Relational morphisms, transductions and operations on languages*

Jean-Éric Pin,
LITP, Université Paris VI and CNRS

Proceedings of the sixteenth Spring School of Theoretical Computer Science, Ramatuelle, France, May 1988

The aim of the article is to present two algebraic tools (the representable transductions and the relational morphisms) that have been used in the past decade to study operations on recognizable languages. This study reserves a few surprises. Indeed, both concepts were originally introduced for other purposes : representable transductions are a formalization of automata with output and have been mainly studied in connection with the theory of context-free languages, while relational morphisms were introduced by Tilson to solve some problems related to the wreath product decomposition of finite semigroups. But it turns out that relational morphisms are a very powerful tool in the study of recognizable languages and that transductions lead to some very nice problems on finite semigroups.

Eilenberg's variety theorem gives a one-to-one correspondence between varieties of semigroups and varieties of languages. Part of the results reviewed in this article show that, in certain cases, this correspondence can be extended to operations. That is, an operation on languages (such as concatenation, length-preserving morphism, etc.) is in correspondence with an operation on semigroups. It is therefore tempting to ask whether the most natural operations on languages (respectively semigroups) have a natural counterpart in terms of semigroups (respectively languages). This leads to a number of difficult problems, some of which are still unsolved.

1 Preliminaries.

The notions of recognizable and rational subsets are not limited to free semigroups. The *rational* subsets of a semigroup (monoid) S form the smallest class $\mathrm{Rat}(S)$ of subsets of S such that

(a) the empty set and every singleton $\{s\}$ belong to $\mathrm{Rat}(S)$,
(b) if X and Y are in $\mathrm{Rat}(S)$, then so are XY and $X \cup Y$,
(c) if X is in $\mathrm{Rat}(S)$, then so is X^+ (respectively X^*), the subsemigroup (submonoid) of S generated by X.

*This work was supported by "Programme de Recherches Coordonnées - Mathématique et Informatique".

A subset P of S is said to be *recognizable* if there exists a finite semigroup F, a semigroup morphism $\varphi : S \to F$ and a subset Q of F such that $P = Q\varphi^{-1}$. We shall denote by $\mathrm{Rec}(S)$ the class of recognizable subsets of S. It is well known that, if A is a finite alphabet, then $\mathrm{Rec}(A^+) = \mathrm{Rat}(A^+)$, where A^+ denotes the free semigroup over A.

Let S be a semigroup. Then $\mathcal{P}(S)$, the set of subsets of S, is a semiring under union as addition and under the multiplication defined, for every $X, Y \in \mathcal{P}(S)$, by

$$XY = \{xy \mid x \in X \text{ and } y \in Y\}.$$

Note that $\mathrm{Rat}(S)$, the set of all the rational subsets of S is a subsemiring of $\mathcal{P}(S)$. This structure of semiring will be useful in the study of representable transductions. Given a positive integer n, the set $\mathcal{P}(S)^{nxn}$ of n by n matrices with entries in $\mathcal{P}(S)$ is also a semiring, under the usual addition and multiplication of matrices over a semiring. Furthermore, every semigroup morphism $\varphi : S \to T$ extends into semiring morphisms $\varphi : \mathcal{P}(S) \to \mathcal{P}(T)$ and $\varphi : \mathcal{P}(S)^{n\times n} \to \mathcal{P}(T)^{n\times n}$.

Let S and T be two semigroups (or monoids). A *relation* $\tau : S \to T$ is a function from S into $\mathcal{P}(T)$. The graph of the relation τ is the subset of $S \times T$

$$\mathrm{graph}(\tau) = \{(s, t) \mid t \in s\tau\}.$$

The inverse of τ is the relation $\tau^{-1} : T \to S$ defined by $t\tau^{-1} = \{s \in S \mid t \in s\tau\}$. The relations τ and τ^{-1} can be extended to functions from $\mathcal{P}(S)$ into $\mathcal{P}(T)$ (respectively from $\mathcal{P}(T)$ into $\mathcal{P}(S)$) by setting

$$X\tau = \bigcup_{x\in X} x\tau \quad \text{and} \quad X\tau^{-1} = \{s \in S \mid s\tau \cap X \neq \emptyset\}.$$

A relation $\tau : S \to T$ is *injective* if, for every $s_1, s_2 \in S$, the condition $s_1\tau \cap s_2\tau \neq \emptyset$ implies $s_1 = s_2$.

Relations are too general for our purpose — the semigroup structure is irrelevant so far — so we shall only consider two useful cases : the relational morphisms and the *representable* transductions.

2 Relational morphisms.

A relational morphism $\tau : S \to T$ is a relation satisfying the following properties:

(1) for every $s \in S$, $s\tau$ is non-empty,

(2) for every $s, t \in S$, $(s\tau)(t\tau) \subset (st)\tau$.

If S and T are monoids, then we require a third condition :

(3) $1 \in 1\tau$.

Example 2.1 Let BA_2 be the semigroup with zero presented over the alphabet $\{a, b\}$ by the relations $a^2 = b^2 = 0$, $aba = a$ and $bab = b$. Then $BA_2 = \{a, b, ab, ba, 0\}$, and the idempotents are ab, ba and 0. This semigroup is called the Brandt Aperiodic semigroup of size 2×2 (this explains the notation BA_2) and is also the transformation semigroup of the automaton represented below.

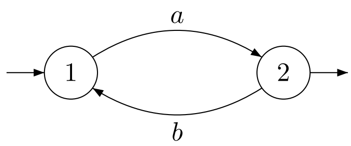


Figure 1:

The $\mathcal{D}$-class structure of BA_2 is

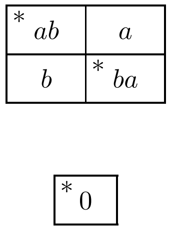


Figure 2:

Let $\tau : BA_2 \to \mathbb{Z}/2\mathbb{Z} = \{-1, 1\}$ be the relation defined by $a\tau = \{-1\}$, $b\tau = \{-1\}$, $(ab)\tau = \{1\}$, $(ba)\tau = \{1\}$ and $0\tau = \{1, -1\}$. Then τ is a relational morphism.

Of course, any morphism (in the usual sense) is also a relational morphism. Furthermore, if $\varphi : S \to T$ is a surjective morphism, then the relation $\tau = \varphi^{-1} : T \to S$ is a relational morphism. Some of the basic properties of morphisms still hold for relational morphisms.

Proposition 2.1

(1) *The composition of two relational morphisms is a relational morphism.*

(2) *Let $\tau : S \to T$ be a relational morphism. If S' is a subsemigroup of S, then $S'\tau$ is a subsemigroup of T. If T' is a subsemigroup of T, then $T'\tau^{-1}$ is a subsemigroup of S.*

A detailed proof of this proposition (and of Propositions 2.2 to 2.5) can be found in [9, p.67] and are not reproduced in this survey.

