

## Aanvullingen cursus A&B

Academiejaar 2024-2025

Vincent Van Schependom

### Pagina 11, bewijs subalgebra:

**Stelling.** RegLan is een subalgebra van  $L_\Sigma$  voor de operaties *unie*, *concatenatie*, *Kleene\** en *complement*.

*Bewijs.* We bewijzen de stelling voor elke operatie apart:

- Unie:  
Zij  $E_1, E_2 \in \text{RegExp}$  de reguliere expressies die respectievelijk de talen  $L_{E_1}$  en  $L_{E_2}$  bepalen, met dus duidelijk  $L_{E_1}, L_{E_2} \in \text{RegLan}$ . Omdat de unie van beide talen wordt bepaald door een reguliere expressie, namelijk door  $(E_1|E_2)$ , geldt dat  $(L_{E_1} \cup L_{E_2}) \in \text{RegLan}$ . We besluiten dat de operatie *unie* inwendig is voor de subalgebra gevormd door RegLan.
- Concatenatie:  
Beschouw  $L_{E_1}$  en  $L_{E_2}$  zoals hierboven beschreven. Omdat de concatenatie van beide reguliere talen wordt bepaald door een reguliere expressie, namelijk door  $(E_1E_2)$ , geldt dat  $(L_{E_1}L_{E_2}) \in \text{RegLan}$ . We besluiten dat de operatie *concatenatie* inwendig is voor de subalgebra gevormd door RegLan.
- Kleene\*  
Beschouw  $L_{E_1}$  zoals hierboven beschreven. Omdat de Kleene\* van deze reguliere taal wordt bepaald door een reguliere , namelijk door  $(E_1)^*$ , geldt dat  $L_1^* \in \text{RegLan}$ . We besluiten dat de operatie *Kleene\** inwendig is voor de subalgebra gevormd door RegLan.
- Complement  
Beschouw de reguliere taal  $L_{E_1}$  zoals hierboven beschreven. Ze wordt bepaald door de reguliere expressie  $E_1$ . Omdat reguliere expressies en NFA's equivalent zijn, kunnen we een NFA  $N$  bouwen die dezelfde taal bepaalt als  $E_1$ . Elke NFA kan omgezet worden in een equivalente DFA, dus dat kunnen we ook hier doen. In de equivalente DFA  $D$  (die dus ook  $L_{E_1}$  bepaalt) maken we niet-aanvaarde toestanden van alle aanvaardende toestanden en vice versa. De bekomen DFA  $D'$  bepaalt nu het complement van  $L_{E_1}$ . We gaan vervolgens omgekeerd te werk: we bouwen een RE op vanuit  $D'$ , door eerst een GNFA te maken en die vervolgens te reduceren tot deze slechts 2 toestanden meer heeft. Tot slot lezen we de reguliere expressie af op de (unieke) boog tussen de start- en eindknoop. De GNFA bepaalt nog steeds  $\bar{L}_{E_1}$ , want deze taal werd ook door de DFA  $D'$  herkend en het procédé paste de taal niet aan. We hebben dus de RE gevonden die het complement van een willekeurige reguliere taal bepaalt. Dit wil precies zeggen dat  $\bar{L}_{E_1} \in \text{RegLan}$ , of nog: ook de operatie *complement* is inwendig voor de subalgebra gevormd door RegLan.

Alternatief: maak een generische product DFA die  $\bar{L} = \Sigma^* \setminus L$  bepaalt:

- DFA<sub>1</sub> is de DFA die  $\Sigma^*$  bepaalt: hij bevat 1 (aanvaardende) toestand, waar twee bogen toekomen: de startboog en de lus met daarop alle symbolen uit het alfabet. Deze laatste boog vertrekt natuurlijk ook uit die enige toestand.
- DFA<sub>2</sub> is de DFA die  $L$  bepaalt.
- $Q = Q_1 \times Q_2$
- $\delta(p \times q, x) = \delta_1(p, x) \times \delta_2(q, x) \iff \delta((p, q), x) = (\delta_1(p, x), \delta_2(q, x))$
- $q_s = (q_{s1}, q_{s2})$
- $F = F_1 \times (Q_2 \setminus F_2)$

□

## Pagina 15, zelf doen 4:

Opgave: Veronderstel dat  $L$  bepaald wordt door de NFA  $M$ , m.a.w. dat  $L = L_M$ . We construeren een NFA  $M_2 = (Q_2, \Sigma, \delta_2, q_{s2}, F_2)$  die de omgekeerde taal  $L^R = \{w^R \mid w \in L\}$  van  $L$  bepaalt.

We bouwen hiervoor eerst een NFA  $M_1 = (Q_1, \Sigma, \delta_1, q_{s1}, F_1)$  die dezelfde taal bepaalt als de NFA  $M$ , maar die slechts één eindtoestand heeft:

- $Q_1 = Q \cup \{q_e\}$
- Overgangsfunctie:  $\delta_1(q, a) = \delta(q, a) \quad \forall q \in Q, \forall a \in \Sigma$   
 $\delta_1(q, \varepsilon) = q_e \quad \forall q \in F$
- $q_{s1} = q_s$
- $F_1 = \{q_e\}$

Deze NFA  $M_1$  vormen we nu om naar een NFA  $M_2$ , zodat  $M_2$  de omgekeerde taal  $L^R$  bepaalt:

- $Q_2 = Q_1$
- Draai alle bogen om:  $\delta_2(q, a) = \{p \mid q \in \delta_1(p, a)\} \quad \forall p \in Q_2, \forall a \in \Sigma_\varepsilon$
- $q_{s2} = q_e$

De bekomen NFA  $M_2$  bepaalt de omgekeerde taal van  $L = L_M$ . Merk op dat – wegens het feit dat deze taal  $L^R$  door een NFA wordt bepaald – dit een reguliere taal is.

