Syntactic Semiring of a Language
(Extended Abstract)

Libor Polák*

Department of Mathematics, Masaryk University
Janáčkovo nám 2a, 662 95 Brno, Czech Republic
polak@math.muni.cz
http://www.math.muni.cz/~polak

Abstract. A classical construction assigns to any language its (ordered) syntactic monoid. Recently the author defined the so-called syntactic semiring of a language. We discuss here the relationships between those two structures. Pin's refinement of Eilenberg theorem gives a one-to-one correspondence between positive varieties of rational languages and pseudovarieties of ordered monoids. The author's modification uses so-called conjunctive varieties of rational languages and pseudovarieties of idempotent semirings. We present here also several examples of our varieties of languages.

Keywords: syntactic semiring, rational languages

1 Introduction

The syntactic monoid is a monoid canonically attached to each language. Certain classes of rational languages can be characterized by the syntactic monoids of their members. The book [3] and the survey [5] by Pin are devoted to a systematic study of this correspondence. One also speaks about the algebraic theory of finite automata.

Recently the author defined in [6] the so-called syntactic semiring of a language. Basic definitions and properties are stated in Sect. 2. In Sect. 3 we recall the main result of [6]. It is an Eilenberg-type theorem relating the so-called conjunctive varieties of rational languages and pseudovarieties of idempotent semirings.

Section 4 is devoted to examples. The first two varieties of languages arise when considering identities satisfied by syntactic semirings. Certain operators on classes of languages yield further examples.

The relationships between the (ordered) syntactic monoid and the syntactic semiring of a given language are studied in Sect. 5 and the last section presents a method for computing the syntactic semiring of a given language.

* The author acknowledges the support of the Grant no. 201/01/0323 of the Grant Agency of the Czech Republic.

J. Sgall, A. Pultr, and P. Kolman (Eds.): MFCS 2001, LNCS 2136, pp. 611–620, 2001.

All the background needed is thoroughly explored in above quoted Pin's sources or in Almeida's book [1]. Our contribution is meant as an extended abstract, we present full proofs only in Sect. 5.

2 Ordered Syntactic Monoid and Syntactic Semiring

A structure $(O, \cdot, \leq)$ is called an *ordered monoid* if

(i) $(O, \cdot)$ is a monoid with the neutral element 1,
(ii) $(O, \leq)$ is an ordered set,
(iii) $a, b, c \in O,\ a \leq b$ implies $ac \leq bc$ and $ca \leq cb$.

A structure $(S, \cdot, \vee)$ is called a *semilattice-ordered monoid* if

(i) $(S, \cdot)$ is a monoid,
(ii) $(S, \vee)$ is a semilattice,
(iii) $a, b, c \in S$ implies $a(b \vee c) = ab \vee ac$ and $(a \vee b)c = ac \vee bc$.

An alternative name is an *idempotent semiring* which we will prefer here. A semilattice $(S, \vee)$ becomes an ordered set with respect to the relation $\leq$ defined by $a \leq b \ \Leftrightarrow\ a \vee b = b,\ a, b \in S$. Moreover, in this view, any idempotent semiring is an ordered monoid. For an ordered set $(O, \leq)$, a subset H of O is *hereditary* if $a \in H,\ b \in O,\ b \leq a$ implies $b \in H$.

We also denote

2^B the set of all subsets of a set B,
$\mathsf{F}(B)$ the set of all non-empty finite subsets of a set B,
$\varDelta_B = \{(b, b) \mid b \in B\}$ the diagonal relation on a set B,
$(B] = \{a \in O \mid a \leq b$ for some $b \in B\}$ the hereditary subset of $(O, \leq)$ generated by a given $B \subseteq O$,
$\mathsf{H}(O, \leq)$ the set of all non-empty finitely generated hereditary subsets of an ordered set $(O, \leq)$,
$\mathsf{c}(u)$ the set of all letters of A in a word $u \in A^*$.

Recall that an *ideal* of a semilattice $(S, \vee)$ is a hereditary subset closed with respect to the operation $\vee$. *Morphisms* of monoids are semigroup homomorphisms with $1 \mapsto 1$ and morphisms of ordered monoids (idempotent semirings) are isotone monoid morphisms (monoid morphisms which respect also the operation $\vee$).

