Algebraische Dekoderspezifikation: Kopplung zwischen der Theorie formaler Sprachen und dem Statistischen Maschinellen Übersetzen

Verteidigung der Dissertation
"Algebraic decoder specification:
coupling formal-language theory
and statistical machine translation"

Matthias Büchse

TECHNISCHE UNIVERSITÄT DRESDEN

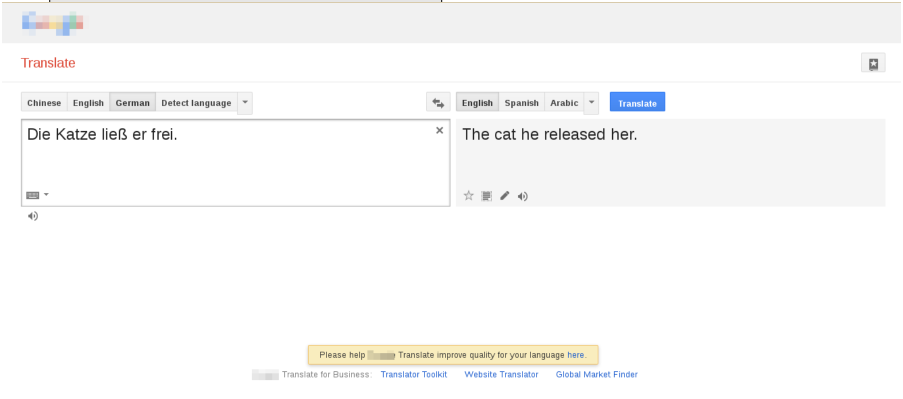
2014-12-18

Übersicht

Statistisches
Maschinelles Übersetzen

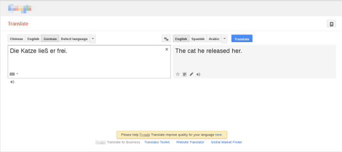
Übersicht

Statistisches
Maschinelles Übersetzen



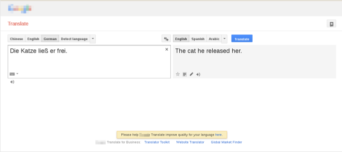
Übersicht

Statistisches
Maschinelles Übersetzen



Übersicht

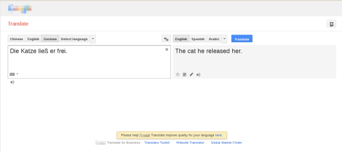
Statistisches
Maschinelles Übersetzen



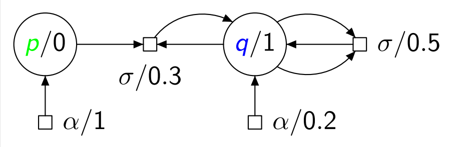
Theorie der
formalen Sprachen

Übersicht

Statistisches
Maschinelles Übersetzen

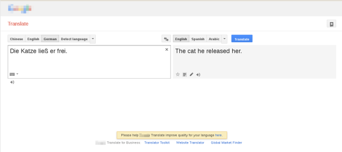


Theorie der
formalen Sprachen

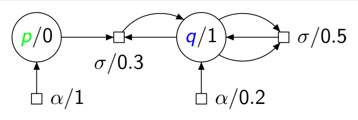


Übersicht

Statistisches Maschinelles Übersetzen

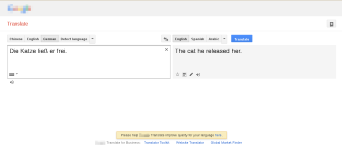


Theorie der formalen Sprachen



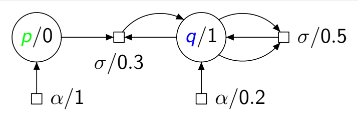
Übersicht

Statistisches
Maschinelles Übersetzen



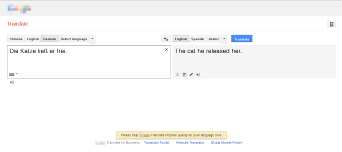
Dekoder-
spezifikation

Theorie der
formalen Sprachen



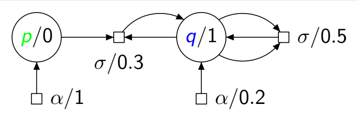
Übersicht

Statistisches Maschinelles Übersetzen



Dekoder-
spezifikation

Theorie der formalen Sprachen



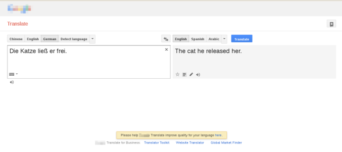
Abschlusseigenschaften

Binarisierung

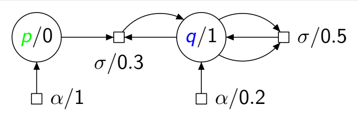
Determinisierung

Übersicht

Statistisches
Maschinelles Übersetzen



Theorie der
formalen Sprachen



Dekoder-
spezifikation

algebraische
Dekoder-
spezifikation

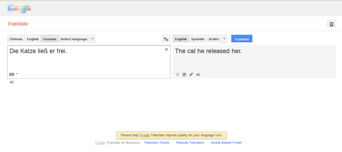
Abschlusseigenschaften

Binarisierung

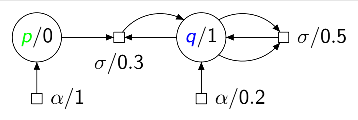
Determinisierung

Übersicht

Statistisches Maschinelles Übersetzen



Theorie der formalen Sprachen



Dekoder-spezifikation

algebraische Dekoder-spezifikation

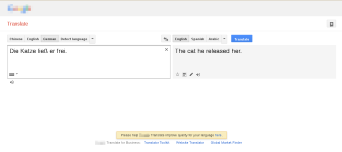
Abschlusseigenschaften

Binarisierung

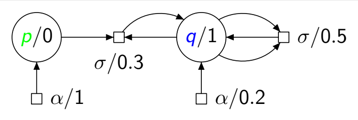
Determinisierung

Übersicht

Statistisches
Maschinelles Übersetzen



Theorie der
formalen Sprachen



1

2

3

Dekoder-
spezifikation

algebraische
Dekoder-
spezifikation

Abschlusseigenschaften

Binarisierung

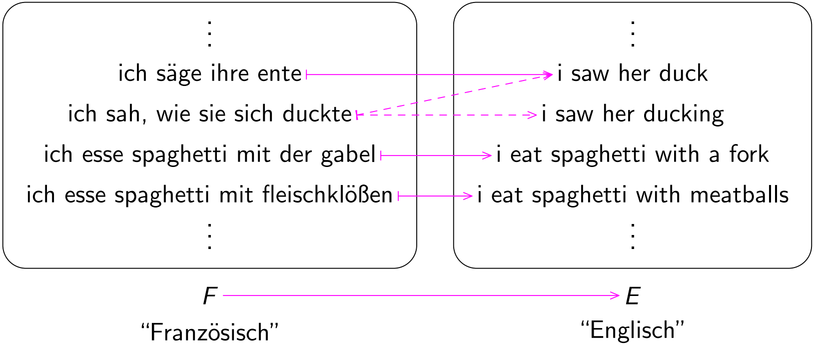
Determinisierung

Gliederung

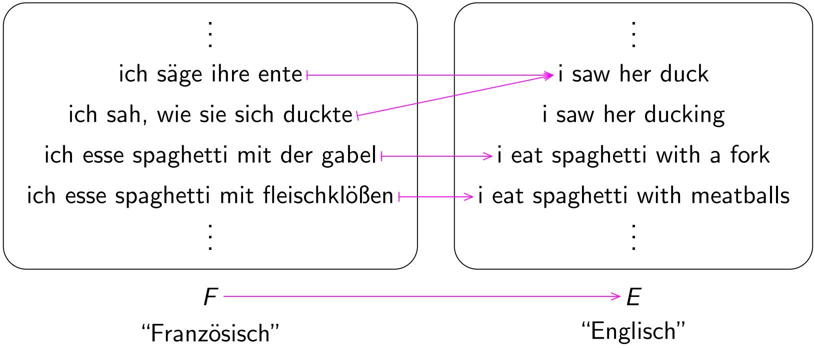
Statistisches Maschinelles Übersetzen

⋮	⋮
ich säge ihre ente	i saw her duck
ich sah, wie sie sich duckte	i saw her ducking
ich esse spaghetti mit der gabel	i eat spaghetti with a fork
ich esse spaghetti mit fleischklößen	i eat spaghetti with meatballs
⋮	⋮
F	E
"Französisch"	"Englisch"

Statistisches Maschinelles Übersetzen



Statistisches Maschinelles Übersetzen



Statistisches Maschinelles Übersetzen

⋮	⋮
ich säge ihre ente	i saw her duck
ich sah, wie sie sich duckte	i saw her ducking
ich esse spaghetti mit der gabel	i eat spaghetti with a fork
ich esse spaghetti mit fleischklößen	i eat spaghetti with meatballs
⋮	⋮

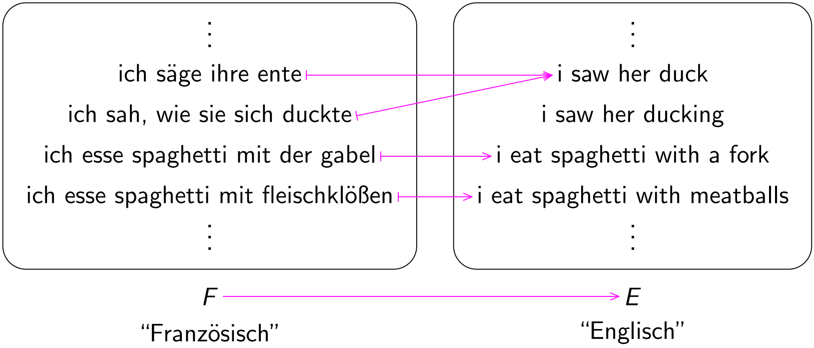
$$F \to E$$

"Französisch" "Englisch"

Dekoder

$\mathbb{D}: \quad F \to E$

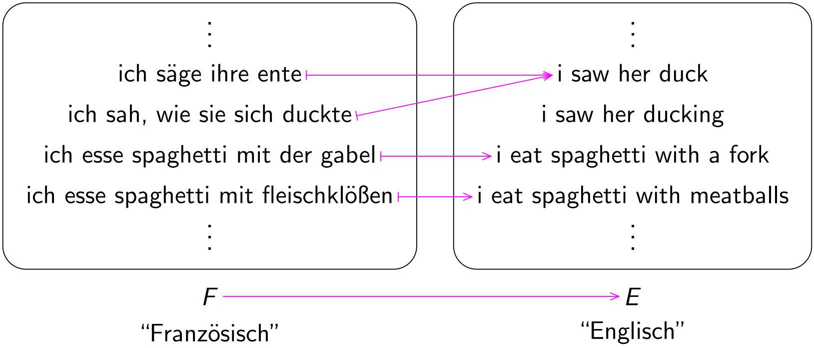
Statistisches Maschinelles Übersetzen



Dekoder

$\mathbb{D}\colon \Omega \to (F \to E)$ $\quad\quad$ Ω ... Parameterraum

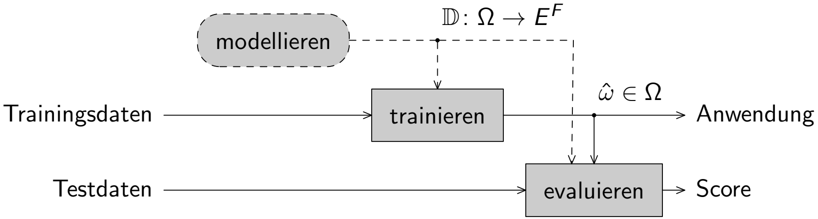
Statistisches Maschinelles Übersetzen



Dekoder

$\mathbb{D}\colon \Omega \to E^F$ $\qquad\qquad$ Ω ... Parameterraum

Entwicklungszyklus



Modellieren

Hierarchische Phrasen

die katze ließ er frei

$$\omega \in \Omega: \quad \begin{array}{rcl} x_1 \text{ ließ } x_2 \text{ frei} & \longleftrightarrow & x_2 \text{ freed } x_1 \\ \text{die katze} & \longleftrightarrow & \text{the cat} \\ \text{er} & \longleftrightarrow & \text{he} \end{array}$$

Modellieren

Hierarchische Phrasen

$$\text{die katze ließ er frei} \quad \rightsquigarrow \quad \overset{\text{die katze}}{\overset{\downarrow}{x_1}} \quad \text{ließ} \quad \overset{\text{er}}{\overset{\downarrow}{x_2}} \quad \text{frei}$$

$$\omega \in \Omega: \quad \begin{array}{rcl} x_1 \text{ ließ } x_2 \text{ frei} & \longleftrightarrow & x_2 \text{ freed } x_1 \\ \text{die katze} & \longleftrightarrow & \text{the cat} \\ \text{er} & \longleftrightarrow & \text{he} \end{array}$$

Modellieren

Hierarchische Phrasen

$$
\text{die katze ließ er frei} \quad \rightsquigarrow \quad
\begin{array}{cccc}
\text{die katze} & & \text{er} & \\
\downarrow & & \downarrow & \\
x_1 & \text{ließ} & x_2 & \text{frei} \\
 & \wr & & \\
x_2 & \text{freed} & x_1 & \\
\uparrow & & \uparrow & \\
\text{he} & & \text{the cat} &
\end{array}
$$

$$
\omega \in \Omega: \qquad
\begin{array}{rcl}
x_1 \text{ ließ } x_2 \text{ frei} & \longleftrightarrow & x_2 \text{ freed } x_1 \\
\text{die katze} & \longleftrightarrow & \text{the cat} \\
\text{er} & \longleftrightarrow & \text{he}
\end{array}
$$

Modellieren

Hierarchische Phrasen

$$
\begin{array}{ccccccc}
 & & \text{die katze} & & \text{er} & \\
 & & \downarrow & & \downarrow & \\
\text{die katze ließ er frei} & \rightsquigarrow & x_1 & \text{ließ} & x_2 & \text{frei} \\
 & & & \wr & & \\
\text{he freed the cat} & \leftsquigarrow & x_2 & \text{freed} & x_1 & \\
 & & \uparrow & & \uparrow & \\
 & & \text{he} & & \text{the cat} &
\end{array}
$$

$$
\omega \in \Omega: \quad
\begin{array}{rcl}
x_1 \text{ ließ } x_2 \text{ frei} & \longleftrightarrow & x_2 \text{ freed } x_1 \\
\text{die katze} & \longleftrightarrow & \text{the cat} \\
\text{er} & \longleftrightarrow & \text{he}
\end{array}
$$

Trainieren

paralleles Korpus

$$d \in (E \times F)^*$$

Resumption of the session	Wiederaufnahme der Sitzungsperiode
I declare resumed the session of the European Parliament adjourned on Friday 17 December 1999 , [. . .] .	Ich erkläre die am Freitag , dem 17. Dezember unterbrochene Sitzungsperiode des Europäischen Parlaments für wiederaufgenommen , [. . .] .
$\vdots$	$\vdots$

EuroParl-Korpus, 11 Sprachen, je 1.5 Mio. Sätze

Trainieren

paralleles Korpus

$$d \in (E \times F)^*$$

$$\downarrow$$

Heuristiken und statistische Methoden anwenden

Trainieren

paralleles Korpus

$d \in (E \times F)^*$

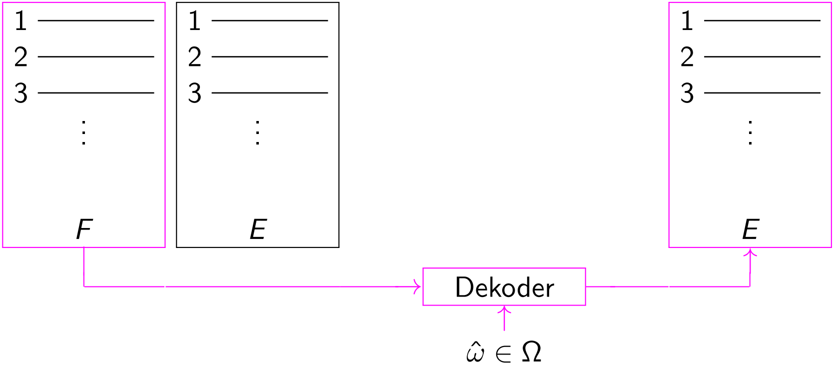
$\downarrow$

Heuristiken und statistische Methoden anwenden

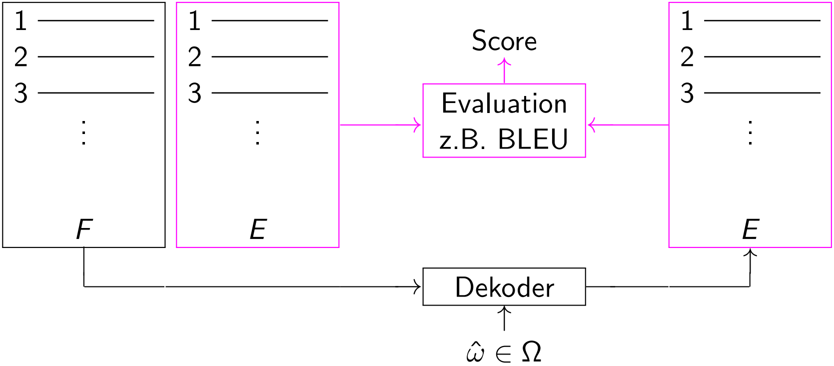
$\downarrow$

$\hat{\omega} \in \Omega$

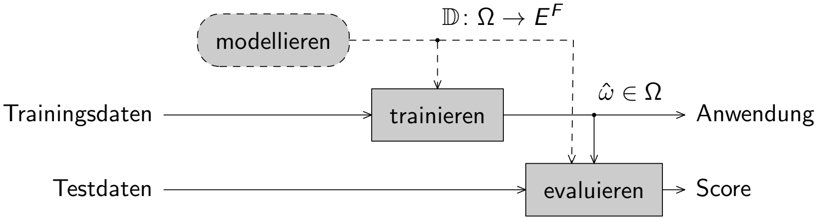
Evaluieren



Evaluieren


1
2
3
F
1
2
3
E
Score
Evaluation
z.B. BLEU
1
2
3
E
Dekoder
$\hat{\omega} \in \Omega$

Entwicklungszyklus



Gliederung

Idee

Hierarchische Phrasen

$$\begin{tikzcd}
 & & \text{die katze} \arrow[d] & & \text{er} \arrow[d] & \\
\text{die katze ließ er frei} & \rightsquigarrow & {\color{blue}x_1} & \text{ließ} \arrow[d, squiggly] & {\color{blue}x_2} & \text{frei} \\
\text{he freed the cat} & \leftsquigarrow & {\color{blue}x_2} & \text{freed} & {\color{blue}x_1} & \\
 & & \text{he} \arrow[u] & & \text{the cat} \arrow[u] &
\end{tikzcd}$$

$$\omega \in \Omega: \quad \begin{array}{rcl} {\color{blue}x_1} \text{ ließ } {\color{blue}x_2} \text{ frei} & \longleftrightarrow & {\color{blue}x_2} \text{ freed } {\color{blue}x_1} \\ \text{die katze} & \longleftrightarrow & \text{the cat} \\ \text{er} & \longleftrightarrow & \text{he} \end{array}$$

Synchrone kontextfreie Grammatiken (SCFGs)

(Lewis und Stearns 1966)

$$\rho_1: \quad S \rightarrow \langle x_1 \text{ ließ er frei}, \text{he freed } x_1 \rangle(NP)$$

$$\rho_2: \quad S \rightarrow \langle \text{die katze ließ } x_1 \text{ frei}, x_1 \text{ let the cat out} \rangle(PPER)$$

$$\rho_3: \quad S \rightarrow \langle x_1 \text{ ließ } x_2 \text{ frei}, x_1 \text{ freed } x_2 \rangle(PPER, NP)$$

$$\rho_4: \quad S \rightarrow \langle x_2 \text{ ließ } x_1 \text{ frei}, x_1 \text{ freed } x_2 \rangle(PPER, NP)$$

$$\rho_5: \quad PPER \rightarrow \langle \text{er}, \text{he} \rangle$$

$$\rho_6: \quad NP \rightarrow \langle \text{die katze}, \text{the cat} \rangle$$

Synchrone kontextfreie Grammatiken (SCFGs)

(Lewis und Stearns 1966)

$$
\begin{array}{lrl}
\rho_1\colon & S \to \alpha_1(NP) & \alpha_1 = \langle x_1 \text{ ließ er frei}, \text{he freed } x_1\rangle \\
\rho_2\colon & S \to \alpha_2(PPER) & \alpha_2 = \langle \text{die katze ließ } x_1 \text{ frei}, x_1 \text{ let the cat out}\rangle \\
\rho_3\colon & S \to \alpha_3(PPER, NP) & \alpha_3 = \langle x_1 \text{ ließ } x_2 \text{ frei}, x_1 \text{ freed } x_2\rangle \\
\rho_4\colon & S \to \alpha_4(PPER, NP) & \alpha_4 = \langle x_2 \text{ ließ } x_1 \text{ frei}, x_1 \text{ freed } x_2\rangle \\
\rho_5\colon & PPER \to \alpha_5 & \alpha_5 = \langle \text{er}, \text{he}\rangle \\
\rho_6\colon & NP \to \alpha_6 & \alpha_6 = \langle \text{die katze}, \text{the cat}\rangle
\end{array}
$$

Synchrone kontextfreie Grammatiken (SCFGs)

(Lewis und Stearns 1966)

$$\begin{array}{lrll}
\rho_1: & S \to \alpha_1(NP) & & \alpha_1 = \langle x_1 \text{ ließ er frei}, \text{he freed } x_1 \rangle \\
\rho_2: & S \to \alpha_2(PPER) & & \alpha_2 = \langle \text{die katze ließ } x_1 \text{ frei}, x_1 \text{ let the cat out} \rangle \\
\rho_3: & S \to \alpha_3(PPER, NP) & & \alpha_3 = \langle x_1 \text{ ließ } x_2 \text{ frei}, x_1 \text{ freed } x_2 \rangle \\
\rho_4: & S \to \alpha_4(PPER, NP) & & \alpha_4 = \langle x_2 \text{ ließ } x_1 \text{ frei}, x_1 \text{ freed } x_2 \rangle \\
\rho_5: & PPER \to \alpha_5 & & \alpha_5 = \langle \text{er}, \text{he} \rangle \\
\rho_6: & NP \to \alpha_6 & & \alpha_6 = \langle \text{die katze}, \text{the cat} \rangle
\end{array}$$

S

Synchrone kontextfreie Grammatiken (SCFGs)

(Lewis und Stearns 1966)

$$
\begin{array}{lrl}
\rho_1: & S \to \alpha_1(NP) & \alpha_1 = \langle x_1 \text{ ließ er frei}, \text{he freed } x_1 \rangle \\
\rho_2: & S \to \alpha_2(PPER) & \alpha_2 = \langle \text{die katze ließ } x_1 \text{ frei}, x_1 \text{ let the cat out} \rangle \\
\rho_3: & S \to \alpha_3(PPER, NP) & \alpha_3 = \langle x_1 \text{ ließ } x_2 \text{ frei}, x_1 \text{ freed } x_2 \rangle \\
\rho_4: & S \to \alpha_4(PPER, NP) & \alpha_4 = \langle x_2 \text{ ließ } x_1 \text{ frei}, x_1 \text{ freed } x_2 \rangle \\
\rho_5: & PPER \to \alpha_5 & \alpha_5 = \langle \text{er}, \text{he} \rangle \\
\rho_6: & NP \to \alpha_6 & \alpha_6 = \langle \text{die katze}, \text{the cat} \rangle
\end{array}
$$

$$S \xRightarrow{\rho_4} \alpha_4(PPER, NP)$$

Synchrone kontextfreie Grammatiken (SCFGs)

(Lewis und Stearns 1966)

$$
\begin{array}{lrl}
\rho_1: & S \to \alpha_1(NP) & \alpha_1 = \langle x_1 \text{ ließ er frei}, \text{he freed } x_1 \rangle \\
\rho_2: & S \to \alpha_2(PPER) & \alpha_2 = \langle \text{die katze ließ } x_1 \text{ frei}, x_1 \text{ let the cat out} \rangle \\
\rho_3: & S \to \alpha_3(PPER, NP) & \alpha_3 = \langle x_1 \text{ ließ } x_2 \text{ frei}, x_1 \text{ freed } x_2 \rangle \\
\rho_4: & S \to \alpha_4(PPER, NP) & \alpha_4 = \langle x_2 \text{ ließ } x_1 \text{ frei}, x_1 \text{ freed } x_2 \rangle \\
\rho_5: & PPER \to \alpha_5 & \alpha_5 = \langle \text{er}, \text{he} \rangle \\
\rho_6: & NP \to \alpha_6 & \alpha_6 = \langle \text{die katze}, \text{the cat} \rangle
\end{array}
$$

$$
S \xRightarrow{\rho_4} \alpha_4(PPER, NP) \xRightarrow{\rho_5} \alpha_4(\alpha_5, NP)
$$

Synchrone kontextfreie Grammatiken (SCFGs)

(Lewis und Stearns 1966)

$$
\begin{array}{lrl}
\rho_1: & S \to \alpha_1(NP) & \alpha_1 = \langle x_1 \text{ ließ er frei}, \text{he freed } x_1 \rangle \\
\rho_2: & S \to \alpha_2(PPER) & \alpha_2 = \langle \text{die katze ließ } x_1 \text{ frei}, x_1 \text{ let the cat out} \rangle \\
\rho_3: & S \to \alpha_3(PPER, NP) & \alpha_3 = \langle x_1 \text{ ließ } x_2 \text{ frei}, x_1 \text{ freed } x_2 \rangle \\
\rho_4: & S \to \alpha_4(PPER, NP) & \alpha_4 = \langle x_2 \text{ ließ } x_1 \text{ frei}, x_1 \text{ freed } x_2 \rangle \\
\rho_5: & PPER \to \alpha_5 & \alpha_5 = \langle \text{er}, \text{he} \rangle \\
\rho_6: & NP \to \alpha_6 & \alpha_6 = \langle \text{die katze}, \text{the cat} \rangle
\end{array}
$$

$$
S \xRightarrow{\rho_4} \alpha_4(PPER, NP) \xRightarrow{\rho_5} \alpha_4(\alpha_5, NP) \xRightarrow{\rho_6} \alpha_4(\alpha_5, \alpha_6)
$$