## Pagina 18-20, de algebra van NFA's

Gegeven  $NFA_1 = (Q_1, \Sigma, \delta_1, q_{s1}, \{q_{f1}\})$  en  $NFA_2 = (Q_2, \Sigma, \delta_2, q_{s2}, \{q_{f2}\})$ .

De concatenatie  $NFA_1 NFA_2$  is de NFA  $= (Q, \Sigma, \delta, q_s, \{q_f\})$  waarbij

- $Q = Q_1 \cup Q_2$
- $q_s = q_{s1}$
- $F = \{q_{f2}\}$
- $\delta$  gedefinieerd als:

$$\begin{aligned} \delta(q_{f1}, x) &= \emptyset & \forall x \in \Sigma \\ \delta(q_{f1}, \varepsilon) &= q_{s2} \\ \delta(q, x) &= \delta_1(q, x) & \forall q \in Q_1 \setminus \{q_{f1}\}, \forall x \in \Sigma_\varepsilon \\ \delta(q, x) &= \delta_2(q, x) & \forall q \in Q_2, \forall x \in \Sigma_\varepsilon \end{aligned}$$

Hierbij moet de eerste regel eigenlijk niet expliciet worden vermeld. We zijn hier bezig met NFA's, dus als er geen overgangsregel voor een bepaald symbool  $x$  gedefinieerd is, wordt er vanuit gegaan dat  $\delta(q, x) = \emptyset$ .

De ster  $(NFA_1)^*$  is de NFA  $= (Q, \Sigma, \delta, q_s, \{q_f\})$  waarbij

- $Q = Q_1 \cup \{q_s, q_f\}$
- $F = \{q_f\}$
- $\delta$  gedefinieerd als:

$$\begin{aligned} \delta(q_s, x) &= \emptyset & \forall x \in \Sigma \\ \delta(q_s, \varepsilon) &= \{q_{s1}, q_{f1}\} \\ \delta(q_{f1}, \varepsilon) &= \{q_s, q_f\} \\ \delta(q_{f1}, x) &= \emptyset & \forall x \in \Sigma \\ \delta(q, x) &= \delta_1(q, x) & \forall q \in Q_1 \setminus \{q_{f1}\}, \forall x \in \Sigma_\varepsilon \end{aligned}$$

## Pagina 21, bewijs structurele inductie

**Stelling.** Onderstaande constructie bewaart de taal, t.t.z.  $L_{\text{NFA}_E} = L_E$

- $\text{NFA}_{E_1 E_2} = \text{concat}(\text{NFA}_{E_1}, \text{NFA}_{E_2})$
- $\text{NFA}_{E_1^*} = \text{ster}(\text{NFA}_{E_1})$
- $\text{NFA}_{E_1 | E_2} = \text{unie}(\text{NFA}_{E_1}, \text{NFA}_{E_2})$

*Bewijs.* We bewijzen eerst volgende hulpstellingen:

- De concatenatie van  $\text{NFA}_1$  en  $\text{NFA}_2$  bepaalt  $L_{\text{NFA}_1} L_{\text{NFA}_2}$ :

We voeren volgende notatie in:

$$\text{NFA } C = \text{concat}(\text{NFA}_1, \text{NFA}_2)$$

Volgens de definitie van de concatenatie van twee talen geldt dat

$$L_{\text{NFA}_1} L_{\text{NFA}_2} = \{xy \mid x \in L_{\text{NFA}_1}, y \in L_{\text{NFA}_2}\}$$

We moeten bewijzen dat

$$s \in L_C \Leftrightarrow s \in L_{\text{NFA}_1} L_{\text{NFA}_2}$$

$\Rightarrow$  Neem aan dat  $s \in L_C$ . Bij het parsen van deze string  $s$  met de machine  $C$  zullen we op een gegeven moment gegarandeerd in de toestand  $q_{f1}$  terechtkomen, aangezien dat de enige toestand is van waaruit we naar de machine  $\text{NFA}_2$  kunnen geraken. Dit gebeurt door een  $\varepsilon$ -boog te nemen naar  $q_{s2}$ . Noem de string die geparst is tijdens deze eerste fase  $x$  en neem de  $\varepsilon$ -boog van  $q_{f1}$  naar  $q_{s2}$ . Er blijft – vanuit deze starttoestand van  $\text{NFA}_2$  – een string  $y$  over. Na het parsen van deze string  $y$  komen we in de toestand  $q_{f2}$  terecht, want  $s \in L_C$  en de enige ( $\varepsilon$ -)boog naar  $q_f$  vertrekt vanuit deze toestand. Omdat  $x \in L_{\text{NFA}_1}$  (na het parsen van  $x$  belanden we in een aanvaardende toestand  $q_{f1}$  van  $\text{NFA}_1$ ) en  $y \in L_{\text{NFA}_2}$  (analoog), geldt dat  $s = xy \in L_{\text{NFA}_1} L_{\text{NFA}_2}$ .

$\Leftarrow$  Als de string  $s \in L_{\text{NFA}_1} L_{\text{NFA}_2}$ , dan bestaat  $s$  uit twee substrings, zodat  $s = xy$  met  $x \in L_{\text{NFA}_1}$  en  $y \in L_{\text{NFA}_2}$ . Dat wil zeggen dat bij het doorlopen van  $C$ , we vanuit  $q_s$  in een eindig aantal stappen in  $q_{f1}$  terechtkomen. Van hieruit nemen we een  $\varepsilon$ -boog naar de begintoestand van  $\text{NFA}_2$ . Vervolgens bereiken we na nog een eindig aantal extra stappen de toestand  $q_{f2}$ , van waaruit we een  $\varepsilon$ -boog nemen naar de aanvaardende toestand  $q_f$ . Hiermee hebben we aangetoond dat de string  $s$  wordt aanvaard door  $\text{NFA } C$ , m.a.w.  $s \in L_C$ .

