Optimization of classifiers for data mining based on combinatorial semigroups

A.V. Kelarev, J.L. Yearwood, P.W. Watters

School of Science, Information Technology and Engineering
University of Ballarat, P.O. Box 663, Ballarat, Victoria 3353, Australia
{a.kelarev,j.yearwood,p.watters}@ballarat.edu.au

Abstract

The aim of the present article is to obtain a theoretical result essential for applications of combinatorial semigroups for the design of multiple classification systems in data mining. We consider a novel construction of multiple classification systems, or classifiers, combining several binary classifiers. The construction is based on combinatorial Rees matrix semigroups without any restrictions on the sandwich-matrix. Our main theorem gives a complete description of all optimal classifiers in this novel construction.

1   Introduction

The present article is devoted to a novel construction of multiple classification systems, or classifiers, combining several binary classifiers and based on combinatorial Rees matrix semigroups. Our main theorem completely describes all optimal classifiers in this novel construction without any restriction on the sandwich-matrices (see Theorem 1 in Section 2).

The design of multiple classifiers combining several binary classifiers is well known in the literature and has valuable applications in data mining (see, for example, [22, 24]). We refer to [18, 19] for more detailed motivation, and to the monographs [11, 21] for preliminaries and standard notation used in this research direction. Recall that a classifier is said to be *binary* if it divides all data into two classes. The set of all outputs of a binary classifier can be regarded as the field of order two, which will be denoted by F throughout. This means that the set of the outputs of m binary classifiers can be regarded as the vector space F^m.

In order to introduce convenient sets of generators for classification systems with outputs of initial classifiers represented as vectors in F^m, further we assume that T is a finite semigroup and the number m of the binary classifiers being combined does not exceed the number of nonzero elements of T. Further, zeros of semigroups are denoted by θ, and 0 stands for the zero of a ring. If T has no zero, then the semigroup ring $F[T]$ is isomorphic to the contracted semigroup ring $F_0[T^0]$, where $T^0 = T \cup \{\theta\}$. In this case we also write $F_0[T]$ for $F[T]$. These constructions are used, for example, in [2, 4, 5, 14]. We identify the standard vector space F^m with $F_0[T]$ by identifying $r = (r_1, \ldots, r_m) \in F^m$ with $\sum_{i=1}^m r_i s_i \in F_0[T]$. Thereby the set $F^m = F_0[T]$ has been endowed with multiplication.

For the design of efficient multiple classifiers it is essential to find sets with small numbers of generators and large minimum distance, see [18], [19] and [22], Section 7.5. The *weight* $\mathrm{wt}(r)$ of $r = \sum_{s \in T} r_s s \in F_0[T]$ is the number of nonzero coefficients r_s, for $s \in S$. The weight of a subset $C \subseteq F^m$ is the minimal weight of a nonzero element in C. The *minimum distance* of C is the

smallest weight of a nonzero difference $u - v$ for $u, v \in C$. If C is a linear space, then the minimum distance of C coincides with its weight. A *class set*, or *centroid set*, generated by the elements $g_1, \ldots, g_k \in F^m$ is the set

$$
\begin{aligned}
C(g_1, \ldots, g_k) &= F^m g_1 F^m + \cdots + F^m g_k F^m \\
&= \left\{ \sum_{j=1}^{m_1} \ell_{1,j} g_1 r_{1,j} + \cdots + \sum_{j=1}^{m_k} \ell_{k,j} g_k r_{k,j} \;\middle|\; \ell_{i,j}, r_{i,j}, \in F^m \cup F \right\}.
\end{aligned}
$$

2   Main Theorem

We use standard concepts and refer the readers to [9, 11, 12, 13, 14, 15, 17, 21, 22, 23] for background information and examples of recent results on related topics. For preliminaries on data mining, multiple classifiers and their class sets let us refer to the book [23] and papers [16, 18, 19, 20, 24].

