

# NFA con $\epsilon$ -mosse

## 1 Idea intuitiva

Gli **NFA con  $\epsilon$ -mosse**, o  **$\epsilon$ -NFA**, sono un'estensione del modello degli NFA in cui sono ammesse transizioni sulla stringa vuota  $\epsilon$ . In altre parole, un automa di questo tipo può effettuare delle mosse (transizioni) *spontaneamente*, senza consumare alcun simbolo di input, e si usa la stringa vuota  $\epsilon$  come simbolo per etichettare le transizioni che possono avvenire spontaneamente.

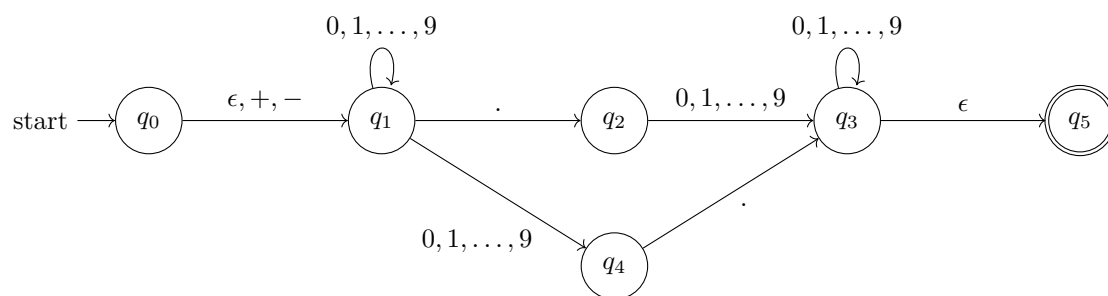
A scopo illustrativo, si consideri il linguaggio dei numeri decimali in notazione anglosassone, che sono composti da:

1. un segno  $+$  o  $-$  opzionale;
2. una prima sequenza di cifre decimali;
3. un punto;
4. una seconda sequenza di cifre decimali.

Una delle due sequenze di cifre può essere vuota, ma non entrambe. Alcune stringhe di questo linguaggio sono:

3.14   +125.   -125.0   .010   +.010

Un  $\epsilon$ -NFA che riconosce questo linguaggio è il seguente:



Per riconoscere, ad esempio, la stringa 3.14, questo automa:

1. parte, come sempre, dallo stato iniziale  $q_0$ ;

2. arriva in  $q_1$  seguendo la transizione etichettata con  $\epsilon$ ,  $q_0 \xrightarrow{\epsilon} q_1$ , cioè senza consumare alcun simbolo di input;
3. consuma il simbolo tre seguendo il coppia  $q_1 \xrightarrow{3} q_1$ ;
4. consuma il punto seguendo la transizione da  $q_1$  a  $q_2$ ;
5. passa a  $q_3$  consumando il simbolo 1;
6. arriva in  $q_3$  sfruttando il coppia  $q_3 \xrightarrow{4} q_3$ ;
7. giunge nello stato finale  $q_5$  mediante un'altra  $\epsilon$ -mossa, cioè senza consumare nulla (tanto è vero che tutti i simboli della in input sono già stati consumati).

In questo percorso di computazione, il ruolo della prima  $\epsilon$ -mossa è stato quello di esprimere l'“opzionalità” del simbolo + o -.

Invece, la stringa +125. viene riconosciuta seguendo il percorso

$$q_0 \xrightarrow{+} q_1 \xrightarrow{1} q_1 \xrightarrow{2} q_1 \xrightarrow{5} q_4 \xrightarrow{\cdot} q_3 \xrightarrow{\epsilon} q_5$$

nel quale la  $\epsilon$ -mossa da  $q_3$  a  $q_5$  cattura (in modo molto semplice, comodo) l'opzionalità della seconda sequenza di cifre.

## 2 Definizione formale

Dato un alfabeto  $\Sigma$ , si indica con  $\Sigma_\epsilon = \Sigma \cup \{\epsilon\}$  l'alfabeto  $\Sigma$  esteso con il simbolo  $\epsilon$ .

