Unit-5

Properties of Context-free Languages (CFL)

Context-free Languages (CFLs) are characterized by their ability to be described by Context Free grammars, which are rules for generating strings in the language. Key properties of CFLs are: a pumping lemma for identifying non-context-free languages, Closure properties of CFL and decisions properties related to membership and emptiness.

5.1. The pumping lemma for CFLs

Theorem:- Let L be the Context Free language and it is finite. Let Z be sufficiently long string and

$$Z \in L \text{ and so that}$$

$$|Z| \leq n$$

Where n is ~~the~~ some positive integer. If the string Z can be decomposed into combinations of strings

$$Z = uvwxy$$

Such that,

$$|vwx| \leq n$$

$$|vx| \geq 1, \text{ then}$$

$$uv^i w x^i y \in L \text{ for } i = 0, 1, 2, 3, \ldots$$

Proof:- According to pumping lemma it is assumed that the string $Z \in L$ is finite and is CFL. This proof leads to following two cases

case1: To generate a sufficiently long string Z, one or more variable must be recursive hence should be apply more than ones. An infinite string can be generated if the grammar has some non-terminals A, such that

$$A \xrightarrow{*} \alpha A B$$

Case2: $Z \in L$ implies that after applying some or all production some no. of time, we get finally string of terminals and the derivation stops.

Proof of case1 :- let us assume that the language is finite and the grammar has finite no. of variables and say all these are useful variables, and each product has finite length. To derive sufficiently long strings using such production the grammar should have one or more recursive variable. Assume that no variable is recursive.

Since variable is recursive,

HK-130.

each variables must be define only interms of terminals and/or other variables. Since those variables are also non-recursive, they have to be defined interms of terminals and other variables. If we keep applying the production ~~hints~~ like this, there are no variables at all and finally we get strings of terminal and generated string is finite. From this it can be ~~contr~~ concluded that here is limit on the length of string i.e. generated from the start symbol S. This contradicts our assumption that language is finite. Therefore, the assumption that one or more variables are non-recursive is incorrect. It means that one or more variables are recursive and hence the proof.

Proof of Case2:-

Let $Z \in L$ is sufficiently long string and therefore the derivation must have involve recursive use of some variables A. And the derivation must have the form

$S \overset{*}{\Rightarrow} VAX$.

Since, A is use recursively the derivation can take the following form:

HK131

$S \xRightarrow{*} uAy \xRightarrow{*} uvAxy \xRightarrow{*} uvwxy \xRightarrow{} z$

This implies that the following derivation

$A \xRightarrow{*} VAX$

and

$A \xRightarrow{*} w$ are also ~~variable~~ possible

From this we can conclude that the derivation

$A \xRightarrow{*}$ ~~vAx~~ vwx which also possible

Next, we need to prove that the longest string vwx is generated without recursion. This can be easily proved. Since CFG that generates CFL doesn't contain any production or unit production. It shows that every derivation step either increases the length of the sentential (using recursive variable) or introduces a terminal. The derivation $A \xRightarrow{*} VAX$ shows that

$$|VX| \geq 1$$

From the derivation $S \xRightarrow{*} uAy \xRightarrow{*} uvAXY$, It should be noted that $uvAXy$ occurs in the derivation and $A \xRightarrow{*} VAX$ and $A \xRightarrow{*} w$ are also possible. This follows that

$\boxed{uv^i w x^i y \in L}$ and ~~also~~ hence the proof.

HK-132

✡ Application of Pumping lemma for CFL

The pumping lemma for CFL is used to prove that certain languages are not context free. It should be noted that the pumping Lemma cannot be used to prove the that certain language are Context free.

The General strategy used to proved that the given language is not Context free is given below:

(i) Assume that the language L is infinite and is Context free.

(ii) Select the string say Z and divide it into sub strings u, v, w, x and y. Such that

$$Z = uvwxy$$

where,

$$|vwx| \leq n$$

$$|vx| \geq 1$$

(iii) Find any i such that $uv^iwx^iy \notin L$.

According to pumping lemma, $uv^iwx^iy \in L$.

So, the result is a contradiction to the assumption that the language is Context free.

Therefore, the given language L is not Context free.

HK-133

Page

Q.1. Show that $L = \{a^n b^n c^n \mid n \geq 0\}$ is not Context free.

Solution:-

Step 1:- Let L is Context free and infinite. Let
$$Z = a^n b^n c^n \in L.$$

Step 2:- Note that $|Z| \geq n$ and so we can split Z into $uvwxy$ i.e. $Z = uvwxy$

such that, $|vwx| \leq n$

and $|vx| \geq 1$ (Case-I: v and x each contain only one type of symbol)

and so, $uv^iwx^iy \in L$ for $i = 0, 1, 2, 3, \dots$

Step 3:- The string vwx is ~~written~~ within a^n.