Synchrone kontextfreie Grammatiken (SCFGs)

(Lewis und Stearns 1966)

$$
\begin{array}{lrl}
\rho_1: & S \to \alpha_1(NP) & \alpha_1 = \langle x_1 \text{ ließ er frei}, \text{he freed } x_1\rangle \\
\rho_2: & S \to \alpha_2(PPER) & \alpha_2 = \langle \text{die katze ließ } x_1 \text{ frei}, x_1 \text{ let the cat out}\rangle \\
\rho_3: & S \to \alpha_3(PPER, NP) & \alpha_3 = \langle x_1 \text{ ließ } x_2 \text{ frei}, x_1 \text{ freed } x_2\rangle \\
\rho_4: & S \to \alpha_4(PPER, NP) & \alpha_4 = \langle x_2 \text{ ließ } x_1 \text{ frei}, x_1 \text{ freed } x_2\rangle \\
\rho_5: & PPER \to \alpha_5 & \alpha_5 = \langle \text{er}, \text{he}\rangle \\
\rho_6: & NP \to \alpha_6 & \alpha_6 = \langle \text{die katze}, \text{the cat}\rangle
\end{array}
$$

$$
S \xRightarrow{\rho_4} \alpha_4(PPER, NP) \xRightarrow{\rho_5} \alpha_4(\alpha_5, NP) \xRightarrow{\rho_6} \alpha_4(\alpha_5, \alpha_6) \qquad \rho_4(\rho_5, \rho_6)
$$

Synchrone kontextfreie Grammatiken (SCFGs)

(Lewis und Stearns 1966)

$$
\begin{array}{lrll}
\rho_1\colon & S \to \alpha_1(NP) & \alpha_1 = \langle x_1 \text{ ließ er frei}, \text{he freed } x_1\rangle \\
\rho_2\colon & S \to \alpha_2(PPER) & \alpha_2 = \langle \text{die katze ließ } x_1 \text{ frei}, x_1 \text{ let the cat out}\rangle \\
\rho_3\colon & S \to \alpha_3(PPER, NP) & \alpha_3 = \langle x_1 \text{ ließ } x_2 \text{ frei}, x_1 \text{ freed } x_2\rangle \\
\rho_4\colon & S \to \alpha_4(PPER, NP) & \alpha_4 = \langle x_2 \text{ ließ } x_1 \text{ frei}, x_1 \text{ freed } x_2\rangle \\
\rho_5\colon & PPER \to \alpha_5 & \alpha_5 = \langle \text{er}, \text{he}\rangle \\
\rho_6\colon & NP \to \alpha_6 & \alpha_6 = \langle \text{die katze}, \text{the cat}\rangle
\end{array}
$$

$$S \xRightarrow{\rho_4} \alpha_4(PPER, NP) \xRightarrow{\rho_5} \alpha_4(\alpha_5, NP) \xRightarrow{\rho_6} \alpha_4(\alpha_5, \alpha_6) \qquad \rho_4(\rho_5, \rho_6)$$

$$h_1(\alpha_4(\alpha_5, \alpha_6))$$

$$h_2(\alpha_4(\alpha_5, \alpha_6))$$

Synchrone kontextfreie Grammatiken (SCFGs)

(Lewis und Stearns 1966)

$$
\begin{array}{lrll}
\rho_1: & S \to \alpha_1(NP) & \alpha_1 = \langle x_1 \text{ ließ er frei}, \text{he freed } x_1 \rangle \\
\rho_2: & S \to \alpha_2(PPER) & \alpha_2 = \langle \text{die katze ließ } x_1 \text{ frei}, x_1 \text{ let the cat out} \rangle \\
\rho_3: & S \to \alpha_3(PPER, NP) & \alpha_3 = \langle x_1 \text{ ließ } x_2 \text{ frei}, x_1 \text{ freed } x_2 \rangle \\
\rho_4: & S \to \alpha_4(PPER, NP) & \alpha_4 = \langle x_2 \text{ ließ } x_1 \text{ frei}, x_1 \text{ freed } x_2 \rangle \\
\rho_5: & PPER \to \alpha_5 & \alpha_5 = \langle \text{er}, \text{he} \rangle \\
\rho_6: & NP \to \alpha_6 & \alpha_6 = \langle \text{die katze}, \text{the cat} \rangle
\end{array}
$$

$$
S \overset{\rho_4}{\Longrightarrow} \alpha_4(PPER, NP) \overset{\rho_5}{\Longrightarrow} \alpha_4(\alpha_5, NP) \overset{\rho_6}{\Longrightarrow} \alpha_4(\alpha_5, \alpha_6) \qquad \rho_4(\rho_5, \rho_6)
$$

$$
h_1(\alpha_4(\alpha_5, \alpha_6)) = h_1(\alpha_6) \text{ ließ } h_1(\alpha_5) \text{ frei}
$$

$$
h_2(\alpha_4(\alpha_5, \alpha_6)) = h_2(\alpha_5) \text{ freed } h_2(\alpha_6)
$$

Synchrone kontextfreie Grammatiken (SCFGs)

(Lewis und Stearns 1966)

ρ_1:	$S \to \alpha_1(NP)$	$\alpha_1 = \langle x_1 \text{ ließ er frei}, \text{he freed } x_1\rangle$
ρ_2:	$S \to \alpha_2(PPER)$	$\alpha_2 = \langle \text{die katze ließ } x_1 \text{ frei}, x_1 \text{ let the cat out}\rangle$
ρ_3:	$S \to \alpha_3(PPER, NP)$	$\alpha_3 = \langle x_1 \text{ ließ } x_2 \text{ frei}, x_1 \text{ freed } x_2\rangle$
ρ_4:	$S \to \alpha_4(PPER, NP)$	$\alpha_4 = \langle x_2 \text{ ließ } x_1 \text{ frei}, x_1 \text{ freed } x_2\rangle$
ρ_5:	$PPER \to \alpha_5$	$\alpha_5 = \langle \text{er}, \text{he}\rangle$
ρ_6:	$NP \to \alpha_6$	$\alpha_6 = \langle \text{die katze}, \text{the cat}\rangle$

$$S \xRightarrow{\rho_4} \alpha_4(PPER, NP) \xRightarrow{\rho_5} \alpha_4(\alpha_5, NP) \xRightarrow{\rho_6} \alpha_4(\alpha_5, \alpha_6) \qquad \rho_4(\rho_5, \rho_6)$$

$$h_1(\alpha_4(\alpha_5, \alpha_6)) = h_1(\alpha_6) \text{ ließ } h_1(\alpha_5) \text{ frei} = \text{die katze ließ er frei}$$

$$h_2(\alpha_4(\alpha_5, \alpha_6)) = h_2(\alpha_5) \text{ freed } h_2(\alpha_6) = \text{he freed the cat}$$

Synchrone kontextfreie Grammatiken (SCFGs)

(Lewis und Stearns 1966)

ρ_1:	$S \to \alpha_1(NP)$	$\alpha_1 = \langle x_1 \text{ ließ er frei}, \text{he freed } x_1\rangle$
ρ_2:	$S \to \alpha_2(PPER)$	$\alpha_2 = \langle \text{die katze ließ } x_1 \text{ frei}, x_1 \text{ let the cat out}\rangle$
ρ_3:	$S \to \alpha_3(PPER, NP)$	$\alpha_3 = \langle x_1 \text{ ließ } x_2 \text{ frei}, x_1 \text{ freed } x_2\rangle$
ρ_4:	$S \to \alpha_4(PPER, NP)$	$\alpha_4 = \langle x_2 \text{ ließ } x_1 \text{ frei}, x_1 \text{ freed } x_2\rangle$
ρ_5:	$PPER \to \alpha_5$	$\alpha_5 = \langle \text{er}, \text{he}\rangle$
ρ_6:	$NP \to \alpha_6$	$\alpha_6 = \langle \text{die katze}, \text{the cat}\rangle$

$$S \xRightarrow{\rho_4} \alpha_4(PPER, NP) \xRightarrow{\rho_5} \alpha_4(\alpha_5, NP) \xRightarrow{\rho_6} \alpha_4(\alpha_5, \alpha_6) \qquad \rho_4(\rho_5, \rho_6)$$

$$h_1(\alpha_4(\alpha_5, \alpha_6)) = h_1(\alpha_6) \text{ ließ } h_1(\alpha_5) \text{ frei} = \text{die katze ließ er frei}$$
$$h_2(\alpha_4(\alpha_5, \alpha_6)) = h_2(\alpha_5) \text{ freed } h_2(\alpha_6) \quad = \text{he freed the cat}$$

Abstrakte Syntaxbäume

$D^S(G) = \{\rho_1(\rho_6), \rho_2(\rho_5), \rho_3(\rho_5, \rho_6), \rho_4(\rho_5, \rho_6)\}$

$D^{PPER}(G) = \{\rho_5\}$, $D^{NP}(G) = \{\rho_6\}$

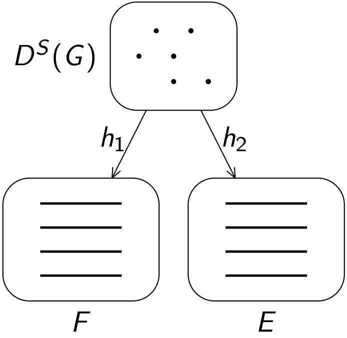
Abstrakte Spezifikation
(Chiang 2007)

$$\Omega = \{\; G \qquad \mid G \text{ SCFG} \qquad\qquad \}$$

$$\mathbb{D}(G \qquad)\colon f \mapsto$$

Abstrakte Spezifikation
(Chiang 2007)

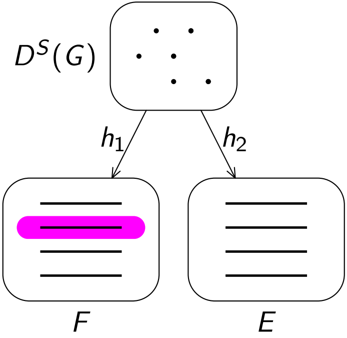
$$\Omega = \{\ G \qquad \mid G \text{ SCFG} \qquad\qquad\qquad \}$$
$$\mathbb{D}(G \qquad): f \mapsto$$



Abstrakte Spezifikation
(Chiang 2007)

$$\Omega = \{\ G \qquad \mid G \text{ SCFG} \qquad\qquad\qquad\qquad\qquad\qquad \}$$

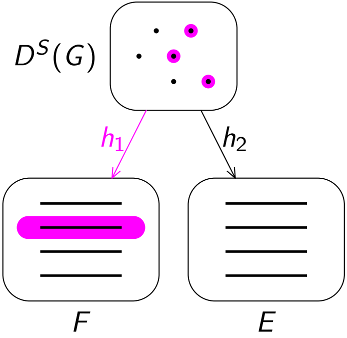
$$\mathbb{D}(G \qquad): f \mapsto$$



Abstrakte Spezifikation
(Chiang 2007)

$$\Omega = \{\ G \qquad \mid G \text{ SCFG} \qquad\qquad\qquad\qquad \}$$

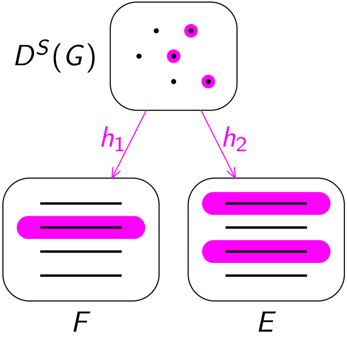
$$\mathbb{D}(G \qquad): f \mapsto$$



Abstrakte Spezifikation
(Chiang 2007)

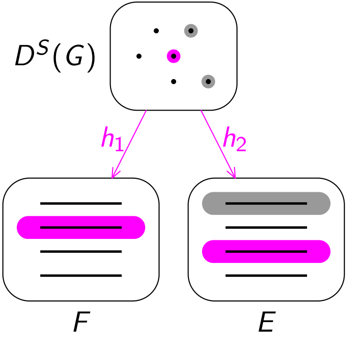
$$\Omega = \{\ G \qquad \mid G \text{ SCFG} \qquad\qquad\qquad\qquad\qquad\qquad \}$$

$$\mathbb{D}(G \qquad) \colon f \mapsto$$



Abstrakte Spezifikation
(Chiang 2007)

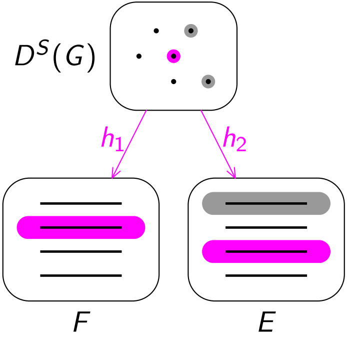
$$\Omega = \{\ G \qquad \mid G \text{ SCFG} \qquad \}$$
$$\mathbb{D}(G \qquad): f \mapsto h_2(\text{select} \qquad (\{d \mid d \in D^S(G), h_1(d) = f\}))$$



Abstrakte Spezifikation

(Chiang 2007)

$$\Omega = \{(G, \mu, \theta) \mid G \text{ SCFG}, \mu \text{ Wahrscheinlichkeitszuweisung}, \theta \in \mathbb{R}^3\}$$
$$\mathbb{D}(G, \mu, \theta)\colon f \mapsto h_2(\text{select}_{G,\mu,\theta}(\{d \mid d \in D^S(G), h_1(d) = f\}))$$

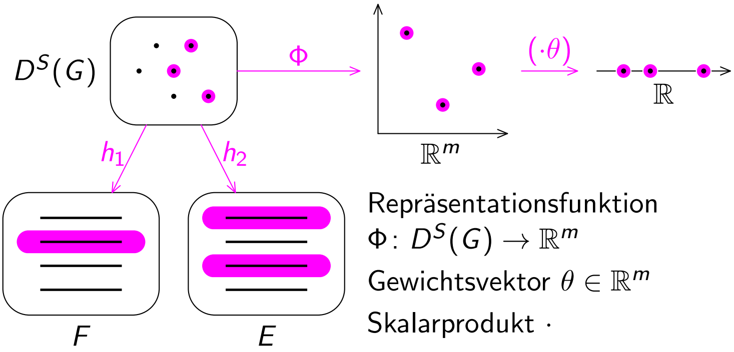


Abstrakte Spezifikation

(Chiang 2007)

$$\Omega = \{(G, \mu, \theta) \mid G \text{ SCFG}, \mu \text{ Wahrscheinlichkeitszuweisung}, \theta \in \mathbb{R}^3\}$$

$$\mathbb{D}(G, \mu, \theta)\colon f \mapsto h_2(\mathrm{select}_{G,\mu,\theta}(\{d \mid d \in D^S(G), h_1(d) = f\}))$$



Repräsentationsfunktion

$\Phi\colon D^S(G) \to \mathbb{R}^m$

Gewichtsvektor $\theta \in \mathbb{R}^m$

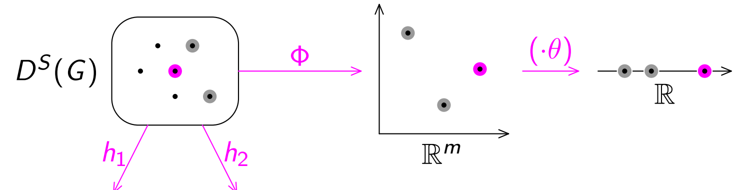
Skalarprodukt $\cdot$

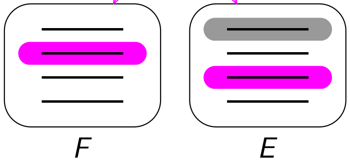
Abstrakte Spezifikation

(Chiang 2007)

$\Omega = \{(G, \mu, \theta) \mid G \text{ SCFG}, \mu \text{ Wahrscheinlichkeitszuweisung}, \theta \in \mathbb{R}^3\}$

$\mathbb{D}(G, \mu, \theta)\colon f \mapsto h_2(\text{select}_{G,\mu,\theta}(\{d \mid d \in D^S(G), h_1(d) = f\}))$





Repräsentationsfunktion

$\Phi\colon D^S(G) \to \mathbb{R}^m$

Gewichtsvektor $\theta \in \mathbb{R}^m$

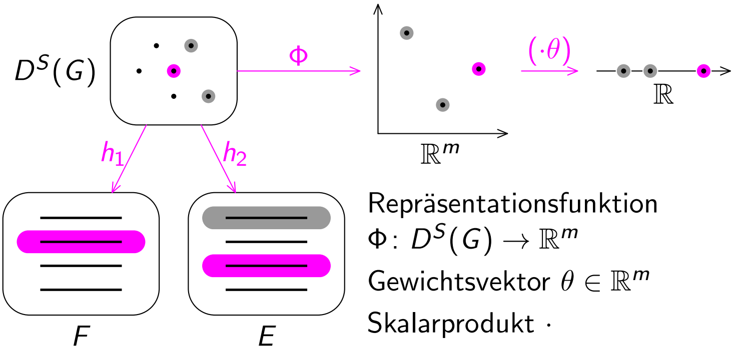
Skalarprodukt $\cdot$

Abstrakte Spezifikation

(Chiang 2007)

$$\Omega = \{(G, \mu, \theta) \mid G \text{ SCFG}, \mu \text{ Wahrscheinlichkeitszuweisung}, \theta \in \mathbb{R}^3\}$$

$$\mathbb{D}(G, \mu, \theta) \colon f \mapsto h_2(\operatorname{argmax}_{d \in D^S(G) \colon h_1(d) = f} \Phi_{G,\mu}(d) \cdot \theta)$$



Repräsentationsfunktion

$\Phi \colon D^S(G) \to \mathbb{R}^m$

Gewichtsvektor $\theta \in \mathbb{R}^m$

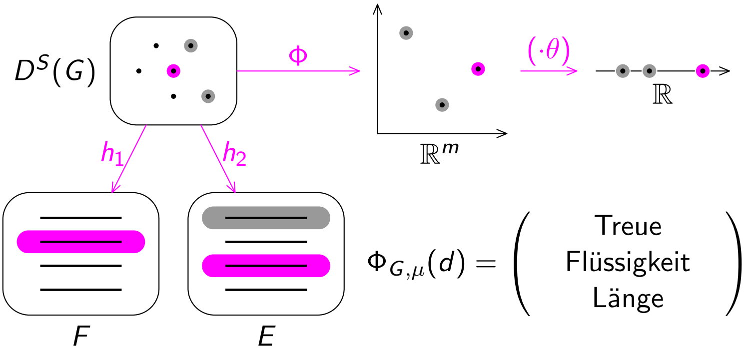
Skalarprodukt $\cdot$

Abstrakte Spezifikation

(Chiang 2007)

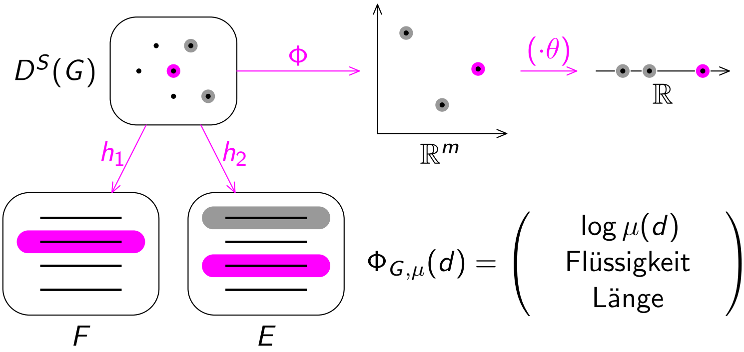
$$\Omega = \{(G, \mu, \theta) \mid G \text{ SCFG}, \mu \text{ Wahrscheinlichkeitszuweisung}, \theta \in \mathbb{R}^3\}$$

$$\mathbb{D}(G, \mu, \theta)\colon f \mapsto h_2(\mathrm{argmax}_{d \in D^S(G)\colon h_1(d) = f}\, \Phi_{G,\mu}(d) \cdot \theta)$$



Abstrakte Spezifikation
(Chiang 2007)

$$\Omega = \{(G, \mu, \theta) \mid G \text{ SCFG}, \mu \text{ Wahrscheinlichkeitszuweisung}, \theta \in \mathbb{R}^3\}$$
$$\mathbb{D}(G, \mu, \theta)\colon f \mapsto h_2(\operatorname{argmax}_{d \in D^S(G)\colon h_1(d)=f} \Phi_{G,\mu}(d) \cdot \theta)$$



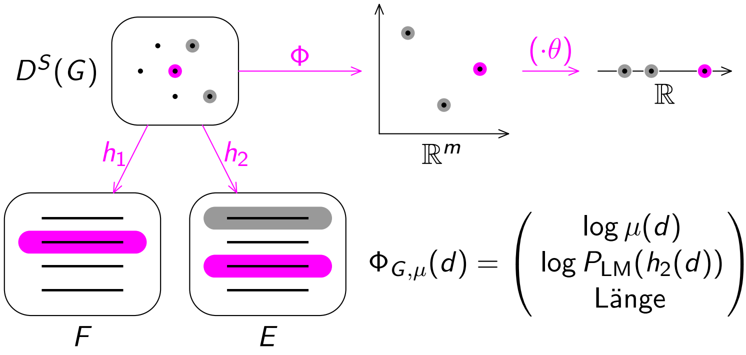
$$\Phi_{G,\mu}(d) = \begin{pmatrix} \log \mu(d) \\ \text{Flüssigkeit} \\ \text{Länge} \end{pmatrix}$$

Abstrakte Spezifikation

(Chiang 2007)

$$\Omega = \{(G, \mu, \theta) \mid G \text{ SCFG}, \mu \text{ Wahrscheinlichkeitszuweisung}, \theta \in \mathbb{R}^3\}$$

$$\mathbb{D}(G, \mu, \theta)\colon f \mapsto h_2(\operatorname{argmax}_{d \in D^S(G)\colon h_1(d) = f} \Phi_{G,\mu}(d) \cdot \theta)$$



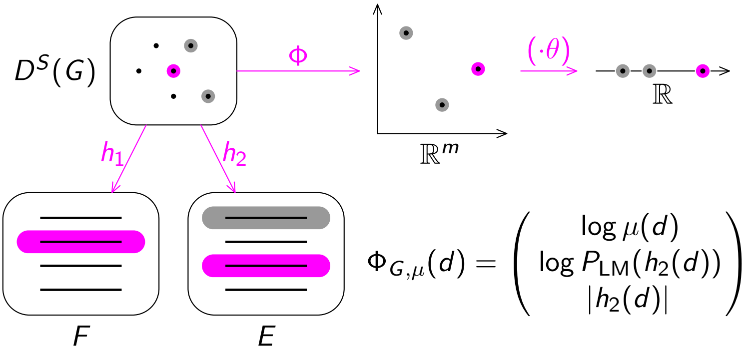
$$\Phi_{G,\mu}(d) = \begin{pmatrix} \log \mu(d) \\ \log P_{\mathrm{LM}}(h_2(d)) \\ \text{Länge} \end{pmatrix}$$

Abstrakte Spezifikation

(Chiang 2007)

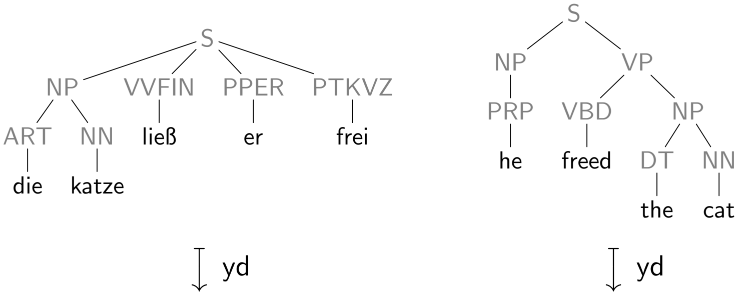
$$\Omega = \{(G, \mu, \theta) \mid G \text{ SCFG}, \mu \text{ Wahrscheinlichkeitszuweisung}, \theta \in \mathbb{R}^3\}$$

$$\mathbb{D}(G, \mu, \theta)\colon f \mapsto h_2(\operatorname{argmax}_{d \in D^S(G)\colon h_1(d) = f} \Phi_{G,\mu}(d) \cdot \theta)$$



$$\Phi_{G,\mu}(d) = \begin{pmatrix} \log \mu(d) \\ \log P_{\mathrm{LM}}(h_2(d)) \\ |h_2(d)| \end{pmatrix}$$

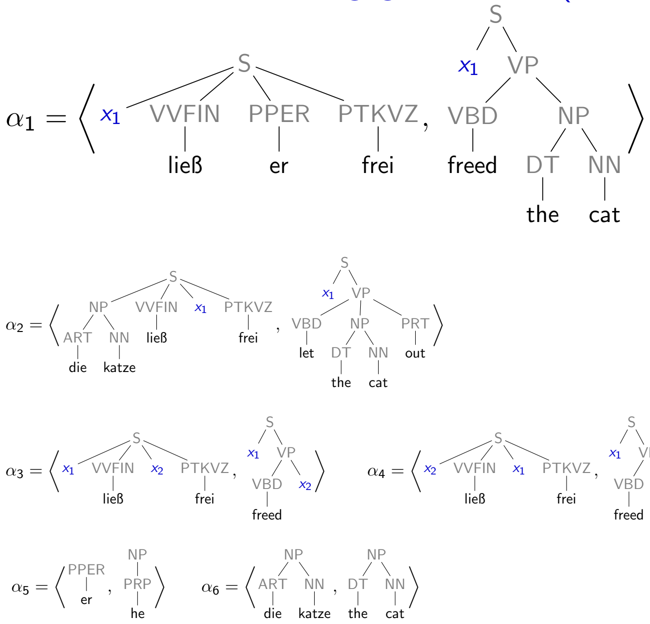
Konstituentenbäume



die katze ließ er frei

he freed the cat

Synchrone Baumersetzungsgrammatik (STSG)



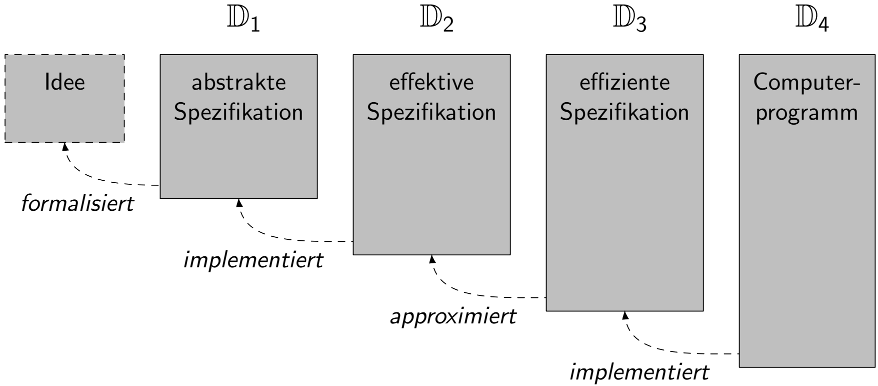
Dekoder mit expliziter Syntax (abstrakte Spezifikation)

