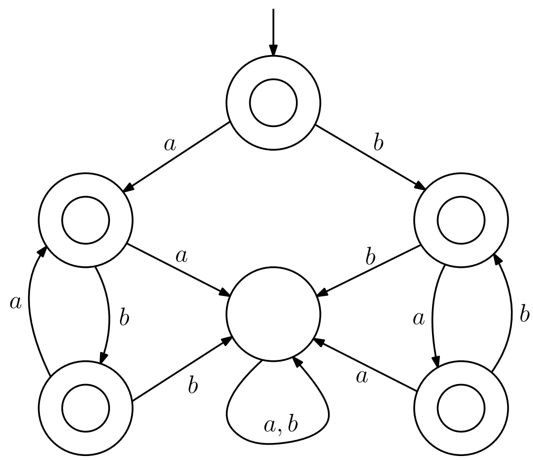
COMP 3803 — Solutions Assignment 1

Question 1: Write your name and student number.

Solution: Erling Haaland, 9

Question 2: What is the language of the following DFA? The alphabet is $\{a, b\}$. Justify your answer.



Solution: Denote the non-accept state by q. The following three claims follow from the state diagram.

- If we reach the state q, then we will stay there forever.
- If we are in any accept state: If we read aa or bb, then we will reach state q.
- The only way to reach state q is by reading aa or bb.

This implies that a string is rejected if and only if it contains aa or bb. And this implies that the DFA accepts the complement of these strings. Equivalently, the DFA accepts all strings in which a's and b's alternate: It accepts

- ε,
- a,
- b,
- $(ab)^k$ for any $k \geq 1$,
- $(ab)^k a$ for any $k \geq 1$,
- $(ba)^k$ for any $k \geq 1$,

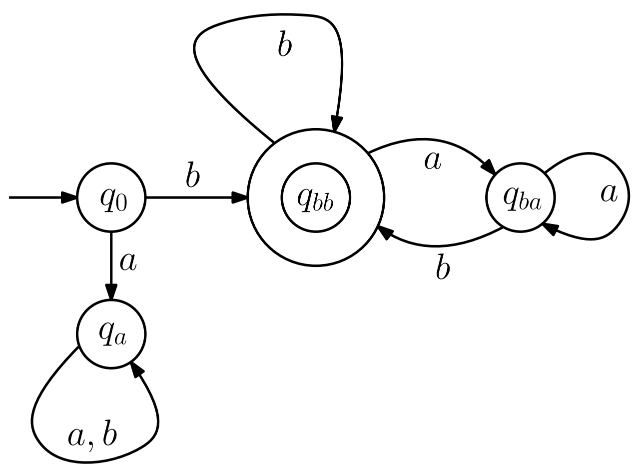
- $(ba)^k b$ for any $k \geq 1$.

Question 3: For each of the following two languages, construct a DFA that accepts the language. In both cases, the alphabet is $\{a, b\}$. For each DFA, justify correctness.
(3.1) The language consisting of all strings $w \in \{a, b\}^*$ that start and end with b.
Note that the string b (having length one) is included in this language.
(3.2) The language consisting of all strings $w \in \{a, b\}^*$ in which the number of a's is even and the number of b's is a multiple of three.
Note that the empty string ε is included in this language.

Solution: For the first part, we will use the following states:

- q_0: We have not read any symbol yet.
- q_a: We have read the first symbol and it is an a.
- q_{bb}: Both the first and last symbols we have read are b; this includes the case when we have read only one symbol (which is b).
- q_{ba}: We have read at least two symbols, the first is b, the last is a.

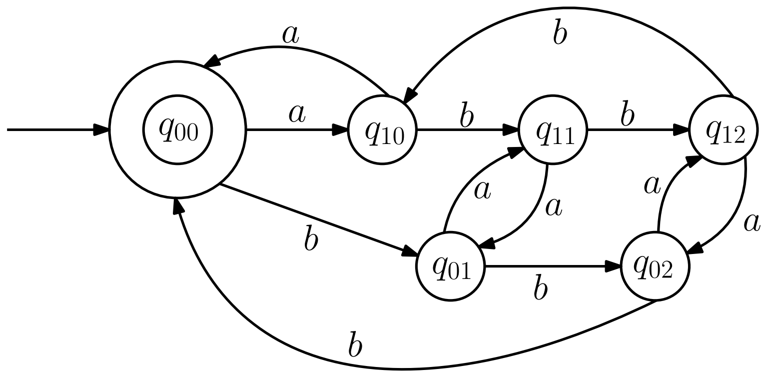
The state q_0 is the start state, the state q_{bb} is the only accept state. The state diagram is given below.



