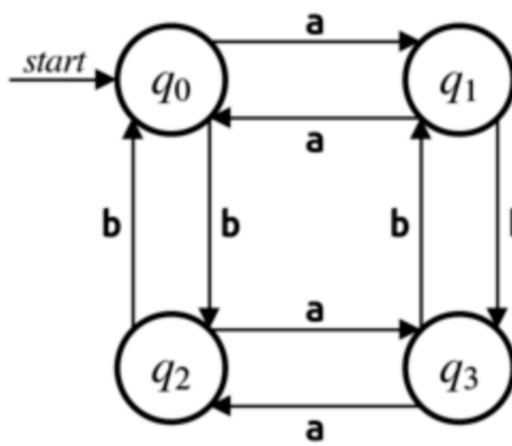
1. Counting Characters

Let $\Sigma = \{a, b\}$. If $w \in \Sigma^*$, we denote by $A(w)$ the number of **a**'s in w and $B(w)$ the numbers of **b**'s in w. For example, we have $A(\mathbf{abbb}) = 1$ and $B(\mathbf{abbb}) = 3$.

Consider the following DFA D:



We can change the language of D by changing which of its states are accepting, leaving the start state and transitions unmodified.

Select the correct option from each of the lists below and fill in the appropriate blanks. No justification is necessary.

a. Let $L = \Sigma^*$

- ☐ We can make $\mathcal{L}(D) = L$ by making states ____________ accepting.
- ☐ No matter which state of D we make accepting, we will never have $\mathcal{L}(D) = L$.

b. Let $L = \{\mathbf{abab}\}$

- ☐ We can make $\mathcal{L}(D) = L$ by making states ____________ accepting.
- ☐ No matter which state of D we make accepting, we will never have $\mathcal{L}(D) = L$.

c. Let $L = \{w \in \{\mathbf{a,b}\}^* | A(w) + B(w) \text{ is even}\}$.

- ☐ We can make $\mathcal{L}(D) = L$ by making states ____________ accepting.
- ☐ No matter which state of D we make accepting, we will never have $\mathcal{L}(D) = L$.

d. Let $L = \{w \in \{\mathbf{a,b}\}^* | (A(w) \text{ is even}) \to (B(w) \text{ is even})\}$.

- ☐ We can make $\mathcal{L}(D) = L$ by making states ____________ accepting.
- ☐ No matter which state of D we make accepting, we will never have $\mathcal{L}(D) = L$.

Here is a theorem about string over $\Sigma = \{\mathbf{a},\mathbf{b}\}$:

> ***Theorem:*** For all $w \in \Sigma^*$, there exist $x, y \in \Sigma^*$ such that $w = xy$ and $A(x) = B(y)$.

Below is a partial proof of the theorem:

Theorem: For all $w \in \Sigma^*$, there exist $x, y \in \Sigma^*$ such that $w = xy$ and $A(x) = B(y)$.
Proof: Let $P(n)$ be the predicate "for all $w \in \Sigma^*$ of length n, the exist $x, y \in \Sigma^*$ where $w = xy$ and $A(x) = B(y)$." We will prove by induction that $P(n)$ holds for all $n \in \mathbf{N}$, from which the theorem follows.

As out base cases, we will prove $P(0)$ and $P(1)$.
[*proof of base case goes here*].

For our inductive step, pick some $k \in \mathbf{N}$ and assume $P(k)$ and $P(k+1)$ are true, meaning that for any string $w \in \Sigma^*$ of length k or $k+1$, there exist strings x and y where $w = xy$ and $A(x) = B(y)$. We will prove $P(k+2)$, that for any string $w \in \Sigma^*$ of length $k+2$, there exist strings x and y where $w = xy$ and $A(x) = B(y)$.
[*proof of inductive step goes here*].

Thus $P(k+2)$ holds, completing the induction.

Your task is to fill in the two missing sections of this proof. We'll begin by having you explain the logic of the base cases $P(0)$ and $P(1)$.

e. Fill in the blanks below. No justification is necessary.

If $w = \varepsilon$, on choice of x and y satisfying the requirements of the theorem is $x =$ __________

and $y =$ __________.

f. Fill in the blanks below. No justification is necessary.

If $w = \mathbf{a}$, one choice of x and y satisfying the requirements of the theorem is $x =$ __________

and $y =$ __________.