$\Omega = \{(G, \mu, \theta) \mid G \text{ STSG}, \mu \text{ Wahrscheinlichkeitszuweisung}, \theta \in \mathbb{R}^3\}$
$\mathbb{D}(G, \mu, \theta)\colon f \mapsto \mathrm{yd}(h_2(\mathrm{argmax}_{d \in D^S(G)\colon \mathrm{yd}(h_1(d)) = f}\ \Phi_{G,\mu}(d) \cdot \theta))$

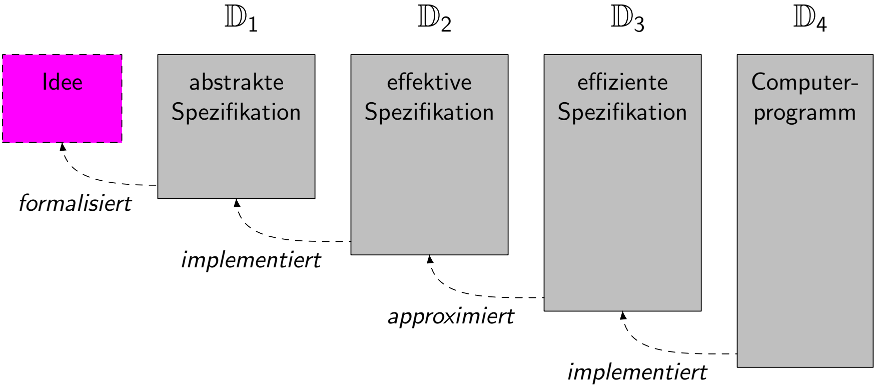
$$\Phi_{G,\mu}(d) = \begin{pmatrix} \log \mu(d) \\ \log P_{\mathrm{LM}}(\mathrm{yd}(h_2(d))) \\ \log P_{\mathrm{parse}}(h_1(d) \mid \mathrm{yd}(h_1(d))) \end{pmatrix}$$

Gliederung

Abstraktionsniveaus (ideal)



Abstraktionsniveaus (ideal)

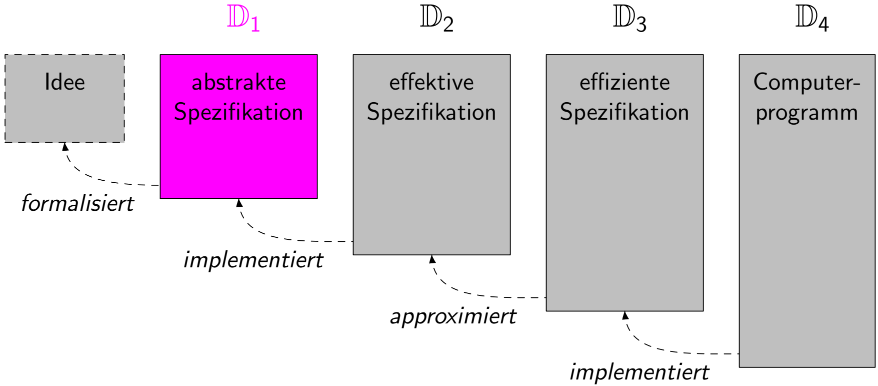


Idee

Prosa, Beispiele

ca. eine Seite

Abstraktionsniveaus (ideal)

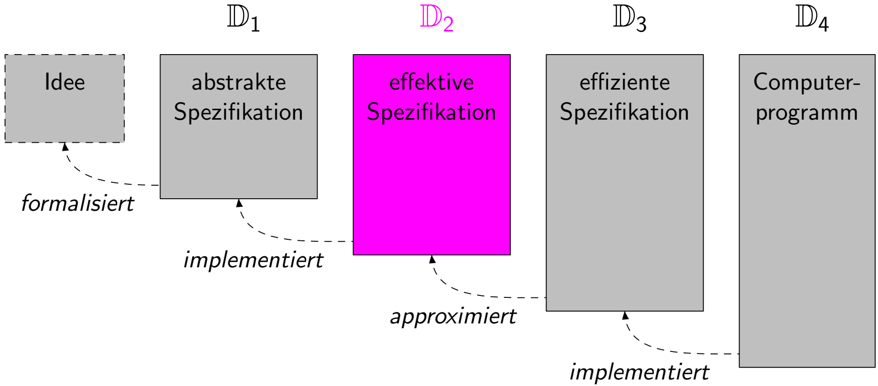


Abstrakte Spezifikation

mathematisch, nicht konstruktiv, nicht eindeutig

wenige Seiten

Abstraktionsniveaus (ideal)

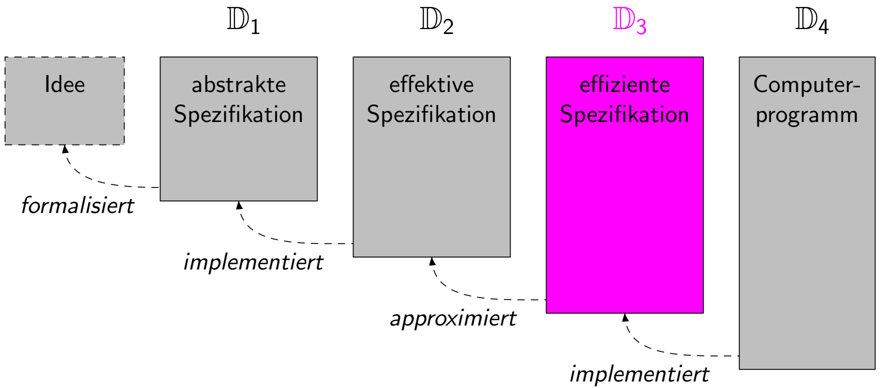


Effektive Spezifikation

mathematisch, suggeriert Algorithmus, ineffizient, nicht eindeutig

ein dünnes Heft

Abstraktionsniveaus (ideal)

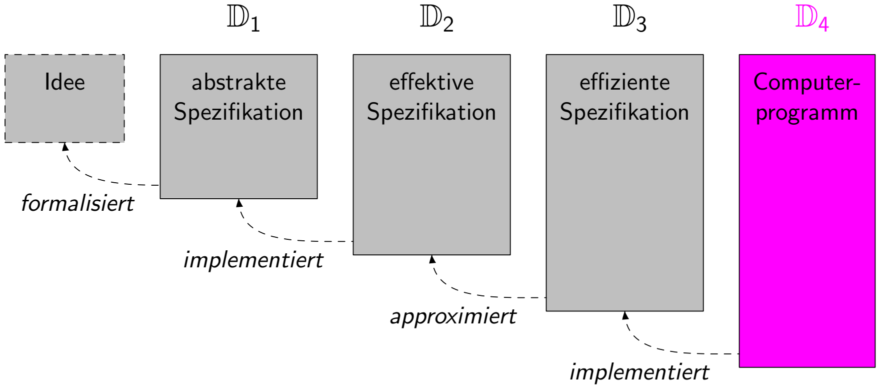


Effiziente Spezifikation

mathematisch, suggeriert Algorithmus, effizient, eindeutig

ein (sehr) dickes Heft

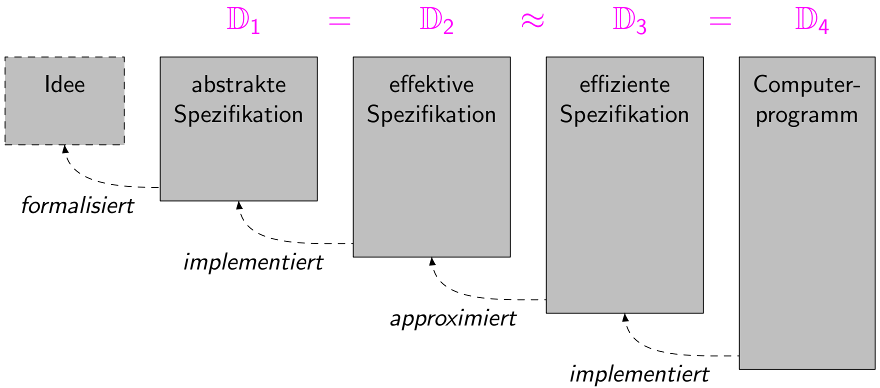
Abstraktionsniveaus (ideal)



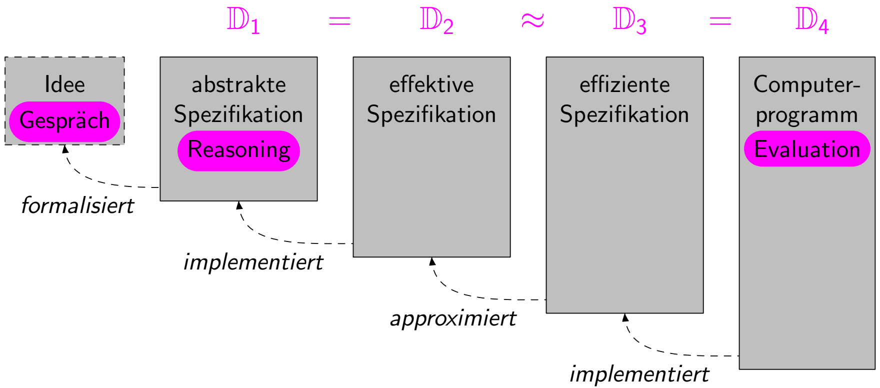
Computerprogramm

Hunderttausende Zeilen Code (C++, Java, Python, Haskell, . . .)

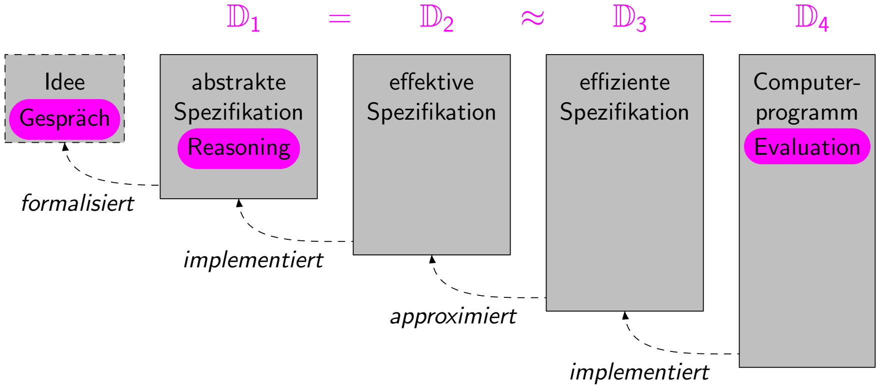
Abstraktionsniveaus (ideal)



Abstraktionsniveaus (ideal)



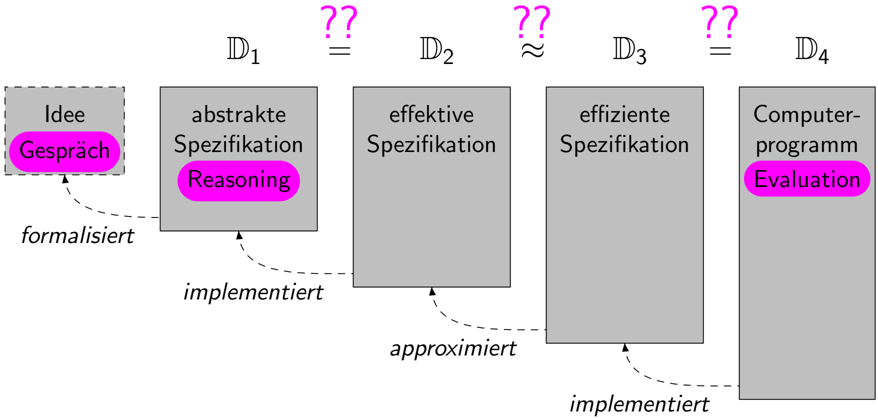
Abstraktionsniveaus (ideal)



Beobachtung

Approximation verschlechtert die Übersetzung (ein wenig)
(Chang und Collins 2011; Chiang 2007; Rush und Collins 2011)

Abstraktionsniveaus (ideal)

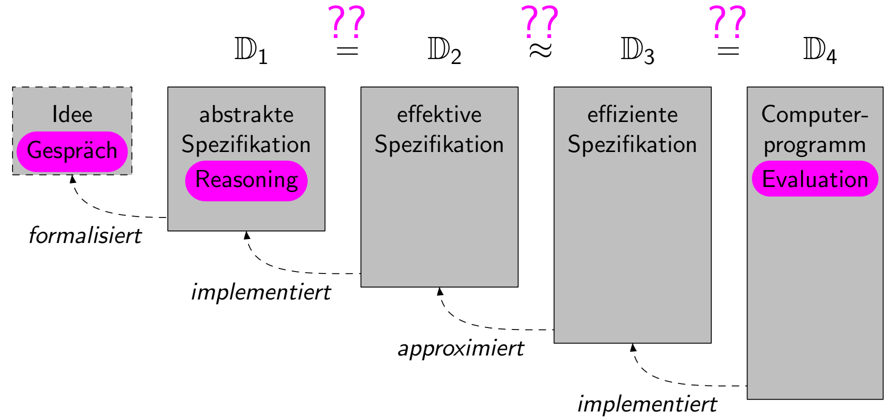


Umsetzung

$\mathbb{D}_2$, $\mathbb{D}_3$ operational, skizzenhaft

Beziehungen werden lediglich suggeriert

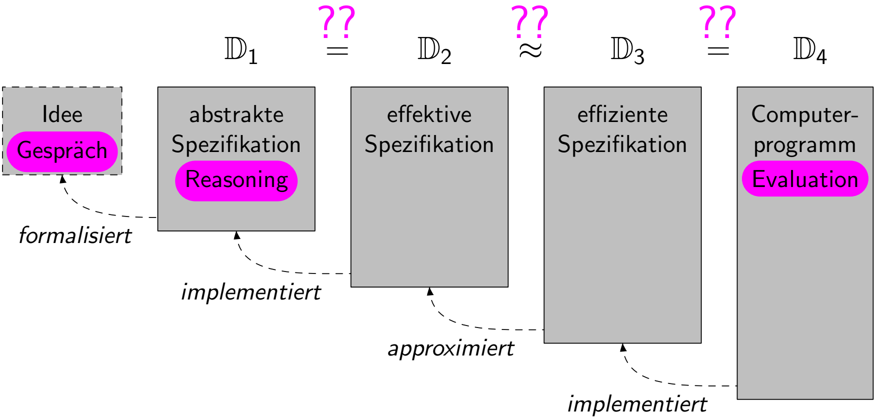
Abstraktionsniveaus (ideal)



Problem

Kompliziert, undurchsichtig, fehlerträchtig,
kaum wiederverwendbar, schwer zu erlernen!

Abstraktionsniveaus (ideal)



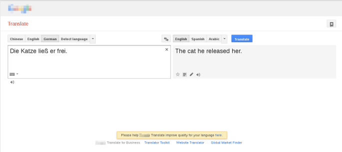
Unser Ziel

Dekoder besser spezifizieren!

Hier nicht: Gute Dekoder entwickeln.

Übersicht

Statistisches Maschinelles Übersetzen



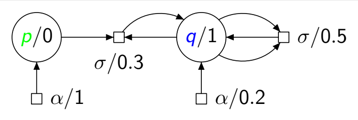
✓

2

Dekoder-spezifikation

algebraische Dekoder-spezifikation

Theorie der formalen Sprachen



3

Abschlusseigenschaften

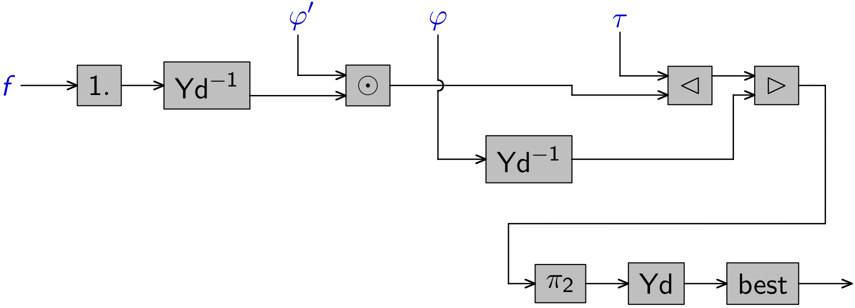
Binarisierung

Determinisierung

STSG-basierter Dekoder

$\Omega = \mathcal{T}_{\mathrm{STSG}} \times \mathcal{K}_{\mathrm{Rec}} \times \mathcal{L}_{\mathrm{Rec}}$, $\mathcal{S} = (\mathbb{R}_{\geq 0}, \max, \cdot, 0, 1)$

$\mathbb{D}(\tau, \varphi, \varphi')\colon \quad f \mapsto \mathrm{best}(\mathrm{Yd}(\pi_2((\mathrm{Yd}^{-1}(1.f) \odot \varphi') \lhd \tau \rhd \mathrm{Yd}^{-1}(\varphi))))$



$\mathcal{K} = (\mathbb{R}_{\geq 0})^{\Sigma^*}$ (gewichtete Stringsprachen)

$\mathcal{L} = (\mathbb{R}_{\geq 0})^{T_\Sigma}$ (gewichtete Baumsprachen)

$\mathcal{T} = (\mathbb{R}_{\geq 0})^{T_\Sigma \times T_\Sigma}$ (gewichtete Baumübersetzungen)

Beispiel f

die katze ließ er frei

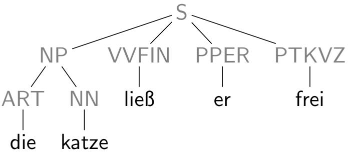
Stringinjektion 1.: $\Sigma^* \to \mathcal{K}$

1.f:

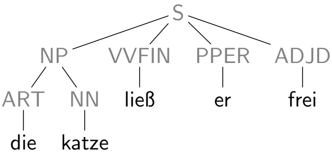
$$\begin{aligned} \ldots &\mapsto 0 \\ \text{die katze ließ er frei} &\mapsto 1 \\ \cdots &\mapsto 0 \end{aligned}$$

Sprachfrontinverse $\mathrm{Yd}^{-1}\colon \mathcal{K} \to \mathcal{L}$

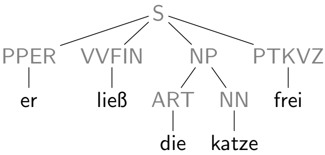
$\mathrm{Yd}^{-1}(1.f)$:



$\mapsto \quad 1$



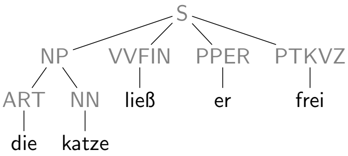
$\mapsto \quad 1$



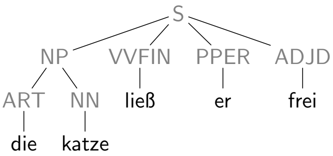
$\mapsto \quad 0$

$\cdots \mapsto \quad 0$

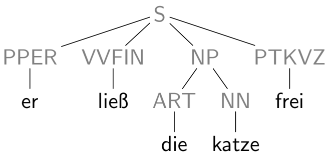
Beispiel φ'



$\mapsto \quad 0.5$



$\mapsto \quad 0.1$

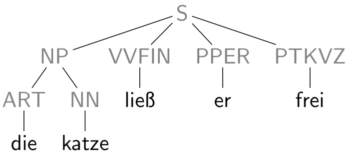


$\mapsto \quad 0.2$

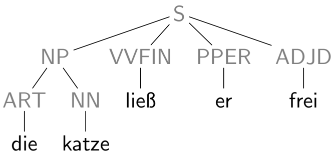
$\cdots \quad \mapsto \quad \cdots$

Hadamardprodukt $\odot$: $\mathcal{L} \times \mathcal{L} \to \mathcal{L}$

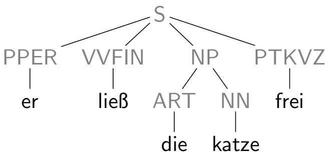
$\mathrm{Yd}^{-1}(1.f) \odot \varphi'$:



$\mapsto$ 0.5



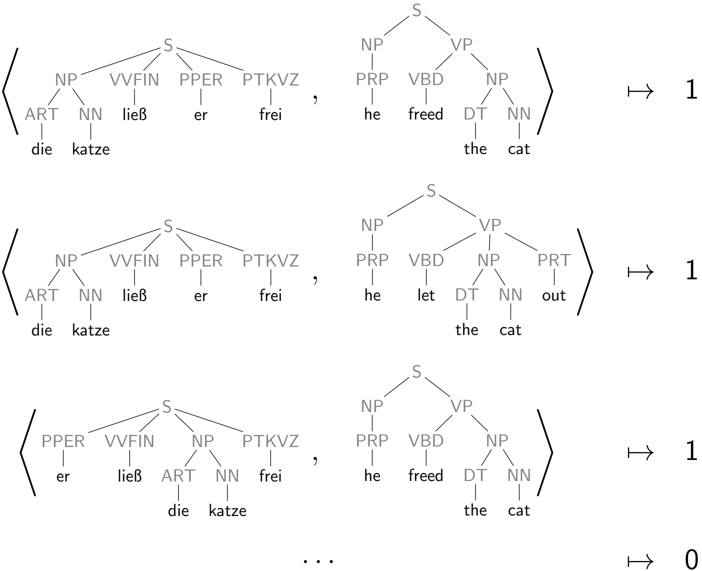
$\mapsto$ 0.1



$\mapsto$ 0

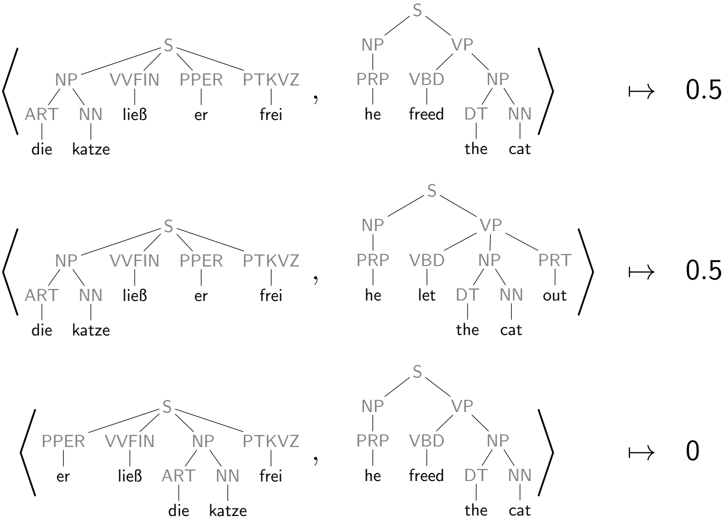
$\cdots \mapsto 0$

Beispiel τ



Eingabeprodukt $\lhd\colon \mathcal{L} \times \mathcal{T} \to \mathcal{T}$

$(\mathrm{Yd}^{-1}(1.f) \odot \varphi') \lhd \tau$:



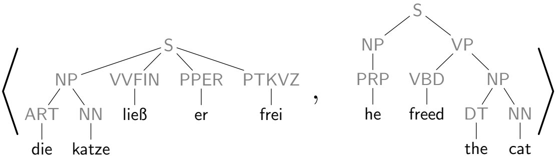
$\ldots \qquad \mapsto \quad 0$

Beispiel φ

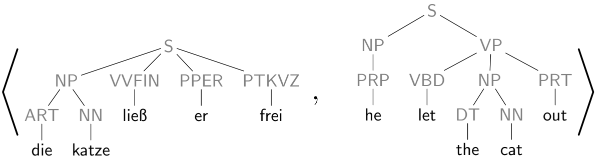
$$\begin{aligned} \text{he freed the cat} &\mapsto 0.5 \\ \text{he let the cat out} &\mapsto 0.1 \\ \dots &\mapsto \dots \end{aligned}$$

Ausgabeprodukt $\triangleright : \mathcal{T} \times \mathcal{L} \to \mathcal{T}$

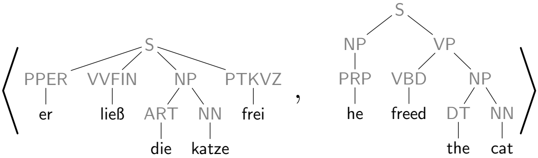
$(\mathrm{Yd}^{-1}(1.f) \odot \varphi') \triangleleft \tau \triangleright \mathrm{Yd}^{-1}(\varphi)$:



$\mapsto$ 0.25



$\mapsto$ 0.05

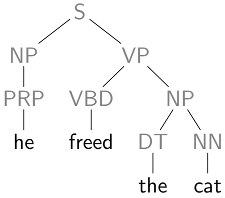


$\mapsto$ 0

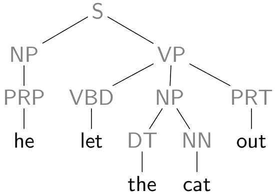
$\ldots \qquad \mapsto$ 0

Ausgabeprojektion $\pi_2 \colon \mathcal{T} \to \mathcal{L}$

$$\pi_2\big((\mathrm{Yd}^{-1}(1.f) \odot \varphi') \lhd \tau \rhd \mathrm{Yd}^{-1}(\varphi)\big):$$



$\mapsto \quad 0.25$



$\mapsto \quad 0.05$

$\cdots \mapsto \quad 0$

Sprachfront Yd: $\mathcal{L} \to \mathcal{K}$

$\mathsf{Yd}\big(\pi_2\big((\mathsf{Yd}^{-1}(1.f) \odot \varphi') \lhd \tau \rhd \mathsf{Yd}^{-1}(\varphi)\big)\big)$:

$$
\begin{aligned}
\text{he freed the cat} &\mapsto 0.25 \\
\text{he let the cat out} &\mapsto 0.05 \\
\ldots &\mapsto 0
\end{aligned}
$$

Bester-Index-Operation best: $\mathcal{K} \to \Sigma^*$

$\mathsf{best}(\mathrm{Yd}(\pi_2\big((\mathrm{Yd}^{-1}(1.f) \odot \varphi') \lhd \tau \rhd \mathrm{Yd}^{-1}(\varphi)\big))):$

he freed the cat

Effektivität durch endliche Darstellung (vgl. Tabelle 1.2)

