BULL. AUSTRAL. MATH. SOC. 08A70,20M35,68Q70
VOL. 62 (2000) [471–477]

SYNTACTIC SEMIGROUPS AND GRAPH ALGEBRAS

A.V. KELAREV AND O.V. SOKRATOVA

We describe all directed graphs with graph algebras isomorphic to syntactic semigroups of languages.

The aim of this paper is to give a complete description of all graph algebras isomorphic to syntactic semigroups of languages. We consider languages which are subsets of the free semigroup A^+ over an alphabet A. It is well known that various important properties of languages can be characterised in terms of their syntactic semigroups (see [**3**, **10**]).

Graph algebras make it possible to apply methods of universal algebra to various problems of discrete mathematics and operations research. They have been investigated by several authors (see, for example, [**2**, **5**, **6**, **7**, **8**, **9**] for references). In this paper the word *graph* means a directed graph without multiple edges but possibly with loops. The *graph algebra* $\mathrm{Alg}(D)$ of a graph $D = (V, E)$ is the set $V \cup \{\infty\}$ endowed with multiplication defined by $xy = x$ if $(x, y) \in E$ and $xy = \infty$ otherwise, and $x\infty = \infty x = \infty\infty = \infty$, for all $x, y \in V$.

Let S be a semigroup with a subset T, and let $x \in S$. Then the *context* of x with respect to T is denoted by $\mathrm{Cont}_T(x)$ or $\mathrm{Cont}_{T,S}(x)$ and is defined as the set

$$\mathrm{Cont}_T(x) = \big\{(a, b) \in S^1 \times S^1 \mid axb \in T\big\}.$$

The syntactic semigroup $\mathrm{Syn}(L)$ of a language $L \subseteq A^+$ is the quotient of A^+ by the congruence σ_L

$$\sigma_L = \big\{(w_1, w_2) \in A^+ \times A^+ \mid \mathrm{Cont}_{L,A^+}(w_1) = \mathrm{Cont}_{L,A^+}(w_2)\big\}.$$

The *in-degree (out-degree)* of a vertex v of the graph $D = (V, E)$ is the number of vertices $u \in V$ such that $(u, v) \in E$ (respectively, $(v, u) \in E$). A graph $D = (V, E)$ is said to be *undirected* if $(x, y) \in E \Rightarrow (y, x) \in E$, for all $x, y \in V$. The graph D is said to be *connected* if its underlying undirected graph is connected. A maximal connected subgraph of a graph D is called a *connected component* of D.

Received 5th April, 2000

The sum $D_1 + D_2$ of graphs $D_1 = (V_1, E_1)$ and $D_2 = (V_2, E_2)$ is the graph with vertex set $V_1 \cup V_2$ and edge set $E_1 \cup E_2 \cup E_{1,2}$, where

$$E_{1,2} = \{(x, y) \mid x \in E_1,\ y \in E_2\}.$$

If D_1 and D_2 have no common vertices, then we say that the sum is *direct*, and we denote it by $D_1 \oplus D_2$. A *null graph* is a graph without edges. The null graph of order m is denoted by N_m. We assume that *complete graphs* contain all edges including loops. The complete graph of order n is denoted by K_n. The direct sum of these graphs will be denoted by

$$K_n^m = N_m \oplus K_n.$$

If G is a graph, then the set of all vertices (edges) of G is denoted by $V(G)$ (respectively, $E(G)$). Now we can state our main theorem.

Theorem 1. *Let $D = (V, E)$ be a directed graph. Then the following conditions are equivalent:*

(1) *the graph algebra* $\mathrm{Alg}(D)$ *is a syntactic semigroup of a language;*

(2) *the graph D satisfies the following properties:*

  (i) *D has at most one isolated vertex;*

  (ii) *D does not contain complete graphs with three vertices;*

  (iii) *D has no connected component with more than one vertex of zero in-degree;*

  (iv) *in each connected component C of D all vertices of nonzero in-degree induce a complete subgraph;*

  (v) *every vertex of zero in-degree is adjacent to all vertices of nonzero in-degree in its connected component.*

(3) *D has at most one isolated vertex, and every other connected component is isomorphic to one of the graphs K_1, K_2, K_1^1, K_2^1.*

Remark 2. There exist non-regular languages with syntactic semigroups isomorphic to graph algebras. Indeed, a language is regular if and only if its syntactic semigroup is finite.

