Question: use the pumping lemma with length to prove that the language $L = \{(aa)^n (bb)^n (aa)^{2n} \mid n \geq 0\}$ over the alphabet $\Sigma = \{a, b\}$ is not context free.

Solution

Introduction

Assume that L is context-free and it is generated by a CFG in CNF with p line productions. Then by pumping lemma with length, any word w in L with length(w) $\geq 2^p$ can be decomposed into 5 parts mentioned on the paper such that $uv^m xy^m z$ is also in L, for every $m > 0$.

Choose a typical word in L, such as $w = (aa)^{2^p} (bb)^{2^p} (aa)^{2*2^p}$, which has length(w) $\geq 2^p$.

Consider all the possible ways in which w can be divided into 5 substrings. Because length(vxy) $\leq 2^p$ there are only 2 cases.

Body of the proof

Case I

The substring vxy consists of a's only or b's only

If vxy consists of a's from the first group, then uvvxyy will contain more than $2^p$ instances of a in the first group, but still only $2^p$ instances of (bb) and still only $2*2^p$ instance of (aa) at the end.

If vxy consists of b's only, then uvvxyyz will contain more than $2^p$ instances of b, but still only $2^p$ groups of (aa) at the beginning and still only $2*2^p$ group of (aa) at the end.

If vxy consists of a's from the last group, then uvvxyyz will contain more than $2*2^p$ a's

at the end, but still only $2^p$ group of (aa) at the front, and still only $2^p$ group of (bb)

Case II.

The substring vxy consists of letters from two adjacent group of a's and b's.

If vxy consists of a's and b's from the first 2 groups of letters, then uvvxyyz would increase the size of one (or both) of these groups to a length greater than $2^p$, while there would be still only $2 * 2^p$ group of (aa) at the end.

If vxy consists of b's followed by a's from the last two groups of letters, then uvvxyyz would increase the size of one (or both) of these groups to more than $2^p$ (bb)'s or more than $2 * 2^p$ (aa)'s, while the first group of (aa) would still occur only $2^p$ times.

Conclusion

We have shown that every possible choice of vxy leads to a contradiction, so our assumption was wrong and L is non context-free.