It is not too difficult to see that if τ is a relational morphism from S into T, then its graph $R \subset S \times T$ is a subsemigroup (resp. submonoid) of $S \times T$, and the projections $S \times T \to S$ and $S \times T \to T$ induce morphisms $\alpha : R \to S$ and $\beta : R \to T$ such that

(i) α is a surjective morphism

(ii) $\tau = \alpha^{-1}\beta$.

$$\begin{tikzcd}
 & R \arrow[dl, "\alpha"'] \arrow[dr, "\beta"] & \\
S \arrow[rr, "\tau = \alpha^{-1}\beta"] & & T
\end{tikzcd}$$

Figure 3:

The factorization represented in Figure 3 is called the canonical factorization of τ. Therefore

Proposition 2.2 *A relation τ is a relational morphism if and only if there exist two morphisms α and β such that $\tau = \alpha^{-1}\beta$.*

Thus, if one has trouble to think in terms of relational morphisms, one can always go back to usual morphisms. In fact, if $\tau = \alpha^{-1}\beta$ is the canonical factorization of τ, a number of properties of b are inherited by τ, and it is usually a good idea to prove a given result on b, and then try to generalize the proof for τ. For instance,

Proposition 2.3 *Let $S \xrightarrow{\alpha^{-1}} R \xrightarrow{\beta} T$ be the canonical factorization of a relational morphism $\tau : S \to T$. Then*
(1) *τ is injective if and only if β is injective,*
(2) *τ is surjective if and only if β is surjective.*

We arrive to an essential definition about relational morphisms, which is similar in its form to the definition of a continuous function. Given a variety of finite semigroups $\mathbf{V}$, we say that a morphism (respectively a relational morphism) is a $\mathbf{V}$-morphism (respectively a relational $\mathbf{V}$-morphism) if, for every subsemigroup T' of T, $T' \in \mathbf{V}$ implies $T'\tau^{-1} \in \mathbf{V}$. If $\mathbf{V}$ is the variety A of aperiodic semigroups, a (relational) $\mathbf{V}$-morphism is simply called an aperiodic (relational) morphism. Similarly, if $\mathbf{V}$ is the variety **LI** of locally trivial semigroups, a (relational) $\mathbf{V}$-morphism is called a locally trivial (relational) morphism. The following proposition was to be expected :

Proposition 2.4 *Let $S \xrightarrow{\alpha^{-1}} R \xrightarrow{\beta} T$ be the canonical factorization of a relational morphism $\tau : S \to T$. Then τ is a $\mathbf{V}$-morphism if and only if β is a $\mathbf{V}$-morphism.*

For certain varieties of semigroups $\mathbf{V}$ (but not for all of them!) there is a simple characterization of the $\mathbf{V}$-morphisms. This is the case in particular for the varieties of aperiodic (respectively locally trivial) semigroups.

Proposition 2.5 *Let $\tau : S \to T$ be a relational morphism. Then τ is an aperiodic (respectively locally trivial) morphism if and only if, for every idempotent $e \in T$, the semigroup $e\tau^{-1}$ is aperiodic (respectively locally trivial).*

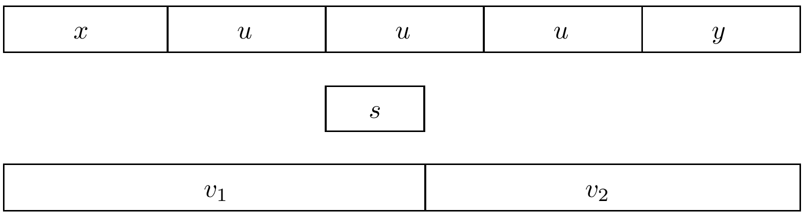
Examples of relational morphisms abound in semigroup theory, but we have selected here some examples that come from language theory. Let L_1 and L_2 be two rational languages of A^*, and let $L = L_1L_2$. We say that the product L_1L_2 is *unambiguous* if every word u of L has a unique factorization as $u = u_1u_2$, where $u_1 \in L_1$ and $u_2 \in L_2$. Let $\eta : A^* \to M$, $\eta_1 : A^* \to M_1$ and $\eta_2 : A^* \to M_2$ be the syntactic morphisms of L, L_1 and L_2 respectively. Let $\eta_1 \times \eta_2 : A^* \to M_1 \times M_2$ be the monoid morphism defined by $u(\eta_1 \times \eta_2) = (u\eta_1, u\eta_2)$. Finally, we put $\tau = \eta^{-1}(\eta_1 \times \eta_2)$. Thus τ is a relational morphism, represented in the following diagram.

$$\begin{tikzcd}
 & A^* \arrow[dl, "\eta"'] \arrow[dr, "\eta_1 \times \eta_2"] & \\
M \arrow[rr, "\tau"] & & M_1 \times M_2
\end{tikzcd}$$

Figure 4:

Theorem 2.6 [17] *The relational morphism $\tau : M \to M_1 \times M_2$ is aperiodic. Furthermore, if the product L_1L_2 is unambiguous, then τ is locally trivial.*

Proof. We only give the proof of the first part of the theorem. Denote by $\sim_L$, $\sim_{L_1}$ and $\sim_{L_2}$ the syntactic congruences of L, L_1 and L_2 respectively. Let e be an idempotent of $M_1 \times M_2$. Then $e = (e_1, e_2)$ where e_1 and e_2 are idempotent in M_1 and M_2, respectively. Let $s \in e\tau^{-1}$. Then there exists a word u such that $u\eta = s$, $u\eta_1 = e_1$ and $u\eta_2 = e_2$. Then by definition, $u \sim_{L_1} u^2$ and $u \sim_{L_2} u^2$. We claim that $u^3 \sim_L u^4$. Indeed, suppose that $xu^3y \in L$, for some $x, y \in A^*$. Then $xu^3y = v_1v_2$ for some $v_1 \in L_1$ and $v_2 \in L_2$, and either xu is a prefix of v_1, or uy is a suffix of v_2.



