Semilinear Sets over Commutative Semirings

Manfred Kudlek

Department Informatik, MIN-Fakultät, Universität Hamburg
Vogt-Kölln-Str. 30, D-22527 Hamburg, Germany
email : kudlek@informatik.uni-hamburg.de

Abstract. It is shown that the semilinear sets over commutative semirings are exactly the algebraic/rational sets over such structures.

1 Introduction

Multisets play an important role in the theory of Petri nets and membrane computing [5, 6, 10]. In [3] it was shown that the Parikh sets of context-free languages are semilinear. Context-free languages correspond to least fixed point solutions of algebraic systems of equations over an ω-complete semiring. In the case that the underlying (multiplicative) operation of the semiring is commutative, algebraic and rational solutions coincide [7]. Least fixed point solutions are also equivalent to sets (*languages*) generated by grammars. In the case of classical multisets, namely k-tuples over $\mathbb{N}$ this was shown in [6].

The notion of semilinear sets in [3] can be extended to commutative monoids M. In [2, 1] it has been shown that rational sets are exactly the semilinear sets over such structures. The proofs, however, are incomplete. This has been corrected in [4]. Actually, the statement holds for semilinear sets $\bigcup_{i=1}^{k}\{x_i\} \oplus P_i^{\oplus}$ where $\oplus$ is the monoid operation, $x \in M$, $P_i \subseteq M$, and $|P_i| < \infty$. Therefore, for the monoid $M = (2^{\Sigma^*}, ⧢, \{\lambda\})$, where Σ is an alphabet and ⧢ the shuffle operation, the result holds only for $x = A \in 2^{\Sigma^*}$.

The result can be strengthened in such a way that for a commutative monoid $(M, \oplus, \mathbf{0})$ for the also commutative monoid $(2^M, \oplus, \{\mathbf{0}\})$ the elements $x = \{z\}$ are singletons. $(2^M, \cup, \oplus, \emptyset, \{\mathbf{0}\})$ then is the corresponding ω-complete semiring.

In this paper it is shown that rational least fixed point solutions over a commutative ω-complete semiring are exactly the semilinear sets over that structure.

As a corollary a simple proof of the semilinearity of classical context-free languages is obtained. Furthernore, the result holds also for higher order multisets, e.g. multisets of multisets etc..

Rational/semilinear sets over more complex commutative semirings can be used to define simple subclasses of certain concurrency models as e.g. simple reference nets like MOB's (minimal object based nets) [8].

2 Definitions

Definition 2.1 Let $\mathbf{M} = (M, \oplus, \mathbf{0})$ be a structure where $\oplus$ is an associative and commutative operation with $\oplus : M \times M \to 2^M$, $|\{x\} \oplus \{y\}| < \infty$, and $\{x\} \oplus \{\mathbf{0}\} = \{x\}$.

$\oplus$ can be extended to 2^M by defining $A \oplus B = \bigcup_{x \in A, y \in B}(\{x\} \oplus \{y\})$. This also holds for infinite A or B. $\oplus$ then is associative and commutative on 2^M.

Thus $(2^M, \oplus, \{\mathbf{0}\})$ is a commutative monoid with neutral element $\{\mathbf{0}\}$.

Let be $\mathbf{S} = (2^M, \cup, \oplus, \emptyset, \{\mathbf{0}\})$ the corresponding ω-complete commutative semiring.

Example 2.1 Consider a finite alphabet Σ and the structure $(2^{\Sigma^*}, ⧢, \{\lambda\})$ where ⧢ is the *shuffle* operator which is associative and commutative. Then $(2^{\Sigma^*}, ⧢, \{\lambda\})$ is a commutative monoid, and $(2^{\Sigma^*}, \cup, ⧢, \emptyset, \{\lambda\})$ a commutative semiring.

Example 2.2 Consider $(\mathbb{N}^k, +, \mathbf{0})$, the set of multisets or k-vectors over $\mathbb{N}$. Clearly, $(\mathbb{N}^k, +, \mathbf{0})$ is a commutative monoid, as well as $(2^{\mathbb{N}^k}, +, \{\mathbf{0}\})$, and the structure $(2^{\mathbb{N}^k}, \cup, +, \emptyset, \{\mathbf{0}\})$ is a commutative semiring.