Let G be a group, I and Λ nonempty sets, and let e be the identity of G. Denote by $G^1 = G \cup \{1\}$ and $G^0 = G \cup \{\theta\}$ the group G with identity 1 and, respectively, zero θ adjoined. Let $P = [p_{\lambda i}]$ be a $(\Lambda \times I)$-matrix with entries $p_{\lambda i} \in G^0$, for all $\lambda \in \Lambda$, $i \in I$. The *Rees matrix semigroup* $M(G; I, \Lambda; P)$ *over* G with *sandwich-matrix* P consists of all triples $(g; i, \lambda)$, where $i \in I$, $\lambda \in \Lambda$, $g \in G$, and multiplication is defined by the rule

$$
(g_1; i_1, \lambda_1)(g_2; i_2, \lambda_2) = (g_1 p_{\lambda_1 i_2} g_2; i_1, \lambda_2). \tag{1}
$$

The *Rees matrix semigroup* $M^0(G; I, \Lambda; P)$ *over* G *with zero* is the set consisting of θ and all triples $(g; i, \lambda)$, where $i \in I$, $\lambda \in \Lambda$, $g \in G^0$, where all triples $(\theta; i, \lambda)$ are identified with θ, and multiplication is defined by (1). Let us refer to [1, 3, 5, 6, 7, 8, 10, 19] for relevant recent results.

A semigroup T is said to be *combinatorial* if every subgroup of T is a singleton. Throughout we assume that $T = M^0(G; I, \Lambda; P)$ is a finite combinatorial Rees matrix semigroup over a group. We define subsets

$$
\begin{aligned}
L &= \{\lambda \in \Lambda \;:\; (\forall i \in I)\, p_{\lambda i} = \theta\}, \\
R &= \{i \in I \;:\; (\forall \lambda \in \Lambda)\, p_{\lambda i} = \theta\}.
\end{aligned}
$$

Denote by L_E a subset of $\Lambda \setminus L$ with the largest cardinality and such that, for each $i \in I$, the set $\{\lambda \in L_E \;:\; p_{\lambda i} \neq \theta\}$ has an even number of elements. Let R_E be a subset of $I \setminus R$ with the largest cardinality and such that, for each $\lambda \in \Lambda$, the set $\{i \in R_E \;:\; p_{\lambda i} \neq \theta\}$ has an even number of elements. Denote by D_z a subset of $T \cap G \times (I \setminus R) \times (\Lambda \setminus L)$ with the largest cardinality $|D_z|$ and such that, for all $j, k \in I$, $\mu, \nu \in \Lambda$, the sets

$$
\{i \in I \;:\; (e; i, \nu) \in D_z, p_{\mu i} \neq \theta\} \text{ and } \{\lambda \in \Lambda \;:\; (e; j, \lambda) \in D_z, p_{\lambda k} \neq \theta\}
$$

all have even numbers of elements. We define the following numbers

$$
\begin{aligned}
M_r &= |L| + |L_E|, \\
M_c &= |R| + |R_E|, \\
M_z &= |L| \cdot |R| + |L_E| \cdot |R| + |R_E| \cdot |L| + |D_z|.
\end{aligned}
$$

To introduce the generators of class sets in $F_0[T]$ we take any subset X in T and put

$$g_X = \begin{cases} \sum_{x \in X \setminus \{\theta\}} x & \text{if } X \nsubseteq \{\theta\}, \\ 0 & \text{otherwise.} \end{cases}$$

Further, we use the obvious fact that if two subsets X and Y of T are disjoint, then $g_{X \cup Y} = g_X + g_Y$. Let Q be a subset of T. For $j \in I$, $\mu \in \Lambda$, we use standard notation for the sets

$$Q_{*\mu} = Q \cap \{(e; i, \mu) : i \in I\},$$
$$Q_{j*} = Q \cap \{(e; j, \lambda) : \lambda \in \Lambda\},$$
$$Q_{j\mu} = Q \cap \{(e; j, \mu)\}.$$

For any $X \subseteq I$, $Y \subseteq \Lambda$, let

$$Q_{X*} = \cup_{i \in X} Q_{i*},$$
$$Q_{*Y} = \cup_{\lambda \in Y} Q_{*\lambda},$$
$$Q_{XY} = \cup_{i \in X, \lambda \in Y} Q_{i\lambda}.$$