Un  $\epsilon$ -NFA è allora definito come una quintupla  $A = \langle Q, \Sigma, \delta, q_0, F \rangle$  in cui:

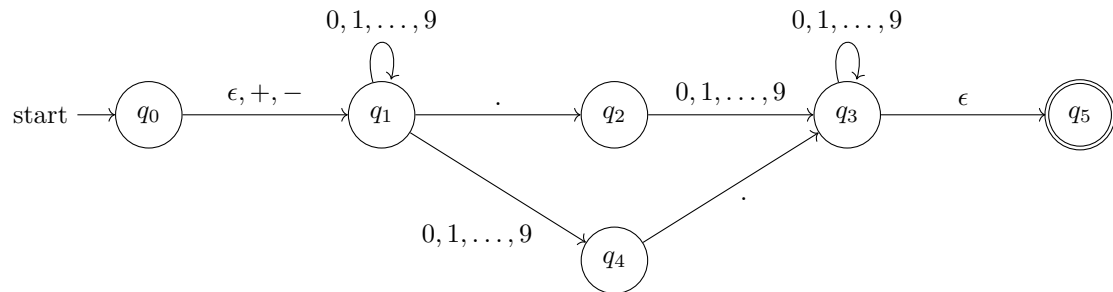
- $Q$  è l'insieme finito e non vuoto degli stati;
- $\Sigma$  è l'alfabeto di input;
- $\delta : Q \times \Sigma_\epsilon \rightarrow 2^Q$  è la funzione di transizione;
- $q_0 \in Q$  è lo stato iniziale;
- $F \subseteq Q$  è l'insieme degli stati finali.

Si osserva che l'unica differenza di questa definizione rispetto a quella di un “normale” NFA è il dominio della funzione di transizione, che qui è  $Q \times \Sigma_\epsilon$  invece di  $Q \times \Sigma$ . L'estensione dell'alfabeto con un simbolo ad-hoc  $\epsilon$  serve appunto a rappresentare le mosse che non consumano l'input.

È importante notare che tecnicamente, in questo contesto,  $\epsilon$  va interpretata come un semplice simbolo, che di per sé non ha alcun significato, e non invece come la stringa vuota ( $w = \epsilon \in \Sigma^*$ , la stringa che non contiene simboli). Volendo, infatti, si potrebbe scegliere qualunque altro simbolo (non già presente nell'alfabeto  $\Sigma$ ) per etichettare le transizioni che non consumano input, ma, proprio per il suo “collegamento” con il concetto di stringa vuota, la scelta di  $\epsilon$  può essere considerata in qualche modo intuitiva.

## 2.1 Esempio

Secondo questa definizione, l'automa introdotto nell'esempio di prima



viene caratterizzato formalmente come segue:

- $Q = \{q_0, q_1, q_2, q_3, q_4, q_5\}$ ;
- $\Sigma = \{+, -, ., 0, 1, \dots, 9\}$ ;
- stato iniziale  $q_0$ ;
- $F = \{q_5\}$ ;
- funzione di transizione  $\delta : Q \times \{\epsilon, +, -, ., 0, 1, \dots, 9\} \rightarrow 2^Q$ , rappresentata dalla seguente tabella:<sup>1</sup>

|                   | $\epsilon$  | $+, -$      | $.$         | $0, 1, \dots, 9$ |
|-------------------|-------------|-------------|-------------|------------------|
| $\rightarrow q_0$ | $\{q_1\}$   | $\{q_1\}$   | $\emptyset$ | $\emptyset$      |
| $q_1$             | $\emptyset$ | $\emptyset$ | $\{q_2\}$   | $\{q_1, q_4\}$   |
| $q_2$             | $\emptyset$ | $\emptyset$ | $\emptyset$ | $\{q_3\}$        |
| $q_3$             | $\{q_5\}$   | $\emptyset$ | $\emptyset$ | $\{q_3\}$        |
| $q_4$             | $\emptyset$ | $\emptyset$ | $\{q_3\}$   | $\emptyset$      |
| $*q_5$            | $\emptyset$ | $\emptyset$ | $\emptyset$ | $\emptyset$      |

Da questa tabella, si nota che la funzione di transizione è definita praticamente come per un normale NFA, senza trattare in modo “particolare” il simbolo  $\epsilon$ . Quello che invece cambierà sarà il modo in cui tale simbolo verrà trattato nelle computazioni.

<sup>1</sup>Per comodità, in questa tabella sono raggruppati i simboli  $+, -$  e analogamente i simboli  $0, 1, \dots, 9$ , dato che la funzione di transizione si comporta sempre allo stesso modo su tutti i simboli di ciascuno di questi gruppi.

