4.1 Axioma is steeds S, we geven alleen de producties, zonder alle componenten van de grammatica te specificeren.

a Genereer $a^+\{a^jb^j \mid j \in \mathbb{N}\}c^*$

$S \to AXC \quad A \to aA \quad A \to a$
$X \to aXb \quad X \to \lambda \quad C \to cC \quad C \to \lambda$

b Genereer $\{a^ib^jc^jc^i \mid i,j \in \mathbb{N}\}$

$S \to aSc \quad S \to T \quad T \to bTc \quad T \to \lambda$

c Genereer $\{a^ib^ib^kc^k \mid i,k \in \mathbb{N}\}$

$S \to XY \quad X \to aXb \quad X \to \lambda \quad Y \to bYc \quad Y \to \lambda$

d $S \to aSa \quad S \to bSb \quad S \to a \quad S \to b \quad S \to \lambda$

e Genereer $uwcvx$ met $x = \mathrm{mir}(w)$, $u,v \in (\{a,b\}^*c)^* = \{a,b,c\}^*c \cup \{\lambda\}$

$S \to RM \quad S \to M$
$R \to aR \quad R \to bR \quad R \to cR \quad R \to c$
dus R genereert $\{a,b,c\}^*c = (\{a,b\}^*c)^+$
$M \to aMa \quad M \to bMb \quad M \to c \quad M \to cR$

f $S \to aSbS \quad S \to bSaS \quad S \to \lambda$
maar: hoe zien we in dat deze producties echt álles in de taal kunnen genereren?

4.2 $S \to AbR \mid BaR$
$A \to LAL \mid aR\$$
$B \to LBL \mid bR\$$
$R \to LR \mid \lambda$
$L \to a \mid b$

4.3a $S \to PC \quad C \to cC \mid \lambda \quad P \to aPa \mid bPb \mid a \mid b \mid \lambda$

b $S \to aSa \mid bSb \mid C \quad C \to cC \mid \lambda$

4.4 Als $L \in \mathsf{RLIN}$, dan is er een rechtslineaire grammatica G voor L. Laat nu G' de grammatica zijn die uit G ontstaat door alle producties van de vorm $A \to aB$ te vervangen door $A \to aBa$, en alle producties van de vorm $A \to \lambda$ te vervangen door $A \to \$$ (hierbij A, B niet-terminaal, a terminaal). Dan is G' een cf grammatica die de taal $\{w\$\,\mathrm{mir}(w) \mid w \in L(G)\}$ voortbrengt.

4.5b Een links-preferente afleiding is: $S \Rightarrow XY \Rightarrow abY \Rightarrow abQYZ \Rightarrow abbYZ \Rightarrow abbbZZ \Rightarrow abbbaQZ \Rightarrow abbbabZ \Rightarrow abbbabaQ \Rightarrow abbbabab$

c Voor Q is er slechts één productie, die terminalen oplevert. We kunnen een eenvoudigere, equivalente, grammatica krijgen door de productie $Q \to b$ te 'substitueren' in de rechterkanten van de overige producties, dwz. we vervangen $X \to aXQ$ door $X \to aXb$, $Y \to QYZ$ door $Y \to bYZ$, en $Z \to aQ$ door $Z \to ab$. Hierna geldt dezelfde observatie voor deze laatste productie.

We vinden zo de volgende grammatica

$S \to XY \quad X \to aXb \quad X \to ab$
$Y \to bYab \quad Y \to bab$

Nu verifiëren we eenvoudig dat $L(G) = \{\ a^m b^m b^n (ab)^n \mid m, n \geq 1\ \}$.

4.6 De verdwijnende niet-terminalen zijn E, W en L. Een met G equivalente λ-vrije cf grammatica heeft de volgende producties:

$S \to$ *begin int* L **;** *real* L **;** *skip end*
$S \to$ *begin int* **;** *real* L **;** *skip end*
$S \to$ *begin int* L **;** *real* **;** *skip end*
$S \to$ *begin int* **;** *real* **;** *skip end*

$L \to WE \quad L \to E \quad L \to W$

$E \to \mathbf{,}WE \quad E \to \mathbf{,}E \quad E \to \mathbf{,}W \quad E \to \mathbf{,}$

$W \to aW \quad W \to a \quad W \to bW \quad W \to b$

4.7a Bepaal de productieve symbolen. Achtereenvolgens vinden we:
(0) a, b (1) A, B (2) S, D, H (3) E.
Verwijder de niet-productieve C en F met de vier producties waar ze in voorkomen.