Suppose, for instance, that $v_1 = xus$, for some $s \in A^*$ such that $sv_2 = u^2y$ (the other case is dual). Then since $u \sim_{L_1} u^2$, we have $xu^2s \in L_1$ and $xu^2sv_2 = xu^4y \in L$. Conversely, if $xu^4y \in L$, then $xu^4y = v_1v_2$ for some $v_1 \in L_1$ and $v_2 \in L_2$. Then either v_1 or v_2 contains an occurrence of u^2, and the argument above shows that $xu^3y \in L$. Thus $xu^3y \in L$ if and only if $xu^4y \in L$, proving the claim. It follows that $s^3 = u^3\eta = u^4\eta = s^4$. Therefore, $e\tau^{-1}$ is an aperiodic semigroup, and τ is aperiodic. $\square$

In particular, if M_1 and M_2 are aperiodic, M is also aperiodic. This theorem motivates the following definition. Given two varieties of finite monoids $\mathbf{V}$ and $\mathbf{W}$, we denote by $\mathbf{V}^{-1}\mathbf{W}$ the class of all monoids M such that there exists a $\mathbf{V}$-relational morphism from M into a monoid N of $\mathbf{W}$. This class $\mathbf{V}^{-1}\mathbf{W}$ is in fact a variety, called the Malcev product of $\mathbf{V}$ and $\mathbf{W}$. Now, Theorem 2.6 can be reformulated as follows: if M_1 and M_2 belong to a variety of finite monoids $\mathbf{V}$, then M belongs to the variety $\mathbf{A}^{-1}\mathbf{W}$. Furthermore, if the product L_1L_2 is unambiguous, then M belongs to $\mathbf{LI}^{-1}\mathbf{W}$. The converse of this result is true, and can be informally summarized by saying that “concatenation corresponds to aperiodic morphisms” and that “non-ambiguous concatenation corresponds to locally trivial morphisms”. More precisely, let $\mathbf{V}$ be a variety of monoids and let $\mathcal{V}$ be the corresponding variety of languages. Let $\mathcal{V}'$ (respectively $\mathcal{V}''$) be the smallest variety of languages such that, for every alphabet A,

(a) $A^*\mathcal{V}'$ (respectively $A^*\mathcal{V}''$) contains $A^*\mathcal{V}$,

(b) $A^*\mathcal{V}'$ (respectively $A^*\mathcal{V}''$) is closed under concatenation (respectively non-ambiguous concatenation).

Then we have the following theorem.

Theorem 2.7 [16, 11] *The variety of monoids corresponding to $\mathcal{V}'$ (respectively $\mathcal{V}$”) is $\mathbf{A}^{-1}\mathbf{V}$ (respectively $\mathbf{LI}^{-1}\mathbf{V}$).*

We say that a variety of languages $\mathbf{V}$ is closed under (unambiguous) concatenation product if, for every alphabet A, the conditions $L_1, L_2 \in A^*\mathbf{V}$ imply $L_1L_2 \in A^*\mathbf{V}$ (if the product is unambiguous).

Corollary 2.8

(1) *A variety of languages $\mathbf{V}$ is closed under concatenation product if and only if $\mathbf{A}^{-1}\mathbf{V} = \mathbf{V}$.*

(2) *A variety of languages $\mathbf{V}$ is closed under non-ambiguous concatenation product if and only if $\mathbf{LI}^{-1}\mathbf{V} = \mathbf{V}$.*

The star operation gives another example of relational morphisms in language theory. Let L be a recognizable language of A^*. A submonoid L^* of A^* is pure if, for every $n > 0$, and every $u \in A^*$, $u^n \in L$ implies $u \in L$. Let $\eta : A^* \to M$ and $\varphi : A^* \to N$ be the syntactic morphisms of L and L^* respectively. Then $\tau = \varphi^{-1}\eta$ is a relational morphism from N into M, represented in the following diagram.

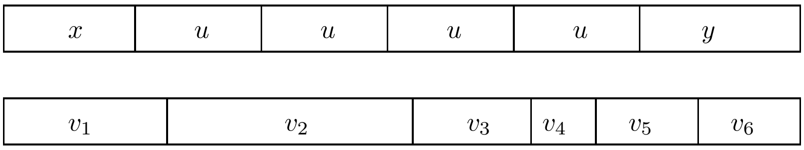
$$\begin{tikzcd}
 & A^* \arrow[dl, "\eta"'] \arrow[dr, "\varphi"] & \\
M \arrow[rr, "\tau"] & & N
\end{tikzcd}$$

Figure 5:

Theorem 2.9 [17] *If L^* is a pure submonoid of A^*, then τ is an aperiodic relational morphism.*

Proof. Denote by $\sim_L$ and $\sim_{L^*}$ the syntactic congruences of L and L^* respectively. Let e be an idempotent of M and let $s \in e\tau^{-1}$. Then there exists a word $u \in A^*$ such that $u\varphi = s$ and $u\eta = e$. In particular, $u^2 \sim_L u$. We claim that, for every $n > |u|$, and every $x, y \in A^*$, $xu^ny \in L^*$ implies $xu^{n+1}y \in L^*$. Indeed, suppose that $xu^ny = v_1v_2 \cdots v_k$ where $k \geq 0$, and $v_1, v_2, \cdots, v_k \in L$. The integers $|v_1v_2 \cdots v_i|$, for $1 \leq i \leq k$, are called the scansions of the factorization $v_1v_2 \cdots v_k$. We consider two cases.

(a) There exists an index i such that v_i contains an occurrence of u (as v_2 in the diagram below).



Then $v_i = s_i u t_i$, and since $u^2 \sim_L u$, $s_i u t_i \in L$. Therefore $xu^{n+1}y \in L^*$.

(b) Every occurrence of u contains a scansion of the factorization $v_1 v_2 \cdots v_k$. Therefore, there exists a sequence $j_0 < j_1 < \cdots < j_{n-1}$ such that, for $0 \leq i \leq n-1$, we have $xu^i p_i = v_1 v_2 \cdots v_{j_i}$ for some prefix p_i of u (different from u, but possibly empty). Now since $n > |u|$, there exist two indices $i < i'$ such that $p_i = p_{i'}$. Let $u = p_i s_i$. This yields

$$\begin{aligned} x_1 &= v_1 v_2 \cdots v_{j_i} = xu^i pi, \\ x_2 &= v_{i+1} \cdots v_{i'} = s_i u^{i'-i-1} p_i = (s_i p_i)^{i'-i} \in L^*, \\ x_3 &= v_{i'} \cdots v_k = s_i u^{n-i'-1} y, \end{aligned}$$

and since L^* is pure, $s_i p_i \in L^*$. Therefore $(s_i p_i)^{i'-i+1} \in L^*$, and thus $xu^{n+1}y = x_1 (s_i p_i)^{i'-i+1} x_3 \in L^*$.