Abschlusseigenschaft	Referenzen	Komplexität
$1.(\Sigma^*) \subseteq \mathcal{K}_{\mathrm{Rec}}$	(Berstel und Reutenauer 1988; Schützenberger 1961)	$O(n)$
$\mathrm{Yd}(\mathcal{L}_{\mathrm{Rec}}) \subseteq \mathcal{K}_{\mathrm{CF}}$	(Thatcher 1967; Ésik und Kuich 2003)	$O(r)$
$\mathrm{Yd}^{-1}(\mathcal{K}_{\mathrm{Rec}}) \subseteq \mathcal{L}_{\mathrm{Rec}}$	(Maletti und Satta 2009)	$O(p^k)$
$\mathcal{L}_{\mathrm{Rec}} \odot \mathcal{L}_{\mathrm{Rec}} \subseteq \mathcal{L}_{\mathrm{Rec}}$	(Borchardt 2004, Cor. 3.9)	$O(r_1 \cdot r_2)$
$\mathcal{L}_{\mathrm{Rec}} \lhd \mathcal{T}_{\mathrm{STSG}} \subseteq \mathcal{T}_{\mathrm{STSG}}$	(Maletti 2010)	$O(r_2 \cdot p_1^{k_2})$
$\mathcal{T}_{\mathrm{STSG}} \rhd \mathcal{L}_{\mathrm{Rec}} \subseteq \mathcal{T}_{\mathrm{STSG}}$	(Maletti 2010)	$O(r_1 \cdot p_2^{k_1})$
$\pi_2(\mathcal{T}_{\mathrm{STSG}}) \subseteq \mathcal{L}_{\mathrm{Rec}}$	(Fülöp u. a. 2011)	$O(r)$
$\mathrm{best}(\mathcal{K}_{\mathrm{CF}}) \subseteq \Sigma^*$	(Knuth 1977; Huang und Chiang 2005; Büchse u. a. 2010)	$O(r \cdot \log p)$

Effektivität durch endliche Darstellung (vgl. Tabelle 1.2)

Abschlusseigenschaft	Referenzen	Komplexität
$1.(\Sigma^*) \subseteq \mathcal{K}_{\text{Rec}}$	(Berstel und Reutenauer 1988; Schützenberger 1961)	$O(n)$
$\text{Yd}(\mathcal{L}_{\text{Rec}}) \subseteq \mathcal{K}_{\text{CF}}$	(Thatcher 1967; Ésik und Kuich 2003)	$O(r)$
$\text{Yd}^{-1}(\mathcal{K}_{\text{Rec}}) \subseteq \mathcal{L}_{\text{Rec}}$	(Maletti und Satta 2009)	$O(p^k)$
$\mathcal{L}_{\text{Rec}} \odot \mathcal{L}_{\text{Rec}} \subseteq \mathcal{L}_{\text{Rec}}$	(Borchardt 2004, Cor. 3.9)	$O(r_1 \cdot r_2)$
$\mathcal{L}_{\text{Rec}} \lhd \mathcal{T}_{\text{STSG}} \subseteq \mathcal{T}_{\text{STSG}}$	(Maletti 2010)	$O(r_2 \cdot p_1^{k_2})$
$\mathcal{T}_{\text{STSG}} \rhd \mathcal{L}_{\text{Rec}} \subseteq \mathcal{T}_{\text{STSG}}$	(Maletti 2010)	$O(r_1 \cdot p_2^{k_1})$
$\pi_2(\mathcal{T}_{\text{STSG}}) \subseteq \mathcal{L}_{\text{Rec}}$	(Fülöp u. a. 2011)	$O(r)$
$\text{best}(\mathcal{K}_{\text{CF}}) \subseteq \Sigma^*$	(Knuth 1977; Huang und Chiang 2005; Büchse u. a. 2010)	$O(r \cdot \log p)$

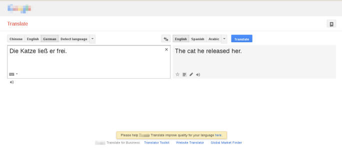
Effektivität durch endliche Darstellung (vgl. Tabelle 1.2)

Abschlusseigenschaft	Referenzen	Komplexität
$1.(\Sigma^*) \subseteq \mathcal{K}_{\text{Rec}}$	(Berstel und Reutenauer 1988; Schützenberger 1961)	$O(n)$
$\text{Yd}(\mathcal{L}_{\text{Rec}}) \subseteq \mathcal{K}_{\text{CF}}$	(Thatcher 1967; Ésik und Kuich 2003)	$O(r)$
$\text{Yd}^{-1}(\mathcal{K}_{\text{Rec}}) \subseteq \mathcal{L}_{\text{Rec}}$	(Maletti und Satta 2009)	$O(p^k)$
$\mathcal{L}_{\text{Rec}} \odot \mathcal{L}_{\text{Rec}} \subseteq \mathcal{L}_{\text{Rec}}$	(Borchardt 2004, Cor. 3.9)	$O(r_1 \cdot r_2)$
$\mathcal{L}_{\text{Rec}} \triangleleft \mathcal{T}_{\text{STSG}} \subseteq \mathcal{T}_{\text{STSG}}$	(Maletti 2010)	$O(r_2 \cdot p_1^{k_2})$
$\mathcal{T}_{\text{STSG}} \triangleright \mathcal{L}_{\text{Rec}} \subseteq \mathcal{T}_{\text{STSG}}$	(Maletti 2010)	$O(r_1 \cdot p_2^{k_1})$
$\pi_2(\mathcal{T}_{\text{STSG}}) \subseteq \mathcal{L}_{\text{Rec}}$	(Fülöp u. a. 2011)	$O(r)$
$\text{best}(\mathcal{K}_{\text{CF}}) \subseteq \Sigma^*$	(Knuth 1977; Huang und Chiang 2005; Büchse u. a. 2010)	$O(r \cdot \log p)$

Kopplung zwischen dem Statistischen Maschinellen Übersetzen und der Theorie der formalen Sprachen!

Übersicht

Statistisches Maschinelles Übersetzen

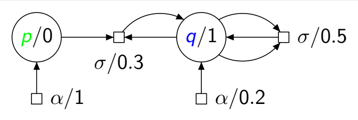


✓ ✓

Dekoder-spezifikation

algebraische Dekoder-spezifikation

Theorie der formalen Sprachen



3

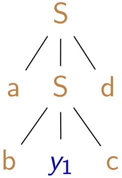
Abschlusseigenschaften

Binarisierung

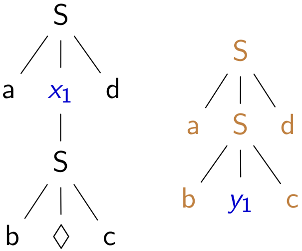
Determinisierung

Gliederung

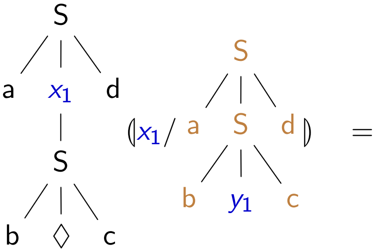
Substitution zweiter Ordnung



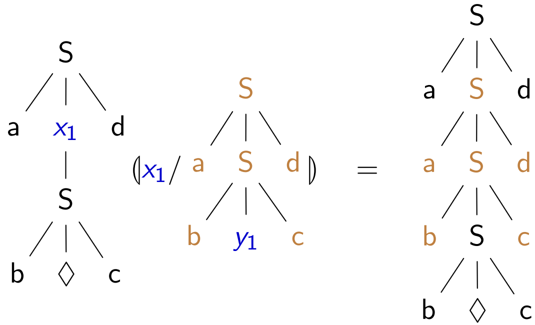
Substitution zweiter Ordnung



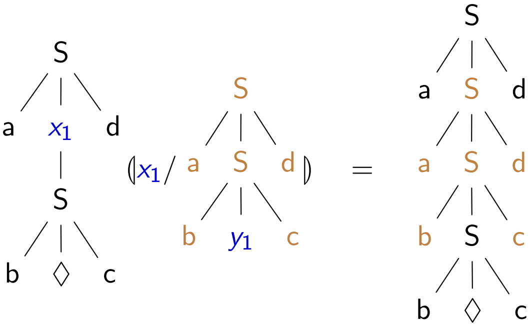
Substitution zweiter Ordnung



Substitution zweiter Ordnung



Substitution zweiter Ordnung



Wichtig, um Phänomene der realen Welt zu beschreiben!

(Shieber 2007; Kallmeyer u. a. 2009; Gildea 2010; Kaeshammer 2013)

Gewichtete synchrone kontextfreie Baumgrammatik (WSCFTG)

$$\mathrm{rk}(q) = 0, \quad \mathrm{rk}(f) = 1$$

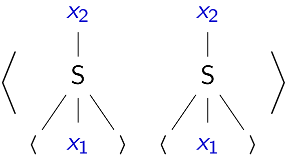
$\rho_1\colon \quad q \to \alpha_1(q, f) \quad \#\ 0.3$

$\rho_2\colon \quad q \to \alpha_2() \quad \#\ 1$

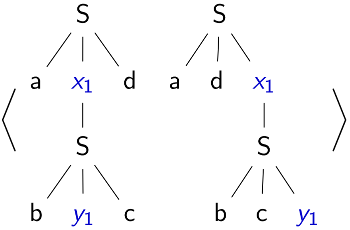
$\rho_3\colon \quad f \to \alpha_3(f) \quad \#\ 0.7$

$\rho_4\colon \quad f \to \alpha_4() \quad \#\ 1$

$\alpha_1 =$



$\alpha_3 =$

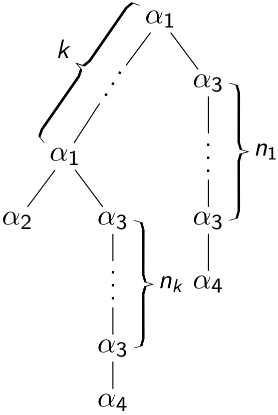


$\alpha_2 = \langle\ \Diamond \quad \Diamond\ \rangle$

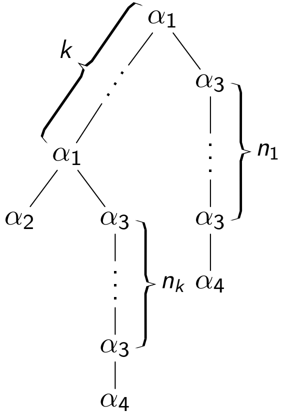
$\alpha_4 = \langle\ y_1 \quad y_1\ \rangle$

Zentralbäume und abgeleitete Bäume

(Arnold und Dauchet 1976)



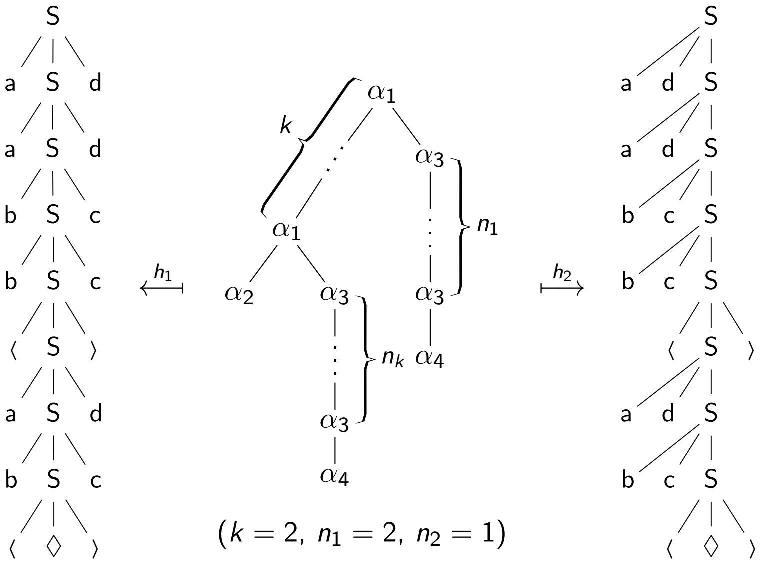
Zentralbäume und abgeleitete Bäume
(Arnold und Dauchet 1976)



$$\left(k = 2,\ n_1 = 2,\ n_2 = 1\right)$$

Zentralbäume und abgeleitete Bäume

(Arnold und Dauchet 1976)



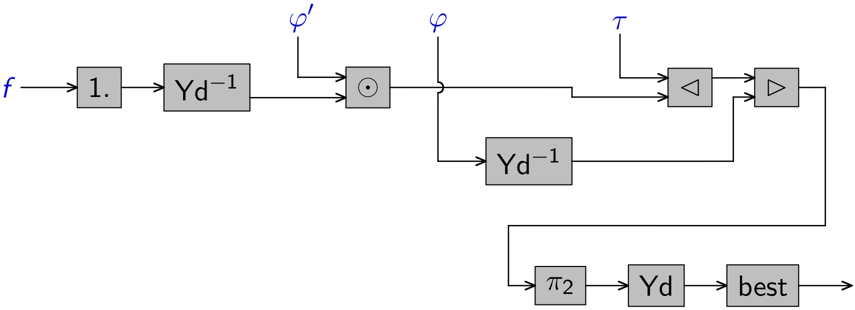
Teilklassen von WSCFTG für Dekoder

- Synchrone Baumersetzungsgrammatik (STSG) (Eisner 2003),
- Synchrone Baumadjunktionsgrammatik (STAG) (Shieber und Schabes 1990; Abeillé u. a. 1990),
- Synchrone Baumeinfügungsgrammatik (STIG) (Schabes und Waters 1994; Nesson 2009; DeNeefe 2011).

WSCFTG-basierter Dekoder

$\Omega = \mathcal{T}_{\text{WSCFTG}} \times \mathcal{K}_{\text{Rec}} \times \mathcal{L}_{\text{Rec}}$, $\mathcal{S} = (\mathbb{R}_{\geq 0}, \max, \cdot, 0, 1)$

$\mathbb{D}(\tau, \varphi, \varphi')$: $f \mapsto \text{best}(\text{Yd}(\pi_2((\text{Yd}^{-1}(1.f) \odot \varphi') \lhd \tau \rhd \text{Yd}^{-1}(\varphi))))$

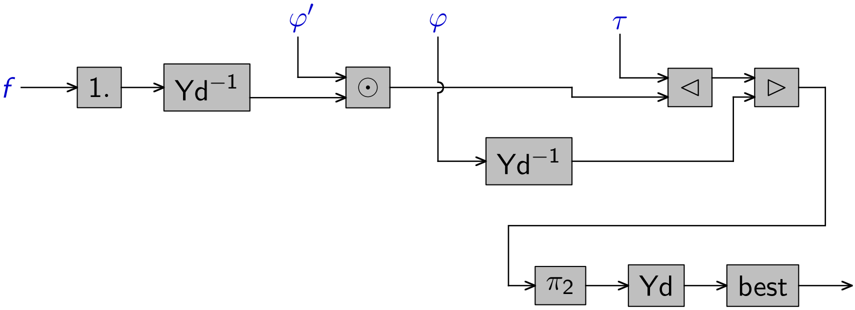


Effektiv? Tabelle erweitern!

WSCFTG-basierter Dekoder

$\Omega = \mathcal{T}_{\text{WSCFTG}} \times \mathcal{K}_{\text{Rec}} \times \mathcal{L}_{\text{Rec}}$, $\qquad \mathcal{S} = (\mathbb{R}_{\geq 0}, \max, \cdot, 0, 1)$

$\mathbb{D}(\tau, \varphi, \varphi')$: $\quad f \mapsto \text{best}(\text{Yd}(\pi_2((\text{Yd}^{-1}(1.f) \odot \varphi') \lhd \tau \rhd \text{Yd}^{-1}(\varphi))))$



Effektiv? Tabelle erweitern! Beweisen der Inklusionen

$$\mathcal{L}_{\text{Rec}} \lhd \mathcal{T}_{\text{WSCFTG}} \subseteq \mathcal{T}_{\text{WSCFTG}} , \quad \mathcal{T}_{\text{WSCFTG}} \rhd \mathcal{L}_{\text{Rec}} \subseteq \mathcal{T}_{\text{WSCFTG}} ,$$

gegebenenfalls unter geeigneten Voraussetzungen.

Ergebnisse I: Abschlusseigenschaft

Theorem 3.3.3

Für kommutatives $\mathcal{S}$ gilt

$$\mathcal{L}_{\mathsf{Rec}} \lhd \mathcal{T}_{\mathsf{WSCFTG}} \subseteq \mathcal{T}_{\mathsf{WSCFTG}}\,, \quad \mathcal{T}_{\mathsf{WSCFTG}} \rhd \mathcal{L}_{\mathsf{Rec}} \subseteq \mathcal{T}_{\mathsf{WSCFTG}}\,.$$

Ergebnisse I: Abschlusseigenschaft

Theorem 3.3.3

Für kommutatives $\mathcal{S}$ gilt

$$\mathcal{L}_{\mathsf{Rec}} \lhd \mathcal{T}_{\mathsf{WSCFTG}} \;\subseteq\; \mathcal{T}_{\mathsf{WSCFTG}}\,, \quad \mathcal{T}_{\mathsf{WSCFTG}} \rhd \mathcal{L}_{\mathsf{Rec}} \;\subseteq\; \mathcal{T}_{\mathsf{WSCFTG}}\,.$$

Lemma 3.3.2 *There is effectively an admissible WSCFTG* $M \lhd G = (Q', R', \mu', \nu')$ *over* Σ *and* $\mathcal{S}$ *such that* $M \lhd G$ *is also a WTA over* Γ *and* $\mathcal{S}$*,* $Q' = \bigcup_m Q^{(m)} \times P \times P^m$ *with the ranks carried over from* Q*,* $\nu'_{(q,p,\varepsilon)} = \nu_q \cdot (\nu_M)_p$*, and the following holds. Let* $\xi \in T_\Gamma$ *be type conformant,* $s = h_1(\xi)$*, and* $s \in C_\Sigma(m)$*. Then there are families* $(\equiv_{(p,p')} \mid p \in P, p' \in P^m)$ *and* $(\pi_{q'} \mid q' \in Q'^{(m)})$ *such that*

- $\equiv_{(p,p')}$ *is an equivalence relation on* $D^{(p,p')}(M, s)$*,*
- $\pi_{(q,p,p')} \colon D^{(q,p,p')}(M \lhd G, \xi) \to D^q(G, \xi) \times D^{(p,p')}(M, s)/_{\equiv_{(p,p')}}$ *is bijective,*
- $\pi_{q'}(d') = (d, D)$ *implies* $\langle d' \rangle = \langle d \rangle \cdot \sum_{e \in D} \langle e \rangle$*.*

Ergebnisse I: Abschlusseigenschaft

Theorem 3.3.3

Für kommutatives $\mathcal{S}$ gilt

$$\mathcal{L}_{\text{Rec}} \lhd \mathcal{T}_{\text{WSCFTG}} \subseteq \mathcal{T}_{\text{WSCFTG}}, \quad \mathcal{T}_{\text{WSCFTG}} \rhd \mathcal{L}_{\text{Rec}} \subseteq \mathcal{T}_{\text{WSCFTG}}.$$

Lemma 3.3.2 *There is effectively an admissible WSCFTG $M \lhd G = (Q', R', \mu', \nu')$ over Σ and $\mathcal{S}$ such that $M \lhd G$ is also a WTA over Γ and $\mathcal{S}$, $Q' = \bigcup_m Q^{(m)} \times P \times P^m$ with the ranks carried over from Q, $\nu'_{(q,p,\varepsilon)} = \nu_q \cdot (\nu_M)_p$, and the following holds. Let $\xi \in T_\Gamma$ be type conformant, $s = h_1(\xi)$, and $s \in C_\Sigma(m)$. Then there are families $(\equiv_{(p,p')} \mid p \in P, p' \in P^m)$ and $(\pi_{q'} \mid q' \in Q'^{(m)})$ such that*

- $\equiv_{(p,p')}$ *is an equivalence relation on* $D^{(p,p')}(M, s)$,
- $\pi_{(q,p,p')} \colon D^{(q,p,p')}(M \lhd G, \xi) \to D^q(G, \xi) \times D^{(p,p')}(M, s)/_{\equiv_{(p,p')}}$ *is bijective,*
- $\pi_{q'}(d') = (d, D)$ *implies* $\langle d' \rangle = \langle d \rangle \cdot \sum_{e \in D} \langle e \rangle$.

Zeitkomplexität der Konstruktion: $O(|R| \cdot |P|^C)$ mit
$C = \max\{\mathrm{rk}(q_0) + \cdots + \mathrm{rk}(q_l) + l + 1 \mid (q_1 \cdots q_l, \alpha, q_0) \in R\}$.

Ergebnisse II: Algorithmus

Algorithmus 3.1

Auf Earley-Parsing (Earley 1970) basierender Algorithmus, der $M \lhd G$ berechnet und dabei nutzlose Transitionsregeln vermeidet.

Ergebnisse II: Algorithmus

Algorithmus 3.1

Auf Earley-Parsing (Earley 1970) basierender Algorithmus, der $M \lhd G$ berechnet und dabei nutzlose Transitionsregeln vermeidet.

(1) $$\frac{}{(q_0, p_0)}$$

(2) $$\frac{(q, p)}{(\rho, \varepsilon, p)}$$

(3) $$\frac{(q, p) \qquad [\rho, \varepsilon, p, \theta]}{[q, p, p']} \;\Big\{\; p' = (\theta(y_1), \ldots, \theta(y_m))$$

(4) $$\frac{(\rho, w, p)}{[\rho, w, 0, p, p']} \;\Big\{\; (p', \zeta(w), p) \in R_M^{(\mathrm{rk}_\zeta(w))}$$

(5) $$\frac{[\rho, w, j, p, (p_1, \ldots, p_k)]}{(\rho, w(j+1), p_{j+1})} \;\Big\{\; 0 \le j < k$$

(6) $$\frac{[\rho, w, j, p, (p_1, \ldots, p_k)] \qquad [\rho, w(j+1), p_{j+1}]}{[\rho, w, j+1, p, (p_1, \ldots, p_k)]} \;\Big\{\; 0 \le j < k$$

(7) $$\frac{[\rho, w, p, \theta]}{[\rho, w, p]}$$

(8) $$\frac{(\rho, w, p)}{(q_\iota, p)} \;\Big\{\; \zeta(w) = x_\iota$$

(9) $$\frac{(\rho, w, p) \qquad [q_\iota, p, p']}{[\rho, w, 0, p, p']} \;\Big\{\; \zeta(w) = x_\iota$$

(10) $$\frac{(\rho, w, p)}{[\rho, w, p, \{y_\iota \mapsto p\}]} \;\Big\{\; \zeta(w) = y_\iota$$

(11) $$\frac{[\rho, w, k, p, p'] \qquad \begin{array}{c}[\rho, w1, p_1, \theta_1] \\ \cdots\ [\rho, wk, p_k, \theta_k]\end{array}}{[\rho, w, p, \theta \cup \theta_1 \cup \cdots \cup \theta_k]} \;\Big\{\; p' = (p_1, \ldots, p_k)$$

where $\theta = \{\zeta(w) \mapsto (p, p')\}$ if $\zeta(w) \in X$, and
$\theta = \emptyset$ otherwise

Note: we assume that $\rho \in R$, $\rho = (q_1 \cdots q_l, \langle \zeta\zeta' \rangle, q)$, and $q \in Q^{(m)}$.

Ergebnisse II: Algorithmus

Algorithmus 3.1

Auf Earley-Parsing (Earley 1970) basierender Algorithmus, der $M \lhd G$ berechnet und dabei nutzlose Transitionsregeln vermeidet.

Theorem 3.4.7

Der Algorithmus ist korrekt und vollständig.

(Er berechnet keine falschen Regeln und mindestens die nützlichen.)

$$(1)\quad \frac{}{(q_0, p_0)}$$

$$(2)\quad \frac{(q,p)}{(\rho,\varepsilon,p)}$$

$$(3)\quad \frac{(q,p) \qquad [\rho,\varepsilon,p,\theta]}{[q,p,p']} \ \left\{\ p' = (\theta(y_1),\dots,\theta(y_m))\right.$$

$$(4)\quad \frac{(\rho,w,p)}{[\rho,w,0,p,p']} \ \left\{\ (p',\zeta(w),p) \in R_M^{(\mathrm{rk}_\zeta(w))}\right.$$

$$(5)\quad \frac{[\rho,w,j,p,(p_1,\dots,p_k)]}{(\rho,w(j+1),p_{j+1})} \ \left\{\ 0 \le j < k\right.$$

$$(6)\quad \frac{[\rho,w,j,p,(p_1,\dots,p_k)] \qquad [\rho,w(j+1),p_{j+1}]}{[\rho,w,j+1,p,(p_1,\dots,p_k)]} \ \left\{\ 0 \le j < k\right.$$

$$(7)\quad \frac{[\rho,w,p,\theta]}{[\rho,w,p]}$$

$$(8)\quad \frac{(\rho,w,p)}{(q_\iota,p)} \ \left\{\ \zeta(w) = x_\iota\right.$$

$$(9)\quad \frac{(\rho,w,p) \qquad [q_\iota,p,p']}{[\rho,w,0,p,p']} \ \left\{\ \zeta(w) = x_\iota\right.$$

$$(10)\quad \frac{(\rho,w,p)}{[\rho,w,p,\{y_\iota \mapsto p\}]} \ \left\{\ \zeta(w) = y_\iota\right.$$

$$(11)\quad \frac{[\rho,w,k,p,p'] \qquad \begin{matrix}[\rho,w1,p_1,\theta_1] \\ \cdots\ [\rho,wk,p_k,\theta_k]\end{matrix}}{[\rho,w,p,\theta\cup\theta_1\cup\dots\cup\theta_k]} \ \left\{\ p' = (p_1,\dots,p_k)\right.$$

where $\theta = \{\zeta(w) \mapsto (p,p')\}$ if $\zeta(w) \in X$, and
$\theta = \emptyset$ otherwise

Note: we assume that $\rho \in R$, $\rho = (q_1 \cdots q_l, \langle\zeta\zeta'\rangle, q)$, and $q \in Q^{(m)}$.

Ergebnisse II: Algorithmus

Algorithmus 3.1

Auf Earley-Parsing (Earley 1970) basierender Algorithmus, der $M \lhd G$ berechnet und dabei nutzlose Transitionsregeln vermeidet.

