Formele talen

12

uitgebreid

strings

Een *alfabet* is een eindige, niet-lege, verzameling *letters*.

Σ = { a,b,c }
B = { 0,1 }
C = { а,б,в,г,д,е,ж,з,и,й,к,л, … э,ю,я }
P = { <u>if</u>, <u>else</u>, <u>while</u>, <u>do</u>, … }
V = { … , appel, koek, ei , … }

strings

Σ alfabet.
Een *string/woord* (over Σ) is een eindig geordend rijtje letters uit Σ.

ab, abca, abcbabcba
0, 1, 00, 01, 10, 11, 000, 001, …
Приключения, Астерикса
“de appel valt niet ver”

Σ^*, lege string λ, lengte $|x|$

$|abcbac| = 6$ $\quad |\lambda| = 0$ $\quad a^6b^3 = aaaaaabbb$
$B^* = \{ \lambda, 0, 1, 00, 01, 10, 11, 100, 101, \dots \}$

taal

Een *taal* (over Σ) is een verzameling strings over Σ

Σ^* alle strings

PAL = { λ, aa, bb, abba, baab, abaaba, …}

BIN = { 0, 1, 10, 11, 100, 101, 110, 111, 1000, … }

K = { a, aa, ba, aaa, aba, baa, bba, aaaa, aaba, … }

= { $x \in \{a,b\}^*$ | x eindigt op een a }

L = { λ, aa, ab, ba, bb, aaaa, aaab, aaba, aabb, … }

= { $x \in \{a,b\}^*$ | x heeft even lengte }

$\varnothing$

$\mathcal{P}(\Sigma^*)$ talen over Σ

concatenatie

$\cdot : \Sigma^* \times \Sigma^* \rightarrow \Sigma^*$
$a_1 \ldots a_m \cdot b_1 \ldots b_n = a_1 \ldots a_m b_1 \ldots b_n$

λ één $\lambda \cdot x = x = x \cdot \lambda$
associatief $(x \cdot y) \cdot z = x \cdot (y \cdot z)$
niet commutatief bei·aard ≠ aard·bei
$|x \cdot y| = |x| + |y|$
$\cdot$ wordt vaak weggelaten

prefix

$x, y \in \Sigma^*$
x is een *deelwoord* van y als $y = u \cdot x \cdot v$ voor $u, v \in \Sigma^*$
prefix … als $y = x \cdot v$ voor $v \in \Sigma^*$
suffix … als $y = u \cdot x$ voor $u \in \Sigma^*$

óók: *sub*woord

lekkerkerker deelwoord kerk (twee *voorkomens*)

spotterspot spot prefix én suffix (als λ)

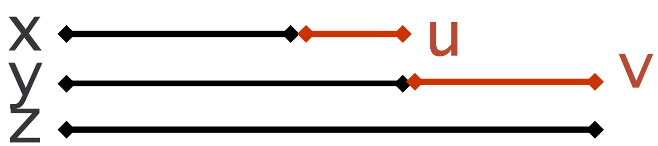
eigenschappen van prefix

$x, y \in \Sigma^*$
x is een *deelwoord* van y als $y = u \cdot x \cdot v$ voor $u, v \in \Sigma^*$
prefix … als $y = x \cdot v$ voor $v \in \Sigma^*$
suffix … als $y = u \cdot x$ voor $u \in \Sigma^*$

reflexief $x \preceq x$ voor alle x
anti-symmetrisch als $x \preceq y$ en $y \preceq x$ dan $x=y$
transitief als $x \preceq y$ en $y \preceq z$ dan $x \preceq z$

$\Rightarrow$ partiële ordening

$x = x \cdot \lambda$ prefix
$y = x \cdot u$ en $x = y \cdot v$ dan $|x| \leq |y|$ en $|y| \leq |x|$
daarom $|x| = |y|$ en dus $u = v = \lambda$, $x = y$
$y = x \cdot u$ en $z = y \cdot v$ dan $z = (x \cdot u) \cdot v = x \cdot (uv)$



macht

$x^n = x \cdot \ldots \cdot x$ (n maal)

inductief:

