Opuscula Mathematica • Vol. 28 • No. 2 • 2008

Wit Foryś

ON A COMPLETE LATTICE OF RETRACTS OF A FREE MONOID GENERATED BY THREE ELEMENTS

Abstract. We prove that the family of retracts of a free monoid generated by three elements, partially ordered with respect to the inclusion, is a complete lattice.

Keywords: free monoid, retract, lattice, complete lattice.

Mathematics Subject Classification: 68Q05.

1. INTRODUCTION

We consider a family of retracts of a free monoid partially ordered with respect to the inclusion. It is known fact that the family is a lattice if the considerations are limited to retracts of a free monoid A^* generated by at most three-element alphabet A. The paper sharpens this result a bit. Namely, it is proven that the family is a complete lattice. The presented proofs are independent and different from the former ones connected with the lattice property and due to T. Head [6].

If A has at least four elements then some counterexamples for the lattice property may be constructed [6].

2. BASIC NOTIONS AND DEFINITIONS

Let A be any finite set and let A^* denote a free monoid generated by A^*. A retraction $r : A^* \to A^*$ is a morphism for which $r \circ r = r$. A retract of A^* is the image of A^* by a retraction.

Definition 1. *A word $w \in A^*$ is called a key-word if there is at least one letter in A that occurs exactly once in w. A letter that occurs once in a key-word w is called a key of w. A set $C \subset A^*$ of key-words is called key-code if there exists an injection $i : C \to A$ such that:*

1) *for any $w \in C$, $i(w)$ is a key of w;*
2) *the letter $i(w)$ occurs in no word of C other than w itself.*

The following characterization of retracts due to T. Head [6] is basic for our research.

Theorem 2 ([6]). *$R \subset A^*$ is a retract of A^* if and only if $R = C^*$, where C is a key-code.*

In the sequel, we use the following notation. Let C_1, C_2 denote key-codes of retracts R_1, R_2, respectively. The intersection of the retracts $R_1 \cap R_2 = C_1^* \cap C_2^*$ is a free submonoid of A^*. Denote by C the basis of the submonoid (minimal set of generators). Any word in $R_1 \cap R_2$ has two factorizations, one in key-words of C_1 and the second in key-words of C_2. In general, C is not a key-code [4] and this was a reason for starting a research of semiretracts [1, 3].

Definition 3. *A code $C \subset A^*$ is an infix code if for all $u, v, w \in A^*$, $v \in C$ and $uvw \in C$ implies that $u = w = 1$.*

Definition 4. *A code $C \subset A^*$ is comma free if for all $v \in C^*$, $u, w \in A^*$ and $uvw \in C^*$ implies that $u, w \in C^*$.*

The theorem of Tarski and Knaster [7] is essential for our final result.

Theorem 5 (Tarski, Knaster). *Let D be a complete lattice and $f : D \to D$ a monotonic function. Then a set $Fp\, f = \{x \in D : f(x) = x\}$ of all fixed points of f forms a complete sublattice of D.*

In what follows, we limit our considerations to retracts of a free monoid generated by exactly three-element alphabet A, denoted in the sequel with A_3. Let us denote by $RET\, A_3^*$ the family of all retracts of A_3^* partially ordered with respect to the inclusion.

3. RESULTS

Let A be a finite or infinite alphabet and let us denote by $(RET\ A^*, \subset)$ the family of all retracts of A^* partially ordered with respect to the inclusion. Define for any $X \subset A^*$ the family of retracts

$$L_X = \{R \subset A^* : X \subset R,\ R \text{ is a retract of } A^*\}.$$

L_X is not empty because $A^* \in L_X$ for any X and as a subset of $RET\ A^*$ is partially ordered too.

