Notes on maximal congruence relations, automata and related topics

By H. ANDRÉKA*, S. HORVÁTH**, I. NÉMETI*

Abstract

The paper starts from the fact that if r_0 is an equivalence relation on a free semigroup **A**, then (uniquely) exists a greatest right compatible refinement of r_0 (see e.g. [3, chapter 9] and [4, 1. §]).

In Part 1, the authors generalize the above question and investigate it in the case when **A** is an arbitrary semigroup. They present a constructive proof for one of the concerning theorems (Theorem 1′) e.g. they show that if r_0 is an equivalence relation on **A**, then the relation

$$r_m \overset{\text{def}}{=} \{\langle x, y\rangle | \langle x, y\rangle \in r_0 \wedge (\forall a, b)[a, b \in \mathbf{A} \Rightarrow (\langle ax, ay\rangle, \langle xb, yb\rangle, \langle axb, ayb\rangle \in r_0)]\}$$

is the greatest congruent refinement of r_0 in the sense that whenever r_1 is a congruence relation on **A** and $r_1 \subset r_0$, then $r_1 \subset r_m$.

In an interesting way, it turns out that in the definition of r_m, requiring $\langle axb, ayb\rangle \in r_0$ too (in addition to $\langle ax, ay\rangle, \langle xb, yb\rangle \in r_0$), is not superfluous: generally it does not follow from the other two.

The most general theorem of Part 1 is proved by using lattice-theoretical considerations (Theorem 1).

In Part 2, it is proved (Theorems 2 and 2′) that a partial reverse of Theorem 1 is equivalent to **A** having some sort of the special "quasi-trivial" structure (Definition 1).

In part 3, we represent every equivalence class of initially connected Moore automata, the elements of which induce the same automaton mapping $\bar{f}$, by the function f, derived from $\bar{f}$ by putting for every $w \in X^*$ (X is the input alphabet)

$$f(w) \overset{\text{def}}{=} \text{"the last letter of } \bar{f}(w)\text{"}.$$

These functions f we simply call automata. We draw a short parallel between the notion of an automaton f and the classical notion of a Moore automaton. During this the theorems of Part 1 prove to be directly applicable to the automata f, and in this way classical results concerning Moore automata can be deduced (e.g. the Corollary of Statement 1).

As a generalization of the fact that the semigroup of a finite Moore automaton is also finite, we prove (Statement 2) that if r is a right congruence relation of finite

index, on a semigroup $\mathbf{A}$, then r can always be refined into a congruence relation of finite index.

In connection with the general investigation of the semigroup of the so-called semigroup-machine $\langle \mathbf{A}, \mathbf{A}, \delta \rangle$, where $\mathbf{A}$ is an arbitrary semigroup and $(\forall a, b \in \mathbf{A})$ $\delta(a, b) \stackrel{\text{def}}{=} ab$; we introduce the "congruence relations of right uniformity, left uniformity and uniformity" (Def. 8).

At the end of Part 3, we prove that the possibility of simulating an automaton f by an automaton g, depends essentially on the semigroups of f and g, and is independent of their input alphabets which may be different.

1. Maximal compatible refinements of equivalence relations; generalizations

In this paper by the word *relation* we shall always mean a binary relation r over some nonvoid set $\mathbf{A}$ i.e.

$$r \subset \mathbf{A} \times \mathbf{A} = \mathbf{A}^2.$$

If we define an associative binary operation "$\circ$" on $\mathbf{A}$, we have the semigroup $\langle \mathbf{A}, \circ \rangle$. For the sake of simplicity, we shall refer to $\mathbf{A}$ *as a semigroup* simply by the same letter $\mathbf{A}$, instead of $\langle \mathbf{A}, \circ \rangle$ and instead of $x \circ y$ we shall write xy. If an equivalence relation r on $\mathbf{A}$ has the property

$$(\forall x, y, u, w)[(\langle x, y \rangle \in r \wedge \langle u, w \rangle \in r) \Rightarrow \langle xu, yw \rangle \in r], \tag{1.1}$$

we call it a *congruence relation* on $\mathbf{A}$ (as a semigroup). If we regard only "one half" of (1.1), namely

$$(\forall x, y, u)[(\langle x, y \rangle \in r \wedge u \in \mathbf{A}) \Rightarrow \langle xu, yu \rangle \in r] \tag{1.2}$$

or

$$(\forall x, y, u)[(\langle x, y \rangle \in r \wedge u \in \mathbf{A}) \Rightarrow \langle ux, uy \rangle \in r], \tag{1.3}$$

then we call r a *right* or *left congruence relation* respectively. Of course, a congruence relation is at the same time a right congruence relation as well as a left one. Conversely, because of the transitivity of r (as an equivalence relation)

$$(\forall r)[((1.2) \wedge (1.3)) \Rightarrow (1.1)].$$

Hence

$$(\forall r)[((1.2) \wedge (1.3)) \Leftrightarrow (1.1)]$$

i.e. r is a congruence relation iff

$$(\forall x, y, u)[(\langle x, y \rangle \in r \wedge u \in A) \Rightarrow (\langle xu, yu \rangle, \langle ux, uy \rangle \in r)]. \tag{1.4}$$

We shall always use (1.4) instead of (1.1).

The following notations will also prove useful

$$\mathscr{C}\mathbf{A} \stackrel{\text{def}}{=} \{r \mid r \text{ is an equivalence relation on } \mathbf{A}\},$$

$$\mathscr{R}\mathbf{A} \stackrel{\text{def}}{=} \{r \mid r \text{ is a reflexive relation on } \mathbf{A}\},$$

$$\mathscr{S}\mathbf{A} \stackrel{\text{def}}{=} \{r \mid r \text{ is a symmetric relation on } \mathbf{A}\},$$

$$\mathscr{T}\mathbf{A} \stackrel{\text{def}}{=} \{r \mid r \text{ is a transitive relation on } \mathbf{A}\},$$

$$\mathcal{TS}\mathbf{A} \overset{\text{def}}{=} \mathcal{T}\mathbf{A} \cap \mathcal{S}\mathbf{A},$$
$$\mathcal{TR}\mathbf{A} \overset{\text{def}}{=} \mathcal{T}\mathbf{A} \cap \mathcal{R}\mathbf{A},$$
$$\mathcal{SR}\mathbf{A} \overset{\text{def}}{=} \mathcal{S}\mathbf{A} \cap \mathcal{R}\mathbf{A},$$
$$\mathcal{C}_\Omega\mathbf{A} \overset{\text{def}}{=} \{r | r \text{ is a congruence relation on } \mathbf{A}\},$$
$$\mathcal{C}_{\Omega R}\mathbf{A} \overset{\text{def}}{=} \{r | r \text{ is a right congruence relation on } \mathbf{A}\},$$
$$\mathcal{C}_{\Omega L}\mathbf{A} \overset{\text{def}}{=} \{r | r \text{ is a left congruence relation on } \mathbf{A}\}.$$

Of course, by definition, $\mathcal{C}\mathbf{A} = \mathcal{R}\mathbf{A} \cap \mathcal{S}\mathbf{A} \cap \mathcal{T}\mathbf{A}$ and by the equivalence of (1.1) and (1.4)

$$\mathcal{C}_\Omega\mathbf{A} = \mathcal{C}_{\Omega R}\mathbf{A} \cap \mathcal{C}_{\Omega L}\mathbf{A}.$$

Further notations

$$\pi\mathbf{X} \overset{\text{def}}{=} \{\mathbf{Y} | \mathbf{Y} \subset \mathbf{X}\}$$

(here and all along the symbol "$\subset$" may stand for "$=$" too),

$$1_\mathbf{X} \overset{\text{def}}{=} \{\langle z, z\rangle | z \in \mathbf{X}\}.$$

If $r \subset \mathbf{A}^2$ and n is a natural number, the n-th power of r we define as

$$r^n \overset{\text{def}}{=} \{\langle x, y\rangle | (\exists z_0, z_1, \ldots, z_n)[(z_0, z_1, \ldots, z_n \in A) \wedge$$
$$\wedge z_0 = x \wedge z_n = y \wedge (\langle z_0, z_1\rangle, \ldots, \langle z_{n-1}, z_n\rangle \in r)]\}$$

and the *transitive closure* of r is

$$\hat{r} \overset{\text{def}}{=} \bigcup_{i=1}^{\infty} r^i. \tag{1.5}$$

As is well known, for any set $\mathbf{X}$, $\pi\mathbf{X}$ forms a complete lattice with respect to the partial ordering $\subset$.