Let $v = a^j$ and $x = a^k$ where,
$$|vx| = j + k \geq 1$$
$$|vwx| \leq n$$

which can be shown pictorially as

$$\underbrace{\overbrace{a \cdots a}^{n}}_{uvwx}\ \underbrace{\overbrace{b \cdots b}^{n}\ \overbrace{c \cdots c}^{n}}_{y}$$

If $n = 4$ then
$$Z = a^4 b^4 c^4$$
i.e.
$$\underbrace{a}_{u}\underbrace{aa}_{v}\underbrace{abbbbc}_{w}\underbrace{c}_{x}\underbrace{cc}_{y}$$

for $i = 2$ uv^iwx^iy becomes uv^2wx^2y

i.e. aaaaaabbbbccccc

i.e. $a^6 b^4 c^5 \notin L$

i.e. Contradiction

Now, according to pumping lemma $uv^2wx^2y \in L$ for $i = 2$ which can be pictorially shown as

$$\underbrace{\overbrace{a \cdots\cdots a}^{n+j+k}}_{uvwx}\ \underbrace{\overbrace{b \cdots b}^{n}\ \overbrace{c \cdots c}^{n}}_{y}$$

Case II: E... more than one kind of symbols.

If $n=4$

$\underbrace{aa}_{u}\underbrace{aab}_{v}\underbrace{b}_{w}\underbrace{b}_{x}\underbrace{b\,cccc}_{z}$ $\Rightarrow$ uv^iwx^iy $(i=2)$

uv^2wx^2y

then $z = aaaabbaabbbbbcccc$

$= a^4b^2a^2b^3c^4 \notin L$

It should be noted that $uv^iwx^iy = a^{n+j+k}b^nc^n$. Here, $uv^2wx^2y \in a^{n+j+k}b^nc^n$ which is not the case. Therefore it is the contradiction to the assumption that the language is Context free. So, the given language $L = \{a^nb^nc^n \mid n \geq 0\}$ is not Context free.

Q.2. Show that $L = \{ww \mid w \in (a+b)^*\}$ is not Context free.

Soln → Step 1: Let L is Context Free and is infinite. Let $z = ww \in L$ or $z = a^nb^na^nb^n \in L$ where $w = a^nb^n$

Step 2: Note that $|z| > n$ and so we can split z into $uvwxy$ such that

$$|vwx| \leq n$$

and $|vx| \geq 1$

and so, $uv^iwx^iy \in L$ for $i = 0, 1, 2, 3, \ldots$

Step 3:- The string vwx is within w. Let $v = a^j$ and $x = a^k$. where,

$$|vx| = j+k \geq 1$$

$$|vwx| \leq n$$

which can be shown pictorially as

$$\underbrace{\overbrace{a\ldots\ldots a}^{n}}_{uvwx}\underbrace{\overbrace{b\ldots\ldots b}^{n}\overbrace{a\ldots\ldots a}^{n}\overbrace{b\ldots\ldots b}^{n}}_{y}$$

Now, according to pumping lemma, $uv^2wx^2y \in L$ for $i = 2$ which can be pictorially shown as

$$\underbrace{\overset{n+j+k}{a \ldots\ldots a}}_{uv^2wx^2}\ \underbrace{\overset{n}{b \ldots\ldots b}\ \overset{n}{a \ldots\ldots a}\ \overset{n}{b \ldots\ldots b}}_{y}$$

It should be noted that
$uv^iwx^iy = a^{n+j+k}b^na^nb^n = uv^2wx^2y \notin L$
when $|j+k| \geq 1$ because the string ww ~~of the~~ is not generated. But by pumping $uv^2wx^2y \in L$ which is the contradiction to the assumption that the language is Context Free. Therefore the given language $L = \{ww \mid w \in (a+b)^*\}$ is not Context free.

✡ Closure Properties of CFL

Theorem :- If L_1 and L_2 are in CFLs, then $L_1 \cup L_2$, $L_1 \cdot L_2$ and L_1^* also denote the CFL and so the context free languages are closed under union, concatenation and star(*) closure.

Proof :- Let L_1 and L_2 are two CFLs generated by the CFGs.

$G_1 = (V_1, T_1, P_1, S_1)$
$G_2 = (V_2, T_2, P_2, S_2)$

respectively and assume that V_1 and V_2 are disjoint.

classmate
Date
Page

case1: Union of two CFLs is CFLs. Consider the language L_3 generated by the grammar $G_3 = (V_1 \cup V_2 \cup S_3, T_1 \cup T_2, P_3, S_3)$, where S_3 is the start symbol for the grammar G_3 and $S_3 \notin (V_1 \cup V_2)$ and $P_3 = P_1 \cup P_2 \cup \{S_3 \longrightarrow S_1 / S_2\}$

It is clear from this grammar that the grammar G_3 is Context free and the languages generated by this grammar is context free and so $L_3 = L_1 \cup L_2$

case2: Concatenation of two CFLs is CFL.
Let the language L_4 is generated by the grammar $G_4 = (V_1 \cup V_2 \cup S_4, T_1 \cup T_2, P_4, S_4)$. Where, S_4 is the start symbol for the grammar G_4 and $S_4 \notin (V_1 \cup V_2)$ and $P_4 = P_1 \cup P_2 \cup \{S_4 \rightarrow S_1 S_2\}$

