René Maseli
Prof. Dr. Roland Meyer
TU Braunschweig

Kontextsensitive Sprachen

Hier sehen wir, dass Satzformen gezielt umgeordnet werden können:

$L_1 := \{w\#w \mid w \in \{a,b\}\} \subseteq a, b, \#^*$ ist kontextsensitiv.

Um das zu zeigen, kann man eine Typ-1-Grammatik G_1 angeben, die $\mathcal{L}(G_1) = L_1$ erfüllt.

Idee: Produziere zuerst kontext*frei* das Wort $w_1w_2 \dots w_n\#W_n \dots W_2W_1$, wobei jedes Nichtterminal $W_i \in \{A, B\}$ zu dem Terminal $w_i \in \{a, b\}$ passt. Dann übersetze jeweils nur an dem Trennsymbol und bewege die daraus entstehenden Terminale an das rechte Ende. Erhalte dabei die Reihenfolge der Nichtterminale untereinander und die der Terminale untereinander.

Betrachte die Grammatik $G_1 = \langle\{S, A, B\}, \{a, b, \#\}, P_1, S\rangle$ mit folgenden Produktionen:

$$\begin{array}{llll} S \to aSA \mid bSB \mid \# & \#A \to \#a & aA \to Aa & bA \to Ab \\ & \#B \to \#b & aB \to Ba & bB \to Bb \end{array}$$

Es gilt $\mathcal{L}_1$: Nehme ein Wort $w\#w \in L_1$ mit $w = w_1w_2 \dots w_{n-1}w_n$. Dieses kann von G_1 wiefolgt abgeleitet werden (Für alle $1 \le i \le n$ sei dabei W_i das entsprechende Nichtterminal zu w_i). $S \to^n w\#W_nW_{n-1} \dots W_2W_1 \to w\#W_{n-1} \dots W_2W_1w_n \to \dots \to w\#W_1w_2 \dots w_{n-1}w_n \to w\#w$.

Beispiel: $abab\#abab \in \mathcal{L}(G_1)$:

$S \to^4 ababSBABA \to abab\#BABA \to abab\#bABA \to^3 abab\#ABAb \to abab\#aBAb \to^2 abab\#BAab \to abab\#bAab \to abab\#Abab \to abab\#abab$

Es gilt $\mathcal{L}(G_1) \subseteq L_1$: Nehme ein Wort $w \in \mathcal{L}(G_1)$ und betrachte eine Ableitung $S \to^* w$. Es musste die Produktion $S \to \#$ genau einmal angewandt worden sein: $S \to^* uSv \to u\#v \to^* w$. Da vorher nur zwei Produktionen benutzt werden konnten, ist v genau das Rückwärtige von u in Großbuchstaben. Nach derAnwendung von $S \to \#$ kann der linke Teil der Satzform nicht verändert werden. Stattdessen ist der rechte Teil in jedem Schritt ein ‚Shuffle' aus einem Suffix von v und dem Präfix von u mit der passenden Länge, da nur die letzten Produktionen benutzt werden können. Das einzige ableitbare Terminal-Wort ist daher $w = u\#u \in L_1$.

$L_2 := \{ww \mid w \in a, b^*\} \subseteq \{a, b\}^*$ ist kontextsensitiv.

Idee: Nutze zum umsortieren immer noch ein Spezialsymbol und übersetze es anschließend in den letzten Buchstaben von w. Speichere dazu die Information, was der letzte Buchstabe ist, im Symbol. Um das leere Wort zu akzeptieren, nutze ein weiteres Nichtterminal.

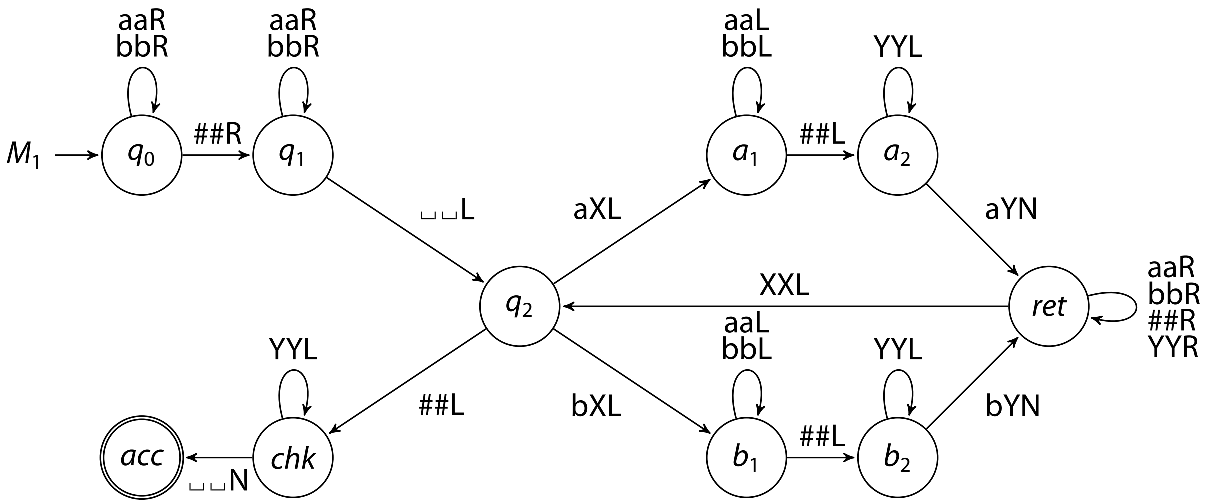
Betrachte die Grammatik $G_1 = \langle\{S, T, A, B, \#_a, \#_b\}, \{a, b\}, P_2, S\rangle$ mit folgenden Produktionen:

$S \to \varepsilon \mid T$	$\#_a A \to \#_a a$	$\#_b A \to \#_b a$	$aA \to Aa$	$bA \to Ab$
$T \to aTA \mid bTB \mid \#_a a \mid \#_b b$	$\#_a B \to \#_a b$	$\#_b B \to \#_b b$	$aB \to Ba$	$bB \to Bb$

Turing-Maschinen

Betrachte die folgende Turing-Maschine $M_1 = \langle\{q_0, q_1, q_2, a_1, a_2, b_1, b_2, ret, chk, acc\}, \{a, b, \#\}, \{a, b, \#, \sqcup, X, Y$
mit $\mathcal{L}(M_1) = L_1$. Die Transitionsrelation δ ist durch den folgenden **Transitionsgraphen** gegeben.

Idee: Prüfe zunächst, ob die Eingabe die reguläre Form $\{a, b\}^*.\{\#\}.\{a, b\}^*$ hat. Vergleiche anschließend buchstabenweise von rechts nach links und markiere die geprüften Zeichen. Zuletzt prüfe, dass alle Buchstaben markiert wurden.



Bemerkung: Mehrere Label können sich den selben Pfeil teilen. Für die Lesbarkeit wurde auf Klammern und Trennsymbole bei den Labeln verzichtet.

Beispiel: $ab\#ab \in \mathcal{L}(M_1)$:

$q_0 ab\#ab \to^* ab\#q_1 ab \to^* ab\#aq_2 b\sqcup \to ab\#b_1 aX\sqcup \to^* ab_2 b\#aX\sqcup \to^* aretY\#aX\sqcup \to^* aY\#q_2 aX\sqcup$
$\to aYa_1\#XX\sqcup \to^* aa_2 Y\#XX\sqcup \to^* retYY\#XX\sqcup \to^* YYq_2\#XX\sqcup \to YchkY\#XX\sqcup \to^* acc\sqcup YY\#XX\sqcup$

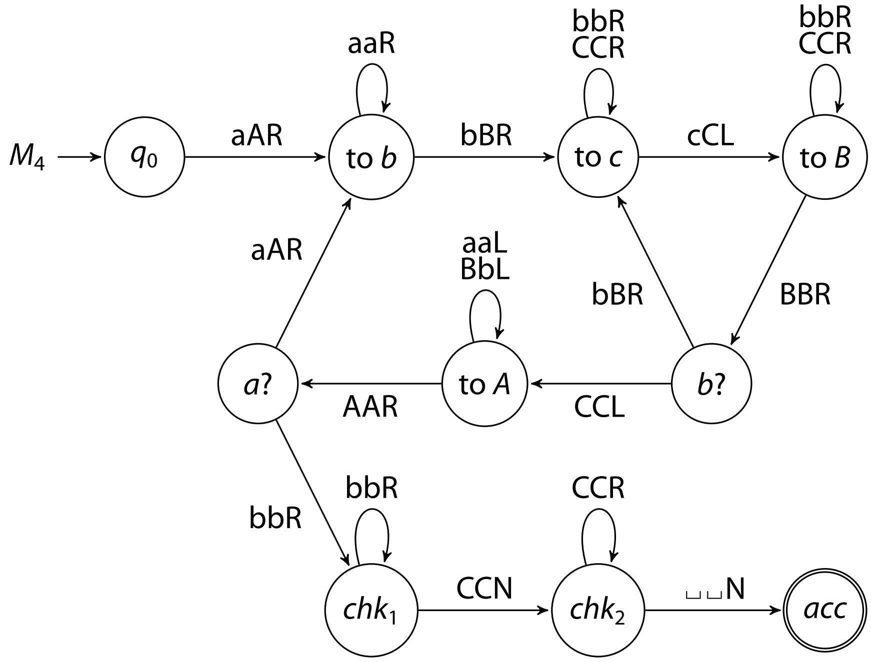
$L_3 := \{a^m b^n c^{m+n} \mid m, n \geq 1\}$ ist bekannterweise kontextfrei.

$L_4 := \{a^m b^n c^{mn} \mid m, n \geq 1\}$ ist kontextsensitiv.

Hier greifen wir voraus, dass die Sprachen von linear-beschränkten Turing-Maschinen genau die kontextsensitiven Sprachen sind, und konstruieren eine Turing-Maschine M_4 mit $\mathcal{L}(M_4) = L_4$, die während ihrer Berechnung keinen weiteren Speicher nutzt. (Auf die gleiche Weise eignet sich auch M_1 oben für einen Beweis der Kontextsensitivität.)

Idee:

Erwarte mindestens ein *a* und ein *b*.
for *a* in *w* **do**
    for *b* in *w* **do**
        Markiere ein *c*.
    end for
end for
Zuletzt prüfe, dass sich hinter den *c* keine weiteren Zeichen befinden.



(Zusatz:) Eine Typ-1-Grammatik G_4 mit $\mathcal{L}(G_4) = L_4$. Sie schickt für jedes *a* eine ‚Nachricht' nach rechts, die für jedes *b* ein neues *c* erzeugt.

$S \to TB$	$U \to aUA \mid aA$	$Cb \to bC$	$CB \to Bc$
$T \to Tb \mid U$	$B \to b$	$Ab \to bAC$	$AB \to Bc$