In this case, the meet and join operations are the following

$$(\forall \mathbf{Z} \subset \pi\mathbf{X}) \begin{cases} \bigcap\limits_{z \in \mathbf{Z}} z \overset{\text{def}}{=} \bigcap\limits_{z \in \mathbf{Z}} z & (1.6) \\ \text{and} \\ \bigcup\limits_{z \in \mathbf{Z}} z \overset{\text{def}}{=} \bigcup\limits_{z \in \mathbf{Z}} z & (1.7) \end{cases}$$

where $\frown$, $\smile$ denote the lattice-theoretical operations and $\bigcap$, $\bigcup$ are the usual symbols of the set-theoretical intersection and union respectively. We agree (as usual) that

$$\bigcap_{z \in \emptyset} z = \mathbf{X}, \quad \bigcup_{z \in \emptyset} z = \emptyset.$$

E.g. if $\mathbf{X} = \mathbf{A}^2$, $\pi\mathbf{A}^2$ is a complete lattice with meet operation (1.6) and join operation (1.7). However, if we replace $\pi\mathbf{A}^2$ with $\mathcal{T}\mathbf{A}$, we must modify the join operation of (1.7) for $\mathcal{T}\mathbf{A}$ to be a complete lattice (under the partial ordering $\subset$)

$$\smile_{r \in \mathbf{Z}} r \overset{\text{def}}{=} \coprod_{r \in \mathbf{Z}} r \overset{\text{def}}{=} \widehat{\bigcup_{r \in \mathbf{Z}} r}. \tag{1.8}$$

The reason why transitive closure (1.5) has entered is just the transitivity of the elements of $\mathcal{T}\mathbf{A}$. It can easily be checked that with the operations $\cap$ and $\amalg$, $\langle \mathcal{T}\mathbf{A}, \subset \rangle$ is indeed a complete lattice.

Using the following notation for any two lattices V and W, $V \to W =$ "W is a complete sublattice of V", the following "directed graph" is valid

$$\begin{array}{ccccccccc} & \nearrow & \mathcal{T}\mathcal{S}\mathbf{A} & \searrow & & \nearrow & \mathcal{C}_{\Omega R}\mathbf{A} & \searrow & \\ \mathcal{T}\mathbf{A} & & & & \mathcal{C}\mathbf{A} & & & & \mathcal{C}_{\Omega}\mathbf{A}. \\ & \searrow & \mathcal{T}\mathcal{R}\mathbf{A} & \nearrow & & \searrow & \mathcal{C}_{\Omega L}\mathbf{A} & \nearrow & \end{array} \tag{1.9}$$

(The relation "$\to$" is itself a partial ordering over the complete sublattices of any complete lattice, as it is reflexive, antisymmetric and transitive.)

The "edges" in (1.9) between $\mathcal{T}\mathbf{A}$ and $\mathcal{C}\mathbf{A}$ may be verified simply by using definitions (1.6) and (1.8), while for those between $\mathcal{C}\mathbf{A}$ and $\mathcal{C}_{\Omega}\mathbf{A}$ we must take into account (1.2), (1.3) and (1.4) also (to show that the meet and join operations always result in an appropriate — belonging to $\mathcal{C}_{\Omega R}\mathbf{A}$ etc. — relation). This is a routine calculation. ($\pi\mathbf{A}^2 \to \mathcal{T}\mathbf{A}$ is not true, because the join operation in $\mathcal{T}\mathbf{A}$ (see (1.8)) differs from that in $\pi\mathbf{A}^2$ (see (1.7))).

The common unit element of all these complete lattices is $\mathbf{A}^2$, while the zero element of $\mathcal{T}\mathcal{R}\mathbf{A}$, $\mathcal{C}\mathbf{A}$, $\mathcal{C}_{\Omega R}\mathbf{A}$, $\mathcal{C}_{\Omega L}\mathbf{A}$ and $\mathcal{C}_{\Omega}\mathbf{A}$ is $1_{\mathbf{A}}$, and that of $\mathcal{T}\mathbf{A}$ and $\mathcal{T}\mathcal{S}\mathbf{A}$ is $\emptyset$. For any two relations r, r_1 for which $r_1 \subset r$, we say that r_1 is less than or equal to r, or r is greater than or equal to r_1, or (equivalently) r_1 is a refinement of r.

Theorem 1. If $\mathbf{A}$ is a semigroup and

(a) $$r_0 \in \mathcal{T}\mathcal{R}\mathbf{A} \quad \text{and} \quad \mathbf{M} \in \{\mathcal{C}\mathbf{A}, \mathcal{C}_{\Omega L}\mathbf{A}, \mathcal{C}_{\Omega R}\mathbf{A}, \mathcal{C}_{\Omega}\mathbf{A}\},$$

or

(b) $$r_0 \in \mathcal{T}\mathbf{A}, \quad \pi r_0 \cap \mathcal{S}\mathbf{A} \neq \emptyset \quad \text{and} \quad \mathbf{M} = \mathcal{T}\mathcal{S}\mathbf{A},$$

then the set $\mathbf{H} \overset{\text{def}}{=} \mathbf{M} \cap \pi r_0$ has a (unique) greatest element r_g

$$(\exists r_g \in \mathbf{H})(\forall r)[r \in \mathbf{H} \Rightarrow r \subset r_g]. \tag{1.10}$$

Proof

(a) By the definition of r_0, $1_{\mathbf{A}} \subset r_0$, so $\mathbf{H} \neq \emptyset$ (the case is not trivial). Being $\mathbf{M}$ a complete lattice and $\mathbf{H} \subset \mathbf{M}$, there is in $\mathbf{M}$ a least upper bound of $\mathbf{H}$ (see (1.8) and (1.9))

$$r_g \overset{\text{def}}{=} \coprod_{r \in \mathbf{H}} r \tag{1.11}$$

for which $r \in \mathbf{H} \Rightarrow r \subset r_g$ of (1.10) holds. So we have only to prove that

$$r_g \in \pi r_0. \tag{1.12}$$

Being r_g the transitive closure of a subset $\left(\bigcup_{r \in \mathbf{H}} r\right)$ of r_0 (see (1.11), (1.8) and (1.5)) and r_0 being transitive,

$$r_g \subset r_0 \tag{1.13}$$

i. e. (1.12) holds.

(b) Again $\mathbf{H} \neq \emptyset$ (the case is not trivial) and by an argument, similar to that of (a), we again have (1.13) i.e. (1.12).

Now we proceed by giving a *constructive proof* for a special case of Theorem 1, part (a).

Theorem 1'. If $\mathbf{A}$ is a semigroup, $r_0 \in \mathscr{C}\mathbf{A}$, $\mathbf{M} \in \{\mathscr{C}_{\Omega R}\mathbf{A}, \mathscr{C}_{\Omega L}\mathbf{A}, \mathscr{C}_{\Omega}\mathbf{A}\}$, and $\mathbf{H} \overset{\text{def}}{=} \mathbf{M} \cap \cap \pi r_0$, then (1.10) holds.

Proof. First we deal with the case $\mathbf{M} = \mathscr{C}_{\Omega}\mathbf{A}$ and then point out the obvious differences for the case $\mathbf{M} \in \{\mathscr{C}_{\Omega R}\mathbf{A}, \mathscr{C}_{\Omega L}\mathbf{A}\}$.