For any $i \in I$, $\lambda \in \Lambda$, we put

$$Q^{(ei)} = \{\mu \in \Lambda : (e; i, \mu) \in Q\},$$
$$Q^{(e\lambda)} = \{j \in I : (e; j, \lambda) \in Q\}.$$

For any $\lambda \in \Lambda$, denote by $\mathcal{P}_c(\lambda, T)$ the set of all subsets X of $G \times (I \setminus R) \times \{\lambda\} \subseteq T$ such that $|X^{(e\lambda)}| = |R_E|$ and the set $\{i \in X^{(e\lambda)} : p_{\lambda i} \neq \theta\}$ has an even number of elements. Denote by $\mathcal{G}_c$ the set of all elements $g = g_{T_{R\lambda} \cup X} = g_{T_{R\lambda}} + g_X$, for all $\lambda \in \Lambda$ and all subsets $X \in \mathcal{P}_c(\lambda, T)$. If $|R_E| > 0$, then it is easily seen that the set $\mathcal{G}_c$ is nonempty and contains only nonzero elements.

For any $i \in I$, denote by $\mathcal{P}_r(i, T)$ the set of all subsets Y of $G \times \{i\} \times (\Lambda \setminus L) \subseteq T$ such that $Y^{(ei)} = |L_E|$ and the set $\{\lambda \in L_E : p_{\lambda i} \neq \theta\}$ has an even number of elements. Denote by $\mathcal{G}_r$ the set of all elements $g = g_{T_{iL} \cup Y} = g_{T_{iL}} + g_Y$ for all $i \in I$ and all subsets $Y \in \mathcal{P}_r(i, T)$. If $|L_E| > 0$, then it is easy to see that the set $\mathcal{G}_r$ is nonempty and contains only nonzero elements.

Let $\mathcal{P}_{cz}$ be the set of all subsets X of $G \times (I \setminus R) \times L \subseteq T$ such that, for all $\lambda \in L$, $\mu \in \Lambda$, the set $\{i \in X^{(e\lambda)} : p_{\mu i} \neq \theta\}$ contains an even number of elements and $|X^{(e\lambda)}| = |R_E|$. Denote by $\mathcal{P}_{rz}$ the set of all subsets Y of $G \times R \times (\Lambda \setminus L) \subseteq T$ such that, for all $i \in R$, $j \in I$, the set $\{\lambda \in Y^{(ei)} : p_{\lambda j} \neq \theta\}$ has an even number of elements and $|Y^{(ei)}| = |L_E|$. Denote by $\mathcal{P}_{ez}$ the set of all subsets V of $G \times (I \setminus R) \times (\Lambda \setminus L) \subseteq T$ such that, for all $j \in I$ and $\mu \in \Lambda$, the set $\{(e; i, \lambda) \in V : p_{\mu i}, p_{\lambda j} \neq \theta\}$ has an even number of elements $|V| = |D_z|$. Let $\mathcal{G}_z$ be the set of all elements $g_{T_{RL} \cup X \cup Y \cup V} = g_{T_{RL}} + g_X + g_Y + g_V$, for all subsets $X \in \mathcal{P}_{cz}$, $Y \in \mathcal{P}_{rz}$, $V \in \mathcal{P}_{ez}$. It is routine to verify that, if $M_z > 0$, then $\mathcal{G}_z$ is nonempty and contains only nonzero elements.

Theorem 1 *Let $T = M^0(G; I, \Lambda; P)$ be a finite combinatorial Rees matrix semigroup over a group with zero. Then the following conditions are satisfied:*

(i) $\mathrm{wt}(C(g)) = \mathrm{wt}(g) = M_c$, *for all $g \in \mathcal{G}_c$.*

(ii) $\mathrm{wt}(C(g)) = \mathrm{wt}(g) = M_r$, *for all $g \in \mathcal{G}_r$.*

(iii) $\mathrm{wt}(C(g)) = \mathrm{wt}(g) = M_z$, *for all* $g \in \mathcal{G}_z$.