For the second part, we are going to keep track of the number of a's read modulo 2 and the number of b's read modulo 3. Thus, we will use $2 \cdot 3 = 6$ states:

- For $i = 0, 1$ and $j = 0, 1, 2$, we are in state q_{ij} if the number of a's read modulo 2 is equal to i, and the number of b's read modulo 3 is equal to j.

The state q_{00} is the start state and it is also the only accept state. The state diagram is given below.



Question 4: Construct an NFA with four states whose language is the set of all strings $w \in \{a, b\}^*$ such that

- $w = a^k$, for some integer $k \geq 0$, or
- $w = (ab)^k$, for some integer $k \geq 0$.

As always, justify correctness.

Solution: We define the languages

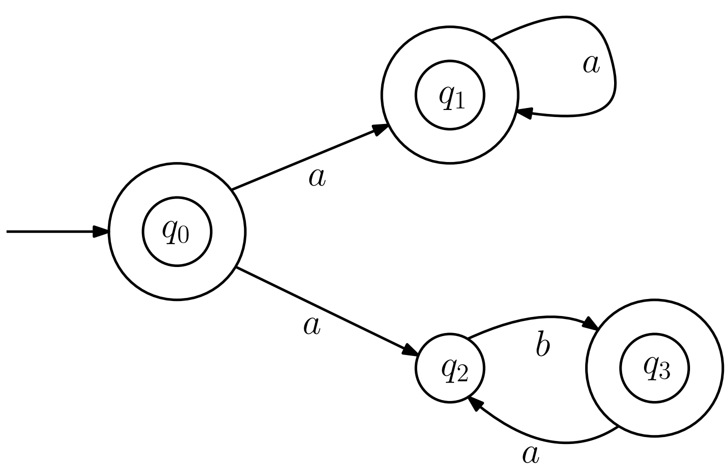
$$A = \{a^k : k \geq 0\}$$

and

$$B = \{(ab)^k : k \geq 0\}.$$

The question asks for an NFA that accepts the union of A and B. The first idea is to construct an NFA that accepts A (this can be done using one state) and an NFA that accepts B (this can be done using three states). Then we apply the union construction, giving an NFA with five states (because we add a new start state and give it ε-transitions to the start states of the two NFAs).

We obtain an NFA with four states in the following way: From the start state, if the first symbol is a, then we "guess" whether we are going to check if the string is in A or in B. Consider the following state diagram.



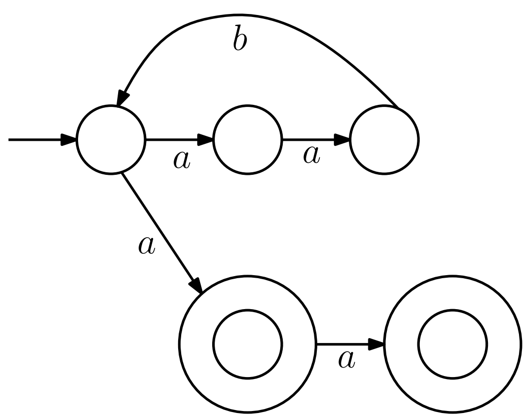
- The empty string is accepted; it is in both A and B.
- A non-empty string that starts with b cannot be read; thus it is rejected.
- If the first symbol is a and we move to state q_1, then the string can only be accepted if this first a is followed by zero or more a's. In this way, we accept all strings of the form a^k for some $k \geq 1$.
- If the first symbol is a and we move to state q_2, then the string can only be accepted if this first a is followed by a string of the form $(ba)^{k-1}b$ for some $k \geq 1$. In this way, we accept all strings of the form $a(ba)^{k-1}b = (ab)^k$ for some $k \geq 1$.

