

## SAMENVATTING VAN

# AUTOMATEN EN BEREKENBAARHEID



W. Van Onsem  
KU Leuven  
Department Computerwetenschappen  
Declaratieve Talen en Artificialle Intelligentie (DTAI)

Disclaimer: Er is geen garantie dat de volledige inhoud gebruikt wordt of correct wordt voorgesteld.

### Orakel (§3.17)

$O^L$

**Orakelmachine:** Een Orakelmachine  $O^L$  is een TM met een orakel voor de taal  $L$ . Men kan dit voorstellen als een TM met een extra band (de orakel-tape) en drie extra toestanden  $q_{or}$ ,  $q_{ye}$  en  $q_{no}$ . De TM raadpleegt het orakel door de string op de orakel-band te schrijven en in toestand  $q_{or}$  te gaan. Het orakel brengt wanneer de string tot  $L$  behoort de machine in toestand  $q_{ye}$  en anders in toestand  $q_{no}$ .  $L$  is mogelijk onbeslisbaar.

Reducties

$$L_1 \leq_T L_2$$

**Turing-reductie van talen:** Een taal  $L_1$  is Turingreducerbaar naar taal  $L_2$  indien  $L_1$  beslisbaar is relatief ten opzichte van  $L_2$ , dit wil zeggen: er bestaat een orakelmachine  $O^{L_2}$  die  $L_1$  beslist. We schrijven  $L_1 \leq_T L_2$ .

Toepassingen:  
• Filosofie  
• Complexiteitstheorie



### Turing (§3.1-3.9)

TM

**Turingmachine:** Een TM is een 7-tal  $(Q, \Sigma, \Gamma, \delta, q_0, q_{acc}, q_{rej})$  met:  
•  $Q$  een eindige verzameling toestanden is;  
•  $\Sigma$  een alfabet dat niet  $\emptyset$  bevat;  
•  $\Gamma$  een tape alfabet met  $\# \in \Gamma$  en  $\Sigma \subseteq \Gamma$ ;  
•  $q_0 \in Q$  de start-toestand,  $q_{acc} \in Q$  de accepterende eindtoestand,  $q_{rej} \in Q$  de verwerpende eindtoestand met  $q_{acc} \neq q_{rej}$ ;  
•  $\delta: Q \times \Gamma \rightarrow Q \times \Gamma \times \Gamma \times \{L, S, R\}$  de totale transitiefunctie.

$s \in L_{TM}$

**String door een TM geaccepteerd:** De TM wordt geïnitieerd in toestand  $q_0$  met de invoerstring op de tape, op de overige plaatsen staat een  $\#$ , en de leeskop op het meest linkse teken. Zolang de TM niet in  $q_{acc}$  en  $q_{rej}$  zit, wordt het karakter uitgelezen onder de leeskop. Op basis van de transitiefunctie  $\delta$  wordt een karakter onder de leeskop gezet, de leeskop naar links/rechts verplaatst en komt de machine in een nieuwe toestand. Als de TM na een eindig aantal stappen in  $q_{acc}$  komt, wordt de string aanvaard.

$s \in \infty TM$

**String waarvoor de TM niet stopt:** Een string  $s$  behoort tot  $\infty TM$  indien bij invoer van  $s$  de TM nooit in  $q_{acc}$  of  $q_{rej}$  terecht komt.

Herk

**Herkennen:** Een Turingmachine TM herkent  $TM$  indien  $s \in L_{TM}$  voor alle  $s$  waarvoor  $s \in L$ . Een TM bestaat noemt men een herkennbare taal.

Besl

**Beslissen:** Een TM beslist een taal  $L$  indien de TM  $L$  herkent en bovendien  $\infty TM = \emptyset$ . Een taal  $L$  waarvoor zo'n TM bestaat noemt men een beslisbare taal.