(iv) *Let* $C = C(g_1, \ldots, g_k)$ *be a class set with the largest weight in* $F_0[T]$, *and let*

$$W = \max\{1, M_r, M_c, M_z\}.$$

Then the weight of C *is equal to* W. *Furthermore, if* $W > 1$, *then* C *contains an element of weight* W *belonging to the union of the sets* $\mathcal{G}_c$, $\mathcal{G}_r$ *and* $\mathcal{G}_z$.

The next corollary follows immediately.

Corollary 2 *Let* $T = M^0(G; I, \Lambda; P)$ *be a finite combinatorial Rees matrix semigroup over a group with zero. Then the largest number of errors of the binary classifiers, which can be corrected by a multiple classifier with class set of the form* $C(g_1, \ldots, g_k)$ *in* $F_0[T]$, *is equal to*

$$E = \max\left\{0, \lfloor (M_r - 1)/2 \rfloor, \lfloor (M_c - 1)/2 \rfloor, \lfloor (M_z - 1)/2 \rfloor\right\}.$$

3   Proofs

For any $r = \sum_{t \in T} r_t t \in F_0[T]$, $j \in I$ and $\mu \in \Lambda$, put $r_j = \sum_{\lambda \in \Lambda} r_{(e;j,\lambda)}$ and $r_\mu = \sum_{i \in I} r_{(e;i,\mu)} \in F_0[T]$. The *support* of r is the set

$$\mathrm{supp}(r) = \{t \in T \ : \ r_t \neq 0\}.$$

Let us define the sets

$$\begin{aligned}
\mathrm{Ann}_\ell(F_0[T]) &= \{x \in F_0[T] \ : \ xF_0[T] = 0\}, \\
\mathrm{Ann}_r(F_0[T]) &= \{x \in F_0[T] \ : \ F_0[T]x = 0\}, \\
\mathrm{Ann}(F_0[T]) &= \mathrm{Ann}_\ell(F_0[T]) \cap \mathrm{Ann}_r(F_0[T]).
\end{aligned}$$

Let Z_c be the set of all subsets X of T such that, for all $\lambda, \mu \in \Lambda$, the set $\{i \in X^{(e\lambda)} \ : \ p_{\mu i} \neq \theta\}$ has an even number of elements. Denote by Z_r the set of all subsets X of T such that, for all $i, j \in I$, the set $\{\lambda \in X^{(ei)} \ : \ p_{\lambda j} \neq \theta\}$ has an even number of elements.

Lemma 3 *Let* $T = M^0(G; I, \Lambda; P)$ *be a combinatorial Rees matrix semigroup over a group with zero. Then the following conditions hold:*

(i) $\mathrm{Ann}_\ell(F_0[T]) = \{g_X \ : \ X \in Z_r\}$,

(ii) $\mathrm{Ann}_r(F_0[T]) = \{g_X \ : \ X \in Z_c\}$,

(iii) $\mathrm{Ann}(F_0[T]) = \{g_X \ : \ X \in Z_c \cap Z_r\}$.

Proof. First, let us show that equality (i) holds. Consider any $r \in \mathrm{Ann}_\ell(F_0[T])$. Put $X = \mathrm{supp}(r)$. Since $F = GF(2)$, we see that $r = g_X$. Choose an arbitrary element $y = (e; j, \mu)$ in T. We will also write y for $1y \in F_0[T]$.

By the choice of r, we get $g_X \cdot y = 0$ and $X \subseteq I \times \Lambda$. Since $y \in T$, we see that $j \in I$ and $\mu \in \Lambda$. Pick any $i \in I$. It follows that $(ry)_i = r_i y = \sum_{\lambda \in X^{(ei)}} (e; i, \lambda) p_{\lambda j}(e; j, \mu)$. Since $(ry)_i = 0$ and F is a field of characteristic 2, we see that the number of elements $\lambda \in X^{(ei)}$, satisfying $p_{\lambda j} \neq \theta$, is even. Since i and j are arbitrary elements of I, this means that X belongs to Z_r. Therefore $\mathrm{Ann}_\ell(F_0[T]) \subseteq \{g_X \ : \ X \in Z_r\}$.

