Solution to Problem 15

Task. Prove that the language consisting of all the words in the alphabet $\{A, B, C\}$ that have twice as many B's as C's and three times as many A's than C'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^ixy^iz also belongs to the language L.

Let us take the word

$$w = A^{3p}B^{2p}C^p = A \ldots AB \ldots BC \ldots C,$$

where A is repeated $3p$ times, B is repeated $2p$ times, and C is repeated p times. The length of this word – i.e., the number of symbols in this word – is equal to $3p + 2p + p = 6p$. Clearly, $6p \geq p$, 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 all three types of symbols: A's, B's, and C's – since then it would have to include all $2p$ letters B plus additional A and C 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 A's;
2. it can be in A's and B's;
3. it can be in B's;
4. it can be in B's and C's; and
5. it can be in C's.

Let us consider these cases one by one.

Case 1. If vxy is in 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 B's or C's. In the original word $w = A^{3p}B^{2p}C^p$, there were exactly three times as many A's as C's. When we add more A's, the balance is disrupted. Since the language L only contains the words which have three times as many A's as C's, the word $uvvxyyz$ cannot belong to the language L.

Case 2. If vxy is in A's and in B's, this means that the parts v and y contain only A's and B'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 B's – but we do not add any C's. In the original word $w = A^{3p}B^{2p}C^p$, there were exactly three times as many A's as C's, and exactly two times as many B's as C's. When we add more A's and B's, the balance is disrupted. Since the language L only contains the words which have three times as many A's as C's and two times as many B's as C's, the word $uvvxyyz$ cannot belong to the language L.

Case 3. If vxy is in B's, this means that the parts v and y contain only B's. Thus, when we pump, i.e., when we go from the original word $uvxyz$ to the word $uv^2xy^2z = uvvxyyz$, we add B's – but we do not add any A's or C's. In the original word $w = A^{3p}B^{2p}C^p$, there were exactly twice as many B's as C's. When we add more B's, the balance is disrupted. Since the language L only contains the words which have exactly twice as many B's as C's, the word $uvvxyyz$ cannot belong to the language L.

Case 4. If vxy is in B's and C's, this means that the parts v and y contain only B's and C's. Thus, when we pump, i.e., when we go from the original word $uvxyz$ to the word $uv^2xy^2z = uvvxyyz$, we add B's and C's – but we do not add any A's. In the original word $w = A^{3p}B^{2p}C^p$, there were exactly three times as many A's as C's and exactly 1.5 times more A's than B's. When we add more B's and/or C's, the balance is disrupted. Since the language L only contains the words which have exactly three times as many A's as C's and exactly 1.5 times more A's than B's, the word $uvvxyyz$ cannot belong to the language L.

Case 5. If vxy is in C's, this means that the parts v and y contain only C's. Thus, when we pump, i.e., when we go from the original word $uvxyz$ to the word $uv^2xy^2z = uvvxyyz$, we add C's – but we do not add any A's or B's. In the original word $w = A^{3p}B^{2p}C^p$, there were exactly three times as many A's as C's. When we add more C's, the balance is disrupted. Since the language L only contains the words which have exactly three times more A's than C's, the word $uvvxyyz$ cannot belong to the language L.

So, in all five cases, we 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.