Theorem 3.4.7

Der Algorithmus ist korrekt und vollständig.

Zeitkomplexität des Algorithmus:

$O(|G|_{\text{in}} \cdot |R_M| \cdot |P|^C)$

$$(1) \quad \frac{}{(q_0, p_0)}$$

$$(2) \quad \frac{(q, p)}{(\rho, \varepsilon, p)}$$

$$(3) \quad \frac{(q, p) \qquad [\rho, \varepsilon, p, \theta]}{[q, p, p']} \; \{ \; p' = (\theta(y_1), \ldots, \theta(y_m))$$

$$(4) \quad \frac{(\rho, w, p)}{[\rho, w, 0, p, p']} \; \Big\{ \; (p', \zeta(w), p) \in R_M^{(\mathrm{rk}_\zeta(w))}$$

$$(5) \quad \frac{[\rho, w, j, p, (p_1, \ldots, p_k)]}{(\rho, w(j+1), p_{j+1})} \; \{ \; 0 \leq j < k$$

$$(6) \quad \frac{[\rho, w, j, p, (p_1, \ldots, p_k)] \qquad [\rho, w(j+1), p_{j+1}]}{[\rho, w, j+1, p, (p_1, \ldots, p_k)]} \; \{ \; 0 \leq j < k$$

$$(7) \quad \frac{[\rho, w, p, \theta]}{[\rho, w, p]}$$

$$(8) \quad \frac{(\rho, w, p)}{(q_\iota, p)} \; \{ \; \zeta(w) = x_\iota$$

$$(9) \quad \frac{(\rho, w, p) \qquad [q_\iota, p, p']}{[\rho, w, 0, p, p']} \; \{ \; \zeta(w) = x_\iota$$

$$(10) \quad \frac{(\rho, w, p)}{[\rho, w, p, \{y_\iota \mapsto p\}]} \; \{ \; \zeta(w) = y_\iota$$

$$(11) \quad \frac{[\rho, w, k, p, p'] \qquad \begin{matrix} [\rho, w1, p_1, \theta_1] \\ \cdots \; [\rho, wk, p_k, \theta_k] \end{matrix}}{[\rho, w, p, \theta \cup \theta_1 \cup \cdots \cup \theta_k]} \; \{ \; p' = (p_1, \ldots, p_k)$$

where $\theta = \{\zeta(w) \mapsto (p, p')\}$ if $\zeta(w) \in X$, and $\theta = \emptyset$ otherwise

Note: we assume that $\rho \in R$, $\rho = (q_1 \cdots q_l, \langle \zeta\zeta' \rangle, q)$, and $q \in Q^{(m)}$.

Gliederung

Effektivität durch endliche Darstellung (vgl. Tabelle 1.2)

Abschlusseigenschaft	Referenzen	Komplexität
$1.(\Sigma^*) \subseteq \mathcal{K}_{\text{Rec}}$	(Berstel und Reutenauer 1988; Schützenberger 1961)	$O(n)$
$\text{Yd}(\mathcal{L}_{\text{Rec}}) \subseteq \mathcal{K}_{\text{CF}}$	(Thatcher 1967; Ésik und Kuich 2003)	$O(r)$
$\text{Yd}^{-1}(\mathcal{K}_{\text{Rec}}) \subseteq \mathcal{L}_{\text{Rec}}$	(Maletti und Satta 2009)	$O(p^k)$
$\mathcal{L}_{\text{Rec}} \odot \mathcal{L}_{\text{Rec}} \subseteq \mathcal{L}_{\text{Rec}}$	(Borchardt 2004, Cor. 3.9)	$O(r_1 \cdot r_2)$
$\mathcal{L}_{\text{Rec}} \lhd \mathcal{T}_{\text{STSG}} \subseteq \mathcal{T}_{\text{STSG}}$	(Maletti 2010)	$O(r_2 \cdot p_1^{k_2})$
$\mathcal{T}_{\text{STSG}} \rhd \mathcal{L}_{\text{Rec}} \subseteq \mathcal{T}_{\text{STSG}}$	(Maletti 2010)	$O(r_1 \cdot p_2^{k_1})$
$\pi_2(\mathcal{T}_{\text{STSG}}) \subseteq \mathcal{L}_{\text{Rec}}$	(Fülöp u. a. 2011)	$O(r)$
$\text{best}(\mathcal{K}_{\text{CF}}) \subseteq \Sigma^*$	(Knuth 1977; Huang und Chiang 2005; Büchse u. a. 2010)	$O(r \cdot \log p)$

Begriffe

Binarisierungsabbildung

bin: $\mathcal{F} \to \mathcal{F}$

- bedeutungstreu,
- "lineares Wachstum": $|R'| \leq k \cdot |R|$

Begriffe

Binarisierungsabbildung

bin: $\mathcal{F} \to \mathcal{F}$

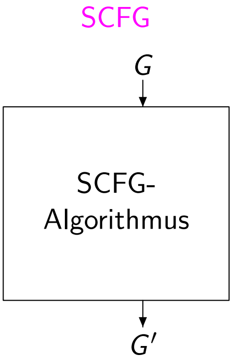
- bedeutungstreu,
- “lineares Wachstum”: $|R'| \leq k \cdot |R|$

Binarisierungsdomäne

$\mathrm{bdom}(\mathrm{bin}) = \{G \mid G \in \mathcal{F}, \mathrm{rk}(\mathrm{bin}(G)) \leq 2\}$

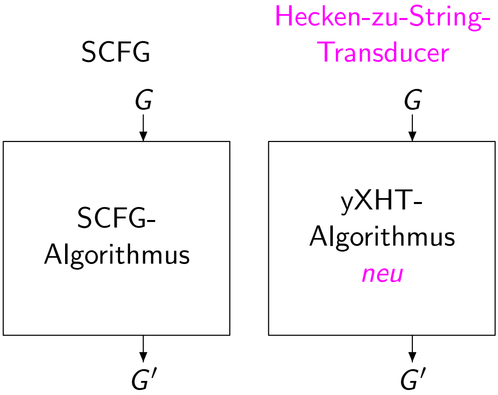
Ergebnis: generischer Algorithmus + Instanzen

Regel-für-regel-vollständige Binarisierungsabbildungen



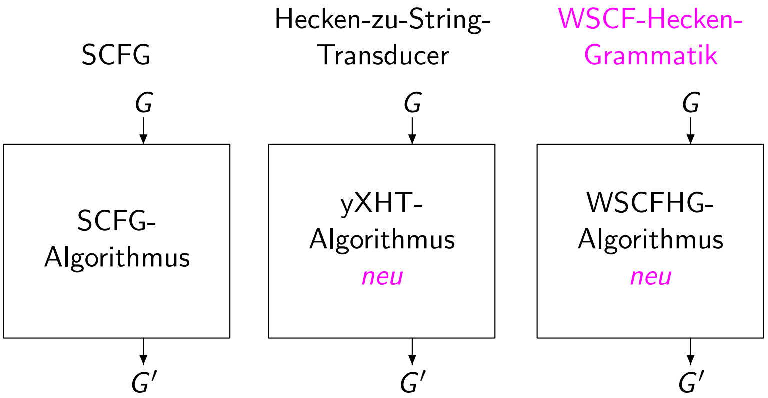
Ergebnis: generischer Algorithmus + Instanzen

Regel-für-regel-vollständige Binarisierungsabbildungen



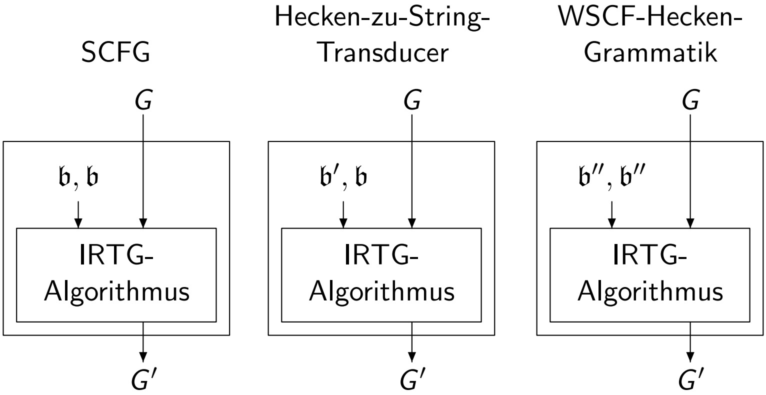
Ergebnis: generischer Algorithmus + Instanzen

Regel-für-regel-vollständige Binarisierungsabbildungen



Ergebnis: generischer Algorithmus + Instanzen

Regel-für-regel-vollständige Binarisierungsabbildungen



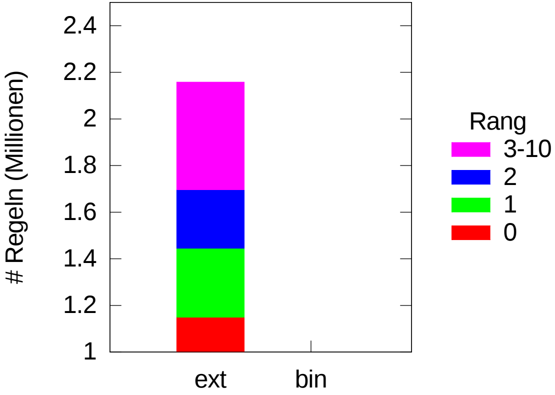
... mit lediglich drei b-rules $\mathfrak{b}, \mathfrak{b}', \mathfrak{b}''$!

Experiment: Hecken-zu-String-Transducer

- extrahiert anhand 1 Mio. Satzpaare (Europarl Eng-Ger)

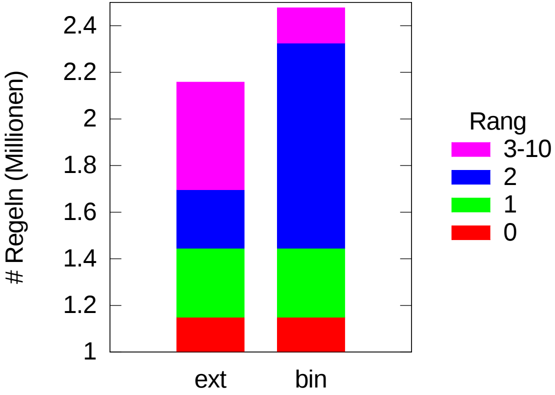
Experiment: Hecken-zu-String-Transducer

- extrahiert anhand 1 Mio. Satzpaare (Europarl Eng-Ger)
- 2.15 Mio. Regeln bis Rang 10



Experiment: Hecken-zu-String-Transducer

- extrahiert anhand 1 Mio. Satzpaare (Europarl Eng-Ger)
- 2.15 Mio. Regeln bis Rang 10
- 67 % konnten verarbeitet werden



Gliederung

Ergebnisse I: Determinisierung

Formalismus	Einschränkung	Halbringeinschränkung
	Literatur	

Ergebnisse I: Determinisierung

Formalismus	Einschränkung	Halbringeinschränkung	
		Literatur	
FSA	–	Boolesch	
WSA	Zwillingseigenschaft	tropisch	
WSA	Zwillingseigenschaft	kommutativ, extremal	(1)

Legende: (1) erfordert eine maximale Faktorisierung vom Nutzer

Ergebnisse I: Determinisierung

Formalismus	Einschränkung	Halbringeinschränkung	
		Literatur	
FSA	–	Boolesch	
WSA	Zwillingseigenschaft	tropisch	
WSA	Zwillingseigenschaft	kommutativ, extremal	(1)
FTA	–	Boolesch	
WTA	–	lokal endlich, Halbkörper	
WTA	–	lokal endlich	
WTA	azyklisch	nichtnegative reelle Zahlen	(2)

Legende: (1) erfordert eine maximale Faktorisierung vom Nutzer
(2) entbehrt eines Beweises

Ergebnisse I: Determinisierung

Formalismus	Einschränkung	Halbringeinschränkung	
		Literatur	
FSA	–	Boolesch	
WSA	Zwillingseigenschaft	tropisch	
WSA	Zwillingseigenschaft	kommutativ, extremal	(1)
FTA	–	Boolesch	
WTA	–	lokal endlich, Halbkörper	
WTA	–	lokal endlich	
WTA	azyklisch	nichtnegative reelle Zahlen	(2)
		Theorem 5.5.3	
WTA	Zwillingseigenschaft	kommutativ, extremal	(1)
WTA	–	lokal endlich	
WTA	azyklisch	kommutativ	

Legende: (1) erfordert eine maximale Faktorisierung vom Nutzer
(2) entbehrt eines Beweises

Ergebnisse I: Determinisierung

Formalismus	Einschränkung	Halbringeinschränkung	
		Literatur	
FSA	–	Boolesch	
WSA	Zwillingseigenschaft	tropisch	
WSA	Zwillingseigenschaft	kommutativ, extremal	(1)
FTA	–	Boolesch	
WTA	–	lokal endlich, Halbkörper	
WTA	–	lokal endlich	
WTA	azyklisch	nichtnegative reelle Zahlen	(2)
		Theorem 5.5.3	
WTA	Zwillingseigenschaft	kommutativ, extremal	(1)
WTA	–	lokal endlich	
WTA	azyklisch	kommutativ	

Legende: (1) erfordert eine maximale Faktorisierung vom Nutzer
(2) entbehrt eines Beweises

Ergebnisse I: Determinisierung

Formalismus	Einschränkung	Halbringeinschränkung	
		Literatur	
FSA	–	Boolesch	
WSA	Zwillingseigenschaft	tropisch	
WSA	Zwillingseigenschaft	kommutativ, extremal	(1)
FTA	–	Boolesch	
WTA	–	lokal endlich, Halbkörper	
WTA	–	lokal endlich	
WTA	azyklisch	nichtnegative reelle Zahlen	(2)
		Theorem 5.5.3	
WTA	Zwillingseigenschaft	kommutativ, extremal	(1)
WTA	–	lokal endlich	
WTA	azyklisch	kommutativ	

Legende: (1) erfordert eine maximale Faktorisierung vom Nutzer
(2) entbehrt eines Beweises

Ergebnisse I: Determinisierung

Formalismus	Einschränkung	Halbringeinschränkung	
		Literatur	
FSA	–	Boolesch	
WSA	Zwillingseigenschaft	tropisch	
WSA	Zwillingseigenschaft	kommutativ, extremal	(1)
FTA	–	Boolesch	
WTA	–	lokal endlich, Halbkörper	
WTA	–	lokal endlich	
WTA	azyklisch	nichtnegative reelle Zahlen	(2)
		Theorem 5.5.3	
WTA	Zwillingseigenschaft	kommutativ, extremal	(1)
WTA	–	lokal endlich	
WTA	azyklisch	kommutativ	

Legende: (1) erfordert eine maximale Faktorisierung vom Nutzer
(2) entbehrt eines Beweises

Ergebnisse II: Entscheidbarkeit

Literatur

Zwillingseigenschaft ist entscheidbar für die Klassen

- der zykeleindeutigen WSA über kommutativen kürzbaren Halbringen,
- der WSA über dem tropischen Halbring.

Ergebnisse II: Entscheidbarkeit

Literatur

Zwillingseigenschaft ist entscheidbar für die Klassen

- der zykeleindeutigen WSA über kommutativen kürzbaren Halbringen,
- der WSA über dem tropischen Halbring.

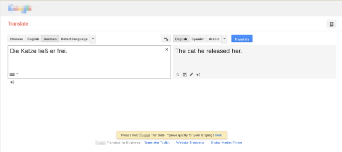
Theoreme 5.5.4, 5.5.5

Zwillingseigenschaft ist entscheidbar für die Klassen

- der zykeleindeutigen WTA über kommutativen nullsummenfreien nullteilerfreien Halbringen,
- der WTA über extremalen Halbkörpern.

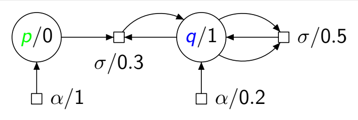
Zusammenfassung

Statistisches
Maschinelles Übersetzen



✓ ✓

Theorie der
formalen Sprachen



✓

Dekoder-
spezifikation

algebraische
Dekoder-
spezifikation

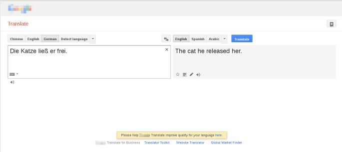
Abschlusseigenschaften

Binarisierung

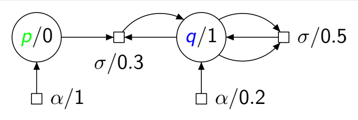
Determinisierung

Zusammenfassung

Statistisches Maschinelles Übersetzen



Theorie der formalen Sprachen



✓ ✓ ✓

Dekoder-spezifikation

algebraische Dekoder-spezifikation

Abschlusseigenschaften

Binarisierung

Determinisierung

Vielen Dank für die Aufmerksamkeit.

Referenzen I

- Abeillé, Anne, Yves Schabes und Aravind K. Joshi (1990). “Using lexicalized tags for machine translation”. In: Proceedings of the Thirteenth International Conference on Computational Linguistics (COLING ’90). Bd. 3, S. 1–6.
- Aho, Alfred V. und Jeffrey D. Ullman (1969). “Syntax directed translations and the pushdown assembler”. In: Journal of Computer and System Sciences 3, S. 37–56.
- Alexandrakis, Athanasios und Symeon Bozapalidis (1987). “Weighted grammars and Kleene’s theorem”. In: Information Processing Letters 24.1, S. 1–4.
- Arnold, André und Max Dauchet (1976). “Bi-transduction de forêts”. In: Proc. 3rd Int. Coll. Automata, Languages and Programming, S. 74–86.
- Berstel, Jean und Christophe Reutenauer (1988). Rational Series and Their Languages. Bd. 12. EATCS Monographs on Theoretical Computer Science. Springer.

Referenzen II

- Borchardt, Björn (2004). “A pumping lemma and decidability problems for recognizable tree series”. In: Acta Cybernet. 16.4, S. 509–544.
- Büchse, Matthias, Daniel Geisler, Torsten Stüber und Heiko Vogler (2010). “n-Best Parsing Revisited”. In: Proceedings of the 2010 Workshop on Applications of Tree Automata in Natural Language Processing, ACL 2010. Uppsala, Sweden, 16 July 2010, S. 46–54.
- Chang, Yin-Wen und Michael Collins (2011). “Exact Decoding of Phrase-Based Translation Models through Lagrangian Relaxation”. In: Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing, S. 26–37. URL: `http://www.aclweb.org/anthology/D11-1003`.
- Chiang, David (2007). “Hierarchical Phrase-Based Translation”. In: Comp. Ling. 33.2, S. 201–228.

Referenzen III

- Chiang, David, Adam Lopez, Nitin Madnani, Christof Monz, Philip Resnik und Michael Subotin (2005). "The Hiero machine translation system: extensions, evaluation, and analysis". In: HLT '05: Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing, S. 779–786.
- DeNeefe, Steve (2011). "Tree-Adjoining Machine Translation". Diss. University of Southern California.
- Dyer, Chris, Adam Lopez, Juri Ganitkevitch, Jonathan Weese, Ferhan Ture, Phil Blunsom, Hendra Setiawan, Vladimir Eidelman und Philip Resnik (2010). "cdec: A Decoder, Alignment, and Learning Framework for Finite-State and Context-Free Translation Models". In: Proceedings of the ACL 2010 System Demonstrations, S. 7–12. URL: http://www.aclweb.org/anthology/P10-4002.

Referenzen IV

Earley, Jay (1970). "An Efficient Context-Free Parsing Algorithm". In: Communications of the ACM 13.2, S. 94–102.

Eisner, Jason (2003). "Learning non-isomorphic tree mappings for machine translation". In: Proceedings of the 41st Annual Meeting on Association for Computational Linguistics - Volume 2. ACL '03, S. 205–208.

Ésik, Zoltán und Werner Kuich (2003). "Formal Tree Series". In: J. Autom. Lang. Combin. 8.2, S. 219–285.

Fülöp, Zoltán, Andreas Maletti und Heiko Vogler (2011). "Weighted Extended Tree Transducers". In: Fundam. Inform. 111.2, S. 163–202.

Fülöp, Zoltán und Heiko Vogler (2009). "Weighted tree automata and tree transducers". In: Handbook of Weighted Automata. Hrsg. von Manfred Droste, Werner Kuich und Heiko Vogler. EATCS Monographs in Theoretical Computer Science. Kap. 9.

Referenzen V

- Gildea, Daniel (2010). “Optimal Parsing Strategies for Linear Context-Free Rewriting Systems”. In: Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics, S. 769–776. URL: http://www.aclweb.org/anthology/N10-1118.
- Goguen, Joseph A., Jim W. Thatcher, Eric G. Wagner und Jesse B. Wright (1977). “Initial algebra semantics and continuous algebras”. In: J. ACM 24, S. 68–95.
- Golan, Jonathan S. (1999). Semirings and their Applications. Kluwer Academic.
- Hebisch, Udo und Hanns Joachim Weinert (1998). Semirings: Algebraic Theory and Applications in Computer Science. Bd. 5. Series in Algebra. World Scientific.

Referenzen VI

Huang, Liang und David Chiang (2005). "Better k-best parsing". In: Parsing '05: Proceedings of the Ninth International Workshop on Parsing Technology, S. 53–64. URL: http://www.cis.upenn.edu/~lhuang3/huang-iwpt-correct.pdf.

Huang, Liang, Hao Zhang, Daniel Gildea und Kevin Knight (2009). "Binarization of Synchronous Context-Free Grammars". In: Comp. Ling. 35.4, S. 559–595. URL: http://www.aclweb.org/anthology/J/J09/J09-4009.pdf.

Kaeshammer, Miriam (2013). "Synchronous Linear Context-Free Rewriting Systems for Machine Translation". In: Proceedings of the Seventh Workshop on Syntax, Semantics and Structure in Statistical Translation, S. 68–77. URL: http://www.aclweb.org/anthology/W13-0808.

Referenzen VII

- Kallmeyer, Laura, Wolfgang Maier und Giorgio Satta (2009). "Synchronous Rewriting in Treebanks". In: Proceedings of the 11th International Conference on Parsing Technologies (IWPT'09), S. 69–72. URL: http://www.aclweb.org/anthology/W09-3810.
- Knuth, Donald E. (1977). "A Generalization of Dijkstra's Algorithm". In: Inform. Process. Lett. 6.1, S. 1–5.
- Koehn, Philipp, Hieu Hoang, Alexandra Birch, Chris Callison-Burch, Marcello Federico, Nicola Bertoldi, Brooke Cowan, Wade Shen, Christine Moran, Richard Zens, Chris Dyer, Ondřej Bojar, Alexandra Constantin und Evan Herbst (2007). "Moses: open source toolkit for statistical machine translation". In: Proceedings of the 45th Annual Meeting of the ACL on Interactive Poster and Demonstration Sessions. ACL '07, S. 177–180. URL: http://www.aclweb.org/anthology/P07-2045.

Referenzen VIII

Koller, Alexander und Marco Kuhlmann (2011). "A Generalized View on Parsing and Translation". In: Proceedings of the 12th International Conference on Parsing Technologies, S. 2–13. URL: http://www.aclweb.org/anthology/W11-2902.

Kuich, Werner (1998). "Formal power series over trees". In: 3rd International Conference on Developments in Language Theory, DLT 1997, Thessaloniki, Greece, Proceedings. Hrsg. von Symeon Bozapalidis, S. 61–101.

Lewis, Philip M. und Richard E. Stearns (1966). "Syntax directed transduction". In: Foundations of Computer Science, IEEE Annual Symposium on, S. 21–35.

Referenzen IX

- Li, Zhifei, Chris Callison-Burch, Chris Dyer, Juri Ganitkevitch, Sanjeev Khudanpur, Lane Schwartz, Wren N. G. Thornton, Jonathan Weese und Omar F. Zaidan (2009). "Joshua: an open source toolkit for parsing-based machine translation". In: Proceedings of the Fourth Workshop on Statistical Machine Translation. StatMT '09, S. 135–139. URL: http://dl.acm.org/citation.cfm?id=1626431.1626459.
- Maletti, Andreas (2010). "Input and Output Products for Weighted Extended Top-down Tree Transducers". In: Proc. 14th Int. Conf. Developments in Language Theory. Hrsg. von Yuan Gao, Hanlin Lu, Shinnosuke Seki und Sheng Yu. Bd. 6224. LNCS, S. 316–327.
- Maletti, Andreas und Giorgio Satta (2009). "Parsing Algorithms based on Tree Automata". In: Proc. 11th Int. Conf. Parsing Technologies, S. 1–12.

Referenzen X

May, Jonathan und Kevin Knight (2006). "A better N-best list: practical determinization of weighted finite tree automata". In: Proceedings of the main conference on Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics, S. 351–358.

Nesson, Rebecca (2009). "Synchronous and Multicomponent Tree-Adjoining Grammars: Complexity, Algorithms and Linguistic Applications". Diss. Harvard University.

Rush, Alexander M. und Michael Collins (2011). "Exact Decoding of Syntactic Translation Models through Lagrangian Relaxation". In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Techologies, S. 72–82. URL: http://www.aclweb.org/anthology/P11-1008.

Referenzen XI

- Schabes, Yves und Richard C. Waters (1994). “Tree insertion grammars: a cubic-time, parsable formalism that lexicalizes context-free grammar without changing the trees produced”. In: Comput. Linguist. 21, S. 479–513.
- Schützenberger, Marcel-Paul (1961). “On the definition of a family of automata”. In: Information and Control 4, S. 245–270.
- Shieber, Stuart M. (2007). “Probabilistic Synchronous Tree-Adjoining Grammars for Machine Translation: The Argument from Bilingual Dictionaries”. In: Proceedings of the Workshop on Syntax and Structure in Statistical Translation. Hrsg. von Dekai Wu und David Chiang.
- Shieber, Stuart M. und Yves Schabes (1990). “Synchronous Tree-Adjoining Grammars”. In: Proceedings of the 13th International Conference on Computational Linguistics (COLING '90). Bd. 3, S. 253–258.

Referenzen XII

Thatcher, James W. (1967). “Characterizing derivation trees of context-free grammars through a generalization of finite automata theory.” In: J. Comput. System Sci. 1.4, S. 317–322.

Gliederung

Ziele konkreter

Langfristziel: jede Methode sollte ...

(a)

(b)

(c)

(d)

(e)

Ziele konkreter

Langfristziel: jede Methode sollte ...

