Language Operations and a Structure Theory of ω-Languages

July 30, 2012

Introduction: $\mathcal{P}(\Sigma^*) \to \mathcal{P}(\Sigma^*)$

We have the common $\mathcal{P}(\Sigma^*) \to \mathcal{P}(\Sigma^*)$ language operators:

1. $\mathrm{ext}(L) := \{\alpha \in \Sigma^\omega \mid \exists n\colon \alpha[0,n] \in L\} = L \cdot \Sigma^\omega$
2. $\widehat{\mathrm{ext}}(L) := \{\alpha \in \Sigma^\omega \mid \forall n\colon \alpha[0,n] \in L\}$
3. $\lim(L) := \{\alpha \in \Sigma^\omega \mid \forall N\colon \exists n > N\colon \alpha[0,n] \in L\} =$
   $\{\alpha \in \Sigma^\omega \mid \exists^\omega n\colon \alpha[0,n] \in L\}$
4. $\widehat{\lim}(L) := \{\alpha \in \Sigma^\omega \mid \exists N\colon \forall n > N\colon \alpha[0,n] \in L\}$

Introduction: $\mathcal{P}(\mathcal{P}(\Sigma^*)) \to \mathcal{P}(\mathcal{P}(\Sigma^*))$

From these, define language class operators:

1. $\mathrm{ext}(\mathcal{L}) := \{\lim L \mid L \in \mathcal{L}\}$
2. $\widehat{\mathrm{ext}}(\mathcal{L}) := \left\{\widehat{\mathrm{ext}}\, L \,\middle|\, L \in \mathcal{L}\right\}$
3. $\lim(\mathcal{L}) := \{\lim L \mid L \in \mathcal{L}\}$
4. $\widehat{\lim}(\mathcal{L}) := \left\{\widehat{\lim}\, L \,\middle|\, L \in \mathcal{L}\right\}$

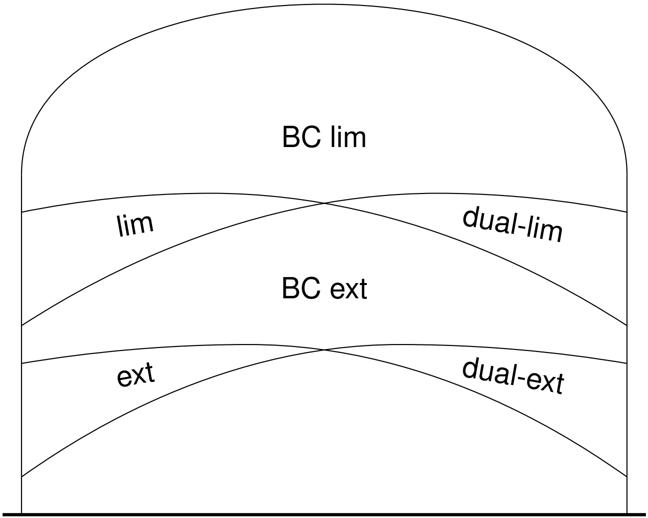
We combine these operators via union or intersection, e.g.

$$\mathrm{ext} \cup \widehat{\mathrm{ext}}\, \mathcal{L} := \mathrm{ext}\, \mathcal{L} \cup \widehat{\mathrm{ext}}\, \mathcal{L}.$$

Or boolean combinations:

1. $\mathrm{BC}\,\mathrm{ext}\,\mathcal{L} = \mathrm{BC}(\mathrm{ext}(\mathcal{L}))$
2. $\mathrm{BC}\lim\mathcal{L} = \mathrm{BC}(\lim(\mathcal{L}))$

$\mathcal{L}^*(\mathrm{reg})$ inclusion diagram



Questions

- instead of the class of regular $*$-languages, look at other $*$-language classes, e.g. starfree, LT, PT, or any arbitrary $*$-language class $\mathcal{L}$
- does it result in the same relations as in the diagram? are the enclosures strict?