This proves the claim. But since L^* is recognizable, there exist two integers $m \geq 0$ and $p > 0$ such that, for every $n \geq m$, $u^n \sim_{L^*} u^{n+p}$. If we take $n > \max(m, |u|)$, then $xu^n y \in L^*$ implies $xu^{n+1}y \in L^*$ by the claim. Conversely, $xu^{n+1}y \in L^*$ implies $xu^{n+p}y \in L^*$ by the claim, and since $u^n \sim_{L^*} u^{n+p}$, $xu^{n+p}y \in L^*$ implies $xu^n y \in L^*$. Therefore $u^n \sim_{L^*} u^{n+1}$ and hence $s^n = u^n\varphi = u^{n+1}\varphi = s^{n+1}$. Thus $e\tau^{-1}$ is aperiodic, and τ is an aperiodic relational morphism. □

Corollary 2.10 [17]. *If L is star-free, and if L^* is a pure submonoid of A^*, then L^* is star-free.*

Proof. We keep the previous notations. If L is star-free, then M is aperiodic by Schützenberger's theorem. Now, since L^* is pure, τ is an aperiodic relational morphism, so that $N = M\tau^{-1}$ is aperiodic. Therefore, L^* is star-free by Schützenberger's theorem. □

Corollary 2.11 *Let $\mathcal{V}$ be a variety closed under concatenation product. Then $\mathcal{V}$ is closed under the operation $L \to L^*$ when L^* is pure.*

Proof. Let $\mathbf{V}$ be the variety of monoids corresponding to $\mathcal{V}$. Then, by Theorem 2.8, $\mathbf{A}^{-1}\mathbf{V} = \mathbf{V}$. Therefore, by Theorem 2.9, $\mathbf{V}$ is closed under the operation $L \to L^*$ when L^* is pure. □

H. Straubing asked whether the converse of Corollary 2.11 was true, but this problem is still open. The next result is a very special case of this problem.

Theorem 2.12 [8]. *The star-free languages form the smallest variety closed under the operation $L \to L^*$ when L^* is pure.*

To conclude this section, we shall give an application of relational morphisms to sets recognized by finite groups. Rational sets are preserved under monoid morphisms, but recognizable sets are not in general. However, we have

Theorem 2.13 *Let $\pi : M \to H$ be a surjective monoid morphism from a monoid M onto a group H. If L is a subset of M recognized by a finite group G, then $L\pi$ is a recognizable subset of H recognized by a quotient of G.*

We need the following result on relational morphisms.

Lemma 2.14 *Let G be a finite group, H an arbitrary group, and $\tau : G \to H$ be a surjective relational morphism. Then $G' = 1\tau^{-1}$ is a normal subgroup of G, $H' = 1\tau$ is a normal subgroup of H, and $G/G' = H/H'$. Furthermore, for every $g \in G$, $g\tau$ is a coset of H'.*

Proof. We first claim that for every $g \in G$, $h \in g\tau$ implies $h^{-1} \in g^{-1}\tau$. Indeed, since τ is surjective, there exists $x \in G$ such that $h^{-1} \in x\tau$. Let n be the order of G. Then $h^{-1} = (h^{-1})^n h^{n-1} \in (x\tau)^n (g\tau)^{n-1} \subset (x^n g^{n-1})\tau = g^{-1}\tau$ since $x^n = g^n = 1$.

In particular, $H' = 1\tau$ is a subgroup of H. Furthermore, if $h \in g\tau$ and $x \in 1\tau$, then by the claim, $hxh^{-1} \in (g\tau)(1\tau)(g^{-1}\tau) \subset (gg^{-1})\tau = 1\tau$.Thus H' is normal in H. Furthermore, the following property holds :

$$\text{(1) for every } h \in g\tau,\ g\tau = H'h.$$

Indeed, $H'h \subset (1\tau)(g\tau) \subset g\tau$ and, on the other hand, $(g\tau)h^{-1} \subset (g\tau)(g^{-1}\tau) \subset 1\tau$. Thus τ induces a group morphism from G onto H/H', the kernel of which is the group

$$N = \{g \in G \mid g\tau = 1\tau\}.$$

We claim that $N = G' = \{g \in G \mid 1 \in g\tau\}$. Indeed, since $1 \in 1\tau$, N is contained in G'. Conversely, if $g \in G'$, then $1 \in g\tau$, whence $g\tau = (1\tau).1 = 1\tau$ by (1). Thus $N = G'$, G' is normal in G and $G/G' = H/H'$. □

We now prove Theorem 2.13. Let $\varphi : M \to G$ be the monoid morphism recognizing L, and let $P \subset G$ be such that $P\varphi^{-1} = L$. We may assume that φ is surjective (otherwise it suffices to replace G by $M\varphi$, since a submonoid of a finite group is a group). Then $\tau = \varphi^{-1}\pi : G \to H$ is a surjective relational morphism and we may apply Lemma 2.14. In particular $L\pi = P\tau$ is a finite union of cosets of $H' = 1\tau$. Thus L is recognized by $H/H' = G/G'$. □

A recognizable language is called a *group language* if it is recognized by a finite group.

Corollary 2.15 *Let $\pi : A^* \to H$ be a surjective monoid morphism from A^* onto a group H, and let S be a subset of H. Then*

(a) *S is rational if and only if there exists a rational language $R \subset A^*$ such that $S = R\pi$,*

(b) *S is recognizable if and only if there exists a group language $L \subset A^*$ such that $S = L\pi$.*

3 Representable transductions.

Besides relational morphisms, one mainly considers relations of the form $\tau : A^* \to M$ where A^* is a finitely generated free monoid and M is a monoid. Relations of this form are called *transductions*. The definition of a representable transduction is a little more abstract. Informally, a transduction $\tau : A^* \to M$ admits a representation $\mu : A^* \to \mathcal{P}(M)^{n\times n}$ if

(a) μ is a morphism: for every $u_1, u_2 \in A^*$, $(u_1u_2)\mu = (u_1\mu)(u_2\mu)$,

(b) for every $u \in A^*$, $u\tau$ can be expressed from the $u\mu_{i,j}$ $(1 \le i, j \le n)$.

The following list of examples should help the reader to understand the meaning of the expression "expressed from the $u\mu_{i,j}$". Here, $A = \{a, b\}$ and $\tau_1, \cdots, \tau_4$ are transductions from A^* into A^*:

(1) $u\tau_1 = u\mu_{1,1} + u\mu_{1,2} + u\mu_{2,1}$,

(2) $u\tau_2 = a^*(u\mu_{1,1})aba + (u\mu_{1,2})A^*babA^* + \{a, ba\}^*(u\mu_{2,2})$,

(3) $u\tau_3 = bab^*(u\mu_{1,1})3(a + ba)(u\mu_{1,2})b + (u\mu_{2,1})\{ua^{|u|} \mid u \in A^*\}$,

(4) $u\tau_4 = \sum_{n\ge 0} a^n(u\mu_{1,2})b^{2n}(u\mu_{1,1})a^{n^2}$.