(a) Stand der Technik unterbringen

(b)

(c)

(d)

(e)

Ziele konkreter

Langfristziel: jede Methode sollte ...

(a) Stand der Technik unterbringen

(b) Verfeinerungsprozess (Kaskade) vereinfachen

(c)

(d)

(e)

Ziele konkreter

Langfristziel: jede Methode sollte . . .

(a) Stand der Technik unterbringen

(b) Verfeinerungsprozess (Kaskade) vereinfachen

(c) formale Beziehungen zw. Abstraktionsniveaus enthalten

(d)

(e)

Ziele konkreter

Langfristziel: jede Methode sollte ...

(a) Stand der Technik unterbringen

(b) Verfeinerungsprozess (Kaskade) vereinfachen

(c) formale Beziehungen zw. Abstraktionsniveaus enthalten

(d) gegenseitige Befruchtung von Theorie und Anwendung fördern

(e)

Ziele konkreter

Langfristziel: jede Methode sollte . . .

(a) Stand der Technik unterbringen

(b) Verfeinerungsprozess (Kaskade) vereinfachen

(c) formale Beziehungen zw. Abstraktionsniveaus enthalten

(d) gegenseitige Befruchtung von Theorie und Anwendung fördern

(e) leicht zu lernen und beizubehalten sein

Ziele konkreter

Langfristziel: jede Methode sollte ...

(a) Stand der Technik unterbringen

(b) Verfeinerungsprozess (Kaskade) vereinfachen

(c) formale Beziehungen zw. Abstraktionsniveaus enthalten

(d) gegenseitige Befruchtung von Theorie und Anwendung fördern

(e) leicht zu lernen und beizubehalten sein

Ergebnisse

	Langfristziel: jede Methode sollte . . .	*Ergebnis:* das algebraische Framework . . .
(a)	Stand der Technik unterbringen	
(b)	Verfeinerungsprozess (Kaskade) vereinfachen	
(c)	formale Beziehungen zw. Abstraktionsniveaus enthalten	
(d)	gegenseitige Befruchtung von Theorie und Anwendung fördern	
(e)	leicht zu lernen und beizubehalten sein	

Ergebnisse

	Langfristziel: jede Methode sollte ...	*Ergebnis:* das algebraische Framework ...
(a)	Stand der Technik unterbringen	kann zeitgenössische syntax-basierte Dekoder darstellen
(b)	Verfeinerungsprozess (Kaskade) vereinfachen	
(c)	formale Beziehungen zw. Abstraktionsniveaus enthalten	
(d)	gegenseitige Befruchtung von Theorie und Anwendung fördern	
(e)	leicht zu lernen und beizubehalten sein	

Ergebnisse

	Langfristziel: jede Methode sollte …	*Ergebnis:* das algebraische Framework …
(a)	Stand der Technik unterbringen	kann zeitgenössische syntax-basierte Dekoder darstellen
(b)	Verfeinerungsprozess (Kaskade) vereinfachen	lässt jede Operation isoliert verfeinern (abstrakte, effektive Spezifikation)
(c)	formale Beziehungen zw. Abstraktionsniveaus enthalten	
(d)	gegenseitige Befruchtung von Theorie und Anwendung fördern	
(e)	leicht zu lernen und beizubehalten sein	

Ergebnisse

	Langfristziel: jede Methode sollte …	*Ergebnis:* das algebraische Framework …
(a)	Stand der Technik unterbringen	kann zeitgenössische syntax-basierte Dekoder darstellen
(b)	Verfeinerungsprozess (Kaskade) vereinfachen	lässt jede Operation isoliert verfeinern (abstrakte, effektive Spezifikation)
(c)	formale Beziehungen zw. Abstraktionsniveaus enthalten	garantiert Äquivalenz zwischen beiden Niveaus
(d)	gegenseitige Befruchtung von Theorie und Anwendung fördern	
(e)	leicht zu lernen und beizubehalten sein	

Ergebnisse

	Langfristziel: jede Methode sollte ...	*Ergebnis:* das algebraische Framework ...
(a)	Stand der Technik unterbringen	kann zeitgenössische syntax-basierte Dekoder darstellen
(b)	Verfeinerungsprozess (Kaskade) vereinfachen	lässt jede Operation isoliert verfeinern (abstrakte, effektive Spezifikation)
(c)	formale Beziehungen zw. Abstraktionsniveaus enthalten	garantiert Äquivalenz zwischen beiden Niveaus
(d)	gegenseitige Befruchtung von Theorie und Anwendung fördern	ist Schnittstelle zur Theorie gewichteter Baumautomaten
(e)	leicht zu lernen und beizubehalten sein	

Ergebnisse

	Langfristziel: jede Methode sollte ...	*Ergebnis:* das algebraische Framework ...
(a)	Stand der Technik unterbringen	kann zeitgenössische syntax-basierte Dekoder darstellen
(b)	Verfeinerungsprozess (Kaskade) vereinfachen	lässt jede Operation isoliert verfeinern (abstrakte, effektive Spezifikation)
(c)	formale Beziehungen zw. Abstraktionsniveaus enthalten	garantiert Äquivalenz zwischen beiden Niveaus
(d)	gegenseitige Befruchtung von Theorie und Anwendung fördern	ist Schnittstelle zur Theorie gewichteter Baumautomaten
(e)	leicht zu lernen und beizubehalten sein	ist schwer; enthält viele fortgeschrittene Konzepte

Ergebnisse

	Langfristziel: jede Methode sollte ...	*Ergebnis:* das algebraische Framework ...
(a)	Stand der Technik unterbringen	kann zeitgenössische syntax-basierte Dekoder darstellen
(b)	Verfeinerungsprozess (Kaskade) vereinfachen	lässt jede Operation isoliert verfeinern (abstrakte, effektive Spezifikation)
(c)	formale Beziehungen zw. Abstraktionsniveaus enthalten	garantiert Äquivalenz zwischen beiden Niveaus
(d)	gegenseitige Befruchtung von Theorie und Anwendung fördern	ist Schnittstelle zur Theorie gewichteter Baumautomaten
(e)	leicht zu lernen und beizubehalten sein	ist schwer; enthält viele fortgeschrittene Konzepte

Gliederung

Halbring

(Hebisch und Weinert 1998; Golan 1999)

$\mathcal{S} = (S, +, \cdot, 0, 1)$ heiße Halbring, falls

Halbring

(Hebisch und Weinert 1998; Golan 1999)

$\mathcal{S} = (S, +, \cdot, 0, 1)$ heiße Halbring, falls

- Trägermenge: S Menge

Halbring

(Hebisch und Weinert 1998; Golan 1999)

$\mathcal{S} = (S, +, \cdot, 0, 1)$ heiße Halbring, falls

- Trägermenge: S Menge
- Addition: $+$ binäre Operation auf S, assoziativ, kommutativ

$$(s_1 + s_2) + s_3 = s_1 + (s_2 + s_3)$$
$$s_1 + s_2 = s_2 + s_1$$

Halbring

(Hebisch und Weinert 1998; Golan 1999)

$\mathcal{S} = (S, +, \cdot, 0, 1)$ heiße Halbring, falls

- Trägermenge: S Menge
- Addition: $+$ binäre Operation auf S, assoziativ, kommutativ
- Multiplikation: $\cdot$ binäre Operation auf S, assoziativ

$$(s_1 \cdot s_2) \cdot s_3 = s_1 \cdot (s_2 \cdot s_3)$$

Halbring

(Hebisch und Weinert 1998; Golan 1999)

$\mathcal{S} = (S, +, \cdot, 0, 1)$ heiße Halbring, falls

- Trägermenge: S Menge
- Addition: $+$ binäre Operation auf S, assoziativ, kommutativ
- Multiplikation: $\cdot$ binäre Operation auf S, assoziativ
- Null: 0 neutrales Element der Addition

$$0 + s = s$$

Halbring

(Hebisch und Weinert 1998; Golan 1999)

$\mathcal{S} = (S, +, \cdot, 0, 1)$ heiße Halbring, falls

- Trägermenge: S Menge
- Addition: $+$ binäre Operation auf S, assoziativ, kommutativ
- Multiplikation: $\cdot$ binäre Operation auf S, assoziativ
- Null: 0 neutrales Element der Addition
- Eins: 1 neutrales Element der Multiplikation

$$1 \cdot s = s = s \cdot 1$$

Halbring

(Hebisch und Weinert 1998; Golan 1999)

$\mathcal{S} = (S, +, \cdot, 0, 1)$ heiße Halbring, falls

- Trägermenge: S Menge
- Addition: $+$ binäre Operation auf S, assoziativ, kommutativ
- Multiplikation: $\cdot$ binäre Operation auf S, assoziativ
- Null: 0 neutrales Element der Addition
- Eins: 1 neutrales Element der Multiplikation
- Distributivität: Multiplikation distribuiert über Addition:

$$s \cdot (s_1 + s_2) = s \cdot s_1 + s \cdot s_2$$
$$(s_1 + s_2) \cdot s = s_1 \cdot s + s_2 \cdot s$$

Halbring

(Hebisch und Weinert 1998; Golan 1999)

$\mathcal{S} = (S, +, \cdot, 0, 1)$ heiße Halbring, falls

- Trägermenge: S Menge
- Addition: $+$ binäre Operation auf S, assoziativ, kommutativ
- Multiplikation: $\cdot$ binäre Operation auf S, assoziativ
- Null: 0 neutrales Element der Addition
- Eins: 1 neutrales Element der Multiplikation
- Distributivität: Multiplikation distribuiert über Addition:
- die Null ist annihilierend bezüglich der Multiplikation:

$$0 \cdot s = 0 = s \cdot 0$$

Halbring

Beispiele

Halbring

Beispiele

- $(\bar{\mathbb{R}}, \max, +, -\infty, 0)$ arktischer Halbring

Halbring

Beispiele

- $(\bar{\mathbb{R}}, \mathsf{max}, +, -\infty, 0)$ arktischer Halbring
- $(\mathbb{R}_{\geq 0}, +, \cdot, 0, 1)$ nichtnegative reelle Zahlen

Halbring

Beispiele

- $(\bar{\mathbb{R}}, \max, +, -\infty, 0)$ arktischer Halbring
- $(\mathbb{R}_{\geq 0}, +, \cdot, 0, 1)$ nichtnegative reelle Zahlen
- $(\{0, 1\}, \vee, \wedge, 0, 1)$ Boolescher Halbring

Halbring

Beispiele

- $(\bar{\mathbb{R}}, \mathsf{max}, +, -\infty, 0)$ arktischer Halbring
- $(\mathbb{R}_{\geq 0}, +, \cdot, 0, 1)$ nichtnegative reelle Zahlen
- $(\{0, 1\}, \vee, \wedge, 0, 1)$ Boolescher Halbring
- $(\mathcal{P}(\Sigma^*), \cup, \cdot, \emptyset, \{\varepsilon\})$ formale Sprachen

Halbring

Beispiele

- $(\bar{\mathbb{R}}, \max, +, -\infty, 0)$ arktischer Halbring
- $(\mathbb{R}_{\geq 0}, +, \cdot, 0, 1)$ nichtnegative reelle Zahlen
- $(\{0,1\}, \vee, \wedge, 0, 1)$ Boolescher Halbring
- $(\mathcal{P}(\Sigma^*), \cup, \cdot, \emptyset, \{\varepsilon\})$ formale Sprachen
- $(\bar{\mathbb{R}}^m, \max, +, \widetilde{-\infty}, \widetilde{0})$ Produkt-Halbring (Vektoren)

Halbring

Beispiele

- $(\bar{\mathbb{R}}, \max, +, -\infty, 0)$ arktischer Halbring
- $(\mathbb{R}_{\geq 0}, +, \cdot, 0, 1)$ nichtnegative reelle Zahlen
- $(\{0, 1\}, \vee, \wedge, 0, 1)$ Boolescher Halbring
- $(\mathcal{P}(\Sigma^*), \cup, \cdot, \emptyset, \{\varepsilon\})$ formale Sprachen
- $(\bar{\mathbb{R}}^m, \max, +, \widetilde{-\infty}, \widetilde{0})$ Produkt-Halbring (Vektoren)
- $(\bar{\mathbb{R}}^{T_\Sigma}, \max, +, \widetilde{-\infty}, \widetilde{0})$ (gewichtete Baumsprachen)

(Halbring-) gewichtete Baumsprache

Seien Σ Alphabet, $\mathcal{S}$ Halbring.

$\varphi\colon T_\Sigma \to S$ heiße gewichtete Baumsprache über Σ und $\mathcal{S}$

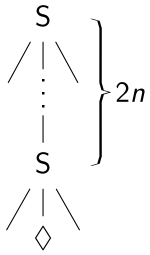
(Halbring-) gewichtete Baumsprache

Seien Σ Alphabet, $\mathcal{S}$ Halbring.
$\varphi\colon T_\Sigma \to S$ heiße gewichtete Baumsprache über Σ und $\mathcal{S}$

Beispiel

$\Sigma = \{\mathsf{S}, \mathsf{a}, \mathsf{b}, \mathsf{c}, \mathsf{d}, \langle, \rangle, \Diamond\}$, $\mathcal{S} = (\mathbb{R}_{\geq 0}, +, \cdot, 0, 1)$

$$\varphi(t) = \begin{cases} 0.5^{2n} \cdot 0.2^{4n} & \text{falls } t \text{ von der Form} \\ 0 & \text{anderenfalls.} \end{cases}$$



(Halbring-) gewichteter Baumautomat (WTA)

(Alexandrakis und Bozapalidis 1987; Kuich 1998; Ésik und Kuich 2003; Fülöp und Vogler 2009)

Seien Σ Alphabet, $\mathcal{S}$ Halbring.
$M = (Q, R, \mu, \nu)$ heiße WTA über Σ und $\mathcal{S}$, falls

(Halbring-) gewichteter Baumautomat (WTA)

(Alexandrakis und Bozapalidis 1987; Kuich 1998; Ésik und Kuich 2003; Fülöp und Vogler 2009)

Seien Σ Alphabet, $\mathcal{S}$ Halbring.
$M = (Q, R, \mu, \nu)$ heiße WTA über Σ und $\mathcal{S}$, falls

- Zustände: Q nichtleere, endliche Menge

(Halbring-) gewichteter Baumautomat (WTA)

(Alexandrakis und Bozapalidis 1987; Kuich 1998; Ésik und Kuich 2003; Fülöp und Vogler 2009)

Seien Σ Alphabet, $\mathcal{S}$ Halbring.
$M = (Q, R, \mu, \nu)$ heiße WTA über Σ und $\mathcal{S}$, falls

- Zustände: Q nichtleere, endliche Menge
- Transitionsregeln: $R \subseteq Q^* \times \Sigma \times Q$ endlich

(Halbring-) gewichteter Baumautomat (WTA)

(Alexandrakis und Bozapalidis 1987; Kuich 1998; Ésik und Kuich 2003; Fülöp und Vogler 2009)

Seien Σ Alphabet, $\mathcal{S}$ Halbring.
$M = (Q, R, \mu, \nu)$ heiße WTA über Σ und $\mathcal{S}$, falls

- Zustände: Q nichtleere, endliche Menge
- Transitionsregeln: $R \subseteq Q^* \times \Sigma \times Q$ endlich
- Gewichtszuweisung: $\mu\colon R \to S$

(Halbring-) gewichteter Baumautomat (WTA)

(Alexandrakis und Bozapalidis 1987; Kuich 1998; Ésik und Kuich 2003; Fülöp und Vogler 2009)

Seien Σ Alphabet, $\mathcal{S}$ Halbring.
$M = (Q, R, \mu, \nu)$ heiße WTA über Σ und $\mathcal{S}$, falls

- Zustände: Q nichtleere, endliche Menge
- Transitionsregeln: $R \subseteq Q^* \times \Sigma \times Q$ endlich
- Gewichtszuweisung: $\mu \colon R \to S$
- Wurzelgewichtsabbildung: $\nu \colon Q \to S$

(Halbring-) gewichteter Baumautomat (WTA)

Seien Σ Alphabet, $\mathcal{S}$ Halbring.
$M = (Q, R, \mu, \nu)$ heiße WTA über Σ und $\mathcal{S}$, falls

- Zustände: Q nichtleere, endliche Menge
- Transitionsregeln: $R \subseteq Q^* \times \Sigma \times Q$ endlich
- Gewichtszuweisung: $\mu\colon R \to S$
- Wurzelgewichtsabbildung: $\nu\colon Q \to S$

(Halbring-) gewichteter Baumautomat (WTA)

Seien Σ Alphabet, $\mathcal{S}$ Halbring.
$M = (Q, R, \mu, \nu)$ heiße WTA über Σ und $\mathcal{S}$, falls

- Zustände: Q nichtleere, endliche Menge
- Transitionsregeln: $R \subseteq Q^* \times \Sigma \times Q$ endlich
- Gewichtszuweisung: $\mu \colon R \to S$
- Wurzelgewichtsabbildung: $\nu \colon Q \to S$

Beispiel

$\Sigma = \{\mathsf{S}, \mathsf{a}, \mathsf{b}, \mathsf{c}, \mathsf{d}, \langle, \rangle, \Diamond\}$, $\mathcal{S} = (\mathbb{R}_{\geq 0}, +, \cdot, 0, 1)$

Q, ν: $\quad 0 \mapsto 1\ , \quad 1 \mapsto 0\ , \quad * \mapsto 0\ .$

$$R, \mu: \quad \begin{aligned} (*1*, \mathsf{S}, 0) &\mapsto 0.5\ , \\ (*0*, \mathsf{S}, 1) &\mapsto 0.5\ , \\ (\varepsilon, \Diamond, 0) &\mapsto 1\ , \\ (\varepsilon, x, *) &\mapsto 0.2\ . \qquad x \in \{\mathsf{a}, \mathsf{b}, \mathsf{c}, \mathsf{d}, \langle, \rangle\} \end{aligned}$$

(Halbring-) gewichteter Baumautomat (WTA)

Seien Σ Alphabet, $\mathcal{S}$ Halbring.
$M = (Q, R, \mu, \nu)$ heiße WTA über Σ und $\mathcal{S}$, falls

- Zustände: Q nichtleere, endliche Menge
- Transitionsregeln: $R \subseteq Q^* \times \Sigma \times Q$ endlich
- Gewichtszuweisung: $\mu \colon R \to S$
- Wurzelgewichtsabbildung: $\nu \colon Q \to S$

Semantik

$D(M) \subseteq T_{Q \cup R}$ Menge der (gültigen) Läufe,
intuitiv:

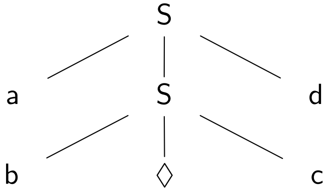
(Halbring-) gewichteter Baumautomat (WTA)

Seien Σ Alphabet, $\mathcal{S}$ Halbring.
$M = (Q, R, \mu, \nu)$ heiße WTA über Σ und $\mathcal{S}$, falls

- Zustände: Q nichtleere, endliche Menge
- Transitionsregeln: $R \subseteq Q^* \times \Sigma \times Q$ endlich
- Gewichtszuweisung: $\mu \colon R \to S$
- Wurzelgewichtsabbildung: $\nu \colon Q \to S$

Semantik

$D(M) \subseteq T_{Q \cup R}$ Menge der (gültigen) Läufe,
intuitiv:



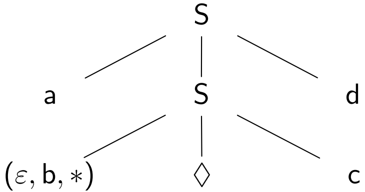
(Halbring-) gewichteter Baumautomat (WTA)

Seien Σ Alphabet, $\mathcal{S}$ Halbring.
$M = (Q, R, \mu, \nu)$ heiße WTA über Σ und $\mathcal{S}$, falls

- Zustände: Q nichtleere, endliche Menge
- Transitionsregeln: $R \subseteq Q^* \times \Sigma \times Q$ endlich
- Gewichtszuweisung: $\mu \colon R \to S$
- Wurzelgewichtsabbildung: $\nu \colon Q \to S$

Semantik

$D(M) \subseteq T_{Q \cup R}$ Menge der (gültigen) Läufe,
intuitiv:



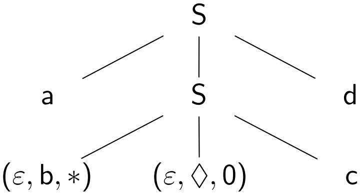
(Halbring-) gewichteter Baumautomat (WTA)

Seien Σ Alphabet, $\mathcal{S}$ Halbring.
$M = (Q, R, \mu, \nu)$ heiße WTA über Σ und $\mathcal{S}$, falls

- Zustände: Q nichtleere, endliche Menge
- Transitionsregeln: $R \subseteq Q^* \times \Sigma \times Q$ endlich
- Gewichtszuweisung: $\mu\colon R \to S$
- Wurzelgewichtsabbildung: $\nu\colon Q \to S$

Semantik

$D(M) \subseteq T_{Q \cup R}$ Menge der (gültigen) Läufe,
intuitiv:



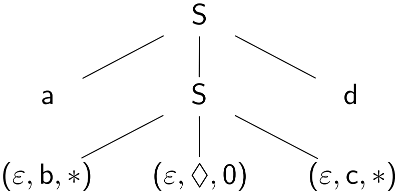
(Halbring-) gewichteter Baumautomat (WTA)

Seien Σ Alphabet, $\mathcal{S}$ Halbring.
$M = (Q, R, \mu, \nu)$ heiße WTA über Σ und $\mathcal{S}$, falls

- Zustände: Q nichtleere, endliche Menge
- Transitionsregeln: $R \subseteq Q^* \times \Sigma \times Q$ endlich
- Gewichtszuweisung: $\mu\colon R \to S$
- Wurzelgewichtsabbildung: $\nu\colon Q \to S$

Semantik

$D(M) \subseteq T_{Q \cup R}$ Menge der (gültigen) Läufe,
intuitiv:



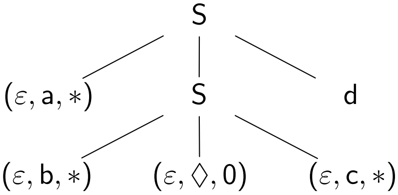
(Halbring-) gewichteter Baumautomat (WTA)

Seien Σ Alphabet, $\mathcal{S}$ Halbring.
$M = (Q, R, \mu, \nu)$ heiße WTA über Σ und $\mathcal{S}$, falls

- Zustände: Q nichtleere, endliche Menge
- Transitionsregeln: $R \subseteq Q^* \times \Sigma \times Q$ endlich
- Gewichtszuweisung: $\mu\colon R \to S$
- Wurzelgewichtsabbildung: $\nu\colon Q \to S$

Semantik

$D(M) \subseteq T_{Q \cup R}$ Menge der (gültigen) Läufe,
intuitiv:



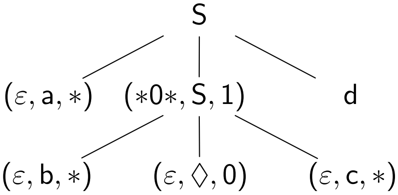
(Halbring-) gewichteter Baumautomat (WTA)

Seien Σ Alphabet, $\mathcal{S}$ Halbring.
$M = (Q, R, \mu, \nu)$ heiße WTA über Σ und $\mathcal{S}$, falls

- Zustände: Q nichtleere, endliche Menge
- Transitionsregeln: $R \subseteq Q^* \times \Sigma \times Q$ endlich
- Gewichtszuweisung: $\mu \colon R \to S$
- Wurzelgewichtsabbildung: $\nu \colon Q \to S$

Semantik

$D(M) \subseteq T_{Q \cup R}$ Menge der (gültigen) Läufe,
intuitiv:



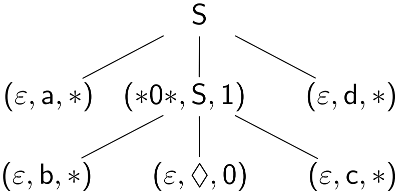
(Halbring-) gewichteter Baumautomat (WTA)

Seien Σ Alphabet, $\mathcal{S}$ Halbring.
$M = (Q, R, \mu, \nu)$ heiße WTA über Σ und $\mathcal{S}$, falls

- Zustände: Q nichtleere, endliche Menge
- Transitionsregeln: $R \subseteq Q^* \times \Sigma \times Q$ endlich
- Gewichtszuweisung: $\mu\colon R \to S$
- Wurzelgewichtsabbildung: $\nu\colon Q \to S$

Semantik

$D(M) \subseteq T_{Q \cup R}$ Menge der (gültigen) Läufe,
intuitiv:



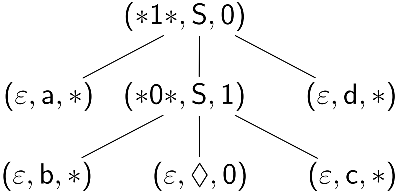
(Halbring-) gewichteter Baumautomat (WTA)

Seien Σ Alphabet, $\mathcal{S}$ Halbring.
$M = (Q, R, \mu, \nu)$ heiße WTA über Σ und $\mathcal{S}$, falls

- Zustände: Q nichtleere, endliche Menge
- Transitionsregeln: $R \subseteq Q^* \times \Sigma \times Q$ endlich
- Gewichtszuweisung: $\mu\colon R \to S$
- Wurzelgewichtsabbildung: $\nu\colon Q \to S$

Semantik

$D(M) \subseteq T_{Q \cup R}$ Menge der (gültigen) Läufe,
intuitiv:



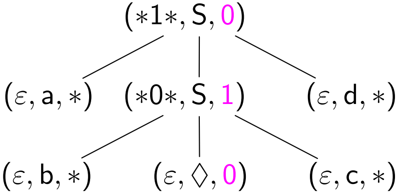
(Halbring-) gewichteter Baumautomat (WTA)

Seien Σ Alphabet, $\mathcal{S}$ Halbring.
$M = (Q, R, \mu, \nu)$ heiße WTA über Σ und $\mathcal{S}$, falls

- Zustände: Q nichtleere, endliche Menge
- Transitionsregeln: $R \subseteq Q^* \times \Sigma \times Q$ endlich
- Gewichtszuweisung: $\mu \colon R \to S$
- Wurzelgewichtsabbildung: $\nu \colon Q \to S$

Semantik

$D(M) \subseteq T_{Q \cup R}$ Menge der (gültigen) Läufe,
intuitiv:



(Halbring-) gewichteter Baumautomat (WTA)

Seien Σ Alphabet, $\mathcal{S}$ Halbring.
$M = (Q, R, \mu, \nu)$ heiße WTA über Σ und $\mathcal{S}$, falls

- Zustände: Q nichtleere, endliche Menge
- Transitionsregeln: $R \subseteq Q^* \times \Sigma \times Q$ endlich
- Gewichtszuweisung: $\mu \colon R \to S$
- Wurzelgewichtsabbildung: $\nu \colon Q \to S$

Semantik

$D(M) \subseteq T_{Q \cup R}$ Menge der (gültigen) Läufe,
formal: kleinste Menge $D \subseteq T_{Q \cup R}$, so dass

- $Q \subseteq D$,
- $\rho(d_1, \ldots, d_k) \in D$, falls
  - $\rho \in R$, $\rho = (q_1 \cdots q_k, \sigma, q)$,
  - $d_1, \ldots, d_k \in D$ und
  - $\pi_Q(d_j) = q_k$.