Lemma 6. *For any alphabet A, whether finite or infinite, there exists a mapping $\rho : \wp(A^*) \to \wp(A^*)$ which maps any subset X of A^* into a retract R_X such that $X \subset R_X$ and R_X is minimal.*

Proof. First, we establish the fact that there exists a minimal element in $(L_X, \subset)$ for any X of A^*. Let us fix a descending chain of retracts $\{R_i\}_{i \in I}$ in $(L_X, \subset)$. We claim that there exists in L_X a lower bound of the chain. For this purpose, let us consider the intersection $\bigcap_I R_i$. If the chain $\{R_i\}_{i \in I}$ is finite our claim is trivially true. Thus let us assume that $\{R_i\}_{i \in I}$ is infinite and let B denote the base of the submonoid

$\bigcap_I R_i$. First observe that for any $w \in B$ there exists a retract R_k in the considered chain such that w is an element of the key-code that generates R_k. To justify this observation, one can take into account the fact that w has finitely many factorizations into subwords. w is then an element of the key-code of any R_l for $l \geq k$. Now let us assume, for the contrary, that B is not a key-code. Hence there exists a word $w \in B$ such that for any $a \in A$ that occurs exactly once in w there exists $v \in B$ such that the letter a occurs in v at least once. For any a such that $\sharp_a w = 1$, let us fix exactly one word $v \in B$ which has the above property. Denote by B_w the set of all v chosen in that way. B_w is non-empty and finite. Observe that there exists $k \in I$ such that B_w is included in the key-code of R_i for any $i \geq k$. This contradicts the fact that R_i is a retract for $i \in I$. Finally, B is a key-code, and from the Zorn-Kuratowski Lemma it follows that for each $X \subset A^*$ there exists a minimal element in L_X. Now define a mapping $\rho : \wp(A^*) \to \wp(A^*)$ putting for every $X \in \wp(A^*)$ $\rho(X) = R_X$ where R_X is minimal for X (Axiom of Choice). Hence the proof is finished. □

Now we limit the considerations to the three-element alphabet A_3. Let C be the base (minimal set of generators) of a submonoid obtained as the intersection of retracts $R_1 \cap R_2 = C_1^* \cap C_2^*$ where $R_1, R_2 \in RET\ A_3^*$. We will analyze all possible forms of C (according to the form of C_1 and C_2) and conclude that, in any case, C is a key-code. It is not difficult to observe that among the all possibilities for C_1 and C_2 one case is not obvious only, and needs to be considered. Namely, for $C_1 = \{u_1, u_2\}$, and $C_2 = \{v_1, v_2\}$, where u_1 and v_1 have the same key, say a, and keys of u_2 and v_2 are different and equal to b and c, respectively. According to the symmetry of C_1 and C_2, it is sufficient to consider the following two cases:

1) $u_1 = v_1 = a$, u_2 without restrictions,
2) $|u_1| > 1$, u_2 without restrictions.

Assumptions of the first case imply that the base C of the semiretract $R_1 \cap R_2$ is equal to $\{u_1\}$ or $\{u_1, u_2\}$. Hence, the intersection $R_1 \cap R_2$ is in fact a retract according to Theorem 2.

We start to consider the second case with the following lemmas.

Lemma 7. *If $w = \ldots u_1^k \ldots \in C$, then $k = 1$.*

Proof. Let us assume for the contrary that $k > 1$. It means that between the key a in the first (from the left) u_1 and the key a in the second u_1 there exists at least one c. It implies that $v_2 = c$ and leads to the conclusion that w can be represented as a catenation of two words in C^+, which contradicts the code property of C. □

Lemma 8. *There is no word of the form $\ldots u_1 u_2^k u_1 \ldots$ in C for any $k \geq 0$.*

Proof. In view of the above lemma, we can consider the case $k \geq 1$. Observe the following property of words in C. Any word in C is expressible as catenation of elements of C_1 and elements of C_2 as well. Imagine that w is expressed in two lines: the upper line uses words from C_1, lower one from C_2. If u_1 occurs in the upper line, then it enforces the occurrence of v_1 in the lower line. The condition $|u_1| > 1$ enforces the occurrence of v_2 to the left or to the right of v_1. If v_2 is a one letter word, it is possible to fulfil the lower line to obtain a suffix equal to u_1 exactly. It also means that a prefix of w which ends at the first u_1 is in C^+, which contradicts the code property of C and the fact that $w \in C$. If v_2 is not a one letter word, the occurrence of u_2 in the upper line is enforced. While repeating the above reasoning, notice that the process of enforcing finishes by the conclusion that w is in C. Now consider the first and the second occurrence (from the left) of u_1 in w. Based on the described property, observe that a word x, the prefix of w which ends at the first u_1, is also a suffix of $u_2^k u_1$. The same word x composed of v's occurs two times in the lower line at the same positions. This implies that w can be factorized into words from C^+; again a contradiction with the code property of C. □

The following result is the direct corollary of the above lemmas.

