Azzolini Riccardo 2020-10-29

NFA con ϵ-mosse

1 Idea intuitiva

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

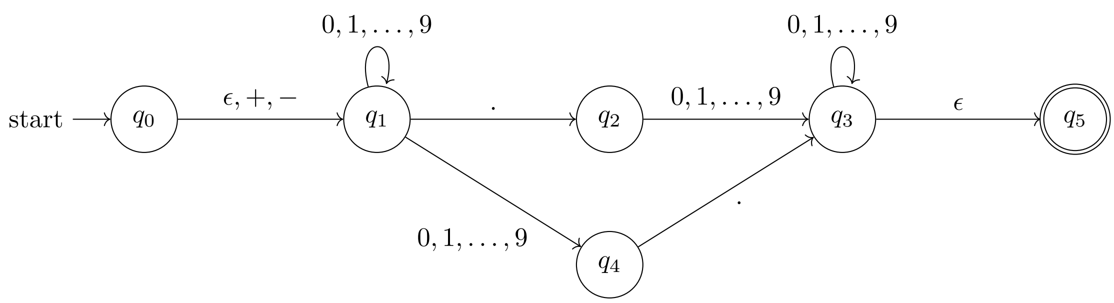
A scopo illustrativo, si consideri il linguaggio dei numeri decimali in notazione anglosassone, che sono composti da:

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:

$$3.14 \quad +125. \quad -125.0 \quad .010 \quad +.010$$

Un ϵ-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 ϵ, $q_0 \xrightarrow{\epsilon} q_1$, cioè senza consumare alcun simbolo di input;
3. consuma il simbolo tre seguendo il cappio $q_1 \xrightarrow{3} q_1$;
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 cappio $q_3 \xrightarrow{4} q_3$;
7. giunge nello stato finale q_5 mediante un'altra ϵ-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 ϵ-mossa è stato quello di esprimere l'"opzionalità" del simbolo $+$ o $-$.

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

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

nel quale la ϵ-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 Σ, si indica con $\Sigma_\epsilon = \Sigma \cup \{\epsilon\}$ l'alfabeto Σ esteso con il simbolo ϵ, che rappresenta la stringa vuota.

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

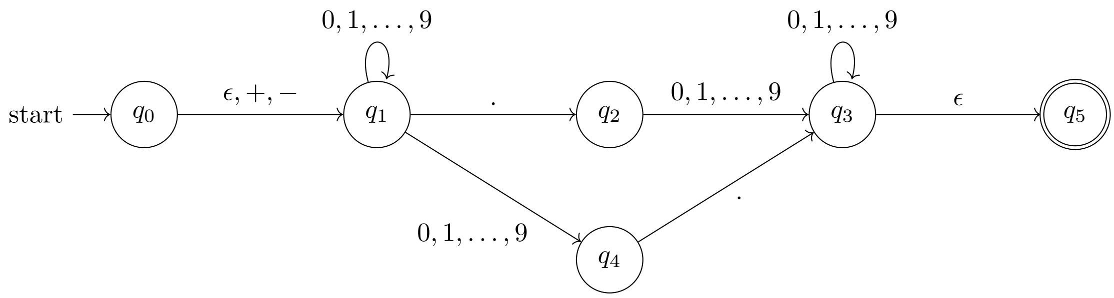
- Q è l'insieme finito e non vuoto degli stati;
- Σ è l'alfabeto di input;
- $\delta : Q \times \Sigma_\epsilon \to 2^Q$ è la funzione di transizione;
- $q_0 \in Q$ è lo stato iniziale;
- $F \subseteq Q$ è l'insieme degli stati finali.

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 ϵ[1] serve appunto a rappresentare le mosse che non consumano l'input.

[1] In teoria, si potrebbe scegliere qualunque simbolo, ma ϵ, rappresentando la stringa vuota, risulta in qualche modo intuitivo.

2.1 Esempio

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



viene caratterizzato formalmente come segue:

- $Q = \{q_0, q_1, q_2, q_3, q_4, q_5\}$;
- $\Sigma = \{+, -, ., 0, 1, \ldots, 9\}$;
- stato iniziale q_0;
- $F = \{q_5\}$;
- funzione di transizione $\delta : Q \times \{\epsilon, +, -, ., 0, 1, \ldots, 9\} \to 2^Q$, rappresentata dalla seguente tabella:[2]

	ϵ	$+, -$	$.$	$0, 1, \ldots, 9$
$\to q_0$	$\{q_1\}$	$\{q_1\}$	$\varnothing$	$\varnothing$
q_1	$\varnothing$	$\varnothing$	$\{q_2\}$	$\{q_1, q_4\}$
q_2	$\varnothing$	$\varnothing$	$\varnothing$	$\{q_3\}$
q_3	$\{q_5\}$	$\varnothing$	$\varnothing$	$\{q_3\}$
q_4	$\varnothing$	$\varnothing$	$\{q_3\}$	$\varnothing$
$*q_5$	$\varnothing$	$\varnothing$	$\varnothing$	$\varnothing$

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 ϵ. Quello che invece cambierà sarà il modo in cui tale simbolo verrà trattato nelle computazioni.