(Halbring-) gewichteter Baumautomat (WTA)

Seien Σ Alphabet, $\mathcal{S}$ Halbring.
$M = (Q, R, \mu, \nu)$ heiße WTA über Σ und $\mathcal{S}$, falls

- Zustände: Q nichtleere, endliche Menge
- Transitionsregeln: $R \subseteq Q^* \times \Sigma \times Q$ endlich
- Gewichtszuweisung: $\mu \colon R \to S$
- Wurzelgewichtsabbildung: $\nu \colon Q \to S$

Semantik

$D(M) \subseteq T_{Q \cup R}$ Menge der (gültigen) Läufe,
$\langle . \rangle_\mu \colon T_{Q \cup R} \to S$ Laufgewicht,

$$\langle q \rangle_\mu = 1 \ ,$$
$$\langle \rho(d_1, \ldots, d_k) \rangle_\mu = \langle d_1 \rangle_\mu \cdots \langle d_k \rangle_\mu \cdot \mu(\rho) \ .$$

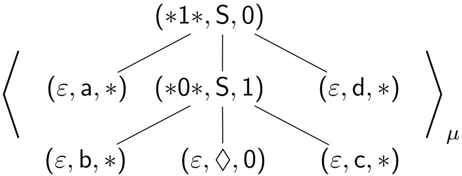
(Halbring-) gewichteter Baumautomat (WTA)

Seien Σ Alphabet, $\mathcal{S}$ Halbring.
$M = (Q, R, \mu, \nu)$ heiße WTA über Σ und $\mathcal{S}$, falls

- Zustände: Q nichtleere, endliche Menge
- Transitionsregeln: $R \subseteq Q^* \times \Sigma \times Q$ endlich
- Gewichtszuweisung: $\mu\colon R \to S$
- Wurzelgewichtsabbildung: $\nu\colon Q \to S$

Semantik

$D(M) \subseteq T_{Q \cup R}$ Menge der (gültigen) Läufe,
$\langle . \rangle_\mu\colon T_{Q \cup R} \to S$ Laufgewicht,



(Halbring-) gewichteter Baumautomat (WTA)

Seien Σ Alphabet, $\mathcal{S}$ Halbring.
$M = (Q, R, \mu, \nu)$ heiße WTA über Σ und $\mathcal{S}$, falls

- Zustände: Q nichtleere, endliche Menge
- Transitionsregeln: $R \subseteq Q^* \times \Sigma \times Q$ endlich
- Gewichtszuweisung: $\mu\colon R \to S$
- Wurzelgewichtsabbildung: $\nu\colon Q \to S$

Semantik

$D(M) \subseteq T_{Q \cup R}$ Menge der (gültigen) Läufe,
$\langle . \rangle_\mu\colon T_{Q \cup R} \to S$ Laufgewicht,

$$= \langle (\varepsilon, \mathsf{a}, *) \rangle_\mu \cdot \Big\langle (*0*, \mathsf{S}, 1)\big((\varepsilon, \mathsf{b}, *),\ (\varepsilon, \Diamond, 0),\ (\varepsilon, \mathsf{c}, *)\big) \Big\rangle_\mu$$

$$\cdot \langle (\varepsilon, \mathsf{d}, *) \rangle_\mu \cdot \mu(*1*, S, 0)$$

(Halbring-) gewichteter Baumautomat (WTA)

Seien Σ Alphabet, $\mathcal{S}$ Halbring.
$M = (Q, R, \mu, \nu)$ heiße WTA über Σ und $\mathcal{S}$, falls

- Zustände: Q nichtleere, endliche Menge
- Transitionsregeln: $R \subseteq Q^* \times \Sigma \times Q$ endlich
- Gewichtszuweisung: $\mu\colon R \to S$
- Wurzelgewichtsabbildung: $\nu\colon Q \to S$

Semantik

$D(M) \subseteq T_{Q \cup R}$ Menge der (gültigen) Läufe,
$\langle . \rangle_\mu \colon T_{Q \cup R} \to S$ Laufgewicht,

$$= \langle (\varepsilon, \mathsf{a}, *) \rangle_\mu \cdot \langle (\varepsilon, \mathsf{b}, *) \rangle_\mu \cdot \langle (\varepsilon, \Diamond, 0) \rangle_\mu \cdot \langle (\varepsilon, \mathsf{c}, *) \rangle_\mu \\ \cdot \mu(*0*, \mathsf{S}, 1) \cdot \langle (\varepsilon, \mathsf{d}, *) \rangle_\mu \cdot \mu(*1*, \mathsf{S}, 0)$$

(Halbring-) gewichteter Baumautomat (WTA)

Seien Σ Alphabet, $\mathcal{S}$ Halbring.
$M = (Q, R, \mu, \nu)$ heiße WTA über Σ und $\mathcal{S}$, falls

- Zustände: Q nichtleere, endliche Menge
- Transitionsregeln: $R \subseteq Q^* \times \Sigma \times Q$ endlich
- Gewichtszuweisung: $\mu\colon R \to S$
- Wurzelgewichtsabbildung: $\nu\colon Q \to S$

Semantik

$D(M) \subseteq T_{Q \cup R}$ Menge der (gültigen) Läufe,
$\langle . \rangle_\mu\colon T_{Q \cup R} \to S$ Laufgewicht,

$$
\begin{aligned}
&= \mu(\varepsilon, \mathsf{a}, *) \cdot \mu(\varepsilon, \mathsf{b}, *) \cdot \mu(\varepsilon, \Diamond, 0) \cdot \mu(\varepsilon, \mathsf{c}, *) \\
&\quad \cdot \mu(*0*, \mathsf{S}, 1) \cdot \mu(\varepsilon, \mathsf{d}, *) \cdot \mu(*1*, \mathsf{S}, 0)
\end{aligned}
$$

(Halbring-) gewichteter Baumautomat (WTA)

Seien Σ Alphabet, $\mathcal{S}$ Halbring.
$M = (Q, R, \mu, \nu)$ heiße WTA über Σ und $\mathcal{S}$, falls

- Zustände: Q nichtleere, endliche Menge
- Transitionsregeln: $R \subseteq Q^* \times \Sigma \times Q$ endlich
- Gewichtszuweisung: $\mu \colon R \to S$
- Wurzelgewichtsabbildung: $\nu \colon Q \to S$

Semantik

$D(M) \subseteq T_{Q \cup R}$ Menge der (gültigen) Läufe,
$\langle . \rangle_\mu \colon T_{Q \cup R} \to S$ Laufgewicht,

$$= 0.5^2 \cdot 0.2^4 \ .$$

(Halbring-) gewichteter Baumautomat (WTA)

Seien Σ Alphabet, $\mathcal{S}$ Halbring.
$M = (Q, R, \mu, \nu)$ heiße WTA über Σ und $\mathcal{S}$, falls

- Zustände: Q nichtleere, endliche Menge
- Transitionsregeln: $R \subseteq Q^* \times \Sigma \times Q$ endlich
- Gewichtszuweisung: $\mu\colon R \to S$
- Wurzelgewichtsabbildung: $\nu\colon Q \to S$

Semantik

$D(M) \subseteq T_{Q \cup R}$ Menge der (gültigen) Läufe,
$\langle . \rangle_\mu\colon T_{Q \cup R} \to S$ Laufgewicht,
$[\![M]\!]\colon T_\Sigma \to S$ Bedeutung von M,

(Halbring-) gewichteter Baumautomat (WTA)

Seien Σ Alphabet, $\mathcal{S}$ Halbring.
$M = (Q, R, \mu, \nu)$ heiße WTA über Σ und $\mathcal{S}$, falls

- Zustände: Q nichtleere, endliche Menge
- Transitionsregeln: $R \subseteq Q^* \times \Sigma \times Q$ endlich
- Gewichtszuweisung: $\mu\colon R \to S$
- Wurzelgewichtsabbildung: $\nu\colon Q \to S$

Semantik

$D(M) \subseteq T_{Q \cup R}$ Menge der (gültigen) Läufe,
$\langle . \rangle_\mu\colon T_{Q \cup R} \to S$ Laufgewicht,
$[\![M]\!]\colon T_\Sigma \to S$ Bedeutung von M,

$$[\![M]\!](t) = \sum_{q \in Q} \left(\sum_{\substack{d \in D(M)\colon \\ \pi_\Sigma(d) = t, \\ \pi_Q(d) = q}} \langle d \rangle_\mu \right) \cdot \nu(q) \, .$$

(Halbring-) gewichteter Baumautomat (WTA)

Seien Σ Alphabet, $\mathcal{S}$ Halbring.
$M = (Q, R, \mu, \nu)$ heiße WTA über Σ und $\mathcal{S}$, falls

- Zustände: Q nichtleere, endliche Menge
- Transitionsregeln: $R \subseteq Q^* \times \Sigma \times Q$ endlich
- Gewichtszuweisung: $\mu\colon R \to S$
- Wurzelgewichtsabbildung: $\nu\colon Q \to S$

Semantik

$D(M) \subseteq T_{Q \cup R}$ Menge der (gültigen) Läufe,
$\langle . \rangle_\mu\colon T_{Q \cup R} \to S$ Laufgewicht,
$[\![M]\!]\colon T_\Sigma \to S$ Bedeutung von M,

$$[\![M]\!]\Big(\mathrm{S}\big(\mathrm{a},\ \mathrm{S}(\mathrm{b}, \Diamond, \mathrm{c}),\ \mathrm{d}\big)\Big) = \Big\langle (*1*, \mathrm{S}, 0)\big((\varepsilon, \mathrm{a}, *),\ (*0*, \mathrm{S}, 1)\big((\varepsilon, \mathrm{b}, *), (\varepsilon, \Diamond, 0), (\varepsilon, \mathrm{c}, *)\big),\ (\varepsilon, \mathrm{d}, *)\big) \Big\rangle_\mu = 0.5^2 \cdot 0.2^4 \,.$$

(Halbring-) gewichtete synchrone kontextfreie Baumgrammatik (WSCFTG)

$C_\Sigma(m, r_1, \ldots, r_l) \subseteq T_{\Sigma \cup X_l \cup Y_m}$, so dass für jedes Element t:

- jede Variable kommt genau einmal in t vor,
- ein mit x_j beschrifteter Knoten hat genau r_j Nachfolger,
- ein mit y_j beschrifteter Knoten hat genau 0 Nachfolger.

(Halbring-) gewichtete synchrone kontextfreie Baumgrammatik (WSCFTG)

$C_\Sigma(m, r_1, \ldots, r_l) \subseteq T_{\Sigma \cup X_l \cup Y_m}$, so dass für jedes Element t:

- jede Variable kommt genau einmal in t vor,
- ein mit x_j beschrifteter Knoten hat genau r_j Nachfolger,
- ein mit y_j beschrifteter Knoten hat genau 0 Nachfolger.

Sei $\Gamma \subseteq \{\langle \zeta_1 \zeta_2 \rangle \mid \exists m, k, r_1, \ldots, r_l \colon \zeta_1, \zeta_2 \in C_\Sigma(m, r_1, \ldots, r_l)\}$ und $G = (Q, R, \mu, \nu)$ WTA über Γ und $\mathcal{S}$.

(Halbring-) gewichtete synchrone kontextfreie Baumgrammatik (WSCFTG)

$C_\Sigma(m, r_1, \ldots, r_l) \subseteq T_{\Sigma \cup X_l \cup Y_m}$, so dass für jedes Element t:

- jede Variable kommt genau einmal in t vor,
- ein mit x_j beschrifteter Knoten hat genau r_j Nachfolger,
- ein mit y_j beschrifteter Knoten hat genau 0 Nachfolger.

Sei $\Gamma \subseteq \{\langle \zeta_1 \zeta_2 \rangle \mid \exists m, k, r_1, \ldots, r_l \colon \zeta_1, \zeta_2 \in C_\Sigma(m, r_1, \ldots, r_l)\}$ und $G = (Q, R, \mu, \nu)$ WTA über Γ und $\mathcal{S}$.

G heiße WSCFTG über Σ und $\mathcal{S}$, falls es eine Abbildung

$$\mathrm{rk} \colon Q \to \mathbb{N}$$

gibt, so dass

$$\begin{aligned} &(q_1 \cdots q_k, \langle \zeta_1 \zeta_2 \rangle, q) \in R \\ &\implies \zeta_1, \zeta_2 \in C_\Sigma(\mathrm{rk}(q), \mathrm{rk}(q_1), \ldots, \mathrm{rk}(q_k)) \,, \\ &\nu(q) \neq 0 \implies \mathrm{rk}(q) = 0 \,. \end{aligned}$$

Beobachtung zum Typ

Seien $\xi \in T_\Gamma$, $d \in D(G)$, $\xi = \pi_\Gamma(d)$, $q = \pi_Q(d)$.

Dann gilt $h_i(\xi) \in C_\Sigma(\mathrm{rk}(q))$.

Semantik einer WSCFTG

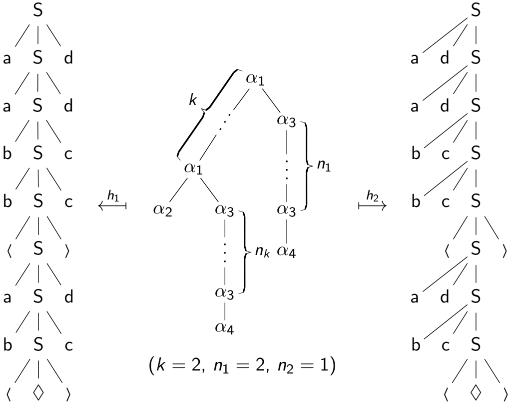
G WSCFTG über Σ und $\mathcal{S}$

$[\![G]\!]' \colon T_\Sigma \times T_\Sigma \to S$ WSCFTG-Bedeutung von G

Semantik einer WSCFTG

G WSCFTG über Σ und $\mathcal{S}$

$[\![G]\!]' \colon T_\Sigma \times T_\Sigma \to S$ WSCFTG-Bedeutung von G

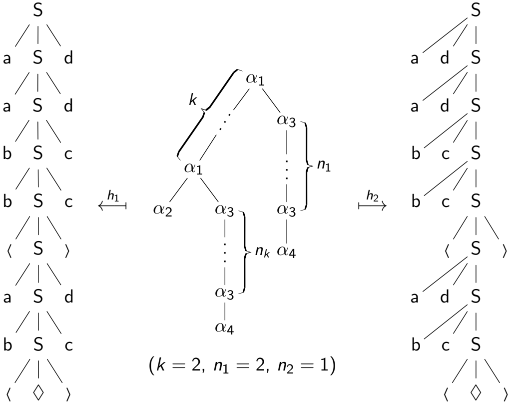


$(k = 2,\ n_1 = 2,\ n_2 = 1)$

Semantik einer WSCFTG

G WSCFTG über Σ und $\mathcal{S}$

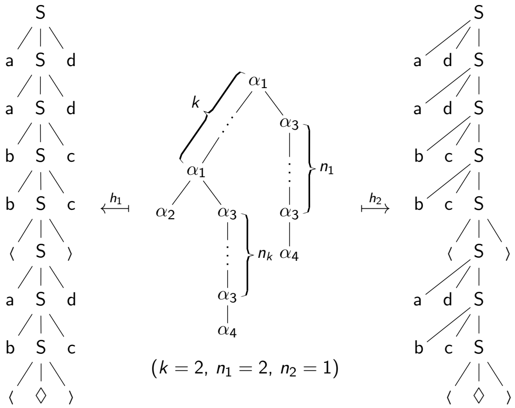
$[\![G]\!]' \colon T_\Sigma \times T_\Sigma \to S$ WSCFTG-Bedeutung von G
$[\![G]\!]'(t_1, t_2) = \sum_{\xi \in T_\Gamma \colon h_i(\xi) = t_i} [\![G]\!](\xi)$.



Semantik einer WSCFTG

G WSCFTG über Σ und $\mathcal{S}$, $\langle \zeta_1 \zeta_2 \rangle \in \Gamma \implies \zeta_i \notin \{x_1, x_1(y_1)\}$

$[\![G]\!]' \colon T_\Sigma \times T_\Sigma \to S$ WSCFTG-Bedeutung von G
$[\![G]\!]'(t_1, t_2) = \sum_{\xi \in T_\Gamma \colon h_i(\xi) = t_i} [\![G]\!](\xi)$.



$(k = 2,\ n_1 = 2,\ n_2 = 1)$

Gliederung

Zu zeigen

Es existiert eine WSCFTG $M \lhd G$ mit $[\![M]\!] \lhd [\![G]\!] = [\![M \lhd G]\!]$.

Zu zeigen

Es existiert eine WSCFTG $M \lhd G$ mit $[\![M]\!] \lhd [\![G]\!] = [\![M \lhd G]\!]$.

Idee: Zentralbäume umgewichten. Konstruiere $M \lhd G$ mit

$$
\begin{aligned}
[\![M]\!](t_1) \cdot [\![G]\!](t_1, t_2) &= \\
&= \\
&= \\
&= \\
&= \\
&= [\![M \lhd G]\!](t_1, t_2)\ .
\end{aligned}
$$

Zu zeigen

Es existiert eine WSCFTG $M \lhd G$ mit $[\![M]\!] \lhd [\![G]\!] = [\![M \lhd G]\!]$.

Idee: Zentralbäume umgewichten. Konstruiere $M \lhd G$ mit

$$
\begin{aligned}
[\![M]\!](t_1) \cdot [\![G]\!](t_1, t_2) &= [\![M]\!](t_1) \cdot \sum\nolimits_{\xi\colon h_i(\xi) = t_i} [\![G]\!](\xi) \\
&= \\
&= \\
&= \\
&= \\
&= [\![M \lhd G]\!](t_1, t_2) \; .
\end{aligned}
$$

Zu zeigen

Es existiert eine WSCFTG $M \lhd G$ mit $[\![M]\!] \lhd [\![G]\!] = [\![M \lhd G]\!]$.

Idee: Zentralbäume umgewichten. Konstruiere $M \lhd G$ mit

$$\begin{aligned}
[\![M]\!](t_1) \cdot [\![G]\!](t_1, t_2) &= [\![M]\!](t_1) \cdot \sum\nolimits_{\xi\colon\, h_i(\xi)=t_i} [\![G]\!](\xi) \\
&= \sum\nolimits_{\xi\colon\, h_i(\xi)=t_i} [\![M]\!](t_1) \cdot [\![G]\!](\xi) \\
&= \\
&= \\
&= \\
&= [\![M \lhd G]\!](t_1, t_2)\ .
\end{aligned}$$

Zu zeigen

Es existiert eine WSCFTG $M \lhd G$ mit $[\![M]\!] \lhd [\![G]\!] = [\![M \lhd G]\!]$.

Idee: Zentralbäume umgewichten. Konstruiere $M \lhd G$ mit

$$\begin{aligned}
[\![M]\!](t_1) \cdot [\![G]\!](t_1, t_2) &= [\![M]\!](t_1) \cdot \sum\nolimits_{\xi \colon h_i(\xi) = t_i} [\![G]\!](\xi) \\
&= \sum\nolimits_{\xi \colon h_i(\xi) = t_i} [\![M]\!](t_1) \cdot [\![G]\!](\xi) \\
&= \sum\nolimits_{\xi \colon h_i(\xi) = t_i} [\![M]\!](h_1(\xi)) \cdot [\![G]\!](\xi) \\
&= \\
&= \\
&= [\![M \lhd G]\!](t_1, t_2) \ .
\end{aligned}$$

Zu zeigen

Es existiert eine WSCFTG $M \lhd G$ mit $[\![M]\!] \lhd [\![G]\!] = [\![M \lhd G]\!]$.

Idee: Zentralbäume umgewichten. Konstruiere $M \lhd G$ mit

$$
\begin{aligned}
[\![M]\!](t_1) \cdot [\![G]\!](t_1, t_2) &= [\![M]\!](t_1) \cdot \sum\nolimits_{\xi\colon\, h_i(\xi)=t_i} [\![G]\!](\xi) \\
&= \sum\nolimits_{\xi\colon\, h_i(\xi)=t_i} [\![M]\!](t_1) \cdot [\![G]\!](\xi) \\
&= \sum\nolimits_{\xi\colon\, h_i(\xi)=t_i} [\![M]\!](h_1(\xi)) \cdot [\![G]\!](\xi) \\
&= \cdots \\
&= \\
&= [\![M \lhd G]\!](t_1, t_2)\ .
\end{aligned}
$$

Zu zeigen

Es existiert eine WSCFTG $M \lhd G$ mit $[\![M]\!] \lhd [\![G]\!] = [\![M \lhd G]\!]$.

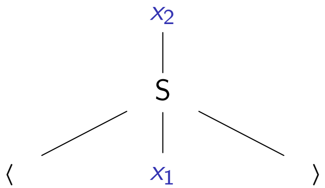
Idee: Zentralbäume umgewichten. Konstruiere $M \lhd G$ mit

$$
\begin{aligned}
[\![M]\!](t_1) \cdot [\![G]\!](t_1, t_2) &= [\![M]\!](t_1) \cdot \sum\nolimits_{\xi \colon h_i(\xi) = t_i} [\![G]\!](\xi) \\
&= \sum\nolimits_{\xi \colon h_i(\xi) = t_i} [\![M]\!](t_1) \cdot [\![G]\!](\xi) \\
&= \sum\nolimits_{\xi \colon h_i(\xi) = t_i} [\![M]\!](h_1(\xi)) \cdot [\![G]\!](\xi) \\
&= \cdots \\
&= \sum\nolimits_{\xi \colon h_i(\xi) = t_i} [\![M \lhd G]\!](\xi) \\
&= [\![M \lhd G]\!](t_1, t_2) \, .
\end{aligned}
$$

Beispiel: $M \lhd G$

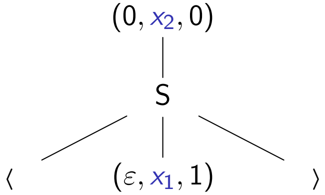
$$q \qquad \rightarrow \alpha_1(\; q \qquad , \; f \qquad) \quad \# \; 0.3$$

Beispiel: $M \lhd G$



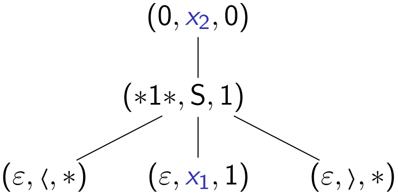
$$q \quad \rightarrow \alpha_1(\, q \quad , \; f \quad) \quad \# \; 0.3$$

Beispiel: $M \lhd G$



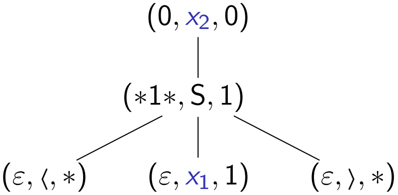
$$q \quad \to \alpha_1(\; q \quad , \; f \quad) \quad \# \; 0.3$$

Beispiel: $M \lhd G$



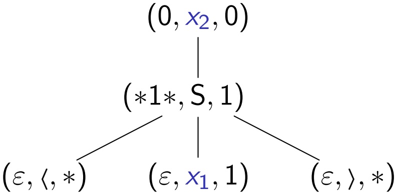
$$q \quad \to \alpha_1(\ q \quad , \ f \quad) \quad \#\ 0.3 \cdot (0.5 \cdot 0.2^2)$$

Beispiel: $M \lhd G$



$$(q, 0, \varepsilon) \to \alpha_1\big((q, 1, \varepsilon), (f, 0, 0)\big) \quad \#\ 0.3 \cdot (0.5 \cdot 0.2^2)$$

Beispiel: $M \lhd G$



$$
\begin{array}{ll}
(q, 0, \varepsilon) \to \alpha_1\big((q, 1, \varepsilon), (f, 0, 0)\big) & \#\ 0.3 \cdot (0.5 \cdot 0.2^2) \\
(q, 1, \varepsilon) \to \alpha_1\big((q, 0, \varepsilon), (f, 1, 1)\big) & \#\ 0.3 \cdot (0.5 \cdot 0.2^2) \\
(q, 0, \varepsilon) \to \alpha_2() & \#\ 1.0 \\
(f, 0, 0) \to \alpha_3\big((f, 1, 1)\big) & \#\ 0.7 \cdot (0.5^2 \cdot 0.2^4) \\
(f, 1, 1) \to \alpha_3\big((f, 0, 0)\big) & \#\ 0.7 \cdot (0.5^2 \cdot 0.2^4) \\
(f, 0, 0) \to \alpha_4() & \#\ 1.0 \\
(f, 1, 1) \to \alpha_4() & \#\ 1.0
\end{array}
$$

Allgemein: raten und M für $C_\Sigma(m, r_1, \ldots, r_l)$ hochrüsten

$\theta\colon X_l \cup Y_m \cup \{\diamond\} \to P \times P^* \cup P$ heiße Zuweisung vom Typ $(m, r_1, \ldots, r_l)$, falls

- $\theta(x_j) \in P \times P^{r_j}$,
- $\theta(y_j) \in P$,
- $\theta(\diamond) \in P$.

Allgemein: raten und M für $C_\Sigma(m, r_1, \ldots, r_l)$ hochrüsten

$\theta\colon X_l \cup Y_m \cup \{\diamond\} \to P \times P^* \cup P$ heiße Zuweisung vom Typ $(m, r_1, \ldots, r_l)$, falls

- $\theta(x_j) \in P \times P^{r_j}$,
- $\theta(y_j) \in P$,
- $\theta(\diamond) \in P$.