Example 2.3 Consider multisets of k-tuples $\tau \in \mathbb{N}^k$. This can be represented by a mapping $\mu : \mathbb{N}^k \to \mathbb{N}$. Denote that set of multisets by $\mathcal{M}(\mathbb{N}^k)$. If $\mu_1, \mu_2 \in \mathcal{M}(\mathbb{N}^k)$, an operation $+$ can be defined on $\mathcal{M}(\mathbb{N}^k)$ by $(\mu_1+\mu_2)(\tau) = \mu_1(\tau)+\mu_2(\tau)$ for all $\tau \in \mathcal{M}(\mathbb{N}^k)$. The operation $+$ is associative and commutative. Therefore $(\mathcal{M}(\mathbb{N}^k), +, O)$ is a commutative monoid, with neutral element O represented by $o(\tau) = 0$ for all $\tau \in \mathbb{N}^k$.

Lemma 2.1 Let $A, B, C, D \in 2^M$. Then
$(A \subseteq B \wedge C \subseteq D) \Rightarrow A \oplus C \subseteq B \oplus D$.

Proof. Trivial.

Lemma 2.2 Let $A, B \in 2^M$. Then $A \subseteq B \Rightarrow A^{\oplus} \subseteq B^{\oplus}$.

Proof. Trivial.

Lemma 2.3 Let $A \in 2^M$. Then $(A \cup \{\mathbf{0}\})^{\oplus} = A^{\oplus}$.

Proof. Straightforward.

Lemma 2.4 Let $A, B, C \in 2^M$. Then $(A \cup B) \oplus C = (A \oplus C) \cup (B \oplus C)$.

Proof. By the property of the semiring $\mathbf{S}$.

Lemma 2.5 Let $A_i, B \in 2^M$. Then $(\bigcup_{i \in I} A_i) \oplus B = \bigcup_{i \in I}(A_i \oplus B)$. I can also be infinite.

Proof. By the property of the semiring $\mathbf{S}$.

Lemma 2.6 Let $A, B \in 2^M$. Then $(A \cup B)^{\oplus} = A^{\oplus} \oplus B^{\oplus}$.

Proof. a) $x \in (A \cup B)^{\oplus}$

$$\begin{aligned} &\Rightarrow x \in \textstyle\bigoplus_{j=1}^{k}\{c_j\} \quad (\ c_j \in A \cup B\ ) \\ &\qquad = \textstyle\bigoplus_{r=1}^{k'}\{a_r\} \oplus \bigoplus_{s=1}^{k''}\{b_s\} \quad (\ a_r \in A, b_s \in B, k' + k'' = k\ ) \\ &\qquad \subseteq A^{\oplus} \oplus B^{\oplus}. \end{aligned}$$

b) $x \in A^{\oplus} \oplus B^{\oplus} \Rightarrow x \in \bigoplus_{r=1}^{k'}\{a_r\} \oplus \bigoplus_{s=1}^{k''}\{b_s\} \quad (\ a_r \in A, b_s \in B\ )$

$$\begin{aligned} &= \textstyle\bigoplus_{j=1}^{k'+k''}\{c_j\}\ (\ c_j \in A \cup B\ ) \\ &\subseteq (A \cup B)^{\oplus}. \end{aligned}$$

Lemma 2.7 Let $A_i \in 2^M$ ($i \in \{1, \cdots, k\}$).
Then $(\bigcup_{i=1}^{k} A_i)^{\oplus} = \bigoplus_{i=1}^{k} A_i^{\oplus}$.

Proof. By induction on k.

$A^{\oplus} = A^{\oplus}$.

$$\begin{aligned} (\textstyle\bigcup_{i=1}^{k+1} A_i)^{\oplus} &= (\textstyle\bigcup_{i=1}^{k} A_i) \cup A_{k+1})^{\oplus} \\ &= (\textstyle\bigoplus_{i=1}^{k} A_I^{\oplus}) \oplus A_{k+1}^{\oplus} \\ &= \textstyle\bigoplus_{i=1}^{k+1} A_I^{\oplus}. \end{aligned}$$

Lemma 2.8 Let $A \in 2^M$. Then $(A \cup \{\mathbf{0}\})^{\oplus} = A^{\oplus}$.