Let

$$r_m \overset{\text{def}}{=} \{\langle x, y\rangle | \langle x, y\rangle \in r_0 \wedge (\forall a, b)[a, b \in \mathbf{A} \Rightarrow (\langle ax, ay\rangle, \langle xb, yb\rangle, \langle axb, ayb\rangle \in r_0)]\}. \tag{1.14}$$

Obviously, $r_m \subset r_0$ i.e. $r_m \in \pi r_0$ and it can easily be verified that r_m satisfies condition (1.4), so

$$r_m \in \mathscr{C}_{\Omega}\mathbf{A} \cap \pi r_0 = \mathbf{H}.$$

It remained to prove (1.10) for r_m in place of r_g

$$(\forall r)[r \in \mathbf{H} \Rightarrow r \subset r_m]. \tag{1.15}$$

By the definition of $\mathbf{H}$

$$r \in \mathbf{H} \Leftrightarrow \left\{\begin{matrix} \text{(i)} \quad r \in \mathscr{C}_{\Omega}\mathbf{A} \\ \text{and} \\ \text{(ii)} \quad r \subset r_0 \end{matrix}\right\}. \tag{1.16}$$

From (i) of (1.16) follows (see (1.4)) that

$$(\forall a, b, x, y \in \mathbf{A})(\forall r \in \mathbf{H})[\langle x, y\rangle \in r \Rightarrow (\langle ax, ay\rangle, \langle xb, yb\rangle, \langle axb, ayb\rangle \in r)]. \tag{1.17}$$

From (1.17), (ii) of (1.16), and (1.14) we get that

$$(\forall x, y \in \mathbf{A})(\forall r \in \mathbf{H})[\langle x, y\rangle \in r \Rightarrow \langle x, y\rangle \in r_m],$$

which is equivalent to (1.15).

If e.g. $\mathbf{M} = \mathscr{C}_{\Omega L}\mathbf{A}$, then b, $\langle xb, yb\rangle$ and $\langle axb, ayb\rangle$ above must be deleted etc.

Remark. Condition (1.4) suggests that requiring $\langle axb, ayb\rangle \in r_0$ too in (1.14) is perhaps superfluous, but this is not at all the case

Fact. In definition (1.14), condition $(\forall a, b \in \mathbf{A})[\langle axb, ayb\rangle \in r_0]$ does not follow from

$$(\forall a, b \in \mathbf{A})[\langle ax, ay\rangle, \langle xb, yb\rangle \in r_0].$$

Proof. We construct an example. Let $\mathbf{A} = \{1, 2\}^*$ (the free monoid, generated by the set $\{1, 2\}$) and

$$r_0 \overset{\text{def}}{=} \{\langle u\alpha v, w\alpha z\rangle | u, v, w, z \in \{1, 2\} \wedge \alpha \in \{1, 2\}^*\} \cup \{\langle \Lambda, \Lambda\rangle, \langle 1, 1\rangle, \langle 1, 2\rangle, \langle 2, 1\rangle, \langle 2, 2\rangle\},$$

where

$$\Lambda \overset{\text{def}}{=} \text{the empty word (of any free monoid).}$$

It can easily be seen that $r_0 \in \mathscr{C}\mathbf{A}$, because $\langle \beta, \gamma\rangle \in r_0$ means that (denoting the length of the words in $\mathbf{A}$ by "lg") $\lg(\beta) = \lg(\alpha)$ and if $\lg(\beta) > 2$, then removing the first and last symbols from β and γ, the remaining word will be the same.

Constructing from this r_0 relations

$$r'_m \stackrel{\text{def}}{=} \{\langle x, y\rangle | \langle x, y\rangle \in r_0 \wedge (\forall a \in A)[\langle ax, ay\rangle, \langle xa, ya\rangle \in r_0]\}$$

and r_m — the latter according to (1.14) —, then by an easy calculation we get that $r'_m \supset r_m$, $r'_m \neq r_m$, $r'_m \in \mathscr{C}\mathbf{A}$, $r_m \notin \mathscr{C}_\Omega \mathbf{A}$. Namely, $r_m = 1_\mathbf{A}$ and $r'_m - r_m = \{\langle 1, 2\rangle, \langle 2, 1\rangle\}$.

Remark. If **A** is a monoid (i.e. a semigroup, having a unit element) then (1.14) becomes simpler

$$r_m \stackrel{\text{def}}{=} \{\langle x, y\rangle | (\forall a, b)[a, b \in \mathbf{A} \Rightarrow \langle axb, ayb\rangle \in r_0]\}. \tag{1.18}$$

2. A characterization of quasi-trivial semigroups

We shall introduce the following

Definition 1. We call the semigroup **A** *right quasi-trivial* iff $|\mathbf{A}| \geqq 3$ and there is a decomposition of $\mathbf{A}: \mathbf{A} = \mathbf{A}_{1R} \cup \mathbf{A}_{2R}$, $\mathbf{A}_{1R} \cap \mathbf{A}_{2R} = \emptyset$, for which there is a function $f_{R\mathbf{A}}: \mathbf{A}_{2R} \to \mathbf{A}_{2R}$ (in case $A_{2R} \neq \emptyset$) and $f_{R\mathbf{A}} \upharpoonright \mathscr{R}(f_{R\mathbf{A}}) = 1_{\mathscr{R}(f_{R\mathbf{A}})}$[1] and

$$(\forall x \in \mathbf{A})(\forall y \in \mathbf{A}) \left[xy = \begin{cases} x, & \text{if } y \in \mathbf{A}_{1R} \\ f_{R\mathbf{A}}(y), & \text{if } y \in \mathbf{A}_{2R} \end{cases} \right].$$

We analogously interpret the *left quasi-trivial* property. We can refer to both of right and left quasi-triviality by saying simply *quasi-trivial.* We call the semigroup **A** *strongly quasi-trivial,* iff the structure of **A** is one of the following three alternatives

$$\begin{aligned} &\text{(i)} \quad (\forall x, y)[x, y \in \mathbf{A} \Rightarrow xy = x], \\ &\text{(ii)} \quad (\forall x, y)[x, y \in \mathbf{A} \Rightarrow xy = y], \\ &\text{(iii)} \quad (\exists c \in \mathbf{A})(\forall x, y)[x, y \in \mathbf{A} \Rightarrow xy = c]. \end{aligned} \tag{2.1}$$

Obviously, if **A** is strongly quasi-trivial, then it is quasi-trivial also, but the converse is not true. Further, it can easily be checked that the quasi-trivial structure is associative.

As a characterization of quasi-trivial and strongly quasi-trivial semigroups, we prove the following theorem, which is a partial reverse of Theorem 1'.

Theorem 2. If **A** is a semigroup with $|\mathbf{A}| \geqq 3$, $r_0 \subset \mathbf{A}^2$ and $\mathbf{H} \stackrel{\text{def}}{=} \mathbf{M} \cap \pi r_0$, where

(i) $\mathbf{M} = \mathscr{C}_{\Omega R} \mathbf{A}$,

(ii) $\mathbf{M} = \mathscr{C}_{\Omega L} \mathbf{A}$,

(iii) $\mathbf{M} = \mathscr{C}_\Omega \mathbf{A}$,

then the needful and sufficient condition of

$$(\forall r_0)[((r_0 \in \mathscr{S}\mathscr{R}\mathbf{A}) \wedge (1.10)) \Rightarrow r_0 \in \mathscr{T}\mathbf{A}] \tag{2.2}$$

[1] As usual, $\mathscr{D}$ and $\mathscr{R}$ stand for "domain" and "range" respectively. The symbol "$\upharpoonright$" is used to denote the restriction of functions.

is that

$$\begin{aligned}&\text{(i)}\quad \mathbf{A} \text{ is right quasi-trivial,}\\ &\text{(ii)}\quad \mathbf{A} \text{ is left quasi-trivial,}\\ &\text{(iii)}\quad \mathbf{A} \text{ is strongly quasi-trivial,}\end{aligned} \tag{2.3}$$

respectively.

Remark. If in (2.2) we change "$\Rightarrow r_0 \in \mathcal{T}\mathbf{A}$" into "$\Rightarrow r_0 \in \mathcal{C}\mathbf{A}$", then (2.2) remains the same.

