14.03.2022

Uwaga Klasa języków regularnych nad Σ jest najmniejszą klasą języków nad Σ, która:
- zawiera wszystkie języki skończone (i) $(*)$
- jest zamknięta na sumę, konkatenację (ii) i gwiazdkę Kleene'go (iii) $(L \rightsquigarrow L^*)$

Istnienie: proste, Najmniejszość: pokazać, że dowolna klasa języków spełniająca powyższe warunki zawiera języki regularne.

GRAMATYKI BEZKONTEKSTOWE

Def. **Gramatyka bezkontekstowa** (CFG) to krotka $\langle N, \Sigma, S, \Pi \rangle$, $S \in N$,

$\Pi \underset{\text{skończone}}{\subseteq} N \times (N \cup \Sigma)^*$, $N \cap \Sigma = \emptyset$

CFG – Context Free Grammar

Dygresja $\mathcal{A}$: alfabet, $\Pi \subseteq \mathcal{A}^* \times \mathcal{A}^*$ skończone

Dla $w, v \in \mathcal{A}^*$ definiujemy $w \longrightarrow_\Pi v$ gdy istnieją słowa $x, y \in \mathcal{A}^*$ i para $\langle l, r \rangle \in \Pi$ t. że $w = xly$ oraz $v = xry$

Po ludzku: bierzemy w, znajdujemy w nim infiks l, zamieniamy go na r i dostajemy v.

Relacje $\xrightarrow{*}_\Pi$ i $\xleftrightarrow{*}_\Pi$ definiujemy odpowiednio jako tranzytywne i równoważnościowe domknięcie relacji $\longrightarrow_\Pi$.

- $\xrightarrow{*}_\Pi$ osiągalność w grafie
- $\xleftrightarrow{*}_\Pi$ bycie w jednej spójnej składowej (osiągalność w grafie bez skierowania)

KONIEC DYGRESJI

Dla danej CFG $G = \langle N, \Sigma, S, \Pi \rangle$ przez $\overline{L}_G$ oznaczamy $\{ w \in (N \cup \Sigma)^* : S \xrightarrow{*}_{\Pi} w \}$, przez $L_G = \overline{L}_G \cap \Sigma^*$

<u>Def.</u> $L \subseteq \Sigma^*$ jest **bezkontekstowy** (CFL), gdy istnieje CFG G t.że $L = L_G$

<u>Obserwacja</u> Każdy język regularny jest bezkontekstowy.

<u>Dowód</u> Klasa CFL spełnia warunki z uwagi (*), (i), (ii), (iii):

(*): proste, dodajemy reguły $S \to w$ dla w z języka

(i): bierzemy sumę rozłączną dwóch gramatyk i dodajemy reguły $\langle S, S' \rangle$, $\langle S, S'' \rangle$

nowy S

(ii): $G' = \langle N', \Sigma, S', \Pi' \rangle$, $G'' = \langle N'', \Sigma, S'', \Pi'' \rangle$,

konstruujemy $G = \langle N' \sqcup N'' \sqcup \{S\}, \Sigma, S,$
$\Pi' \sqcup \Pi'' \sqcup \langle S, S'S'' \rangle \rangle$

(iii) Podobnie gwiazdka.

<u>Przykład</u> • Notacja: $S \to aSa \mid bSb \mid \varepsilon$
oznacza, że $\Pi = \{ \langle S, aSa \rangle, \langle S, bSb \rangle, \langle S, \varepsilon \rangle \}$
(to konkretnie daje język palindromów parzystej długości).

• $S \to SS \mid \varepsilon \mid aSb \mid bSa$

$\leadsto$ język słów, które mają tyle samo liter a co b.

• $S \to (S) \mid [S] \mid SS \mid \varepsilon$

poprawne nawiasowanie z (,), [,].

Konwencja notacyjna (niemoralna Marcinkowskiego):

Dla języków L, L' piszemy $L = L'$ gdy $L \dot{-} L' \subseteq \{\varepsilon\}$

różnica symetryczna

Def CFG $G = \langle N, \Sigma, S, \Pi \rangle$ jest postaci **normalnej Chomskiego**, gdy każda produkcja z Π jest postaci $A \longrightarrow BC$ dla $A, B, C \in N$ lub $A \longrightarrow a$ dla $A \in N$, $a \in \Sigma$.

TW. (Chomskiego o postaci normalnej)
Dla każdego CFL L istnieje CFG G w postaci Chomsky'ego t. że $L = L_G$.

Dowód: Nudny :(

Lemat (o pompowaniu dla CFG)
Dla każdego CFL L istnieje $n \in \mathbb{N}$ t. że dla każdego $w \in L$, $|w| \geq n$, istnieją słowa

s, z, t, y, x t. że $|zty| \le n$, $|zy| > 0$, $w = sztyx$ dla każdego $k \in \mathbb{N}$ $sz^k t y^k x \in L$.

<u>Uwaga</u> Czy język $\{a^i b^i c^j : i, j \in \mathbb{N}\}$ jest CFG?

Odp.: TAK (proste)

A co z $\{a^i b^j c^j : i, j \in \mathbb{N}\}$?

Odp.: TAK (to samo)

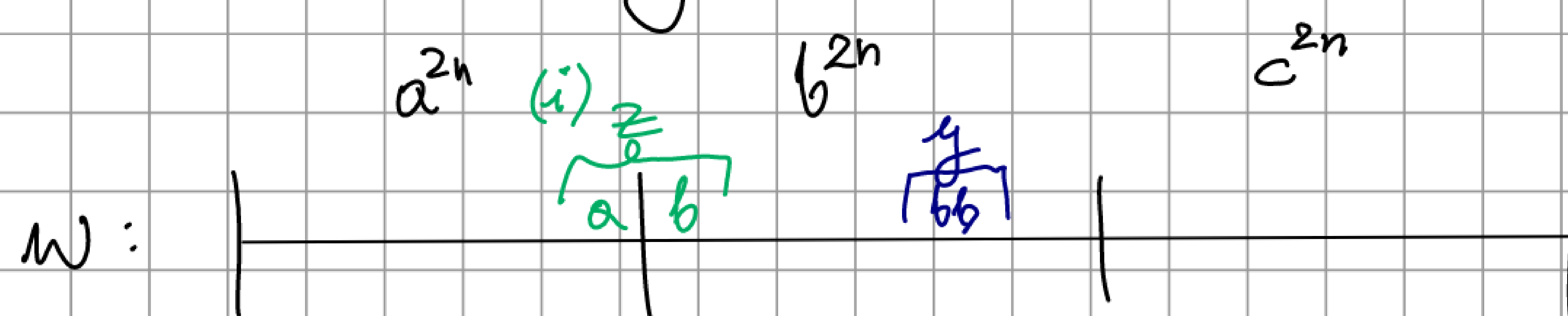
A czym jest przekrój tych języków?

Odp.: $\{a^i b^i c^i : i \in \mathbb{N}\}$

Ten język jednak nie jest CFL!

<u>Dowód</u> Załóżmy, że L jest CF. Niech n: stała z lematu o pompowaniu.

Niech $w = a^{2n} b^{2n} c^{2n}$. Wtedy są stałe s, z, t, y, x z lematu.



Przypadki: (i) z lub y zawiera dwie różne literki, to dla $k=2$ wygrywamy

(ii) jeżeli x, y mają tylko po jednej literce, to $k=0$ wygrywamy bo tej trzeciej literki jest więcej.