Enumereerbaar

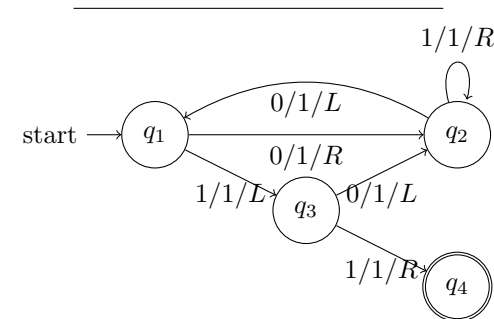
EM

**Enumeratormachine:** Een EM is een TM met volgende uitbreidingen:  
• Een enumeratortoestand  $q_e$ ;  
• Een uitvoerband met uitvoermarker;  
• De transitiefunctie  $\delta: Q \times \Gamma \rightarrow Q \times \Gamma \times \Gamma \times \{L, S, R\}$ . De machine start met lege band en lege uitvoerband, in de  $q_e$  toestand en werkt als een TM. Wanneer iets op de uitvoer wordt geschreven verschuift de schrijfkop naar rechts. Wanneer de machine in toestand  $q_e$  komt, wordt op de uitvoer de uitvoermarker geschreven en komt de EM opnieuw in  $q_e$  terecht.

Herk  $\equiv$  EM

**Enumereren equivalent met herkennen:** Een taal genummerd door een EM is herkennbaar en elke herkennbare taal wordt door een EM genummerd.

turingmachine



Rice

**Stelling van Rice:** Voor elke niet-triviale taal-invariantie eigenschap  $P$  van Turingmachines geldt  $L_P = \{ \langle M \rangle \mid M \text{ voldoet aan } P \}$  is niet beslisbaar.

Turing-berekenbaar

**Turing-berekenbare functie:** Een functie  $f$  is Turing-berekenbaar indien er een Turingmachine bestaat die bij input  $s$  uiteindelijk stopt met  $f(s)$  op de band.

$$L_1 \leq_m L_2$$

**Reductie van talen:** Een taal  $L_1$  (over  $\Sigma_1$ ) is reducerbaar naar een taal  $L_2$  (over  $\Sigma_2$ ) indien er een afbeelding  $f: \Sigma_1^* \rightarrow \Sigma_2^*$  bestaat zodat  $f(L_1) \subseteq L_2$  en  $f(L_1) \cap L_2^c = \emptyset$ . We noteren dat door  $L_1 \leq_m L_2$ .

$\Sigma(n)$

**Besize bever:** De besize bever is een totale functie  $\Sigma: \mathbb{N} \rightarrow \mathbb{N}$  die met invoer  $n$ , het maximale aantal 1'en uitgeschreven door een Turingmachine met  $n$  toestanden en een initieel lege invoer associeert.

Toepassingen:  
• Programmatalen  
• Levens redden (WO-II)



### Context-sensitief (§2.26)

CSG

**Context-sensitieve grammatica over een alfabet  $\Sigma$ :** Een CSG over  $\Sigma$  is een 4-tal  $(V, \Sigma, R, S)$  met:  
•  $V$  een eindige verzameling niet-eindigheidsymbolen;  
•  $\Sigma$  een eindig alfabet van eindigheidsymbolen disjunct met  $V$ ;  
•  $R$  een eindige verzameling regels een regel is een koppelpaar  $\alpha A \beta \rightarrow \alpha \gamma \beta$  met  $\alpha, \beta, \gamma \in (\Sigma \cup V)^*$  en  $A \in V$ ;  
•  $S \in V$  is het startsymbool.

CSL

**Context-sensitieve taal:** Een taal  $L$  is context-sensitief indien er een CSG bestaat zodat  $L = L_{CSG}$ . De verzameling van context-sensitieve talen duiden we aan met CSL.

LBA

**Lineair begrensde automaat:** Een LBA is een TM die niet best of eindig buiten het deel van de band dat initieel invoer bevat.

CSG  $\equiv$  LBA

**Equivalentie van CSG en LBA:** Elke LBA bepaalt een context-sensitieve taal en elke context-sensitieve taal wordt bepaald door een LBA.