Bepaal de bereikbare symbolen. Achtereenvolgens vinden we:
(1) S (2) B, A (3) a, D, b.
Verwijder de onbereikbare E en H met hun twee producties.

b Een equivalente gereduceerde grammatica heeft axioma S, terminalen a, b, en producties:
$S \to SBS$, $S \to A$, $A \to aBD$, $A \to a$, $B \to bB$, $B \to b$, $D \to AA$.

Een equivalente sterk gereduceerde grammatica heeft axioma S', terminalen a, b, en producties:
$S' \to SBS$, $S' \to A$, $S \to SBS$, $S \to A$,
$A \to aBD$, $A \to a$, $B \to bB$, $B \to b$, $D \to AA$.

c E is productief: zie **a**. E is bereikbaar via $S \to EC$ in G. E is niet terminaal bereikbaar want de tegelijk geïntroduceerde niet-terminaal C is niet productief.

4.8 Verwijder de keten-regels:
$\langle\text{sum}\rangle \to \langle\text{term}\rangle$, $\langle\text{term}\rangle \to \langle\text{fact}\rangle$, $\langle\text{fact}\rangle \to \langle\text{var}\rangle$.
Voeg dan alle 'doorgeschoven' producties toe:
$\langle\text{sum}\rangle \to \langle\text{term}\rangle \otimes \langle\text{fact}\rangle$, $\langle\text{sum}\rangle \to \langle\text{term}\rangle / \langle\text{fact}\rangle$, $\langle\text{sum}\rangle \to \mathbf{(}\langle\text{sum}\rangle\mathbf{)}$,
$\langle\text{sum}\rangle \to a$, $\langle\text{sum}\rangle \to b$,
$\langle\text{term}\rangle \to \mathbf{(}\langle\text{sum}\rangle\mathbf{)}$, $\langle\text{term}\rangle \to a$, $\langle\text{term}\rangle \to b$, $\langle\text{fact}\rangle \to a$, $\langle\text{fact}\rangle \to b$.

Merk op dat $\langle\text{var}\rangle$ onbereikbaar is geworden.

4.9a Omdat iedere rechts-lineaire grammatica tot de klasse $\mathcal{KL}$ behoort, geldt voor de bijbehorende talen $\mathsf{REG} \subseteq \mathsf{KL}$.

Laat nu G een cf grammatica uit $\mathcal{KL}$ zijn. De met G equivalente keten-vrije grammatica G' die verkregen wordt met de constructie van Stelling 4.34 is een rechtslineaire grammatica. En dus geldt ook $\mathsf{KL} \subseteq \mathsf{REG}$.

b Analoog aan het bewijs van Stelling 3.66 kan bewezen worden dat eindige automaten met λ-takken corresponderen met cf grammatica's uit $\mathcal{KL}$.

Opmerking. De constructie in het bewijs van Stelling 4.34 komt overeen met de λe-methode voor het verwijderen van λ-takken.

4.10a $2k - 1$

b $\frac{1}{2}k(k+1)$

4.11a De grammatica volgt de inductieve definitie; het bewijs is dus tamelijk rechttoerechtaan, en daarmee wat saai. Laat $\Delta = \{(,)\}$.

$L(G) \subseteq D$. Bewijs met inductie naar de lengte k van de afleiding.
basis. $k = 1$: Als $S \Rightarrow w$ met $w \in \Delta^*$, dan moet de productie $S \to \lambda$ toegepast zijn; dus $w = \lambda \in D$.

inductiestap. Neem aan: als $S \Rightarrow^k w \in \Delta^*$ met $k \leq n$ dan $w \in D$. Stel $S \Rightarrow^{n+1} w \in \Delta^*$. Dan, afhankelijk van de eerste productie, ofwel $S \Rightarrow SS \Rightarrow^n w_1 w_2 = w$, ofwel $S \Rightarrow (S) \Rightarrow^n (w') = w$. De afleidingen $S \Rightarrow^* w_1$, $S \Rightarrow^* w_2$ resp. $S \Rightarrow^* w'$ zijn alle ten hoogste n stappen lang, en dus geldt w_1, w_2 resp. $w' \in D$. Met de definitie van D volgt dat $w = w_1 w_2 \in D$, resp. $w = (w') \in D$.