The following known lemma is used in our proof.

Lemma 3. [1, Proposition 2.3] *A semigroup S is the syntactic semigroup of some language if and only if S contains a subset T such that every two distinct elements of S have different contexts with respect to T.*

A *path* of length n in a graph $D = (V, E)$ is a path in the underlying undirected graph of D, that is, a sequence of vertices $v_0, v_1, \ldots, v_n$ such that $(v_i, v_{i+1}) \in E$ or $(v_{i+1}, v_i) \in E$, for $i = 0, 1, \ldots, n-1$.

PROPOSITION 4. *For any directed graph $D = (V, E)$, the following conditions are equivalent:*

(1) *the graph algebra* $\mathrm{Alg}(D)$ *is associative;*

(2) *for all $(x, y) \in E$ and $z \in V$,*

$$(x, z) \in E \Leftrightarrow (y, z) \in E;$$

(3) *each connected component of D is isomorphic to N_1, or a complete graph, or a direct sum of a null graph and a complete graph.*

PROOF: (1) $\Rightarrow$ (2): Consider any $(x, y) \in E$ and $z \in V$. Since $xy = x$, and $(xy)z = x(yz)$ by (1), we get

$$(x, z) \in E \Leftrightarrow (xy)z = x \Leftrightarrow x(yz) = x \Leftrightarrow (y, z) \in E.$$

(2) $\Rightarrow$ (1): Take any three elements x, y and $z \in V$. If $(x, y) \notin E$, then $(xy)z = \infty = x(yz)$. If, however, $(x, y) \in E$, then $(xy)z = xz$, and so by (2) we get

$$(xy)z = x \Leftrightarrow (x, z) \in E \Leftrightarrow (y, z) \in E \Leftrightarrow x(yz) = x.$$

Thus (1) holds in both the cases.

(2) $\Rightarrow$ (3): Let C be a connected component of D with at least two vertices. Denote by K the subgraph induced by all vertices of C with nonzero in-degrees.

Pick any vertex $v \in V(K)$. There exists an edge (u, v), where $u \in V(C)$. By (2),

$$(u, v) \in E \Leftrightarrow (v, v) \in E.$$

Therefore all vertices in K have loops.

Consider an edge (u, v), where $u, v \in V(K)$. In view of (2),

$$(u, u) \in E \Leftrightarrow (v, u).$$

Since u has a loop, $(v, u) \in E$. Thus the subgraph K is undirected.

In order to prove (3) it suffices to verify that C contains edges (x, y), for all vertices $x \in V(C)$ and $y \in V(K)$.

If $x = y$, then $(x, y) \in E$, because K contains all loops. Let y, z_1, z_2, $\ldots$, z_{n-1}, x be the shortest path from y to x, where $n > 0$.

By the definition of a path $(z_1, y) \in E$ or $(y, z_1) \in E$. If $(y, z_1) \in E$, then $z_1 \in K$; whence $(z_1, y) \in E$, because K is undirected. Therefore (z_1, y) is an edge in any case.

Suppose that $n > 1$. If $(z_2, z_1) \in E$, then $(z_2, y) \in E$, because (2) shows that E is a transitive relation. If, however, $(z_1, z_2) \in E$, then $(z_2, y) \in E$ by (2), again. Thus $y, z_2, \ldots, z_{n-1}, x$ is a shorter path from y to x. This contradicts the minimality of n and shows that $n = 1$.

Therefore $x = z_1$, and so $(x,y) = (z_1,y) \in E$, as claimed. It follows that K is complete, and moreover (3) holds.