- De ster van  $\text{NFA}_1$  bepaalt  $L_{\text{NFA}_1}^*$ :

We voeren volgende notatie in:

$$\text{NFA } S = \text{ster}(\text{NFA}_1)$$

De Kleene-ster van een taal  $L$  is de unie van alle talen  $L^n$  die ontstaan wanneer we deze taal  $n$  keer concateneren met zichzelf ( $n \in \mathbb{N}$ ). Per definitie geldt dat  $\varepsilon \in L^*$ , want er geldt dat  $L^0 = \{\varepsilon\}$ . We moeten bewijzen dat

$$s \in L_S \Leftrightarrow s \in L_{\text{NFA}_1}^*$$

Voor elk accepterend pad in  $L_S$ , zijn de enige bogen die karakters uit  $\Sigma$  bevatten de bogen uit  $\text{NFA}_1$ . Bovendien: voor elke toestand  $q$  in  $\text{NFA}_1$ , gaat elk pad van deze toestand  $q$  naar de toestand  $q_f$  door  $q_{f1}$ . Met andere woorden: de enige strings die in  $L_S$  zitten zijn  $\varepsilon$  en  $x_1, x_2, x_3, \dots$  (met  $x_i \in L_{\text{NFA}_1}$ ). Dit zijn precies die strings uit  $L_{\text{NFA}_1}^*$ .

- De unie van  $\text{NFA}_1$  en  $\text{NFA}_2$  bepaalt  $L_{\text{NFA}_1} \cup L_{\text{NFA}_2}$ :

We voeren volgende notatie in:

$$\text{NFA } U = \text{unie}(\text{NFA}_1, \text{NFA}_2)$$

Volgens de definitie van de unie van talen geldt dat

$$L_{\text{NFA}_1} \cup L_{\text{NFA}_2} = \{s \mid s \in L_{\text{NFA}_1} \vee s \in L_{\text{NFA}_2}\}$$

We moeten bewijzen dat

$$s \in L_U \Leftrightarrow s \in L_{\text{NFA}_1} \cup L_{\text{NFA}_2}$$

- $\Rightarrow$  Neem aan dat  $s \in L_U$ . Als we deze string parsen met de machine  $C$ , maken we in het begin de keuze om vanuit  $q_s$  de  $\varepsilon$ -boog te nemen naar ofwel  $q_{s1}$ , ofwel  $q_{s2}$ . We veronderstellen het eerste geval, namelijk de keuze voor de starttoestand van  $\text{NFA}_1$ , het andere geval verloopt analoog. Bij het parsen van de  $s$  belanden we uiteindelijk in  $q_f$ , want dit is een aanvaarde string. Het bereiken van die toestand kan enkel met een  $\varepsilon$ -boog vanuit  $q_{f1}$  of  $q_{f2}$ . Aangezien we in het begin gekozen hebben voor  $q_{s1}$  (en dus ook voor  $\text{NFA}_1$ ), kan dat enkel vanuit  $q_{f1}$  gebeurd zijn. Het bereiken van  $q_{f1}$ <sup>1</sup> wil precies zeggen dat  $s \in L_{\text{NFA}_1}$  en dus ook  $c \in L_{\text{NFA}_1} \cup L_{\text{NFA}_2}$ . We kunnen, zoals gezegd, hetzelfde aantonen voor de keuze van  $\text{NFA}_2$  in het begin.
- $\Leftarrow$  Als de string  $s \in L_{\text{NFA}_1} \cup L_{\text{NFA}_2}$ , dan geldt dat ofwel  $s \in L_{\text{NFA}_1}$ , ofwel  $s \in L_{\text{NFA}_2}$ . Veronderstel het eerste geval. Dan kunnen we bij het parsen van  $s$  aan de hand van de machine  $U$  de  $\varepsilon$ -boog naar  $q_{s1}$  nemen, waarna we de string  $s$  helemaal parsen, tot we in  $q_{f1}$  terechtkomen. Van hieruit kunnen we de  $\varepsilon$ -boog naar  $q_f$  nemen. We vinden dus dat  $s$  wordt aanvaard door  $U$  en dus dat  $s \in L_U$ . Het andere geval (namelijk dat  $s \in L_{\text{NFA}_2}$ ) loopt nu volledig analoog.

We bewijzen nu de oorspronkelijke stelling aan de hand van structurele inductie:

- Basisstap: We bewijzen dat de stelling geldt voor volgende basisgevallen:
  - Als  $E = \varepsilon$ , dan is  $L_E = \{\varepsilon\}$ . Kijkend naar de constructie van de NFA voor dit basisgeval op pagina 21, zien we duidelijk dat deze NFA dezelfde taal bepaalt als  $E$  en dus geldt dat  $L_{\text{NFA}_E} = L_E = \{\varepsilon\}$ .
  - Als  $E = \phi$ , dan is  $L_E = \emptyset$ . Kijkend naar de constructie van de NFA voor dit basisgeval op pagina 21, zien we duidelijk dat deze NFA dezelfde taal bepaalt als  $E$  en dus geldt dat  $L_{\text{NFA}_E} = L_E = \emptyset$ .
  - Als  $E = a \in \Sigma$ , dan is  $L_E = \{a\}$ . Kijkend naar de constructie van de NFA voor dit basisgeval op pagina 21, zien we duidelijk dat deze NFA dezelfde taal bepaalt als  $E$  en dus geldt dat  $L_{\text{NFA}_E} = L_E = \{a\}$ .
- Inductiestap: neem aan dat de stelling geldt voor reguliere expressies  $E_1$  en  $E_2$ :

$$L_{\text{NFA}_{E_1}} = L_{E_1}, \quad L_{\text{NFA}_{E_2}} = L_{E_2}$$

Dan bewijzen we dat de stelling ook geldt ( $L_{\text{NFA}_E} = L_E$ ) voor de ster van  $E_1$ , alsook voor de unie en concatenatie van beide RE's:

