Solution to Homework 5

Task: A student who only has As and Bs has a GPA higher than 3.5 if and only if this student has more As than Bs. Prove that the language L of all A-B sequences like AAB that correspond to students with GPA higher than 3.5 is not regular. .

Solution. We will prove this result by contradiction. Let us assume that the language L is regular, and let us show that this assumption leads to a contradiction.

Since this language is regular, according to the Pumping Lemma, there exists an integer p such that every word from L whose length $\text{len}(w)$ is at least p can be represented as a concatenation $w = xyz$, where:

- y is non-empty;
- the length $\text{len}(xy)$ does not exceed p, and
- for every natural number i, the word $xy^iz \stackrel{\text{def}}{=} xy\ldots yz$, in which y is repeated i times, also belongs to the language L.

Let us take the word

$$w = \text{B}^p\text{A}^{p+1} = \text{B}\ldots\text{BA}\ldots\text{A},$$

in which first the grade B is repeated p times and then the grade A is repeated $p+1$ times. The length of this word is $p+p+1 = 2p+1 > p$. So, by pumping lemma, this word can be represented as $w = xyz$ with $\text{len}(xy) \le p$. The word $w = xyz$ starts with xy, and the length of xy is smaller than or equal to p. Thus, xy is among the first p symbols of the word w – and these symbols are all Bs. So, the word y only has Bs.

In the original word $w = xyz$, we had B repeated p times and A repeated $p+1$ times. When we go from the word $w = xyz$ to the word $xyyz$, we add Bs, and we do not add any As. Thus, we still have A repeated $p+1$ times, but the number of times B is repeated is now larger than p, i.e., larger than or equal to $p+1$. In any word from the language L, we should have more As than Bs. In the word $xyyz$, this condition is not satisfied. Thus, the word $xyyz$ cannot be in the language L.

On the other hand, by Pumping Lemma, the word $xyyz$ must be in the language L. So, we proved two opposite statements:

- that this word *is not* in L and

- that this word *is* in L.

This is a contradiction.

The only assumption that led to this contradiction is that L is a regular language. Thus, this assumption is false, so the language L is not regular.