$x^0 = \lambda$

$x^{n+1} = x^n \cdot x$

$x^{m+n} = x^m \cdot x^n$

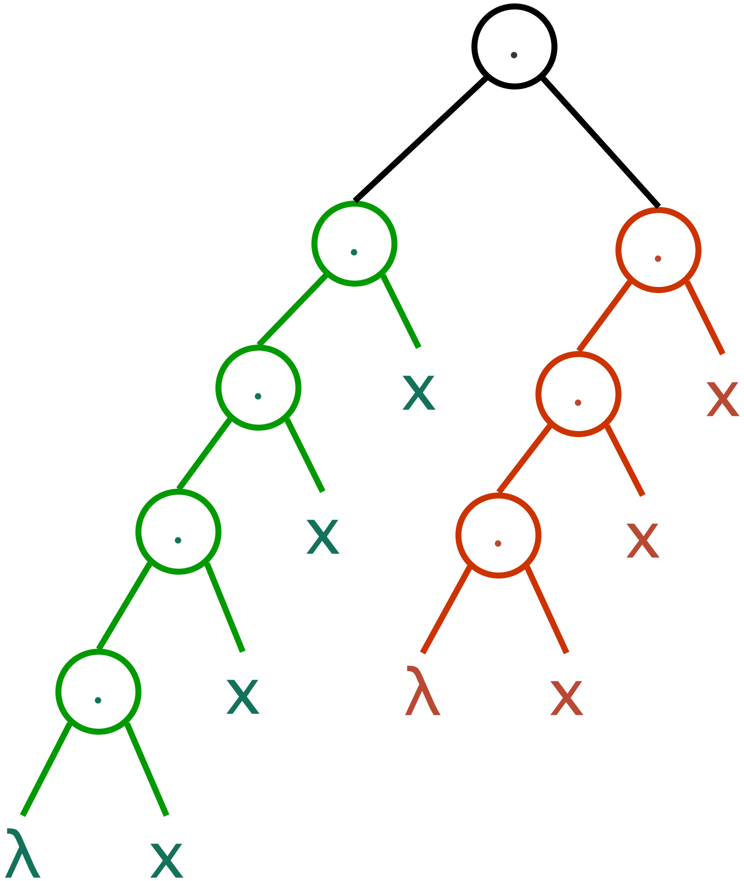
$|x^n| = n \cdot |x|$

$(ker)^5 = kerker \cdot kerk \cdot erker$

intuïtie formeel gemaakt

inductief:

$$x^0 = \lambda$$
$$x^{n+1} = x^n \cdot x$$



$x^4 \cdot x^3 = x^7$?

$((((\lambda x)x)x)x) \cdot$

$(((\lambda x)x)x)$

associatief !

spiegelbeeld

inductief:

$$\mathrm{mir}(\lambda) = \lambda$$
$$\mathrm{mir}(xa) = a \cdot \mathrm{mir}(x)$$

ook x^R

palindroom $x = \mathrm{mir}(x)$

snorfrons netebeten

$\mathrm{mir}(\mathrm{mir}(x)) = x$

$\mathrm{mir}(xy) = \mathrm{mir}(y) \cdot \mathrm{mir}(x)$

$\mathrm{mir}(x^n) = \mathrm{mir}(x)^n$

taal: Boolese operaties

Een *taal* (over Σ) is een verzameling strings over Σ

$K = \{$ a, aa, ba, aaa, aba, baa, bba, aaaa, aaba, … $\}$
$\quad = \{ x \in \{a,b\}^* \mid x$ eindigt op een a $\}$
$L = \{ \lambda,$ aa, ab, ba, bb, aaaa, aaab, aaba, aabb, … $\}$
$\quad = \{ x \in \{a,b\}^* \mid x$ heeft even lengte $\}$

$\mathcal{P}(\Sigma^*)$ talen over Σ
vereniging, doorsnede, complement (tov. Σ^*)

$K \cap L = \{$ aa, ba, aaaa, aaba, abaa, abba, baaa, … $\}$
$K - L = \{$ a, aaa, aba, baa, bba, aaaaa, aaaba, … $\}$
$L - K = \{ \lambda,$ ab, bb, aaab, aabb, abab, abbb, baab, … $\}$
$\{a,b\}^* - (K \cup L) = \{$ b, aab, abb, bab, bbb, aaaab, … $\}$