In the two first examples, $u\tau$ is a fixed linear combination of the $u\mu_{i,j}$. We say in this case that τ admits a *linear representation*. Formally, a linear representation of a transduction $\tau : A^* \to M$ is a triple (X, μ, Y), where $X, Y \in \mathcal{P}(M)^n$, such that, for every $u \in A^*$,

$$(*) \qquad u\tau = \sum_{1\le i,j\le n} X_i(u\mu_{i,j})Y_j.$$

This linear representation is rational if μ is a map from A^* into $\mathrm{Rat}(M)^{n\times n}$ and if $X, Y \in \mathrm{Rat}(M)^n$. The general definition of a representation, as given in [12, 13], replaces linear expressions by polynomials, as in (3), or even by power series, as in (4), but we shall not use it in this survey. Representable transductions cover a great variety of situations, and are related to recognizable sets by the following proposition.

Proposition 3.1 [12, 13] *Let $\tau : A^* \to M$ be a representable transduction. Then for every recognizable subset P of M, the language $P\tau^{-1}$ is recognizable.*

More precisely, suppose that a transduction $\tau : A^* \to M$ admits a representation $\mu : A^* \to \mathcal{P}(M)^{n\times n}$. Let $\varphi : M \to N$ be a monoid morphism, and let P be a subset of M recognized by φ. Then φ induces a semigroup morphism $\varphi : \mathcal{P}(M)^{n\times n} \to \mathcal{P}(N)^{n\times n}$, and we have

Proposition 3.2 [12, 13] *The language $P\tau^{-1}$ is recognized by the monoid $A^*\mu\varphi$.*

The precise description of this monoid $A^*\mu\varphi$ is the key to understand several constructions, that otherwise would seem awkward. Some examples are given below.

A transduction which admits a linear rational representation is called a *rational transduction*. It is not difficult to see from the definition that every monoid morphism is a rational transduction. Similarly, for every rational subset R of A^*, the transduction $Id_R : A^* \to A^*$, defined by

$$u\tau = \begin{cases} u & \text{if } u \in R \\ \text{undefined} & \text{otherwise} \end{cases}$$

is rational. Rational transductions have been extensively studied in connection with the theory of context-free languages [1] and admit various characterizations.

Theorem 3.3 *Let $\tau : A^* \to M$ be a transduction. Then the following conditions are equivalent :*
(1) *τ is a rational transduction,*
(2) *the graph of τ is a rational subset of $A^* \times M$,*
(3) *there exist an alphabet C, two morphisms $\varphi : C^* \to A^*$ and $\psi : C^* \to M$, and a rational set $R \subset C^*$ such that $\tau = \varphi^{-1} Id_R \psi$.*

In particular, if $\tau : A^* \to B^*$ is a rational transduction, then the transduction $\tau^{-1} : B^* \to A^*$ is also rational. Every rational transduction is representable, but the converse is not true (see example (b) below). Condition (3) of Theorem 3.3 is known as Nivat's theorem, and should be compared with the canonical factorization of relational morphisms. The transductions considered in the examples (a), (c), (d), (e), (f) below are rational. Other examples of rational transductions can be found in [1].

We now give some examples of application of Proposition 3.2.

(a) Inverse morphisms.

Let $\varphi : A^* \to M$ be a morphism. Then $\mu = \varphi$ is a linear representation for φ, and the construction above shows that if L is a subset of B^* recognized by a monoid N, then $L\varphi^{-1}$ is also recognized by N — a well-known result.

(b) Inverse substitutions.

Recall that a substitution s from A^* into M is a monoid morphism $\sigma : A^* \to \mathcal{P}(M)$. Therefore $\mu = \sigma$ is a linear representation for σ, and the construction above shows that if L is a subset of B^* recognized by a monoid N, then $L\sigma^{-1}$ is recognized by $\mathcal{P}(N)$.

(c) Length preserving morphisms.

Let $\varphi : A^* \to B^*$ be a length preserving morphism. Then the transduction $\varphi^{-1} : B^* \to A^*$ is a substitution. Thus, by (b), if L be a subset of B^* recognized by a monoid N, then $L\varphi$ is recognized by $\mathcal{P}(N)$. This result can be extended to varieties as follows. Given a variety of monoids $\mathbf{V}$, denote by $\mathbf{PV}$ the variety of monoids generated by all monoids of the form $\mathcal{P}(M)$, where $M \in \mathbf{V}$. Given a variety of languages $\mathbf{V}$, let $\mathbf{V}'$ be the smallest variety of languages such that $L\varphi \in B^*\mathbf{V}'$ for every $L \in A^*\mathbf{V}$ and every length-preserving morphism $\varphi : A^* \to B^*$.

Theorem 3.4 [14, 15] *Let $\mathbf{V}$ be a variety of monoids, and let $\mathbf{V}$ be the corresponding variety of languages. Then $\mathbf{V}'$ is the variety of languages corresponding to $\mathbf{PV}$.*

A similar result was obtained for inverse substitutions [14]. Thus, roughly speaking, power varieties correspond to length preserving morphisms or to inverse substitutions. Theorem 3.4 has been the starting point of the study of the operation $\mathbf{V} \to \mathbf{PV}$, and more generally, of the classification of the varieties of the type $\mathbf{PV}$. Although a number of interesting results have been obtained, this classification is far from being complete. See the survey [10] and the recent (or forthcoming) articles by J. Almeida for more information.

(d) Shuffle product.

Recall that the *shuffle* of two words u and v is the set

$$u \circ v = \{u_1 v_1 u_2 v_2 \cdots u_n v_n \mid u_1 u_2 \cdots u_n = u \text{ and } v_1 v_2 \cdots v_n = v\}.$$

More generally, the shuffle of two languages L_1 and L_2 is the language

$$L_1 \circ L_2 = \bigcup_{u_1 \in L_1, u_2 \in L_2} u_1 \circ u_2.$$

Now, if $\tau : A^* \to A^* \times A^*$ is the transduction defined by

$$u\tau = \{(u_1, u_2) \in A^* \times A^* \mid u \in u_1 \circ u_2\},$$

then $(L_1 \times L_2)\tau^{-1} = L_1 \circ L_2$, and τ is a substitution, defined, for every $a \in A$, by $a\tau = \{(a, 1), (1, a)\}$. Thus, by (b), if the languages L_1 and L_2 are recognized by the monoids M_1 and M_2, respectively, then $L_1 \circ L_2$ is recognized by $\mathcal{P}(M_1 \times M_2)$.

However, contrary to the case of length-preserving morphisms or inverse substitutions, it is not known whether this result can be extended to varieties. For instance, the following problem is still open.

Conjecture *If a variety of languages is closed under shuffle and contains a non-commutative language*[1]*, then it contains all rational languages.*

(e) Concatenation product.