$D \subseteq L(G)$. Bewijs met inductie naar de definitie van D.
basis. (i) $w = \lambda$ wordt in G gegenereerd door de productie $S \to \lambda$ toe te passen.

inductiestap. Te bewijzen: (ii) als $w \in L(G)$ dan $(w) \in L(G)$, en
(iii) als $w_1, w_2 \in L(G)$, dan $w_1 w_2 \in L(G)$.

Dat is eenvoudig. Als $w \in L(G)$ dan bestaat een afleiding $S \Rightarrow^* w \in \Delta^*$, dus ook $S \Rightarrow (S) \Rightarrow^* (w)$. Evenzo, als $w_1, w_2 \in L(G)$, dan $S \Rightarrow^* w_1$, $S \Rightarrow^* w_2$ en daarmee $S \Rightarrow SS \Rightarrow^* w_1 S \Rightarrow^* w_1 w_2$.

b Modulo λ equivalente grammatica, zonder λ- en keten-regels:
$S \to SS, \quad S \to (S), \quad S \to ()$.

Chomsky normaalvorm:
$S \to SS, \quad S \to AX, \quad X \to SB, \quad S \to AB, \quad A \to (, \quad B \to)$.

Greibach normaalvorm:
We gebruiken de methode uit het bewijs van Stelling 4.39 en passen deze toe op de grammatica met producties $S \to SS$, $S \to (S)$, $S \to ()$. Deze grammatica heeft geen λ-producties, en slechts één niet-terminaal.

Verwijder linksrecursie voor S door nieuwe X te introduceren:
$S \to (S) \quad S \to (S)X \quad S \to () \quad S \to ()X$
$X \to S \quad X \to SX$

Substitueer in rechterzijden voor X die met een niet-terminaal beginnen:
$S \to (S) \quad S \to (S)X \quad S \to () \quad S \to ()X$

$X \to (S) \quad X \to (S)X \quad X \to () \quad X \to ()X$
$X \to (S)X \quad X \to (S)XX \quad X \to ()X \quad X \to ()XX$

Merk op dat hier twee producties voor X dubbel staan vermeld; ze ontstaan op twee verschillende manieren. Tenslotte voeren we de nieuwe niet-terminaal B in om) ter rechterzijde te representeren:

$S \to (SB \quad S \to (SBX \quad S \to (B \quad S \to (BX$
$X \to (SB \quad X \to (SBX \quad X \to (B \quad X \to (BX$
$X \to (SBXX \quad X \to (BXX$
$B \to)$

Alternatief. Het kan bewezen worden dat de taal D ook gegenereerd wordt door de cf grammatica met producties $S \to (S)S$ en $S \to \lambda$. Verwijdering van λ-producties en invoering van niet-terminaal B geeft de volgende cf grammatica in Greibach normaalvorm:
$S \to (SBS, \quad S \to (BS, \quad S \to (SB, \quad S \to (B, \quad B \to).$

4.12a In de oorspronkelijke opgave zijn de ketens al weggewerkt. Voer nu variabelen in om de terminalen te genereren:
$P \to \oplus, \quad T \to \otimes, \quad M \to \ominus, \quad D \to /, \quad A \to (, \quad B \to)$