Proof. a) $A \subseteq A \cup \{\mathbf{0}\}$ is trivial.

b) $x \in (A \cup \{\mathbf{0}\})^{\oplus} \Rightarrow x \in \bigoplus_{j=1}^{s}\{a_j\} \quad (\ a_j \in A \cup \{\mathbf{0}\}\ )$

$$\begin{aligned} &\Rightarrow x \in \textstyle\bigoplus_{i=1}^{r}\{a_i\} \quad (\ a_i \neq \mathbf{0}\ ) \\ &\qquad \subseteq A^{\oplus}. \end{aligned}$$

Lemma 2.9 Let $A \in 2^M$. Then $(A^{\oplus})^{\oplus} = A^{\oplus}$.

Proof. a) $A^{\oplus} \subseteq (A^{\oplus})^{\oplus}$ is trivial.

b) $x \in (A^{\oplus})^{\oplus} \Rightarrow x \in \bigoplus_{i=1}^{r}\{b_i\} \quad (\ b_j \in A^{\oplus}\ )$

$$\begin{aligned} &\subseteq \textstyle\bigoplus_{i=1}^{r} \bigoplus_{j=1}^{s(i)}\{a_{ij}\} \quad (\ a_{ij} \in A\ ) \\ &\subseteq A^{\oplus}. \end{aligned}$$

Definition 2.2 $A \in 2^M$ is called *linear* with respect to $\oplus$ if there exist $z \in M$ and $P \in 2^M$ with $|P| < \infty$ such that $A = \{z\} \oplus P^{\oplus}$.

$A \in 2^M$ is called *semilinear* with repect to $\oplus$ if A is a finite union of linear sets :

$A = \bigcup_{j=1}^{k}(\{z_j\} \oplus P_j^{\oplus})$.

Definition 2.3 Let $\mathcal{X} = \{X_1, \cdots, X_m\}$ be a finite set of formal *variables*, and $\mathcal{C} = \{c_1, \cdots, c_n\}$ be a finite set of constants with $c_i \in M$.

Then any expression $m \in (\mathcal{X} \cup \mathcal{C})^+$ with operation $\oplus$ is called a *monomial*, and any finite union p of monomials (i.e. using the operation $\cup$) is called a *polynomial.* A *system of equations* $\mathcal{E}$ consists of equations $X_i = p_i(\mathbf{X})$ where $\mathbf{X} = (X_1, \cdots, X_m)$ and p_i is a polynomial in $\mathbf{X}$. Such a system is called algebraic if the monomials are arbitrary, linear if all monomials have the form $\{c\} \oplus Y \oplus \{c'\}$

or $\{c\}$, and rational if either all have the form $Y \oplus \{c\}$ or $\{c\}$, or all $\{c\} \oplus Y$ or $\{c\}$ with $c, c' \in \mathcal{C}$.

Such a system of equations $\mathbf{X} = \mathbf{p}(\mathbf{X})$ has a unique solution $\mathbf{X}$ as least fixed point which can be achieved by iteration :

$\mathbf{X}^0 = (\emptyset, \cdots, \emptyset)$, $\mathbf{X}^{t+1} = \mathbf{p}(\mathbf{X}^t)$.

Let the classes of sets defined in this way be denoted by $\mathbf{RAT}(\oplus)$, $\mathbf{LIN}(\oplus)$, and $\mathbf{ALG}(\oplus)$, respectively. More precisely, these classes depend on $\mathcal{C}$.

If the underlying operation $\oplus$ is commutative then the classes of rational, linear, and algebraic $\oplus$-languages coincide, i.e.

$\mathbf{RAT}(\oplus) = \mathbf{LIN}(\oplus) = \mathbf{ALG}(\oplus)$.

If the following condition $\mathbf{0} \in A \oplus B \;\Rightarrow\; \mathbf{0} \in A \wedge \mathbf{0} \in B$ is fulfilled, then to each system of equations there exists another one, possibly with additional variables, having the same solutions in the original variables, and all monomials in the normal forms ([5, 9]) :

$\{\mathbf{0}\}$ occurring with at most one equation of the form $X_0 = \{\mathbf{0}\}$ if $\mathbf{0} \in \mathcal{C}$, $Y \oplus Z$, or $\{c\}$ with $c \in \mathcal{C}$.

