Postsches Korrespondenzproblem

Simon Schrodi

DHBW Karlsruhe

12.04.2019

Motivation

Gegeben zwei kontextfreie Grammatiken G_1, G_2, so sind folgende Probleme unentscheidbar:

- Ist $L(G_1) \cap L(G_2) = \emptyset$? (Schnittproblem)
- Ist $|L(G_1) \cap L(G_2)| = \infty$? (Endlichkeitsproblem)
- Ist $L(G_1) \subseteq L(G_2)$? (Inklusionsproblem)
- Ist $L(G_1) = L(G_2)$? (Äquivalenzproblem)
- Ist G_1 mehrdeutig? (Mehrdeutigkeitsproblem)
- ...

Wiederholung: Reduktion

Seien $A \subseteq \Sigma^*$ und $B \subseteq \Gamma^*$ Sprachen. $A \leq B$ gdw. es gibt eine berechenbare Funktion $f : \Sigma^* \to \Gamma^*$, so dass $\forall x \in \Sigma^*$ gilt:

$$x \in A \Leftrightarrow f(x) \in B. \tag{0}$$

Falls $A \leq B$ und A unentscheidbar, so ist auch B unentscheidbar.

Definition: Postsches Korrespondenzproblem (PKP)

Eine Instanz des PKP besteht aus einer **endlichen Folge**

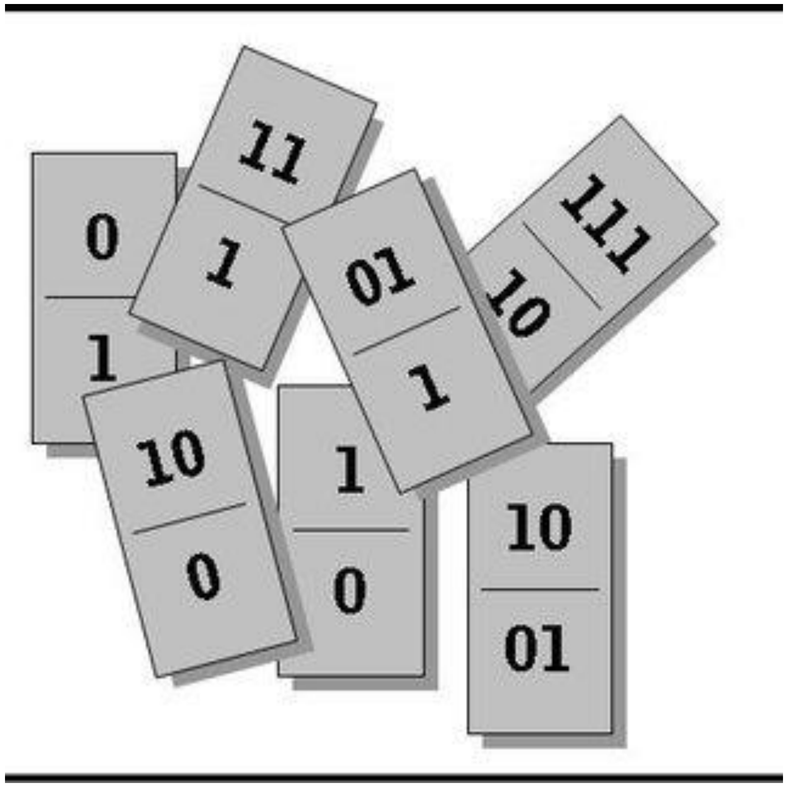
$$K = [(x_1, y_1), ..., (x_k, y_k)], \tag{1.1}$$

wobei $x_i, y_i \neq \epsilon$ über einem endlichen Alphabet Σ sind. Es soll entschieden werden, ob es eine **korrespondierende Folge**

$$i_1, ..., i_n \in [1, ..., k], n \geq 1 \tag{1.2}$$

von Indizes, genannt **Lösung**, gibt, so dass gilt

$$x_{i_1} x_{i_2} ... x_{i_n} = y_{i_1} y_{i_2} ... y_{i_n}. \tag{1.3}$$



Bildquelle: https://www.researchgate.net/profile/Klaus_Jantke/publication/297212574/figure/fig2/AS:337034053996545@1457366588223/Abbildung-23-Unloesbare-Problemstellung-Dass-es-fuer-den-Pool-von-Wortpaaren-in-der-Abb_Q320.jpg