Een passende grammatica ontstaat door tenslotte alle rechterkanten op juiste lengte te brengen:
$\langle\text{sum}\rangle \to \langle\text{sum}\rangle X, \quad X \to P\langle\text{term}\rangle, \quad X \to M\langle\text{term}\rangle, \quad \langle\text{sum}\rangle \to M\langle\text{term}\rangle$
$\langle\text{sum}\rangle \to \langle\text{term}\rangle Y, \quad \langle\text{term}\rangle \to \langle\text{term}\rangle Y, \quad Y \to T\langle\text{fact}\rangle, \quad Y \to D\langle\text{fact}\rangle$
$\langle\text{sum}\rangle \to AZ, \quad \langle\text{term}\rangle \to AZ, \quad \langle\text{fact}\rangle \to AZ, \quad Z \to \langle\text{sum}\rangle B$
$\langle\text{sum}\rangle \to a, \langle\text{sum}\rangle \to b, \quad \langle\text{term}\rangle \to a, \langle\text{term}\rangle \to b, \quad \langle\text{fact}\rangle \to a, \langle\text{fact}\rangle \to b.$

c Merk eerst op dat ieder woord in $L(G)$ evenveel linker- als rechterhaakjes heeft. Pas dan het pomplemma toe op het woord $z = (^n a)^n$.

4.14:2 We tonen aan dat $K_2 = \{\ ss \mid s \in \{a,b\}^*\ \}$ niet context-vrij is door te laten zien dat er geen constante voor K_2 bestaat zoals bedoeld in het pomplemma: voor elke n bestaat er een woord $z \in K_2$ met $|z| > n$ dat niet binnen de taal gepompt kan worden.

Kies n als in het pomplemma. Neem $z = a^n b^n a^n b^n$, dus $z \in K_2$ en $|z| > n$. Dan is $z = uvwxy$ volgens het pomplemma. z bestaat uit vier 'segmenten' ter lengte n. Omdat het 'pompstuk' vwx niet langer dan n mag zijn, overlapt dit met ten hoogste twee naast elkaar gelegen segmenten. We gaan de mogelijke gevallen na, en bekijken steeds $z_0 = uv^0wx^0y = uwy$.

i. eerste twee segmenten: $z_0 = a^k b^\ell a^n b^n$ met $n \le k + \ell < 2n$. $z_0 \notin K_2$, immers, als $ss = z_0$, dan is $|s| > n$ en moet (de tweede) s eindigen op ab^n, en dus zou $s = a^n b^n$ zijn.
ii. middelste twee segmenten: $z_0 = a^n b^k a^\ell b^n$ met $n \le k + \ell < 2n$. $z_0 \notin K_2$.
iii. laatste twee segmenten: symmetrisch met eerste.

Opmerking. Het is makkelijker om te bewijzen dat de taal $L = \{\ a^k b^m a^k b^m \mid k, m \geq 0\ \}$ niet context-vrij is. Daaruit concluderen we dan (via de afsluitingseigenschappen van CF) dat K_2 niet context-vrij is omdat $L = K_2 \cap a^* b^* a^* b^*$.

:4 Stel K_4 is context-vrij. Laat n de constante zijn voor K_4 als bedoeld in het pomplemma voor context-vrije talen.
Beschouw nu het woord $z = a^n b^n c^n \in K_4$. Dan is $|z| > n$ en dus zijn er volgens het pomplemma woorden u, v, w, x en y met $z = uvwxy$, $|vwx| \leq n$ en $vx \neq \lambda$, zodat $uv^i wx^i y \in K_4$ voor alle $i \geq 0$. Omdat $|vwx| \leq n$ is $vx \in \{a, b\}^* \cup \{b, c\}^*$.
Als $vx \in \{a, b\}^*$, dan is uwy van de vorm $a^{n-k} b^{n-l} c^n$ voor zekere k, l met $0 < k + l \leq n$. Maar dan hebben we een tegenspraak: $uwy \notin K_4$.
Als $vx \notin \{a, b\}^*$, dan is $uv^2 wx^2 y$ van de vorm $a^n b^{n+k} c^{n+l}$ voor zekere k, l met $0 < l \leq n$. Maar dan hebben we wederom een tegenspraak: $uv^2 wx^2 y \notin K_4$.
K_4 is dus *niet* context-vrij.

4.16:1 $L = \{a^m b^m c^m \mid m \in \mathbb{N}\} = \{a^i b^i c^k \mid i, k \in \mathbb{N}\} \cap \{a^j b^i c^i \mid j, i \in \mathbb{N}\}$. De twee talen zijn context-vrij. Laat dit zien door er een grammatica voor te geven.