Corollary 9. *If $w \in C$, then $w = u_2^k u_1 u_2^l$ for some $k, l \geq 0$ or $w = u_2$.*

Let us remind the following simple and easily proved fact.

Fact 10 ([1]). *The base C of $R_1 \cap R_2$ is an infix code and comma free code.*

The above considerations allow us to formulate

Lemma 11. *Let C be the base of a semiretract $R_1 \cap R_2$. C has one of the following forms:*

$$\{u_1\}, \{u_2\}, \{u_1, u_2\}, \{u_2^k u_1 u_2^l\}$$

where $k, l \geq 0$.

Proof. The statement of the lemma follows from the preceding lemmas and Fact 10. Notice that the property of words in C formulated and applied in Lemma 8 implies that if $u_2^k u_1 u_2^l$ is in C with $k + l \geq 1$, then $C = \{u_2^k u_1 u_2^l\}$. □

We summarize the above results in the following lemma.

Lemma 12. *The intersection of two retracts $R_1, R_2 \in RET\, A_3^*$ is a retract $R_1 \cap R_2$ of A^*. The cardinality of the base (key code) C of the retract $R_1 \cap R_2$ is at most three.*

Finally we formulate a theorem containing the main result of the paper.

Theorem 13. *The family $RET\, A_3^*$, partially ordered with respect to the inclusion, is a complete lattice.*

Proof. Consider the mapping $\rho : \wp(A^*) \to \wp(A^*)$ defined in Lemma 6. Let $X, Y \in \wp(A^*)$ be two non-empty sets, $X \subset Y$ and $\rho(X) = R_X$, $\rho(Y) = R_Y$. There is $R_Y \in L_X$ and $X \subset R_Y$. According to Lemma 12, $R_X \cap R_Y$ is a retract and obviously $R_X \cap R_Y$ is in L_X. The inclusions $X \subset R_X \cap R_Y \subset R_X$ imply $R_X \cap R_Y = R_X$ and finally $R_X \subset R_Y$. Hence we come to the conclusion that the mapping ρ is monotonic. Now from the Tarski-Knaster theorem it follows that $(Fp(\rho), \subset)$ is a complete lattice, where $Fp(\rho)$ denotes the family of all fixed points of the mapping ρ, that is the family of sets $S \in \wp(A^*)$ such that $\rho(S) = S$. The observation that $Fp(\rho) = RET\, A_3^*$ finishes the proof. □

REFERENCES

[1] J.A. Anderson, *Semiretracts of a free monoid,* Theoretical Computer Science, 134, 1994.

[2] J.A. Anderson, *The intersection of retracts of A^*,* Theoretical Computer Science, 237, 2000.

[3] J.A. Anderson, W. Foryś, *Regular languages and semiretracts,* Proccedings of Conference "Words, Languages and Combinatorics", Kyoto, 2000.

[4] J.A. Anderson, W. Foryś, T. Head, *Retracts and semiretracts of free monoids,* AMS Meeting (Proccedings), San Francisco, 1991.

[5] W. Foryś, T. Head, *The poset of retracts of a free monoid,* Intern. J. Computer Math., 37, 1990.

[6] T. Head, *Expanded subalphabets in the theories of languages and semigroups,* Intern. J. Computer Math., 12, 1982.

[7] A. Tarski, *A lattice-theoretic fixpoint theorem and its applications,* Pacific J. Math., 5, 1955.

Wit Foryś
forysw@ii.uj.edu.pl

PWSZ Nowy Sącz,
Institute of Economics,
ul. Jagiellońska 61, 33-300 Nowy Sącz, Poland

Received: November 5, 2006.
Revised: October 10, 2007.
Accepted: October 12, 2007.