Problem 2.46. Consider the following CFG G:

$$S \to SS \mid T$$
$$T \to aTb \mid ab$$

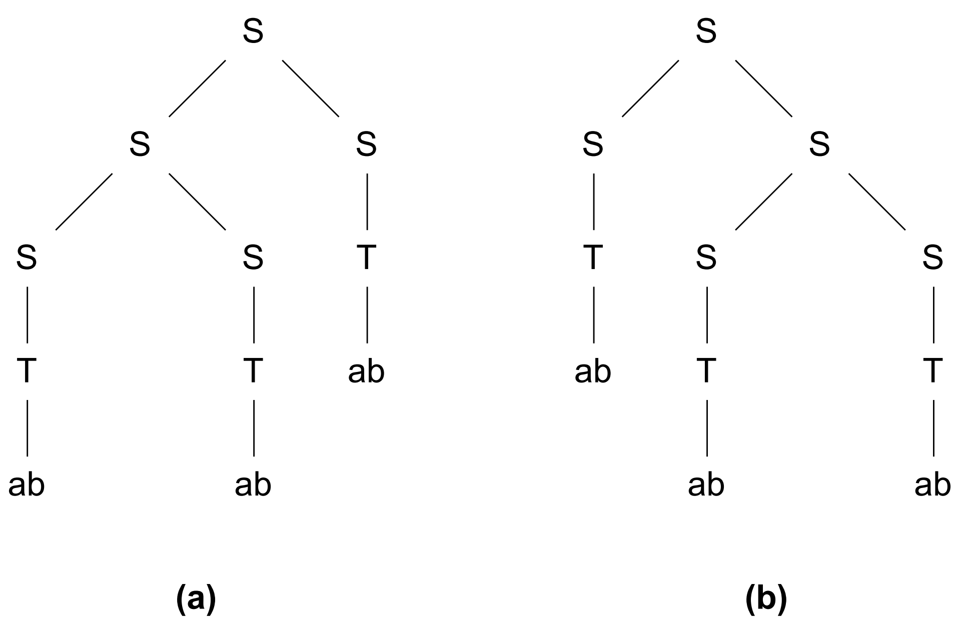
Describe $L(G)$ and show that G is ambiguous. Give an unambiguous grammar H where $L(H) = L(G)$ and sketch a proof that H is unambiguous.

Part a. Describe $L(G)$.

Let $A = \{a^n b^n \mid n \geq 1\}$, then $L(G) = \{w \mid w \in A^+\}$.

Part b. Show that G is ambiguous.

Then CFG G is ambiguous, because the string $ababab$ is a member of $L(G)$ and it has more than 1 parse trees in G.



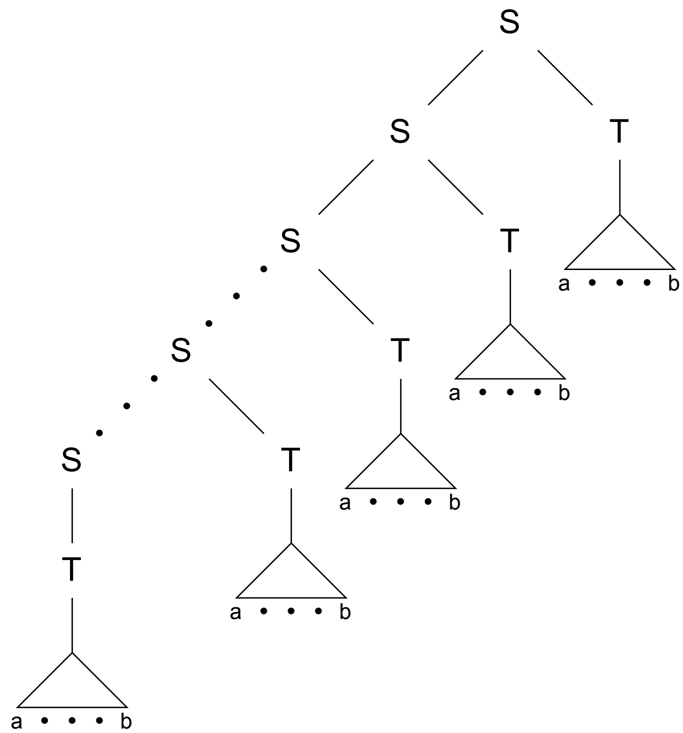
Two parse trees of the string $ababab$.

Part c. Give an unambiguous grammar H where $L(H) = L(G)$.

$$S \to ST \mid T$$
$$T \to aTb \mid ab$$

Part d. Sketch a proof that H is unambiguous.

Proof. All strings that contain two or more segments of $a \cdots b$, have the same parse tree structure in H as shown in the following diagram. As the number of $a \cdots b$ segments grows, the corresponding parse tree grows only to the left. If a string has n segments of $a \cdots b$, where $n \geq 2$, then the $S \rightarrow ST$ rule is applied $n-1$ times to generate the n number of $T's$ required to generate the string. As this is the only way to generate strings containing two or segments of $a \cdots b$, therefore H is unambiguous.



Structure of parse trees in H.

$\square$