

Solution to Problem 13

Task. Show that the language L of all the words that have 10 times more a 's than t and 5 times more h 's than t 's is not context-free.

Solution. Let us prove that this language is not context-free.

The proof will be by contradiction. Let us assume that this language is context-free. Then, by pumping lemma, there exists an integer p such that every word w from the language L whose length is at least p can be represented as $w = uvxyz$, where:

- $\text{len}(vxy) \leq p$;
- $\text{len}(vy) > 0$, and
- for all natural numbers i , the word $uv^i xy^i z$ also belongs to the language L .

Let us take the word

$$w = a^{10p} h^{5p} t^p = a \dots a h \dots h t \dots t,$$

where first a is repeated $10p$ times, then h is repeated $5p$ times, and finally t is repeated p times. The length of this word – i.e., the number of symbols in this word – is equal to $10p + 5p + p = 16p$. Clearly, $16p \geq p$, so, according to the Pumping Lemma, this word can be described as $uvxyz$ with the above properties.

Where can the central part vxy of this word be? We know that the length $\text{len}(vxy)$ of this part cannot exceed p . Thus, it cannot contain three different types of symbols: a 's, h 's, and t 's – since then it would have to include all $5p$ symbols h plus additional a and t symbols, so its length would have been larger than $5p$. So, there are only 5 cases remaining for the location of the part vxy :

1. it can be in the a 's;
2. it can be in a 's and h 's;
3. it can be in h 's;
4. it can be in h 's and t 's;
5. it can be in t 's.

Let us consider these cases one by one.

Case 1. If vxy is in the a 's, this means that the parts v and y contain only a 's. Thus, when we pump, i.e., when we go from the original word $uvxyz$ to the word $uv^2xy^2z = uvvxyyz$, we add a 's – but we do not add any h 's or t 's. In the original word $w = a^{10p}h^{5p}t^2$, there was a balance between letters of the three types. When we add more a 's, the balance is disrupted: the ratio between number of a 's and number of t 's is now different from 10. Since the language L only contains the words for which the above proportions have to be satisfied, the word $uvvxyyz$ cannot belong to the language L .

Case 2. If vxy is in a 's and in h 's, this means that the parts v and y contain only a 's and h 's. Thus, when we pump, i.e., when we go from the original word $uvxyz$ to the word $uv^2xy^2z = uvvxyyz$, we add a 's and/or h 's – but we do not add any t 's. In the original word $w = a^{10p}h^{5p}t^p$, there was the desired balance between numbers of letters of all three types. When we add more a 's and/or h 's, the balance is disrupted, so the word $uvvxyyz$ cannot belong to the language L .

Similarly, we can see that in the other 3 cases, we also get a contradiction. This means that the original assumption – that the language L is context-free – is wrong. Thus, the language L is not context-free.