My Diplom thesis:

- Chapter 3: general results on arbitrary $\mathcal{L}$, given some introduced properties on $\mathcal{L}$
- Chapter 4: concrete $*$-language classes

Properties on $\mathcal{L}$

Let $L, A, B \in \mathcal{L}$.

1. **Closure under suffix-independence**: $L \cdot \Sigma^* \in \mathcal{L}$
2. **Closure under union, intersection**: $A \cup B \in \mathcal{L}$, $A \cap B \in \mathcal{L}$
3. **Closure under negation**: $-L \in \mathcal{L}$
4. **Closure under change of final states**: Let $\mathcal{A}_L = (Q, \Sigma, q_0, \delta, F_L)$ be the minimal deterministic automaton for L, i.e. with $L^*(\mathcal{A}_L) = L$. Then, for all $F' \subseteq Q$, we have $L^*((Q, \Sigma, q_0, \delta, F')) \in \mathcal{L}$.
5. **Closure under alphabet permutation**: For all permutations $\sigma : \Sigma \to \Sigma$, we have $L_\sigma := \{\sigma(w) \mid w \in L\} \in \mathcal{L}$

General results

- Lemma 3.3: Closure under suffix-independence $\Rightarrow$ $\mathrm{ext}\,\mathcal{L} \subseteq \lim \cap \widehat{\lim}\,\mathcal{L}$ (but $\not\Leftarrow$)
- Lemma 3.8: Closure under suffix-independence and negation $\Rightarrow$
$\mathrm{ext} \cup \widehat{\mathrm{ext}}\,\mathcal{L} \subseteq \lim \cap \widehat{\lim}\,\mathcal{L}$

Separating language for $\mathrm{ext} \cup \widehat{\mathrm{ext}} \subsetneqq \mathrm{BC}\,\mathrm{ext}$, $\lim \cup \widehat{\lim} \subsetneqq \mathrm{BC}\lim$:
$\Sigma := \{a, b, c\}$, $L_a := \Sigma^* a$, $L_b := \Sigma^* b$.
$\tilde{L}_1 := \mathrm{ext}\, L_a \cap -\mathrm{ext}\, L_b$, $\tilde{L}_2 := \lim L_a \cap -\lim L_b$.
$\tilde{L}_1 \notin \mathrm{ext} \cup \widehat{\mathrm{ext}}\,\mathcal{L}$ but $\tilde{L}_1 \in \mathrm{BC}\,\mathrm{ext}\,\mathcal{L}$.
$\tilde{L}_2 \notin \lim \cup \widehat{\lim}\,\mathcal{L}$ but $\tilde{L}_2 \in \mathrm{BC}\lim\mathcal{L}$.
More general:

General results

Definition 3.12. A language $L \subseteq \Sigma^* \cup \Sigma^\omega$ is called *M*-**invariant** for $M \subseteq \Sigma$ iff for all $w \in \Sigma^* \cup \Sigma^\omega$,

$$w \in L \quad \Leftrightarrow \quad w|_M \in L,$$

where $w|_M$ is the word w with all letters from M removed. There is always exactly one **maximum invariant alphabet set** $M_m \subseteq \Sigma$ **of** L such that L is M_m-invariant. Then call $\Sigma - M_m$ the **non-invariant alphabet set of** L.

Theorem 3.15. Let $\mathcal{L}$ be closed under negation and under alphabet permutation. Let $\{a, b, c\} \subseteq \Sigma$. Let there be $L_a \in \mathcal{L}$. Let $\{a\}$ be the *non-invariant alphabet set of* L_a and let L_a be $\{b, c\}$*-invariant*. Then

$$\mathrm{ext}\, L_a \notin \widehat{\mathrm{ext}}\, \mathcal{L}^*(\mathrm{reg}) \quad \Rightarrow \quad \mathrm{ext} \cup \widehat{\mathrm{ext}}\, \mathcal{L} \subsetneq \mathrm{BC}\, \mathrm{ext}\, \mathcal{L}$$

and

$$\lim L_a \notin \widehat{\lim}\, \mathcal{L}^*(\mathrm{reg}) \quad \Rightarrow \quad \lim \cup \widehat{\lim}\, \mathcal{L} \subsetneq \mathrm{BC} \lim \mathcal{L}.$$