(3) $\Rightarrow$ (2): Take any $(x,y) \in E$ and $z \in V$. Let C by the connected component of D containing x and y. Clearly, we may assume that $C \neq N_1$. Denote by K the complete subgraph of C mentioned in (3). Since K contains all vertices of nonzero in-degree, we get

$$(x,z) \in E \Leftrightarrow z \in K \Leftrightarrow (y,z) \in E,$$

that is (2) holds. ∎

Proposition 4 immediately gives us the following corollary:

COROLLARY 5. *If the graph algebra of a directed graph D is associative, then for all $x,y,z \in V$, the following conditions hold:*

(i) $(x,x),(x,y) \in E \Rightarrow (y,x) \in E$;

(ii) $(x,y) \in E \Rightarrow (y,y) \in E$.

PROOF OF THEOREM 1: (1) $\Rightarrow$ (2): Suppose that the graph algebra $\mathrm{Alg}(D)$ of D is a syntactic semigroup. Then $\mathrm{Alg}(D)$ is associate, and by Proposition 4 every connected component C of D with more than one vertex has a complete subgraph K_C such that each vertex of $C \setminus K$ has in-degree zero and is adjacent to all vertices in K_C. This establishes conditions (iv) and (v).

Further, by Lemma 3 there exists $T \subsetneq \mathrm{Alg}(D)$ such that $\mathrm{Cont}_T(x) \neq \mathrm{Cont}_T(y)$, for all $x \neq y \in \mathrm{Alg}(D)$. Note that the set $\mathrm{Alg}(D) \setminus T$ also satisfies this condition. Therefore we may always assume that $\infty \notin T$.

First, we prove the auxiliary fact that all vertices of in-degree zero belong to T. Indeed, since $\infty \notin T$, the context of each vertex with in-degree zero which is not in T, equals $\emptyset = \mathrm{Cont}_T(\infty)$, and so all these vertices belong to T.

Second, we verify that D has not more than one isolated vertex. All isolated vertices belong to T, as we have seen above. Therefore the context of every isolated vertex consists of one and the same pair $(1,1)$. Therefore condition (i) follows.

Next, we show that the maximal complete subgraph of each connected component of D has at most two vertices. Take a connected component C of D with more than one vertex, and let K_C be the complete subgraph of C induced by vertices of nonzero in-degree.

Since $\infty \notin T$, all elements of $K_C \cap T$ have the same context $\big(V(C) \cap T\big)^1 \times V(K_C)^1$; whence $\big|V(K_C) \cap T\big| \leqslant 1$. Similarly, the context of every element in $V(K_C) \setminus T$ equals $\big(V(C) \cap T\big) \times V(C)^1$, and hence $\big|V(K_C) \setminus T\big| \leqslant 1$. Thus K_C has at most two vertices, that is, condition (ii) holds.

Finally, consider any connected component C of D with more than one vertex, and denote by K_C the complete subgraph induced in C by all vertices of nonzero in-degree.

Since every vertex of C with in-degree zero belongs to T, its context is equal to $\{1\} \times V(K_C)^1$. Therefore each connected component of D has at most one vertex with in-degree zero. Thus condition (iii) holds, too.

(2) $\Rightarrow$ (3): Let us take a connected component C of D. If C has one vertex, then it is isomorphic to N_1 or K_1, and so by (i) we may assume that C has more than one vertex. Denote by K_C the subgraph induced in C by all vertices of nonzero in-degree. By (iv), K_C is a complete graph. It follows from (ii) that K_C is isomorphic to K_1 or K_2. In view of (iii), we have $|V(C \setminus K_C)| \leqslant 1$. If $|V(C \setminus K_C)| = 1$, then by (v), the only vertex of $C \setminus C_K$ is adjacent to all vertices of K_C. Therefore C is isomorphic to K_1, or K_2, or K_1^1, or K_2^1.

(3) $\Rightarrow$ (1): Suppose that D has at most one isolated vertex and every other connected component is isomorphic to K_1, K_2, K_1^1, or K_2^1. Then the graph algebra $\mathrm{Alg}(D)$ is a semigroup by Proposition 4, and so we can use Lemma 3.

Let us define a subset T of $\mathrm{Alg}(D)$. We include all vertices of zero in-degree in T. For each connected component C of D, we choose one vertex in the maximal connected subgraph K_C of C, and put it in T.