verzamelingsrekenregels

commutatief	$A \cap B = B \cap A$	$A \cup B = B \cup A$
associatief	$(A \cap B) \cap C = A \cap (B \cap C)$	$(A \cup B) \cup C = A \cup (B \cup C)$
distributief	$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$	$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
wetten van De Morgan	$(A \cap B)^C = A^C \cup B^C$	$(A \cup B)^C = A^C \cap B^C$
absorptie	$A \cap (A \cup B) = A$	$A \cup (A \cap B) = A$
idempotentie	$A \cap A = A$	$A \cup A = A$
nulelement	$A \cap \varnothing = \varnothing$	$A \cup \varnothing = A$
éénelement	$A \cap U = A$	$A \cup U = U$
dubbel complement	$(A^c)^c = A$	
complementregels	$A \cap A^c = \varnothing$	$A \cup A^c = U$

concatenatie

$\cdot : \mathcal{P}(\Sigma^*) \times \mathcal{P}(\Sigma^*) \to \mathcal{P}(\Sigma^*)$

$K \cdot L = \{ x{\cdot}y \mid x \in K \text{ en } y \in L \}$

{ a,ab }·{ a,ba } = { a·a,a·ba,ab·a,ab·ba }
= { aa,aba,abba }

{ λ,a,b,aa,ab,ba,bb,aaa,aab, … }·{ a } =
{ a,aa,ba,aaa,aba,baa,bba,aaaa,aaba, … }

{ a,ab,abb,abbb, … }·{ a,ba,aba,baba,ababa, … }
= { aa,aba,aaba,abba,ababa,abbba,aababa,abbaba, … }

	a,	ba,	aba,	baba,	ababa,	…
a	aa,	aba,	aaba,	ababa,	aababa,	…
ab	aba,	abba,	ababa,	abbaba,	abababa,	…
abb	abba,	abbba,	abbaba,	abbbaba,	abbababa,	…
abbb	abbba,	abbbba,	abbbaba,	abbbbaba,	…	
…						

concatenatie

$\cdot : \mathcal{P}(\Sigma^*) \times \mathcal{P}(\Sigma^*) \rightarrow \mathcal{P}(\Sigma^*)$

$K \cdot L = \{ x \cdot y \mid x \in K \text{ en } y \in L \}$

$\{\lambda\}$ één $\quad \{\lambda\} \cdot L = L \cdot \{\lambda\} = L$

$\varnothing$ nul $\quad \varnothing \cdot L = L \cdot \varnothing = \varnothing$

associatief $\quad (K \cdot L) \cdot M = K \cdot (L \cdot M)$

niet commutatief

macht

$K^n = K \cdot \ldots \cdot K$ (n maal)
inductief:
$K^0 = \{ \lambda \}$
$K^{n+1} = K^n \cdot K$

$\{\lambda, a, ab\}^1 = \{\lambda, a, ab\}$

$\{\lambda, a, ab\}^2 =$
$\{\lambda, a, aa, ab, a \cdot ab, aba, abab\}$

$\{\lambda, a, ab\}^3 =$
$\{\lambda, a, aa, ab, aaa, aab, aba, aaab, a \cdot ab \cdot a, abaa,$
$abab, aabab, abaab, ababa, ababab\}$

ster & plus

$K^n = K \cdot \ldots \cdot K$ (n maal)
inductief:
$K^0 = \{ \lambda \}$
$K^{n+1} = K^n \cdot K$

$K^* = \bigcup_{n\geq 0} K^n$, $K^+ = \bigcup_{n>0} K^n$

$ab^* = \{a\}\cdot\{b\}^* = \{ a,ab,abb,abbb,abbbb \ldots \}$
$(ab^*)^* = \{ a,ab,abb,abbb,abbbb \ldots \}^* =$
$= \{ \lambda,a,aa,ab,aaa,aab,aba,abb,aaaa,aaab,\ldots \}$
$= \{\lambda\} \cup \{a\}\cdot\{a,b\}^*$

$\{\lambda,a,ab\}^* = \{a,ab\}^* =$
$\{ x \mid$ voor elke b staat een a $\}$

ster & plus

$K^n = K \cdot \ldots \cdot K$ (n maal)
inductief:
$K^0 = \{ \lambda \}$
$K^{n+1} = K^n \cdot K$

