Regexp-B-NFA example:

$$((ab)^* + a^*) \cdot (ba + bb)^*$$

The whole expression parses as $$((r_1 + r_2) \cdot r_3$$

$$r_3$$ could be

uses $$ba + bb = b(a+b)$$

Dead strings: baa, bba, more strings of length ≥4

Does N accept $$\varepsilon$$? Yes because the Epsilon-closure $$E(5)$$ includes the accepting state f.

$$N':$$

$$L(N') = \{ b \cdot (\varepsilon \cup (aa + a)^*b) \}$$

Rem does not allow bba

The E-closure $$E(p)$$ of a state p is the set of all states reachable by zero or more E-arcs out of p.

$$E(f) = \{ f, g \}$$
Theorem: For any NFA $N = (Q, \Sigma, E, s, F)$, we can build a DFA $M = (P(Q), \Sigma, \Delta, S, F)$ s.t. $L(M) = L(N)$.

Abstract idea: States of $M$ are subsets of possible states of $N$ at any particular step $i$ of processing an input $x = x_1 \cdots x_i \cdots x_n$.

Proof: starts in the state $S = E(s)$.

We will ensure that every intermediate state $P, R, \ldots$ and our final state $T$ are also $E$-closed.

Defn: The $E$-closure $E(P)$ of a set of states $P$ is the union of $E(p)$ over all $p \in P$. $E(P) = \bigcup_{p \in P} E(p)$.

Proof: Build $S = E(s)$, $\Delta(P, c) = E(s \triangleleft E(c \triangleright P))$ for some $p \in P$ and for any $p \in Q, c \in \Sigma$:

Claim: For each $i$, the set $R$ of states that $N$ can be in after processing $x_i \cdots x_i$ equals $\Delta(P, x_i)$ where $P$ is the set of possible states after processing $x_i \cdots x_{i-1}$.
The claim is by induction. Base case $i=0$ holds since $S = E(s)$. For inductive hypothesis, suppose it holds for $p$, i.e., $P = $ \( \exists \rho : N \text{ can process } x_1 \cdots x_{i-1} \text{ from } s \text{ to } \rho \).

We've defined $R = E(\exists r : \text{ for some } p \in P, (p, r) \in S)$

We need to show $\Delta(P, c) = 0$ where $c = x_i$.

How $R = \exists r : N \text{ can process } x_1 \cdots x_i \text{ from } s \text{ to } r$.

Let any $r \in R$ be given as defined. Then $N$ can process $x_1 \cdots x_i$ from $s$ to $r$.

Suppose $N$ can process $x_1 \cdots x_i$ from $s$ to $r$. Show $r \in R$.

**Proof of (a):** By defn of $r \in R$, there is some state $p \in P$ such that $N$ can process $x_1 \cdots x_i \text{ from } s \text{ to } p$ and instruction $(p, c, q) \in S$ such that $r \in E(q)$. Then $N$ can process:

- $x_1 \cdots x_i \text{ from } s \text{ to } p$ (by Ind. Hyp).
- take the arc $p \xrightarrow{q} x_i$
- take 0 or more $\varepsilon$-arcs from q to r.

**In the processing of $r = X_1 \cdots x_i \cdots$, there must have been some legal step $(p, x_i, q) \in S$.

So $p \in P$, $(p, x_i, q) \in S$, and only $\varepsilon$-arcs were used from $q$ to $r$.

Thus $r \in E(\exists q = (p, x_i, q) \in S$ for some $p \in P) = \Delta(P, x_i)$. \( \square \)
This claim finally shows that the end state $T$ of the DFA equals $\exists r \in Q: N$ can process all of $x$ from $s$ to $r$.

By definition $x \in L(N)$, $N$ accepts $x$ if and only if some such $r$ is a final state, i.e., $r \in F$.

Hence $T \in \mathcal{F}$ iff $T$ includes some state in $F$.

So define $\mathcal{F} = \{ T \in Q: F \cap T \neq \emptyset \}$.

Then $L(M) = L(N)$. \qed