DFA to Regex

Noah Buchanan

CS 4043: Formal Languages
University of Arkansas - Fort Smith
Fall 2021

October 27, 2021

Before we begin converting the DFA to a regular expression I will lay out all the necessary rules and steps. First of all it is important to keep in mind the following identites:

$$R \cup \emptyset = R$$
$$R \circ \epsilon = R$$
$$R \cup \epsilon \text{ may not equal } R$$
$$R \circ \emptyset \text{ may not equal } R$$

The last two identities are possibly not R for the third identity only in the case that R does not already have ϵ in its language, and for the fourth identity only in the case that R is not already the empty set. The empty set concatenated to anything is the empty set. The star operation of an empty set is the empty string, ϵ.

To begin converting a DFA to a regular expression the DFA must first be converted to a GNFA. GNFAs must conform to the following special conditions:

- The start state has transition arrows going to every other state but no arrows coming in from any other state.
- There is only a single accept state, and it has arrows coming in from every other state but no arrows going to any other state. Furthermore, the accept and start states must be mutually exclusive.
- Except for the start and accept states, one arrow goes from every state to every other state and also from each state to itself.

Let me clarify the last point: An arrow must point to every other state and itself excluding the accept and start states, however in the case that there is not a transition already from one state to another we simply denote that transition as $\emptyset$.

Lastly we need a way to find the new regular expression between two nodes where we are ripping out the middle man between them that connects them. In the original DFA, if

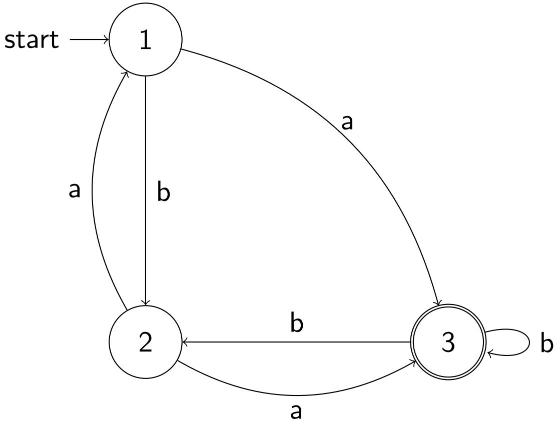
1. q_i goes to q_{rip} with an arrow labeled R_1,
2. q_{rip} goes to itself with an arrow labeled R_2,
3. q_{rip} goes to q_j with an arrow labeled R_3, and
4. q_i goes to q_j with an arrow labeled R_4,

then in the new GNFA we are simplifying, the new arrow from q_i to q_j takes the following form

$$(R_1) \circ (R_2)^* \circ (R_3) \cup (R_4)$$

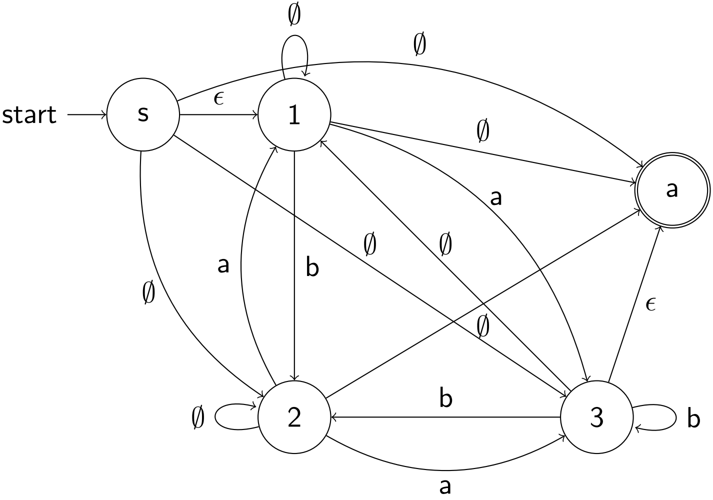
We know that all these cases are possible for any GNFA because of the $\emptyset$ transitions we add to conform to GNFAs special form. Now it is simply a matter of putting it together.

Here is the DFA we will be converting:



You will notice that it does not conform to a GNFAs special form as of right now. We must add the necessary pieces.