Proof. First of all, transform (2.2) into an equivalent form

$$(\forall r_0)\left[((r_0 \in \mathcal{S}\mathcal{R}\mathbf{A}) \wedge (r_0 \notin \mathcal{T}\mathbf{A})) \Rightarrow \neg(1.10)\right]. \tag{2.4}$$

If (2.4) is true, then it must hold for every r_0 of the form

$$\begin{aligned} r_0' &\overset{\text{def}}{=} 1_{\mathbf{A}} \cup \{\langle a, b\rangle, \langle b, a\rangle, \langle a, c\rangle, \langle c, a\rangle\},\\ a, b, c &\in \mathbf{A}, \qquad a \neq b \neq c \neq a. \end{aligned} \tag{2.5}$$

(Evidently, for any such r_0', $r_0' \in \mathcal{S}\mathcal{R}\mathbf{A}$ and $r_0' \notin \mathcal{T}\mathbf{A}$.)

(i) $\mathbf{M} = \mathcal{C}_{\Omega R}\mathbf{A}$. If one of the two equivalence relations

$$r_b \overset{\text{def}}{=} 1_{\mathbf{A}} \cup \{\langle a, b\rangle, \langle b, a\rangle\} (\subset r_0')$$

and

$$r_c \overset{\text{def}}{=} 1_{\mathbf{A}} \cup \{\langle a, c\rangle, \langle c, a\rangle\} (\subset r_0')$$

is not a right congruence relation, then (2.4) does not hold for $r_0 = r_0'$. (Because if e.g. $r_b \notin \mathcal{C}_{\Omega R}\mathbf{A}$ and $r_c \in \mathcal{C}_{\Omega R}\mathbf{A}$, then taking $r_g = r_c$, (1.10) will hold; and if also $r_c \notin \mathcal{C}_{\Omega R}\mathbf{A}$, then $r_g = 1_{\mathbf{A}}$ will satisfy (1.10).) On the other hand, if $r_b, r_c \in \mathcal{C}_{\Omega R}\mathbf{A}$, then $\sup(\mathbf{H})$, by virtue of its belonging to $\mathcal{T}\mathbf{A}$, must contain $\langle b, c\rangle$ (as $\langle b, a\rangle, \langle a, c\rangle \in r_0'$) therefore in this case $\sup(\mathbf{H}) \in \mathbf{H}$, i.e. (2.4) holds. This argument is valid for any r_0' of the type (2.5), so an equivalent transcription of (2.4) is the following

$$(\forall a, b)[(a, b \in A) \Rightarrow (1_{\mathbf{A}} \cup \{\langle a, b\rangle, \langle b, a\}) \in \mathcal{C}_{\Omega R}\mathbf{A}]. \tag{2.6}$$

Using criterion (1.2), (2.6) is further equivalent to

$$(\forall a, b, x)\left[(a, b, x \in A) \Rightarrow \left(ax = bx \vee (\{ax, bx\} \subset \{a, b\})\right)\right]. \tag{2.7}$$

Now we shall deduce (2.3) (i) from (2.7) (the converse is obvious: if $\mathbf{A}$ is right quasi-trivial, then (2.7) holds). Indeed, define the subset $\mathbf{A}_{1R}$ of $\mathbf{A}$ so

$$\mathbf{A}_{1R} \overset{\text{def}}{=} \{x \mid x \in \mathbf{A} \wedge (\forall y)[y \in \mathbf{A} \Rightarrow yx = y]\} \tag{2.8}$$

(obviously, $\mathbf{A}_{1R}$ may be empty), and let

$$\mathbf{A}_{2R} \overset{\text{def}}{=} \mathbf{A} - \mathbf{A}_{1R}. \tag{2.9}$$

Fix an arbitrary

$$x \in \mathbf{A}_{2R} \quad (\text{if } \mathbf{A}_{2R} \neq \emptyset).$$

By the definition of $\mathbf{A}_{2R}$, there is a $y \in \mathbf{A}$, for which

$$yx \neq y, \quad \text{say} \quad yx = z. \tag{2.10}$$

Then because of $|\mathbf{A}| \geqq 3$, there is a $u \in \mathbf{A}$, $u \neq y$, $u \neq z$. According to (2.7)

$$(z =)yx = ux \vee (\{yx, ux\} \subset \{y, u\}).$$

As $z \neq y$ (see (2.10)) and, by its choosing, $z \neq u$, $yx = z \notin \{y, u\}$, so

$$(\forall u)[(u \neq y \wedge u \neq z) \Rightarrow ux = z]. \tag{2.11}$$

Let us now examine zx. On the basis of (2.7), if $u \neq y$ and $u \neq z$

$$(zx = yx \vee (\{zx, yx\} \subset \{z, y\})) \wedge (zx = ux \vee (\{zx, ux\} \subset \{z, u\})). \tag{2.12}$$

As $yx = ux = z$ — from (2.10) and (2.11) —, (2.12) is not other than

$$zx = z \vee (\{zx, z\} \subset (\{z, y\} \cap \{z, u\})(= \{z\}))$$

i.e.

$$zx = z.$$

Summing up, if $x \in \mathbf{A}_{2R}$, then the value of wx does not depend on w

$$(\exists f_{\mathbf{RA}}: \mathbf{A}_{2R} \to \mathbf{A})(\forall w, x)[(w \in \mathbf{A} \wedge x \in \mathbf{A}_{2R}) \Rightarrow wx = f_{\mathbf{RA}}(x)]. \tag{2.13}$$

Taking now into consideration that the structure of $\mathbf{A}$ is associative; if $x \in \mathbf{A}_{2R}$ and $w, s \in \mathbf{A}$, then

$$f_{\mathbf{RA}}(x) = (ws)x = w(sx) = wf_{\mathbf{RA}}(x), \quad wf_{\mathbf{RA}}(x) = f_{\mathbf{RA}}(x),$$

independently of w, i.e. $f_{\mathbf{RA}}(x) \in \mathbf{A}_{2R}$ and

$$f_{\mathbf{RA}}(f_{\mathbf{RA}}(x)) = wf_{\mathbf{RA}}(x) = f_{\mathbf{RA}}(x),$$

from which we conclude, that in (2.13)

$$(\mathscr{R}(f_{\mathbf{RA}}) \subset \mathbf{A}_{2R}) \wedge (f_{\mathbf{RA}} \restriction \mathscr{R}(f_{\mathbf{RA}}) = 1_{\mathscr{R}(f_{\mathbf{RA}})}) \tag{2.14}$$

i.e. $\mathbf{A}$ is right quasi-trivial.

(ii) $\mathbf{M} = \mathscr{C}_{\Omega L}\mathbf{A}$. The argument is analogous to that of case (i).

(iii) $\mathbf{M} = \mathscr{C}_{\Omega}\mathbf{A}$. The left counterpart of (2.7) being

$$(\forall a, b, x)[a, b, x \in \mathbf{A} \Rightarrow (xa = xb \vee (\{xa, xb\} \subset \{a, b\}))], \tag{2.15}$$

we get in a similar way as in case (i), that in case (iii) — using condition (1.4) among others — (2.2) is equivalent to (2.7)$\wedge$(2.15), i.e.

$$(\forall a, b, x)[a, b, x \in \mathbf{A} \Rightarrow ((ax = bx \vee (\{ax, bx\} \subset \{a, b\})) \wedge \wedge (xa = xb \vee (\{xa, xb\} \subset \{a, b\})))]. \tag{2.16}$$

There are two distinct (disjoint) subcases of case (iii)

$$(\alpha) \quad (\forall a, b)[(a, b \in \mathbf{A} \wedge a \neq b) \Rightarrow ab \subset \{a, b\}],$$

$$(\beta) \quad (\exists a, b)[a, b \in \mathbf{A} \wedge a \neq b \wedge ab \notin \{a, b\}] \quad (\text{i.e.} \neg(\alpha)).$$