It is clear from this that the grammar G_4 is context free and the language generated by this grammar is Context free and so, $L_4 = L_1 \cdot L_2$

Case3:- CFL are closed under star closure.
Consider the language L_5 generated by the grammar $G_5 = (V_1 \cup S_5, T_1, P_5, S_5)$
where, S_5 is the start symbol for the grammar G_5 and $P_5 = P_1 \cup \{S_5 \rightarrow S_1 S_5 | \epsilon\}$

AK-137

It is clear from the above production that the grammar G_5 is Context free and the language generated by this grammar is Context free and So, $L_5 = L_1^*$, this it is proved that the Context free languages are closed under union, ~~cont~~ concatenation and star closure.

✡ Decision algorithm for CFL

Given Context free Grammar $G = (V, T, P, S)$ Decision algorithm for CFL are given below

1. Algorithm for deciding whether a CFL L is empty:

Let $G = (V, T, P, S)$ be CFG then their excist an algorithm for deciding whether $L(G)$ is empty or not. Assume that $\epsilon \notin L(G)$. Then, use the algorithm for removing useless symbols and productions. If S is found to be useless the $L(G)$ is empty, if not, then $L(G)$ contains at least one element.

2. Algorithm for deciding whether a Context Free language L is Finite

Assume that a given CFG G contains a

MK-138

It is clear from the above production that the grammar G_5 is Context free and the language generated by this grammar is Context free and So, $L_5 = L_1^*$, this it is proved that the Context free languages are closed under union, ~~con~~ concatenation and star closure.

Note:- CFL are not closed under:- Intersection, Complement Reversal, Homomorphism & Inverse Homomorphism

Decision algorithm for CFL

Given Context free Grammar $G=(V,T,P,S)$ Decision algorithm for CFL are given below

1. Algorithm for deciding whether a CFL L is empty:

Let $G=(V,T,P,S)$ be CFG then there exist an algorithm for deciding whether $L(G)$ is empty or not. Assume that $\epsilon \notin L(G)$. Then, use the algorithm for removing useless symbols and productions. If S is found to be useless the $L(G)$ is empty, if not, then $L(G)$ contains at least one element.

2. Algorithm for deciding whether a Context Free language L is Finite

Assume that a given CFG G contains no ϵ production, not unit

MK-138

production and no useless symbols. Suppose the grammar G has a repeating variable in the sense that there exist some $A \in V$ for which there is a derivation $A \overset{*}{\Rightarrow} xAy$.
Since, A is neither nullable nor useless symbol, we have,

$$S \overset{*}{\Rightarrow} uAv \overset{*}{\Rightarrow} w$$

and $A \overset{*}{\Rightarrow} z$ where u, v and z are in T^* but then $S \overset{*}{\Rightarrow} uAv \overset{*}{\Rightarrow} ux^nAy^nv$
Is possible for all n, show that L(G) is infinite.

If no variable can ever repeat, the length of any derivation is bounded by $|V|$. In this case, L(G) is finite.

Thus to get an algorithm for determining L(G) is finite, we need only to determine whether the grammar G has some repeating i.e. recursive variables or not. If not then the given language is finite.

Q.1. Discuss the decision algorithms for Context Free languages.

Q.2. Discuss the closure properties of CFL

Solu^n → A Context-free language is closed under the following:

Union

In case L1 and L2 are two Context-free languages, L1UL2 will ~~be~~ also be Context free.

Example

In case $L1 = \{a^n b^n \mid n \geq 0\}$

The G1 corresponding grammar would have

$P: S1 \rightarrow aAb \mid ab$

In case $L2 = \{c^m d^m \mid m \geq 0\}$

The G2 corresponding grammar would have

$P: S2 \rightarrow cBd \mid \epsilon$

The union of L1 and L2 would be

$L = L1 \cup L2 = \{a^n b^n\} \cup \{c^m d^m\}$

Here, the G corresponding grammar would have the additional production, that is, $S \rightarrow S1 \mid S2$

Concatenation

In case L1 and L2 are CFLs then L1.L2 would also be Context free.

Example

The union of the languages L1 and L2

HK1440

would be $L = L1 \cdot L2 = \{a^n b^n c^m d^m\}$
The G corresponding grammar would have the additional production, that is, $S \rightarrow S_1 . S_2$

Kleen Star

In case L is a CFL, then L^* would also be Context free.

Example:-

In case $L = \{a^n b^n \mid n \geq 0\}$

Then, the G corresponding grammar would have
$P: S \rightarrow aAb \mid \varepsilon$

Thus, the Kleen Star $L1 = \{a^n b^n\}^*$
Here, the G_1 corresponding grammar would have additional productions, and they are
$S_1 \rightarrow SS1 \mid \varepsilon$