

Language Operations and a Structure Theory of ω -Languages

Albert Zeyer

13. Mai 2012

Inhaltsverzeichnis

1	Introduction	3
2	regular ω-languages	4
3	$*$ $\rightarrow \omega$	5
3.1	language operators	5
3.2	$\mathcal{L}^*(reg)$	5
3.3	Questions	6
4	Generic results	7
4.1	\exists^ω power	7
4.2	non suffix sensitive	7
5	*-language classes	8
5.1	Overview	8
5.2	FO[i] / starfree	8
5.3	FO[+1]	8
5.4	FO[]	8
5.5	piece-wise testable	8
5.6	positive piece-wise testable	8
5.7	locally testable	8
5.8	endwise testable	8
5.9	local	8
5.10	finite / co-finite	8
5.11	dot-depth- n	8
5.12	L -trivial	8
5.13	R -trivial	8
5.14	locally modulo testable	8
5.15	context free	9

6	Lemmas	10
6.1	piece-wise testable	10
6.2	extension of $\mathcal{L}^*(\text{FO}[+1])$	10
6.3	limit of $\mathcal{L}^*(\text{FO}[<])$	11
6.4	$\text{BC ext } \mathcal{L}^*(\text{FO}[<]) \subsetneq \text{BC lim } \mathcal{L}^*(\text{FO}[<])$	12
6.5	locally testable	12
6.6	positive piece-wise testable	13
6.7	pos-PT and PT	13

1 Introduction

Language theory is strongly connected to the theory of automata. With some interpretation of run-acceptance in an automaton, we canonically get a language.

We call languages over infinite words the $*$ -languages and often use \mathcal{L}^* or some variant for such language class. Likewise, ω -languages (\mathcal{L}^ω) are over infinite words. The acceptance-condition in automata for $*$ -languages is straight-forward. If we look at ω -languages, several different types of automata and their acceptance have been thought of. For the class of regular languages, we see that many of them are equivalent.

For all types, we can also argue with equivalent language-theoretical operators which operate on a $*$ -language. We will study the equivalences in more detail.

Depending on the $* \rightarrow \omega$ language operator (or the ω -automaton acceptance condition), we get different ω -language classes. This was studied earlier already in detail for the class of regular $*$ -languages.

When we look at other $*$ -language classes, we might get different results. This study is the main topic of this thesis.

2 regular ω -languages

The class of regular ω -languages can be defined in many different ways. We will use one common definition and show some equivalent descriptions.

$$\mathcal{L}^\omega(\text{reg}) := \{\cup_i U_i \cdot V_i^\omega \mid U_i, V_i \in \mathcal{L}^*(\text{reg})\}$$

A different, very common description is in terms of automata.

An automaton $\mathcal{A} = (Q, \Sigma, E, I, F)$ **Büchi-accepts** a word $\alpha = (a_0, a_1, a_2, \dots) \in \Sigma^\omega$ iff there is an infinite run $q_0 \xrightarrow{a_0} q_1 \xrightarrow{a_1} q_2 \xrightarrow{a_2} q_3 \dots$ with $q_0 \in I$ and $\{q_i \mid q_i \in F\}$ infinite, i.e. which reaches a state in F infinitely often.

The language $\mathcal{L}^\omega(\mathcal{A})$ is defined as the set of all infinite words which are Büchi-accepted by \mathcal{A} .

An automaton \mathcal{A} is a Büchi automaton iff we use the Büchi-acceptance.

The set of all languages accepted by a non-deterministic Büchi automaton is exactly $\mathcal{L}^\omega(\text{reg})$. (S218,R101) Deterministic Büchi automata are less powerful, e.g. they cannot recognise $(a+b)^*b^\omega$.

There are some different forms of ω -automata, e.g. the Rabin automata and the Muller automata. We see that the class of languages accepted by non-deterministic Büchi automata is equal to deterministic Rabin automata and deterministic Muller automata. (S407)

We also see that this is equal to boolean combinations of languages accepted by deterministic Büchi automata. Under this regard, an operator of interest is $\lim(L) := \{\alpha \in \Sigma^\omega \mid \exists^\omega n: \alpha[0, n] \in L\}$. We see that $\lim(\mathcal{L}^\omega)$ is equal to the languages accepted by deterministic Büchi automata. (S407) Thus:

$$\text{BC } \lim \mathcal{L}^*(\text{reg}) = \mathcal{L}^\omega(\text{reg})$$

Some other descriptions: $\mathcal{L}^\omega(\text{reg}) = \{\cup_i U_i \cdot \lim V_i \mid U_i, V_i \in \mathcal{L}^*(\text{reg})\}$ (S218,S411,R107)
 $\mathcal{L}^\omega(\text{reg}) = \{A \subset \Sigma^\omega \mid A \text{ definable in } L_2(\Sigma)\}$

We will formulate some properties of interest in a general form for a $*$ -language class \mathcal{L} which all hold for $\mathcal{L}^*(\text{reg})$. Let $L, A, B \in \mathcal{L}$.

