9. Übung (July 4, 2016)

Formale Baumsprachen

Thirteenth MT Marathon 2018. Organized by the Institute of Formal and Applied Linguistics, Charles University, Machine Translation Marathon 2018 is a week long gathering of machine translation researchers, developers, students and users. It features:

- MT Lectures and Labs covering the basics and tutorials.
- Keynote Talks from experienced researchers and practitioners.
- Presentations of open source tools related to MT.
- Hacking Projects to advance tools or research in one week or start new collaborations.

Task 20 (construction for Rec $\subseteq$ Rat*)*

Consider the ranked alphabet $\Sigma = \{\alpha^{(0)}, \gamma^{(1)}\}$.

(a) Give sets N and P such that the regular tree grammar $G = (N, \Sigma, Z, P)$ is in normal form and recognizes

$$L = \{\xi \in T_\Sigma \mid \text{the number of occurrences of } \gamma \text{ in } \xi \text{ is } \textit{not} \text{ divisible by } 3\}.$$

(b) Convince yourself that $L^N_{Z,\emptyset} = L$ using the following definition and property:

Definition. For every $Q, K \subseteq N$ such that $Q \cap K = \emptyset$, and for every $A \in N$:

$$L^Q_{A,K} = \{\xi \in T_\Sigma(K) \mid \text{there is a derivation } A \Rightarrow_G \xi_1 \Rightarrow_G \ldots \Rightarrow_G \xi_n \Rightarrow_G \xi_{n+1} = \xi \text{ with } n \geq 0 \text{ such that for every } i \in [n]\colon \xi_i \in T_\Sigma(Q \cup K) \text{ and a rule with left-hand side in } Q \text{ is applied to } \xi_i \text{ to obtain } \xi_{i+1}\}$$

Property. For every $Q, K \subseteq N$ and $A, B \in N$ such that $B \in N \setminus Q$ and $(Q \cup \{B\}) \cap K = \emptyset$:

$$L^{Q \cup \{B\}}_{A,K} = L^Q_{A,K \cup \{B\}} \cdot_B (L^Q_{B,K \cup \{B\}})^*_B \cdot_B L^Q_{B,K}$$

Task 21 (local tree languages [Com+08, Exercise 2.5])

Definition. Let Σ be a ranked alphabet. For every $\xi \in T_\Sigma$, the *fork of* ξ is the set

$$\text{fork}(\xi) = \{(\sigma, \sigma_1 \cdots \sigma_k) \in \Sigma \times \Sigma^* \mid \rho \in \text{pos}(\xi), \xi(\rho) = \sigma, k = \text{rank}(\sigma), \forall i \in [k]\colon \xi(\rho i) = \sigma_i\}.$$

A tree language $L \subseteq T_\Sigma$ is called *local* if there are sets $F \subseteq \Sigma$ and $G \subseteq \text{fork}(T_\Sigma)$ such that $\xi \in L$ iff $(\xi(\varepsilon) \in F) \wedge (\text{fork}(\xi) \subseteq G)$. □

(a) Prove that every local tree language is regular.

(b) Prove that a language is local iff it is the set of parse trees of a context-free string grammar.

Task 22 (path languages [Com+08, Exercise 2.8])

Definition. Let Σ be a ranked alphabet. For every $\xi = \sigma(\xi_1, ..., \xi_k) \in T_\Sigma$, the *set of paths of* ξ is recursively defined by

$$\begin{aligned} \mathrm{Paths}(\sigma) &= \{\sigma\}, && \text{and} \\ \mathrm{Paths}(\sigma(\xi_1, ..., \xi_k)) &= \{\sigma\} \cdot \bigcup_{i \in [k]} \mathrm{Paths}(\xi_k) && \text{if } k > 0. \end{aligned}$$

(a) Prove that $\mathrm{Paths}(L) = \bigcup_{\xi \in L} \mathrm{Paths}(\xi)$ is regular for every regular tree language L.

(b) What about the converse?

References

[Com+08] Hubert Comon, Max Dauchet, Rémi Gilleron, Christof Löding, Florent Jacquemard, Denis Lugiez, Sophie Tison, and Marc Tommasi. *Tree Automata Techniques and Applications*. Nov. 18, 2008. url: `http://tata.gforge.inria.fr/`.