

It is often the case that we want to know how we can manipulate a class of objects and still end up with something in the same class (see Group Theory/Abstract Algebra). We call these manipulations (operations) closed on a class if the above property holds.

Ex) The integers are closed over addition.

$$\forall x, y \in \mathbb{Z}, x + y \in \mathbb{Z}.$$

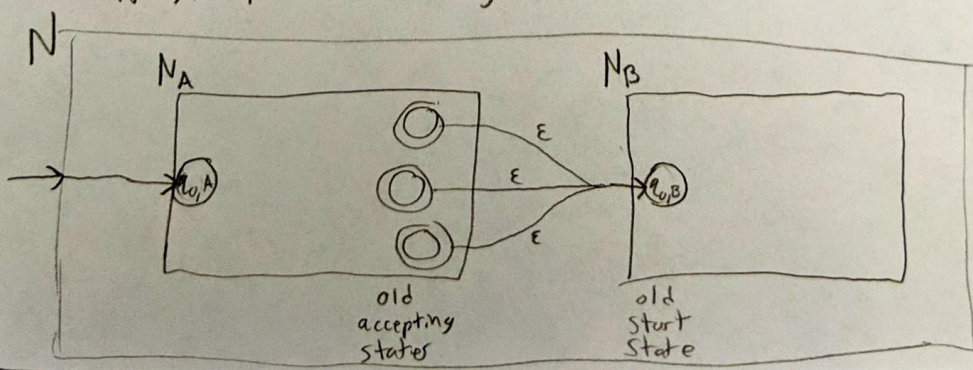
The regular languages are closed under a number of interesting properties. We've seen a few in exercises so far for specific cases. Here, we'll formally prove three particularly important closure properties.

Thm) The regular languages are closed under concatenation.

Pf) Let  $A, B$  be regular languages. Then there are NFAs  $N_A$  and  $N_B$  which accept  $A, B$  respectively. We construct a new NFA  $N$  to recognize  $A \circ B$  as follows.

For  $x, y \in \Sigma^*$   
 $x \circ y = xy$ ,  
that is  $x$   
followed by  $y$

For  $A, B$   
regular,  
 $A \circ B =$   
 $\{ab \mid a \in A, b \in B\}$



Let  $N_A = (Q_A, \Sigma_A, \delta_A, q_{0,A}, F_A)$  and  $N_B = (Q_B, \Sigma_B, \delta_B, q_{0,B}, F_B)$ .

Define  $N = (Q_A \cup Q_B, \Sigma_A \cup \Sigma_B, \delta, q_{0,A}, F_B)$ , where for  $c \in (\Sigma_A \cup \Sigma_B) \setminus \epsilon$   
and  $q \in Q_A \cup Q_B$ ,

$$\delta(q, c) = \begin{cases} \delta_A(q, c) & q \in Q_A \setminus F_A \text{ and } c \in (\Sigma_A) \setminus \epsilon \\ \delta_A(q, c) & q \in F_A \text{ and } c \in \Sigma_A \\ \{q_{0,B}\} \cup \delta_A(q, c) & q \in F_A \text{ and } c = \epsilon \\ \delta_B(q, c) & q \in Q_B \text{ and } c \in (\Sigma_B) \setminus \epsilon \end{cases}$$

It remains to show  $L(N) = A \circ B$ .

Let  $w \in L(N)$ . Then  $\exists$  an accepting path  $p$  from  $q_{0,A}$  to some  $q_F \in F_B$  on input  $w$ . By construction,  $q_{0,B}$  must be in  $p$  and  $p$  gets to  $q_{0,B}$  via an  $\epsilon$  transition from a state  $q'_F \in F_A$ .

Thus  $p = rs$  where  $r$  is a path starting from  $q_{0,A}$  and ending in  $q'_F$  while  $s$  is a path starting from  $q_{0,B}$  and ending at  $q_F$ .

As such,  $r$  accepts some string in  $A$  and  $s$  accepts some string in  $B$ , hence  $w \in A \circ B$ .

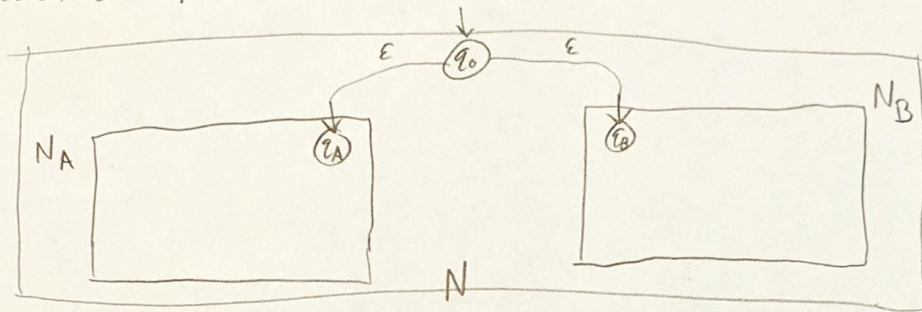
Now let  $w \in A \circ B$ .