Let $\tau : A^* \to A^* \times A^* \times \cdots \times A^*$ be the transduction defined by

$$u\tau = \{(u_1, u_2, \cdots, u_n) \in A^* \times A^* \times \cdots \times A^* \mid u_1 u_2 \cdots u_n = u\}.$$

Then, for every $L_1, \ldots, L_n \subset A^*$, $(L_1 \times \cdots \times L_n)\tau^{-1} = L_1 \cdots L_n$. Furthermore, τ is a representable transduction since, for every $u \in A^*$, $u\tau = u\mu_{1,n}$ where $\mu : A^* \to P(A^*)^{n \times n}$ is the morphism defined by

$$u\mu_{i,j} = \begin{cases} 0 & \text{if } i > j \\ \{(1, \cdots, 1, u_i, u_{i+1}, \cdots, u_j, 1, \cdots, 1) \mid u_i u_{i+1} \cdots u_j = u\} & \text{if } i \leq j. \end{cases}$$

Now, if the languages $L_1, \ldots, L_n$ of A^* are recognized by the morphisms $\varphi_1 : A^* \to M_1$, $\ldots$, $\varphi_n : A^* \to M_n$, respectively, the product $L_1 \cdots L_n$ is recognized by the monoid $A^*\mu\varphi$, where $\varphi = \varphi_1 \times \cdots \times \varphi_n$. This monoid is called the Schützenberger product of the monoids $M_1, \cdots, M_n$, and is denoted $\diamondsuit(M_1, \cdots, M_n)$.

Formally, let K be the semiring $\mathcal{P}(M_1 \times \cdots \times M_n)$ and let $M_n(K)$ be the semiring of matrices of size n by n over K. Then $\diamondsuit(M_1, \cdots, M_n)$ is the multiplicative submonoid of $K_{n \times n}$ consisting of all matrices p satisfying the following conditions.

[1] A language L is non-commutative if there exist two words u and v such that $uv \in L$ and $vu \notin L$

(1) For $i > j$, $p_{i,j} = 0$,
(2) for $i = j$, $p_{i,i} = \{(1, \dots, 1, mi, 1, \dots, 1)\}$ for some $m_i \in M_i$,
(3) for $i < j$, $p_{i,j} \subset \{1\} \times \cdots \times \{1\} \times M_i \times \cdots \times M_j \times \{1\} \times \cdots \times \{1\}$.

Condition (1) says that p is upper-triangular, condition (2) enables us to identify $p_{i,i}$ with an element of M_i, and condition (3) states that $p_{i,j}$ can be identified to a subset of $M_i \times \cdots \times M_j$. Finally, we have the following result, due to Schützenberger for $n = 2$ and to Straubing for the general case.

Theorem 3.5 *Let $L_1, \dots, L_n$ be languages recognized by monoids $M_1, \dots, M_n$. Then the language $L_1 \cdots L_n$ is recognized by the monoid $\diamondsuit(M_1, \cdots, M_n)$.*

The converse of Theorem 3.5 is not true, but a very similar statement admits a partial converse (Theorem 3.7 below, due to Reutenauer for $n = 1$ and to the author in the general case).

Theorem 3.6 *Let $L_0, \dots, L_n$ be languages recognized by monoids $M_0, \dots, M_n$ and $a_1, a_2, \cdots, a_n$ be letters. Then the language $L_0 a_1 L_1 a_2 \cdots a_n L_n$ is recognized by the monoid $\diamondsuit(M_0, \cdots, M_n)$.*

Theorem 3.7 *If a language L of A^* is recognized by $\diamondsuit(M_0, \cdots, M_n)$, then L is a boolean combination of languages of the form $L_{i_0} a_1 L_{i_1} a_2 \cdots a_r L_{i_r}$ where $0 \le i_0 < i_1 < \cdots < i_r \le n$, and, for $1 \le k \le r$, $a_k \in A$ and L_{i_k} is a language recognized by M_{i_k}.*

We refer the reader to the article of P. Weil in this volume for other applications of the Schützenberger product.

(f) Sequential functions.

A *sequential function* $\sigma : A^* \to B^*$ is a function from A^* to B^* whose behaviour is described by a machine called a "sequential transducer". Formally, a *sequential transducer* is a set $\mathcal{T} = (Q, A, B, q_0, ., *)$ where $(Q, A, .)$ is a finite automaton, q_0 is an element of Q called the initial state and $*$ is an output function, i.e. a function $(q, a) \to q * a$ from $Q \times A$ into B^* which can be extended to a function

$$Q \times A^* \to B^*$$

by putting $q * 1 = 1$ and, for $u \in A^*$ and $a \in A$, $q * ua = (q * u)((q.u) * a)$.

A function $\sigma : A^* \to B^*$ is called *sequential* if there exists a sequential transducer $\mathcal{T} = (Q, A, B, q_0, ., *)$ such that, for every word $u \in A^*$, $u\sigma = q_0 * u$. A sequential function is a rational transduction, which admits the linear representation $\mu : A^* \to \mathcal{P}(B^*)^{Q \times Q}$ defined, for every $a \in A$, by

$$a\mu_{q,q'} = \begin{cases} q * a & \text{if } q.a = q' \\ \emptyset & \text{otherwise.} \end{cases}$$

Indeed, $u\sigma = \sum_{q \in Q} u\mu_{q_0,q}$, and thus, if a language L of B^* is recognized by a morphism $\varphi : B^* \to M$, the language $L\sigma^{-1}$ is recognized by the monoid $A^*\mu\varphi$. Now, if $M(\sigma)$ is the transition monoid of the automaton $(Q, A, .)$, the monoid $A^*\mu\varphi$ is a submonoid of the "wreath product" $M \circ M(\sigma)$. We recall the definition of this important construction.

Let M and N be two monoids. We write M additively (although M is not assumed to be commutative) and N multiplicatively. The wreath product $M \circ N$ is the monoid defined on the set $M^N \times N$ by the multiplication given by the following formula (where f_1, f_2 are applications from M into N, and n_1, n_2 are elements of N)

$$(f_1, n_1)(f_2, n_2) = (f, n_1 n_2)$$

where f is the application from M into N defined, for all $n \in N$, by $nf = nf_1 + (nn_1)f_2$.

Therefore, we can state

Theorem 3.8 *Let $\sigma : A^* \to B^*$ be a sequential function realized by a sequential transducer $\mathcal{T} = (Q, A, B, q_0, ., *)$ and let $M(\sigma)$ be the transition monoid of the automaton $(Q, A, .)$. If a language L of B^* is recognized by a monoid M, then the language $L\sigma^{-1}$ is recognized by the monoid $M \circ M(\sigma)$.*

Again, this statement admits a partial converse, called the "wreath product principle", and stated for the first time by Straubing. Let M and N be two monoids, and let $\eta : A^* \to M \circ N$ be a morphism. Denote by $\pi : M \circ N \to N$ the morphism defined by $(f, n)\pi = n$ and let $\varphi = \eta\pi : A^* \to N$. Let $B = N \times A$ and $\sigma : A^* \to B^*$ be the sequential function defined by $(a_1 \cdots a_n)\sigma = (1, a_1)(a_1\varphi, a_2) \cdots ((a_1 \cdots a_{n-1})\varphi, a_n)$. Then we have:

Theorem 3.9 *If a language L is recognized by $\eta : A^* \to M \circ N$, then L is a finite boolean combination of languages of the form $X \cap Y\sigma^{-1}$ where $Y \subset B^*$ is recognized by M and where $X \subset A^*$ is recognized by N.*

Despite its technical appearance, the wreath product principle, with its variants, is one of the most useful tools of the theory of finite automata.