- Concatenatie: De operatie wordt als volgt beschreven:

$$\text{NFA}_{E_1 E_2} = \text{concat}(\text{NFA}_{E_1}, \text{NFA}_{E_2})$$

Uit bovenstaande hulpstelling voor de concatenatie volgt dat deze NFA de concatenatie bepaalt van de talen bepaald door  $\text{NFA}_{E_1}$  en  $\text{NFA}_{E_2}$ . Verder gebruiken we ook de inductiehypothese:

$$L_{\text{NFA}_{E_1 E_2}} \stackrel{\text{hulpstelling}}{=} L_{\text{NFA}_{E_1}} L_{\text{NFA}_{E_2}} \stackrel{\text{IH}}{=} L_{E_1} L_{E_2}$$

Omdat volgens de definitie van *de taal bepaald door een RE* geldt dat  $L_{E_1} L_{E_2} = L_{E_1 E_2}$ , volgt het te bewijzen nu direct:  $L_{\text{NFA}_{E_1 E_2}} = L_{E_1 E_2}$

<sup>1</sup>Hiermee wordt bedoeld dat de toestand bereikt wordt zonder dat er nog symbolen overschieten in  $s$  die nog geparst moeten worden.

- Ster: Het operatie wordt als volgt beschreven:

$$\text{NFA}_{E_1^*} = \text{ster}(\text{NFA}_{E_1})$$

Uit bovenstaande hulpstelling voor de ster volgt dat deze NFA de taal  $L_{\text{NFA}_{E_1^*}}^* \stackrel{\text{IH}}{=} L_{E_1}^*$  bepaalt. Dat wil precies zeggen dat  $L_{\text{NFA}_{E_1^*}} = L_{E_1}^* = L_{E_1^*}^2$ , zoals bewezen moest worden.

- Unie: De operatie wordt als volgt beschreven:

$$\text{NFA}_{E_1|E_2} = \text{unie}(\text{NFA}_{E_1}, \text{NFA}_{E_2})$$

Uit bovenstaande hulpstelling voor de unie volgt dat deze NFA de taal

$$L_{\text{NFA}_{E_1}} \cup L_{\text{NFA}_{E_2}} \stackrel{\text{IH}}{=} L_{E_1} \cup L_{E_2}$$

bepaalt. Dat wil precies zeggen dat  $L_{\text{NFA}_{E_1|E_2}} = L_{E_1} \cup L_{E_2} = L_{E_1|E_2}^3$ , zoals bewezen moest worden.

□

## Pagina 26

**Stelling.** De omzetting van NFA naar GNFA wijzigt de verzameling aanvaarde strings niet.

*Bewijs.*

- Stel dat de NFA dezelfde taal  $L_E$  bepaalt als een reguliere expressie  $E$ . Een nieuwe begintoestand toevoegen met een  $\varepsilon$  boog naar de oude starttoestand van de NFA, staat gelijk aan de expressie  $\varepsilon E$ , dewelke gelijk is aan  $E$ .
- Stel dat de NFA dezelfde taal  $L_E$  bepaalt als een reguliere expressie  $E$ . Een nieuwe eindtoestand toevoegen met een  $\varepsilon$  bogen van de oude eindtoestand van de NFA, staat gelijk aan de expressie  $E\varepsilon$ , dewelke gelijk is aan  $E$ .
- Het toevoegen van de extra bogen om de GNFA te vervolledigen, wijzigt de verzameling aanvaarde strings niet. Je kan zo'n  $\phi$ -bogen wel volgen, maar als je zo een boog volgt, zal geen enkele string behoren tot de taal bepaald door die reguliere expressie.
- Indien we  $n$  parallel gerichte bogen met labels  $a_i \in \Sigma$  ( $i \in \{1, \dots, n\}$ ) samennemen als een unie van die labels, dan verandert de verzameling aanvaarde strings niet. In de nieuw gevormde reguliere expressie  $a_1|a_2|\dots|a_n$  moeten we immers een keuze maken bestaande uit één symbool, hetgeen equivalent is met het kiezen van één boog in de DFA. De keuze van zulke boog maakt niet uit, aangezien ze allemaal naar dezelfde toestand leiden.

□

<sup>2</sup>Gebruik hier ook de definitie van een taal bepaald door een reguliere expressie.