- E1: $L \cdot \Sigma^* \in \mathcal{L}$ (not suffix sensitive)
- E2a: $A \cup B \in \mathcal{L}$
- E2b: $A \cap B \in \mathcal{L}$
- E3: $-L \in \mathcal{L}$ (closed under complementation) (S303.E3, S218, R101)

3 $\ast \rightarrow \omega$

3.1 language operators

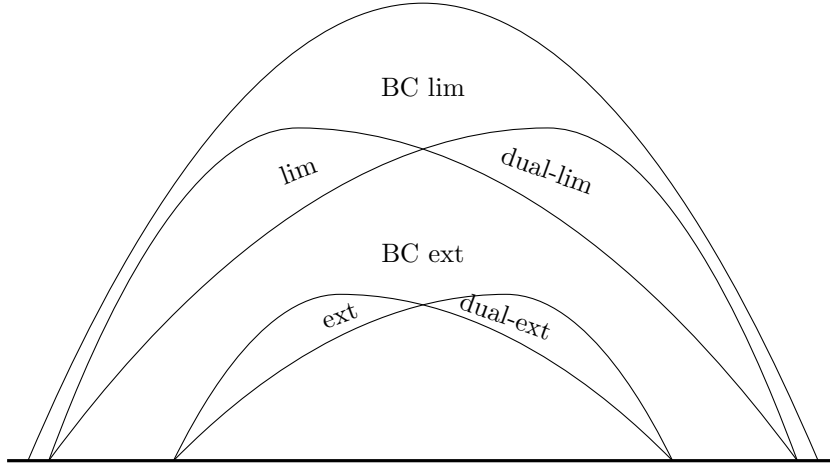
We already introduced \lim . We can define a family of language operators, partly also derived from the study of $\mathcal{L}^\omega(\text{reg})$. Some of these operators operate on a single language and not on the class. Let \mathcal{L} be a \ast -language class. Let $L \in \mathcal{L}$.

1. $\text{ext}(L) := \{\alpha \in \Sigma^\omega \mid \exists n: \alpha[0, n] \in L\} = L \cdot \Sigma^\omega$
2. $\overline{\text{ext}}(L) := \{\alpha \in \Sigma^\omega \mid \forall n: \alpha[0, n] \in L\} = L \cdot \Sigma^\omega$
3. BC ext
4. $\lim(L) := \{\alpha \in \Sigma^\omega \mid \forall N: \exists n > N: \alpha[0, n] \in L\} = \{\alpha \in \Sigma^\omega \mid \exists^\omega n: \alpha[0, n] \in L\}$
5. $\overline{\lim}(L) := \{\alpha \in \Sigma^\omega \mid \exists N: \forall n > N: \alpha[0, n] \in L\}$
6. BC lim
7. Kleene-Closure of \mathcal{L} : $\{\bigcup_{i=1}^n U_i \cdot V_i^\omega \mid U_i, V_i \in \mathcal{L}\}$
8. $\{\bigcup_{i=1}^n U_i \cdot \lim V_i \mid U_i, V_i \in \mathcal{L}\}$

From language operators, we get language class operators in a canonical way, e.g. $\lim(\mathcal{L}) := \{\lim L \mid L \in \mathcal{L}\}$.