4 Transductions and decidability problems on semigroups.

Transductions can lead to some very difficult problems of semigroup theory. As an example, we would like to mention some decidability problems recently solved by Hashiguchi. Let $\mathcal{L} = \{L_1, \cdots, L_n\}$ be a finite set of rational languages. Denote by $\mathrm{Rat}(\mathcal{L})$ the smallest set of languages containing L and closed under finite union, product and star operation. Similarly, denote by $\mathrm{Pol}(\mathcal{L})$ the polynomial closure of $\mathcal{L}$, that is, the smallest set of languages containing L and closed under finite union and product. Therefore, a language L belongs to $\mathrm{Pol}(\mathcal{L})$ if and only if it is a finite union of products of the form $L_{i_1} L_{i_2} \cdots L_{i_k}$, where $L_{i_1}, L_{i_2}, \cdots L_{i_k} \in \mathcal{L}$. Hashiguchi studied the following decidability problems.

Problem 1 *Given $\mathcal{L}$ and a rational language L, is it decidable whether $L \in$* $\mathrm{Rat}(\mathcal{L})$*?*

Problem 2 *Given $\mathcal{L}$ and a rational language L, is it decidable whether $L \in Pol(\mathcal{L})$?*

Both problems are decidable, but the first problem is easy to solve, while the second one is much more difficult. We introduce a second alphabet $B = \{1, \cdots, n\}$ and the substitution $\sigma : B^* \to A^*$ defined, for $1 \leq i \leq n$, by $i\sigma = L_i$. Clearly, $L \in \mathrm{Rat}(\mathcal{L})$ (respectively $\mathrm{Pol}(\mathcal{L})$) if and only if there exists a rational (respectively finite) subset R of B^* such that $R\sigma = L$. We first examine the language $R(L) = \{u \in B^* \mid u\sigma \subset L\}$. This language can be computed effectively:

Lemma 4.1 *For every rational language L, $R(L) = B^* \setminus (A^* \setminus L)\sigma^{-1}$. In particular, $R(L)$ is a rational subset of B^*.*

Proof. We have, by definition,

$$(A^*\ L)\sigma^{-1} = \{i_1 \cdots i_k \in B^* \mid L_{i_1} L_{i_2} \cdots L_{i_k} \cap (A^* \setminus L) \neq \emptyset\}.$$

Therefore,

$$\begin{aligned} B^* \setminus (A^* \setminus L)\sigma^{-1} &= \{i_1 \cdots i_k \in B^* \mid L_{i_1} L_{i_2} \cdots L_{i_k} \cap (A^* \setminus L) = \emptyset\} \\ &= \{i_1 \cdots i_k \in B^* \mid L_{i_1} L_{i_2} \cdots L_{i_k} \subset L\} \\ &= \{u \in B^* \mid u\sigma \subset L\} = R(L). \end{aligned}$$

Now L is rational, and so is $(A^* \setminus L)$, since the rational sets of A^* are closed under complement. It follows that $(A^* \setminus L)\sigma^{-1}$ is a rational set of B^* (see Section 3.b above). Therefore $B^* \setminus (A^* \setminus L)\sigma^{-1}$ is rational. □

This gives the solution to the first problem.

Proposition 4.2 *A rational language L belongs to* $\mathrm{Rat}(\mathcal{L})$ *if and only if* $(R(L))\sigma = L$. *Problem 1 is decidable.*

Proof. By Lemma 4.1, $R(L)$ is a rational set. Therefore, $(R(L))\sigma \in \mathrm{Rat}(\mathcal{L})$. Thus if $(R(L))\sigma = L$, then $L \in \mathrm{Rat}(\mathcal{L})$. Conversely, if $L \in \mathrm{Rat}(\mathcal{L})$, then there exists $K \in \mathrm{Rat}(B^*)$ such that $K\sigma = L$. Now, if $u \in K$, then $u\sigma \subset K$, and therefore $u \in R(L)$. It follows that K is contained in $R(L)$, so that $L = K\sigma \subset (R(L))\sigma \subset L$. Therefore $(R(L))\sigma = L$. This establishes the first sentence of the proposition.

Given L and $\mathcal{L}$, there exists an algorithm to compute the rational set $(R(L))\sigma$ and to test whether $L = (R(L))\sigma$. Therefore Problem 1 is decidable. □

We now consider Problem 2. To each word $v \in A^*$, associate the minimal length of all the words $u = i_1 \cdots i_k \in R(L)$ such that $v \in u\sigma = L_{i_1} L_{i_2} \cdots L_{i_k}$ (if no such word exists, we map v onto ∞). This defines a mapping $\gamma : A^* \to N \cup \{\infty\}$, and, for every $v \in A^*$,

$$v\gamma = \min\{|u| \mid u \in B^*, u\sigma \subset L \text{ and } v \in u\sigma\}.$$

Note that $v\gamma < \infty$ if and only if there exists $u \in R(L)$ such that $v \in u\sigma$, that is, if and only if $v \in R(L)\sigma$. Put $S = A^*\gamma$. The next proposition shows that Problem 2 reduces to deciding whether S is finite.

Proposition 4.3 *L belongs to* $\mathrm{Pol}(\mathcal{L})$ *if and only if* $L \in \mathrm{Rat}(\mathcal{L})$ *and S is finite.*

Proof. If $L \in \mathrm{Pol}(\mathcal{L})$, then $L \in \mathrm{Rat}(\mathcal{L})$ and $L = (R(L))\sigma$ by Proposition 4.2. Furthermore, there exists a finite set $F \subset B^*$ such that $L = F\sigma$. Let n be the maximum length of the words of F. If $v \in L$, then $v \in u\sigma$ for some $u \in F$ and thus $v\gamma \leq n$. On the other hand, $v\gamma = \infty$ if $v \notin L$, since $L = (R(L))\sigma$. Thus S is contained in $\{0, 1, \cdots, n\} \cup \{\infty\}$ and is finite.

