

Solution to Homework 4

Task: Prove that the language L of all the sequences that have exactly the same number of ups and down 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 the language L 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 $xy^iz \stackrel{\text{def}}{=} xy \dots yz$, in which y is repeated i times, also belongs to the language L .

Let us take the word

$$w = u^p d^p = u \dots u d \dots d,$$

in which first the symbol u is repeated p times and then the symbol d is repeated p times. The length of this word is $p + p = 2p > p$. So, by pumping lemma, this word can be represented as $w = xyz$ with $\text{len}(xy) \leq p$. 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 u 's. So, the word y only has u 's.

In the original word $w = xyz$, we had u repeated p times and d repeated p times. When we go from the word $w = xyz$ to the word $xyyz$, we add u 's, and we do not add any d 's. Thus, we still have d repeated p times, but the number of times u is repeated is now larger than p . In any word from the language L , we should have exactly as many u 's as d 's. 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.