Consider any two distinct elements x, y of $\mathrm{Alg}(D)$. We are going to show that $\mathrm{Cont}_T(x) \neq \mathrm{Cont}_T(y)$.

If D has an isolated vertex v, then v is the only element of $\mathrm{Alg}(D)$ with $\mathrm{Cont}_T(v) = \{(1,1)\}$. Besides, ∞ is the only element of $\mathrm{Alg}(D)$ with empty context with regard to T. Therefore we may assume that neither x nor y is an isolated vertex of D, and $x, y \neq \infty$.

Suppose that x and y belong to connected components C and C' of D, respectively. Let K_C and $K_{C'}$ be the maximal complete subgraphs of these components. Denote by ℓ the element of T chosen in K_C. The following four cases are possible:

CASE 1. $x, y \in K_C$. Then only one of them, say x, belongs to T. Hence $(1,1) \in \mathrm{Cont}_T(x) \setminus \mathrm{Cont}_T(y)$.

CASE 2. $x \in K_C$, $y \notin K_C$. Since K_C is a complete graph, we get $(\ell, x) \in E$. Hence $\ell x = \ell \in T$. Since $y \notin K_C$, there is no edge (ℓ, y), and so $\ell y = \infty \notin T$. Therefore $(\ell, 1) \in \mathrm{Cont}_T(x) \setminus \mathrm{Cont}_T(y)$.

CASE 3. $x \in C \setminus K_C$, $y \in K_{C'}$. This case is similar to Case 2.

CASE 4. $x \in C \setminus K_C$, $y \in C' \setminus K_{C'}$. Then x is adjacent to all vertices of K_C, and so $(x, \ell) \in E$. It follows that $(1, \ell) \in \mathrm{Cont}_T(x) \setminus \mathrm{Cont}_T(y)$.

Thus $\mathrm{Cont}_T(x) \neq \mathrm{Cont}_T(y)$ in all cases. By Lemma 3, the graph algebra of D is a syntactic semigroup. ∎

COROLLARY 6. *Let $D = (V, E)$ be an undirected graph. Then the graph algebra $\mathrm{Alg}(D)$ is a syntactic semigroup if and only if D has at most one isolated vertex and all other connected components of D are complete graphs with not more than two vertices.*

A groupoid S with zero 0 is said to be a 0-*direct union* of its subgroupoids S_i, where

$i \in I$, if and only if $S = \bigcup_{i \in I} S_i$ and $S_i S_j = S_i \cap S_j = 0$, for all $i \neq j$.

Lemma 7. *A 0-direct union of syntactic semigroups is a syntactic semigroup.*

Proof: Let $S = \bigcup_{i \in I} S_i$ be a 0-direct union of semigroups S_i, where $i \in I$. Suppose that all the semigroups S_i are syntactic. Then by Lemma 3 every semigroup S_i contains a subset T_i such that every two distinct elements of S_i have different contexts with respect to T_i. Note that we can always assume that $0 \notin T$, for all $i \in I$. (Indeed, if $0 \in T_i$ for some $i \in I$, we can take the subset $S_i \setminus T_i$ instead of T.) Put $T = \bigcup_{i \in I} S_i$. The context of 0 with regard to T is empty, as well as with regard to all T_i, and for any element $x \in S_i$, $x \neq 0$, $\mathrm{Cont}_{T_i}(x) = \mathrm{Cont}_T(x)$. Therefore $\mathrm{Cont}_T(x) \neq \mathrm{Cont}_T(y)$ whenever $x \neq y$. ∎

Proposition 8. *The graph algebra of an undirected graph is a syntactic semigroup if and only if it is isomorphic to a 0-direct union of subdirectly irreducible semigroups.*

Proof: It is easily seen that $\mathrm{Alg}(D)$ is a 0-direct union of the graph algebras of all connected components of the graph D.