$s \in L_{CSG}$  wordt beslist in

$$O(c^n)$$

Toepassingen:  
• Natuurlijke taalverwerking  
• Netwerken  
• Gegevensbanken



### Context-vrij (§2.19-2.25)

CFG

**Contextvrije grammatica over een alfabet  $\Sigma$ :**  $(V, \Sigma, R, S)$  is een CFG over  $\Sigma$  indien:  
•  $V$  een eindige verzameling niet-eindigheidsymbolen is;  
•  $\Sigma$  een eindig alfabet van eindigheidsymbolen disjunct met  $V$ ;  
•  $R$  een eindige verzameling regels; een regel is een koppelpaar niet-eindigheidsymbolen en een string van elementen uit  $V \cup \Sigma$ ; we schrijven de twee delen van zulk een koppelpaar met een  $\rightarrow$  ertussen;  
•  $S \in V$  is het startsymbool.

CFL

**Contextvrije taal:** Een taal  $L$  is contextvrij indien er een CFG bestaat zodat  $L = L_{CFG}$ . De verzameling van contextvrije talen duiden we aan met CFL.

$$b \Rightarrow_{CFG}^* f$$

**Afbeelding met behulp van een contextvrije grammatica:** Gegeven een CFG  $(V, \Sigma, R, S)$ . Een string  $f$  over  $V \cup \Sigma$  wordt afgeleid uit een string  $b$  uit  $V \cup \Sigma$  met behulp van de CFG indien er een eindige rij strings  $s_0, s_1, \dots, s_n$  bestaat zodat:  
•  $s_0 = b \wedge s_n = f$ ;  
•  $s_{i+1}$  verkregen wordt uit  $s_i$  (voor  $i < n$ ) door in  $s_i$  een niet-eindigheidsymbolen  $X$  te vervangen door de rechterkant van een regel waarin  $X$  links voorkomt.  
We noteren  $s_i \Rightarrow_{CFG} s_{i+1}$  en  $b \Rightarrow_{CFG}^* f$ .

$$CFG \sim L_{CFG}$$

**Een contextvrije grammatica bepaalt een taal:** De taal  $L_{CFG}$  bepaald door een CFG  $(V, \Sigma, R, S)$  is gedefinieerd als:  $L_{CFG} = \{ s \in \Sigma^* \mid s \Rightarrow_{CFG}^* s \}$ .

PDA

**Push-down automaat:** Een PDA is een 5-tal  $(Q, \Sigma, \Gamma, \delta, q_0, F)$  met:  
•  $Q$  een eindige verzameling toestanden;  
•  $\Sigma$  een eindig inputalfabet;  
•  $\Gamma$  een eindig stapelalfabet;  
•  $\delta: Q \times \Sigma \times \Gamma \rightarrow P(Q \times \Gamma \times \{L, S, R\})$  de overgangsfunctie;  
•  $q_0 \in Q$  de starttoestand;  
•  $F \subseteq Q$  de verzameling eindtoestanden.

$s \in L_{PDA}$

**String aanvaard door een PDA:** Een string  $s$  wordt aanvaard door een PDA  $(Q, \Sigma, \Gamma, \delta, q_0, F)$  indien  $s$  kan worden opgesplitst in delen  $u_1, u_2, \dots, u_n$  ( $u_i \in \Sigma^*$ ) en toestanden  $q_1, q_2, \dots, q_n$  ( $q_i \in Q$ ) zijn, en stapels  $s_1, s_2, \dots, s_n$  ( $s_i \in \Gamma^*$ ) zodat:  
•  $s = u_1 u_2 \dots u_n$ ;  
•  $(q_{i-1}, u_i) \in \delta(q_{i-1}, u_{i-1}, s_i)$  waarbij  $x, y \in \Gamma^*$ , en  $s_{i+1} = xs_i$ ,  $s_{i+1} = ys_i$  of  $s_{i+1} = ys_i$  met  $t \in \Gamma^*$ .

