Voorbereiding Gequoteerde Oefenzitting 1

Academiejaar 2024-2025 Vincent Van Schependom

$$\text{RegLan} \subset \text{DCFL} \subset \text{niet-ambigue CFL} \subset \text{CFL} \subset \mathcal{P}(\Sigma^*)$$

Algemeenheden

- Een oneindige taal is aftelbaar, want

$$L \in L_\Sigma \Leftrightarrow L \in \mathcal{P}(\Sigma^*) \Leftrightarrow L \subseteq \Sigma^* \quad \text{en} \quad \Sigma^* \text{ is aftelbaar oneindig}$$

- Twee verschillende reguliere expressies kunnen dezelfde reguliere taal bepalen
  - $R_1 = a^*$
  - $R_2 = a^*|\varepsilon$
  - $L_{R_1} = L_{R_2} = \{a^n \mid n \in \mathbb{N}\}$
- Equivalentie van NFA's
  - Indien deze dezelfde taal bepalen
  - Bepaalt een equivalentierelatie
  - Elke equivalentieklasse komt overeen met één (reguliere) taal
- Een DFA is een NFA dus de taal bepaald door een DFA is ook regulier. Maar wordt – omgekeerd – elke reguliere taal ook bepaald door een DFA? We weten dat elke reguliere taal bepaald wordt door een NFA, dus het volstaat om aan te tonen dat NFA's om te zetten zijn in equivalente DFA's om te bewijzen dat elke reguliere taal bepaald wordt door een DFA. Dit bewijst dan (samen met het eerste) ineens dat DFA's en NFA's equivalent zijn: ze bepalen dezelfde talen, namelijk de reguliere talen.
- We kunnen makkelijk inzien dat er voor een reguliere taal een DFA bestaat met het minimale aantal toestanden (rangschik de DFA's naar $|Q|$), maar de constructie van de minimale DFA (aan de hand van de equivalentierelatie f-gelijk) is de echte vraag.
- De verzameling niet-reguliere talen is overaftelbaar, want RegLan is aftelbaar, $\text{RegLan} \cup \overline{\text{RegLan}}$ is precies gelijk aan $\mathcal{P}(\Sigma^*)$ en die laatste verzameling is overaftelbaar.

Alle talen

- Gesloten onder uiteenlopende operaties:
  unie, doorsnede, verschil, complement, omkering, concatenatie, Kleene ster, ...
- Oneindige talen zijn altijd aftelbaar, want elke taal is een deelverzameling van Σ^* en dat is een aftelbaar oneindige verzameling.

Reguliere talen

- *Gesloten onder / algebra voor* de operaties...
  - Unie
  - Concatenatie
  - Kleene ster
  - Complement
- (Verrassende) voorbeelden van reguliere talen:
  - Σ^*
    kan beschreven worden met de RegExp $(a_1|a_2|...|a_n)^*$ met $a_i \in \Sigma, n = |\Sigma|, a_i \neq a_j$ voor $i \neq j$
  - Eindige talen
    kunnen beschreven worden met een reguliere expressie die alle symbolen uit het alfabet oplijst met daartussen een |
  - De lege taal
    Bouw een simpele DFA met 1 niet-accepterende toestand met 1 lus naar zichzelf met daarop elk symbool uit het alfabet.
- RegLan $\subset$ CFL, want
  - RegLan $\subseteq$ CFL: elke NFA is een speciaal type PDA (namelijk eentje zonder stack)
  - $\exists$CFL $L : L \notin$ RegLan, bijvoorbeeld $L = \{0^n1^n \mid n \in \mathbb{N}\}$
- Er zijn oneinig veel reguliere talen
  - Er zijn oneindig veel strings over $\Sigma \Rightarrow \Sigma^*$ is een oneindige verzameling
  - Elke eindige deelverzameling van Σ^* is regulier
  - $\uparrow$ want we kunnen zo'n deelverzameling = taal $L = \{s_1, ..., s_n\}$ beschrijven met een reguliere expressie $(s_1|...|s_n)$.
- $L \notin$ RegLan $\Rightarrow \overline{L} \notin$ RegLan

Contextvrije talen

- Elke reguliere taal is contextvrij.
  - Een FSA is een PDA waarbij de stapel niet meespeelt.
  - Alternatieve manier om dit in te zien: stel een CFG op voor de reguliere taal
    * Op basis van de reguliere expressie:
      · Basisgevallen:

$$\begin{aligned} a &\rightarrow \{a\} \\ \varepsilon &\rightarrow \{\varepsilon\} \\ \phi &\rightarrow \emptyset \end{aligned}$$

      · Recursieve gevallen:

$$\begin{aligned} S &\rightarrow S_1S_2 && \text{voor de concatenatie } E_1E_2 \\ S &\rightarrow S_1|S_2 && \text{voor de 'keuze' } E_1|E_2 \\ S &\rightarrow S_1S|\varepsilon && \text{voor de ster } E_1^* \end{aligned}$$