3.2 $\mathcal{L}^\ast(\text{reg})$

Considering $\mathcal{L} := \mathcal{L}^\ast(\text{reg})$, we get a language diagram like:



where all inclusions are strict. In more detail:

- P1: $\text{ext } \mathcal{L} \cap \overline{\text{ext}} \mathcal{L} \neq \emptyset$
 Proof: $\tilde{L}_1 := a\Sigma^\omega \in \text{ext} \cap \overline{\text{ext}} \mathcal{L}$ with $\tilde{L}_1 = \text{ext}(a)$ and $\tilde{L}_1 = \overline{\text{ext}}(a\Sigma^*)$. (R101, prop, p.38)

- P2a: $\text{ext } \mathcal{L} \cap \overline{\text{ext } \mathcal{L}} \subsetneq \text{ext } \mathcal{L}$
 Proof: $\tilde{L}_{2a} := \text{ext}(a^*b) = a^*b\Sigma^\omega \in \text{ext } \mathcal{L}$. Assume some A-automaton \mathcal{A} with n states accepts \tilde{L}_{2a} . \mathcal{A} would also accept $a^n b^\omega$. I.e. the $(n+1)$ th state after the run of a^n would also accept a , i.e. \mathcal{A} would accept a^{n+1} . By inclusion, \mathcal{A} would accept a^ω . That is a contradiction. Thus, there is no such A-automaton. Thus, $\tilde{L}_{2a} \notin \overline{\text{ext } \mathcal{L}}$.
- P2b: $\text{ext } \mathcal{L} \cap \overline{\text{ext } \mathcal{L}} \subsetneq \overline{\text{ext } \mathcal{L}}$
 Proof: $\tilde{L}_{2b} := -\tilde{L}_{2a} \in \overline{\text{ext } \mathcal{L}}$, $\tilde{L}_{2b} \notin \text{ext } \mathcal{L}$.
- P3: $\text{ext } \mathcal{L} \neq \overline{\text{ext } \mathcal{L}}$
 Proof: Follows directly from P2a and P2b.
- P4: $\text{ext } \mathcal{L} \cup \overline{\text{ext } \mathcal{L}} \subsetneq \text{BC ext } \mathcal{L}$
 Proof: $\tilde{L}_4 := \Sigma^* a \Sigma^\omega \cap -(\Sigma^* b \Sigma^\omega)$, $\Sigma = \{a, b, c\}$. Then we have $\tilde{L}_4 \notin \text{ext } \cup \overline{\text{ext } \mathcal{L}}$, $\tilde{L}_4 \in \text{BC ext } \mathcal{L}$. (R101, p.38)
- P5: $\text{BC ext } \mathcal{L} = \lim \mathcal{L} \cap \overline{\lim \mathcal{L}}$
 Proof: S405 / Staigner-Wagner-recognizable
- P6a: $\lim \mathcal{L} \cap \overline{\lim \mathcal{L}} \subsetneq \lim \mathcal{L}$
 Proof: $\tilde{L}_{6a} := \lim(\Sigma^* a) = (\Sigma^* a)^\omega$. Assume there is $L \subseteq \Sigma^*$ with $\lim(L) = -\tilde{L}_{6a}$. Let $(w_0, w_1, w_2, \dots) \in (\Sigma^*)^\mathbb{N}$ so that $w_0 \in L, w_0 a w_1 \in L, \dots, w_0 \prod_{i=0}^n a w_i \in L \forall n \in \mathbb{N}$. Thus, $\alpha := w_0 \prod_{i \in \mathbb{N}} a w_i \in \lim L$. But $\alpha \notin -\tilde{L}_{6a}$. That is a contradiction. Thus, $-\tilde{L}_{6a} \notin \lim \mathcal{L}$. With E3, we get $\tilde{L}_{6a} \notin \overline{\lim \mathcal{L}}$.
- P6b: $\lim \mathcal{L} \cap \overline{\lim \mathcal{L}} \subsetneq \overline{\lim \mathcal{L}}$
 Proof: Analog to P6a with $\tilde{L}_{6b} := -\tilde{L}_{6a}$.
- P7: $\lim \mathcal{L} \neq \overline{\lim \mathcal{L}}$ Proof: Follows directly from P6a and P6b.
- P8: $\lim \mathcal{L} \cup \overline{\lim \mathcal{L}} \subsetneq \text{BC lim } \mathcal{L}$
 Proof: (S408, S409, ...)
- P9: $\text{BC lim } \mathcal{L} = \{\bigcup_{i=1}^n U_i \cdot V_i^\omega \mid U_i, V_i \in \mathcal{L}\}$
 Proof: (S402, S407, ...)
- P10: $\text{BC lim } \mathcal{L} = \{\bigcup_{i=1}^n U_i \cdot \lim V_i \mid U_i, V_i \in \mathcal{L}\}$
 Proof: (S403, S411, ...)

3.3 Questions

This was studied in detail for $\mathcal{L}^*(\text{reg})$. We are now studying relations of resulting ω -language classes for different $*$ -language classes.