CFG  $\equiv$  PDA

**Equivalentie van CFG en PDA:** Elke PDA bepaalt een contextvrije taal en elke contextvrije taal wordt bepaald door een PDA.

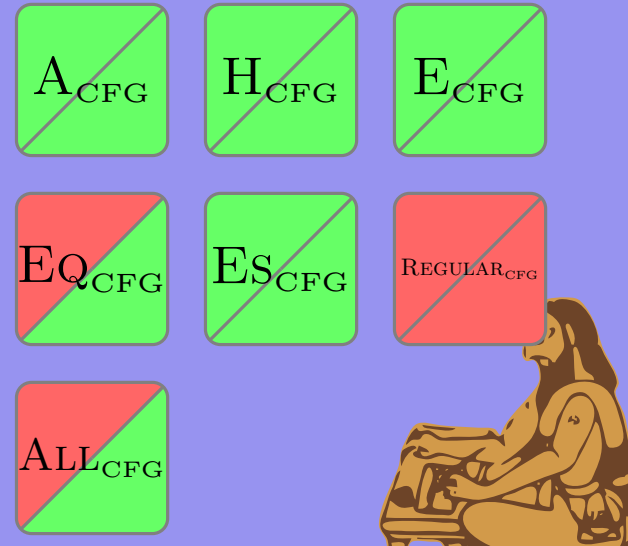
$L \in CFL$

**Pompen lemma voor CFL:** Voor een contextvrije taal  $L$  bestaat er een pomplengte  $p$  zodat elke string  $s \in L$  met  $|s| \geq p$  kan opgesplitst worden in  $s = xyz$  met  $|xy| \leq p$  en  $xy^iz \in L$  voor alle  $i \geq 0$ .  
•  $|xy| \geq 1$ ;  
•  $|xy| \leq p$ ;  
•  $|xy| \geq 1$ .

$s \in L_{CFG}$  wordt beslist in

$$O(n^3)$$

Toepassingen:  
• Extraheren van informatie  
• Compilers  
• Natuurlijke taalverwerking



### Regulier (§2.4-2.18)

$RE \sim L_{RE}$

RE

**Reguliere expressie over een alfabet  $\Sigma$ :**  $E$  is een RE over  $\Sigma$  indien  $E$  van de vorm is:  
•  $\emptyset$ ;  
•  $a$  met  $a \in \Sigma$ ;  
•  $(E_1 E_2)$  met  $E_1$  en  $E_2$  RE's zijn over  $\Sigma$ ;  
•  $(E_1 | E_2)$  met  $E_1$  en  $E_2$  RE's zijn over  $\Sigma$ ;  
•  $(E_1)^*$  met  $E_1$  een RE is over  $\Sigma$ , dan is  $L_E = L_{E_1} \cup L_{E_2}$ ;  
•  $(E_1)^+$  met  $E_1$  een RE is over  $\Sigma$ , dan is  $L_E = L_{E_1}$ .

RegLang

**Reguliere taal:** Een taal die door een reguliere expressie bepaald wordt is een reguliere taal. De verzameling van reguliere talen duiden we aan met RegLang.

NFA

**Niet-deterministische eindige toestandsautomat:** Een NFA is een 5-tal  $(Q, \Sigma, \delta, q_0, F)$  met:  
•  $Q$  een eindige verzameling toestanden;  
•  $\Sigma$  een eindig alfabet;  
•  $\delta: Q \times \Sigma \rightarrow P(Q)$  de overgangsfunctie van de automaat;  
•  $q_0 \in Q$  de starttoestand;  
•  $F \subseteq Q$  de verzameling eindtoestanden.

$s \in L_{NFA}$

**String aanvaard door een NFA:** Een string  $s$  wordt aanvaard door een NFA  $(Q, \Sigma, \delta, q_0, F)$  indien  $s$  kan geschreven worden als  $a_1 a_2 \dots a_n$  met  $a_i \in \Sigma$ , en er een rij toestanden  $t_0, t_1, \dots, t_n$  bestaat zodat:  
•  $t_0 = q_0$ ;  
•  $t_{i+1} \in \delta(t_i, a_{i+1})$ ;  
•  $t_n \in F$ .