Question 5: Construct an NFA whose language is the set of all strings $w \in \{a, b\}^*$ such that

- $w = (aab)^k a$, for some integer $k \geq 0$, or
- $w = (aab)^k aa$, for some integer $k \geq 0$.

As always, justify correctness.

Solution: Consider the following state diagram.



First note that both strings $a = (aab)^0 a$ and $aa = (aab)^0 aa$ are accepted.

In the top part, we verify that the input string starts with $(aab)^k$ for some $k \geq 1$. If this is the case, then the bottom part verifies that the string ends with a or aa.

The only way to reach the leftmost accept state is if the input string is of the form $(aab)^k a$ for some $k \geq 0$.

The only way to reach the rightmost accept state is if the input string is of the form $(aab)^k aa$ for some $k \geq 0$.

Question 6: Professor Justin Bieber claims to have proved the following result:

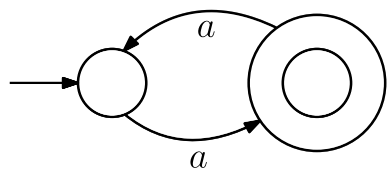
Bieber's Theorem: Let M be an arbitrary NFA with alphabet $\{a, b\}$ that has exactly one accept state q_f, and let A be the language accepted by M. Let B be the concatenation of A and $\{b\}^*$, i.e.,

$$B = \{vw : v \in A, w \in \{b\}^*\}.$$

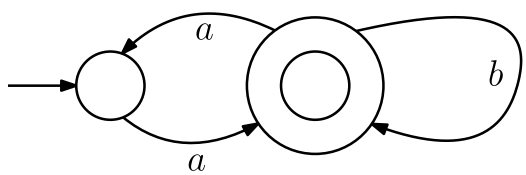
Let M' be the NFA obtained by making a copy of M and adding a b-transition from q_f to q_f. Then this new NFA M' accepts the language B.

Is Bieber's Theorem correct? As always, justify your answer.

Solution: Professor Bieber should fail COMP 3803. Let M be the following NFA, which accepts the language $A = \{a^k : \ k \text{ is odd}\}$.



If we apply the Bieber construction, we obtain the following NFA, which accepts the string $abaa$, which is not in B.



Question 7: Let A be an arbitrary language over the alphabet $\{a, b\}$. We define the language

$$A' = \{v \in \{a, b\}^* : \text{ there exists a string } w \in A \text{ such that } v \text{ and } w \text{ have the same length and differ in at most one position}\}.$$

For example, if $abba \in A$, then this string gives rise to the five strings $abba$, $bbba$, $aaba$, $abaa$, and $abbb$ in A'.

Prove that if A is regular, then A' is also regular.

Hint: Take a DFA that accepts A. Make multiple copies of its state diagram, and connect the copies with the original state diagram. It is possible to do this without using ε-transitions.

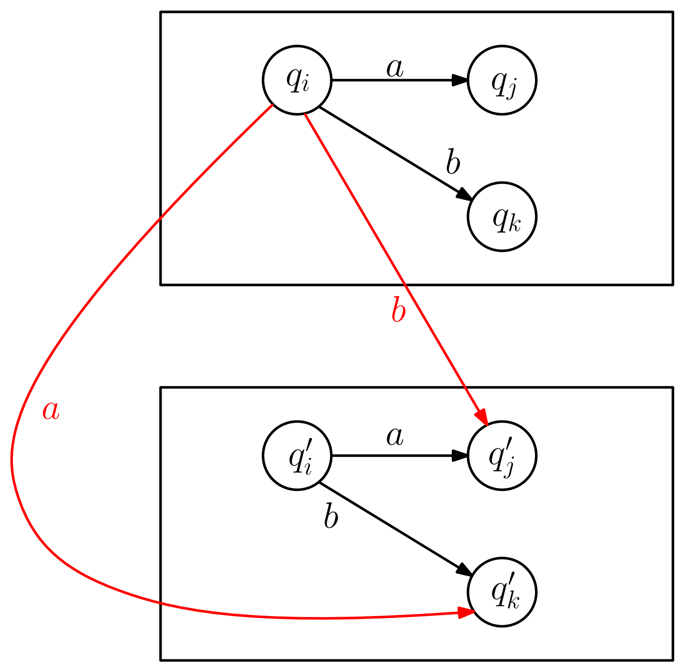
Solution: Let M be a DFA that accepts the language A, and let $Q = \{q_1, q_2, \ldots, q_s\}$ be its set of states.