General results

- **Theorem 3.19.** (Staiger-Wagner 1) $\mathcal{L}$ closed under change of final states. Then
$$\lim \cap \widehat{\lim}\, \mathcal{L} \subseteq \mathrm{BC}\,\mathrm{ext}\, \mathcal{L}.$$
- **Theorem 3.20.** (Staiger-Wagner 2) $\mathcal{L}$ closed under suffix-independence, negation, union and change of final states. Then
$$\mathrm{BC}\,\mathrm{ext}\, \mathcal{L} \subseteq \lim \cap \widehat{\lim}\, \mathcal{L}.$$
- **Theorem 3.22.** $\mathcal{L}$ closed under suffix-independence, negation, union, change of final states and alphabet permutation. Then we have
$$\mathrm{ext} \cap \widehat{\mathrm{ext}}\, \mathcal{L} \overset{(1.)}{\subseteq} \mathrm{ext} \cup \widehat{\mathrm{ext}}\, \mathcal{L} \overset{(2.)}{\subseteq} \mathrm{BC}\,\mathrm{ext}\, \mathcal{L} \overset{(3.)}{=}$$
$$\lim \cap \widehat{\lim}\, \mathcal{L} \overset{(4.)}{\subseteq} \lim \cup \widehat{\lim}\, \mathcal{L} \overset{(5.)}{\subseteq} \mathrm{BC}\,\lim \mathcal{L}.$$

With $\mathcal{L}$-ext-$\widehat{\mathrm{ext}}$-separating language L_a, the inclusions in (1) and (2) are strict. With $\mathcal{L}$-lim-$\widehat{\lim}$-separating language L'_a, the inclusions in (4) and (5) are strict.

General results: Kleene closure

$$\text{Kleene}(\mathcal{L}) := \left\{ \bigcup_{i=1}^{n} U_i \cdot V_i^\omega \;\middle|\; U_i, V_i \subseteq \Sigma^*, U_i \cdot V_i^* \in \mathcal{L}, n \in \mathbb{N}_0 \right\}$$

- **Lemma 3.24.** $\mathcal{L}$ closed under change of final states for all deterministic simplified automata. Then

  $$\text{Kleene}\,\mathcal{L} \subseteq \text{BC}\lim\mathcal{L}.$$

  (The closure of final states here is stronger.) (The idea in the proof can probably be generalized into a general non-deterministic Büchi to deterministic Muller automaton conversion.)
- **Lemma 3.25.** $\mathcal{L}$ closed under change of final states. Then

  $$\lim\mathcal{L} \subseteq \text{Kleene}\,\mathcal{L}.$$

General results: congruence based classes

Let $R \subseteq \Sigma^* \times \Sigma^*$ be a congruence relation.

$$\mathcal{L}^*(R) := \{L \subseteq \Sigma^* \mid L \text{ is finite union of } R\text{-equivalence-classes}\} .$$

There is a canonical deterministic automaton with states $S_R := \Sigma^*/R$. We call it the R-automaton.

- Lemma 3.28. $\mathcal{L}(R)$ is *closed under change of final states*.
- Lemma 3.28. $\mathcal{L}(R)$ is *closed under negation, union and intersection*.
- Example 3.29. *Closure under suffix-independence* doesn't directly follow from this.
- Lemma 3.30. $\mathcal{L}^\omega_E(\mathcal{A}_R) = \text{ext}\,\mathcal{L}(R)$
- Lemma 3.31. $\mathcal{L}^\omega_{\text{Büchi}}(\mathcal{A}_R) = \lim \mathcal{L}(R)$
- Lemma 3.32. $\mathcal{L}^\omega_{\text{Muller}}(\mathcal{A}_R) = \text{BC} \lim \mathcal{L}(R)$
- Lemma 3.33. $\text{BC} \lim \mathcal{L}(R) \cap \text{ext}\,\mathcal{L}^*(\mathrm{reg}) = \text{ext}\,\mathcal{L}(R)$

