CS 205
Lecture 26
17/3/21

Closure Properties of CFLs

Let L_i be a CFL generated by CFG $G_i = (V_i, T_i, S_i, P_i)$ for $i = 1, 2$. Assume : $V_1 \cap V_2 = \emptyset$.

Closure properties

1. Union : Is $L_1 \cup L_2$ a CFL ?
   Yes $S \to S_1 \mid S_2$

2. Concatenation : $L_1 L_2$ a CFL ?
   Yes $S \to S_1 S_2$

3. Kleene Closure : L_1^* a CFL ?
   Yes $S \to S_1 S \mid \varepsilon$

4. Intersection : $L_1 \cap L_2$ need not be a CFL.
   Ex $L_1 = \{a^i b^i c^j \mid i, j \geq 0\}$ is a CFL.
   $L_2 = \{a^i b^j c^j \mid i, j \geq 0\}$ " " ".
   But $L_1 \cap L_2 = \{a^n b^n c^n \mid n \geq 0\}$ is a CFL.

5. Intersection with a regular language :
   L is a CFL (accepted by a PDA P by final state)
   R " " regular language (acc. by a DFA M)
   Then $L \cap R$ is a CFL.
   Proof : Take the product $P \times M$

6. Complementation : L is a CFL. Is $\bar{L}$ a CFL ?
   No. $A \cap B = \overline{\bar{A} \cup \bar{B}}$ Contradiction.
   Ex $L = \{x \mid x \text{ is not of the form } ww\}$ is a CFL. (Exercise !)
   $\bar{L} = \{ww \mid w \in \{0,1\}^*\}$ is not a CFL.
   If $\bar{L}$ were a CFL then
   $L' = \bar{L} \cap a^* b^* a^* b^* = \{a^i b^j a^i b^j \mid i, j \geq 0\}$
   would be a CFL, but it is not.

7. Set Difference : No $\bar{L} = \Sigma^* \setminus L$

Decision Problems

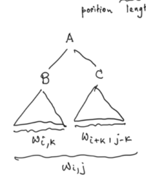
(1) Emptiness : Given CFG G, is $L(G)$ empty ?
   Soln : Check whether S is generating.

(2) Membership : Given CFG G in CNF and string w, is $w \in L(G)$?
   Parsing

Idea : Dynamic Programming
Let w_{ij} be the substring of w that starts at position i and is of length j.
Key idea : For every $A \in V$ and every substring $w_{i,j}$ of w, determine if $A \stackrel{*}{\Rightarrow} w_{i,j}$
Then $w \in L(G)$ iff $S \stackrel{*}{\Rightarrow} w_{1,n}$ where $n = |w|$.
Q : How do we determine if $A \stackrel{*}{\Rightarrow} w_{i,j}$?
A : Dynamic Programming

Base Case : $A \stackrel{*}{\Rightarrow} w_{i,1}$ iff $A \to w_{i,1}$ (true for CFG in CNF)
So for each i, $1 \leq i \leq |w| = n$, we can determine if $A \stackrel{*}{\Rightarrow} w_{i,1}$.

Inductive Step : Assume that we know for every variable X and every $w_{i,l}$ (where $l < j$) if $X \stackrel{*}{\Rightarrow} w_{i,l}$. We want to determine if $A \stackrel{*}{\Rightarrow} w_{i,j}$ for a given $A \in V$.
Now, $A \stackrel{*}{\Rightarrow} w_{i,j}$ iff there are variables B and C and some $1 \leq k < j$ s.t.
$A \to BC$ is a prodn., and
$B \stackrel{*}{\Rightarrow} w_{i,k}$ and
$C \stackrel{*}{\Rightarrow} w_{\underbrace{i+k}_{\text{position}}, \underbrace{j-k}_{\text{length}}}$



Cocke – Younger – Kasami (CYK) Algorithm

Maintains $X_{i,j} = \{A \mid A \stackrel{*}{\Rightarrow} w_{i,j}\}$
Initialize : $X_{i,1} = \{A \mid A \to w_{i,1} \text{ is a prodn.}\}$ for $1 \leq i \leq n$.

```
for j ← 2 to n do
    for i ← 1 to n-j+1 do
        X_{i,j} ← ∅
        for k ← 1 to j-1 do
            X_{i,j} ← X_{i,j} ∪ {A | A → BC,
                                B ∈ X_{i,k} and
                                C ∈ X_{i+k, j-k}}
```

Correctness : After each iteration of the outermost loop $X_{i,j}$ contains exactly the set of variables A that can derive $w_{i,j}$ for each i.

Complexity : $O(n^3)$

Undecidable problems about CFGs

1. Is $L(G) = T^*$?
2. Is $L(G_1) \cap L(G_2) = \emptyset$?
3. Is $L(G_1) = L(G_2)$?
4. Is G ambiguous ?
5. Is $L(G)$ inherently ambiguous ?