NFA  $\equiv$  NFA

**Equivalentie van twee NFA's:** Twee NFA's worden equivalent genoemd als ze dezelfde taal besleiden.

GNFA

**Generaliseerde niet-deterministische eindige toestandsautomat:** Een GNFA is een NFA  $(Q, \Sigma, \delta, q_0, F)$  met:  
• slechts één eindtoestand  $F = \{q_f\}$  zodat  $q_f \neq q_0$ ;  
• de leden bevatten een reguliere expressie als label;  
•  $\forall q_i, q_j \in Q, (q_i, q_j) : \exists e \in \text{RegExp}(\Sigma) : \delta(q_i, e) = q_j$ ;  
•  $\forall q_i \in Q, (q_i, q_i) : \exists e \in \text{RegExp}(\Sigma) : \delta(q_i, e) = q_i$ ;  
•  $\forall q_i \in Q, (q_i, q_f) : \exists e \in \text{RegExp}(\Sigma) : \delta(q_i, e) = q_f$ .

Reduce

**Reductie-stap:** Neem een toestand  $x \in Q \setminus \{q_0, q_f\}$  en toestanden  $a, b \in Q$  en voer volgende reductie uit:

DFA

**Deterministische eindige toestandsautomat:** Een DFA is een 5-tal  $(Q, \Sigma, \delta, q_0, F)$  met:  
•  $Q$  een eindige verzameling toestanden;  
•  $\Sigma$  een eindig alfabet;  
•  $\delta: Q \times \Sigma \rightarrow Q$  de overgangsfunctie van de automaat;  
•  $q_0 \in Q$  de starttoestand;  
•  $F \subseteq Q$  de verzameling eindtoestanden.

NFA  $\rightarrow$  DFA

**NFA naar DFA:** De gelijkenheden geven voor elke NFA  $(Q_1, \Sigma, \delta_1, q_{01}, F_1)$  een equivalente DFA  $(Q_2, \Sigma, \delta_2, q_{02}, F_2)$ :  
 $Q_2 = P(Q_1)$   
 $q_{02} = \{q_{01}\}$   
 $\delta_2((q, a), b) = \bigcup_{q' \in \delta_1(q, a)} \{q'\}$

DFA  $\cong$  DFA

**Isomorf DFA:** Twee DFA's  $(Q_1, \Sigma, \delta_1, q_{01}, F_1)$  en  $(Q_2, \Sigma, \delta_2, q_{02}, F_2)$  zijn isomorf indien er een bijectie  $b: Q_1 \rightarrow Q_2$  bestaat met:  
•  $b(F_1) = F_2 \wedge b(q_{01}) = q_{02}$ ;  
•  $\forall q \in Q_1, \forall a \in \Sigma: b(\delta_1(q, a)) = \delta_2(b(q), a)$ .

$f \sim$

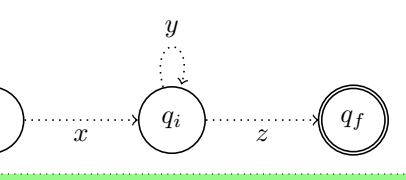
**$f$ -gelijk:** Twee toestanden  $q_1, q_2 \in Q$  in een DFA zijn  $f$ -gelijk indien:  
 $\forall w \in \Sigma^*: \delta^*(q_1, w) \in F \leftrightarrow \delta^*(q_2, w) \in F$   
Anders zijn ze  $f$ -verschillend.

Min DFA

**Minimale DFA:** Gegeven een DFA zonder onbereikbare toestanden en waarbij elke twee verschillende toestanden,  $f$ -verschillend zijn, hiervoor bestaat er geen DFA met strikt minder toestanden die dezelfde taal besleiden.