$M = (P, R_M, \mu_M, \nu_M) \rightsquigarrow M\theta = (P, R_M \cup R', \mu_M \cup \mu', \nu')$ mit

$$\nu'(p) = \begin{cases} 1 & \text{falls } p = \theta(\diamond) \text{ ,} \\ 0 & \text{anderenfalls.} \end{cases}$$

R', μ':

$$\begin{aligned} (p', x_j, p) &\mapsto 1 & \qquad \theta(x_j) &= (p, p') \\ (\varepsilon, y_j, p) &\mapsto 1 & \qquad \theta(y_j) &= p \end{aligned}$$

Allgemein: Konstruktion von $M \lhd G$

$G = (Q, R, \mu, \nu) \rightsquigarrow M \lhd G = (Q \times P \times P^*, R', \mu', \nu')$ mit

Allgemein: Konstruktion von $M \lhd G$

$$G = (Q, R, \mu, \nu) \rightsquigarrow M \lhd G = (Q \times P \times P^*, R', \mu', \nu') \text{ mit}$$

$\rho \in R$, $\rho = (q_1 \cdots q_l, \langle \zeta_1 \zeta_2 \rangle, q)$,
θ Zuweisung für ρ, d. h. vom Typ $(\mathrm{rk}(q), \mathrm{rk}(q_1), \ldots, \mathrm{rk}(q_l))$

$$\rho\theta = \Big((q_1, \theta(x_1)) \cdots (q_l, \theta(x_l)), \langle \zeta_1 \zeta_2 \rangle, \big(q, (\theta(\diamond), \theta(y_1) \cdots \theta(y_m))\big) \Big)$$

Allgemein: Konstruktion von $M \lhd G$

$G = (Q, R, \mu, \nu) \rightsquigarrow M \lhd G = (Q \times P \times P^*, R', \mu', \nu')$ mit

R', μ': $\quad \rho\theta \mapsto \mu(\rho) \cdot [\![M\theta]\!](\zeta_1)$ $\qquad$ ζ_1 Eingabebaum von ρ

$\rho \in R$, $\rho = (q_1 \cdots q_l, \langle \zeta_1 \zeta_2 \rangle, q)$,
θ Zuweisung für ρ, d. h. vom Typ $(\mathrm{rk}(q), \mathrm{rk}(q_1), \dots, \mathrm{rk}(q_l))$

$$\rho\theta =$$
$$\Big((q_1, \theta(x_1)) \cdots (q_l, \theta(x_l)), \langle \zeta_1 \zeta_2 \rangle, (q, (\theta(\diamond), \theta(y_1) \cdots \theta(y_m))) \Big)$$

Allgemein: Konstruktion von $M \lhd G$

$G = (Q, R, \mu, \nu) \rightsquigarrow M \lhd G = (Q \times P \times P^*, R', \mu', \nu')$ mit

R', μ': $\quad \rho\theta \mapsto \mu(\rho) \cdot [\![M\theta]\!](\zeta_1)$ $\qquad$ ζ_1 Eingabebaum von ρ

$$\nu'(q, p, p) = \begin{cases} \nu(q) \cdot \nu_M(p) & \text{falls } p' = \varepsilon \text{ ,} \\ 0 & \text{anderenfalls.} \end{cases}$$

$\rho \in R$, $\rho = (q_1 \cdots q_l, \langle \zeta_1 \zeta_2 \rangle, q)$,
θ Zuweisung für ρ, d. h. vom Typ $(\mathrm{rk}(q), \mathrm{rk}(q_1), \ldots, \mathrm{rk}(q_l))$

$$\rho\theta =$$
$$\Big((q_1, \theta(x_1)) \cdots (q_l, \theta(x_l)), \langle \zeta_1 \zeta_2 \rangle, (q, (\theta(\diamond), \theta(y_1) \cdots \theta(y_m))) \Big)$$

Gliederung

Bekannte Technik: Regel-für-Regel-Binarisierung

SCFG G: (Rang 4)

$$S \to \langle x_1 x_2 x_3 x_4, x_3 x_1 x_4 x_2 \rangle (A, B, C, D)$$
$$A \to \langle \mathrm{a}, \mathrm{a} \rangle$$
$$B \to \langle \mathrm{b}, \mathrm{b} \rangle$$
$$C \to \langle \mathrm{c}, \mathrm{c} \rangle$$
$$D \to \langle \mathrm{d}, \mathrm{d} \rangle$$

Bekannte Technik: Regel-für-Regel-Binarisierung

SCFG G: (Rang 4)

$S \to \langle x_1 x_2 x_3 x_4, x_3 x_1 x_4 x_2 \rangle (A, B, C, D)$
$A \to \langle \mathrm{a}, \mathrm{a} \rangle$
$B \to \langle \mathrm{b}, \mathrm{b} \rangle$
$C \to \langle \mathrm{c}, \mathrm{c} \rangle$
$D \to \langle \mathrm{d}, \mathrm{d} \rangle$

Bekannte Technik: Regel-für-Regel-Binarisierung

SCFG G: (Rang 4)

$S \rightarrow \langle x_1 x_2 x_3 x_4, x_3 x_1 x_4 x_2 \rangle (A, B, C, D)$

$A \rightarrow \langle \mathrm{a}, \mathrm{a} \rangle$

$B \rightarrow \langle \mathrm{b}, \mathrm{b} \rangle$

$C \rightarrow \langle \mathrm{c}, \mathrm{c} \rangle$

$D \rightarrow \langle \mathrm{d}, \mathrm{d} \rangle$

SCFG G': (Rang 2)

$S \rightarrow \langle x_1 x_2, \lightning \rangle (A, [BCD])$

Bekannte Technik: Regel-für-Regel-Binarisierung

SCFG G: (Rang 4)

$S \rightarrow \langle x_1 x_2 x_3 x_4, x_3 x_1 x_4 x_2 \rangle (A, B, C, D)$

$A \rightarrow \langle \mathrm{a}, \mathrm{a} \rangle$

$B \rightarrow \langle \mathrm{b}, \mathrm{b} \rangle$

$C \rightarrow \langle \mathrm{c}, \mathrm{c} \rangle$

$D \rightarrow \langle \mathrm{d}, \mathrm{d} \rangle$

SCFG G': (Rang 2)

$S \rightarrow \langle x_1 x_2, ↯ \rangle (A, [BCD])$

$S \rightarrow \langle x_1 x_2, ↯ \rangle ([ABC], D)$

Bekannte Technik: Regel-für-Regel-Binarisierung

SCFG G: (Rang 4)

$S \to \langle x_1 x_2 x_3 x_4, x_3 x_1 x_4 x_2 \rangle (A, B, C, D)$

$A \to \langle \mathrm{a}, \mathrm{a} \rangle$

$B \to \langle \mathrm{b}, \mathrm{b} \rangle$

$C \to \langle \mathrm{c}, \mathrm{c} \rangle$

$D \to \langle \mathrm{d}, \mathrm{d} \rangle$

SCFG G': (Rang 2)

$S \to \langle x_1 x_2, \lightning \rangle (A, [BCD])$

$S \to \langle x_1 x_2, \lightning \rangle ([ABC], D)$

$S \to \langle x_1 x_2, \lightning \rangle ([AB], [CD])$

Bekannte Technik: Regel-für-Regel-Binarisierung

SCFG G: (Rang 4)

$S \rightarrow \langle x_1 x_2 x_3 x_4, x_3 x_1 x_4 x_2 \rangle (A, B, C, D)$

$A \rightarrow \langle \mathrm{a}, \mathrm{a} \rangle$

$B \rightarrow \langle \mathrm{b}, \mathrm{b} \rangle$

$C \rightarrow \langle \mathrm{c}, \mathrm{c} \rangle$

$D \rightarrow \langle \mathrm{d}, \mathrm{d} \rangle$

SCFG G': (Rang 2)

$S \rightarrow \langle x_1 x_2, \lightning \rangle (A, [BCD])$

$S \rightarrow \langle x_1 x_2, \lightning \rangle ([ABC], D)$

$S \rightarrow \langle x_1 x_2, \lightning \rangle ([AB], [CD])$

Fazit: Im Allgemeinen nicht erfolgreich!

Bekannte Technik: Regel-für-Regel-Binarisierung

SCFG G: (Rang 4)

$S \to \langle x_1 x_2 x_3 x_4, x_3 x_1 x_4 x_2 \rangle (A, B, C, D)$
$A \to \langle \mathrm{a}, \mathrm{a} \rangle$
$B \to \langle \mathrm{b}, \mathrm{b} \rangle$
$C \to \langle \mathrm{c}, \mathrm{c} \rangle$
$D \to \langle \mathrm{d}, \mathrm{d} \rangle$

SCFG G': (Rang 2)

$S \to \langle x_1 x_2, \lightning \rangle (A, [BCD])$
$S \to \langle x_1 x_2, \lightning \rangle ([ABC], D)$
$S \to \langle x_1 x_2, \lightning \rangle ([AB], [CD])$

Fazit: Im Allgemeinen nicht erfolgreich!
Algorithmus von Huang, Zhang u. a. (2009)

Bekannte Technik: Regel-für-Regel-Binarisierung

SCFG G: (Rang 4)

$S \to \langle x_1 x_2 x_3 x_4, x_3 x_1 x_4 x_2 \rangle (A, B, C, D)$
$A \to \langle \mathrm{a}, \mathrm{a} \rangle$
$B \to \langle \mathrm{b}, \mathrm{b} \rangle$
$C \to \langle \mathrm{c}, \mathrm{c} \rangle$
$D \to \langle \mathrm{d}, \mathrm{d} \rangle$

SCFG G': (Rang 2)

$S \to \langle x_1 x_2, \lightning \rangle (A, [BCD])$
$S \to \langle x_1 x_2, \lightning \rangle ([ABC], D)$
$S \to \langle x_1 x_2, \lightning \rangle ([AB], [CD])$

Fazit: Im Allgemeinen nicht erfolgreich!
Algorithmus von Huang, Zhang u. a. (2009)

Vergleiche: $S \to \langle \mathrm{abcd}, \mathrm{cadb} \rangle$

Bekannte Technik: Regel-für-Regel-Binarisierung

SCFG G: (Rang 4)

$S \to \langle x_1x_2x_3x_4, x_3x_1x_4x_2\rangle(A, B, C, D)$
$A \to \langle \mathrm{a}, \mathrm{a}\rangle$
$B \to \langle \mathrm{b}, \mathrm{b}\rangle$
$C \to \langle \mathrm{c}, \mathrm{c}\rangle$
$D \to \langle \mathrm{d}, \mathrm{d}\rangle$

SCFG G': (Rang 2)

$S \to \langle x_1x_2, \lightning\rangle(A, [BCD])$
$S \to \langle x_1x_2, \lightning\rangle([ABC], D)$
$S \to \langle x_1x_2, \lightning\rangle([AB], [CD])$

Fazit: Im Allgemeinen nicht erfolgreich!
Algorithmus von Huang, Zhang u. a. (2009)

Vergleiche: $S \to \langle \mathrm{abcd}, \mathrm{cadb}\rangle$

Wiederum: Im Allgemeinen gibt es keine äquivalente SCFG mit Rang 2. (Aho und Ullman 1969)

Bekannte Technik: Regel-für-Regel-Binarisierung

SCFG G: (Rang 4)

$S \to \langle x_1 x_2 x_3 x_4, x_3 x_1 x_4 x_2 \rangle (A, B, C, D)$
$A \to \langle \mathrm{a}, \mathrm{a} \rangle$
$B \to \langle \mathrm{b}, \mathrm{b} \rangle$
$C \to \langle \mathrm{c}, \mathrm{c} \rangle$
$D \to \langle \mathrm{d}, \mathrm{d} \rangle$

SCFG G': (Rang 2)

$S \to \langle x_1 x_2, \lightning \rangle (A, [BCD])$
$S \to \langle x_1 x_2, \lightning \rangle ([ABC], D)$
$S \to \langle x_1 x_2, \lightning \rangle ([AB], [CD])$

Fazit: Im Allgemeinen nicht erfolgreich!
Algorithmus von Huang, Zhang u. a. (2009)

Vergleiche: $S \to \langle \mathrm{abcd}, \mathrm{cadb} \rangle$

Wiederum: Im Allgemeinen gibt es keine äquivalente SCFG mit Rang 2. (Aho und Ullman 1969)

Binarisierung für andere Formalismen (generisch)?

Andere Formalismen: formell

Formalismus: endliche Kontrolle, Generation

$$\begin{tikzcd} & T_\Gamma \arrow[r, "[\![M]\!]"] \arrow[dl, "h_1"'] \arrow[dr, "h_2"] & S \\ T_\Sigma & & T_\Sigma \end{tikzcd}$$

Zentralbäume

Abgeleitete Bäume

Andere Formalismen: formell

Formalismus: endliche Kontrolle, Generation

Formalisieren über universelle Algebren (Goguen u. a. 1977)

$$
\begin{tikzcd}
 & T_\Gamma \arrow[r, "\llbracket M \rrbracket"] \arrow[dl, "h_1"'] \arrow[dr, "h_2"] & S & \text{Zentralbäume} \\
T_\Sigma & & T_\Sigma & \text{Abgeleitete Bäume}
\end{tikzcd}
$$

Andere Formalismen: formell

Formalismus: endliche Kontrolle, Generation

Formalisieren über universelle Algebren (Goguen u. a. 1977)

$$
\begin{tikzcd}
 & T_\Gamma \arrow[r, "[\![M]\!]"] \arrow[dl, "(.)^{\mathcal{A}_1}"'] \arrow[dr, "(.)^{\mathcal{A}_2}"] & S \\
A_1 & & A_2
\end{tikzcd}
$$

Zentralbäume

Abgeleitete Objekte

Andere Formalismen: formell

Formalismus: endliche Kontrolle, Generation

Formalisieren über universelle Algebren (Goguen u. a. 1977)

Problem: black box

$$
\begin{tikzcd}
 & T_\Gamma \arrow[r, "\llbracket M \rrbracket"] \arrow[dl, "(.)^{\mathcal{A}_1}"'] \arrow[dr, "(.)^{\mathcal{A}_2}"] & S & \text{Zentralbäume} \\
A_1 & & A_2 & \text{Abgeleitete Objekte}
\end{tikzcd}
$$

Andere Formalismen: formell

Formalismus: endliche Kontrolle, Generation

Formalisieren über universelle Algebren (Goguen u. a. 1977)

Problem: black box

Interpretierte reguläre Baumgrammatiken (IRTGs, Koller und Kuhlmann (2011))

$$
\begin{tikzcd}
 & T_\Gamma \arrow[r, "{[\![M]\!]}"] \arrow[dl, "(.)^{\mathcal{A}_1}"'] \arrow[dr, "(.)^{\mathcal{A}_2}"] & S \\
A_1 & & A_2
\end{tikzcd}
$$

Zentralbäume

Abgeleitete Objekte

Andere Formalismen: formell

Formalismus: endliche Kontrolle, Generation

Formalisieren über universelle Algebren (Goguen u. a. 1977)

Problem: black box

Interpretierte reguläre Baumgrammatiken (IRTGs, Koller und Kuhlmann (2011))

$$
\begin{tikzcd}
 & T_\Gamma \arrow[r, "\llbracket M \rrbracket"] \arrow[dl, "h_1"'] \arrow[dr, "h_2"] & S & \text{Zentralbäume} \\
T_{\Delta_1} \arrow[d, "(.)^{\mathcal{A}_1}"'] & & T_{\Delta_2} \arrow[d, "(.)^{\mathcal{A}_2}"] & \text{Semantische Terme} \\
A_1 & & A_2 & \text{Abgeleitete Objekte}
\end{tikzcd}
$$

Andere Formalismen: formell

Formalismus: endliche Kontrolle, Generation

Formalisieren über universelle Algebren (Goguen u. a. 1977)

Problem: black box

Interpretierte reguläre Baumgrammatiken (IRTGs, Koller und Kuhlmann (2011))

$$
\begin{tikzcd}
 & T_\Gamma \arrow[r, "{[\![M]\!]}"] \arrow[dl, "h_1"'] \arrow[dr, "h_n"] & S & \text{Zentralbäume} \\
T_{\Delta_1} \arrow[d, "(.)^{\mathcal{A}_1}"'] & \cdots & T_{\Delta_n} \arrow[d, "(.)^{\mathcal{A}_n}"] & & \text{Semantische Terme} \\
A_1 & \cdots & A_n & & \text{Abgeleitete Objekte}
\end{tikzcd}
$$

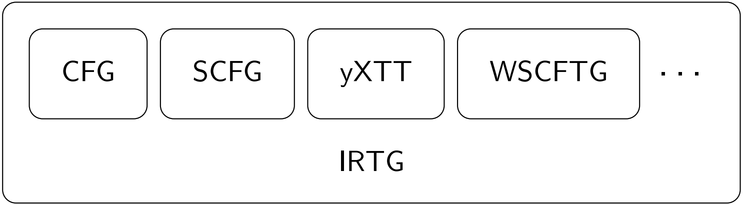
Andere Formalismen: formell

Formalismus: endliche Kontrolle, Generation

Formalisieren über universelle Algebren (Goguen u. a. 1977)

Problem: black box

Interpretierte reguläre Baumgrammatiken (IRTGs, Koller und Kuhlmann (2011))



Ergebnisse I: Generische Binarisierung

- Binarisierungsabbildung und deren Eigenschaften für IRTGs formalisiert

Ergebnisse I: Generische Binarisierung

- Binarisierungsabbildung und deren Eigenschaften für IRTGs formalisiert
- Regel-für-regel-vollständige Binarisierungsabbildung für IRTGs konstruiert

Ergebnisse I: Generische Binarisierung

- Binarisierungsabbildung und deren Eigenschaften für IRTGs formalisiert
- Regel-für-regel-vollständige Binarisierungsabbildung für IRTGs konstruiert
- Problem: Äquivalenz in beliebigen Algebren unentscheidbar!

Ergebnisse I: Generische Binarisierung

- Binarisierungsabbildung und deren Eigenschaften für IRTGs formalisiert
- Regel-für-regel-vollständige Binarisierungsabbildung für IRTGs konstruiert
- Problem: Äquivalenz in beliebigen Algebren unentscheidbar!
- "Outsourcen" des nicht berechenbaren Teils: b-rules

Ergebnisse I: Generische Binarisierung

- Binarisierungsabbildung und deren Eigenschaften für IRTGs formalisiert
- Regel-für-regel-vollständige Binarisierungsabbildung für IRTGs konstruiert
- Problem: Äquivalenz in beliebigen Algebren unentscheidbar!
- "Outsourcen" des nicht berechenbaren Teils: b-rules
- vollständige b-rules führen wieder zu regel-für-regel-vollständiger Binarisierungsabbildung

Ergebnisse I: Generische Binarisierung

- Binarisierungsabbildung und deren Eigenschaften für IRTGs formalisiert
- Regel-für-regel-vollständige Binarisierungsabbildung für IRTGs konstruiert
- Problem: Äquivalenz in beliebigen Algebren unentscheidbar!
- "Outsourcen" des nicht berechenbaren Teils: b-rules
- vollständige b-rules führen wieder zu regel-für-regel-vollständiger Binarisierungsabbildung
- Laufzeit in $O(|R| \cdot c^n)$

Gliederung

Vergleich

$$\Omega = \{(G, \mu) \mid G \text{ STSG}, \mu \text{ Wahrscheinlichkeitszuweisung}\}$$

$$\mathbb{D}_1, \mathbb{D}_2 \colon \Omega \to E^F \ ,$$

$$\mathrm{yd}(\mathrm{best}(\pi_2(\mathrm{Yd}^{-1}(1.f) \lhd [\![(G, \mu)]\!])))$$
$$= \begin{cases} \mathbb{D}_1(G, \mu)(f) & \text{falls } \mathcal{S} = (\mathbb{R}_{\geq 0}, \max, \cdot, 0, 1), \\ \mathbb{D}_2(G, \mu)(f) & \text{falls } \mathcal{S} = (\mathbb{R}_{\geq 0}, +, \cdot, 0, 1). \end{cases}$$

Vergleich

$$\Omega = \{(G, \mu) \mid G \text{ STSG}, \mu \text{ Wahrscheinlichkeitszuweisung}\}$$

$$\mathbb{D}_1, \mathbb{D}_2 \colon \Omega \to E^F \ ,$$

$$\mathrm{yd}(\mathrm{best}(\pi_2(\mathrm{Yd}^{-1}(1.f) \lhd [\![(G, \mu)]\!])))$$

$$= \begin{cases} \mathbb{D}_1(G, \mu)(f) & \text{falls } \mathcal{S} = (\mathbb{R}_{\geq 0}, \max, \cdot, 0, 1), \\ \mathbb{D}_2(G, \mu)(f) & \text{falls } \mathcal{S} = (\mathbb{R}_{\geq 0}, +, \cdot, 0, 1). \end{cases}$$

May und Knight (2006): $\mathbb{D}_2$ übersetzt besser als $\mathbb{D}_1$!

Vergleich

$$\Omega = \{(G, \mu) \mid G \text{ STSG}, \mu \text{ Wahrscheinlichkeitszuweisung}\}$$

$$\mathbb{D}_1, \mathbb{D}_2 \colon \Omega \to E^F \ ,$$

$$\mathrm{yd}(\mathrm{best}(\pi_2(\mathrm{Yd}^{-1}(1.f) \lhd [\![(G, \mu)]\!])))$$

$$= \begin{cases} \mathbb{D}_1(G, \mu)(f) & \text{falls } \mathcal{S} = (\mathbb{R}_{\geq 0}, \max, \cdot, 0, 1), \\ \mathbb{D}_2(G, \mu)(f) & \text{falls } \mathcal{S} = (\mathbb{R}_{\geq 0}, +, \cdot, 0, 1). \end{cases}$$

May und Knight (2006): $\mathbb{D}_2$ übersetzt besser als $\mathbb{D}_1$!

Problem: Berechnen von best.

Vergleich

$$\Omega = \{(G, \mu) \mid G \text{ STSG}, \mu \text{ Wahrscheinlichkeitszuweisung}\}$$

$$\mathbb{D}_1, \mathbb{D}_2 \colon \Omega \to E^F \ ,$$

$$\text{yd}(\text{best}(\pi_2(\text{Yd}^{-1}(1.f) \lhd [\![(G, \mu)]\!])))$$

$$= \begin{cases} \mathbb{D}_1(G, \mu)(f) & \text{falls } \mathcal{S} = (\mathbb{R}_{\geq 0}, \max, \cdot, 0, 1), \\ \mathbb{D}_2(G, \mu)(f) & \text{falls } \mathcal{S} = (\mathbb{R}_{\geq 0}, +, \cdot, 0, 1). \end{cases}$$

May und Knight (2006): $\mathbb{D}_2$ übersetzt besser als $\mathbb{D}_1$!

Problem: Berechnen von best.
Argument von best ist regulär; Berechnung:

Vergleich

$$\Omega = \{(G, \mu) \mid G \text{ STSG}, \mu \text{ Wahrscheinlichkeitszuweisung}\}$$

$$\mathbb{D}_1, \mathbb{D}_2 \colon \Omega \to E^F \ ,$$

$$\begin{aligned}
&\mathrm{yd}(\mathrm{best}(\pi_2(\mathrm{Yd}^{-1}(1.f) \lhd [\![(G, \mu)]\!])))\\
&= \begin{cases} \mathbb{D}_1(G, \mu)(f) & \text{falls } \mathcal{S} = (\mathbb{R}_{\geq 0}, \max, \cdot, 0, 1), \\ \mathbb{D}_2(G, \mu)(f) & \text{falls } \mathcal{S} = (\mathbb{R}_{\geq 0}, +, \cdot, 0, 1). \end{cases}
\end{aligned}$$

May und Knight (2006): $\mathbb{D}_2$ übersetzt besser als $\mathbb{D}_1$!

Problem: Berechnen von best.
Argument von best ist regulär; Berechnung:

2. besten Lauf bestimmen (kürzeste Wege, Knuth (1977))

Vergleich

$$\Omega = \{(G, \mu) \mid G \text{ STSG}, \mu \text{ Wahrscheinlichkeitszuweisung}\}$$
$$\mathbb{D}_1, \mathbb{D}_2 \colon \Omega \to E^F \ ,$$

$$\mathrm{yd}(\mathrm{best}(\pi_2(\mathrm{Yd}^{-1}(1.f) \lhd [\![(G, \mu)]\!])))$$
$$= \begin{cases} \mathbb{D}_1(G, \mu)(f) & \text{falls } \mathcal{S} = (\mathbb{R}_{\geq 0}, \max, \cdot, 0, 1), \\ \mathbb{D}_2(G, \mu)(f) & \text{falls } \mathcal{S} = (\mathbb{R}_{\geq 0}, +, \cdot, 0, 1). \end{cases}$$

May und Knight (2006): $\mathbb{D}_2$ übersetzt besser als $\mathbb{D}_1$!

Problem: Berechnen von best.
Argument von best ist regulär; Berechnung:

1. für $\mathbb{D}_2$: determinisieren (May und Knight 2006, Algorithmus 1)
2. besten Lauf bestimmen (kürzeste Wege, Knuth (1977))

Vergleich

$$\Omega = \{(G,\mu) \mid G \text{ STSG}, \mu \text{ Wahrscheinlichkeitszuweisung}\}$$

$$\mathbb{D}_1, \mathbb{D}_2 \colon \Omega \to E^F \ ,$$

$$\begin{aligned}&\mathrm{yd}(\mathrm{best}(\pi_2(\mathrm{Yd}^{-1}(1.f) \lhd [\![(G,\mu)]\!])))\\ &= \begin{cases}\mathbb{D}_1(G,\mu)(f) & \text{falls } \mathcal{S} = (\mathbb{R}_{\geq 0}, \max, \cdot, 0, 1),\\ \mathbb{D}_2(G,\mu)(f) & \text{falls } \mathcal{S} = (\mathbb{R}_{\geq 0}, +, \cdot, 0, 1).\end{cases}\end{aligned}$$

May und Knight (2006): $\mathbb{D}_2$ übersetzt besser als $\mathbb{D}_1$!

Problem: Berechnen von best.
Argument von best ist regulär; Berechnung:

1. für $\mathbb{D}_2$: determinisieren (May und Knight 2006, Algorithmus 1)
2. besten Lauf bestimmen (kürzeste Wege, Knuth (1977))

Offen: formaler Beweis zur Korrektheit von Algorithmus 1