We say that a monoid $(M, \cdot)$ *divides* a monoid $(N, \cdot)$ if $(M, \cdot)$ is a morphic image of a submonoid of $(N, \cdot)$; similarly for ordered monoids and semirings.

A hereditary subset H of an ordered monoid $(O, \cdot, \leq)$ defines a relation $\approx_H$ on O by

$$a \ \approx_H \ b \ \text{ if and only if } \ (\ \forall\, p, q \in O\ )\ (\ paq \in H \Leftrightarrow pbq \in H\ )\ .$$

This relation is a congruence of $(O, \cdot)$ and the corresponding factor-structure is called the *syntactic monoid* of H in $(O, \cdot, \leq)$. It is ordered by

$$a \approx_H \ \leq \ b \approx_H \ \text{ if and only if } \ (\ \forall\, p, q \in O\ )\ (\ pbq \in H \ \Rightarrow\ paq \in H\ )$$

and we speak about the *ordered syntactic monoid*. We also write $a \preceq_H b$ instead of $a \approx_H \ \leq \ b \approx_H$. Let σ_H denote the mapping $a \mapsto a \approx_H, \ a \in O$. It is a surjective ordered monoid morphism.

Similarly, an ideal I of an idempotent semiring $(S, \cdot, \vee)$ defines a relation $\sim_I$ on the set S by

$$a \ \sim_I \ b \ \text{ if and only if } \ (\ \forall\, p, q \in S\ )\ (\ paq \in I \Leftrightarrow pbq \in I\ )\ .$$

This relation is a congruence of $(S, \cdot, \vee)$ and the corresponding factor-structure is called the *syntactic semiring* of I in $(S, \cdot, \vee)$. Let ρ_I denote the mapping $a \mapsto a \sim_I, \ a \in S$. It is a surjective semiring morphism.

Any language L over A is a hereditary subset of the trivially ordered monoid $(A^*, \cdot, =)$ and the set $\mathsf{F}(L)$ is an ideal of $(\mathsf{F}(A^*), \cdot, \cup)$ where

$$U \cdot V = \{u \cdot v \mid u \in U, \ v \in V\}\ .$$

The above constructions gives the *ordered syntactic monoid* and *syntactic semiring of the language* L; we denote them by $\mathsf{O}(L)$ and $\mathsf{S}(L)$. We also put $\mathsf{o}(L) = \sigma_L(L)$ and $\mathsf{s}(L) = \rho_{\mathsf{F}(L)}(\mathsf{F}(L))$.

For $L \subseteq A^*$, the congruence $\sim_{\mathsf{F}(L)}$ can be also expressed as

$$\{u_1, \dots, u_k\} \sim_{\mathsf{F}(L)} \{v_1, \dots, v_l\} \text{ if and only if}$$

$$(\ \forall\, x, y \in A^*\ )\ (\ xu_1y, ..., xu_ky \in L \ \Leftrightarrow \ xv_1y, ..., xv_ly \in L\ )\ .$$

Notice that $(A^*, \cdot)$ is a free monoid over the set A and that $(\mathsf{F}(A^*), \cdot, \cup)$ is a free idempotent semiring over A. Further, any ideal of $(\mathsf{F}(A^*), \cdot, \cup)$ is of the form $\mathsf{F}(L)$ for $L \subseteq A^*$.

A language $L \subseteq A^*$ is *recognizable* by a monoid $(M, \cdot)$ with respect to its subset B (by an ordered monoid $(O, \cdot, \leq)$ with respect to its hereditary subset H, by an idempotent semiring $(S, \cdot, \vee)$ with respect to its ideal I) if there exists a monoid morphism $\alpha : A^* \to M$ such that $L = \alpha^{-1}(B)$ ($\alpha : A^* \to O$ such that $L = \alpha^{-1}(H)$, $\alpha : A^* \to S$ such that $L = \alpha^{-1}(I)$).

Clearly, the last notion can be rephrased as follows: A language $L \subseteq A^*$ is recognizable by an idempotent semiring $(S, \cdot, \vee)$ with respect to its ideal I if there exists a semiring morphism $\beta : (\mathsf{F}(A^*), \cdot, \cup) \to (S, \cdot, \vee)$ such that $\mathsf{F}(L) = \beta^{-1}(I)$.