Definition: Modifiziertes PKP (MPKP)

- Wie PKP, außer dass Indexfolge mit $\boldsymbol{i_1 = 1}$ beginnen muss

Beispiel PKP 1

Gegeben: $K = [(10111, 10), (1, 111), (10, 0)]$
Gesucht: Korrespondierende Indexfolge

Lösung: Die Indexfolge ist $(1, 2, 2, 3)$:

$$\underbrace{10111}_{x_1}\underbrace{1}_{x_2}\underbrace{1}_{x_2}\underbrace{10}_{x_3} = 101111110 = \underbrace{10}_{y_1}\underbrace{111}_{y_2}\underbrace{111}_{y_2}\underbrace{0}_{y_3}$$

Beispiel PKP 2

Gegeben: $K = [(10, 101), (011, 11), (101, 011)]$

Lösung: Es gibt keine passende Indexfolge (**Zugzwangargument**)

- Jede potentielle Lösung muss mit $i_1 = 1$ beginnen
- Immer wenn y-Sequenz eine 1 Vorsprung hat, ist die einzig mögliche Fortsetzung:

$$x - Sequenz: \ ... \ \underbrace{101}_{x_3}$$

$$y - Sequenz: \ ... \ 1\,\underbrace{011}_{y_3}$$

Die y-Sequenz hat eine 1 Vorsprung

Beispiel PKP 3

Gegeben: $K = [(001, 0), (01, 011), (01, 101), (10, 001)]$

Lösung: $(2, 4, 3, 4, 4, 2, 1, 2, 4, 3, 4, 3, ...)$ mit 66 Indizes

Satz: MPKP ist semi-entscheidbar

- Kombinatorische Entscheidungssuche
- Tiefensuche oder Breitensuche?

Lemma: MPKP ist unentscheidbar

Zeige:

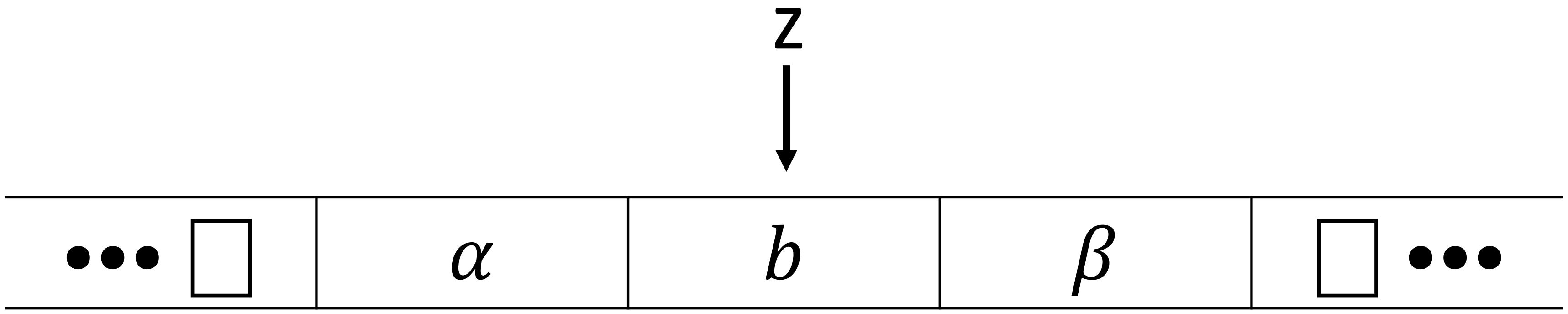
$$H \leq MPKP, \tag{2.1}$$

indem eine Reduktionsabbildung f gegeben wird, die die Eingaben vom allgemeinen Halteproblem $H = \{(p, w) | \text{Turingprogramm p angesetzt auf w hält}\}$, so auf Eingaben von MPKP abbildet, dass:

$$(p, w) \in H(M_p) \Leftrightarrow f(p, w) \in MPKP \tag{2.2}$$

Konfiguration einer Turingmaschine M_w

- Konfiguration einer Turingmaschine ist ein Wort $k = \alpha z b \beta \in \Gamma^* Z \Gamma^*$
- Momentaufnahme der Turingmaschine:



Beweisidee

- Stelle x- und y-Sequenzen als Konfigurationsfolgen von TM dar
- Die y-Sequenz hat immer eine Konfiguration Vorsprung
- Die x-Sequenz holt nach dem Stoppen von TM den Vorsprung ein

Beispiel der Simulation

Sei eine Turingmaschine M mit Eingabe $w = 01$ gegeben:

$$M = (\underbrace{\{z_1, z_2, z_f\}}_{Z}, \underbrace{\{0,1\}}_{\Sigma}, \underbrace{\{0,1,\square\}}_{\Gamma}, \delta, \underbrace{z_1}_{z_1 \in Z}, \square, \underbrace{\{z_f\}}_{E})$$

δ	0	1	$\square$
z_1	$(z_2, 1, R)$	$(z_2, 0, L)$	$(z_2, 1, L)$
z_2	$(z_f, 0, L)$	$(z_1, 0, R)$	$(z_2, 0, R)$
z_f	—	—	—

Beispiel der Simulation: Lösung 1

- Regelfolge: **A(nfangsregel)**

- x-Sequenz: #
- y-Sequenz: $\#z_1 01\#$

- M_w: $z_1 01$

x- und y-Sequenz als Konfigurationsfolge

y-Sequenz hat immer eine Konfiguration Vorsprung

δ	0	1	$\square$
z_1	$(z_2, 1, R)$	$(z_2, 0, L)$	$(z_2, 1, L)$
z_2	$(z_f, 0, L)$	$(z_1, 0, R)$	$(z_2, 0, R)$
z_f	—	—	—

Beispiel der Simulation: Lösung 2

- Regelfolge: A, **(R)echts-Überführungsregel**

- x-Sequenz: $\#\boldsymbol{z_1 0}$
- y-Sequenz: $\#z_1 01\#\boldsymbol{1 z_2}$

y-Sequenz hat immer eine Konfiguration Vorsprung

- $M_w: z_1 01 \vdash \boldsymbol{1 z_2} \boldsymbol{1}$

δ	0	1	$\square$
z_1	$(z_2, 1, R)$	$(z_2, 0, L)$	$(z_2, 1, L)$
z_2	$(z_f, 0, L)$	$(z_1, 0, R)$	$(z_2, 0, R)$
z_f	—	—	—

Beispiel der Simulation: Lösung 3

- Regelfolge: A, R,**(K)opierregel**

- x-Sequenz: $\#z_1 0\mathbf{1}$
- y-Sequenz: $\#z_1 01\#1z_2\mathbf{1}$

y-Sequenz hat immer eine Konfiguration Vorsprung

- M_w: $z_1 01 \vdash \mathbf{1z_2}\mathbf{1}$

δ	0	1	$\square$
z_1	$(z_2, 1, R)$	$(z_2, 0, L)$	$(z_2, 1, L)$
z_2	$(z_f, 0, L)$	$(z_1, 0, R)$	$(z_2, 0, R)$
z_f	—	—	—

Beispiel der Simulation: Lösung 4

- Regelfolge: A, R, K, **K**

- x-Sequenz: $\#z_1 01\#$
- y-Sequenz: $\#z_1 01\#1z_2 1\#$

y-Sequenz hat immer eine Konfiguration Vorsprung

- $M_w: z_1 01 \vdash \mathbf{1z_2 1}$

δ	0	1	$\square$
z_1	$(z_2, 1, R)$	$(z_2, 0, L)$	$(z_2, 1, L)$
z_2	$(z_f, 0, L)$	$(z_1, 0, R)$	$(z_2, 0, R)$
z_f	—	—	—

