

Huiswerkgave 3 Automata Theory, najaar 2021

Gepubliceerd, woensdag 17 november 2021.

Uiterste inleverdatum: woensdag 1 december 2021, 23.59 uur.

De opgave moet individueel gemaakt worden. Antwoorden in te leveren via Brightspace. Lever één bestand (eventueel een zip) in. Vermeld in je inzending ook je naam en studentnummer. Je mag je antwoorden zowel intypen als handmatig schrijven. Lever in het laatste geval een goed leesbare scan / goed leesbare foto's in.

1. Laat

$$L = \{x \in \{a, b\}^* \mid x = a^i(ba)^j a^k \text{ met } i, j, k \geq 0 \\ \text{en } x \text{ bevat evenveel voorkomens van de substring } ba \text{ als van de substring } aa\}$$

Een string $x \in L$ moet dus aan alle vermelde eisen voldoen. Merk op dat voorkomens kunnen overlappen. Bijvoorbeeld $x = aaa$ bevat twee voorkomens van de substring aa .

- Geef de eerste zeven elementen in de canonieke volgorde van L .
- Geef een context-vrije grammatica G , zó dat $L(G) = L$.

Leg ook uit wat de functie is van de diverse variabelen en producties in G voor het genereren van L .

2. Opgave 4.38d in het boek gaat over de context-vrije grammatica G met als enige variabele (en dus ook startvariabele) S en de volgende producties:

$$S \rightarrow aSb \mid abS \mid \Lambda$$

Deze grammatica is dubbelzinnig, want bijvoorbeeld de string $ab \in L(G)$ heeft twee verschillende afleidingsbomen.

- Het is eenvoudig in te zien (en met inductie te bewijzen) dat voor elke string $x \in L(G)$ geldt dat
 - $n_a(x) = n_b(x)$
 - en dat voor elke prefix y van x geldt dat $n_a(y) \geq n_b(y)$ (ofwel: de a komt altijd vóór de corresponderende b).

Het is echter niet zo dat elke string $x \in \{a, b\}^*$ die aan die twee eisen voldoet, in $L(G)$ zit.

Geef een voorbeeld van een string $x \in \{a, b\}^*$, die aan de twee eisen voldoet, maar niet in $L(G)$ zit. Beredeneer ook *waarom* x niet in $L(G)$ zit.

- Bij het werkcollege hebben we ook een alternatieve grammatica G' gezien, met $L(G') = L(G)$. Deze grammatica G' heeft startvariabele S' en de volgende producties:

$$S' \rightarrow \Lambda \mid S \quad S \rightarrow aSb \mid abS \mid ab$$

Toon aan dat deze grammatica G' ondubbelzinnig is.

Toon daartoe eerst aan dat er voor elke string $x \in \{a, b\}^*$ die vanuit variabele S is af te leiden, maar één afleiding vanuit S bestaat (en dus ook maar één linkspreferente afleiding). Doe dit met behulp van inductie naar de lengte van een string $x \in \{a, b\}^*$ die af te leiden is vanuit S . Je mag gebruiken (zonder bewijs) dat x aan de twee bij (a) vermelde eisen voldoet.

Toon vervolgens aan dat er voor elke string $x \in L(G')$ maar één afleiding vanuit startvariabele S' bestaat (en dus ook maar één linkspreferente afleiding).

3. Beschouw de context-vrije grammatica G met startvariabele S en de volgende producties:

$$S \rightarrow \Lambda \mid T \mid aTb \quad T \rightarrow abS \mid a$$

Zet deze grammatica stap voor stap om in een context-vrije grammatica G' in Chomsky normaalvorm, zó dat $L(G') = L(G) - \{\Lambda\}$. Dat wil zeggen:

- Bepaal de nullable variabelen in G .
- Geef de context-vrije grammatica G_1 die uit G ontstaat door de Λ -producties te elimineren.
- Bepaal voor elke variabele X de X -derivable variabelen.
- Geef de context-vrije grammatica G_2 die uit G_1 ontstaat door unitproducties te elimineren.
- Geef de context-vrije grammatica G_3 die uit G_2 ontstaat door voorkomens van terminalen σ in de rechterkant van producties te beperken tot producties van de vorm $X \rightarrow \sigma$.
- Geef de context-vrije grammatica G' die uit G_3 ontstaat door te lange rechterkanten van producties op te splitsen.