:2 Een woord behoort niet tot L als (1) de letters niet in de juiste volgorde staan, of (2) wel op volgorde, maar niet in gelijke aantallen voorkomen. Dus $\{a, b, c\}^* - L = K \cup M$, met $K = \{a, b, c\}^* - a^* b^* c^*$, en $M = \{a^i b^j c^k \mid i \neq j \vee j \neq k\}$.

K is context-vrij want zelfs regulier. M is te schrijven als samenstelling van simpele talen, waarvan duidelijk in te zien is dat ze context-vrij zijn:
$M = \{a^i b^j c^k \mid i < j\} \cup \{a^i b^j c^k \mid i > j\} \cup \{a^i b^j c^k \mid j < k\} \cup \{a^i b^j c^k \mid j > k\}$
$\quad = \{a^n b^n \mid n \in \mathbb{N}\} b^+ c^* \cup a^+ \{a^n b^n \mid n \in \mathbb{N}\} c^* \cup a^* \{b^n c^n \mid n \in \mathbb{N}\} c^+ \cup a^* b^+ \{b^n c^n \mid n \in \mathbb{N}\}$.

Merk op dat de context-vrije talen gesloten zijn onder de gebruikte operaties vereniging en concatenatie.

4.18* Deze opgave, die vraagt om het bewijs van Stelling 4.52(5), laat zich intuitief eenvoudiger bewijzen mbv. automaten, zie volgende hoofdstuk. Voor de gemotiveerden (*koersen op kwaliteit*):

Laat K gegeven zijn door een cf grammatica $G = (\Sigma, \Delta, P, S)$ in Chomsky normaalvorm, en laat R gegeven zijn door een eindige automaat $\mathcal{A} = (Q, \Delta, \delta, I, F)$. (Aangenomen is alvast dat zowel K als R talen over Δ zijn.)

We construeren een nieuwe grammatica met niet-terminalen van de vorm $[pAq]$, $p, q \in Q$ en $A \in \Sigma - \Delta$. Voor $w \in \Delta^*$ moet gelden $[pAq] \Rightarrow^* w$ dan en slechts dan als $A \Rightarrow^* w$ én $q \in \delta(p, w)$. Met andere woorden: A leidt w af in de cf grammatica, en de automaat heeft een pad van p naar q gelabeld met w.

De nieuwe grammatica heeft axioma Z en de producties:

$Z \to [pSq]$ voor alle $p \in I$ en $q \in F$ (S is axioma van G),
$[pAq] \to [pBr][rCq]$ voor alle $A \to BC \in P$ en $p, q, r \in Q$,
$[pAq] \to a$ voor alle $A \to a \in P$ en $p, q \in Q$ met $(p, a, q) \in \delta$.

4.19a Er geldt dan $\Sigma = \Sigma_1 = \Sigma_2$, en $\phi_1(w) = \phi_2(w) = w$ voor alle $w \in \Sigma^*$. Vullen we dit in de definitie in dan zien we dat weave gelijk is aan de doorsnede: $K \omega L = \{w \in \Sigma^* \mid w \in K \text{ en } w \in L\} = K \cap L$.

b Gegeven automaten $\mathcal{A}_1 = (Q_1, \Sigma_1, \delta_1, I_1, F_1)$ voor $K \subseteq \Sigma_1^*$ en $\mathcal{A}_2 = (Q_2, \Sigma_2, \delta_2, I_2, F_2)$ voor $L \subseteq \Sigma_2^*$. Construeer een automaat $\mathcal{A}$ op een manier die erg lijkt op de productconstructie: de automaat simuleert $\mathcal{A}_1$ en $\mathcal{A}_2$ in parallel op letters die in beide alfabetten voorkomen. Voor letters die in slechts één van de alfabetten zitten wordt slechts de bijbehorende automaat nagedaan.