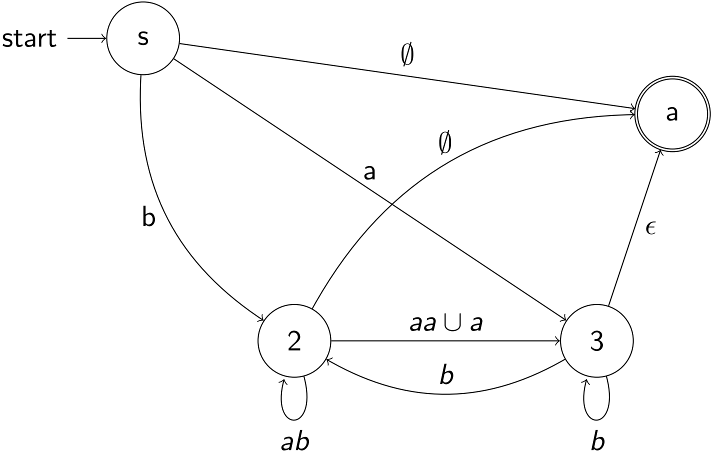
Lets add the new start and accept states and add any necessary but not present transitions.



We will use the steps specified on the fourth slide to create new connections for the in and out arrows on the node we are ripping e.g. for this first item s will be q_i, 1 is q_{rip}, and 2 is q_j.

- $s \rightarrow 1 \rightarrow 2 : (\epsilon) \circ (\emptyset)^* \circ (b) \cup (\emptyset) = b$
- $s \rightarrow 1 \rightarrow 3 : (\epsilon) \circ (\emptyset)^* \circ (a) \cup (\emptyset) = a$
- $s \rightarrow 1 \rightarrow a : (\epsilon) \circ (\emptyset)^* \circ (\emptyset) \cup (\emptyset) = \emptyset$
- $2 \rightarrow 1 \rightarrow 3 : (a) \circ (\emptyset)^* \circ (a) \cup (a) = aa \cup a$
- $3 \rightarrow 1 \rightarrow 3 : (\emptyset) \circ (\emptyset)^* \circ (a) \cup (b) = b$
- $2 \rightarrow 1 \rightarrow 2 : (a) \circ (\emptyset)^* \circ (b) \cup (\emptyset) = ab$
- $2 \rightarrow 1 \rightarrow a : (a) \circ (\emptyset)^* \circ (\emptyset) \cup (\emptyset) = \emptyset$
- $3 \rightarrow 1 \rightarrow 2 : (\emptyset) \circ (\emptyset)^* \circ (b) \cup (b) = b$
- $3 \rightarrow 1 \rightarrow a : (\emptyset) \circ (\emptyset)^* \circ (\emptyset) \cup (\epsilon) = \epsilon$

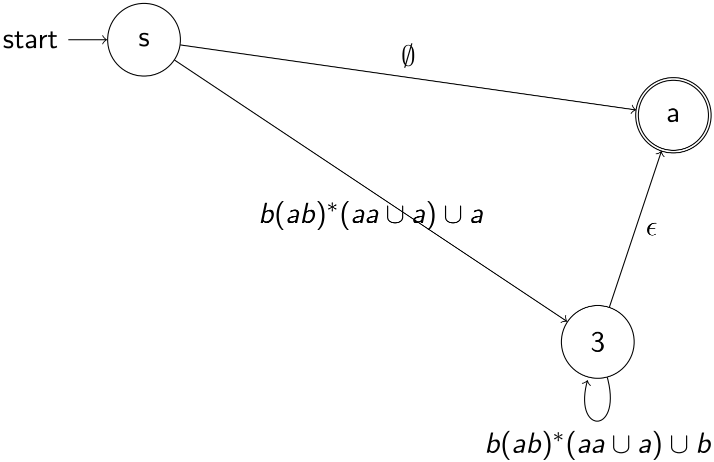
Now we create the connections for our GNFA with one less state.



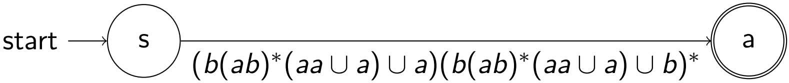
Now we must choose a new rip node, we will use 2 now.

- $s \to 2 \to 3 : (b) \circ (ab)^* \circ (aa \cup a) \cup (a) = b(ab)^*(aa \cup a) \cup a$
- $s \to 2 \to a : (b) \circ (ab)^* \circ (\emptyset) \cup (\emptyset) = \emptyset$
- $3 \to 2 \to 3 : (b) \circ (ab)^* \circ (aa \cup a) \cup (b) = b(ab)^*(aa \cup a) \cup b$
- $3 \to 2 \to a : (b) \circ (ab)^* \circ (\emptyset) \cup (\epsilon) = \epsilon$

New GNFA with one less state.



Now finally, we remove the last node between s and a:
$s \to 3 \to a : (b(ab)^*(aa \cup a) \cup a) \circ (b(ab)^*(aa \cup a) \cup b)^* \circ (\epsilon) \cup (\emptyset) =$
$(b(ab)^*(aa \cup a) \cup a)(b(ab)^*(aa \cup a) \cup b)^*$.



Thus we have our final regular expression to represent the original DFA.