<sup>3</sup>Idem

## Pagina 28, bewijs DFA

**Stelling.**  $(Q', \Sigma, \delta', q'_s, F')$  is een DFA equivalent met de NFA  $(Q, \Sigma, \delta, q_s, F)$

*Bewijs.* Uit de constructie op pagina 27-28 volgt duidelijk dat de geconstrueerde automaat een DFA is:

- Er zijn geen  $\varepsilon$ -bogen, want  $\delta'(S, a)$  is enkel gedefinieerd voor  $a \in \Sigma$  (en  $S \in Q'$ ), m.a.w.  $a \neq \varepsilon$
- De functie  $\delta' : Q' \times \Sigma \rightarrow Q'$  is een totale functie: ze is overal goed gedefinieerd.
  - $\delta'(S, a)$  is gesloten onder  $\varepsilon$ -bogen. Stel immers dat  $p \in \delta'(S, a)$ . Dit wil zeggen dat  $\exists p_{-1} \in S : p_{-1} \xrightarrow{a} p$ . Stel nu dat er voor een  $q \in Q$  geldt dat  $p \xrightarrow{\varepsilon} q_1 \xrightarrow{\varepsilon} \dots \xrightarrow{\varepsilon} q_n$  met  $n \in \mathbb{N}$  (we mogen willekeurig veel  $\varepsilon$ -bogen nemen). Dan is duidelijk dat ook  $p_{-1} \xrightarrow{a} q_n$ , want  $a$  is een  $\varepsilon$ -compressie van  $a\varepsilon^n$ . En dus geldt ook dat  $q_n \in \delta'(S, a)$ . Dit wil precies zeggen dat  $\delta'(S, a)$  gesloten is onder  $\varepsilon$ -bogen.
  - Stel dat  $S_w = \{q \mid q_s \xrightarrow{w} q\}$ .  $S_{wa}$  is de verzameling toestanden die we bereiken door uit toestanden  $p \in S_w$  één  $a$ -boog te volgen gevolgd door een willekeurig aantal  $\varepsilon$ -bogen. Het is dan duidelijk dat  $\delta'(S_w, a) = S_{wa}$ .

Wat betreft de equivalentie, moeten we verifiëren dat

$$\forall w \in \Sigma^* : q'_s \xrightarrow{w} F' \text{ (in de DFA)} \iff q_s \xrightarrow{w} F \text{ (in de NFA)}$$

We bewijzen beide richtingen.

$\Rightarrow$  Deze implicatie volgt uit iets algemeners dat we nu zullen bewijzen: zij  $S$  een deelverzameling van  $Q$  gesloten onder  $\varepsilon$ -bogen. Dan geldt

$$\forall w \in \Sigma^* : q'_s \xrightarrow{w} S \text{ (in de DFA)} \implies \forall p \in S : q_s \xrightarrow{w} p \text{ (in de NFA)}$$

Dit bewijzen we per inductie op de lengte van  $w$ .

- Basisstap: Als  $|w| = 0$ , dan geldt dat  $w = \varepsilon =$  de lege string. Neem aan dat  $q'_s \xrightarrow{w} S$ . Dan is  $S = q'_s = \{q_s, \text{toestanden bereikbaar vanuit } q_s \text{ met } \varepsilon\}$ . We zien nu duidelijk in dat de implicatie geldt, want elke toestand in  $S = q'_s$  is evident bereikbaar vanuit  $q_s$  met  $\varepsilon$ .
- Inductiehypothese: veronderstel dat de stelling geldt voor alle strings  $w$  van hoogstens lengte  $|w| = n$ .
- Inductiestap: Beschouw een string  $w' = wa$  (met  $a \in \Sigma$ ) van lengte  $n + 1$ .

We willen aantonen dat als  $q'_s \xrightarrow{wa} S$ , dan geldt  $\forall p \in S : q_s \xrightarrow{wa} p$ . Zij  $S_{-1}$  de toestand in de DFA zodat  $q'_s \xrightarrow{w} S_{-1}$ . Wegens de inductiehypothese geldt nu

$$\forall p \in S_{-1} : q_s \xrightarrow{w} p$$

We kunnen in de DFA in  $S$  geraken door een pijl met label  $a$  te volgen vanuit toestand  $S_{-1}$ . Dit betekent precies dat  $S$  de verzameling is van alle toestanden die we in de NFA kunnen bereiken door vanuit een toestand in  $S_{-1}$  een pijl te nemen met een label  $a$  erop, gevolgd door eventueel een aantal  $\varepsilon$ -bogen. En dus geldt voor iedere  $p \in S$  dat  $q_s \xrightarrow{wa=w'} p$ , hetgeen we wilden bewijzen.

Nu volgt de rechterimplicatie

$$q'_s \xrightarrow{w} F' \xrightarrow{\text{def. } F'} \exists S \in F' : q'_s \xrightarrow{w} S \xrightarrow{\text{hierboven}} \forall p \in S : q_s \xrightarrow{w} p \xrightarrow{S \in F'} \exists p \in S : p \in F$$

Voor die laatste  $p$  geldt dus ook dat  $q_s \xrightarrow{w} p$  en dus  $q_s \xrightarrow{w} F$ , q.e.d..

$\Leftarrow$  Als  $q_s \xrightarrow{w} F$ , dan bestaat er een  $p \in F$  zodat  $q_s \xrightarrow{w} p$ . Er bestaat in de NFA dus een accepterend pad  $q_s, q_1, q_2, \dots, q_n, p$ . Voor een toestand  $q$  in de NFA, zij  $S(q)$  de grootste verzameling die  $q$  bevat en gesloten is onder  $\varepsilon$ -bogen. Nu is  $S(q_s), S(q_1), S(q_2), \dots, S(q_n), S(p)$  een accepterend pad in de DFA. En dus geldt dat  $q'_s \xrightarrow{w} S(p)$  – en dus ook dat  $q'_s \xrightarrow{w} F'$ , want het beschreven pad met verzamelingen  $S(q)$ , die gesloten zijn onder  $\varepsilon$ -bogen, is een accepterend pad.

