EQUIVALENZA TRA DFA ED ESPRESSIONI REGOLARI (CONT.)

Teorema. *Sia M un DFA. Allora esiste una espressione regolare r tale che $L(M) = L(r)$.*

Dimostrazione: Sia $M = \langle Q, \Sigma, \delta, q_0, F \rangle$ con $Q = \{1, 2, \ldots, n\}$ e $q_0 = 1$.

Con $r_{ij}^{(k)}$ denotiamo l'espressione regolare che definisce l'insieme di stringhe che permettono di spostarsi dallo stato i allo stato j nel DFA M, senza passare attraverso stati maggiori di k (salvo i e j).

Definiamo induttivamente $r_{ij}^{(k)}$:

Base: con $k = 0$ ci sono solo due possibilità:

1. una transizione dallo stato i allo stato j;
2. un cammino di lunghezza 0 e $i = j$.

Equivalenza tra DFA ed espressioni regolari (cont.)

Se $i \neq j$ allora è possibile solo il primo caso, quindi:

i. se non c'è nessuna transizione da i a j, allora $r_{ij}^{(0)} = \emptyset$;

ii. se c'è esattamente una transizione da i a j per il simbolo a, allora $r_{ij}^{(0)} = a$;

iii. se ci sono $p > 1$ simboli $a_1, a_2, \ldots, a_p$ corrispondenti ad altrettante transizioni da i a j, allora $r_{ij}^{(0)} = a_1 + a_2 + \cdots + a_p$.

Se $i = j$, allora è possibile avere un cammino di lunghezza 0. Le definizioni vanno dunque così modificate:

i. $r_{ij}^{(0)} = \varepsilon$ (nessun simbolo a);

ii. $r_{ij}^{(0)} = \varepsilon + a$ (un solo simbolo a);

iii. $r_{ij}^{(0)} = \varepsilon + a_1 + a_2 + \cdots + a_p$ (p simboli).

EQUIVALENZA TRA DFA ED ESPRESSIONI REGOLARI (CONT.)

Passo: Sia $k > 0$. Per ogni coppia di stati i e j, consideriamo tutti i cammini da i a j che attraversano stati non maggiori di k (possono essere infiniti). Per ognuno di questi ci sono due casi:

a. Il cammino non passa attraverso lo stato k: dunque, per ipotesi induttiva, la stringa corrispondente è in $r_{ij}^{(k-1)}$;

b. Il cammino passa attraverso lo stato k almeno una volta. In questo caso spezzeremo il cammino in tre parti:

   1. il primo andrà dallo stato i allo stato k senza passare attraverso k, e corrisponderà per i. i. ad una stringa in $r_{ik}^{(k-1)}$;
   2. l'ultimo andrà dallo stato k allo stato j senza passare attraverso k, e corrisponderà per i. i. ad una stringa in $r_{kj}^{(k-1)}$;

EQUIVALENZA TRA DFA ED ESPRESSIONI REGOLARI (CONT.)

3. la parte centrale sarà a sua volta la concatenazione di zero o più frammenti che vanno da k a k *senza* passare da k. Ognuno di questi frammenti corrisponde per i. i. ad una stringa in $r_{kk}^{(k-1)}$. La parte centrale corrisponderà quindi ad una stringa in $\left(r_{kk}^{(k-1)}\right)^*$.

Combinando i casi a. e b., otteniamo

$$r_{ij}^{(k)} = r_{ij}^{(k-1)} + r_{ik}^{(k-1)}\left(r_{kk}^{(k-1)}\right)^* r_{kj}^{(k-1)}.$$

Se $F = \{f_1, \ldots, f_h\}$, l'espressione cercata è

$$r = r_{1f_1}^{(n)} + \cdots + r_{1f_h}^{(n)}.$$