Theoretische Informatik 2
Große Übung 1

$L_1 := \{w\#w \mid w \in \{a,b\}^*\} \subseteq \{a,b,\#\}^*$ 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 auf kontext*freie* Art das Wort $w_1W_1w_2W_2\dots w_nW_n\#$, wobei jedes Nichtterminal $W_i \in \{A,B\}$ zu dem Terminal $w_i \in \{a,b\}$ passt. Dann bewege die Nichtterminale nach rechts, ohne deren Reihenfolge zu ändern.

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

$$S \to \# \mid aAS \mid bBS \qquad Aa \to aA \qquad Ab \to bA \qquad A\# \to \#a$$
$$Ba \to aB \qquad Bb \to bB \qquad B\# \to \#b$$

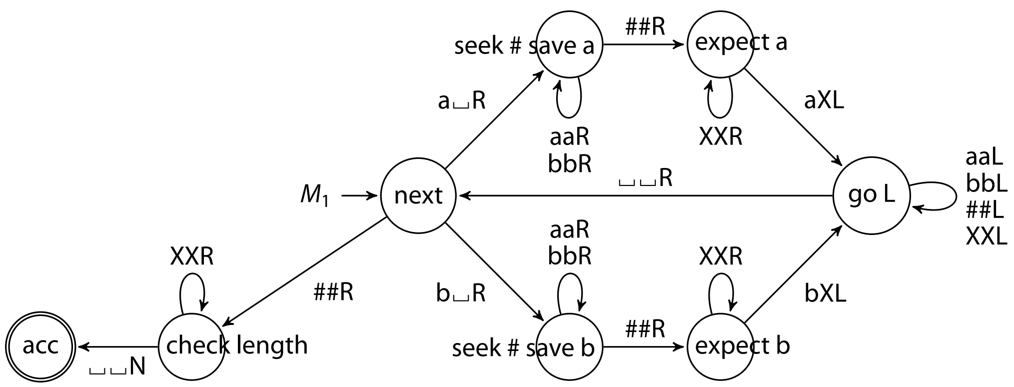
Es gilt $L_1 \subseteq \mathcal{L}(G_1)$: Nehme ein Wort $w\#w \in L_1$ mit $w = w_1\dots w_n$. Dieses kann von G_1 wiefolgt abgeleitet werden (Für alle $i \leq n$ sei dabei W_i das entsprechende Nichtterminal zu w_i).
$S \to^{n+1} w_1W_1\dots w_nW_n\# \to^{n(n-1)/2} w_1\dots w_nW_1\dots W_n\# \to^n w_1\dots w_n\#w_1\dots w_n = w\#w$.

Beispiel: $S \to^5 aAbBaAbB\# \to^6 ababABAB\# \to^4 abab\#abab$

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: $S \to^* xS \to x\# \to^* w$. Betrachte $\Gamma = \{a,b,A,B\}$ und die Homomorphismen $g,h : \Gamma^* \to \{a,b\}^*$ mit $g(abAB) = ab\varepsilon\varepsilon$ und $h(abAB) = \varepsilon\varepsilon ab$. Jede Satzform vor $S \to \#$ hat die Form xS mit $x \in \Gamma^*$ und $g(x) = h(x)$ (Beweis durch Induktion). Nachher hat jede Satzform die Form $y\#z$ mit $x \in \Gamma^*, y \in \{a,b\}^*$ und $g(x) = h(x).z$ (auch induktiv). Da dies $w \in \{a,b,\#\}^*$ einschließt, folgt $w \in L_1$.

Man kann einen LBA aus G_1 mit dem Satz von Kuroda erzeugen. Stattdessen konstruieren wir eine gewöhnliche Turingmaschine direkt.

Idee: Vergleiche buchstabenweise von links nach rechts und markiere die geprüften Zeichen. Zuletzt prüfe, ob alle Buchstaben markiert wurden.