□

## Pagina 30

De overgangsfunctie

$$\delta^* : Q \times \Sigma^* \rightarrow Q$$

van een DFA, is een functie die een koppel  $(q, s)$  afbeeldt op de unieke toestand  $p$  zodat  $q \xrightarrow{s} p$ . We kunnen deze functie ook inductief definiëren, en wel als volgt:

1.  $\delta^*(q, \varepsilon) = q$
2.  $\delta^*(q, aw) = \delta^*(\delta(q, a), w) \quad \forall a \in \Sigma, w \in \Sigma^*$

**Stelling.** In een DFA geldt dat

$$\delta^*(q, wa) = \delta(\delta^*(q, w), a) \quad \text{voor } a \in \Sigma, w \in \Sigma_\varepsilon^*$$

*Bewijs.* We bewijzen dit per inductie op de lengte van  $w$ .

- Basisstap: ingeval de lengte van  $w$  gelijk is aan 0, geldt dat  $w = \varepsilon$ . In dat geval geldt dat

$$\begin{aligned} \delta^*(q, wa) &= \delta^*(q, \varepsilon a) & w &= \varepsilon \\ &= \delta^*(\delta(q, \varepsilon), a) & (2) \\ &= \delta^*(q, a) \\ &= \delta(q, a) & |a| &= 1 \\ &= \delta(\delta^*(q, \varepsilon), a) & (1) \\ &= \delta(\delta^*(q, w), a) & w &= \varepsilon \end{aligned}$$

- Inductiehypothese: stel dat de stelling geldt voor strings  $w$  van hoogstens lengte  $|w| = n$ , m.a.w. dat voor zulke strings geldt dat  $\delta^*(q, wa) = \delta(\delta^*(q, w), a)$  voor  $a \in \Sigma, w \in \Sigma_\varepsilon^*$ .
- Inductiestap: we bewijzen dat de stelling ook geldt voor strings van lengte  $n + 1$ . Zo'n string  $w'$  kunnen we schrijven als  $w' = bw$  met  $|w| = n, b \in \Sigma$ . Nu geldt dat

$$\begin{aligned} \delta^*(q, w'a) &= \delta^*(q, bwa) & w' &= sw \\ &= \delta^*(\delta(q, b), wa) & (1) \\ &= \delta(\delta^*(\delta(q, b), w), a) & \text{inductiehypothese} \\ &= \delta(\delta^*(q, bw), a) & (2) \text{VRNL} \\ &= \delta(\delta^*(q, w'), a) & w' &= bw \end{aligned}$$

□

## Pagina 34, bewijs $\text{DFA}_{\min}$

**Stelling.**  $\text{DFA}_{\min}$  is een unieke DFA, equivalent met DFA, en alle toestanden zijn f-verschillend.

*Bewijs.*  $\text{DFA}_{\min}$  is een DFA:

- Er zijn geen  $\varepsilon$ -bogen
- 2 verschillende bogen met hetzelfde symbool vanuit  $p$  en  $q$  versmelten wanneer de twee toestanden zelf versmelten door f-gelijkheid: stel namelijk dat  $p$  en  $q$  f-gelijk zijn. Dan zijn ook  $p' = \delta(p, a)$  en  $q' = \delta(q, a)$  f-gelijk. We bewijzen dat.

De f-strings van  $p$  en  $q$  zijn gelijk, dus ook hun f-strings van de vorm  $as$ . De f-strings van  $p'$  zijn de strings  $s$  zodat  $as$  een f-string is van  $p$ . Hetzelfde geldt voor  $q'$ . Bijgevolg hebben  $p'$  en  $q'$  dezelfde f-strings en zijn ze f-gelijk.

De equivalentie van DFA en  $\text{DFA}_{\min}$  bewijzen we door per inductie aan te tonen dat

$$w \text{ is een f-string van } Q_i \text{ (in } \text{DFA}_{\min}) \iff w \text{ is een f-string van alle } q \in Q_i \text{ (in } \text{DFA}_{\text{origineel}})$$

- Basisstap: als de lengte van de string  $w$  gelijk is aan 0, geldt dat  $w = \varepsilon$ . Nu geldt dat

$$\begin{aligned} \varepsilon \text{ is een f-string van } Q_i &\iff Q_i \in \tilde{F} \\ &\iff Q_i \subseteq F && \text{definitie } \tilde{F} \\ &\iff \forall q \in Q_i : q \in F \\ &\iff \varepsilon \text{ is een f-string van alle } q \in Q_i \end{aligned}$$

- Inductiehypothese: stel dat de stelling geldt voor strings  $w$  van hoogstens lengte  $|w| = n$ .
- Inductiestap: beschouw de string  $w' = bw$ . We tonen aan dat de stelling ook geldt voor deze string van lengte  $|w'| = n + 1$ , m.a.w. we tonen aan dat

$$w' = bw \text{ is een f-string van } Q_i \iff w' = bw \text{ is een f-string van alle } q \in Q_i$$

Er geldt dat

$$\begin{aligned} bw \text{ is een f-string van } Q_i &\iff \tilde{\delta}^*(Q_i, bw) \in \tilde{F} \\ &\iff \tilde{\delta}^*(\tilde{\delta}(Q_i, b), w) \in \tilde{F} && (\text{ind. definitie } \tilde{\delta}^*) \\ &\iff w \text{ is een f-string van } \tilde{\delta}(Q_i, b) \\ &\iff w \text{ is een f-string van alle } q \in \tilde{\delta}(Q_i, b) && (\text{inductiehypothese}) \\ &\iff bw \text{ is een f-string van alle } q \in Q_i \end{aligned}$$

Zie p45 voor het bewijs van de uniciteit.

□



## Pagina 43, equivalentie $MN(L)$ en DFA

**Stelling.** De overgangen van DFA naar een  $MN(L)$ -relatie ( $op_1$ ), en van de  $MN(L)$ -relatie naar een DFA ( $op_2$ ), zijn elkaars inversen – op DFA-isomorfisme na.