We say that a language L is *recognizable* by a monoid $(M, \cdot)$ (by an ordered monoid $(O, \cdot, \leq)$, by an idempotent semiring $(S, \cdot, \vee)$) if it is recognizable with respect to some subset (hereditary subset, ideal). Clearly, the recognizability by $(M, \cdot)$ gives the recognizability by $(M, \cdot, =)$ and conversely the recognizability by $(M, \cdot, \leq)$ yields the recognizability by $(M, \cdot)$. Furthermore, the *recognizability* means the recognizability by a finite monoid.

Proposition 1. *Let $(O, \cdot, \leq)$ be an ordered monoid. Then $(\mathsf{H}(O, \leq), \circ, \cup)$, where*

$$H \circ K = \{\ c \in O \mid \ \text{there exist } a \in H, \ b \in K \text{ such that } c \leq a \cdot b\ \}\ ,$$

is an idempotent semiring and the mapping $\iota : a \mapsto (a], \ a \in O$, is an injective monoid morphism of $(O, \cdot, \leq)$ to $(\mathsf{H}(O, \leq), \circ, \cup)$ satisfying $a \leq b$ if and only if

$\iota(a) \subseteq \iota(b),\ a, b \in O$. Moreover, the latter structure is a free idempotent semiring over $(O, \cdot, \leq)$ with respect to ι, that is, for any idempotent semiring $(S, \cdot, \vee)$ and an ordered monoid morphism

$$\alpha : (O, \cdot, \leq) \to (S, \cdot, \vee)$$

there exists exactly one semiring morphism

$$\beta : (\mathsf{H}\,(O, \leq), \circ, \cup) \to (S, \cdot, \vee)$$

such that $\beta \circ \iota = \alpha$.

Proof. Obviously $d \in (H \circ K) \circ L$ iff there exist $a \in H,\ b \in K,\ c \in L$ such that $d \leq abc$ and the same holds for $H \circ (K \circ L)$. Also the validity of the distributive laws is immediate. Further, $a \leq b$ iff $(a] \subseteq (b]$ (in particular, ι is injective) and $(ab] = (a] \cdot (b]$ for any $a, b \in O$.

Given $(S, \cdot, \vee)$ and α, define $\beta((a_1, \ldots, a_k]) = \alpha(a_1) \vee \ldots \vee \alpha(a_k)$ which yields the rest of the proposition. □

Proposition 2. *A language $L \subseteq A^*$ is recognizable by a finite idempotent semiring if and only if it is recognizable.*

Proof. If L is recognizable by a finite ordered monoid $(O, \cdot, \leq)$ with respect to its hereditary subset H then L is also recognizable by the idempotent semiring $(\mathsf{H}\,(O, \leq), \circ, \cup)$ with respect to its ideal $(H]$ of all hereditary subsets of H.

The opposite implication is trivial. □

The following result is classical (see [5]).

Proposition 3. *(i) Let A be a finite set, let $(O, \cdot, \leq)$ be an ordered monoid, H its hereditary subset, $\alpha : (A^*, \cdot) \to (O, \cdot)$ a surjective monoid morphism, and let $L = \alpha^{-1}(H)$. Then*

$$((O, \cdot)/\approx_H, \leq) \text{ is isomorphic to } (\mathsf{O}\,(L), \cdot, \leq),\ \mathsf{o}\,(L) \text{ being the image of } \sigma_H(H)$$

$$\text{and } \preceq_{\mathsf{o}\,(L)} \ = \ \leq \text{ on } \mathsf{O}\,(L)\ .$$

(ii) A language L over a finite set is recognizable by a finite ordered monoid $(O, \cdot, \leq)$ if and only if $(\mathsf{O}\,(L), \cdot, \leq)$ divides $(O, \cdot, \leq)$. □

An analogy of the last result follows.

