En *regulær* grammatik $G = (V, T, S, P)$ består af

- V et alfabet (vocabulary).

- $T \subset V$ mængde af elementer kaldet terminalsymboler.
  $N = V \setminus T$ er mængden af nonterminaler.

- $S \in N$ et startsymbol.

- P mængden af produktioner.

En produktion er et udtryk på en af følgende former

- $A \to aB$
- $A \to a$
- $S \to \lambda$

hvor $A, B \in N$, $a \in T$ og λ er den tomme streng.

Hvis $A \to z$ er en produktion så skriver vi $\ell Ar \Rightarrow \ell zr$. Hvis $w_0 \Rightarrow w_1, w_1 \Rightarrow w_2, \ldots w_{n-1} \Rightarrow w_n$ så skriver vi $w_0 \Rightarrow^* w_n$ og kalder dette en derivation.

Når G er en regulær grammatik så vil enhver streng $w \in V^*$, der kan deriveres fra startsymbolet (altså $S \Rightarrow^* w$), være på formen

- en streng af terminalsymboler eller

- en streng af terminalsymboler efterfulgt af én non-terminal.

Sproget frembragt af G:

$$L(G) = \{w \in T^* \mid S \Rightarrow^* w\}.$$

$L(G)$ kaldes et regulært sprog når G er en regulær grammatik.

En nondeterministisk endelig automat $M = (S, I, f, s_0, F)$ består af

- S en endelig mængde af tilstande.
- I et inputalfabet.
- $f : S \times I \to P(S)$ en (transitions-) funktion. Hvis vi er i tilstand s og læser et symbol a fra inputstrengen så går vi til en tilstand i mængden $f(s, a)$.
- $s_0 \in S$ en starttilstand.

- $F \subseteq S$ mængden af sluttilstande.

$P(S)$ betegner mængden af alle delmængder af S.

En streng $w \in I^*$ accepteres af M hvis læsning af w *kan* føre M fra s_0 til en tilstand i F.
Sproget accepteret af M:

$L(M) =$ mængden af strenge $w \in I^*$ der accepteres af M.

En endelig automat er deterministisk hvis $|f(s, a)| = 1$ for alle $s \in S, a \in I$.

Sætning.
Hvis $G = (V, T, S_0, P)$ er en regulær grammatik så findes der en (nondeterministisk) endelig automat $M = (S, I, f, s_0, F)$ så $L(G) = L(M)$.

- $S = \{s_A \mid A \in N\} \cup \{s_F\}$
  hvor $N = V \setminus T$ og F nyt symbol.

- $I = T$

- $f(s_A, a) = \begin{cases} \{s_B \mid A \rightarrow aB\} \cup \{s_F\}, & \text{hvis} \quad A \rightarrow a \\ \{s_B \mid A \rightarrow aB\} & \text{ellers} \end{cases}$
  $f(s_F, a) = \emptyset$

- $s_0 = s_{S_0}$

- $F = \begin{cases} \{s_F, s_0\} & \text{hvis } S_0 \rightarrow \lambda \\ \{s_F\} & \text{ellers} \end{cases}$

Sætning.
Hvis $M_0 = (S, I, f_0, s_0, F_0)$ er en nondeterministisk endelig automat så findes der en (deterministisk) endelig automat $M_1 = (P(S), I, f_1, \{s_0\}, F_1)$ så $L(M_1) = L(M_0)$.

$f_1(S', a) = \bigcup_{s \in S'} f_0(s, a)$, for $S' \subseteq S, a \in I$.

$F_1 = \{S' \in P(S) \mid S' \cap F_0 \neq \emptyset\}$.

Tilstande, der ikke kan nås fra $\{s_0\}$ kan udelades.

Sætning.
Hvis $M = (S, I, f, s_0, F)$, $S = \{s_0, s_1, \ldots, s_n\}$ er en (deterministisk) endelig automat så findes der en regulær grammatik $G = (V, T, A_0, P)$ så $L(G) = L(M)$.

Kan antage: $f(s, a) \neq s_0$ for alle $s \in S, a \in I$.

- $T = I$

- $N = \{A_i \mid i = 0, 1, \ldots, n\}$ en nonterminaler for hver tilstand.
  $V = T \cup N$

- A_0 (svarende til starttilstand) er startsymbol

Der er følgende produktioner i P:

- Hvis $f(s_i, a) = s_j$ så er $A_i \to aA_j$ i P.
- Hvis $f(s_i, a) = s_j$ og $s_j \in F$ så er $A_i \to a$ i P.
- Hvis $s_0 \in F$ så er $A_0 \to \lambda$ i P.

I: et alfabet.
Regulære udtryk over I defines rekursivt ved:

- $\emptyset$ og λ er regulære udtryk.
  For alle $x \in I$ er $\mathbf{x}$ et regulært udtryk.

- Hvis $\mathbf{A}$ og $\mathbf{B}$ er regulære udtryk så er $(\mathbf{AB})$, $(\mathbf{A} \cup \mathbf{B})$ og $\mathbf{A}^*$ også regulære udtryk.

Et regulært udtryk repræsenterer en mængde (delmængde af I^*) som kaldes en regulær mængde.

Sætning
En delmængde $L \subseteq I^*$ er en regulær mængde hvis og kun hvis der findes en endelig automat M så $L = L(M)$.

Bemærk: beviserne er konstruktive.