We kunnen de stelling ook anders formuleren. Hiervoor beschouwen we twee operaties:

- $op_1(\text{DFA})$  levert als output een  $MN(L)$ -relatie  $\sim_D$ .
- $op_2(MN(L))$  levert ons als output een DFA.

**Stelling.** Voor elke DFA  $D$  geldt dat de DFA  $D' = op_2(op_1(\text{DFA}))$  isomorf is met  $D$ .

*Bewijs.* We bewijzen eerst dat  $op_1$  en  $op_2$  effectief steeds de gewenste output hebben:

1. Elke DFA  $D$  bepaalt een  $MN(L_D)$  (equivalentie)relatie  $\sim_D$  op  $\Sigma^*$ .

Definieer voor elke toestand volgende deelverzameling van  $\Sigma^*$ :

$$\text{reach}(q) = \{w \in \Sigma^* \mid \delta^*(q_s, w) = q\}$$

De verzameling met als elementen al deze verzamelingen  $\text{reach}(q_i)$  vormt een partitie van  $\Sigma^*$ :

- Er bestaat geen  $\text{reach}(q) = \emptyset$ . Elke string  $s \in \Sigma^*$  zit namelijk in een of andere  $\text{reach}(q)$ . De overgangsfunctie  $\delta$  van een DFA is totaal, dus bij het parsen van  $s$  kunnen voor elk symbool een boog volgen in de DFA, zodat we uiteindelijk in een of andere toestand terechtkomen. De string behoort dan precies tot de  $\text{reach}(q)$  van deze toestand.
- De  $\text{reach}(q)$ 's zijn disjunct. Een string  $s \in \Sigma^*$  kan namelijk niet in twee  $\text{reach}(q)$ 's zitten, want we hebben in een DFA nooit een keuze naar welke toestand we zullen overgaan: er zijn nooit twee verschillende bogen met eenzelfde symbool. Bij het parsen van  $s$  belanden we dus in een unieke toestand  $q$  en bijgevolg geldt dat  $s \in \text{reach}(q)$ .
- De unie van alle  $\text{reach}(q)$ 's is precies  $\Sigma^*$ .

Omdat partities equivalentierelaties induceren en vice versa, kunnen we dus ook de geïnduceerde equivalentierelatie  $\sim_D$  beschouwen:

$$x \sim_D y \iff x \text{ en } y \text{ behoren tot dezelfde } \text{reach}(q) \iff \delta^*(q_s, x) = \delta^*(q_s, y)$$

We tonen nu aan dat deze equivalentierelatie  $\sim_D$  een  $MN(L)$ -relatie is. Herinner: **een equivalentierelatie  $\sim$  tussen strings is een Myhill-Nerode relatie voor  $L$  als  $\sim$  voldoet aan 3 voorwaarden.** We checken deze 3 voorwaarden nu voor  $\sim_D$ :

- (a) De partitie is eindig. Inderdaad: DFA's hebben een eindig aantal toestanden en bijgevolg zijn er dus ook een eindig aantal  $\text{reach}(q)$ 's.
- (b) Rechtscongruentie: we willen aantonen dat

$$x \sim_D y \implies xa \sim_D ya$$

Stel dat  $x \sim_D y$ . Dan geldt volgens de definitie van onze equivalentierelatie  $\sim_D$  dat beide strings behoren tot  $\text{reach}(q)$  voor een  $q \in Q$ , of nog dat  $\delta^*(q_s, x) = \delta^*(q_s, y) = q$ . Vanuit deze toestand  $q$  hebben we voor elk symbool  $a \in \Sigma$  slechts één keuze met betrekking tot de boog die we nemen om over te gaan naar een nieuwe toestand. Noem deze nieuwe toestand  $q' = \delta(q, a)$ . We hebben nu met de strings  $xa$  en  $ya$  dezelfde toestand  $q'$  bereikt, wat precies wil zeggen dat  $xa \sim_D ya$ .

- (c)  $\sim_D$  verfijnt de partitie  $\{L, \bar{L}\}$ . We willen aantonen dat

$$x \sim_D y \implies (x \in L \iff y \in L)$$

Stel dat  $x \sim_D y$ . Als  $x \in L$ , dan wil dat zeggen dat  $\delta^*(q_s, x) \in F$ . Omdat  $x \sim_D y$  geldt dat  $\delta^*(q_s, x) = \delta^*(q_s, y)$  en dus geldt ook dat  $y \in L$ . We kunnen de andere richting analoog bewijzen.

## 2. Elke $MN(L)$ -relatie $\sim$ op $\Sigma^*$ bepaalt een DFA $D$ zodat $L = L_D$ :

Gegeven een taal  $L \in L_\Sigma$ . We construeren de DFA  $(Q, \Sigma, \delta, q_s, F)$  als volgt:

- $Q = \{x_\sim \mid x \in \Sigma^*\}$
- $q_s = \varepsilon_\sim$
- $F = \{x_\sim \mid x \in L\}$
- $\delta(x_\sim, a) = (xa)_\sim$

Dit is inderdaad een DFA:

- $Q$  en  $F$  hebben slechts een eindig aantal toestanden, omdat een  $MN(L)$ -relatie geassocieerd is met een eindige partitie. Er zijn dus slechts een eindig aantal equivalentieklassen.
- De overgangsfunctie  $\delta$  is goed gedefinieerd.

Als  $y, z \in \Sigma^*$  tot dezelfde equivalentieklasse  $x_\sim$  behoren, dan behoren ze tot eenzelfde toestand  $q \in Q$ . Na het volgen van een boog met een symbool  $a \in \Sigma$  vanuit deze toestanden, moeten we in de DFA voor beide strings in een eenzelfde nieuwe toestand  $q'$  terechtkomen, anders zouden er meerdere bogen met dat symbool  $a$  bestaan.