Beispiel der Simulation: Lösung 5

- Regelfolge: A, R, K, K, **K**

- x-Sequenz: $\#z_1 01\#\mathbf{1}$
- y-Sequenz: $\#z_1 01\#1z_2 1\#\mathbf{1}$

y-Sequenz hat immer eine Konfiguration Vorsprung

- M_w: $z_1 01 \vdash 1z_2 1 \vdash \mathbf{10}\boldsymbol{z_1}$

δ	0	1	$\square$
z_1	$(z_2, 1, R)$	$(z_2, 0, L)$	$(z_2, 1, L)$
z_2	$(z_f, 0, L)$	$(z_1, 0, R)$	$(z_2, 0, R)$
z_f	—	—	—

Beispiel der Simulation: Lösung 6

- Regelfolge: A, R, K, K, K, **R**

- x-Sequenz: $\#z_1 01\#1\boldsymbol{z_2 1}$
- y-Sequenz: $\#z_1 01\#1z_2 1\#1\boldsymbol{0z_1}$

y-Sequenz hat immer eine Konfiguration Vorsprung

- M_w: $z_1 01 \vdash 1z_2 1 \vdash \boldsymbol{10z_1}$

δ	0	1	$\square$
z_1	$(z_2, 1, R)$	$(z_2, 0, L)$	$(z_2, 1, L)$
z_2	$(z_f, 0, L)$	$(z_1, 0, R)$	$(z_2, 0, R)$
z_f	—	—	—

Beispiel der Simulation: Lösung 7

- Regelfolge: A, R, K, K, K, R, K, K, **(S1)-Sonderregel**

- x-Sequenz: $\#z_1 01\#1z_2 1\boldsymbol{\#10z_1\#}$
- y-Sequenz: $\#z_1 01\#1z_2 1\#10z_1\boldsymbol{\#1z_2 01\#}$

y-Sequenz hat immer eine Konfiguration Vorsprung

- $M_w: z_1 01 \vdash \cdots \vdash 10z_1 \vdash \boldsymbol{1z_2 01}$

δ	0	1	$\square$
z_1	$(z_2, 1, R)$	$(z_2, 0, L)$	$(z_2, 1, L)$
z_2	$(z_f, 0, L)$	$(z_1, 0, R)$	$(z_2, 0, R)$
z_f	—	—	—

Beispiel der Simulation: Lösung 8

- Regelfolge: A, R, K, K, K, R, K, K, S1, **(L)inks-Überführungsregel**, **K**, **K**

- x-Sequenz: $\#z_1 01\#1z_2 1\#10z_1\#\mathbf{1z_2 01}\#$
- y-Sequenz: $\#z_1 01\#1z_2 1\#10z_1\#1z_2 01\#\boldsymbol{z_f}\mathbf{101}\#$

- M_w: $z_1 01 \ \vdash \cdots \vdash 1z_2 01 \vdash \boldsymbol{z_f}\mathbf{101}$

δ	0	1	$\square$
z_1	$(z_2, 1, R)$	$(z_2, 0, L)$	$(z_2, 1, L)$
z_2	$(z_f, 0, L)$	$(z_1, 0, R)$	$(z_2, 0, R)$
z_f	—	—	—