Neem voor het gemak aan dat $\mathcal{A}_1$ en $\mathcal{A}_2$ totaal deterministisch zijn. Definieer nu $\mathcal{A} = (Q, \Sigma, \delta, I, F)$ met $Q = Q_1 \times Q_2$, $I = I_1 \times I_2$, $F = F_1 \times F_2$, en, voor alle $(p_1, p_2) \in Q_1 \times Q_2$ en $a \in \Sigma$, $\delta((p_1, p_2), a) = (q_1, q_2)$ met

$q_1 = \delta_1(p_1, a)$ als $a \in \Sigma_1$, en $q_1 = p_1$ anders, en
$q_2 = \delta_2(p_2, a)$ als $a \in \Sigma_2$, en $q_2 = p_2$ anders.

Met inductie naar de lengte van w kan nu bewezen worden dat $\delta((p_1, p_2), w) = (\delta_1(p_1, \phi_1(w)), \delta_2(p_2, \phi_2(w)))$ en hieruit volgt eenvoudig dat $L(\mathcal{A}) = L(\mathcal{A}_1)\omega L(\mathcal{A}_2)$:

$w \in L(\mathcal{A}) \iff \delta((p_{01}, p_{02}), w) \in F \iff \delta_1(p_{01}, \phi_1(w)) \in F_1$ én $\delta_2(p_{02}, \phi_2(w)) \in F_2 \iff \phi_1(w) \in L(\mathcal{A}_1)$ én $\phi_2(w) \in L(\mathcal{A}_2) \iff w \in L(\mathcal{A}_1)\omega L(\mathcal{A}_2)$,
waarbij $I_1 = \{p_{01}\}$ en $I_2 = \{p_{02}\}$.

Opmerking. Aangezien REP gesloten is onder inverse homomorfismen (Stelling 3.24) en doorsnede (Stelling 3.20), volgt **b.** ook uit het feit dat $K\omega L = \phi_1^{-1}(K) \cap \phi_2^{-1}(L)$.

Als je inverse homomorfismen niet wilt gebruiken omdat die afsluitingseigenschap nergens netjes bewezen wordt, dan volstaat het te bewijzen dat $\phi_1^{-1}(K) \in \mathsf{REP}$. Construeer daartoe de automaat $\mathcal{A}_1' = (Q_1, \Sigma, \delta_1', I_1, F_1)$ met $\delta_1' = \delta_1 \cup \{(q, \sigma_2, q) \mid q \in Q_1, \sigma \in \Sigma_2\}$. Dan geldt $\delta_1'(p, w) = \delta_1(p, \phi_1(w))$ voor $p \in Q_1$, $w \in \Sigma^*$.

c Kies $\Sigma_1 = \{a, b\}$, $\Sigma_2 = \{b, c\}$. Laat $K = \{a^n b^n \mid n \geq 1\}$ en $L = \{b^n c^n \mid n \geq 1\}$. Dan $K\omega L = \{a^n b^n c^n \mid n \geq 1\}$, hoewel dat nog wel enige argumentatie vergt.

Opmerking. Dat CF niet gesloten is onder weave volgt ook direct uit **a** en het feit dat CF niet gesloten is onder doorsnede (zie Opgave 4.16).

4.20a $S \to XS \mid SX \mid Y$
$X \to aXb \mid \lambda$
$Y \to aYb \mid Xa$

b Laat $G = (\Sigma, \Delta, P, S)$. De nieuwe grammatica G_σ voor $\mathrm{inc}_\sigma(L(G))$ heeft als axioma $\overline{S}$. Verder heeft G_σ alle producties van G en bovendien voor elke productie $A \to BC$ van G de producties $\overline{A} \to \overline{B}C$ en $\overline{A} \to B\overline{C}$, en voor elke productie $A \to a$ de producties $\overline{A} \to \sigma a$ en $\overline{A} \to a\sigma$.

d $\mathrm{inc}_\sigma(K \cdot L) = (\mathrm{inc}_\sigma(K) \cdot L) \cup (K \cdot \mathrm{inc}_\sigma(L))$, $\mathrm{inc}_\sigma(K^*) = \{\sigma\} \cup K^* \cdot \mathrm{inc}_\sigma(K) \cdot K^*$.

4.21 (a) Kies een kopie van het alphabet, $\overline{\Delta} = \{\overline{a} \mid a \in \Delta\}$. Pas op $L \subseteq \Delta^*$ de eindige substitutie $\varphi(a) = \{a, \overline{a}\}$, $a \in \Delta$ toe.

