#1 Linguistic Manipulation

As an exercise in reading, exploring, and understanding definitions, let's analyze a definition we will see next week when we study regular languages, the *star operator*.

Definition 1 (star operator)**.** *Let Σ be an alphabet. Then Σ^* is the set of all possible strings formed from letters drawn from Σ:*

$$\Sigma^* = \{\, x_1 x_2 \cdots x_k \mid k \geq 0, x_i \in \Sigma \ \ \forall i \in [1, k] \,\}.$$

1. Let $\Sigma = \{\, 0, 1 \,\}$. Give three examples of strings found in Σ^*.
2. Let ϵ (lowercase Greek epsilon) be the empty string, i.e., the string with no characters. Is $\epsilon \in \Sigma^*$? Why or why not?
3. What is the size or *cardinality* of Σ^*. Justify your answer in a sentence or two.
4. Does your answer to the previous part hold for *any* choice of Σ? If so, justify your answer in a sentence. If not, provide a counterexample Σ where your answer to the previous part does not hold.

#2 Distinguishability

In this problem, we'll consider a property related to strings, *distinguishability.* While this concept will become useful to us when we discuss irregularity in a few weeks, its formal definition serves as a useful exercise in breaking part and analyzing mathematics.

Definition 2 (Distinguishability)**.** *Let L be a set of strings and x and y be strings (not necessarily in L).*

- *We say that x and y are* distinguishable by L *if there exists a string z (also not necessarily from L itself) such that exactly one of xz and yz are in L.*
- *We say that x and y are* indistinguishable by L *if they are not distinguishable.*

Furthermore, if x and y are strings indistinguishable by L, we write $x \equiv_L y$.

(Note: we use L as the variable for our set of strings because we will eventually call this set our language once we introduce finite automata.)

1. Let $\Sigma = \{\, a, b \,\}$ and $L_1 = \{\, abba, baba, aabb, bbbb \,\}$. Give an example pair of strings that is distinguishable by L_1 and a pair of strings that are indistinguishable by L_1. (Hint: again, note that these strings need not necessarily be in L_1.)

2. Now let L_2 be the infinite set of strings where the first and last characters are different:

$$L_2 = \{\, w \mid w \in \Sigma^*, w = x_1 \cdots x_k, x_1 \neq x_k \,\}$$

   Construct two examples of pairs of strings that are distinguishable and three pairs of strings that are indistinguishable by L_2.

3. Using the intuition you built up from the previous part, try to describe what it means to be indistinguishable at a high-level in a sentence or two.

4. **(Optional)** Recall the definition of equivalence relation from the reading. Argue why the indistinguishability relation ($\equiv_L$) is an equivalence, appealing to the formal definition of an equivalence relation. If you'd like, also formally prove this fact.