$K^* = \bigcup_{n \geq 0} K^n$, $K^+ = \bigcup_{n > 0} K^n$

alternatief:
K^* is de kleinste verzameling die λ bevat en gesloten is onder concatenatie met K

oftewel (inductief):
- $\lambda \in K^*$
- als $x \in K^*$ en $w \in K$ dan $xw \in K^*$

rekenkundige expressies

$D = \{ 0,1,2, \dots , 9 \}$

i. ieder element van $D^* - \{\lambda\}$ behoort tot R

ii. als $x \in R$, dan $(-x) \in R$

als $x \in R$ en $y \in R$,

dan $(x+y) \in R$, $(x-y) \in R$, $(x \cdot y) \in R$, $(x/y) \in R$

alfabet $\{ 0,1,2, \dots ,9 \} \cup \{ +,-,\cdot,/ \} \cup \{ (,) \}$

syntax: geheel getal

$\langle$geheel$\rangle$ ::= $\langle$teken$\rangle\langle$natuurlijk$\rangle$ | $\langle$natuurlijk$\rangle$
$\langle$natuurlijk$\rangle$ ::= $\langle$cijfer$\rangle$ | $\langle$cijfer$\rangle\langle$natuurlijk$\rangle$
$\langle$cijfer$\rangle$ ::= 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9
$\langle$teken$\rangle$ ::= + | -

$$\{ +,- \} \cdot \{ 0,1, \dots ,9 \}^+ \cup \{ 0,1, \dots ,9 \}^+ =$$
$$\{ +,-,\lambda \} \cdot \{ 0,1, \dots ,9 \}^+$$

reguliere expressies in PERL

```
$greeting = "World";
print "It matches\n" if "Hello World" =~ /$greeting/;

"cats and dogs" =~ /dog|cat|bird/; # matches "cat"

/item[0-9]/; # matches 'item0' or ... or 'item9' \d
/(a|b)b/;    # matches 'ab' or 'bb'
   a* = match 'a' 0 or more times
   a? = match 'a' 1 or 0 times
   period '.' matches any character but "\n" (eol)

$time =~ /(\d\d):(\d\d):(\d\d)/; # match hh:mm:ss
$hours = $1; $minutes = $2; $seconds = $3;

# match a number, get $1 = whole number, $2 = exponent
/([+-]?\ *(?:\d+(?:\.\d*)?|\.\d+)(?:[eE]([+-]?\d+))?)/;
```

rekenregels 'algebra'

$\varnothing \cdot L = L \cdot \varnothing = \varnothing$

$\{\lambda\} \cdot L = L \cdot \{\lambda\} = L$

$(K \cdot L) \cdot M = K \cdot (L \cdot M)$

$K \cdot (L \cup M) = K \cdot L \cup K \cdot M$

$(K \cup L) \cdot M = K \cdot M \cup L \cdot M$

$(K^*)^n = K^*$ voor alle $n \geq 1$

$(K^*)^* = K^*$

$(K^* L^*)^* = (K \cup L)^* = (K^* \cup L^*)^*$

$K \subseteq L$

dan $K \cup M \subseteq L \cup M$

$K \cdot M \subseteq L \cdot M$, $M \cdot K \subseteq M \cdot L$

$K^n \subseteq L^n$ voor alle $n \geq 1$

$K^* \subseteq L^*$

$\boxed{K^*K^* = K^*}$

'$\subseteq$'

$z \in K^*K^*$ dan $z = xy$ met $x,y \in K^*$,

zeg $x \in K^m$ en $y \in K^n$,

dus $z = xy \in K^{m+n} \subseteq K^*$.

'$\supseteq$'

$K^* = K^*\{\lambda\} \subseteq K^*K^*$

$$\boxed{\bigcup_{n\geq 0} K^n}$$

$$K^*K^* = (K^0 \cup K^1 \cup K^2 \cup K^3 \cup \cdots)(K^0 \cup K^1 \cup K^2 \cup K^3 \cup \cdots) = \bigcup_{m,n\geq 0} K^m \cdot K^n = \bigcup_{m,n\geq 0} K^{m+n} = \bigcup_{s\geq 0} K^s = K^*$$

oneindige vereniging … ?

eindig:associatief!