We are going to construct an NFA that accepts the language A'. This will imply that A' is regular.

The NFA consists of the following:

- It contains the DFA M. The start state of M is the start state of the NFA.
- It contains a copy M' of M. We denote the set of states of M' by $Q' = \{q'_1, q'_2, \ldots, q'_s\}$. The start state of M' is not a start state in the NFA.
- For every state q_i in Q, do the following:
  - Let j be the index such that, if M is in state q_i and reads the symbol a, it switches to state q_j. (Note that j may be equal to i.)
  - Let k be the index such that, if M is in state q_i and reads the symbol b, it switches to state q_k. (Note that k may be equal to i.)
  - We add to the NFA the transition "if in state q_i and read b, switch to state q'_j.
  - We add to the NFA the transition "if in state q_i and read a, switch to state q'_k.

The figure below illustrates this.

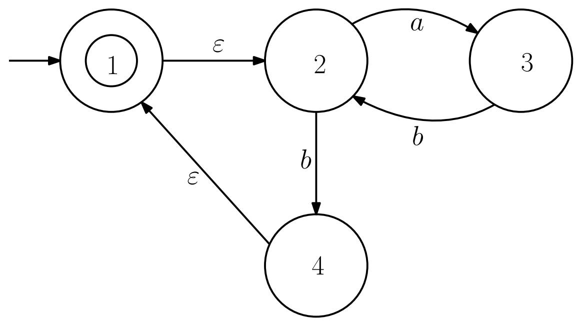


Why does this NFA accept the language A':

- For every input string, The NFA starts in the start state of the DFA M. There are two options:
  - The NFA always stays within the DFA M. In this way, all strings in A are accepted (and nothing else).

– Initially, the NFA stays within the DFA M and at some moment, it switches to the copy M'. This moment indicates that the NFA switches an a for a b, or a b for an a. Afterwards, the NFA stays withing the copy M' and does the same as the DFA M would have done. Notice that, after switching to the copy M', the NFA cannot go back to the original DFA M. Thus, in this second option, the NFA accepts all strings that differ in exactly one position from a string in A.

Question 8: Use the construction given in class to convert the following NFA (with alphabet $\{a, b\}$) to an equivalent DFA.



Show the full state diagram of the DFA; it has $2^4 = 16$ states. Afterwards, simplify the diagram by removing states that cannot be reached from the start state (in case this is possible).

Solution: Each state of the DFA corresponds to a subset of $\{1, 2, 3, 4\}$. Thus, the DFA has $2^4 = 16$ states. The accept states of the DFA are all states that contain the accept state 1 of the NFA. Thus, the DFA has $2^3 = 8$ accept states.

What is the start state of the DFA: We take the start state 1 of the NFA and add to it all states that can be reached by making zero or more ϵ-transitions. Since there is exactly one ϵ-transition leaving state 1, the start state of the DFA is $\{1, 2\}$.