Similarly to before,  $w = uv$  such that  $u \in A$  and  $v \in B$ , so  $\exists$  accepting paths  $r$  and  $s$  for  $u$  and  $v$  in  $N_A$  and  $N_B$  respectively. In  $N$ ,  $r$  and  $s$  is bridged via an  $\epsilon$  transition by construction, so  $rs$  is a path in  $N$ . Moreover, since  $s$  end in a state in  $F_B$ ,  $rs$  is an accepting path, thus  $rs \in L(N)$ .

□

Thm) The regular languages are closed under union.

Pf) Let  $A, B$  be regular languages. Then there are NFAs  $N_A = (Q_A, \Sigma_A, \delta_A, q_A, F_A)$  and  $N_B = (Q_B, \Sigma_B, \delta_B, q_B, F_B)$  such that  $L(N_A) = A$  and  $L(N_B) = B$ .  
We construct NFA  $N$  to accept  $A \cup B$  as follows.



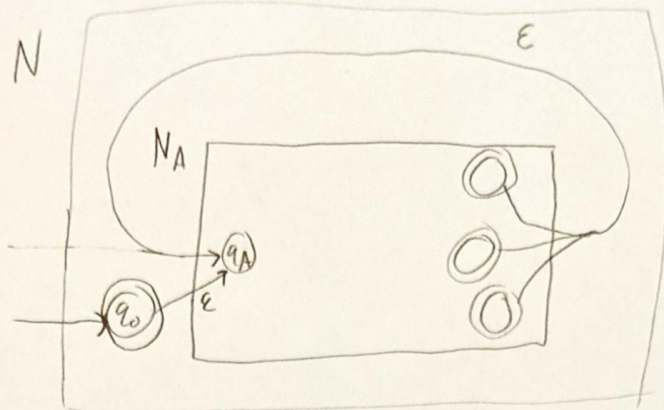
We define  $N = (Q_A \cup Q_B, \Sigma_A \cup \Sigma_B, \delta, q_0, F_A \cup F_B)$ , where for  $q \in Q_A \cup Q_B$  and  $c \in (\Sigma_A \cup \Sigma_B)_\epsilon$ ,

$$\delta(q, c) = \begin{cases} \{q_A, q_B\} & q = q_0 \text{ and } c = \epsilon \\ q_A(q, c) & q \in Q_A \text{ and } c \in (\Sigma_A)_\epsilon \\ q_B(q, c) & q \in Q_B \text{ and } c \in (\Sigma_B)_\epsilon \end{cases}$$

It is routine to verify that  $L(N) = A \cup B$ .

Thm) The regular languages are closed under <sup>□</sup> (Kleene) star. def on page 4

Pf) Let  $A$  be a regular language. Then there is an NFA  $N_A = (Q_A, \Sigma_A, \delta_A, q_A, F_A)$  such that  $L(N_A) = A$ .  
We construct NFA  $N$  to accept  $A^*$  as follows.



We define  $N = (Q_A \cup \{q_0\}, \Sigma_A, \delta, q_0, F_A \cup \{q_0\})$  where for  $q \in Q_A \cup \{q_0\}$  and  $c \in (\Sigma_A)_\epsilon$ ,

$$\delta(q, c) = \begin{cases} \delta(q, c) & q \in Q_A \setminus F_A \\ \delta(q, c) & q \in F_A \text{ and } c \in \Sigma_A \\ \delta(q, c) \cup \{q_A\} & q \in F_A \text{ and } c = \epsilon \\ \{q_A\} & q = q_0 \text{ and } c = \epsilon. \end{cases}$$

Note that because we must accept  $\epsilon$  and don't know if  $N_A$  contains a cycle starting at  $q_A$ , we must add accepting start state  $q_0$ .

From here it is routine to verify that  $L(N) = A^*$ . □

Def) For a language  $A$  over an alphabet  $\Sigma$ ,

$$A^* = \{w_1 w_2 w_3 \dots w_k \mid \forall i, w_i \in \Sigma^+ \text{ and } 0 \leq k\}$$

← when  $k=0$ , we get the empty string

For  $w \in \Sigma^*$ ,  $w^* = \{w\}^*$ .