Esp.:

- $\text{BC ext } \mathcal{L} \stackrel{?}{=} \text{BC lim } \mathcal{L}$

4 Generic results

S401, S302/S303

Let \mathcal{L} be a *-language class.

4.1 \exists^ω power

$\exists \tilde{L} \in \lim \mathcal{L}, \tilde{L} \neq \Sigma^\omega, \forall n: \tilde{L}[0, n] = \Sigma^n \Rightarrow \tilde{L} \notin \text{BC ext}(\mathcal{L}) \Rightarrow \text{BC ext}(\mathcal{L}) \neq \text{BC lim}(\mathcal{L})$

4.2 non suffix sensitive

$\forall L \in \mathcal{L}: L \cdot \Sigma^* \in \mathcal{L} \Rightarrow \text{ext}(\mathcal{L}^*) \subset \lim(\mathcal{L}^*)$

5 *-language classes

5.1 Overview

... We already showed many results for $\mathcal{L}^*(reg)$.

5.2 FO[i] / starfree

W161,S210,S211,S101a,S51,S18

5.3 FO[+1]

S213,215,S103,S17

5.4 FO[]

5.5 piece-wise testable

S212,S101,S52

5.6 positive piece-wise testable

S214,S101,S53

5.7 locally testable

S16

5.8 endwise testable

S104,S20

5.9 local

5.10 finite / co-finite

5.11 dot-depth- n

S21

5.12 L -trivial

5.13 R -trivial

S23

5.14 locally modulo testable

S24

5.15 context free

6 Lemmas

6.1 piece-wise testable

Theorem 6.1.

$$\text{BC ext } \mathcal{L}^*(\text{piece-wise testable}) = \text{BC lim } \mathcal{L}^*(\text{piece-wise testable})$$

Proof. L piece-wise testable $\Leftrightarrow L$ is a boolean algebra of $\Sigma^* a_1 \Sigma^* a_2 \cdots a_n \Sigma^*$

\subseteq : It is sufficient to show $\text{ext}(\mathcal{L}^*(\text{piece-wise testable})) \subseteq \text{BC lim } \mathcal{L}^*(\text{piece-wise testable})$.
By complete induction:

$$\begin{aligned} \text{ext}(\Sigma^* a_1 \Sigma^* a_2 \cdots a_n \Sigma^*) &= \Sigma^* a_1 \Sigma^* a_2 \cdots a_n \Sigma^\omega = \lim(\Sigma^* a_1 \Sigma^* a_2 \cdots a_n \Sigma^*) \\ \text{ext}(\neg(\Sigma^* a_1 \Sigma^* a_2 \cdots a_n \Sigma^*)) &= \Sigma^\omega = \lim(\Sigma^*) \\ \text{ext}(\emptyset) &= \emptyset = \lim(\emptyset) \end{aligned}$$

It is sufficient to show negation only for such ground terms because we can always push the negation down.

$$\begin{aligned} \text{ext}(A \cup B) &= \text{ext}(A) \cup \text{ext}(B) \\ \text{ext}(A \cap B) &= \text{ext}(A) \cap \text{ext}(B) \end{aligned}$$

This makes the induction complete.

\supseteq : It is sufficient to show $\lim(\mathcal{L}^*(\text{piece-wise testable})) \subseteq \text{BC ext } \mathcal{L}^*(\text{piece-wise testable})$.

$$\begin{aligned} \lim(\emptyset) &= \text{ext}(\emptyset), \quad \lim(\Sigma^* a_1 \Sigma^* a_2 \cdots a_n \Sigma^*) = \text{ext}(\Sigma^* a_1 \Sigma^* a_2 \cdots a_n \Sigma^*) \quad (\text{see above}) \\ \lim(\neg(\Sigma^* a_1 \Sigma^* a_2 \cdots a_n \Sigma^*)) &= \{ \alpha \in \Sigma^\omega \mid \exists^\omega n: \alpha[0, n] \notin \Sigma^* a_1 \Sigma^* a_2 \cdots a_n \Sigma^* \} \\ &= \{ \alpha \in \Sigma^\omega \mid \forall n: \alpha[0, n] \notin \Sigma^* a_1 \Sigma^* a_2 \cdots a_n \Sigma^* \} \\ &= \neg \text{ext}(\Sigma^* a_1 \Sigma^* a_2 \cdots a_n \Sigma^*) \\ \lim(A \cup B) &= \{ \alpha \in \Sigma^\omega \mid \exists^\omega n: \alpha[0, n] \in A \cup B \} = \lim(A) \cup \lim(B) \\ \lim(A \cap B) &= \{ \alpha \in \Sigma^\omega \mid \exists^\omega n: \alpha[0, n] \in A \cap B \} \end{aligned}$$

and because A, B are piece-wise testable

$$= \{ \alpha \in \Sigma^\omega \mid \exists n: \forall m > n: \alpha[0, m] \in A \cap B \} = \lim(A) \cap \lim(B)$$