Beispiel der Simulation: Lösung 9

- Regelfolge: A, R, K, K, K, R, K, K, S1, L, K, K, **(Lö)sch-Regeln**

- x-Sequenz: $\#z_1 01\#1z_2 1\#10z_1\#1z_2 01\#\boldsymbol{z_f 1}$
- y-Sequenz: $\#z_1 01\#1z_2 1\#10z_1\#1z_2 01\#z_f 101\#\boldsymbol{z_f}$

x-Sequenz holt nach Stopp Vorsprung auf

Beispiel der Simulation: Lösung 10

- Regelfolge: A, R, K, K, K, R, K, K, S1, L, K, K, L, **K**, **K**, **K**, **Lö**, **K**, **K**, **Lö**, **K**, **(Ab)schlussregel**
- x-Sequenz: $\#z_1 01\#1z_2 1\#10z_1\#1z_2 01\#z_f 1\mathbf{01}\#\boldsymbol{z_f}\mathbf{01}\#\boldsymbol{z_f}\mathbf{1}\#\boldsymbol{z_f}\#\#$
- y-Sequenz: $\#z_1 01\#1z_2 1\#10z_1\#1z_2 01\#z_f 101\#z_f\mathbf{01}\#\boldsymbol{z_f}\mathbf{1}\#\boldsymbol{z_f}\#\#$

Simulation der TM

(i) **Anfangsregel:** $(\#, \#z_1 w\#)$

(ii) **Kopierregeln:** $\forall a \in \Gamma \cup \{\#\} : (a, a)$

(iii) **Überführungsregeln:** $\forall z \in Z \backslash E; \forall z' \in Z; \forall a, c \in \Gamma \backslash \{\Box\}$:

$$(za, cz'),\ \ falls\ \delta(z, a) = (z', c, R)$$
$$(bza, z'bc),\ \ falls\ \delta(z, a) = (z', c, L),\ \ \forall b \in \Gamma$$
$$(z\#, cz'\#),\ \ falls\ \delta(z, \Box) = (z', c, R)$$
$$(bz\#, z'bc\#),\ \ falls\ \delta(z, \Box) = (z', c, L),\ \ \forall b \in \Gamma \backslash \{\Box\}$$

(iv) **Löschregeln:** $\forall z_f \in E; \forall a \in \Gamma \backslash \{\Box\}$: (az_f, z_f), $(z_f a, z_f)$

(v) **Abschlussregeln:** $\forall z_f \in E : (z_f\#\#, \#)$

Nachweis der Reduktion: „=>“

- $(p, w) \in H \Rightarrow f(p, w) \in MPKP$
  - Falls $(p, w) \in H$, erhalten wir irgendwann eine Lösung der Form $(k, k\alpha z_f \beta\#)$ mit $z_f \in E, \alpha, \beta \in \Gamma^*$
  - Mittels Kopierpaare und Löschpaare kann der Vorsprung $\alpha z_f \beta\#$ vermindert werden,
  - bis das Abschlusspaar angewendet werden kann, so dass $(k' z_f \#\#, k' z_f \#\#)$

Nachweis der Reduktion: „<="

- $f(p, w) \in MPKP \Rightarrow (p, w) \in H$
  - Es sei angenommen: $(p, w) \notin H \Rightarrow f(p, w) \notin MPKP$
  - Da kein Zustand $z_f \in E$ erreicht wird, werden keine Löschpaare angewendet, und
  - somit hat die y-Sequenz stets eine Konfiguration Vorsprung

