L_2 = \{a^i b^j c^j : i, j \geq 0\}.$$

Both are context-free (think why), but

$$L_1 \cap L_2 = \{a^n b^n c^n : n \geq 0\},$$

which is not context-free. Hence

CFLs are not closed under intersection.

6. CFLs are not closed under complement

Suppose CFLs were closed under complement. Since they are closed under union,

$$\overline{L_1} \cup \overline{L_2}$$

would be context-free whenever L_1, L_2 are context-free. Taking complements again,

$$\overline{\overline{L_1} \cup \overline{L_2}}$$

would be context-free. By De Morgan's law,

$$\overline{\overline{L_1} \cup \overline{L_2}} = L_1 \cap L_2,$$

contradicting the previous section. Therefore

CFLs are not closed under complement.

7. Using closure to prove a language is not context-free

A useful method is:

To prove that L is not context-free, choose a regular language R such that $L \cap R$ is a known non-context-free language.

Consider

$$L = \{w \in \{a, b, c\}^* : n_a(w) = n_b(w) = n_c(w)\}.$$

Take the regular language

$$R = a^*b^*c^*.$$

Then

$$L \cap R = \{a^n b^n c^n : n \geq 0\}.$$

If L were context-free, then $L \cap R$ would be context-free because CFLs are closed under intersection with regular languages. This is a contradiction. Hence

L is not context-free.

The regular language R acts as a *filter*: it restricts arbitrary strings to the block form $a^*b^*c^*$, exposing a known non-CFL.

8. Summary

Operation	Are CFLs closed under it?
Union	✓
Concatenation	✓
Kleene star	✓
Intersection with a regular language	✓
Intersection	×
Complement	×