To prove the reversed inclusion, choose any X in Z_r and consider the element $r = g_X$. Take any $y \in F_0[T]$. We can write it down as $y = \sum_{j \in I, \mu \in \Lambda}(y_{j\mu}; j, \mu)$, where $y_{j\mu} \in \{\theta, e\}$. By the definition of Z_r, the number $n = |\{\lambda \in X^{(ei)} \ : \ p_{\lambda j} \neq \theta\}|$ is even, for each $i, j \in I$. Hence we get $r_i y_j = \sum_{\lambda \in X^{(ei)}} r_{(e;i,\lambda)} \sum_{\mu \in \Lambda}(y_{j\mu}; j, \mu) = \sum_{\mu \in \Lambda} \sum_{\lambda \in X^{(ei)}}(p_{\lambda j} y_{j\mu}; i, \mu) = \sum_{\mu \in \Lambda} n(y_{j\mu}; i, \mu) = 0$, because $n = 0$ in F. Since $y = \sum_{j \in I} y_j$, we get $ry = 0$. Hence $r \in \mathrm{Ann}_\ell(F_0[T])$. Therefore $\mathrm{Ann}_\ell(F_0[T]) \supseteq \{g_X \ : \ X \in Z_r\}$. Thus equality (i) holds true.

The proof of equality (ii) is dual and we omit it. Equality (iii) follows from (i) and (ii). This completes the proof. $\square$

Lemma 4 *Let $T = M^0(G; I, \Lambda; P)$ be a combinatorial Rees matrix semigroup over a group with zero. Then the following conditions hold:*

(i) $FT_{*L} \cdot F_0[T] = 0$,

(ii) $F_0[T] \cdot FT_{R*} = 0$,

(iii) $FT_{RL} \cdot F_0[T] = F_0[T] \cdot FT_{RL} = 0$,

(iv) $F_0[T] \cdot F\mathcal{G}_c = 0$,

(v) $F\mathcal{G}_r \cdot F_0[T] = 0$,

(vi) $F_0[T] \cdot F\mathcal{G}_z = F\mathcal{G}_z \cdot F_0[T] = 0$.

Proof. Equalities (i) and (ii) follow from the definitions of L and R. Equality (iii) follows from (i) and (ii). The definitions of R, $\mathcal{P}_c(\lambda, T)$, $\mathcal{G}_c$ and Z_c show that

$$\mathcal{G}_c \subseteq \{g_X \ : \ X \in Z_c\}.$$

By Lemma 3(ii), we get $\mathcal{G}_c \subseteq \mathrm{Ann}_r(F_0[T])$. Therefore condition (iv) follows.

The definitions of L, $\mathcal{P}_{ri}$, $\mathcal{G}_r$ and Z_r imply that $\mathcal{G}_r \subseteq \{g_X \ : \ X \in Z_r\}$. By Lemma 3(i), we get $\mathcal{G}_r \subseteq \mathrm{Ann}_\ell(F_0[T])$. Therefore condition (v) follows.

It remains to verify equality (vi). Fix any element x in $\mathcal{G}_z$. The definition of $\mathcal{G}_z$ shows that x can be recorded as $x = g_{T_{RL}} + g_X + g_Y + g_V$, for some $X \in \mathcal{P}_{cz}$, $Y \in \mathcal{P}_{rz}$ and $V \in \mathcal{P}_{ez}$. Choose any homogenous element $y = (e; k, \mu)$ in $F_0[T]$, where $k \in I$, $\mu \in \Lambda$.

Condition (iii) immediately yields $y g_{T_{RL}} = g_{T_{RL}} y = 0$. Since $X \in \mathcal{P}_{cz}$, we get $X \subseteq G \times (I \setminus R) \times L \subseteq T$. Therefore condition (i) implies that $g_X y = 0$. On the other hand, the definition of $\mathcal{P}_{cz}$ shows that the number $n = |\{i \in X^{(e\lambda)} \ : \ p_{\mu i} \neq \theta\}|$ is even, for all $\lambda \in L$. Therefore $(y g_X)_\lambda = y \cdot (g_X)_\lambda = \sum_{i \in X^{(e\lambda)}}(p_{\mu i}; k, \lambda) = n(e; k, \lambda) = 0$. Hence $y g_X = 0$. It follows that $g_X \in \mathrm{Ann}(F_0[T])$. Similarly, we get $g_Y \in \mathrm{Ann}(F_0[T])$.

