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

- $\text{NFA}_{E_1E_2} = \text{concat}(\text{NFA}_{E_1}, \text{NFA}_{E_2})$
- $\text{NFA}_{E_1^*} = \text{ster}(\text{NFA}_{E_1})$
- $\text{NFA}_{E_1|E_2} = \text{unie}(\text{NFA}_{E_1}, \text{NFA}_{E_2})$

Ofwel: talen bepaald door een reguliere expressie, worden herkend door een NFA (1. RE $\to$ NFA)

Bewijs. We bewijzen eerst volgende hulpstellingen:

- De concatenatie van NFA_1 en 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 NFA_2 kunnen geraken. Dit gebeurt door een ε-boog te nemen naar q_{s2}. Noem de string die geparst is tijdens deze eerste fase x en neem de ε-boog van q_{f1} naar q_{s2}. Er blijft – vanuit deze starttoestand van 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 (ε-)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 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 ε-boog naar de begintoestand van NFA_2. Vervolgens bereiken we na nog een eindig aantal extra stappen de toestand q_{f2}, van waaruit we een ε-boog nemen naar de aanvaardende toestand q_f. Hiermee hebben we aangetoond dat de string s wordt aanvaard door NFA C, m.a.w. $s \in L_C$.

- De ster van 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 Σ bevatten de bogen uit NFA_1. Bovendien: voor elke toestand q in 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 ε en $x_1, x_2, x_3, ...$ (met $x_i \in L_{\text{NFA}_1}$). Dit zijn precies die strings uit $L^*_{\text{NFA}_1}$.

- De unie van NFA_1 en 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 ε-boog te nemen naar ofwel q_{s1}, ofwel q_{s2}. We veronderstellen het eerste geval, namelijk de keuze voor de starttoestand van 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 ε-boog vanuit q_{f1} of q_{f2}. Aangezien we in het begin gekozen hebben voor q_{s1} (en dus ook voor NFA$_1$), kan dat enkel vanuit q_{f1} gebeurd zijn. Het bereiken van q_{f1}[1] 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 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 ε-boog naar q_{s1} nemen, waarna we de string s helemaal parsen, tot we in q_{f1} terechtkomen. Van hieruit kunnen we de ε-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:

    $$\text{NFA}_{E_1E_2} = \text{concat}(\text{NFA}_{E_1}, \text{NFA}_{E_2})$$

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

    $$L_{\text{NFA}_{E_1E_2}} \overset{\text{hulpstelling}}{=} L_{\text{NFA}_{E_1}} L_{\text{NFA}_{E_2}} \overset{\text{IH}}{=} L_{E_1} L_{E_2}$$

    Omdat volgens de definitie van *de taal bepaald door een RE* geldt dat $L_{E_1} L_{E_2} = L_{E_1E_2}$, volgt het te bewijzen nu direct: $L_{\text{NFA}_{E_1E_2}} = L_{E_1E_2}$

[1] 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:

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

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

– Unie: De operatie wordt als volgt beschreven:

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

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

$$L_{\mathrm{NFA}_{E_1}} \cup L_{\mathrm{NFA}_{E_2}} \overset{\mathrm{IH}}{=} L_{E_1} \cup L_{E_2}$$

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

□

[2]Gebruik hier ook de definitie van een taal bepaald door een reguliere expressie.
[3]Idem