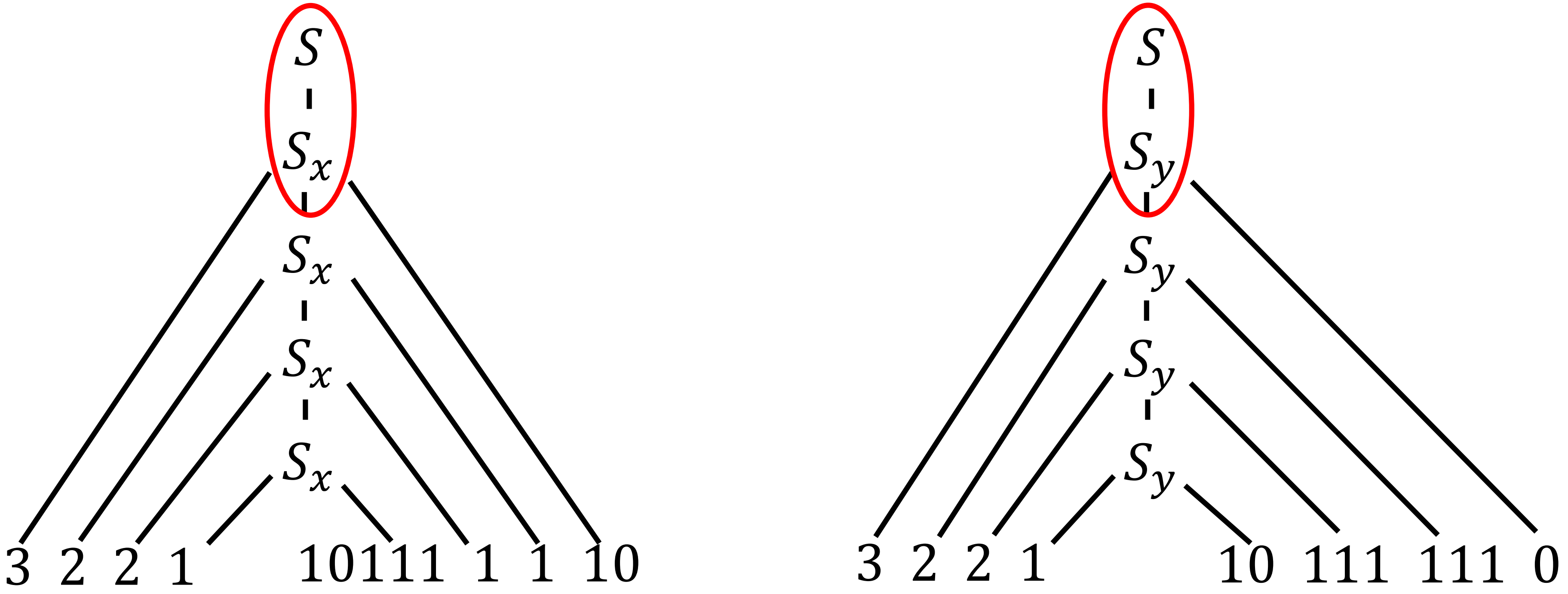
Satz: Ist G mehrdeutig, ist unentscheidbar

- Gegeben eine Instanz des MPKP mit K=$[(x_1, y_1), ..., (x_k, y_k)]$ über einem endlichen Alphabet Σ und $I = \{i_1, ..., i_k\} \notin \Sigma$
- Konstruiere kontextfreie Grammatik $G_x = (V_x, T, P_x, S_x)$ mit
  - $T = \Sigma \cup I$
  - $P_x = \{S_x \to i_1 S_x x_1 | ... | i_k S_x x_k | i_1 x_1\}$

  und eine Grammatik G_y analog mit y_i statt x_i
- Sei $L(G_z) = L(G_x) \cup L(G_y)$ mit $P_z = \{S \to S_x | S_y\} \cup P_x \cup P_y$

Satz: Ist G mehrdeutig, ist unentscheidbar

- K hat eine Lösung $\Rightarrow$ G ist mehrdeutig
  - MPKP aus Beispiel 1: $K = [(10111, 10), (1, 111), (10, 0)]$ hat Lösung $(1, 2, 2, 3)$



Satz: Ist G mehrdeutig, ist unentscheidbar

- G ist mehrdeutig $\Rightarrow$ K hat eine Lösung
  - G_z hat zwei **verschiedene** Ableitungsbäume mit **gleichem** Wort w
  - Fall 1: $S \to S_x \to w$ für beide Ableitungsbäume
    - D.h. S_x ist mehrdeutig
    - Aber G_x ist $LL(2)$ ↯
  - Fall 2: $S \to S_x \to w$ bzw. $S \to S_y \to w$
    - Erster Teil der Worte ist gleich (Indexfolge)
    - Zweiter Teil der Worte ist gleich $\to$ Lösung für K