$L \in \text{RegLang}$

**Pompen lemma voor RegLang:** Voor een reguliere taal  $L$  bestaat een pomplengte  $p$  zodat elke string  $s \in L$  met  $|s| \geq p$  kan opgesplitst worden in  $s = xyz$  met  $|xy| \leq p$  en  $xy^iz \in L$  voor alle  $i \geq 0$ .  
•  $|xy| \geq 1$ ;  
•  $|xy| \leq p$ ;  
•  $|xy| \geq 1$ .



Myhill-Nerode relaties

$\sim_{mn}$

**Myhill-Nerode relatie:** Een Myhill-Nerode relatie over een taal  $L$  is een equivalentierelatie  $\sim_{mn}$  die aan volgende eigenschappen voldoet:  
•  $\forall x, y \in \Sigma^*, a \in \Sigma: x \sim_{mn} y \Rightarrow xa \sim_{mn} ya$ ;  
•  $\forall x, y \in \Sigma^*, x \sim_{mn} y \Rightarrow x \in L \leftrightarrow y \in L$ ;  
•  $\forall x, y \in \Sigma^*, x \sim_{mn} y \Rightarrow |x| = |y|$ ;  
•  $\sim_{mn}$  heeft een eindig aantal equivalentieklassen.

$\sim_{mn} \rightarrow$  DFA

**MN relatie naar DFA:** Gegeven een taal  $L$  over een alfabet  $\Sigma$  en een MN-relatie  $\sim_{mn}$  over  $\Sigma^*$ . Dan is  $(Q, \Sigma, \delta, q_0, F)$  een DFA die  $L$  besleiden met:  $Q = \{x, [x] \in \Sigma^*\}$ ,  $q_0 = [x_0]$ ,  $F = \{[x], x \in L\}$  en  $\delta([x, a], a) = ([xa])$ .

$P \leq P$

**Finere partitie:** Een partitie  $P_1$  is finer dan  $P_2$  indien elk element van  $P_1$  vervat zit in een element van  $P_2$ .

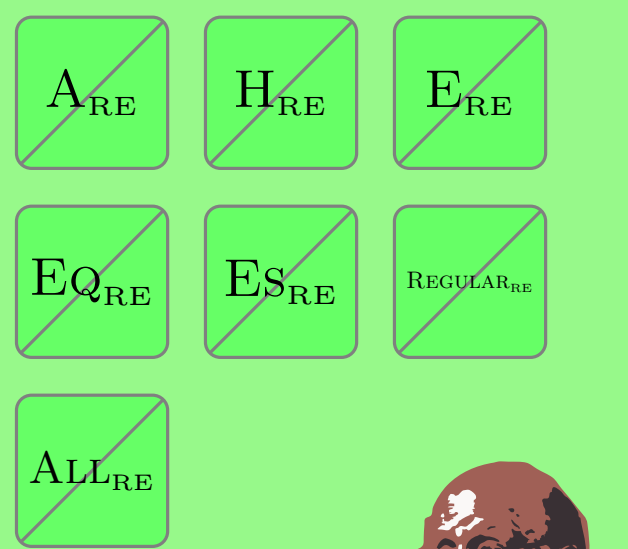
DFA  $\equiv \sim_{mn}$

**MN relaties tegenover DFA:** De overgangen van DFA naar de MN relatie, en van de MN relatie naar een DFA zijn elkaars inverse - op isomorfie na.

$s \in L_{RE}$  wordt beslist in

$$O(n)$$

Toepassingen:  
• Extraheren van informatie  
• Compilers  
• Internetprotocollen  
• Digitale elektronica



### Beschrijven

### Automaten

### Representatiekracht

### Tijdscomplexiteit Berekenbaarheid

### Herschrijfsystemen ( $\lambda$ -calculus; §4)

$A \sim B$

**Equivalentie termen:** Twee termen  $A$  en  $B$  zijn equivalent indien er een sequentie van herschrijfsregels bestaat zodat  $A \rightarrow^* B$  en  $B \rightarrow^* A$ .

