Proving a Language is Not Regular

We’ve seen in class one method to prove that a language is not regular, by proving that it does not satisfy the pumping lemma. This method works often but not always.

A second method (which also doesn’t always work), is by using closure properties of regular languages, and relying on the fact that we already know that some other language is not regular. The proof would go along the following lines: Assume towards contradiction that \( L \) is regular. Apply operations that regular languages are closed under (e.g., union, concatenation, star, intersection, or complement) on \( L \) and other regular languages, to reach a language that is not regular. Contradiction. Conclude that \( L \) is not regular. Here are two examples.

**Claim 1.** \( L_1 = \{ w \in \{a, b\}^*: w \text{ has the same number of } a\text{s and } b\text{s} \} \) is not regular.

**Proof.** Note that \( L_1 \cap a^* b^* = \{a^i b^j : i \geq 0 \} \). Now assume towards contradiction that \( L_1 \) is regular. Since \( a^* b^* \) is regular, and regular languages are closed under intersection, then the intersection is also regular. But we know that \( \{a^i b^j : i \geq 0 \} \) is not regular – contradiction.

In the above example, \( L_1 \) could also be proved non-regular using the pumping lemma. This is not the case for the next example (the adversary would always win / pumping lemma holds for \( L_2 \)). However, using closure properties, we can prove the following example is not regular (try to do this yourself before reading the solution!)

**Claim 2.** \( L_2 = \{a^i b^j : i \neq j \} \) is not regular.

**Proof:** Note that \( \mathcal{L} \cap a^* b^* = \{0 \leq i : \exists q \in Q \} \). Now assume towards contradiction that \( \mathcal{L} \) is also regular. But then, since regular languages are closed under intersection (and \( a^* b^* \) is regular), we get that \( \mathcal{L} \) is regular. But then, since regular languages are closed under complement, \( \mathcal{L} \) is also regular, because \( a^* b^* \) is regular, and thus must be regular. Contradiction. Conclude that \( \mathcal{L} \) is not regular.

A third method for proving a language is not regular, is using the Myhill-Nerode theorem, which is not part of the class material (interested students can read about it in the handout on the class webpage). This has the advantage that it is a characterization (an if and only if condition), and thus can in principle be used for every non-regular language. However, depending on the language, it may be hard to apply this theorem (namely to find the equivalence classes of the language).