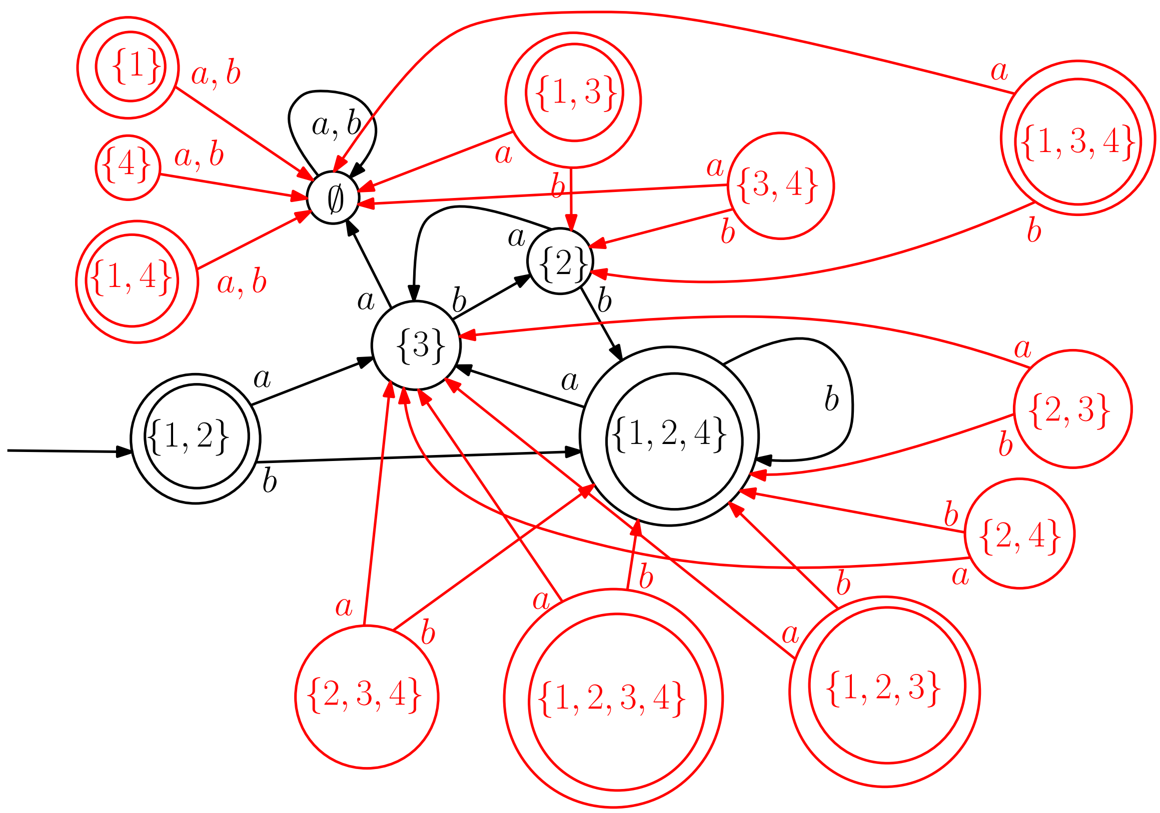
Now we need the transitions of the DFA. Remember, to simulate one step of the NFA, the DFA reads one symbol (a or b) and then follows zero or more ϵ-transitions.

- State $\emptyset$, read a or b: stay in state $\emptyset$.
- State $\{1\}$, read a or b: In the NFA, there are no a- or b-transitions leaving state 1. Thus, we will go to the state $\emptyset$.
- State $\{2\}$, read a: In the NFA, start in state 2 and read a, which takes us to state 3. There are no ϵ-transitions leaving state 3. Thus, the DFA goes to the state $\{3\}$.
- State $\{2\}$, read b: In the NFA, start in state 2 and read b, which takes us to state 4. From state 4, by only making ϵ-transitions, we can reach the states 1 and 2. Thus, the DFA goes to the state $\{1, 2, 4\}$.