Conversely, suppose $L = (R(L))\sigma$ and S finite. Then S is contained in the set $\{0, 1, \cdots, n\} \cup \{\infty\}$ for some n. Therefore $v\gamma \leq n$ if and only if $v \in L$. It follows that $F\sigma = L$, where $F = \{u \in R(L) \mid |u| \leq n\}$. Thus $L \in \mathrm{Pol}(\mathcal{L})$. $\square$

It remains to compute the set S. Since $R(L)$ is rational, and since σ is a rational substitution, the transduction $\tau : A^* \to B^*$ defined by $v\tau = R(L) \cap v\sigma^{-1}$ is rational. By Theorem 3.3, it admits a rational linear representation (X, μ, Y), where $\mu : A^* \to \mathrm{Rat}(B^*)^{n \times n}$ is a monoid morphism, and $X, Y \in \mathrm{Rat}(B^*)^n$. Furthermore, this linear representation can effectively be computed, given L and $\mathcal{L}$. Thus, for every $v \in A^*$,

$$v\tau = \sum_{1 \leq i,j \leq n} X_i (v\mu_{i,j}) Y_j.$$

Let K be the semiring $(\mathbb{N} \cup \{\infty\}, \min, +)$ and let $\varphi : \mathrm{Rat}(B^*) \to K$ be the mapping defined by

$$X\varphi = \min\{|x| \mid x \in X\} \text{ and } \emptyset\varphi = \infty.$$

Then φ is a semiring morphism, and it induces a monoid morphism

$$\varphi : \mathrm{Rat}(B^*)^{n \times n} \to K^{n \times n}.$$

Now, $\gamma = \tau\varphi$, and hence, for every $v \in A^*$,

$$v\gamma = \min_{i,j} \{(X_i\varphi) + (v\mu\varphi_{i,j}) + (Y_j\varphi)\}.$$

Note that this formula is very close to the relation $(*)$ defining linear representations, except that the semiring K is not of the required form. This obstacle can be overcome by noticing that K is isomorphic to the semiring $\mathcal{P}(M)$ where M is the monoid of all subsets of a^* of the form $a^n a^*$ under concatenation (consider the application $n \to a^n a^*$). Thus γ is in fact a representable transduction from A^* into M.

Put $T = A^*\mu\varphi$: T is a submonoid of $K^{n \times n}$ generated by the finite set $\{a\mu\varphi \mid a \in A\}$. Setting

$$I = \{i \mid X_i \neq \emptyset\} = \{i \mid X_i\varphi < \infty\} \text{ and } J = \{j \mid Y_j \neq \emptyset\} = \{j \mid Y_j\varphi < \infty\},$$

we obtain

Proposition 4.4 *The set S is finite if and only if there exists $i \in I$ and $j \in J$ such that the set $\{m_{i,j} \mid m \in T\}$ is finite.*

Therefore, Problem 2 reduces to deciding whether, for a given pair i, j, the set $\{m_{i,j} \mid m \in T\}$ is finite or not. This problem is now a problem of pure semigroup theory, that has been solved positively by Hashiguchi.

Theorem 4.5 [4] *Given i, j, n such that $0 \leq i, j \leq n$, given a finite set F of matrices of $K^{n \times n}$ generating a submonoid T, one can effectively decide whether the set $\{m_{i,j} \mid m \in T\}$ is finite or not.*

The proof is too difficult to be given here, but has motivated some very interesting developments.

Corollary 4.6 *Problem 2 is decidable.*

5 Conclusion.

We have seen that, for the most part, natural operations on (varieties of) languages correspond to natural operations on (varieties of) monoids : concatenation corresponds to aperiodic relational morphisms and to Schützenberger products, non ambiguous concatenation to locally trivial relational morphisms, length preserving morphisms to power monoids, sequential functions to wreath products, etc. In some cases, the corresponding operation is still unknown : for instance, it would be interesting to find an operation on languages corresponding to nilpotent relational morphisms, and an operation on monoids corresponding to shuffle. This correspondence between languages and monoids has motivated numerous research articles, and a number of problems are still open : on the semigroup side, the complete classification of the varieties of the form $\mathbf{PV}$ or the (many) decidability problems about varieties of the form $\mathbf{V}^{-1}\mathbf{W}$ or $\mathbf{V} * \mathbf{W}$; on the language side, the characterisation of varieties closed under shuffle or under pure star, and all problems related to the star operation and to the concatenation product (see the articles of P. Weil and K. Hashiguchi in this volume).

References

[1] J. Berstel, *Transductions and Context-free Languages*, Teubner, Stuttgart, 1979.

[2] S. Eilenberg, *Automata, Languages and Machines*, Academic Press, New York, Vol. A, 1974; Vol B, 1976.

[3] K. Hashiguchi, A decision procedure for the order of regular events, *Theoret. Comput. Sci.* **8**, (1979), 69–72.

[4] K. Hashiguchi, Limitedness theorem on finite automata with distance functions, *J. Comput. System Sci.* **24**, (1982), 232–244.

[5] K. Hashiguchi, Representation theorems on regular languages, *J. Comput. System Sci.* **27**, (1983), 101–115.

[6] K. Hashiguchi, Regular languages of star-height one, *Information and Control* **53**, (1982), 199–210.

[7] K. Hashiguchi, Improved limitedness theorems on finite automata with distance functions, *Rapport LITP* **86-72**, (1986).

[8] J.-É. Pin, Langages reconnaissables et codage préfixe pur. 8th ICALP, *Lecture Notes in Comp. Sci.* **115**, (1981), 184–192.

[9] J.-É. Pin, *Varieties of formal languages*, Masson, Paris (1984), North Oxford Academic, London and Plenum, New-York (1986).

[10] J.-É. Pin, Power semigroups and related varieties of finite semigroups, *Semigroups and Their Applications*, S.M. Goberstein and P.M. Higgins, eds., D. Reidel (1986), 139–152.

[11] J.-É. Pin, H. Straubing and D. Thérien, Locally trivial categories and unambiguous concatenation, *J. Pure Appl. Algebra* **52**, (1988), 297–311.

[12] J.-É. Pin and J. Sakarovitch, Operations and transductions that preserve rationality, 6th GI Conference, *Lecture Notes in Comp. Sci.* **145**, (1983), 617–628.

[13] J.-É. Pin and J. Sakarovitch, Une application de la représentation matricielle des transductions, *Theoret. Comput. Sci.* **35**, (1985), 271–293.

[14] Ch. Reutenauer, Sur les variétés de langages et de monoïdes, *Lecture Notes in Comp. Sci.* **67**, Springer, Berlin, (1979), 260–265.

[15] H. Straubing, Recognizable sets and power sets of finite semigroups, *Semigroup Forum* **18**, (1979), 331–340.

[16] H. Straubing, Aperiodic homomorphisms and the concatenation product of recognizable sets, *J. Pure Appl. Algebra* **15**, (1979), 319–327.

[17] H. Straubing, Relational morphisms and operations on recognizable sets, *RAIRO Inf. Théor.* **15**, (1981), 149–159.

[18] B. Tilson, Chapters XI and XII of reference [2].