### 3 $\epsilon$ -chiusura di uno stato

Per definire le computazioni di un  $\epsilon$ -NFA, bisogna introdurre il concetto di  **$\epsilon$ -chiusura** di uno stato e di un insieme di stati.

Intuitivamente, dato un  $\epsilon$ -NFA  $A = \langle Q, \Sigma, \delta, q_0, F \rangle$ , l' $\epsilon$ -chiusura  $\text{ECLOSE}(q)$  di uno stato  $q \in Q$  è l'insieme di tutti gli stati raggiungibili da esso utilizzando soltanto  $\epsilon$ -mosse. Formalmente, tale insieme è definito per induzione:

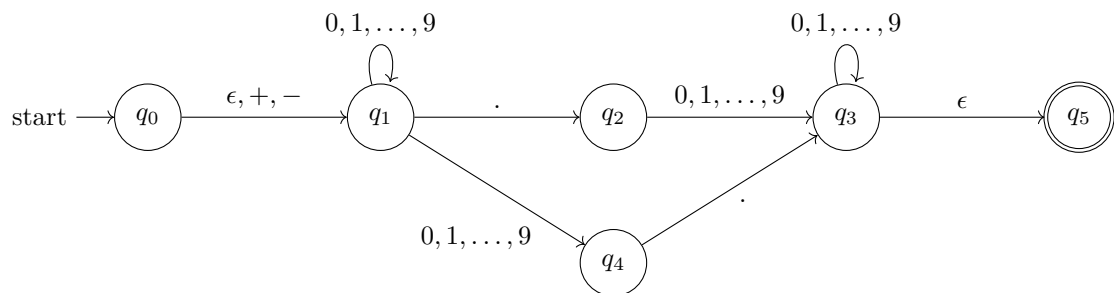
- *Base:*  $q \in \text{ECLOSE}(q)$ .
- *Induzione:* se  $p \in \text{ECLOSE}(q)$  ed esiste una  $\epsilon$ -transizione  $p \xrightarrow{\epsilon} r$  nell'automa, cioè  $r \in \delta(p, \epsilon)$ , allora anche  $r \in \text{ECLOSE}(q)$ . In alternativa, si può affermare che se  $p \in \text{ECLOSE}(q)$  allora  $\delta(p, \epsilon) \subseteq \text{ECLOSE}(q)$ .

Questa nozione può poi essere generalizzata in modo naturale a un insieme di stati  $S \subseteq Q$ , definendo:

$$\text{ECLOSE}(S) = \bigcup_{q \in S} \text{ECLOSE}(q)$$

#### 3.1 Esempio

Considerando ancora l'automa che riconosce i numeri decimali in notazione anglosassone,

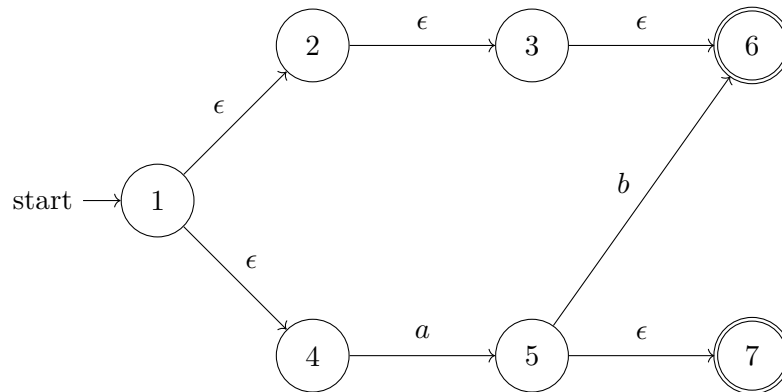


si vuole calcolare la  $\epsilon$ -chiusura per ogni stato dell'automa:

- Nel caso di  $q_0$ , si pone inizialmente  $\text{ECLOSE}(q_0) = \{q_0\}$ , per la base della definizione induttiva. Poi, per il passo induttivo, siccome  $q_0 \xrightarrow{\epsilon} q_1$  si aggiunge alla  $\epsilon$ -chiusura anche  $q_1$ , ottenendo  $\text{ECLOSE}(q_0) = \{q_0, q_1\}$ . Dal "nuovo" stato  $q_1$  non sono raggiungibili altri stati tramite  $\epsilon$ -mosse: avendo trattato tutti gli stati aggiunti a  $\text{ECLOSE}(q_0)$ , la costruzione è conclusa.
- $\text{ECLOSE}(q_1) = \{q_1\}$
- $\text{ECLOSE}(q_2) = \{q_2\}$