Satz: $L(G_1) \cap L(G_2) \stackrel{?}{=} \emptyset$ ist unentscheidbar

- Gegeben eine Instanz des MPKP mit $\{(x_1, y_1), ..., (x_k, y_k)\}$ über einem endlichen Alphabet Σ und $I = \{i_1, ..., i_k\} \notin \Sigma$
- Konstruiere kontextfreie Grammatik $G_x = (V_x, T, P_x, S_x)$ und eine Grammatik G_y analog mit y_i statt x_i
- $L(G_x) \cap L(G_y) \neq \emptyset \iff \exists w \in \Sigma : w \in L(G_x) \wedge w \in L(G_y)$

Satz: $|L(G_1) \cap L(G_2)| \stackrel{?}{=} \infty$ ist unentscheidbar

- Ordne MPKP zwei Grammatiken G_x und G_y zu
- Lösungen von K entsprechen $w \in (L(G_x) \cap L(G_y))$
- Wenn MPKP (mindestens) eine Lösung hat, so hat MPKP unendlich vie Lösungen, indem die Lösungsfolge beliebig oft wiederholt wird
  - MPKP aus Beispiel 1: $K = [(10111, 10), (1, 111), (10, 0)]$ hat Lösur $(1, 2, 2, 3)$
  - Dann ist auch (1,2,2,3,1,2,2,3) eine Lösung
- Also gilt: K hat Lösung $\Longleftrightarrow |L(G_x) \cap L(G_y)| = \infty$

Anhang

Satz: PKP ist unentscheidbar

Zeige:

$$MPKP \leq PKP. \tag{2.0}$$

- Eingabeinstanz von MPKP über Alphabet Σ sei $K = [(x_1, y_1), ..., (x_k, y_k)]$
- Es sei $\#, \$ \notin \Sigma$ ein neues Symbol, so dass gilt:

  $$f(K) = [(x'_0, y'_0), (x'_1, y'_1), ..., (x'_k, y'_k), (x'_{k+1}, y'_{k+1})]$$

  mit

  - $x'_0 = \#x'_1, x_{k+1} = \$, y'_0 = y'_1, y'_{k+1} = \#\$$
  - $\forall i \in \{1, ..., k\}\ gilt\ x'_i = x_i\#\ bzw.\ y'_i = \#y_i$
- Aus der Konstruktion von $f(K)$ ergibt sich leicht $MPKP \leq PKP$

Beispiel $(M)PKP \leq PKP$

Indexfolge für (M)PKP ist $(1, 2, 2, 3)$:

$$\underbrace{10111}_{x_1}\underbrace{1}_{x_2}\underbrace{1}_{x_2}\underbrace{10}_{x_3} = 101111110 = \underbrace{10}_{y_1}\underbrace{111}_{y_2}\underbrace{111}_{y_2}\underbrace{0}_{y_3}$$

Indexfolge für PKP ist $(0, 2, 2, 3, 4)$:

$$\underbrace{\#1\#0\#1\#1\#1\#}_{x'_0}\ \underbrace{1\#}_{x'_2}\ \underbrace{1\#}_{x'_2}\ \underbrace{1\#0\#}_{x'_3}\ \underbrace{\$}_{x'_4} = \underbrace{\#1\#0}_{y'_0}\underbrace{\#1\#1\#1}_{y'_2}\underbrace{\#1\#1\#1}_{y'_2}\underbrace{\#0}_{y'_3}\underbrace{\#\$}_{y'_4}$$

Beispiel $(M)PKP \leq PKP$

Gegeben: $K = [(\underbrace{10111}_{x_1}, \underbrace{10}_{y_1}), (\underbrace{1}_{x_2}, \underbrace{111}_{y_2}), (\underbrace{10}_{x_3}, \underbrace{0}_{y_3})]$.