Since $V \in \mathcal{P}_{ez}$, we see that, for each $\nu \in \Lambda$, the number $n = |\{(e; i, \nu) \in V : p_{\mu i} \neq \theta\}|$ is even. Since $y = (e; k, \mu)$, we get $(yg_V)_\nu = (e; k, \mu) \cdot (g_V)_\nu = (e; k, \mu) \sum_{(e;i,\nu) \in V}(e; i, \nu) = \sum_{(e;i,\nu) \in V}(p_{\mu i}; k, \nu) = n(e; k, \nu) = 0$. Since ν is an arbitrary element, it follows that $yg_V = 0$.

For each $j \in I$, denote by m the number of elements $(e; j, \lambda) \in V$ such that $p_{\lambda k} \neq \theta$. Since $V \in \mathcal{P}_{ez}$, it follows that m is an even number. Now, for any $j \in I$ and $y = (e; k, \mu)$, we obtain $(g_V y)_j = (g_V)_j \cdot (e; k, \mu) = \sum_{(e;j,\lambda) \in V}(e; j, \lambda) \cdot (e; k, \mu) = \sum_{(e;j,\lambda) \in V}(p_{\lambda k}; j, \mu) = m(e; j, \mu) = 0$. Since j is an arbitrary element, we get $g_V y = 0$. Thus $g_V \in \mathrm{Ann}\,(F_0[T])$.

Since $x = g_{T_{RL}} + g_X + g_Y + g_V$, we get $x \in \mathrm{Ann}\,(F_0[T])$. Therefore $\mathcal{G}_z \subseteq \mathrm{Ann}\,(F_0[T])$, and so condition (vi) follows. This completes the proof. $\square$

Lemma 5 *Let $T = M^0(G; I, \Lambda; P)$ be a finite combinatorial Rees matrix semigroup over a group with zero. Then the following equivalences hold:*

(i) $M_c = 0 \Longleftrightarrow \mathcal{G}_c = \emptyset$;

(ii) $M_r = 0 \Longleftrightarrow \mathcal{G}_r = \emptyset$;

(iii) $M_z = 0 \Longleftrightarrow \mathcal{G}_z = \emptyset$.

Proof. Choose any $\lambda \in \Lambda$. The definition of the set R_E implies that the set $G \times R_E \times \{\lambda\}$ belongs to $\mathcal{P}_c(\lambda, T)$. Since $g_{G \times R_E \times \{\lambda\}} = g_{T_{R_E \lambda}}$, we see that $g_{T_{R\lambda}} + g_{T_{R_E \lambda}} \in \mathcal{G}_c$. Hence the equivalence (i) follows.

As above, we can choose any i in I and derive that $g_{T_{iL}} + g_{T_{iL_E}} \in \mathcal{G}_r$. This means that equivalence (ii) follows.

Finally, since $T_{R_E L} \subseteq G \times (I \setminus R) \times L$, it is routine to verify that $T_{R_E L} \in \mathcal{P}_{cz}$. Besides, $T_{RL_E} \subseteq G \times R \times (\Lambda \setminus L)$ implies that $T_{RL_E} \in \mathcal{P}_{rz}$. We get $D_z \subseteq G \times (I \setminus R) \times (\Lambda \setminus L)$. It is straightforward to show that $D_z \in \mathcal{P}_{ez}$. Therefore equivalence (iii) follows. This completes the proof. $\square$

Proof of Theorem 1. First, let us prove condition (i). Take any element g in $\mathcal{G}_c$. There exist $\lambda \in \Lambda$ and $X \in \mathcal{P}_{c\lambda}$ such that $g = g_{T_{R\lambda}} + g_X$. Since $X \subseteq G \times (I \setminus R) \times \{\lambda\}$, we get $\mathrm{wt}(g) = |R| + |R_E| = M_c$. It remains to verify that $\mathrm{wt}(C(g)) = M_c$.