To each linear system of equations there exists another linear one with monomials of the following forms only :

$\{\mathbf{0}\}$ occurring with at most one equation of the form $X_0 = \{\mathbf{0}\}$ if $\mathbf{0} \in \mathcal{C}$, $Y \oplus \{c\}$, $\{c\} \oplus Y$, or $\{c\}$ with $c \in \mathcal{C}$.

To each rational system of equations there exists another rational one with monomials of the following forms only :

$\{\mathbf{0}\}$ occurring with at most one equation of the form $X_0 = \{\mathbf{0}\}$ if $\mathbf{0} \in \mathcal{C}$, $Y \oplus \{c\}$, or $\{c\}$ with $c \in \mathcal{C}$.

$\mathbf{RAT}(\oplus)$ is also identical to the algebraic closure of $\mathcal{C}$ under the operations $\cup, \oplus, {}^{\oplus}$.

3 Results

Theorem 3.1 For a commutative structure $(2^M, \oplus, \{\mathbf{0}\})$ any $A \in \mathbf{RAT}(\oplus)$ is semilinear with respect to $\oplus$.

Proof. By induction on the structure of A.

1) $\{a\}$ with $a \in M$ is semilinear since $\{a\} = \{a\} \oplus \{\mathbf{0}\}^{\oplus}$

Let $A = \bigcup_{i=1}^{k'}(\{z'_i\} \oplus {P'_i}^{\oplus})$, and $B = \bigcup_{i=1}^{k''}(\{z''_j\} \oplus {P''_j}^{\oplus})$.

2) Then $A \cup B = \bigcup_{i=1}^{k'}(\{z'_i\} \oplus {P'_i}^{\oplus}) \cup \bigcup_{i=1}^{k''}(\{z''_j\} \oplus {P''_j}^{\oplus})$, thus semilinear.

3)
$$\begin{aligned}
A \oplus B &= \bigcup_{i=1}^{k'}(\{z'_i\} \oplus {P'_i}^{\oplus}) \oplus \bigcup_{i=1}^{k''}(\{z''_j\} \oplus {P''_j}^{\oplus}) \\
&= \bigcup_{i=1}^{k'} \bigcup_{i=1}^{k''}((\{z'_i\} \oplus {P'_i}^{\oplus}) \oplus (\{z''_j\} \oplus {P''_j}^{\oplus})) \\
&= \bigcup_{i=1}^{k'} \bigcup_{i=1}^{k''}(\{z'_i\} \oplus \{z''_j\} \oplus {P'_i}^{\oplus} \oplus {P''_j}^{\oplus}) \\
&= \bigcup_{i=1}^{k'} \bigcup_{i=1}^{k''}((\{z'_i\} \oplus \{z''_j\}) \oplus (P'_i \cup P''_j)^{\oplus}) \\
&= \bigcup_{i=1}^{k'} \bigcup_{i=1}^{k''} \bigcup_{z \in \{z'_i\} \oplus \{z''_j\}}(\{z\} \oplus (P'_i \cup P''_j)^{\oplus})
\end{aligned}$$

which is semilinear since $\{z'_i\} \oplus \{z''_j\}$ is finite.

4) Let $A = \bigcup_{i=1}^{k}(\{z_i\} \oplus P_i^{\oplus})$.
Consider first the case $k = 1$, i. e. $A = \{z\} \oplus P^{\oplus}$. Then
$A^{\oplus} = (\{z\} \oplus P^{\oplus})^{\oplus} = (\{z\} \oplus (P \cup \{z\})^{\oplus}) \cup \{\mathbf{0}\} = (\{z\} \oplus P^{\oplus} \oplus \{z\}^{\oplus}) \cup \{\mathbf{0}\}$.
The case $x = \mathbf{0}$ is trivial. Assume $x \neq \mathbf{0}$.
a) Assume $x \in (\{z\} \oplus P^{\oplus})^{\oplus}$. Then