Proposition 4. *(i) Let A be a finite set, let $(S, \cdot, \vee)$ be an idempotent semiring, I its ideal, $\beta : (\mathsf{F}\,(A^*), \cdot, \cup) \to (S, \cdot, \vee)$ a surjective semiring morphism, and let*

$$L = \{\ u \in A^* \mid \{u\} \in \beta^{-1}(I)\ \}\ .$$

Then

$$(S, \cdot, \vee)/\sim_I \quad \text{is isomorphic to } (\mathsf{S}\,(L), \cdot, \vee),\ \mathsf{s}\,(L) \text{ being the image of } \rho_I(I)$$

$$\text{and } \sim_{\mathsf{s}\,(L)} \ = \Delta_{\mathsf{S}\,(L)} \text{ on } \mathsf{S}\,(L)\ .$$

(ii) A language L over a finite set is recognizable by a finite idempotent semiring $(S, \cdot, \vee)$ if and only if $(\mathsf{S}\,(L), \cdot, \vee)$ divides $(S, \cdot, \vee)$.

Proof. (i) Here $\{u_1, \ldots, u_k\} \sim_{\mathsf{F}(L)} \quad \mapsto \beta(\{u_1, \ldots, u_k\}) \sim_I$ is the desired isomorphism and the second part is in Lemma 2 of [6]. The item (ii) is Lemma 10 of [6]. □

3 Eilenberg-Type Theorems

For languages $K, L \subseteq A^*$ we define

$$K \cdot L = \{\ uv \mid u \in K,\ v \in L\ \},\ K^* = \{\ u_1 \cdot \ldots \cdot u_k \mid k \geq 0,\ u_1, \ldots, u_k \in K\ \}\ .$$

Recall that the set of all *rational* languages over a finite alphabet A is the smallest family of subsets of A^* containing the empty set, all singletons $\{u\}$, $u \in A^*$, closed with respect to binary unions and the operations $\cdot$ and *.

As well-known, the rational languages are exactly the recognizable ones.

A class of finite monoids (ordered monoids, idempotent semirings) is called a *pseudovariety of monoids (ordered monoids, idempotent semirings)* if it is closed under forming of products of finite families, substructures and morphic images.

For sets A and B, a semiring morphism

$$\psi : (\mathsf{F}(A^*), \cdot, \cup) \to (\mathsf{F}(B^*), \cdot, \cup)$$

and $L \subseteq B^*$ we define

$$\psi^{[-1]}(L) = \{\ u \in A^* \mid \psi(\{u\}) \subseteq L\ \} \text{ and}$$

$$\psi^{(-1)}(L) = \{\ u \in A^* \mid \psi(\{u\}) \cap L \neq \emptyset\ \}\ .$$

A *class* of rational languages is an operator $\mathcal{L}$ assigning to every finite set A a set $\mathcal{L}(A)$ of rational languages over the alphabet A containing both $\emptyset$ and A^*.

Conditions which such a class of languages can satisfy follow.

- $(\cap)$: for every A, the set $\mathcal{L}(A)$ is closed with respect to finite intersections,
- $(\cup)$: for every A, the set $\mathcal{L}(A)$ is closed with respect to finite unions,
- (Q) : for every A, $a \in A$ and $L \in \mathcal{L}(A)$ we have $a^{-1}L,\ La^{-1} \in \mathcal{L}(A)$,
- $(^{-1})$: for every sets A and B, a monoid morphism $\phi : A^* \to B^*$ and $L \in \mathcal{L}(B)$ we have $\phi^{-1}(L) \in \mathcal{L}(A)$,
- $(^{[-1]})$: for every sets A and B, a semiring morphism $\psi : (\mathsf{F}(A^*), \cdot, \cup) \to (\mathsf{F}(B^*), \cdot, \cup)$ and $L \in \mathcal{L}(B)$ we have $\psi^{[-1]}(L) \in \mathcal{L}(A)$,
- $(^{(-1)})$: for every sets A and B, a semiring morphism $\psi : (\mathsf{F}(A^*), \cdot, \cup) \to (\mathsf{F}(B^*), \cdot, \cup)$ and $L \in \mathcal{L}(B)$ we have $\psi^{(-1)}(L) \in \mathcal{L}(A)$,
- (C) : for every A, the set $\mathcal{L}(A)$ is closed with respect to complements.

For a class $\mathcal{L}$ of languages we define its *complement* $\mathcal{L}^{\mathrm{c}}$ by

$$\mathcal{L}^{\mathrm{c}}(A) = \{\ A^* \setminus L \mid L \in \mathcal{L}(A)\ \} \text{ for every finite set } A\ .$$

Clearly, $\mathcal{L}$ satisfies the condition $(\cap)$ if and only if $\mathcal{L}^{\mathrm{c}}$ satisfies $(\cup)$. Similarly, $\mathcal{L}$ satisfies the condition $({}^{[-1]})$ if and only if $\mathcal{L}^{\mathrm{c}}$ satisfies $({}^{(-1)})$. Further, either of the conditions $({}^{[-1]})$ and $({}^{(-1)})$ implies $({}^{-1})$.