(α) Fix a, b for which, say, let

$$ab = a \quad (a, b \in \mathbf{A},\ a \neq b). \tag{2.17}$$

If $c \notin \{a, b\}$, then as now case (α) is valid,

$$cb \in \{c, b\} \tag{2.18}$$

and on the basis of (2.7) and (2.17)

$$cb \in \{a, c\}. \tag{2.19}$$

As $a \neq b$, from (2.18) and (2.19) follows

$$cb = c. \tag{2.20}$$

Further (2.7), (2.17) and (2.20) give that

$$bb \in \{a, b\} \wedge bb \in \{c, b\}$$

i.e. because of $a \neq c$;

$$bb = b.$$

From the above we can conclude that

$$(\forall x)[x \in \mathbf{A} \Rightarrow xb = x]. \tag{2.21}$$

Now let

$$y \neq b \neq x \neq y. \tag{2.22}$$

On the basis of (2.15), (2.21) and (2.22)

$$\big(xy = xb \vee (\{xy, xb\} \subset \{y, b\})\big) \wedge (xb = x \notin \{y, b\})$$

i.e.

$$xy = x \quad (x \neq y). \tag{2.23}$$

From (2.23), quite in a similar way as starting from (2.17), we can deduce that (2.21) is true for y in place of b. And finally, as $y(\neq b)$ was arbitrary, we get

$$(\forall u, w)[u, w \in \mathbf{A} \Rightarrow uw = u]. \tag{2.24}$$

If at the beginning in (2.17) we alter $ab = a$ into $ab = b$, then the final result will be

$$(\forall u, w)[u, w \in A \Rightarrow uw = w]. \tag{2.25}$$

(β) We can start with

$$(a, b, c \in A) \wedge (a \neq b \neq c \neq a) \wedge ab = c. \tag{2.26}$$

From (2.15) and (2.26)

$$\big(aa = ab \vee (\{aa, ab\} \subset \{a, b\})\big) \wedge (ab = c \notin \{a, b\}),$$

from which

$$aa = c. \tag{2.27}$$

Similarly, by means of (2.7) and (2.26), we get

$$bb = c.$$

To determine ac, using (2.15), (2.26) and (2.27), we can write

$$\big(ac = ab \vee (\{ac, ab\} \subset \{c, b\})\big) \wedge \big(ac = aa \vee (\{ac, aa\} \subset \{c, a\})\big) \wedge$$
$$\wedge ab = c \wedge aa = c \wedge (a \neq b \neq c \neq a)$$

i.e.

$$ac = c. \tag{2.28}$$

Likewise

$$bc = ca = cb = c.$$

For ba, using (2.7), (2.26) and (2.27)

$$\big(ba = aa \vee (\{ba, aa\} \subset \{b, a\})\big) \wedge (aa = c \notin \{b, a\})$$

and from this

$$ba = c. \tag{2.29}$$

At last, starting from (2.15), (2.26) and $cb = ca = c$, in the same way as leading to (2.28), we have

$$cc = c. \tag{2.30}$$

Summing up (2.26) to (2.30)

$$(\forall x, y)[(x, y \in \{a, b, c\}) \Rightarrow xy = c]. \tag{2.31}$$

If $z \in \mathbf{A} - \{a, b, c\}$, then in the same fashion as in (2.29) we have

$$zb = za = c.$$

Analogously to deducing (2.31), we conclude, that

$$(\forall u, w)[(u, w \in \{z, b, c\}) \Rightarrow uw = c] \tag{2.32}$$

and

$$(\forall u, w)[(u, w \in \{z, a, c\}) \Rightarrow uw = c]$$

and similarly, if $w \in \mathbf{A} - \{z, a, b, c\}$, then

$$wz = zw = ww = c. \tag{2.33}$$

Summarizing (2.31), (2.32) and (2.33)

$$(\forall x, y)[x, y \in \mathbf{A} \Rightarrow xy = c]. \tag{2.34}$$

As (2.24), (2.25) and (2.34) correspond to (2.1) (i), (2.1) (ii) and (2.1) (iii) respectively, we are ready.

Remark. In the proof of part (iii) and up to (2.13) in that of part (i) (analogous statement holds true of part (ii)) we did not make use of associativity.

In the following we give a second proof for part (iii), on the basis of (2.13) and its left counterpart, without making use of property (2.14) and the left counterpart of it i.e. again not taking into account associativity.

Second proof for part (iii) *of Theorem 2.* Let

$$\mathbf{A} = \mathbf{A}_{1R} \cup \mathbf{A}_{2R} (\mathbf{A}_{1R} \cap \mathbf{A}_{2R} = \emptyset)$$

the decomposition of A, defined by (2.8) and (2.9) (this decomposition exists — and is unique — for any semigroup A), and let

$$\mathbf{A} = \mathbf{A}_{1L} \cup \mathbf{A}_{2L} (\mathbf{A}_{1L} \cap \mathbf{A}_{2L} = \emptyset)$$

be the left counterpart of the former decomposition.

If one of $\mathbf{A}_{1R}$ and $\mathbf{A}_{1L}$ is A itself, we are ready, evidently having (2.1) (i) or (2.1) (ii) respectively.

If

$$\mathbf{A}_{1R} \neq \emptyset \wedge \mathbf{A}_{1L} \neq \emptyset \tag{2.35}$$

then

$$(\forall x, y)[(x \in \mathbf{A}_{1L} \wedge y \in \mathbf{A}_{1R}) \Rightarrow y = xy = x]$$

i.e.

$$|\mathbf{A}_{1R}| = |\mathbf{A}_{1L}| = 1,$$

$$\{e\} \stackrel{\text{def}}{=} \mathbf{A}_{1R} = \mathbf{A}_{1L}$$

and consequently

$$\mathbf{A}_{2R} = \mathbf{A}_{2L} = \mathbf{A} - \{e\}$$

(e is the — unique — identity element of A).

Furthermore

$$(\forall x, y)[(x \in \mathbf{A}_{2L} \wedge y \in \mathbf{A}_{2R}) \Rightarrow xy = f_{L\mathbf{A}}(x) = f_{R\mathbf{A}}(y)]$$

i.e.

$$f_{R\mathbf{A}} = f_{L\mathbf{A}} = \text{constant}. \tag{2.36}$$

Being $|\mathbf{A}| \geqq 3$, $|A_{2R}| (= |A_{2L}|) \geqq 2$, so there are $x, y \in \mathbf{A}_{2R}$, $x \neq y$, for which on one hand $ex = f_{R\mathbf{A}}(x) = f_{R\mathbf{A}}(y) = ey$, while on the other hand $ex = x \neq y = ey$, which is a contradiction, and therefore (2.35) is impossible. Thus, let e.g.

$$\mathbf{A}_{2R} = \mathbf{A} \wedge \mathbf{A}_{2L} \neq \emptyset$$

(the symmetric counterpart is quite analogous).

From this immediately follows (2.36) with $\mathbf{A}_{2R} = \mathbf{A}_{2L} = \mathbf{A}$ i.e. (2.1) (iii).

To close Part 2 of our paper, we formulate the following

Theorem 2'. If A is a semigroup and $|\mathbf{A}| \geqq 3$, then the following three statements are equivalent

$$\text{(a)} \begin{cases} \text{(i)} \ \mathbf{M} = \mathscr{C}_{\Omega R}\mathbf{A}, \\ \text{(ii)} \ \mathbf{M} = \mathscr{C}_{\Omega L}\mathbf{A}, \\ \text{(iii)} \ \mathbf{M} = \mathscr{C}_{\Omega}\mathbf{A}, \end{cases} r_0 \in \mathbf{A}^2, \mathbf{H} \stackrel{\text{def}}{=} \mathbf{M} \cap \pi r_0 \text{ and } (\forall r_0)[(r_0 \in \mathscr{S}\mathscr{R}\mathbf{A} \wedge (1.10)) \Rightarrow r_0 \in \mathscr{T}\mathbf{A}]$$

$$\text{(b)} \begin{cases} \text{(i)} \ \mathbf{A} \text{ is right quasi-trivial,} \\ \text{(ii)} \ \mathbf{A} \text{ is left quasi-trivial,} \\ \text{(iii)} \ \mathbf{A} \text{ is strongly quasi-trivial,} \end{cases}$$

$$\text{(c)} \begin{cases} \text{(i)} \ \mathscr{C}\mathbf{A} = \mathscr{C}_{\Omega R}\mathbf{A}, \\ \text{(ii)} \ \mathscr{C}\mathbf{A} = \mathscr{C}_{\Omega L}\mathbf{A}, \\ \text{(iii)} \ \mathscr{C}\mathbf{A} = \mathscr{C}_{\Omega}\mathbf{A} \end{cases}$$

(i.e. $(a)(x) \Leftrightarrow (b)(x) \Leftrightarrow (c)(x)$ for $x = \text{i, ii, iii}$).

