Solution to Problem 13

Task. A perfect arrangement would be to have 8 hours of study (S), 8 hours of rest (R), and 8 hours of sleep (P). Show that the language L of all the words that have equal number of Ss, Rs, and Ps 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:

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

Let us take the word

$$w = S^pR^pP^p = S\ldots SR\ldots RP\ldots P,$$

where S is repeated p times, R is repeated p times, and P is repeated p times. The length of this word – i.e., the number of symbols in this word – is equal to $p + p + p = 3p$. Clearly, $3p \ge p$, 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: S's, R's, and P's – since then it would have to include all p symbols R plus additional S and P symbols, so its length would have been larger than p. So, there are only 5 cases remaining for the location of the part vxy:

1. it can be in the S's;
2. it can be in S's and R's;
3. it can be in R's;
4. it can be in R's and P's;
5. it can be in P's.

Let us consider these cases one by one.

Case 1. If vxy is in the S's, this means that the parts v and y contain only S's. Thus, when we pump, i.e., when we go from the original word $uvxyz$ to the word $uv^2xy^2z = uvvxyyz$, we add S's – but we do not add any R's or P's. In the original word $w = S^pR^pP^p$, there was a balance between letters of the three types. When we add more S's, the balance is disrupted: the ratio between number of S's and number of R's is now different from 1. 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 S's and in R's, this means that the parts v and y contain only S's and R's. Thus, when we pump, i.e., when we go from the original word $uvxyz$ to the word $uv^2xy^2z = uvvxyyz$, we add S's and/or R's – but we do not add any P's. In the original word $w = S^pR^pP^p$, there was the desired balance between numbers of letters of all three types. When we add more S's and/or R'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.