Es gilt $L_1 \subseteq \mathcal{L}(M_1)$: Nehme ein beliebiges Wort $w\#w \in L_1$ mit $w = w_1 \dots w_n$. Die einzige Berechnung von M_1 ist next $w_1 \dots w_n\#w_1 \dots w_n \to^* \sqcup^n\text{next}\#X^n \to \sqcup^n\#\text{checklength}X^n \to^* \sqcup^n\#X^n\text{acc}$.

Beispiel: $ab\#ab \in \mathcal{L}(M_1)$:
$q_0ab\#ab \to^* ab\#q_1ab \to^* ab\#aq_2b\sqcup \to ab\#b_1aX\sqcup \to^* ab_2b\#aX\sqcup \to^* aretY\#aX\sqcup \to^* aY\#q_2aX\sqcup$
$\to aYa_1\#XX\sqcup \to^* aa_2Y\#XX\sqcup \to^* retYY\#XX\sqcup \to^* YYq_2\#XX\sqcup \to YchkY\#XX\sqcup \to^* acc\sqcup YY\#XX\sqcup$

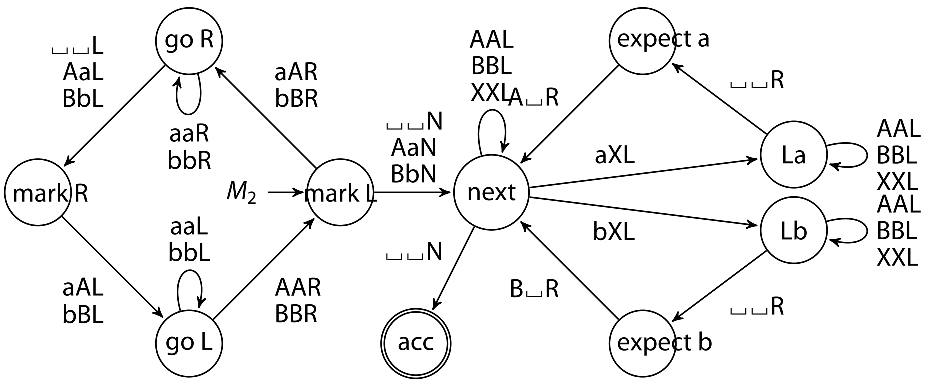
Es gilt $\mathcal{L}(M_1) \subseteq L_1$: Nehme $x \in \mathcal{L}(M_1)$. Da keine Transition $\#$ schreibt, wenn es nicht schon vorher im Band stand, und das Symbol gelesen werden muss, um check length und nachher acc zu erreichen, sei $i + 1$ der erste Index eines $\#$ in x. x muss die Form $w_1 \dots w_n\#y_1 \dots y_m$ mit $w_i \in \{a, b\}$ haben. Wann immer die Berechnung zum i-ten Mal go L verlässt, wird die Konfiguration $\sqcup^i$next $w_{i+1} \dots w_n\#X^iy_{i+1} \dots y_m$ und es gilt dann $w_1 \dots w_i = y_1 \dots y_i$. Sobald check length erreicht wird, ist die Konfiguration $\sqcup^n\#$check length $X^ny_{n+1} \dots y_m$ und es gilt $w_1 \dots w_n = y_1 \dots y_n$. Da acc erreicht wird, folgt nun $n = m$. Also gilt $x = w_1 \dots w_n\#y_1 \dots y_n = w_1 \dots w_n\#w_1 \dots w_n \in L_1$.

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

$$\begin{array}{llllll}
S \to \varepsilon \mid T & Aa \to aA & Ab \to bA & A\#_a \to \#_aa & A\#_b \to \#_ba & \#_a \to a \\
T \to aAT \mid bBT \mid \#_aa \mid \#_bb & Ba \to aB & Bb \to Bb & B\#_a \to \#_ab & B\#_b \to \#_bb & \#_b \to b
\end{array}$$

Idee für TM: Springe nichtdeterministisch zur Mitte des Wortes, oder ermittle sie deterministisch mit Markierungen A, B (wie gezeigt). Erzeuge so ein Bandinhalt der Form $(A+B)^n(a+b)^n$. Anschließend vergleiche buchstabenweise, ähnlich wie M_1.

