CHAPTER 8: PROPERTIES OF CONTEXT-FREE LANGUAGES*

Peter Cappello
Department of Computer Science
University of California, Santa Barbara
Santa Barbara, CA 93106
cappello@cs.ucsb.edu

- Please read the corresponding chapter before attending this lecture.
- These notes are supplemented with figures, and material that arises during the lecture in response to questions.
- Please report any errors in these notes to `cappello@cs.ucsb.edu`. I'll fix them immediately.

---

[*] Based on **An Introduction to Formal Languages and Automata**, 3rd Ed., Peter Linz, Jones and Bartlett Publishers, Inc.

8.2 CLOSURE PROPERTIES & DECISION ALGORITHMS FOR CFLS

CLOSURE OF CONTEXT-FREE LANGUAGES

Thm. 8.3: CFLs are closed under the operators $+$, $\cdot$, and $*$.
Proof:

1. Let L_1 and L_2 be CFLs generated by $G_1 = (V_1, T_1, S_1, P_1)$ and $G_2 = (V_2, T_2, S_2, P_2)$, respectively.
2. Assume WLOG that $V_1 \cap V_2 = \emptyset$.

3. First, we prove that CFLs are closed under $\cup$.

   (a) Let

   $$G_3 = (V_1 \cup V_2 \cup \{S_3\}, T_1 \cup T_2, S_3, P_1 \cup P_2 \cup \{S_3 \to S_1 \mid S_2\}),$$

   where $S_3 \notin V_1 \cup V_2$.

   (b) Clearly, $L(G_3)$ is a CFL and $L_1 \cup L_2 = L(G_3)$.
   (Why did we insist that $V_1 \cap V_2 = \emptyset$?)

4. Now, we show that CFLs are closed under concatenation.
   (a) Let

   $$G_4 = (V_1 \cup V_2 \cup \{S_4\}, T_1 \cup T_2, S_4, P_1 \cup P_2 \cup \{S_4 \to S_1S_2\}),$$

   where $S_4 \notin V_1 \cup V_2$.
   (b) Then, $L(G_4) = L(G_1)L(G_2)$.
5. Finally, we show that CFLs are closed under the Kleene-star operator.
   (a) Let

   $$G_5 = (V_1 \cup \{S_5\}, T_1, S_5, P_1 \cup \{S_5 \to S_1S_5 \mid \lambda\}),$$

   where $S_5 \notin V_1$.
   (b) Then, $L(G_5) = L(G_1)^*$.

Thm. 8.4: CFLs are not closed under intersection and complementation.
Proof:

1. Let

$$\begin{aligned} L_1 &= \{a^n b^n c^m : n, m \geq 0\} = \{a^n b^n : n \geq 0\} \cdot \{c^m : m \geq 0\} \\ L_2 &= \{a^n b^m c^m : n, m \geq 0\} = \{a^n : n \geq 0\} \cdot \{b^m c^m : m \geq 0\} \end{aligned}$$

2. L_1 and L_2 are the concatenation of CFLs, hence are themselves CFLs.
3. $L_1 \cap L_2 = \{a^n b^n c^n : n \geq 0\}$, which is not a CFL.
4. Therefore CFLs are not closed under intersection.
5. Since $L_1 \cap L_2 = \overline{\overline{L_1} \cup \overline{L_2}}$, CFLs are not closed under complementation.

Thm. 8.5: Let L_1 be a CFL and L_2 be regular. Then, $L_1 \cap L_2$ is a CFL.
Proof:

1. Let $M_1 = (Q, \Sigma, \Gamma, \delta_1, q_0, z, F_1)$ be an NPDA that accepts L_1.
2. Let $M_2 = (P, \Sigma, \delta_2, p_0, F_2)$ be a DFA that accepts L_2.
3. We define an NPDA, $\widehat{M}$, that concurrently simulates M_1 and M_2.
   $\widehat{M} = (\widehat{Q}, \Sigma, \Gamma, \hat{\delta}, z, \widehat{F})$, where

$$\begin{aligned} \widehat{Q} &= Q \times P, \\ \widehat{q_0} &= (q_0, p_0), \\ \widehat{F} &= F_1 \times F_2, \end{aligned}$$

   and define $\hat{\delta}$ so that

$$((q_k, p_l), x) \in \hat{\delta}((q_i, p_j), a, b) \Leftrightarrow$$

$$\begin{aligned} (q_k, x) &\in \delta_1(q_i, a, b), \text{ and} \\ \delta_2(p_j, a) &= p_l. \end{aligned}$$

If $a = \lambda$, then $p_j = p_l$.

