Teoremă (Kleene)**.** *Dacă limbajul* L *este acceptat de un* ***DFA*** A*, atunci există o* ***expresie regulată*** E *astfel încât* $L(E) = L$*.*

Fie $A = (Q, \Sigma, \delta, q_0, F)$ pentru care L(A) = L.
Nu este clar cum am putea găsi o expresie regulată care să reprezinte acest DFA. Soluția va fi să spargem problema în componente mai mici (pentru care putem găsi expresii regulate) și să ne folosim de inducție.

Renumerotăm stările cu indici de la 1 la $n = |Q|$ astfel încât starea inițială să fie 1: $A' = (\{1, 2, \ldots, n\}, \Sigma, \delta', 1, F')$.
Fiind doar o redenumire a stărilor, $A' \cong A$, deci $L(A') = L(A) = L$.

Pentru $1 \leq i, j \leq n, 0 \leq k \leq n$ definim $R^k_{i,j}$, o mulțime de cuvinte care îndeplinesc următoarele condiții:

- w este un cuvânt peste alfabetul Σ
- Dacă plecăm din starea i cu cuvântul w ajungem în starea j
- Toate stările intermediare în care ajungem au indicele cel mult k. Această condiție poate fi scrisă formal:

  Pentru orice descompunere $w = xy$, cu $|x| > 0$ și $|x| < |w|$ (adică x este un *prefix propriu* al lui w), avem că $t \leq k$, unde $t = \delta'(i, x)$ (t este o stare intermediară prin care trecem).

Toate aceste condiții puse într-o singură linie:

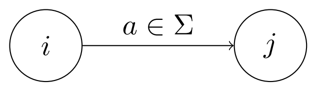
$$R^k_{i,j} = \{w \in \Sigma^* \mid \delta'(i, w) = j \;\; \& \;\; w = xy, 0 < |x| < |w|, \delta'(i, x) = t \leq k\}$$

Cuvintele acceptate de un automat sunt cele care din starea inițială ajung într-o stare finală. Dacă am avea expresiile regulate pentru $R^n_{1,f}$ unde 1 este starea inițială și f este o stare finală, am putea să construim o expresie regulată pentru întreg DFA-ul. Însă această descriere a mulțimii nu este suficient de explicită cât să ne ajute să găsim expresiile corespunzătoare.

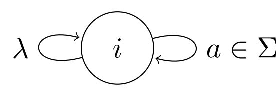
Încercăm să dăm o definiție recursivă a lui $R_{i,j}^k$.

Cazul de bază. Pentru $k = 0$, nu putem avea nicio stare intermediară $t \leq k$, deoarece indicele celei mai mici stări este 1. Deci mulțimea trebuie să fie formată din cuvinte de lungime 1 (adică din litere) pentru care există o tranziție de la i la j, cu observația că și λ este valid dacă $i = j$.

Dacă $i \neq j$: $R_{i,j}^0 = \{\ a \in \Sigma \mid \delta'(i, a) = j\ \}$

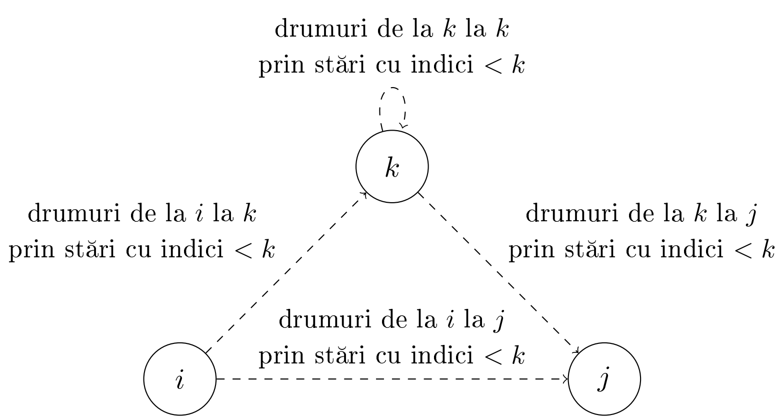


Altfel: $R_{i,i}^0 = \{\ a \in \Sigma \mid \delta'(i, a) = i\ \} \cup \{\ \lambda\ \}$



Ipoteza de inducție. Pentru $k \geq 1$, presupunem că știm deja cât este $R_{i,j}^{k-1}$, $\forall i, j \in \overline{1, n}$. Vrem să scriem $R_{i,j}^k$ în funcție de câteva R^{k-1}.

Pentru a determina toate drumurile de la i la j care trec prin stări cu indice cel mult k, putem începe mai întâi cu toate drumurile $R_{i,j}^{k-1}$, deoarece $k - 1 \leq k$. La acestea se adaugă toate drumurile de la i la j care trec și prin starea intermediară k.



$$R_{i,j}^k = R_{i,j}^{k-1} \cup (R_{i,k}^{k-1} \cdot (R_{k,k}^{k-1})^* \cdot R_{k,j}^{k-1})$$

Acum putem demonstra mult mai ușor că pentru orice mulțime de cuvinte $R^k_{i,j}$ există o expresie regulată $r^k_{i,j}$ astfel încât $L(r^k_{i,j}) = R^k_{i,j}$.

Demonstrație. Demonstrăm prin inducție după k.

Cazul de bază. Când $k = 0$, $R^0_{i,j}$ este o mulțime finită formată din toate literele pentru care există tranziții de la i la j, și eventual λ dacă $i = j$.

$r^0_{i,j}$ poate fi scris ca o disjuncție între aceste litere.

Ipoteza de inducție. Presupunem că pentru orice $R^k_{i,j}$ putem construi o expresie regulată $r^k_{i,j}$ cu $L(r^k_{i,j}) = R^k_{i,j}$.

Definim $r^{k+1}_{i,j} = r^k_{i,j} \cup (r^k_{i,k+1}(r^k_{k+1,k+1})^* r^k_{k+1,j})$.

Arătăm că această expresie regulată într-adevăr corespunde lui $R^k_{i,j}$:

$$
\begin{aligned}
L(r^{k+1}_{i,j}) &= L((r^k_{i,j} \cup (r^k_{i,k+1}(r^k_{k+1,k+1})^* r^k_{k+1,j})) && \text{definiția lui } r^{k+1}_{i,j} \\
&= L(r^k_{i,j}) \cup (L(r^k_{i,k+1})(L(r^k_{k+1,k+1}))^* L(r^k_{k+1,j})) \\
&= R^k_{i,j} \cup (R^k_{i,k+1}(R^k_{k+1,k+1})^* R^k_{k+1,j}) && \text{ipoteza de inducție} \\
&= R^k_{i,j} && \text{definiția lui } R^k_{i,j}
\end{aligned}
$$

□

Un cuvânt este acceptat de DFA dacă este acceptat:

- de expresia pentru cuvintele care merg din starea inițială în prima stare finală (r^n_{1,f_1})
- *sau* de expresia pentru cuvintele care merg din starea inițială în a doua stare finală (r^n_{1,f_2})
- etc.

Expresia finală pentru tot DFA-ul este:

$$E = r^n_{1,f_1} \cup r^n_{1,f_2} \cup \cdots \cup r^n_{1,f_{|F|}} = \bigcup_{f \in F'} r^n_{1,f}$$