$$x \in \bigoplus_{i=1}^{r}(\{z\} \oplus \{b_i\}) \quad (b_i \in P^{\oplus})$$

$$\Rightarrow x \in \bigoplus_{i=1}^{r}\{z\} \oplus \bigoplus_{i=1}^{r} \bigoplus_{j=1}^{m(i)}\{a_{ij}\} \; (m(i) \geq 0,\, a_{ij} \in P, \; \bigoplus_{j=1}^{0}\{a_{ij}\} = \{\mathbf{0}\})$$

$$= \{z\} \oplus \bigoplus_{i=1}^{r-1} \oplus \bigoplus_{j=1}^{s}\{a_j\}$$

$$\subseteq \{z\} \oplus \{z\}^{\oplus} \oplus P^{\oplus}$$

b) Assume $x \in \{z\} \oplus P^{\oplus} \oplus \{z\}^{\oplus}$. Then
$x \in \bigoplus_{i=i}^{r}\{z\} \oplus \bigoplus_{j=1}^{s}\{a_j\} \quad (r \geq 1,\, s \geq 0,\, a_j \in P)$
If $r \leq s$ then trivially $x \in (\{z\} \oplus P^{\oplus})^{\oplus}$.
If $r > s$ then $x \in \bigoplus_{i=i}^{r}\{z\} \oplus \bigoplus_{j=1}^{s}\{a_j\} \oplus \bigoplus_{j=s+1}^{r}\{\mathbf{0}\}$, and therefore
$x \in (\{z\} \oplus P^{\oplus})^{\oplus}$, too.
The general case follows by Lemma 2.7 and 3). Hence $A^{\oplus}$ is semilinear.

Theorem 3.2 $\mathbf{RAT}(\oplus)$ is exactly the family of semilinear sets with respect to $\oplus$.

Proof. a) $A \in \mathbf{RAT}(\oplus)$ is semilinear by Theorem 3.1.
b) $\bigcup_{i=1}^{k}(\{z_i\} \oplus P_i^{\oplus}) \in \mathbf{RAT}(\oplus)$ by the definition of $\mathbf{RAT}(\oplus)$.

Corollary 3.1 The Parikh sets of context-free languages are semilinear.

Proof. Let $L = L(G)$ be generated by a context-free grammar in Chomsky normal form. Interprete that grammar as a multiset grammar as in [6]. Since the underlying structure $(2^M, +, \{\mathbf{0}\})$ is a commutative monoid, Theorem 3.1 holds, and therefore the generated multiset language, being identical to the Parikh set of L, is semilinear in the classical sense.

References

1. J. Berstel : *Transductions and Context-free Languages.* Teubner, 1979.
2. S. Eilenberg, M. P. Schützenberger : *Rational sets in Commutative Monoids.* Journ. of Algebra **13**, pp 173-191, 1969.
3. S. Ginsburg : *The Mathematical Theory of Context-free Languages.* McGraw Hill, 1966.
4. M. Jantzen : *Intersecting Multisets and Applications to Macrosets.* Fachbereich Informatik, Universität Hamburg,, Bericht 247, 2003.
5. M. Kudlek : *Rational, Linear and Algebraic Languages of Multisets.* Pre-Proceedings of the Workshop on Multiset Processing (WMP'2000), ed. C. S. Calude, M. J. Dinneen, G. Păun, CDMTCS-140, pp 138-148, 2000.
6. M. Kudlek, C. Martín Vide, Gh. Păun : *Toward FMT (Formal Macroset Theory).* Pre-Proceedings of the Workshop on Multiset Processing (WMP'2000), ed. C. S. Calude, M. J. Dinneen, G. Păun, CDMTCS-140, pp 149-158, 2000.

7. W. Kuich, A. Salomaa : *Semirings, Automata, Languages*. EATCS Monographs on Theoretical Computer Science 5, Springer, Berlin, 1986.
8. O. Kummer : *Referenznetze.* Logos-Verlag, 2002.
9. J. Mezei, J. B. Wright : *Algebraic Automata and Context-free Sets.* IC **11**, pp 3-29, 1967.
10. Gh. Păun : *Computing with Membranes. An Introduction.* Bulletin of EATCS 67, pp 139-152, 1999.