- $\text{ECLOSE}(q_3) = \{q_3, q_5\}$
- $\text{ECLOSE}(q_4) = \{q_4\}$
- $\text{ECLOSE}(q_5) = \{q_5\}$

Come altro esempio, un po' più articolato, si consideri invece l' $\epsilon$ -NFA rappresentato dal seguente diagramma:



L' $\epsilon$ -chiusura di ciascuno dei suoi stati è:

- $\text{ECLOSE}(1) = \{1, 2, 4, 3, 6\}$
- $\text{ECLOSE}(2) = \{2, 3, 6\}$
- $\text{ECLOSE}(3) = \{3, 6\}$
- $\text{ECLOSE}(4) = \{4\}$
- $\text{ECLOSE}(5) = \{5, 7\}$
- $\text{ECLOSE}(6) = \{6\}$
- $\text{ECLOSE}(7) = \{7\}$

Dato poi l'insieme di stati  $S = \{1, 4, 5\}$ , la sua  $\epsilon$ -chiusura è

$$\begin{aligned}
 \text{ECLOSE}(S) &= \bigcup_{q \in \{1,4,5\}} \text{ECLOSE}(q) \\
 &= \text{ECLOSE}(1) \cup \text{ECLOSE}(4) \cup \text{ECLOSE}(5) \\
 &= \{1, 2, 3, 4, 6\} \cup \{4\} \cup \{5, 7\} \\
 &= \{1, 2, 3, 4, 5, 6, 7\} = Q
 \end{aligned}$$

cioè comprende tutti gli stati dell'automa.

## 4 Computazione di un $\epsilon$ -NFA

Dato un  $\epsilon$ -NFA  $A = \langle Q, \Sigma, \delta, q_0, F \rangle$ , in cui  $\delta : Q \times \Sigma_\epsilon \rightarrow 2^Q$ , si definisce la **funzione di transizione estesa** di  $A$ ,  $\hat{\delta} : Q \times \Sigma^* \rightarrow 2^Q$ , per induzione sulla lunghezza della stringa  $w \in \Sigma^*$ :

- *Base*: se  $|w| = 0$ , cioè  $w = \epsilon$ , si definisce

$$\hat{\delta}(q, \epsilon) = \text{ECLOSE}(q)$$

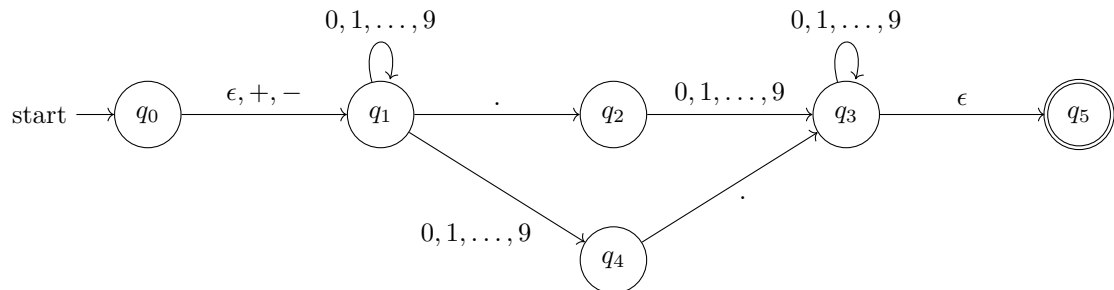
- *Passo induttivo*: se  $|w| > 0$ , allora  $w = xa$ , con  $x \in \Sigma^*$  e  $a \in \Sigma$ . Si definisce allora

$$\hat{\delta}(q, xa) = \text{ECLOSE} \left( \bigcup_{p \in \hat{\delta}(q, x)} \delta(p, a) \right)$$

Intuitivamente,  $\hat{\delta}(q, w)$  è l'insieme degli stati che l'automa può raggiungere partendo da  $q$  e processando l'intera stringa  $w$ , come per gli NFA, ma qui in più si tiene conto anche delle  $\epsilon$ -mosse.