$$\bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup \cdots \cup A_n = ((\cdots(A_1 \cup A_2) \cup \cdots) \cup A_n)$$

$$\boxed{\left(\bigcup_{i=1}^{n} A_i\right)^c = \bigcap_{i=1}^{n} A_i^c}$$

$$\neg(\varphi \vee \psi) \equiv \neg\varphi \wedge \neg\psi$$

$$\left(\bigcup_{i=1}^{n+1} A_i\right)^c = \left(\bigcup_{i=1}^{n} A_i \cup A_{n+1}\right)^c =$$

$$\left(\bigcup_{i=1}^{n} A_i\right)^c \cap A_{n+1}^c = \bigcap_{i=1}^{n} A_i^c \cap A_{n+1}^c = \bigcap_{i=1}^{n+1} A_i^c$$

$$\bigcup_{i=1}^{\infty} A_i = \{x \in U \mid (\exists i)(x \in A_i)\}$$

oneindig ?

$$\boxed{\left(\bigcup_{i=1}^{\infty} A_i\right)^c = \bigcap_{i=1}^{\infty} A_i^c}$$

$$\neg(\exists i)\psi(i) = (\forall i)\neg\psi(i)$$

uitstapje: taal vastleggen

- beschrijven, genereren

  K: w_0, w_1, w_2, w_3, …

  - recursieve definitie $\Rightarrow$
  - syntax-diagram
  - grammatica
  - operaties $*$ $\cup$ $\cdot$
  - logische formule

- herkennen, algoritme

  $w \in K$? ja/nee

  - automaat, Turing machine
  - parser

excursie

grammatica

- MI ∈ L
- als xI ∈ L dan xIU ∈ L ♣
  als Mx ∈ L dan Mxx ∈ L ♦
  als xIIIy ∈ L dan xUy ∈ L ♥
  als xUUy ∈ L dan xy ∈ L ♠
- L bevat geen andere elementen

axioma MI

regels

I	$\rightarrow$	IU	mits uiteinde
Mx	$\rightarrow$	Mxx	mits hele woord (?!)
III	$\rightarrow$	U	in elke context
UU	$\rightarrow$	λ	

MI $\Rightarrow$ MII $\Rightarrow$ MIIII $\Rightarrow$ MUI $\Rightarrow$ MUIUI $\Rightarrow$ MUIUIU $\Rightarrow$...

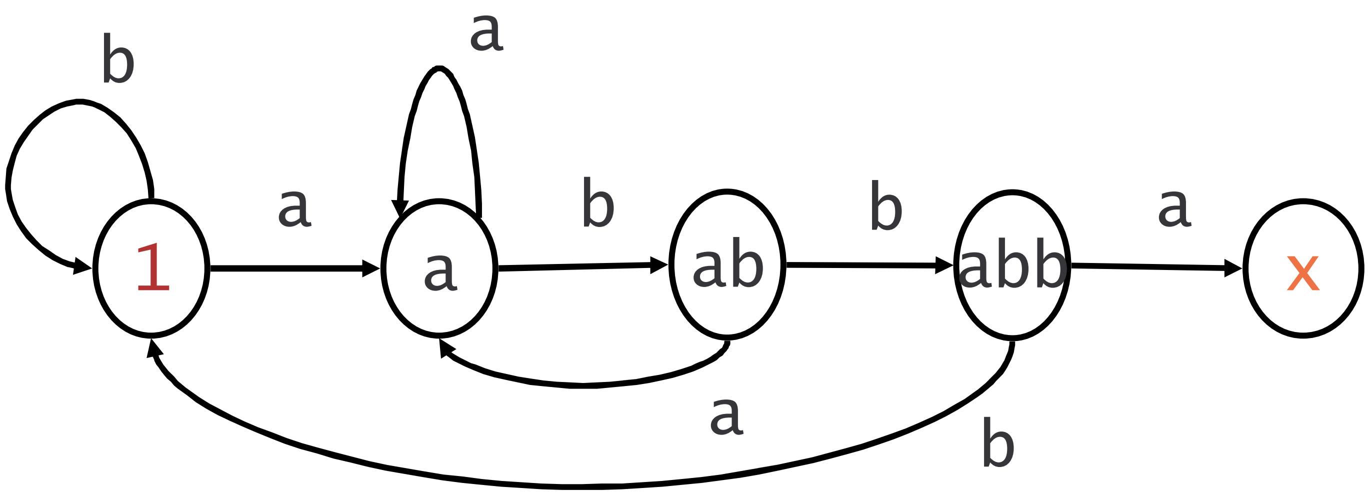
Chomsky: type regels *families*
rechts-lineair, context-vrij, monotoon, type-0

kleinste familie van talen die
- $\varnothing$ $\{\lambda\}$ $\{a\}$ bevat
- gesloten is onder $*$ $\cup$ $\cdot$

