The properties of Context Free Grammar

March 30, 2023

Contents

1 The Normal Form of grammar

There are two kinds of normal forms that we should know, which are **Chomsky Normal Form** and **Greibach Normal Form**. The definitions have been introduced in previous chapter. You should check them out.

2 The simplification of grammar

This section is on its place. However, it has been learned in previous chapter. You may check previous chapter for more info.

3 The Pump Lemma of grammar

3.1 The content of Pump Lemma

Theorem 3.1 (Pump Lemma)**.** Let L be the language of a grammar. Then there exist n that if $|z| \geq n$, then $z = uvwxy$, suit the following conditions:

1. $|vwx| \leq n$. That is to say $|vwx|$ can't be too long.
2. $vx \neq \epsilon$. Since v, x are the pieces to be pumped, v or x should not be zero, that is to say, v and x can be both empty, *and* that is to say $vx \neq \epsilon$.
3. For all $i \geq 0$, $uv^i wx^i y$ is in L.

3.2 The application of Pump Lemma

Similarly, the Pump Lemma is used to prove a language L is not a context free language. Let us restate the lemma using the mathematical logic. If L is a context free language then

$$\exists n \in N \ \forall \alpha \in L (\forall uvwxy = \alpha(|vwx| \leq n, vx \neq \epsilon \rightarrow uv^i wx^i y \in L))$$

Let the proposition above be A, then we have L is CFL $\rightarrow A$. Thus, we have $\neg A \rightarrow L$ is not CFL. And $\neg A$ equates

$$\forall n \in N, \exists \alpha \in L, \exists uvwxy = \alpha(\neg(|vwx| \leq n, vx \neq \epsilon \rightarrow uv^i wx^i y \in L))$$

So we have the procedure here, similar to the one in previous chapter about the formal expressions, that we check for every n in $\mathbb{N}$, considering as a random variable[1], and find a $\alpha \in L$, and prove that for all kinds of $uvwxy$, there exists i s.t. uv^iwx^iy is not in L, where we often discuss about the different conditions.

Example 3.2. Use pump lemma to show that $L = \{0^n1^n2^n \mid n \geq 1\}$ is not context free language.

There is another example to show that the grammar can't describe the string that have two pairs of equal numbers of symbols.

Example 3.3. Let $L = \{0^i1^j2^i3^j\}$. Use pump lemma to prove that L is not context free language.

Mover, since pushdown automata are equivalent to context free grammar, you can easily see that (not prove that) $L = \{ww\}$ is not context free language.

Example 3.4. Let $L = \{ww\}$. Use pump lemma to prove that L is not context free language.

Given a n, we shall prove that if $z \in L$, and $z = uvwxy$, we have that uwy does not in L which leads to contradiction. We shall let $z = 0^n1^n0^n1^n$.

1. Let us talk about vwx first. Since $|vwx| \leq n$, we assume that vwx is all in the first block of 0's. Let $|vx|$ be k. Then, $|uwy| = 4n - k$ and moreover, since vwx is all in the first block, uwy starts with $0^{n-k}1^n$ for sure. If $uwy = tt$ for some t, then $|t| = 2n - k/2 \geq 2n - k$, viz., the length of t is longer than that of $0^{n-k}1^n$, and thus t should end with 0. However, uwy ends with 1. If uwy is tt then it should have ended with 0 since t end with 0, which, leads to a contradiction.

2. Suppose that vwx straddles the first block of 0's and the first block of 1's. The same, we assume that uwy can written as tt. Since k, which is the length of vx, is no bigger than n^2, we have $|uwy| \geq 3n$. Thus $|t| \leq 3n/2$. There are two possiblily: 1. vx contains no 1; 2, vx contains at least one 1. For situation 1., the discussion in (1) is also applied. For situation 2., We assert that t does not end with 1^n, for there is at least one 1 in vx and it is deleted from z, while uwy ends with 1^n, which is for sure.

3. The further discussion is omitted here. You can check page 285 of the textbook for help.

---

[1] not that variable

[2] That is because $|vwx| \leq n$