### 4.1 Esempio

Per mostrare un esempio di computazione, si considera ancora l' $\epsilon$ -NFA che riconosce i numeri decimali.



|                   | $\epsilon$  | $+, -$      | $.$         | $0, 1, \dots, 9$ |
|-------------------|-------------|-------------|-------------|------------------|
| $\rightarrow q_0$ | $\{q_1\}$   | $\{q_1\}$   | $\emptyset$ | $\emptyset$      |
| $q_1$             | $\emptyset$ | $\emptyset$ | $\{q_2\}$   | $\{q_1, q_4\}$   |
| $q_2$             | $\emptyset$ | $\emptyset$ | $\emptyset$ | $\{q_3\}$        |
| $q_3$             | $\{q_5\}$   | $\emptyset$ | $\emptyset$ | $\{q_3\}$        |
| $q_4$             | $\emptyset$ | $\emptyset$ | $\{q_3\}$   | $\emptyset$      |
| $*q_5$            | $\emptyset$ | $\emptyset$ | $\emptyset$ | $\emptyset$      |

Si vuole calcolare la computazione di questo automa sulla stringa 125., cioè il valore della funzione  $\hat{\delta}(q_0, 125.)$ . Come al solito, invece di seguire direttamente la definizione induttiva, si procede per prefissi sempre più lunghi della stringa in input, partendo da  $\epsilon$ :

$$\begin{aligned}
\hat{\delta}(q_0, \epsilon) &= \text{ECLOSE}(q_0) = \{q_0, q_1\} \\
\hat{\delta}(q_0, 1) &= \text{ECLOSE}\left(\bigcup_{p \in \hat{\delta}(q_0, \epsilon)} \delta(p, 1)\right) = \text{ECLOSE}\left(\bigcup_{p \in \{q_0, q_1\}} \delta(p, 1)\right) \\
&= \text{ECLOSE}(\delta(q_0, 1) \cup \delta(q_1, 1)) = \text{ECLOSE}(\emptyset \cup \{q_1, q_4\}) = \text{ECLOSE}(\{q_1, q_4\}) \\
&= \bigcup_{p \in \{q_1, q_4\}} \text{ECLOSE}(p) = \text{ECLOSE}(q_1) \cup \text{ECLOSE}(q_4) = \{q_1\} \cup \{q_4\} \\
&= \{q_1, q_4\} \\
\hat{\delta}(q_0, 12) &= \text{ECLOSE}\left(\bigcup_{p \in \{q_1, q_4\}} \delta(p, 2)\right) = \text{ECLOSE}(\{q_1, q_4\}) = \{q_1, q_4\} \\
\hat{\delta}(q_0, 125) &= \text{ECLOSE}\left(\bigcup_{p \in \{q_1, q_4\}} \delta(p, 5)\right) = \text{ECLOSE}(\{q_1, q_4\}) = \{q_1, q_4\} \\
\hat{\delta}(q_0, 125.) &= \text{ECLOSE}\left(\bigcup_{p \in \{q_1, q_4\}} \delta(p, \cdot)\right) = \text{ECLOSE}(\{q_2, q_3\}) = \{q_2, q_3, q_5\}
\end{aligned}$$

Siccome

$$\hat{\delta}(q_0, 125.) \cap F = \{q_2, q_3, q_5\} \cap \{q_5\} = \{q_5\} \neq \emptyset$$

per definizione la stringa 125. è accettata dall'automa, ovvero appartiene al linguaggio riconosciuto dall'automa.

## 5 Linguaggio accettato da un $\epsilon$ -NFA

Sia  $A = \langle Q, \Sigma, \delta, q_0, F \rangle$  un  $\epsilon$ -NFA. Le definizioni di stringa e linguaggio accettati da  $A$  sono del tutto analoghe a quelle per i normali NFA:

- $A$  **accetta** una stringa  $w$  se e solo se  $\hat{\delta}(q_0, w) \cap F \neq \emptyset$ .
- Il **linguaggio riconosciuto** da  $A$  è l'insieme di tutte le stringhe accettate:

$$L(A) = \{w \in \Sigma^* \mid A \text{ accetta } w\} = \{w \in \Sigma^* \mid \hat{\delta}(q_0, w) \cap F \neq \emptyset\}$$