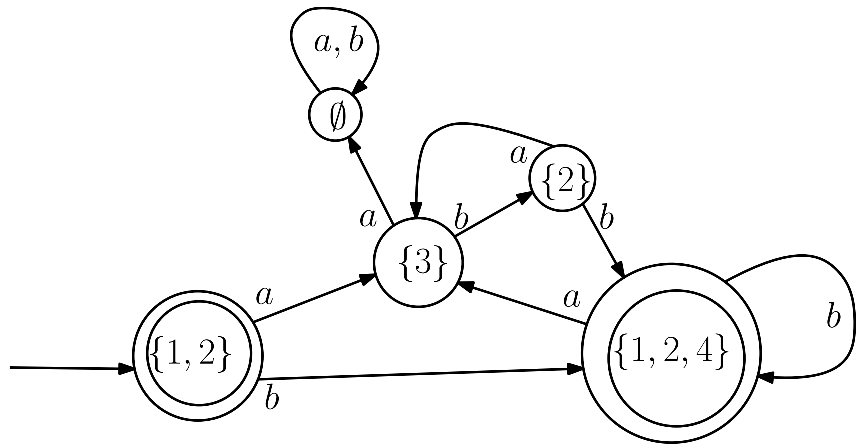
- State $\{3\}$, read a: In the NFA, there is no a-transition leaving state 3. Thus, we will go to the state $\emptyset$.
- State $\{3\}$, read b: In the NFA, start in state 3 and read b, which takes us to state 2. There are no ϵ-transitions leaving state 2. Thus, the DFA goes to the state $\{2\}$.
- State $\{4\}$, read a or b: In the NFA, there are no a- or b-transitions leaving state 4. Thus, we will go to the state $\emptyset$.
- State $\{1, 2\}$, read a: Take the union of the DFA transitions for $\{1\}$ and $\{2\}$, which is $\{3\}$.
- State $\{1, 2\}$, read b: Take the union of the DFA transitions for $\{1\}$ and $\{2\}$, which is $\{1, 2, 4\}$.
- State $\{1, 3\}$, read a: Take the union of the DFA transitions for $\{1\}$ and $\{3\}$, which is $\emptyset$.
- State $\{1, 3\}$, read b: Take the union of the DFA transitions for $\{1\}$ and $\{3\}$, which is $\{2\}$.
- State $\{1, 4\}$, read a or b: Take the union of the DFA transitions for $\{1\}$ and $\{4\}$, which is $\emptyset$.
- State $\{2, 3\}$, read a: Take the union of the DFA transitions for $\{2\}$ and $\{3\}$, which is $\{3\}$.
- State $\{2, 3\}$, read b: Take the union of the DFA transitions for $\{2\}$ and $\{3\}$, which is $\{1, 2, 4\}$.
- State $\{2, 4\}$, read a: Take the union of the DFA transitions for $\{2\}$ and $\{4\}$, which is $\{3\}$.
- State $\{2, 4\}$, read b: Take the union of the DFA transitions for $\{2\}$ and $\{4\}$, which is $\{1, 2, 4\}$.
- State $\{3, 4\}$, read a: Take the union of the DFA transitions for $\{3\}$ and $\{4\}$, which is $\emptyset$.
- State $\{3, 4\}$, read b: Take the union of the DFA transitions for $\{3\}$ and $\{4\}$, which is $\{2\}$.
- State $\{1, 2, 3\}$, read a: Take the union of the DFA transitions for $\{1\}$, $\{2\}$, and $\{3\}$, which is $\{3\}$.
- State $\{1, 2, 3\}$, read b: Take the union of the DFA transitions for $\{1\}$, $\{2\}$, and $\{3\}$, which is $\{1, 2, 4\}$.

- State $\{1, 2, 4\}$, read a: Take the union of the DFA transitions for $\{1\}$, $\{2\}$, and $\{4\}$, which is $\{3\}$.
- State $\{1, 2, 4\}$, read b: Take the union of the DFA transitions for $\{1\}$, $\{2\}$, and $\{4\}$, which is $\{1, 2, 4\}$.
- State $\{1, 3, 4\}$, read a: Take the union of the DFA transitions for $\{1\}$, $\{3\}$, and $\{4\}$, which is $\emptyset$.
- State $\{1, 3, 4\}$, read b: Take the union of the DFA transitions for $\{1\}$, $\{3\}$, and $\{4\}$, which is $\{2\}$.
- State $\{2, 3, 4\}$, read a: Take the union of the DFA transitions for $\{2\}$, $\{3\}$, and $\{4\}$, which is $\{3\}$.
- State $\{2, 3, 4\}$, read b: Take the union of the DFA transitions for $\{2\}$, $\{3\}$, and $\{4\}$, which is $\{1, 2, 4\}$.
- State $\{1, 2, 3, 4\}$, read a: Take the union of the DFA transitions for $\{1\}$, $\{2\}$, $\{3\}$, and $\{4\}$, which is $\{3\}$.
- State $\{1, 2, 3, 4\}$, read b: Take the union of the DFA transitions for $\{1\}$, $\{2\}$, $\{3\}$, and $\{4\}$, which is $\{1, 2, 4\}$.

Here is the state diagram of the DFA:



As you can see in this diagram, the red states cannot be reached from the start state $\{1, 2\}$. Thus, we can remove these, which leads to the final DFA:



This was painful. I need a beer!