A class $\mathcal{L}$ is called a *conjunctive variety of languages* if it satisfies the conditions $(\cap)$, $({}^{[-1]})$ and (Q). Similarly, it is called a *disjunctive variety of languages* if it satisfies $(\cup)$, $({}^{(-1)})$ and (Q).

Further, $\mathcal{L}$ is called *a positive variety of languages* if it satisfies the conditions $(\cap)$, $(\cup)$, $({}^{-1})$, (Q) and such a variety is called *a boolean variety of languages* if it satisfies in addition the condition (C).

We can assign to any class of languages $\mathcal{L}$ the pseudovarieties

$$\begin{aligned}\mathsf{M}\,(\mathcal{L}) &= \langle\{\ (\mathsf{O}\,(L),\cdot) \mid A \text{ a finite set, } L \in \mathcal{L}(A)\ \}\rangle_{\mathsf{M}}\ ,\\ \mathsf{O}\,(\mathcal{L}) &= \langle\{\ (\mathsf{O}\,(L),\cdot,\leq) \mid A \text{ a finite set, } L \in \mathcal{L}(A)\ \}\rangle_{\mathsf{O}}\ ,\\ \mathsf{S}\,(\mathcal{L}) &= \langle\{\ (\mathsf{S}\,(L),\cdot,\vee) \mid A \text{ a finite set, } L \in \mathcal{L}(A)\ \}\rangle_{\mathsf{S}}\end{aligned}$$

of monoids (ordered monoids, idempotent semirings) generated by all syntactic monoids (ordered syntactic monoids, syntactic semirings) of members of $\mathcal{L}$.

Conversely, for pseudovarieties $\mathcal{M}$ of monoids, $\mathcal{P}$ of ordered monoids and $\mathcal{V}$ of idempotent semirings and a finite set A, we put

$$\begin{aligned}(\mathsf{L}\,(\mathcal{M}))(A) &= \{\ L \subseteq A^* \mid (\mathsf{O}\,(L),\cdot) \in \mathcal{M}\ \}\ ,\\ (\mathsf{L}\,(\mathcal{P}))(A) &= \{\ L \subseteq A^* \mid (\mathsf{O}\,(L),\cdot,\leq) \in \mathcal{P}\ \}\ ,\\ (\mathsf{L}\,(\mathcal{V}))(A) &= \{\ L \subseteq A^* \mid (\mathsf{S}\,(L),\cdot,\vee) \in \mathcal{V}\ \}\ .\end{aligned}$$

Theorem 5 (Eilenberg [2], Pin [4], Polák [6]). *(i) The operators* M *and* L *are mutually inverse bijections between boolean varieties of languages and pseudovarieties of monoids.*
(ii) The operators O *and* L *are mutually inverse bijections between positive varieties of languages and pseudovarieties of ordered monoids.*
(iii) The operators S *and* L *are mutually inverse bijections between conjunctive varieties of languages and pseudovarieties of idempotent semirings.*

4 Examples

Let $X = \{x_1, x_2, \ldots\}$ be the set of variables. Any identity of idempotent semirings is of the form

$$f_1 \vee \ldots \vee f_k = g_1 \vee \ldots \vee g_l \text{ where } f_1, \ldots, f_k, g_1, \ldots, g_l \in X^*\ .$$

It is equivalent to the inequalities

$$f_1 \vee \ldots \vee f_k\ \geq\ g_1 \vee \ldots \vee g_l, \quad f_1 \vee \ldots \vee f_k\ \leq\ g_1 \vee \ldots \vee g_l\ .$$

The last inequality is equivalent to k inequalities

$$f_i\ \leq\ g_1 \vee \ldots \vee g_l,\ i = 1, \ldots, k\ .$$

Thus when dealing with varieties of idempotent semirings, it is enough to consider identities of the form

$$f \;\leq\; g_1 \vee \ldots \vee g_l \;\; .$$

Notice that identities of the form $f \leq g$ are not sufficient since, for instance, the identity $xy \leq x \vee yz$ is not equivalent to any set of identities of the previous form. In fact, the variety (pseudovariety) of all (finite) idempotent semirings $(S, \cdot, \vee)$ such that also $(S, \cdot)$ is a semilattice has precisely three proper non-trivial subvarieties (subpseudovarieties). These are given by additional identities $x \leq 1,\; 1 \leq x$ and $xy \leq x \vee yz$, respectively.