(b) Doorsnij de verkregen taal met $\overline{\Delta}^* \cdot R$, zo ontstaat $\{\overline{x}y \mid xy \in L, y \in R\}$.

(c) Pas nu het homomorfisme h van $(\Delta \cup \overline{\Delta})^*$ naar Δ^* toe dat de letters uit Δ verwijdert en de streepjes afhaalt van de letters in $\overline{\Delta}$. Dit levert L/R.

4.22* Nee CF is niet gesloten onder quotiënt.

De taal met woorden van de vorm $(ab)^i c^i$, $i \geq 1$, is natuurlijk context-vrij. De niet context-vrije taal met woorden $a^i b^i c^i$, $i \geq 1$, is hieruit te krijgen door herhaald de verwisseling $ba \to ab$ toe te passen.

$$aaabbbccc \Leftarrow aababbccc \Leftarrow abaabbccc \Leftarrow abababccc$$

Dit principe gebruiken we om te laten zien dat de taal $\{a^i b^i c^i \mid i \geq 1\}$ te schrijven als quotiënt van twee context-vrije talen.

Laat L_1 de context-vrije taal zijn over $\{a, b, c, \$, \#\}$, met de woorden

$$u_0 \$ u_1 \# \cdots \# u_n \# u_{n+1} \# \operatorname{mir}(u_n) \# \cdots \# \operatorname{mir}(u_0),$$

waarbij $u_i \in \{a, b, c\}^*$ $(1 \leq i \leq n)$, $u_0 \in a^* b^* c^*$, en $u_{n+1} = (ab)^i c^i$ met $i \geq 1$.

Axioma S_a, producties: $S_a \to aS_a a$, $S_a \to S_b$, $S_b \to bS_b b$, $S_b \to S_c$, $S_c \to cS_c c$, $S_c \to \$Z\#$, $Z \to X$, $Z \to Y$, $X \to aXa$, $X \to bXb$, $X \to cXc$, $X \to \#Z\#$, $Y \to abYc$, $Y \to abc$.

Vervolgens is L_2 de context-vrije taal over $\{a, b, c, \$, \#\}$, met de woorden

$$\$ v_1 \# \cdots \# v_n \# v_{n+1} \# \operatorname{mir}(w_n) \# \cdots \# \operatorname{mir}(w_0),$$

waarbij $v_{i+1}, w_i \in \{a, b, c\}^*$ $(0 \leq i \leq n)$, en w_i ontstaat uit v_{i+1} door een verwisseling $ba \to ab$.

Producties: $S \to \$X$, $X \to aXa$, $X \to bXb$, $X \to cXc$, $X \to baYba$, $Y \to aYa$, $Y \to bYb$, $Y \to cYc$, $Y \to \#X\#$, $Y \to \#$.

Stel nu dat $w \in L_1/L_2$, dwz. $x_1 = wx_2$ met $x_1 \in L_1, x_2 \in L_2$, zeg met

$$x_1 = u_0 \$ u_1 \# \cdots \# u_n \# u_{n+1} \# \operatorname{mir}(u_n) \# \cdots \# \operatorname{mir}(u_0), \text{ en}$$

$$x_2 = \$ v_1 \# \cdots \# v_m \# v_{m+1} \# \operatorname{mir}(w_m) \# \cdots \# \operatorname{mir}(w_0).$$

Dan moet gelden (let op het unieke voorkomen van '\$'):

$w = u_0 = a^j b^k c^\ell$, $u_{n+1} = (ab)^i c^i$, $n = m$, en $u_i = v_i = w_i$ (voorzover ze bestaan),

en tenslotte: $w_i = u_i$ ontstaat uit $v_{i+1} = u_{i+1}$ door een verwisseling $ba \to ab$.

Hieruit volgt dat $j = k = l = i$ ofwel $w = a^i b^i c^i$.

Voor de liefhebber: dit principe kan worden uitgewerkt. Elke RE taal is te schrijven als het quotiënt van twee CF talen (zelfs 'lineaire' talen) door als hierboven afleidingen in type-0 grammatica's na te doen.