$f x$

**Currying:**  $f$  x kiezen we als  $f$  toegepast op argument  $x$ . Elke functie in  $\lambda$ -calculus heeft maar één argument, maar resulteert soms in een nieuwe functie.

$(\lambda x. x) x$

**Vrij voorkomen:** Een voorkomen van de variabele  $x$  is vrij in de expressie  $E$  indien:  
•  $E$  gelijk is aan  $x$ ;  
•  $E$  gelijk is aan  $(\lambda y. A)$  en het voorkomen van  $x$  is vrij in  $A$  of  $B$ ;  
•  $E$  gelijk is aan  $(\lambda y. A)$  en  $y \neq x$ .

CR-1

**Church-Rosser I:** Gegeven twee expressies  $E_1$  en  $E_2$ . Indien  $E_1 \rightarrow^* E_2$ , dan bestaat er een expressie  $E$  zodat  $E_1 \rightarrow^* E$  en  $E_2 \rightarrow^* E$ . Bijgevolg kan geen expressie geconverteerd worden naar twee verschillende normalisatievormen.

CR-2

**Church-Rosser II:** Gegeven twee expressies  $E$  en  $N$  met  $N$  in normalisatievorm. Dan bestaat er een reductie  $E \rightarrow^* N$  in normalisatievorm. De normalisatievorm bepaalt dat de meest linkse benulste reductie eerst moet worden gereduceerd.

$\rightarrow^\alpha$

**Alfa-conversie:** Herschrijven een variabele in een  $\lambda$ -expressie. Formeel:  
 $\lambda x. E \rightarrow^\alpha \lambda y. E'$   
In  $E'$  worden alle vrije voorkomen van  $x$  vervangen door  $y$  en  $y$  is geen variabele gebonden in  $E$ .

$\rightarrow^\beta$

**Beta-reductie:** Herschrijven naar een kopie van het lijschaam van de  $\lambda$ -expressie waarin de vrije voorkomen van de parameter vervangen zijn door het argument. Formeel:  
 $(\lambda x. E) y \rightarrow^\beta E'$

$\leftarrow^\beta$

**Beta-abstractie:** Herschrijven een term naar een functie die wordt toegepast op een argument.  $\beta$ -abstractie wordt gebruikt om de equivalentie tussen termen aan te tonen. Formeel:  
 $(\lambda x. E) x \rightarrow^\beta E$

$\mathbb{N}$

**Natuurlijke getallen:** Een natuurlijk getal  $n \in \mathbb{N}$  wordt meestal voorgesteld als  $c_n = \lambda f. f^n. x$ . Volgende operaties zijn gedefinieerd:  
• optelling:  $A_1 + A_2 = \lambda x. p. p. x$  ( $p$  is  $\lambda p. p. q$ )  
• versnigvolgving:  $A_1 \times A_2 = \lambda x. y. z. (y. z)$   
• exponent:  $A_1^A_2 = \lambda x. y. y. x$

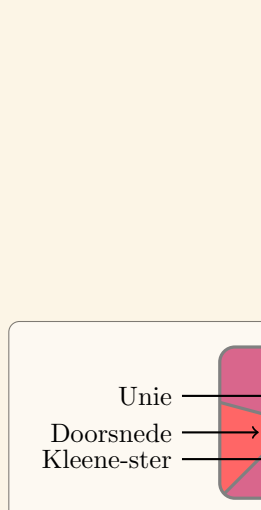
$\mathbb{B}$

**True, False en IF:** Booleaanse waarden  $b \in \mathbb{B}$  worden meestal voorgesteld als volgt:  
• **true** =  $\lambda x. y. x$   
• **false** =  $\lambda x. y. y$   
• **if** =  $\lambda x. y. z. (x. y. z)$



### Legende

Inwendige operator?



### Taal-aspecten (§2.1-2.3)

$\Sigma$

**Alfabet:** Een alfabet  $\Sigma$  is een set karakters die relevant zijn in een gegeven taal  $L$ .

$s \in \Sigma^*$

**String:** Een string over een alfabet