Question:

Show that the language A in Exercise 2.9 is inherently ambiguous.

Answer:

----SETP1----

If a string w has two or more different derivations, it is ambiguous. The ambiguous grammar will generate some string ambiguously.

Some context free grammars can only be generated by the ambiguous grammars. Such languages are called ***inherently ambiguous***.

----SETP2----

Given context-free grammar that generates the language

$A = \left\{a^i b^j c^k \mid i = j \text{ or } j = k \text{ where } i, j, k \geq o\right\}$

Let G be the context-free grammar generating the language A and G is an ambiguous grammar. Now, show that A is inherently ambiguous.

Let p be the pumping length of the context-free grammar G given by the pumping lemma.

Let k be an integer such that $k = p! = p(p-1)(p-2)...1$.

Let s be a string such that $s = a^k b^k c^k$.

To show that A is inherently ambiguous, it is enough to show that s has two or more different leftmost derivations. It is required to know that those parse trees are equivalent to derivations and s has two parse trees.

----SETP3----

Let $s_1 = a^k b^p c^p$ and $s_2 = a^p b^p c^k$ are two strings in A with parse trees $\tau_1 \text{ and } \tau_2$ respectively.

Take τ_1 and remove all nodes that have only $a's$ below them. Since there are $k\ a's$, the sub tree will have only $p\ b's \text{ and } p\ c's$, in total having $2p$ leaves.

Because $2p > p$ it is known that the sub tree contains a path with repeated variable. Let R be the repeated variable. Divide s into $uvxyz$ using R and R can be chosen such that $|vxy| \leq p$. Each of the strings v and y can contain only one of the symbols, if not uv^2xy^2z will not belong to A.

Also y contain no $a's$, because R was on a path to $a, b \text{ or } c$. Hence v must be $b's$ and y must be $c's$ and they must have an equal length l. Consider $d = k/l$ (which must be an integer since $l \leq p \text{ and } k = p!$) the string $s = vu^d xy^d$ has a parse tree where most $b's$ have a parent node in common with $c's$ but not with $a's$.

----SETP4----

Similarly, consider τ_2, the parse tree is obtained for s $b's$ having a parent node in common with $a's$ but not with $c's$.

So s has two distinct parse trees. Hence s has two or more different leftmost derivations and G can be generated by the ambiguous grammars only.

Therefore, the language A is inherently ambiguous.