4. Then,

$$((q_0, p_0), w, z \vdash^*_{\widehat{M}} ((q_r, p_s), \lambda, x),$$

with $q_r \in F_1$ and $p_s \in F_2$ if and only if

$$\begin{array}{rcl} (q_0, w, z) & \vdash^*_{M_1} & (q_r, \lambda, x), \text{ and} \\ \delta_2^*(p_0, w) & = & p_s. \end{array}$$

Thus, $w \in L(\widehat{M}) \Leftrightarrow w \in L(M_1)$ and $w \in L(M_2)$

That is, $L(M_1) \cap L(M_2) = L_1 \cap L_2$.

Example: Is $L = \{a^n b^n : n \geq 0, n \notin [153, 257]\}$ a CFL?

Answer Yes, since

1.
$$L = \{a^n b^n : n \geq 0\} \cap \overline{\{a^n b^n : n \in [153, 257]\}}.$$
2. $\{a^n b^n : n \geq 0\}$ is a CFL.
3. $\{a^n b^n : n \in [153, 257]\}$ is finite, hence regular.
4. Regular languages are closed under complement.
5. The intersection of a CFL with a regular language is a CFL.

Exercise

Let L_1 be a CFL and L_2 be a regular language.
Is $L_1 \cap L_2$ a regular language?

Example 8.8

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

Amswer

No. If it were, then

$$L \cap a^*b^*c^* = \{a^nb^nc^n : n \geq 0\}$$

would be a CFL, since

1. $a^*b^*c^*$ is regular.
2. The intersection of a CFL with a regular language is a CFL.

SOME DECIDABLE PROPERTIES OF CFLS

Thm. 8.6: Given CFG G, there is an algorithm for deciding if $L(G) = \emptyset$.
Proof:

1. Use the algorithm for eliminating useless variables.
2. If the start variable is useless, return true ($L(G) = \emptyset$), else return false.

Thm. 8.7: Given CFG G, $|L(G)| < \infty$ is decidable.
Proof:

1. Let $G = (V, T, S, P)$.

   Assume WLOG that G has no λ-productions, no unit productions, and no useless variables or symbols.

2. Suppose there is an $A \in V$ such that $A \Rightarrow^* xAy$.

3. x and y cannot both be empty, since G has no unit productions.

4. Since G has no useless variables, $S \Rightarrow^* uAv \Rightarrow^* uzy$, where u, v, and $z \in T^+$.

5. Since A is repeating, $S \Rightarrow^* uAv \Rightarrow^* ux^nAy^n \Rightarrow^* ux^nzy^nv$, for $n = 0, 1, \ldots$.

6. Thus, under the assumption that there is a repeating variable, $L(G)$ is infinite.

7. If there is no repeating symbol, then $|L(G)| < \infty$.
8. The question thus reduces to whether there is a repeating variable.
9. There is an algorithm to decide if G has a repeating variable:
   (a) Construct a directed graph where there is a node for each variable, and an directed edge from from node A to node B if there is a production of the form $A \to xBy$.
   (b) If the graph has a cycle, there is a repeating variable.
   (c) There is an algorithm for detecting whether or not a digraph has a cycle (e.g., depth-first search).

Example $G_E = (\{E, T, F\}, \{a, +, *, [,]\}, E, P)$ with productions

$$\begin{aligned} E &\rightarrow E + T \mid T, \\ T &\rightarrow T * F \mid F, \\ F &\rightarrow [\ E\ ] \mid a. \end{aligned}$$

Is $|L(G_E)| < \infty$?

Let G_1 and G_2 be CFGs. The question: “Is $L(G_1) = L(G_2)$?” is not decidable.