For idempotent semirings the identities $x^2 \leq x$ and $xy \leq x \vee y$ are equivalent. The syntactic semiring of a language L over an alphabet A satisfies the identity $xy \leq x \vee y$ if and only if all its quotients $p^{-1}Lq^{-1}$ $(p, q \in A^*)$ satisfy

$$u, v \in p^{-1}Lq^{-1} \text{ implies } u \cdot v \in p^{-1}Lq^{-1} \;\; .$$

This can be generalized to the identities of the form

$$(I_k) \quad x_1 y_1 \ldots x_k y_k \leq x_1 \ldots x_k \vee y_1 \ldots y_k \;\; .$$

Now the syntactic semiring of a language L satisfies (I_k) if and only if every quotient K of L satisfies

$$u_1 \ldots u_k,\; v_1 \ldots v_k \in K \text{ implies } u_1 v_1 \ldots u_k v_k \in K \;\; .$$

Recall that the *shuffle* of words $u, v \in A^*$ is the set $u \sqcup v =$

$$\{\; u_1 v_1 \ldots u_k v_k \,|\, k \in \mathbb{N},\; u = u_1 \ldots u_k,\; v = v_1 \ldots v_k,\; u_1, \ldots, u_k, v_1, \ldots, v_k \in A^* \;\} \; .$$

Thus the system of all $(I_k),\; k \in \mathbb{N}$ characterizes languages all quotients of which are shuffle-closed.

Consider now the identity $x \leq 1 \vee xy$. The syntactic semiring of a language L satisfies this identity if and only if every quotient K of L satisfies

$$1,\; uv \in K \text{ implies } u \in K \;\; .$$

It is equivalent to the following condition on right quotients Lq^{-1} of L: if a deterministic automaton for Lq^{-1} accepts a word p and does not accept pa for a letter a, it does not need to continue reading of the input since any par $(r \in A^*)$ is not accepted.

In certain aspects the disjunctive varieties of languages are more natural than the conjunctive ones. For languages K and L over the alphabet A we define their *shuffle product* by

$$K \sqcup L = \bigcup_{u \in K, v \in L} u \sqcup v \;\; .$$

Let $\mathcal{L}(A)$ consist of finite unions of

$$\{v\} \sqcup C^*,\; vA^* \sqcup C^*,\; A^*vA^* \sqcup C^*,\; A^*v \sqcup C^*, \;\; \text{where } v \in A^*,\; C \subseteq A \;\; .$$

One can show that $\mathcal{L}$ is a disjunctive variety of languages. $\mathcal{L}$ is not a positive variety since, for instance, the language $aA^* \cap A^*b$ over $A = \{a, b\}$ is not from $\mathcal{L}(A)$.

Another method for obtaining a disjunctive variety of languages is to close a positive variety of languages first with respect to $(^{(-1)})$ and then to finite unions. On the other hand, starting from a disjunctive variety $\mathcal{L}$ of languages one gets a positive one closing every $\mathcal{L}(A)$ with respect to finite intersections.

5 Relationships between Syntactic Ordered Monoid and Syntactic Semiring of a Language

Proposition 6 ([6], Lemma 7). *Let L be a language over an alphabet A. The mapping*

$$\iota : u \approx_L \;\; \mapsto \{u\} \sim_{\mathsf{F}(L)},\; u \in A^*$$

is an injective monoid morphism of $(\mathsf{O}(L), \cdot, \leq)$ into $(\mathsf{S}(L), \cdot, \vee)$ satisfying $a \leq b$ if and only if $\iota(a) \leq \iota(b)$, $a, b \in \mathsf{O}(L)$. Moreover, $\iota(\mathsf{o}(L)) = \mathsf{s}(L) \cap \iota(\mathsf{O}(L))$ and $\iota(\mathsf{O}(L))$ contains all join-irreducible elements of $(\mathsf{S}(L), \vee)$.