If $w = \mathbf{b}$, one choice of x and y satisfying the requirements of the theorem is $x =$ __________

and $y =$ __________.

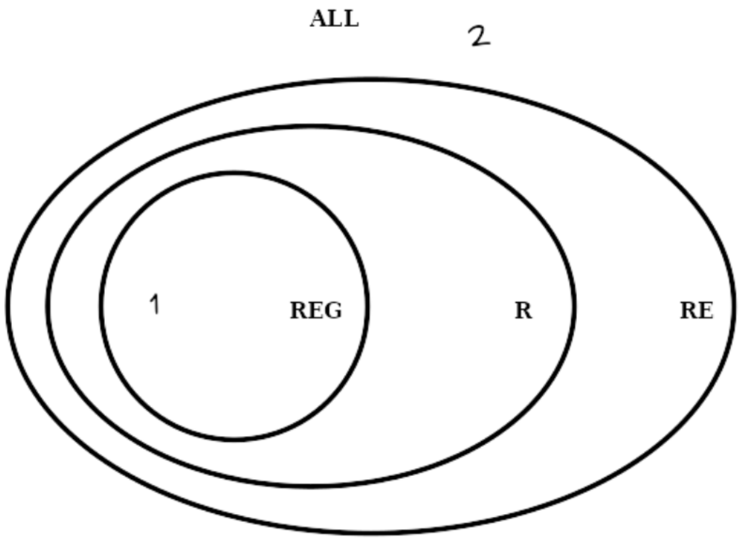
g. Write the proof of the inductive step of the theorem. Some hints:

- Consider the following three cases: the string starts with **b**, the string ends with **a**, or neither of these are true.
- If w is a string that starts with **b**, then there is a string z where $w = \mathbf{b}z$. Similarly, if w is a string that ends in **a**, then there is a string z where $w = z\mathbf{a}$.
- Feel free to use the fact that $A(\mathbf{a}x) = 1 + A(x)$ and $A(\mathbf{b}x) = A(x)$, along with analogous results for B.

Proceed systematically in your proof, making sure not to skip any steps.

2. The Lava Diagram

Below is a Venn diagram showing the overlap of different classes of languages we've studied. We have also provided you a list of numbered languages. For each of those languages, draw where in the Venn diagram that language belongs. As an example, we've indicated where Language 1 and Language 2 should go. No justification is necessary.



a. Σ^*

b. L_D

c. $\{w \in \{\mathbf{a},\mathbf{b}\}^* \mid A(w) + B(w) = 137\}$ ($A(w)$ and $B(w)$ are defined in Problem One)

d. $\{w \in \{\mathbf{a},\mathbf{b}\}^* \mid A(w) - B(w) = 137\}$

e. $\{w \in \{\mathbf{a},\mathbf{b}\}^* \mid A(w) \times B(w) = 137\}$

f. $\{w \in \{\mathbf{a},\mathbf{b}\}^* \mid A(w) \div B(w) = 137\}$

g. The language of this CFG: $S \to \mathbf{a}S \mid S\mathbf{b} \mid \varepsilon$

h. The language of this CFG: $S \to \mathbf{a}S\mathbf{b} \mid \varepsilon$

i. $\{\langle M, w\rangle \mid M \text{ is a TM, } w \text{ is a string, and } M \text{ accepts } w\}$

j. $\{\langle M, w\rangle \mid M \text{ is a TM, } w \text{ is a string, and } M \text{ rejects } w\}$

k. The language of the following TM, which we'll call Z:

```
Start:
Goto Start
```

l. $\{\langle Z\rangle\}$

3. Unary Monoids

In this problem, we'll let $\Sigma = \{\mathbf{a}\}$. As a refresher, a ***monoid*** over Σ is a language $M \subseteq \Sigma^*$ where

$$\varepsilon \in M \quad \text{and} \quad MM \subseteq M$$

Select the correct responses from the lists below and fill in the appropriate blanks. No justification is necessary.