Proof. It follows from the proof of Theorem 2 ($(a) \Leftrightarrow (b)$ is Theorem 2 itself).

3. Some questions of the semigroups and the simulation of automata

In this part of our paper the focus will be on automata, and we shall take known several widely accepted notions and notations of automata theory.

The set of all *initially connected Moore automata*, having the same input alphabet X and output alphabet Y, can be partitioned into equivalence classes, regarding two automata equivalent iff they induce the same *automaton mapping*

$$\bar{f}: X^* \to (Y^* - \{\Lambda\}) \tag{3.1}$$

with the following property

$$(\forall u \in X^*)(\forall w \in X)(\exists z \in Y)[\bar{f}(uw) = \bar{f}(u)z] \wedge f(\Lambda) \in Y.$$

From this easily follows that

$$(\forall u \in X^*)[\lg(\bar{f}(u)) = \lg(z) + 1].$$

As is known, the functions $\bar{f}$ defined in (3.1) are in one-to-one correspondence with the functions

$$f: X^* \to Y \tag{3.2}$$

(if for all $u \in X^*$, $f(u)$ is the last symbol of $\bar{f}(u)$).

In the following — unless otherwise stated — by the word *automaton* we shall always mean a function f of the type (3.2) and the (not necessarily finite) non-void sets X and Y we shall take given.

As a generalization of right and left compatible partitions of the semigroup $\mathbf{A}$, we formulate the following

Definition 2. If $r \in \mathscr{C}_{\Omega R}\mathbf{A}$, the partition p is a *right compatible partition on (the set of classes)* $\mathbf{A}/r$ iff

$$(\forall x)(\forall Z_1, Z_2, W_1, W_2)[(x \in \mathbf{A} \wedge (Z_1, Z_2, W_1, W_2 \in \mathbf{A}/r) \wedge$$
$$\wedge (Z_1\{x\} \subset W_1) \wedge (Z_2\{x\} \subset W_2) \wedge \langle Z_1, Z_2\rangle \in p) \Rightarrow \langle W_1, W_2\rangle \in p].^2$$

The meaning of *left compatible partition on a partition* is analogous.

Remark. "Compatible partition on a compatible partition r" is an ordinary compatible partition on the factor semigroup $\mathbf{A}/r$.

Definition 3. Given a set $\mathbf{Z}$ and $r \in \mathscr{C}\mathbf{Z}$, we call the function

$$\text{nat}\, r: \mathbf{Z} \to \mathbf{Z}/r \tag{3.3}$$

which has the following property

$$(\forall x)[x \in Z \Rightarrow x \in (\text{nat}\, r)(x)],$$

the natural mapping belonging to the partition r.

---

[2] If a binary operation, written as multiplication is defined on a set $\mathbf{S}$, and $\mathbf{T}, \mathbf{U} \subset \mathbf{S}$, then $T \cdot U = TU \stackrel{\text{def}}{=} \{tu \mid t \in \mathbf{T} \wedge u \in \mathbf{U}\}$.

The composition (consecutive application) of two functions f and g we write in the form

$$g\circ f,\ (g\circ f)(x) \stackrel{\text{def}}{=} f(g(x)). \tag{3.4}$$

Definition 4. Given a semigroup **A**, $r\in\mathscr{C}_{\Omega R}\mathbf{A}$ and the set Y, we call the function $k:\mathbf{A}/r\to Y$ *right compatible-free* (in short **RCF**) iff

$$(\forall q, s)[(k = (\text{nat}\, q)\circ s) \Rightarrow q = 1_{\mathbf{A}/r}]$$

where q is a right compatible partition on $\mathbf{A}/r$ and $s:(\mathbf{A}/r)/q\to Y$ (s is uniquely defined by q) (see Definition 2, Definition 3, (3.3) and (3.4)).

The meaning of *left compatible-free* (**LCF**) is analogous. Iff above $r\in\mathscr{C}_{\Omega}\mathbf{A}$ and $q\in\mathscr{C}_{\Omega}(\mathbf{A}/r)$, we call the function k *homomorph-free* (in short **HF**).

For any function f, we define the following equivalence relation

$$f^0 \stackrel{\text{def}}{=} \{\langle x, y\rangle | \langle x, y\rangle\in\mathscr{D}(f)\wedge f(x) = f(y)\}. \tag{3.5}$$

Now we are ready to prove the following

Statement 1. Given a function $f:\mathbf{A}\to Y$ where **A** is a semigroup, the decomposition

$$f = (\text{nat}\, r)\circ k \quad (\text{where } r\in\mathscr{C}\mathbf{A})$$

(exists and) is unique if at least one of the following conditions holds

(i) $r\in\mathscr{C}_{\Omega R}\mathbf{A}$ and k is **RCF**,

(ii) $r\in\mathscr{C}_{\Omega L}\mathbf{A}$ and k is **LCF**,

(iii) $r\in\mathscr{C}_{\Omega}\mathbf{A}$ and k is **HF**

(see (3.3), (3.4) and Definition 4).

Proof

(i) Let r be the greatest ("roughest") right compatible refinement of f^0 (see (3.5)) which exists (and is unique) on the basis of Theorem 1′. If $f = (\text{nat}\,\bar{r})\circ k'$ is another decomposition, for which $\bar{r}\neq r$, then according to Theorem 1′, $\bar{r}\subset r$ and there is a right compatible partition $q\neq 1_{\mathbf{A}/r}$ on $\mathbf{A}/r$ (see Definition 2), for which $k' = (\text{nat}\, q)\circ k''$ (for some k'') i.e. k' is not **RCF**.

(ii) Quite analogous to case (i).

(iii) The argument needs only slight and obvious modifications on that of case (i).

Definition 5. We supply r and k (which we have introduced in Statement 1) with subscripts R, L and C according to cases (i), (ii) and (iii) in Statement 1 respectively and write

(i) $f = (\text{nat}\, r_{Rf})\circ k_{Rf},\ \text{nat}\, r_{Rf} \stackrel{\text{def}}{=} R_f,$

(ii) $f = (\text{nat}\, r_{Lf})\circ k_{Lf},\ \text{nat}\, r_{Lf} \stackrel{\text{def}}{=} L_f,$

(iii) $f = (\text{nat}\, r_{Cf})\circ k_{Cf},\ \text{nat}\, r_{Cf} \stackrel{\text{def}}{=} C_f.$

We call R_f, L_f and C_f the *greatest right compatible*, the *greatest left compatible* and the *greatest homomorphic component of (or contained in) f*, respectively, while r_{Cf} we call the *congruence relation of f*.

Remark. As a consequence of Theorem 1′, for any $f(:A \to Y)$

$$r_{Cf} \subset r_{Rf} \subset f^0 \quad \text{and} \quad r_{Cf} \subset r_{Lf} \subset f^0. \tag{3.6}$$

Corollary of Statement 1, part (i). For any equivalence class K of initially connected Moore automata, the elements of which induce the same automaton mapping $\bar{f}$ (see (3.1)) there is a (unique) automaton $\bar{A}$ in K, which is the state-homomorphic image of all members in K.

Proof. It easily follows from (3.1), (3.2) and part (i) of Statement 1) (cf. [3, Chapter 9], [4, 4. §], [6, § 1.11] and [7, § 3.1]).