Proposition 7. *Let L be a recognizable language over a finite alphabet A. Then*

$$(\mathsf{S}(L), \cdot, \vee) \text{ is isomorphic to } (\mathsf{H}(\mathsf{O}(L), \leq), \circ, \cup)/ \sim_{(\mathsf{o}(L)]} \;\; .$$

Proof. This is a consequence of Prop. 4 (i) since, as mentioned in the proof of Prop. 2, the idempotent semiring $(\mathsf{H}(\mathsf{O}(L), \leq), \circ, \cup)$ recognizes L with respect to $(\mathsf{o}(L)]$. □

Proposition 8. *Let $(S, \cdot, \vee)$ be a finite idempotent semiring and let I be its ideal such that $\sim_I \; = \Delta_S$. Let O be a submonoid of $(S, \cdot)$ containing all join-irreducible elements of $(S, \vee)$. Then there exists a recognizable language L over a finite set A such that*

$$(\mathsf{O}(L), \cdot, \leq) \text{ is isomorphic to } (O, \cdot, \leq) \text{ and } (\mathsf{S}(L), \cdot, \vee) \text{ is isomorphic to } (S, \cdot, \vee)\,.$$

In particular, it is the case for $O = S$.

Proof. Let $a, b \in O$. We show that

$$a \preceq_{O \cap I} b \text{ if and only if } a \vee b = b \text{ in } (S, \vee) \;\; .$$

Really, $a \preceq_{O \cap I} b$ if and only if

$$(\; \forall\, p, q \in O \;)(\; pbq \in O \cap I \quad \Rightarrow \quad paq \in O \cap I \;) \;\; ,$$

$$\text{that is, iff } (\; \forall\, p, q \in O \;)(\; pbq \in I \quad \Rightarrow \quad paq \in I \;) \;\; ,$$

$$\text{that is, iff } (\; \forall\, p, q \in S \;)(\; pbq \in I \quad \Rightarrow \quad paq \in I \;) \;\; ,$$

(since O generates $(S,\vee)$) that is, iff

$$(\forall\, p,q \in S \,)(\; p(a\vee b)q \in I \quad\Leftrightarrow\quad pbq \in I \;) \ ,$$

that is, iff $a \vee b \approx_I\ b$.

Let $\alpha : (A^*,\cdot) \to (O,\cdot)$ be a surjective monoid homomorphism (for an appropriate finite set A). By Prop. 3 (i), $(O,\cdot,\leq)$ is isomorphic to the ordered syntactic monoid of $L = \alpha^{-1}(O \cap I)$.

By Prop. 1, the inclusion $O \subseteq S$ extends to a semiring morphism

$$\beta : (\mathsf{H}\,(O,\leq),\circ,\cup) \to (S,\cdot,\vee),\ (a_1,\ldots,a_k] \mapsto a_1 \vee \ldots \vee a_k \ .$$

We show that the kernel of β is $\sim_{(O\cap I]}$.

Really,

$$(a_1,\ldots,a_k]\ \sim_{(O\cap I]}\ (b_1,\ldots,b_l]$$

if and only if

$$(\forall\, p,q \in O \,)(\; pa_1q,\ldots,pa_kq \in O\cap I \quad\Leftrightarrow\quad pb_1q,\ldots,pb_lq \in O\cap I \;) \ ,$$

that is, iff $(\forall\, p,q \in O \,)(\; pa_1q,\ldots,pa_kq \in I \quad\Leftrightarrow\quad pb_1q,\ldots,pb_lq \in I \;)$,

that is, iff $(\forall\, p,q \in O \,)(\; p(a_1\vee\ldots\vee a_k)q \in I \quad\Leftrightarrow\quad p(b_1\vee\ldots\vee b_l)q \in I \;)$,

which is equivalent to $(\forall\, p,q \in S \,)(\; p(a_1\vee\ldots\vee a_k)q \in I \ \Leftrightarrow\ p(b_1\vee\ldots\vee b_l)q \in I \;)$

since every element of S is a join of elements of O.

Thus, by Prop. 6, $(S,\cdot,\vee)$ is isomorphic to the syntactic semiring of L. $\square$

Consider the language $L = \{u \in A^* \mid \mathsf{c}\,(u) \neq A\}$ over a finite set A.

