York University
CSE 2001 Fall 2017 – Assignment 2 of 4
Instructor: Jeff Edmonds

Family Name: ______________________ Given Name: ______________________

Student #: ______________________ Email: ______________________

Family Name: ______________________ Given Name: ______________________

Student #: ______________________ Email: ______________________

1) Program to DFA	$12 + 12$	
2) Build NFA	$0 + 6 + 6 + 12$	
3) NFA to DFA	$12 + 12 + 12$	
4) NFA into a Reg	16	
0) Art	2	
Total	102 marks	

This exam is designed to be completed in an hour.
Keep your answers short and clear.

0) (2 marks) Art therapy question: When half done the exam, draw a picture of how you are feeling.

1. **Program to DFA:**
   Note in binary if $x = 101_2 = 5$ and $y = 1011_2 = 11$ then $y = 2 \cdot x + 1$.
   Remember $x \ mod \ 3 = 2$ is the remainder when you divide x by 3.
   Consider the following program:

   ```
   q = 0
   loop until no more characters
       get(c)                    % c ∈ {0,1}
       q = (2·q + c) mod 3       % q ∈ {0,1,2}
   end loop
   if q = 0 then
       return("accept")
   else
       return("reject")
   end if
   ```

   (a) Describe this language in one easy to understand English sentence.
   Hint: Look at examples in slides.

   (b) Convert the program into a DFA.

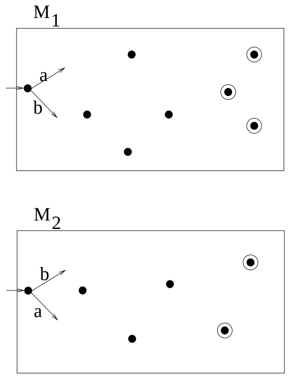
   (c) Convert the DFA into a regular expression.

2. **NFA:**
   Let L_1 and L_2 be arbitrary languages and let M_1 and M_2 below be diagrams representing the NFA for them.

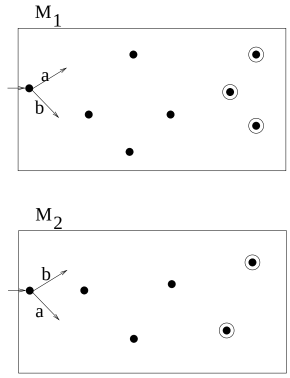
   (a) Explain (as if to a 1030 student) the key differences between the languages $L_1^* \cup L_2^*$ and $(L_1 \cup L_2)^*$. Give an example of a string that is in one but not in the other and vice versa.

   - Answer in Notes: Let $L_1 = \{a\}$ and $L_1 = \{b\}$. $(L_1^* \cup L_2^*)$ contains strings that either only contain a's or only contain b's. On the other hand, $(L_1 \cup L_2)^*$ contains strings that contains only a's and b's. Let $\omega = ab$. It is in the second and but not the first. Everything in the first is in the second.

(b) Draw an NFA for the language $L_1^* \cup L_2^*$ for this generic L_1 and L_2.

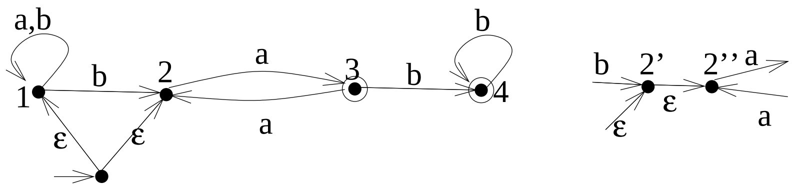


(c) Draw an NFA for the language $(L_1 \cup L_2)^*$.



(d) Explain (as if to a 1030 student) the key differences between the structures of your $L_1^* \cup L_2^*$ and $(L_1 \cup L_2)^*$ NFAs and why these differences cause the difference in the languages accepted. Hint: Describe how clones can travel through the machines. Use the word "commit".

3. Consider the following NFA.



Our goal is to explain in words the language $L(M)$ accepted by this NFA and then to prove by loop invariants (induction) that $L(M) = L$.

Goal: Prove $L(M) = L$ by proving $\forall q_i, M(q_i) = L(q_i)$.

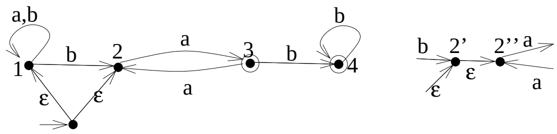
The Machine and its Labeled Paths: $M(q_i) = \{\alpha$ with path to state $q_i\}$.
i.e. it is just the definition of $L(M_i)$ for the NFA M_i where q_i is the only accept state.

The Language and Properties of its Strings: $L(q_i) = \{\alpha$ with some property$\}$.

Suppose the DFA machine has read in the substring "abbaaa" so far. What do you want it to be remembering about this substring? What does Pooh write on his black board? What is the common property of the strings (including this one) that you want to arrive at this state? We will denote the set of strings with this property as $L(q_3)$. Effectively, what is the "meaningful name" you are giving to this state?

(a) For each state q_i, we *guessed* what that set $M(q_i)$ would be with a *Name* $L(q_i)$.
Hint: Split state q_2 as done into two states q_2' and q_2'' and note that this does not change the language.
Hint: Also define the languages $L(q_2''$ to $q_2'')$ and $L(q_2''$ to $q_3)$ in which the start state is q_2'' and accept state is either q_2'' or q_3.

(b) Define the language $L(M)$ accepted by machine M in terms of these state languages $M(q_i)$. In doing so, determine in words what language is accepted by this NFA.

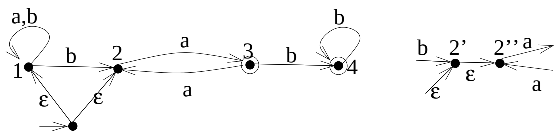


(c) Write an extended regular expression that expresses the same language. Do not do any long conversion.

(d) Our goal is to prove that $\forall i, M(q_i) = L(q_i)$.
Our loop invariant after having read t characters will be
$LI_t =$ "$\forall i, \forall \alpha$ of length t, $\alpha \in M(q_i)$ iff $\alpha \in L(q_i)$".
Assume LI_t. Consider an arbitrary string αc with length $t+1$.
You must prove $\alpha c \in M(q_2)$ iff $\alpha c \in L(q_2)$

(e) State the general *required connections* that must be true about these guessed sets $L(q_i)$?
For each state, state the required connection with respect to $L(q_i)$.
Argue that this condition is in fact true.

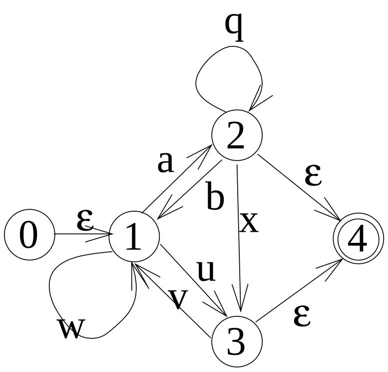
(f) Use loop invariants (induction) and these verified required conditions to prove the machine M computes the stated language, i.e. $M(L) = L$.



(g) Without doing the conversion, design a DFA for this language. Label the states with meaningful names.
Hint: The loop invariant states that what is remembered about the prefix read so far is:
- whether we are working on a block of a's or a block of b's.
- whether the last block of a's has even or odd length.
This implies there are four states.
You don't need to, but my DFA collapses two of these states into one.

(h) Do the steps with the table to convert this NFA into a DFA.

4. Do one step of converting this NFA into a regular expression by ripping out state 2.



Rip 2