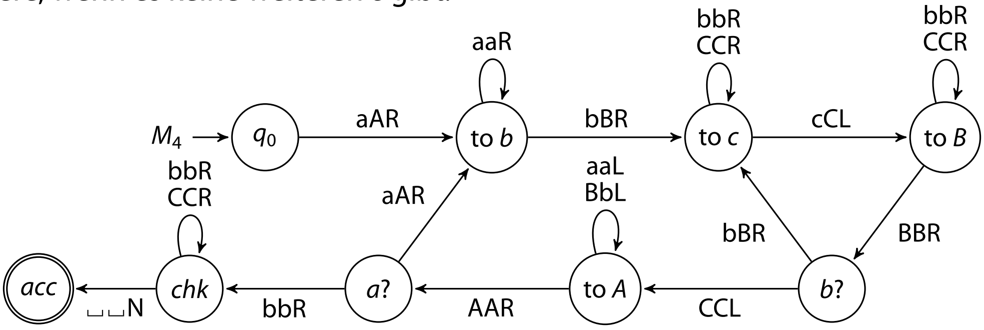


$L_3 := \{a^m b^n c^{m+n} \mid m, n \geq 1\}$ ist bekannterweise kontextfrei.

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

Idee: Implementiere den folgenden Algorithmus.

for *a* in *w* **do**
  for *b* in *w* **do**
    Markiere ein *c*.
  end for
end for
Akzeptiere, wenn es keine weiteren *c* gibt.

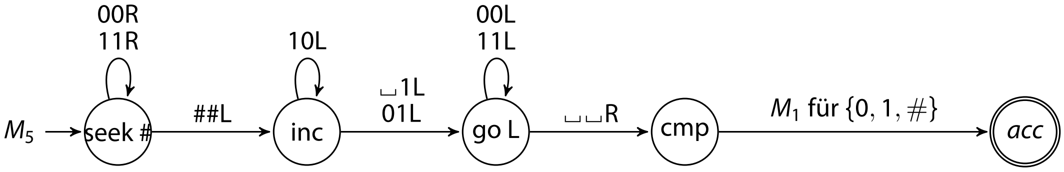


(Zusatz:) Eine Typ-1-Grammatik G_4 mit $\mathcal{L}(G_4) = L_4$. Sie schickt für jedes *a* eine ‚Nachricht' nach rechts, die für jedes *b* ein neues *c* erzeugt.

$$
\begin{array}{llll}
S \to TB & U \to aUA \mid aA & Cb \to bC & CB \to Bc \\
T \to Tb \mid U & B \to b & Ab \to bAC & AB \to Bc
\end{array}
$$

$L_5 := \{\mathrm{bin}(n)\#\mathrm{bin}(n+1) \mid n \in \mathbb{N}\}$ ist kontextsensitiv.

Idee: Erhöhe die linke Zahl um 1, dann vergleiche buchstabenweise.



Bemerkung: Anders als bei M_1, M_2 und M_4, braucht M_5 manchmal eine weitere Bandzelle. Wenn man also mit M_5 als Basis einen LBA konstruieren will, muss man zusätzliche Zustände oder Bandsymbole benutzen.

(Zusatz:) Eine Typ-1-Grammatik G_5 mit $\mathcal{L}(G_5) = L_5$:

$$
\begin{array}{lll}
S \to \#1 \mid 1\underline{1}\underline{0}U \mid 1\underline{1}T & \underline{0}0 \to 0\underline{0} & \underline{1}0 \to 0\underline{1} \\
T \to 0\underline{1}U \mid 0\underline{0}T \mid 1\underline{1}T & \underline{0}1 \to 1\underline{0} & \underline{1}1 \to 1\underline{1} \\
U \to \# \mid 1\underline{0}U & \underline{0}\# \to \#0 & \underline{1}\# \to \#1
\end{array}
$$