Suppose that $\mathrm{Alg}(D)$ is a syntactic semigroup. Theorem 1 implies that D has at most one isolated vertex, and every other connected component is isomorphic to one of the graphs K_1, K_2. Obviously, in each of these components at most one pair of distinct vertices have the same neighbourhoods. [4, Theorem 1] tells us that the graph algebra of an undirected graph is subdirectly irreducible if and only if it is connected and not more than one pair of distinct vertices have identical neighbourhoods. Hence all connected components of D are subdirectly irreducible. Therefore $\mathrm{Alg}(D)$ is isomorphic to a 0-direct union of subdirectly irreducible semigroups.

Conversely, suppose that $\mathrm{Alg}(D)$ is isomorphic to a 0-direct union of subdirectly irreducible semigroups. It is known and easy to verify that every subdirectly irreducible semigroup is a syntactic semigroup of some language (because $\bigcap_{s \in S} \sigma_{\{s\}}$ is the identity relation on S). Therefore all these subsemigroups are syntactic. According to Lemma 7, a 0-direct union of syntactic semigroups is syntactic, as well. ∎

The following example shows that Proposition 8 does not generalise to directed graphs.

Example 9. The graph algebra of the graph K_2^1 is a syntactic semigroup by Theorem 1, but it cannot be represented as a 0-direct union of subdirectly irreducible semigroups. Indeed, first note that $\mathrm{Alg}(K_2^1)$ is indecomposable into 0-direct unions of proper subsemigroups. Second, consider the following two principal congruences $\Theta(b,c)$ and $\Theta(a,\infty)$ on $\mathrm{Alg}(K_2^1)$ with equivalence classes

$$\Theta(b,c) = \big\{\{b,c\},\{a\},\{\infty\}\big\},$$
$$\Theta(a,\infty) = \big\{\{a,\infty\},\{b\},\{c\}\big\},$$

where a is the vertex of in-degree zero in $\mathrm{Alg}(K_2^1)$, and b, c are the other vertices of $\mathrm{Alg}(K_2^1)$. Since $\Theta(b,c) \cap \Theta(a,\infty)$ is the equality congruence, we see that $\mathrm{Alg}(K_2^1)$ is subdirectly reducible.

References

[1] S. Eilenberg, *Automata, languages and machines*, Vol. A, B (Academic Press, New York, 1976).

[2] E.W. Kiss, R. Pöschel and P. Pröhle, 'Subvarieties of varieties generated by graph algebras', *Acta Sci. Math.* **54** (1990), 57–75.

[3] G. Lallement, *Semigroups and combinatorial applications* (J. Wiley, New York, 1979).

[4] G.F. McNulty and C. Shallon, 'Inherently nonfinitely based finite algebras', in *Universal algebra and lattice theory (Puebla, 1982)*, Lecture Notes Math. **1004**, 1983, **pp.** 206–231.

[5] S. Oates-Williams, 'Murski's algebra does not satisfy min', *Bull. Austral. Math.* **22** (1980), 199–203.

[6] S. Oates-Williams, 'Graphs and universal algebra', in *Combinatorial Mathematics III (Geelong 1980)*, Lecture Notes Math. **884** (Springer-Verlag, Berlin, Heidelberg, New York, 1981), **pp.** 351–354.

[7] S. Oates-Williams, 'On the variety generated by Murski's algebra', *Algebra Universalis* **18** (1984), 175–177.

[8] R. Pöschel, 'The equational logic for graph algebras', *Z. Math. Logik Grundlag Math.* **35** (1989), 273–282.

[9] R. Pöschel and W. Wessel, 'Classes of graphs definable by graph algebra identities or quasi-identities', *Comment. Math. Univ. Carolin.* **28** (1987), 581–592.

[10] G. Rozenberg, A. Salomaa (Eds.), *Handbook of formal languages, Vol. 1, 2, 3,* (Springer-Verlag, Berlin, Heidelberg, New York, 1997).

Department of Mathematics
University of Tasmania
G.P.O. Box 252-37, Hobart
Tasmania 7001, Australia
e-mail: Andrei.Kelarev@utas.edu.au

Department of Mathematics
Tartu University
J.Liivi 2, 50409 Tartu
Estonia
e-mail: olga@cs.ut.ee

https://doi.org/10.1017/S0004972700018992 Published online by Cambridge University Press