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

3 ϵ-chiusura di uno stato

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

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

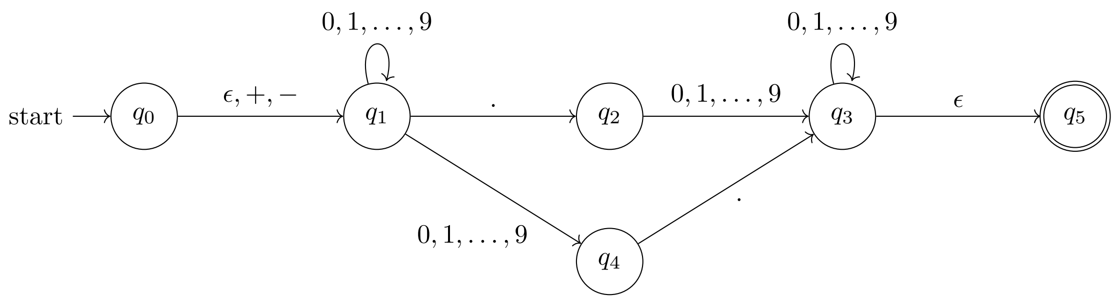
- *Base*: $q \in \textsc{eclose}(q)$.
- *Induzione*: se $p \in \textsc{eclose}(q)$ ed esiste una ϵ-transizione $p \xrightarrow{\epsilon} r$ nell'automa, cioè $r \in \delta(p, \epsilon)$, allora anche $r \in \textsc{eclose}(q)$. In alternativa, si può affermare che se $p \in \textsc{eclose}(q)$ allora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

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

$$\textsc{eclose}(S) = \bigcup_{q \in S} \textsc{eclose}(Q)$$

3.1 Esempio

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

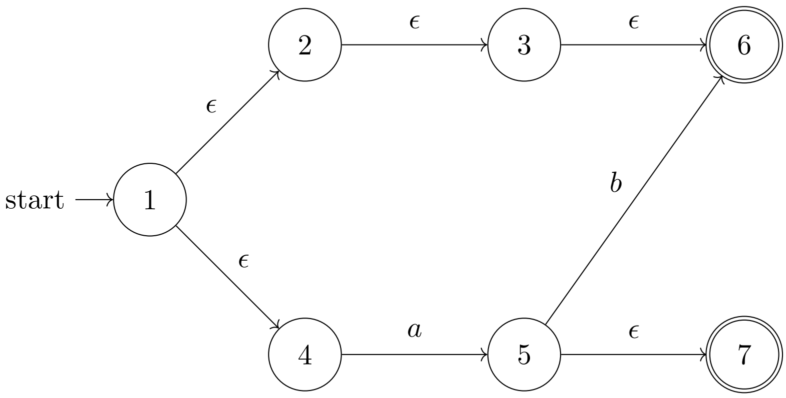


si vuole calcolare la ϵ-chiusura per ogni stato dell'automa:

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

- $\textsc{eclose}(q_3) = \{q_3, q_5\}$
- $\textsc{eclose}(q_4) = \{q_4\}$
- $\textsc{eclose}(q_5) = \{q_5\}$

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



L'ϵ-chiusura di ciascuno dei suoi stati è:

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

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

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

cioè comprende tutti gli stati dell'automa.

4 Computazione di un ϵ-NFA

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

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

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

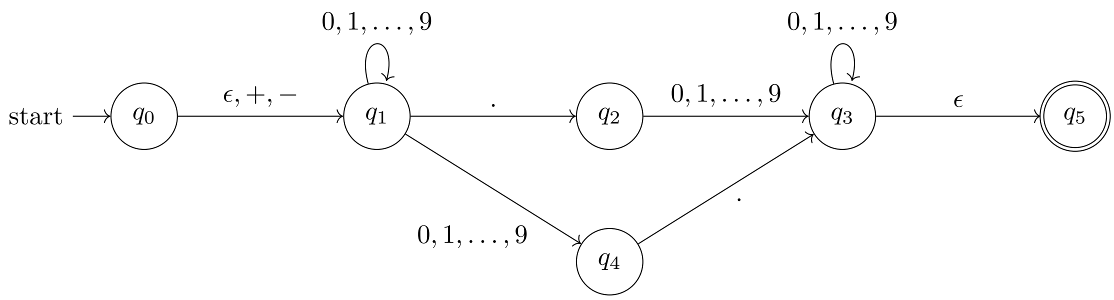
- *Passo induttivo*: se $|w| = 0$, allora $w = xa$, con $x \in \Sigma^*$ e $a \in \Sigma$. Si definisce allora

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

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 ϵ-mosse.

4.1 Esempio

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



	ϵ	$+, -$	.	$0, 1, \dots, 9$
$\to q_0$	$\{q_1\}$	$\{q_1\}$	$\varnothing$	$\varnothing$
q_1	$\varnothing$	$\varnothing$	$\{q_2\}$	$\{q_1, q_4\}$
q_2	$\varnothing$	$\varnothing$	$\varnothing$	$\{q_3\}$
q_3	$\{q_5\}$	$\varnothing$	$\varnothing$	$\{q_3\}$
q_4	$\varnothing$	$\varnothing$	$\{q_3\}$	$\varnothing$
$*q_5$	$\varnothing$	$\varnothing$	$\varnothing$	$\varnothing$

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 ϵ:

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

Siccome

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

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

5 Linguaggio accettato da un ϵ-NFA

Sia $A = \langle Q, \Sigma, \delta, q_0, F\rangle$ un ϵ-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 $\hat{\delta}(q_0, w) \cap F \neq \varnothing$.
- Il **linguaggio riconoscuito** da A è l'insieme di tutte le stringhe accettate:

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