We bewijzen dat we effectief in die toestand  $q'$  terechtkomen voor beide strings. Omdat volgens de  $MN(L)$ -relatie op de strings in  $\Sigma^*$  de rechtscongruentie  $y \sim z \Rightarrow ya \sim za$  geldt, behoren de strings  $ya$  en  $za$  tot dezelfde equivalentieklasse  $(xa)_\sim$ . De definitie  $\delta(x_\sim, a) = (xa)_\sim$  is dus goed.

Tot slot bewijzen we nog dat de DFA de gegeven taal  $L \in L_\Sigma$  effectief bepaalt, of nog dat  $L_{DFA} = L$ :

$$x \in L_{DFA} \stackrel{\Delta}{=} \delta^*(\varepsilon_\sim, x) \in F \iff x_\sim \in F \stackrel{\Delta}{=} x \in L$$

We bewijzen de overgang door per inductie op de lengte van  $x$  aan te tonen dat

$$\delta^*(\varepsilon_\sim, x) = x_\sim$$

- Basisstap: als  $|x| = 0$ , is  $x = \varepsilon$  en geldt per definitie van  $\delta^*$  dat  $\delta^*(\varepsilon_\sim, x) = \varepsilon_\sim$
- Inductiehypothese: stel dat de stelling geldt voor strings  $x$  van lengte  $|x| = n$
- Inductiestap: we bewijzen dat de stelling ook geldt voor strings van lengte  $n + 1$ . Zo'n string  $x'$  kunnen we schrijven als  $x' = xa$  met  $|x| = n, a \in \Sigma$ . Nu geldt dat

$$\begin{aligned} \delta^*(\varepsilon_\sim, x') &= \delta^*(\varepsilon_\sim, xa) & x' &= xa \\ &= \delta(\delta^*(\varepsilon_\sim, x), a) & &\text{eigenschap } \delta^* \\ &= \delta(x_\sim, a) & \delta^*(\varepsilon_\sim, x) &\stackrel{\text{inductiehypothese}}{=} x_\sim \\ &= (xa)_\sim & &\text{definitie } \delta \\ &= (x')_\sim & \text{q.e.d.} & \quad x' = xa \end{aligned}$$

We willen nu bewijzen dat voor iedere DFA  $D$  de DFA  $D' = \text{op}_2(\text{op}_1(D))$  isomorf is met  $D$ . We stellen daarvoor een bijctie op tussen de toestanden van  $D$  en de equivalentieklassen van de bijhorende  $\text{MN}(L_D)$ -relatie:

$$b_1 : Q_D \rightarrow \{x_{\sim_D} \mid x \in \Sigma^*\} : q \mapsto \text{reach}(q)$$

Dan bestaat er omgekeerd dus ook een bijctie

$$b_2 : \{x_{\sim_D} \mid x \in \Sigma^*\} \rightarrow Q_{D'} : x_{\sim_D} \mapsto q' \mid x \in \text{reach}(q')$$

tussen de equivalentieklassen van de  $\text{MN}(L_D)$ -relatie en de toestanden van de DFA  $D'$ . We kunnen nu ook de bijctie

$$b = (b_2 \circ b_1) : Q_D \rightarrow Q_{D'}$$

tussen de toestanden van  $D$  en de toestanden van  $D'$  beschouwen die ontstaat wanneer we de voorgaande twee bijcties samenstellen. We tonen aan dat  $b$  effectief voldoet aan de eigenschappen voor isomorfisme:

- $b(F_D) = F_{D'}$ : er geldt dat  $b(F_D) = b_2(b_1(F_D))$ . De bijctie  $b_1$  mapt elke toestand  $q_f \in F_D$  naar een equivalentieklasse in  $\text{MN}(L_D)$  zodat elke string in die equivalentieklasse tot  $L_D$  behoort. De bijctie  $b_2$  mapt elke equivalentieklasse die strings uit  $L_D$  bevat naar een aanvaardende eindtoestand in  $D'$ . Dus elke aanvaardende eindtoestand uit  $D$  wordt door  $b$  gemapt op een aanvaardende eindtoestand uit  $D'$ . (Met dezelfde redenering zien we dat ook het omgekeerde geldt).
- $b(q_s) = q_{s'}$ : er geldt dat  $b(q_s) = b_2(b_1(q_s))$ . De starttoestand  $q_s$  wordt gemapt op  $\varepsilon_{\sim}$  door  $b_1$  en  $\varepsilon_{\sim}$  wordt gemapt op  $q'_s$  door  $b_2$ , dus  $q_s$  wordt gemapt op  $q'_s$  door  $b$ .
- $b(\delta(q, a)) = \delta'(b(q), a)$ : Voor elke toestand  $q \in Q$  en elk symbool  $a \in \Sigma$  worden  $q$  en  $\delta(q, a)$  gemapt door  $b_1$  op twee equivalentieklassen in  $\text{MN}(L_D)$  zodat alle strings uit de tweede equivalentieklasse kunnen verkregen worden door een  $a$  te zetten achter een string uit de eerste equivalentieklasse. Vervolgens mapt  $b_2$  deze twee equivalentieklassen op twee toestanden in  $D'$  zodat er een  $a$ -overgang is van de eerste toestand naar de tweede toestand. Dit betekent dat  $b(\delta(q, a)) = \delta'(b(q), a)$ .

□

## Pagina 74

**Stelling.** De constructie van een PDA uit een CFG  $G$ , zoals hieronder beschreven, levert een PDA die de taal  $L_G$  accepteert. Met andere woorden: een contextvrije taal wordt bepaald door een PDA.

*Bewijs.*

□