Definition 6. For an automaton f, the factor-semigroup

$$\mathbf{S}_f \overset{\text{def}}{=} X^*/r_{Cf}$$

we call the *semigroup (characteristic semigroup) of f.*[3]

The usual way of defining the semigroups of automata is found in the following

Definition 7. If $M=\langle Q, X, \delta\rangle$ is an automaton without output (with state-set Q, input alphabet X and next-state function δ), the semigroup of M is

$$\mathbf{S}(M) \overset{\text{def}}{=} X^*/\varrho(M), \tag{3.7}$$

where $\varrho(M)$ is the congruence relation of M and

$$\varrho(M) \overset{\text{def}}{=} \{\langle x, y\rangle | x, y \in X^* \wedge (\forall q)[q \in Q \Rightarrow qx = qy]\}. \tag{3.8}$$

(It can easily be checked using (1.4) that indeed $\varrho(M) \in \mathscr{C}_\Omega X^*$.)

Remarks

(a) On the basis of Theorem 1′ (see (1.14) and the end of the proof of Theorem 1′, and (1.18) in the Remark at the end of Part 1) using the notations of Definition 5

$$\begin{aligned} r_{Rf} &= \{\langle x, y\rangle | x, y \in X^* \wedge (\forall a)[a \in X^* \Rightarrow f(xa) = f(ya)]\},\\ r_{Lf} &= \{\langle x, y\rangle | x, y \in X^* \wedge (\forall a)[a \in X^* \Rightarrow f(ax) = f(ay)]\},\\ r_{Cf} &= \{\langle x, y\rangle | x, y \in X^* \wedge (\forall a, b)[a, b \in X^* \wedge f(axb) = f(ayb)]\}. \end{aligned} \tag{3.9}$$

(b) (3.9) is a more explicit formulation of (3.6), and further we can write

$$\begin{aligned} r_{Cf} &= \{\langle x, y\rangle | \langle x, y\rangle \in r_{Rf} \wedge (\forall a)[a \in X^* \Rightarrow \langle ax, ay\rangle \in r_{Rf}]\},\\ r_{Cf} &= \{\langle x, y\rangle | \langle x, y\rangle \in r_{Lf} \wedge (\forall a)[a \in X^* \Rightarrow \langle xa, ya\rangle \in r_{Lf}]\}. \end{aligned} \tag{3.10}$$

[3] See (3.2) and our agreement following it; and Def. 5.

(c) From the Corollary of Statement 1, Definitions 6 and 7, and equations (3.9) and (3.10), easily follows that (if A corresponds to f)

$$\mathbf{S}(\bar{A}) = \mathbf{S}_f \tag{3.11}$$

and

$$\varrho(\bar{A}) = r_{cf}. \tag{3.12}$$

(d) If the state-set Q (of $\bar{A}$) is finite, then we can even deduce from equations (3.7) to (3.12) that $\mathbf{S}(\bar{A})$ is finite too. More generally, in the language of semigroups

Statement 2. If $\mathbf{A}$ is a semigroup, $r \in \mathscr{C}_{\Omega R}\mathbf{A}$ and $|\mathbf{A}/r| < \infty$, then there exists an $r' \subset r$ and $r' \in \mathscr{C}_{\Omega}\mathbf{A}$, for which $|\mathbf{A}/r'| < \infty$. (Analogous statement is true of $r \in \mathscr{C}_{\Omega L}\mathbf{A}$.)

Proof. Let (like (3.10))

$$r' \overset{\text{def}}{=} \{\langle x, y\rangle | \langle x, y\rangle \in r \wedge (\forall a)[a \in \mathbf{A} \Rightarrow \langle ax, ay\rangle \in r]\}, \tag{3.13}$$

from which we can see at once using (1.4) that $r' \in \mathscr{C}_{\Omega}\mathbf{A}$ (and evidently $r' \subset r$). To prove the finiteness of $\mathbf{A}/r'$, we rewrite (3.13) in the following way

$$r' \overset{\text{def}}{=} \{\langle x, y\rangle | \langle x, y\rangle \in r \wedge (\forall a, b)[\langle a, b\rangle \in r \Rightarrow \langle ax, by\rangle \in r]\}. \tag{3.14}$$

(3.14) $\Rightarrow$ (3.13) is obvious. (3.14) can be obtained from (3.13) by taking into account that $r \in \mathscr{C}_{\Omega R}\mathbf{A}$, so $\langle a, b\rangle \in r \Rightarrow \langle ay, by\rangle \in r$ and being $r \in \mathscr{T}\mathbf{A}$, $(\langle ax, ay\rangle \in r \wedge \langle ay, by\rangle \in r) \Rightarrow \Rightarrow \langle ax, by\rangle \in r$. Now, with each element $x \in \mathbf{A}$, we can associate a function

$$(\varphi_x : \mathbf{A}/r \rightarrow \mathbf{A}/r) \wedge (\forall C)[C \in \mathbf{A}/r \Rightarrow C\{x\} \subset \varphi_x(C)] \tag{3.15}$$

(this was hinted by F. Gécseg). With the functions of (3.15), an equivalent form of (3.14) is

$$r' \overset{\text{def}}{=} \{\langle x, y\rangle | \langle x, y\rangle \in r \wedge \varphi_x = \varphi_y\}.$$

By the definition of the functions φ_x (see (3.15))

$$\{\varphi_x | x \in \mathbf{A}\} \subset \{\varphi | \varphi : \mathbf{A}/r \rightarrow \mathbf{A}/r\} \overset{\text{def}}{=} F,$$

so r' can be obtained from r by splitting each class in $\mathbf{A}/r$ into not more than $|F|$ subclasses and therefore

$$|\mathbf{A}/r'| \leqq |\mathbf{A}/r| \cdot |F|. \tag{3.16}$$

Taking

$$|\mathbf{A}/r| \overset{\text{def}}{=} m < \infty,$$

then $|F| = m^m$ and from (3.16) we get

$$|\mathbf{A}/r'| \leqq m \cdot m^m = m^{m+1} < \infty. \tag{3.17}$$

Remarks

(a) (3.17) is also valid for m's of any cardinality, but only $m < \infty$ has practical significance.

(b) Several authors declare that "any semigroup is isomorphic to the semigroup of an automaton" (in the sense of Definition 7), but this is wrong: we must say "any *monoid*" instead of "any semigroup" and so the statement will already be true. This easily follows from (3.7) and (3.8), or more generally from the simple fact: every factor-semigroup of a monoid is again a monoid. The mistake in the "proof" of the former defective assertion, which uses the so-called *semigroup machine*

$$M_{\mathbf{A}} \stackrel{\text{def}}{=} \langle \mathbf{A}, \mathbf{A}, \delta \rangle, \quad (\forall s_1, s_2)[\delta(s_1, s_2) \stackrel{\text{def}}{=} s_1 s_2]$$

(where $\mathbf{A}$ is any semigroup) is that even if $\mathbf{A}$ has no identity element, $\mathbf{A}^*$ does have, when applying (3.8) to $M_{\mathbf{A}}$. We cannot even be sure of

$$\mathrm{S}(M_{\mathbf{A}}) = \mathbf{A}_{\mathrm{I}} \tag{3.18}$$

(for any semigroup $\mathbf{A}$, $\mathbf{A}_{\mathrm{I}} \stackrel{\text{def}}{=} \mathbf{A}$, if $\mathbf{A}$ is a monoid and if not, then $\mathbf{A}_{\mathrm{I}} \stackrel{\text{def}}{=}$ "the monoid which we get by attaching to $\mathbf{A}$ an external unit element"), because if $\mathbf{A}$ is not a monoid, then it can well have *right uniform* elements. The notion of right uniform elements we introduce in the following

Definition 8. In the semigroup $\mathbf{A}$, the elements c and c' are said *right uniform* iff

$$(\forall x)[x \in A \Rightarrow xc = xc']$$

and the relation of right uniformity in the semigroup $\mathbf{A}$ we denote with $u_R(\mathbf{A})$.

The meaning of *left uniformity* is analogous and the notation for the corresponding relation is $u_L(\mathbf{A})$. At last, the relation $u(\mathbf{A}) \stackrel{\text{def}}{=} u_R(\mathbf{A}) \cap u_L(\mathbf{A})$ we call the relation of *uniformity* on $\mathbf{A}$.