□

6.2 extension of $\mathcal{L}^*(\text{FO}[+1])$

Theorem 6.2.

$$\mathcal{L}^\omega(\text{FO}[+1]) = \text{BC ext } \mathcal{L}^*(\text{FO}[+1])$$

Proof. From [Tho96, Theorem 4.8], we know that each formular in $\text{FO}[+1]$ is equivalent (for both finite and infinite words) to a boolean combination of statements “sphere $\sigma \in \Sigma^+$ occurs $\geq n$ times”. That statement can be expressed by a sentence of the form

$$\psi := \exists \bar{x}_1 \cdots \exists \bar{x}_n \varphi(\bar{x}_1, \dots, \bar{x}_n)$$

where each $\overline{x_i}$ is a $|\sigma|$ -tuple of variables and the formula φ states:

$$\bigwedge_{\substack{i,j \in \underline{n}, \\ i \neq j, \\ k,l \in \underline{|\sigma|}}} x_{i,k} \neq x_{j,l} \wedge \bigwedge_{\substack{i \in \underline{n}, \\ k \in \underline{|\sigma|-1}}} x_{i,k+1} = x_{i,k} + 1 \wedge \bigwedge_{\substack{i \in \underline{n}, \\ k \in \underline{|\sigma|}}} Q_{\sigma_k} x_{i,k}$$

For ψ , we have:

$$\alpha \models \psi \Leftrightarrow \exists n: \alpha[0, n] \models \psi \text{ for all } \alpha \in \Sigma^\omega,$$

i.e.

$$L^\omega(\psi) = \text{ext } L^*(\psi).$$

Any formula in $\text{FO}[+1]$ can be expressed as a boolean combination of ψ -like formula. With

$$\begin{aligned} L^\omega(\neg\psi) &= \neg L^\omega(\psi) \\ L^\omega(\psi_1 \wedge \psi_2) &= L^\omega(\psi_1) \cap L^\omega(\psi_2) \\ L^\omega(\psi_1 \vee \psi_2) &= L^\omega(\psi_1) \cup L^\omega(\psi_2) \end{aligned}$$

we get:

$$\mathcal{L}^\omega(\text{FO}[+1]) = \text{BC ext } \mathcal{L}^*(\text{FO}[+1]).$$

□

6.3 limit of $\mathcal{L}^*(\text{FO}[<])$

Theorem 6.3.

$$\mathcal{L}^\omega(\text{FO}[<]) = \text{BC lim } \mathcal{L}^*(\text{FO}[<])$$

Proof. Let $\varphi \in \text{FO}[<]$. By the [Tho81, Normal Form Theorem (4.4)] there are bounded formulas $\varphi_1(y), \dots, \varphi_r(y), \psi_1(y), \dots, \psi_r(y)$ such that for all $\alpha \in \Sigma^\omega$:

$$\alpha \models \varphi \Leftrightarrow \alpha \models \bigvee_{i=1}^r (\forall x \exists y > x: \varphi_i(y)) \wedge \neg (\forall x \exists y > x: \psi_i(y))$$

Thus:

$$\begin{aligned} \alpha \models \varphi &\Leftrightarrow \bigvee_{i=1}^r \underbrace{(\alpha \models \forall x \exists y > x: \varphi_i(y))}_{\Leftrightarrow \forall x \exists y > x: \alpha[0, n] \models \varphi_i(\omega)} \wedge \neg (\alpha \models \forall x \exists y > x: \psi_i(y)) \\ &\Leftrightarrow \exists^\omega n: \alpha[0, n] \models \varphi_i(\omega) \\ &\Leftrightarrow \alpha \in \lim L^*(\varphi_i(\omega)) \end{aligned}$$

where $\varphi_i(\omega)$ stands for φ_i with all bounds removed. I.e. we have

$$L^\omega(\varphi) = \bigcup_{i=1}^r \lim(L^*(\varphi_i(\omega)) \cap \neg \lim(L^*(\psi_i(\omega))),$$

and thus

$$L^\omega(\varphi) \in \text{BC lim } \mathcal{L}^*(\text{FO}[<]).$$

We have proved the \subseteq -direction. For \supseteq :

$$\begin{aligned} \alpha &\in \text{lim}(L^*(\varphi)) \\ \Leftrightarrow \exists^\omega n: \alpha[0, n] \models \varphi \\ \Leftrightarrow \alpha \models \forall x \exists y > x: \varphi(y) \\ \Leftrightarrow \alpha &\in L^\omega(\forall \exists y > x: \varphi(y)) \end{aligned}$$

where $\varphi(y)$ stands for φ with all variables bounded by y . I.e.