4.23 De gevraagde taal bestaat uit alle woorden van de vorm xyz met $x \in \Delta^*$, $y \in (\Delta \cup \Delta')^*$ en $z \in (\Delta')^*$, waarvoor geldt dat $xz \in L(G)$ en ofwel y eindigt op c_i maar begint niet met w_i, ofwel y eindigt op een symbool uit Δ.

Er zijn verschillende manieren waarop y niet met w_i kan beginnen. De meest voor de hand liggende is dat y met een van de beginletters van w_i afwijkt, maar het kan ook zo zijn dat het prefix wel overeenkomt, maar te weinig letters bevat.

Informeel komt deze schrijfwijze overeen met het volgende algoritme: lees herhaald aan de achterzijde een index, en verwijder deze samen met het overeenkomstige woord aan de voorzijde. Accepteer het woord als het gezochte woord aan de voorzijde niet gevonden kan worden, of als geen index aan de achterzijde staat.

a Het axioma van de nieuwe grammatica is S.
De niet-terminaal T genereert $\{a, b, 1, 2, 3\}^*$ met de producties
$T \to aT \mid bT \mid 1T \mid 2T \mid 3T \mid \lambda$.

De niet-terminaal C genereert indices:
$C \to 1 \mid 2 \mid 3$

De producties voor S zijn:

$S \to aabS1 \mid aS2 \mid abaS3$ (genereer xSz)

$S \to bT1 \mid abT1 \mid aaaT1$ (foute letter)
$S \to CT1 \mid aCT1 \mid aaCT1$ (cijfer ipv. letter)
$S \to 1 \mid a1 \mid aa1$ (te kort)

$S \to bT2$
$S \to CT2$
$S \to 2$

$S \to bT3 \mid aaT3 \mid abbT3$
$S \to CT3 \mid aCT3 \mid abCT3$
$S \to 3 \mid a3 \mid ab3$

$S \to Ta \mid Tb$ (y eindigt met een letter)

b Laat $\Sigma = \Delta \cup \Delta'$. Een cf grammatica die het complement van $L(G)$ genereert heeft startsymbool S met producties $S \to w_i S c_i$ voor $1 \le i \le n$, en de productie $S \to F$ waarbij F het startsymbool is van een rechts-lineaire grammatica die de reguliere (!) taal $\Sigma^*\Delta \cup \bigcup_{i=1}^{n}(\Sigma^* - w_i\Sigma^*)c_i$ genereert.

Hierboven is geen rechts-lineaire grammatica geconstrueerd, maar gebruik gemaakt van het feit dat, voor $w = a_1 \dots a_m$, $\Sigma^* - w\Sigma^* = \bigcup_{i=1}^{m} a_1 \dots a_{i-1}(\{\lambda\} \cup (\Sigma - \{a_i\})\Sigma^*)$, uitgesplitst in 'te kort' $= \{\lambda\}$ en 'foute letter' of 'cijfer ipv. letter' $= \Sigma - \{a_i\}$.

c Als $\Sigma^* - L \in \mathsf{CF}$ dan ook $\Sigma^* - (w_1 L c_1) \in \mathsf{CF}$, want die laatste taal is gelijk aan $(\Sigma^* - w_1\Sigma^*) \cup (\Sigma^* - \Sigma^* c_1) \cup w_1(\Sigma^* - L)c_1$.

4.24a Analoog aan Opgave 3.39. Er geldt: $L(G)$ bevat een woord van de vorm $(ab)^n$, $n \ge 1$, desdals $L(G) \cap (ab)^+ \ne \emptyset$. Construeer, om dit te testen, eerst een eindige automaat $\mathcal{A}$ voor de taal $(ab)^+$, en dan een cf grammatica G_d voor de doorsnede $L(G) \cap L(\mathcal{A})$ met de constructie uit het bewijs van Stelling 4.52(5) – zie Opgave 4.18. Dus $L(G_d) = L(G) \cap (ab)^+$. Test nu of $L(G_d)$ niet leeg is, met Stelling 4.56.

b Test of $L(G_d)$ oneindig is, met Stelling 4.56.