General results: $\mathsf{BC}\lim \mathcal{L}(R) \cap \lim \mathcal{L}^*(\mathrm{reg}) = \lim \mathcal{L}(R)$

- **Definition 3.35.** $\mathcal{L}$ is **infinity-postfix-independent.**
  Lemma 3.36. $\mathcal{L}(R)$ is *infinity-postfix-independent* $\Leftrightarrow$ every SCC Q in the R-automata has exactly one looping subset, i.e. Q itself is the only loop in Q.
- **Lemma 3.39** $\mathcal{L}(R)$ *infinity-postfix-independent*. Then

  $$\mathsf{BC}\lim \mathcal{L}(R) \cap \lim \mathcal{L}^*(\mathrm{reg}) = \lim \mathcal{L}(R).$$

  (But $\not\Leftarrow$. Example 3.37 and 3.40.)
- **Definition 3.41.** If there is a SCC $Q \subseteq S_R$ including two loops $P_1, P_2 \subseteq Q$, $P_1 \neq P_2$ with $P_1 \not\subseteq P_2$, $P_2 \not\subseteq P_1$, then call $\mathcal{L}(R)$ **postfix-loop-deterministic.**
- **Theorem 3.44.** $\mathcal{L}(R)$ is not *postfix-loop-deterministic* $\Leftrightarrow$

  $$\mathsf{BC}\lim \mathcal{L}(R) \cap \lim \mathcal{L}^*(\mathrm{reg}) = \lim \mathcal{L}(R).$$

General results: $\mathcal{L}(R)$: Staiger-Wagner

- **Example 3.46.** There is $\mathcal{L}(R)$ infinity-postfix-independent and not postfix-loop-deterministic and

$$\operatorname{ext}\mathcal{L}(R) \not\subseteq \lim\mathcal{L}(R).$$

- **Theorem 3.47.** (Staiger-Wagner) $\mathcal{L}(R)$ not postfix-loop-deterministic. $\mathrm{BC}\operatorname{ext}\mathcal{L}(R) \subseteq \mathrm{BC}\lim\mathcal{L}(R)$. Then

$$\lim\cap\widehat{\lim}\,\mathcal{L}(R) = \mathrm{BC}\operatorname{ext}\mathcal{L}(R)$$

Concrete results

1. $\mathsf{BC}\,\mathsf{ext}\,\mathcal{L}^*(\mathrm{PT}) = \mathsf{BC}\,\mathsf{lim}\,\mathcal{L}^*(\mathrm{PT})$
2. $\mathcal{L}^\omega(\mathrm{FO}[+1]) = \mathsf{BC}\,\mathsf{ext}\,\mathcal{L}^*(\mathrm{FO}[+1])$
3. $\mathcal{L}^\omega(\mathrm{FO}[<]) = \mathsf{BC}\,\mathsf{lim}\,\mathcal{L}^*(\mathrm{FO}[<])$
4. $\mathsf{BC}\,\mathsf{ext}\,\mathcal{L}^*(\mathrm{FO}[<]) \subsetneq \mathsf{BC}\,\mathsf{lim}\,\mathcal{L}^*(\mathrm{FO}[<])$
5. $\mathsf{BC}\,\mathsf{ext}\,\mathcal{L}^*(\mathrm{LT}) \subsetneq \mathsf{BC}\,\mathsf{lim}\,\mathcal{L}^*(\mathrm{LT})$
6. $\mathsf{BC}\,\mathsf{ext}\,\mathcal{L}^*(\text{pos-PT}) = \mathsf{BC}\,\mathsf{lim}\,\mathcal{L}^*(\text{pos-PT})$
7. $\mathsf{BC}\,\mathsf{ext}\,\mathcal{L}^*(\text{pos-PT}) = \mathsf{BC}\,\mathsf{ext}\,\mathcal{L}^*(\mathrm{PT})$