    * Op basis van de DFA: de vorm van de grammatica is dan als volgt

$$\begin{aligned} A &\rightarrow aB && a \in \Sigma, B \in V \\ B &\rightarrow a \end{aligned}$$

    In de eerste regel wordt dan de overgang $\delta(A, a) = B$ gesimuleerd, waarbij $A, B \in Q$
- Er bestaan geen eindige talen die niet contextvrij zijn:
  - Elke eindige taal is regulier
  - Elke reguliere taal is contextvrij
  - Dus elke eindige taal is contextvrij!
- Gesloten onder
  - Unie ($S \rightarrow S_1 \mid S_2$), maar DCFL niet!
  - Concatenatie ($S \rightarrow S_1S_2$)
  - Kleene ster ($S \rightarrow SS_1 \mid \varepsilon$)
- Niet gesloten onder
  - Complement, maar DCFL wel! Het complement van een DCFL is opnieuw DCFL
  - Doorsnede, verschil, ...
  - Truk van productautomaat werkt niet door stacks
  - Je kan wel productautomaat maken tussen PDA en DFA bvb.
- Ambiguïteit:
  - Deterministische talen zijn niet-ambigu
    want een ambigue taal kan onmogelijk deterministisch zijn
  - Er bestaan niet-ambigue talen die niet deterministisch zijn
    * $L = \{s\hat{s} \mid s \in \{a, b\}^*\} = \{\text{palindromen}\}$
    * Niet ambigue grammatica: $S \rightarrow aSa \mid bSb \mid \varepsilon$
    * Maar PDA weet van tevoren niet wat midden is en moet ernaar raden $\rightarrow$ ND

- Hoe niet-deterministische taal vinden?
  - Neem de unie van twee CFL's die overlappen.
    * Stel dat deze unie DCFL is
    * Dan is ook het complement DCFL
    * Bekom een contradictie en besluit dat de unie niet deterministisch kan zijn
  - Neem een $L \notin$ CFL waarvoor $\bar{L} \in$ CFL.
    * Als $\bar{L} \in$ DCFL zou gelden, dan geldt ook $\bar{\bar{L}} = L \in$ DCFL
    * Maar L is niet eens CFL!
    * Dus is $\bar{L} \notin$ DCFL
- De doorsnede van een reguliere taal met een contextvrije taal is contextvrij. Maak daarvoor een generische product-PDA aan de hand van de DFA die de reguliere taal bepaalt en de PDA die de contextvrije taal bepaalt.
- De doorsnede van een DCFL en een reguliere taal is een DCFL: maak een product-DPDA tussen een DPDA en DFA.
- $\{a^n b^n c^n \mid n \in \mathbb{N}\} \notin$ CFL

Het symmetrisch verschil van twee talen L_1 en L_2 is $L_1 \triangle_2$.
De DFA die deze taal bepaalt heeft als verzameling eindtoestanden

$$F = (F_1 \times Q_2) \cup (Q_1 \times F_2) \setminus (F_1 \times F_2)$$

Pompend Lemma Reguliere Talen

We bewijzen dit aan de hand van het pompend lemma voor reguliere talen.
Stel dat er een pomplengte d bestaat zodanig dat elke string $s \in L$ met lengte $|s| \geq d$ kan opgedeeld worden in 3 stukken $x, y, z \in \Sigma^*$ zodanig dat $s = xyz$ en zodat

1. $\forall i \geq 0 : xy^i z \in L$
2. $|y| > 0$
3. $|xy| \leq d$

Beschouw zo'n string $s = ...$ met $s \geq d$. Beschouw een willekeurige opdeling $s = xyz$ met $|y| > 0$ en $|xy| \leq d$. Dan ...
Indien men te string $t = ...$ (kies een i zodat $t = xy^i z$), dan behoort t niet tot de taal en de taal is dus niet regulier.

Pompend Lemma Contextvrije Talen

We bewijzen dit aan de hand van het pompend lemma voor contextvrije talen.
Stel dat er een pomplengte p bestaat zodanig dat elke string $s \in L$ met lengte $|s| > p$ kan opgedeeld worden in 5 stukken $u, v, x, y, z \in \Sigma^*$ zodanig dat $s = uvxyz$ en zodat

1. $\forall i \geq 0 : uv^i xy^i z \in L$
2. $|vy| > 0$
3. $|vxy| \leq p$

Neem zo'n string s die (strikt) langer is dan p, namelijk $s = ...$ Beschouw een willekeurige opdeling $s = uvxyz$ met $|vy| > 0$ en $|vxy| \leq p$. Dan ...
Indien men te string $t = ...$ (kies een i zodat $t = uv^i xy^i z$), dan behoort t niet tot de taal en de taal is dus niet contextvrij.