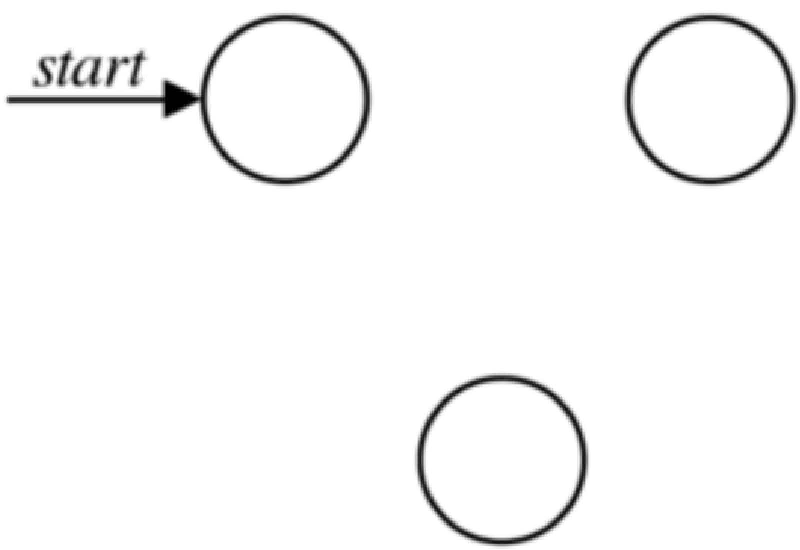
a. Let $M = \{\mathbf{a}^n \mid n \in \mathbf{N} \land n \notin \{1, 2, 5\}\}$. Note that $\varepsilon \in M$. Is $MM \subseteq M$?

- ☐ Yes.
- ☐ No. Pick $x =$ __________ and $y =$ __________. Then $x \in M$ and $y \in M$ but $xy \notin M$.

b. Let $M = \{\mathbf{a}^n \mid n \in \mathbf{N} \land n \notin \{1, 3, 6\}\}$. Note that $\varepsilon \in M$. Is $MM \subseteq M$?

- ☐ Yes.
- ☐ No. Pick $x =$ __________ and $y =$ __________. Then $x \in M$ and $y \in M$ but $xy \notin M$.

c. Below are three states of a DFA, one of which is marked as a start state. Add transitions to the states below, and mark some number of the states as accepting states, so that the result is a DFA for the monoid $\{\mathbf{a}^2, \mathbf{a}^3\}^*$ **(Note the Kleene star)**. Do not introduce any new states. No justification is required.



d. Write a regular expression for the monoid $\{\mathbf{a}^n \mid n \in \mathbf{N} \land n \notin \{1,3\}\}$. No justification is required.

As you saw on Problem Set 8, if M is a monoid over Σ, a ***codeword*** of M is a string $w \in \Sigma^*$ where the following is true:

$$w \in M \quad \land \quad w \neq \varepsilon \quad \land \quad \forall x \in M. \forall y \in M. (w = xy \to x = \varepsilon \lor y = \varepsilon)$$

Below is a list of monoids over $\Sigma = \{\mathbf{a}\}$. For each monoid, we have identifies how many codewords that monoid has. Tell us what the codewords are. No justification is required.

e. The monoid $\{\mathbf{a}^n \mid n \in \mathbf{N} \land n \notin \{1,3\}\}$ has two codewords. What are they?

__________ and __________

f. The monoid $\{\mathbf{a}^n \mid n \in \mathbf{N} \land n \notin \{1,2,4\}\}$ has three codewords. What are they?

__________ __________ and __________

We'd like you to prove a final result about monoids.

g. Let A and B be monoids over $\Sigma = \{\mathbf{a}\}$. Prove that AB is also a monoid over Σ.

In the course of your proof, we expect you to reference the formal definition of language concatenation, which is reprinted below:

$$L_1 L_2 = \{w \mid \exists x_1 \in L_1. \exists x_2 \in L_2. w = x_1 x_2\}$$

Additionally, you should make us of the following two facts, which you may use without proof:

$$(I) \quad \forall w \in \Sigma^*. \exists n \in \mathbf{N}. w = \mathbf{a}^n$$
$$(II) \quad \forall m \in \mathbf{N}. \forall n \in \mathbf{N}. \mathbf{a}^m \mathbf{a}^n = \mathbf{a}^{m+n}$$

Proceed systematically in your proof, making sure not to skip any steps. We are expecting a rigorous proof that calls back to the relevant definitions and facts. (Write proof on next page)