- $\varnothing$ $\{\lambda\}$ $\{a\}$ $\in$ REG
- als $K, L \in$ REG dan
  - $K \cup L \in$ REG
  - $K \cdot L \in$ REG
  - $K^* \in$ REG
- verder geen talen in REG

patroonherkenning

{ x∈{a,b}* | x heeft geen subwoord abba }



van 1 naar 1

({b}∪{a}{a,ba}*{bbb})*

· ({λ}∪{a}{a,ba}*{λ,b,bb})

spiegelbeeld

$\mathrm{mir}(K) = \{ \mathrm{mir}(x) \mid x \in K \}$

$\mathrm{mir}(\mathrm{mir}(K)) = K$
$\mathrm{mir}(K \cup L) = \mathrm{mir}(K) \cup \mathrm{mir}(L)$
$\mathrm{mir}(K \cdot L) = \mathrm{mir}(L) \cdot \mathrm{mir}(K)$
$\mathrm{mir}(K^n) = \mathrm{mir}(K)^n$
$\mathrm{mir}(K^*) = \mathrm{mir}(K)^*$

- $\varnothing$ $\{\lambda\}$ $\{a\}$ $\in$ REG
- als K, $L \in$ REG dan
  - $K \cup L \in$ REG
  - $K \cdot L \in$ REG
  - $K^* \in$ REG

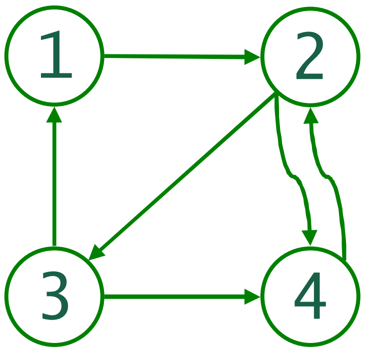
REG is gesloten onder mir
dwz: K regulier dan ook mir(K) regulier

probleem = taal

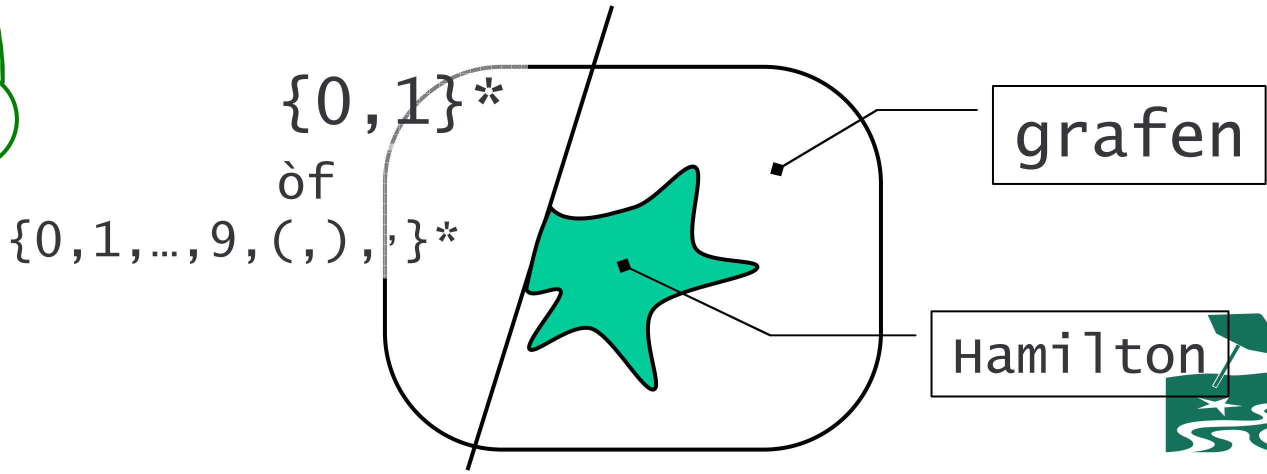
HPP: Hamilton Path Problem
gegeven: gerichte graaf G, begin- & eindpunt
gevr: is er een Hamilton pad (elke knoop éénmaal)

aantal	punten	pijlen					begin & eind
	4	(1,2)	(2,3)	(2,4)	(3,1)	(3,4)	(4,2) 1,2

000 100 001010 010011 010100 011001 011100 100010 000 001010



G graaf -> woord w(G)
Hamilton grafen -> taal Ham



end...