The ordered syntactic monoid of L is isomorphic to $(2^A,\cup,\subseteq)$. Direct calculations yield that the syntactic semiring of L is isomorphic to $(\mathsf{H}\,(2^A,\subseteq),\circ,\cup)$. Denote these structures by $(O_n,\cdot,\leq)$ and by $(S_n,\cdot,\vee)$ in the case of an n-element alphabet A ($n \in \mathbb{N}$). Clearly, any $(O_n,\cdot,\leq)$ generates the pseudovariety J_1^-. Note that J_1^- is given by the identities $x^2 = x,\ xy = yx,\ 1 \leq x$ and that it is an atom in the lattice of pseudovarieties of finite ordered monoids. One can calculate that $(S_n,\cdot,\vee)$ generates the variety of idempotent semirings given by $x^{n+1} = x^n,\ xy = yx,\ 1 \leq\ x$.

Using Prop. 7, we can get languages with their ordered syntactic monoids outside of J_1^- having some of $(S_n,\cdot,\vee)$ as syntactic semiring. For instance, take K consisting of all words over $A = \{a,b,c\}$ having none of ab,ac,ba,bc,ca,cb,cc as a segment has the syntactic semiring isomorphic to $(S_2,\cdot,\vee)$.

We can comment the situation as follows: the ordered syntactic monoid and the syntactic semiring of a language are equationally independent.

6 A Construction of the Syntactic Semiring

We extend here a well-known construction of the syntactic monoid of a rational language to the case of the syntactic semiring.

Let a rational language L over a finite alphabet A be given. Classically one assigns to L its minimal automaton $\mathcal{A}$ using left quotients; namely
$Q = \{ \ u^{-1} \cdot L \mid u \in A^* \ \}$ is the (finite) set of states,
$a \in A$ acts on $u^{-1}L$ by $(u^{-1}L) \cdot a = a^{-1} \cdot (u^{-1}L)$,
$q_0 = L$ is the initial state and $u^{-1}L$ is a final state if and only if $u \in L$.
Now $(\mathsf{O}\,(L), \cdot)$ is the transition monoid of $\mathcal{A}$.
The order on $\mathsf{O}\,(L)$ is given by

$$f \leq \ g \ \text{ if and only if for every } \ q \in Q \text{ we have } \ q \cdot f \ \supseteq \ q \cdot g \ .$$

We extend the set of states to

$$\overline{Q} = \{ \ q_1 \cap \ldots \cap q_m \mid m \in \mathbb{N}, q_1, \ldots, q_m \in Q \ \} \ .$$

The action of a letter $a \in A$ is now given by

$$(q_1 \cap \ldots \cap q_m) \cdot a = q_1 \cdot a \ \cap \ldots \cap \ q_m \cdot a \ .$$

It can be extended to transitions induced by non-empty finite sets of words by

$$q \cdot \{u_1, ..., u_k\} = q \cdot u_1 \cap \ldots \cap q \cdot u_k \text{ for } q \in \overline{Q}, \ u_1, ..., u_k \in A^* \ .$$

It can be shown that this transition monoid with $\vee$ being the union is the syntactic semiring of the language L.

To make the computation finite, one considers instead of words from A^* their representatives in $\mathsf{O}\,(L)$. Moreover, it suffices to take only the actions of hereditary subsets of $(\mathsf{O}\,(L), \leq)$. At present stage we are far from a concrete implementation.

References

1. Almeida, J.; Finite Semigroups and Universal Algebra, World Scientific, 1994 612
2. Eilenberg, S.; Automata, Languages and Machines, Vol. B, Academic Press, 1976 616
3. Pin, J.-E.; Varieties of Formal Languages, Plenum, 1986 611
4. Pin, J.-E.; A variety theorem without complementation, Izvestiya VUZ Matematika **39** (1995) 80–90. English version: Russian Mathem. (Iz. VUZ) **39** (1995) 74–83 616
5. Pin, J.-E.; Syntactic semigroups, Chapter 10 in Handbook of Formal Languages, G. Rozenberg and A. Salomaa eds, Springer, 1997 611, 614
6. Polák, L.; A classification of rational languages by semilattice-ordered monoids, `http://www.math.muni.cz/~polak` 611, 615, 616, 618