
Reguliere expressies: Theorie II – deterministische eindige automaten

Huub de Beer

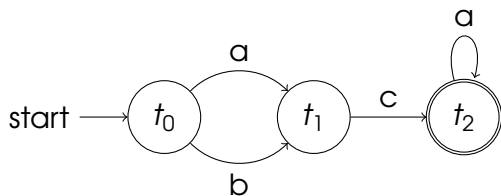
Eindhoven, 4 juni 2011

Reguliere expressies: generator van zinnen

- ▶ Een reguliere expressie geeft een taal aan
- ▶ Je kunt zinnen in die taal *genereren*

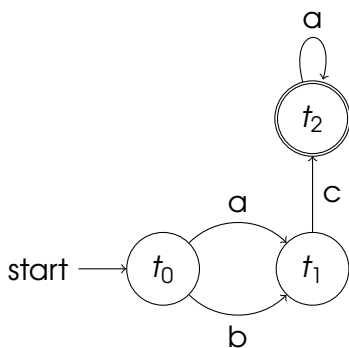
- ▶ Maar hoe *herken* je een zin in een taal?
- ▶ Wanneer voldoet een zin aan een reguliere expressie?
- ▶ Oplossing: deterministische eindige automaten
- ▶ Engels: *deterministic finite acceptor/automaton* (DFA)

Een DFA dat zinnen accepteert in de taal $\mathcal{L}((a|b) \cdot c \cdot a^*)$.



Onderdelen van een DFA

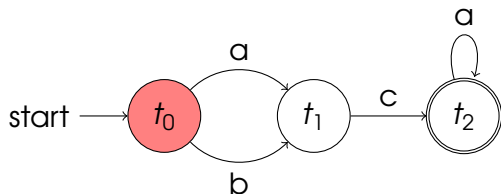
DFA



Onderdelen

- ▶ Toestanden (cirkels met naam)
 - ▶ Normaal
 - ▶ **1** Starttoestand: cirkel met binnenkomende pijl "start"
 - ▶ Eindtoestanden: dubbele cirkel (accepterende toestand)
- ▶ Transitie tussen toestanden: gelabelde pijl
- ▶ Transitie uit alfabet
- ▶ Maar een transitie per symbool per toestand (deterministisch)

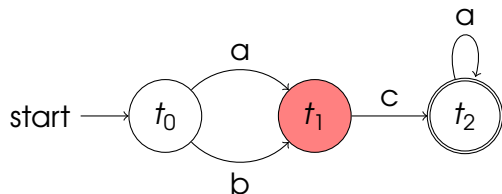
Herkennen zin: stap 1



Invoer: **a**caa

De automaat begint in de starttoestand en de het eerste symbool van de invoer wordt geïnspecteerd: een 'a'. Een 'a'-transitie is mogelijk: ga naar toestand t_1 .

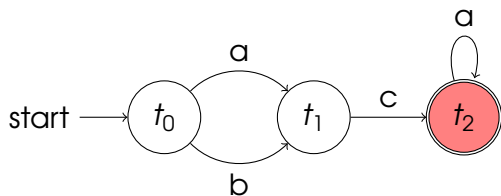
Herkennen zin: stap 2



Invoer: a**C**aa

Het volgende symbool van de invoer wordt geïnspecteerd: een 'c'. In toestand t_1 is er een 'c'-transitie mogelijk: ga naar toestand t_2 .

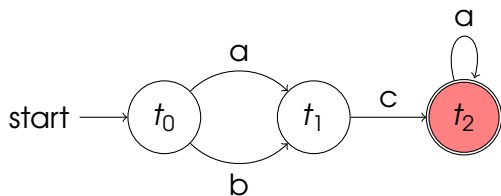
Herkennen zin: stap 3



Invoer: ac**a**a

Toestand t_2 is een eindtoestand: 'ac' is een zin in de taal. Maar er zijn nog meer symbolen in de invoer; het volgende symbool van de invoer wordt geïnspecteerd: een 'a'. In toestand t_2 is er een 'a'-transitie mogelijk: ga naar toestand t_2 .

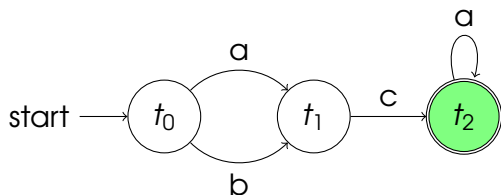
Herkennen zin: stap 4



Invoer: aca□

Toestand t_2 is een eindtoestand: 'aca' is een zin in de taal. Maar er zijn nog meer symbolen in de invoer; het volgende symbool van de invoer wordt geïnspecteerd: een 'a'. In toestand t_2 is er een 'a'-transitie mogelijk: ga naar toestand t_2 .

Herkennen zin: stap 5



Invoer: acaa

Toestand t_2 is een eindtoestand: 'acaa' is een zin in de taal. Er is geen volgend symbool: deze automaat accepteert de invoer 'acaa' als een zin in deze taal.