Remark. Evidently $u_R(\mathbf{A})$, $u_L(\mathbf{A})$, $u(\mathbf{A}) \in \mathscr{C}_\Omega \mathbf{A}$. As an example, suppose $\mathbf{A}$ is right (left) quasi-trivial (see Def. 1), then $u_R(\mathbf{A}) = \mathbf{A}^2_{1R} \cup f^0_{R\mathbf{A}}$ $(u_L(\mathbf{A}) = \mathbf{A}^2_{1L} \cup f^0_{L\mathbf{A}})$ (see (3.5)). In this case $u(\mathbf{A}) \neq 1_{\mathbf{A}}$ iff. $\mathbf{A}_{1R} = \emptyset$ $(\mathbf{A}_{1L} = \emptyset)$ and $f^0_{R\mathbf{A}} = 1_{\mathbf{A}}$ $(f^0_{L\mathbf{A}} = 1_{\mathbf{A}})$, so there exist $\mathbf{A}$'s for which $u_R(\mathbf{A}) \neq u(\mathbf{A})$ $(u_L(\mathbf{A}) = u(\mathbf{A}))$.

A trivial example for uniform elements is the case when A is strongly quasi-trivial and (2.1) (iii) is valid (Def. 1). A less trivial example is the following: take an arbitrary semigroup $\mathbf{A}_0$ and choose a $c \in \mathbf{A}_0$ and let $c' \notin \mathbf{A}_0$, $\mathbf{A} \stackrel{\text{def}}{=} \mathbf{A}_0 \cup \{c'\}$. If we define the operations in $\mathbf{A}$ so

$$(\forall x, y \in \mathbf{A})[x, y \in \mathbf{A}_0 \Rightarrow (xy(\text{in } \mathbf{A}) = xy(\text{in } \mathbf{A}_0) \wedge c'x(\text{in } \mathbf{A}) =$$

$$= cx(\text{in } \mathbf{A}_0) \wedge xc'(\text{in } \mathbf{A}) = xc(\text{in } \mathbf{A}_0) \wedge c'c'(\text{in } \mathbf{A}) = cc(\text{in } \mathbf{A}_0))],$$

then $c \equiv c' \pmod{u(\mathbf{A})}$. Of course, by this method an unbounded number of uniform elements can be achieved. (If, furthermore, we randomly select some pairs $\langle x, y \rangle$ for which $xy = c$ (in $\mathbf{A}$) and change their result into c', then $\mathbf{A}$ will remain a semigroup and c' will play a more active role).

Now if $\mathbf{A}$ has right uniform elements, then (3.18) will not hold, because when forming $\mathrm{S}(M_{\mathbf{A}})$ according to (3.8) and (3.7), the right uniform elements of $\mathbf{A}$ will "coincide" in $\mathrm{S}(M_{\mathbf{A}})$. This can be expressed in the following

Fact. For any semigroup $\mathbf{A}$, $\mathrm{S}(M_{\mathbf{A}}) \cong (\mathbf{A}/u_R(\mathbf{A}))_{\mathrm{I}}$, and $\mathrm{S}(M_{\mathbf{A}}) \cong \mathbf{A}$ iff $\mathbf{A}$ is a monoid (see Def.'s 7 and 8).

Proof. Easy from Def.'s 7 and 8.

Now, let us come to the question of the simulation of automata by each other.

We say that the automaton f *can simulate* (in short: simulates) automaton f' (both f and f' correspond to (3.2)), iff there are suitable functions h and p, for which

$$f' = h \circ f \circ p, \tag{3.19}$$

where (3.19) we interpret in the sense of (3.4). Here

$$f: X^* \to Y \quad \text{and} \quad f': X_1^* \to Y.$$

A glance at (3.1) and (3.2) convinces us that in (3.19)

$$h: X^* \simeq X_1^*$$

("$\simeq$" and "$\cong$" are the usual symbols for denoting homorphic and isomorphic mappings respectively).

First we prove that the possibility of simulation depends essentially on the semigroups of the automata in question, and is independent of the input alphabet

Theorem 3. Let $f: X_f^* \to Y$ and $g: X_g^* \to Y$ two automata, $i: \mathbf{S}_f \cong \mathbf{S}_g$ and

$$k_{C_f} = i \circ k_{C_g}. \tag{3.20}$$

Then f and g can simulate each other.[4]

Proof. It is enough to prove, that g can simulate f. (In the following proof, the definitions, relations etc. mentioned in footnote 4, will be widely used without further explanation.)

Let

$$h_1: X_f \to X_g^* \tag{3.21}$$

be such that

$$(\forall x \in X_f)[h_1(x) \in (C_f \circ i)(x)]. \tag{3.22}$$

From (3.21) easily follows, that h_1 can be uniquely extended into a homomorphism

$$h: X_f^* \to X_g^*,$$

for which automatically $h(\Lambda) = \Lambda$ (otherwise the reader is likely to know the verification of the existence and uniqueness of h, from the theory of free semigroups).

As as consequence of (3.22), it can easily be seen that

$$(\forall w \in X_f^*)[h(w) \in (C_f \circ i)(w)]. \tag{3.23}$$

(It is usual also to require from h_1, that for every $x \in X_f$, $\lg(h_1(x))$ is the least possible, but this is not necessary for our purposes.)

(3.23) eans that

$$(\forall w \in X_f^*)[(h \circ C_g)(w) = (C_f \circ i)(w)],$$

i.e.

$$h \circ C_g = C_f \circ i.$$

[4] See Def.'s 5, 6, equation (3.19) and convention (3.4).

Multiplying this equation with equation

$$k_{Cg} \circ 1_Y = k_{Cg},$$

we get

$$h \circ (C_g \circ k_{Cg}) \circ 1_Y = C_f \circ (i \circ k_{Cg})$$

and taking into account (3.20)

$$h \circ g \circ 1_Y = f,$$

i.e. g can simulate f.

*COMPUTING CENTER OF THE HUNGARIAN MINISTRY OF HEAVY INDUSTRIES, BUDAPEST

**RESEARCH GROUP ON MATHEMATICAL LOGIC AND THE THEORY OF AUTOMATA OF THE HUNGARIAN ACADEMY OF SCIENCES, SZEGED

Замечания о максимальных конгруенциях, автоматах и смежных темах

Статья состоит из трех частей.

В 1-ой части авторы занимаются следующим обобщением: для данной сверх некоторой полугруппы A эквивалентностной реляции однозначно существует уточнение по максимальной конгруенции, доказано, что вместо эквивалентности и для более обобщенных реляций однозначно существуют максимальные уточнения более общего типа, чем конгруенция.

Во 2-ой части показывается, что некоторая возможная инвертность результатов 1-ой части взаимнооднозначно соответствует определенной специальной операционной структуре полугруппы A.

В 3-ей части исследуются вопросы, связанные с полугруппами автоматов Мура и их симуляцией, исходя из эквивалентностных и конгруентных реляций, выходящих из трансформаций авт омата Мура, и используя результаты 1-ой части.

References

[1a] Cohn, P. M., *Universal algebra*, Harper and Row, New York, Evanston and London, 1965.
[1b] Kon, P., *Un'iversalnaya alghebra* (Russian edition), Mir, Moscow, 1968.
[2] Schmidt, E. T., *Kongruenzrelationen Algebraischer Strukturen*, VEB Deutscher Verlag der Wissenschaften, Berlin, GDR, 1969.
[3] Harrison, M. A., *Introduction to switching and automata theory*, McGraw-Hill, New York, London, Sydney, 1965.
[4] Ádám, A., Automata-leképezések, félcsoportok, automaták (in Hungarian, with English and Russian abstracts), *Matematikai Lapok*, Budapest, v. 19, 1968, pp. 327—343.
[5] Arbib, M. A. (editor), *Algebraic theory of machines, languages and semigroups* (Collection of Papers), Academic Press, New York, London, 1968.
[6] Gécseg, F., I. Peák, *Algebraic theory of automata*, Akadémiai Kiadó, Budapest, Hungary, 1972.
[7] Arbib, M. A., *Theories of abstract automata*, Prentice-Hall, Inc., Englewood Cliffs, N. J., 1969.

(Received June 5, 1972)