$$\text{lim } \mathcal{L}^*(\text{FO}[<]) \subseteq \mathcal{L}^\omega(\text{FO}[<]),$$

and thus also

$$\text{BC lim } \mathcal{L}^*(\text{FO}[<]) \subseteq \mathcal{L}^\omega(\text{FO}[<]).$$

Thus we have proved the equality. \square

6.4 $\text{BC ext } \mathcal{L}^*(\text{FO}[<]) \subsetneq \text{BC lim } \mathcal{L}^*(\text{FO}[<])$

Theorem 6.4.

$$\text{BC ext } \mathcal{L}^*(\text{FO}[<]) \subsetneq \text{BC lim } \mathcal{L}^*(\text{FO}[<])$$

Proof. $\subseteq: L \subset \Sigma^\omega \text{ starfree} \Rightarrow L\Sigma^\omega \in \text{lim}(\mathcal{L}^*(\text{FO}[<]))$ \square

Proof. \neq :

$$\begin{aligned} L &:= (\Sigma^* a)^\omega \\ \Rightarrow L &= \text{lim}((\Sigma^* a)^*) \\ \Rightarrow L &= L^\omega(\exists^\omega x: Q_a x) \end{aligned}$$

And we have $L \notin \text{BC ext } \mathcal{L}^*(\text{FO}[<])$. \square

6.5 locally testable

Theorem 6.5.

$$\text{BC ext } \mathcal{L}^*(\text{locally testable}) \subsetneq \text{BC lim } \mathcal{L}^*(\text{locally testable})$$

Proof. Let $w \in \Sigma^+$.

$$\begin{aligned} \text{ext}(w\Sigma^*) &= \text{lim}(w\Sigma^*) \\ \text{ext}(\Sigma^* w) &= \Sigma^* w\Sigma^\omega = \text{lim}(\Sigma^* w\Sigma^*) \\ \text{ext}(\Sigma^* w\Sigma^*) &= \Sigma^* w\Sigma^\omega = \text{lim}(\Sigma^* w\Sigma^*) \end{aligned}$$

Thus we have

$$\text{BC ext } \mathcal{L}^*(\text{locally testable}) \subseteq \text{BC lim } \mathcal{L}^*(\text{locally testable}).$$

But we also have

$$\text{lim}(\Sigma^*) = (\Sigma^* w)^\omega \notin \text{BC ext } \mathcal{L}^*(\text{locally testable}).$$

\square

6.6 positive piece-wise testable

Theorem 6.6.

$$\text{BC ext } \mathcal{L}^*(\text{pos-PT}) = \text{BC lim } \mathcal{L}^*(\text{pos-PT})$$

Proof. \subseteq : Exactly like the proof for PT except that we leave out the negated part. \supseteq : Also like the proof for PT. \square

6.7 pos-PT and PT

Theorem 6.7.

$$\text{BC ext } \mathcal{L}^*(\text{pos-PT}) = \text{BC ext } \mathcal{L}^*(\text{PT})$$

Proof. In the proof of $\text{lim } \mathcal{L}^*(\text{PT}) \subseteq \text{BC ext } \mathcal{L}^*(\text{PT})$ we actually proved $\text{BC lim } \mathcal{L}^*(\text{PT}) \subseteq \text{BC ext } \mathcal{L}^*(\text{pos-PT})$. Similiarly we also proved $\text{BC ext } \mathcal{L}^*(\text{PT}) \subseteq \text{BC lim } \mathcal{L}^*(\text{pos-PT})$.

With 6.6 and 6.1 we get the claimed equality. \square

Literatur

- [Tho81] Wolfgang Thomas. A combinatorial approach to the theory of omega-automata. *Information and Control*, 48(3):261–283, 1981.
- [Tho96] Wolfgang Thomas. Languages, automata, and logic. In *Handbook of Formal Languages*, pages 389–455. Springer, 1996.