So ist

$$f(K) = [(\underbrace{1\#0\#1\#1\#1\#}_{x'_1} \underbrace{\#1\#0}_{y'_1}), (\underbrace{1\#}_{x'_2}, \underbrace{\#1\#1\#1}_{y'_2}), \underbrace{\#1\#0}_{y'_2}), (\underbrace{1\#0\#}_{x'_3}, \underbrace{\#0}_{y'_3})]$$

$$\cup[(\underbrace{\#1\#0\#1\#1\#1\#}_{x'_0}, \underbrace{\#1\#0}_{y'_0}), (\underbrace{\#}_{x'_4}, \underbrace{\#\#}_{y'_4})]$$

Beispiel der Simulation: Konstruktion 1

Regeltyp	Regel.Index	x-Sequenz	y-Sequenz	Quelle
Anfangsregel	(i).0	#	$\#z_0 01\#$	
Kopierregeln	(ii).0	0	0	
	(ii).1	1	1	
	(ii).2	#	#	
Überführungsregeln	(iii).0	$z_1 0$	$1 z_2$	$\delta(z_1, 0) = (z_2, 1, R)$
	(iii).1	$0 z_1 1$	$z_2 00$	$\delta(z_1, 1) = (z_2, 0, L)$
	(iii).2	$1 z_1 1$	$z_2 10$	$\delta(z_1, 1) = (z_2, 0, L)$
	(iii).3	$0 z_1 \#$	$z_2 01\#$	$\delta(z_1, \square) = (z_2, 1, L)$
	(iii).4	$1 z_1 \#$	$z_2 11\#$	$\delta(z_1, \square) = (z_2, 1, L)$
	(iii).5	$0 z_2 0$	$z_f 00$	$\delta(z_2, 0) = (z_f, 0, L)$
	(iii).6	$1 z_2 0$	$z_f 10$	$\delta(z_2, 0) = (z_f, 0, L)$
	(iii).7	$z_2 1$	$0 z_1$	$\delta(z_2, 1) = (z_1, 0, R)$
	(iii).8	$z_2 \#$	$0 z_2 \#$	$\delta(z_2, \square) = (z_2, 0, R)$

Beispiel der Simulation: Konstruktion 2

Regeltyp	Regel.Index	x-Sequenz	y-Sequenz	Quelle
Löschregeln	(iv).0	$0z_f0$	z_f	
	(iv).1	$0z_f1$	z_f	
	(iv).2	$1z_f0$	z_f	
	(iv).3	$1z_f1$	z_f	
	(iv).4	$0z_f$	z_f	
	(iv).5	$1z_f$	z_f	
	(iv).6	z_f0	z_f	
	(iv).7	z_f1	z_f	
Abschlussregeln	(v).0	$z_f\#\#$	$\#$	

Satz: $L(G_1) \subseteq L(G_2)$ und $L(G_1) = L(G_2)$, sind unentscheidbar

- Ordne PKP zwei Grammatiken $G_x = (V_x, \Sigma, P_x, S_x)$ und $G_y(V_y, \Sigma, P_y, S_y)$ zu
- L_x und L_y sind deterministisch kontextfrei
- Es gibt $\bar{G_x}$ sowie $\bar{L_x}$ und es gilt $L(G_{\bar{x}y}) = L(\bar{G_x}) \cup L(G_y)$
- Behauptung: $L(G_x) \cup L(G_y) \stackrel{?}{=} \emptyset \mapsto L(G_{\bar{x}y}) = L(\bar{G_x})$

$$\begin{aligned} L(G_x) \cap L(G_y) = \emptyset &\Leftrightarrow L(G_y) \subseteq L(\bar{G_x}) \\ &\Leftrightarrow L(G_y) \cup L(\bar{G_x}) = L(\bar{G_x}) \\ &\Leftrightarrow L(G_{\bar{x}y}) = L_{\bar{G_x}} \end{aligned}$$

- Daraus folgt Inklusionsproblem und Äquivalenzproblem sind unentscheidbar
- Aber: Äquivalenzproblem für deterministisch kontextfreie Sprachen ist entscheidbar!