Choose a nonzero element y of minimal weight in $C(g)$ and write it down as $y = \sum_{j=1}^n a_j g b_j$, for some $a_j, b_j \in F \cup \{fs : f \in F, s \in T\}$. We may assume that all summands in this expression for y are nonzero. Lemma 4(iv) tells us that $a_j g = 0$ for each $a_j \notin F$. Therefore $a_j = 1$ for all $j = 1, \ldots, n$. Hence

$$y = \sum_{j=1}^n g b_j. \tag{2}$$

If $b_j = 1$, then $gb_j = g$ and $|\mathrm{supp}(gb_j)| = |\mathrm{supp}(g)| = M_c$. Further, consider the case where $0 \neq b_j \notin F$.

Then $b_j = f_j s_j$, for some $f_j \in F, s_j \in T$. There exist $\ell_j \in I, \mu_j \in \Lambda$ such that $s_j = (e; \ell_j, \mu_j)$. Therefore if $p_{\lambda \ell_j} = \theta$, then (1) and $g = g_{T_{R\lambda}} + g_X$ yield us that $gb_j = \theta$ and $\mathrm{supp}(gb_j) = \emptyset$. On

the other hand, if $p_{\lambda \ell_j} \neq \theta$, then $p_{\lambda \ell_j} = e$, and so $g = g_{T_{R\lambda}} + g_X$ implies that

$$\operatorname{supp}(gb_j) = \{(e; i, \mu_j) \; : \; i \in R\} \cup \{(e; i, \mu_j) \; : \; i \in X^{(e\lambda)}\}. \tag{3}$$

Hence $|\operatorname{supp}(gb_j)| = |\{(e; i, \mu_j) : i \in R\}| + |\{(e; i, \mu_j) : i \in X^{(e\lambda)}\}| = |\operatorname{supp}(g)| = M_c$. In addition, notice that if $\mu_{j_1} = \mu_{j_2}$ for $1 \leq j_1, j_2 \leq n$, then the union of two sets occurring in the right hand side of equality (3) for $j = j_1$ coincides with the union corresponding to $j = j_2$. Therefore it follows from (2) and (3) that the weight of y is divisible by M_c. The minimality of the weight of y implies that $\operatorname{wt}(y) \leq \operatorname{wt}(g) = M_c$, and so $\operatorname{wt}(y) = M_c$. Since $\operatorname{wt}(C(g)) = \operatorname{wt}(y)$, we get $\operatorname{wt}(C(g)) = M_c$, as required.

Second, we notice that the proof of condition (ii) is dual to that of (i), and so we omit it.

Third, let us prove condition (iii). Let g be an arbitrary element in $\mathcal{G}_z$. By the definition of $\mathcal{G}_z$ there exist subsets $X \in \mathcal{P}_{cz}$, $Y \in \mathcal{P}_{rz}$, $V \in \mathcal{P}_{ez}$ as required for $\mathcal{G}_z$. In view of the definitions of the sets $\mathcal{P}_{cz}$, $\mathcal{P}_{rz}$, $\mathcal{P}_{ez}$ of subsets of T, it is routine to verify that $\operatorname{wt}(g) = M_z$. On the other hand, it follows from Lemma 4(vi) that g generates a one dimensional cluster set $C(g) = F(g)$. Therefore $\operatorname{wt}(C(g))$ coincides with $\operatorname{wt}(g) = M_z$, as required.

Fourth, let us prove condition (iv). Choose a cluster set $C = C(g_1, \ldots, g_k)$ with the largest possible weight in $F_0[T]$. If $\operatorname{wt}(C) = 1$, then it follows from conditions (i), (ii) and (iii) and the maximality of $\operatorname{wt}(C)$ that M_c, M_r, $M_z \leq 1$. Hence $W = 1$, and so $\operatorname{wt}(C) = W$. Further, we consider the case where $\operatorname{wt}(C) > 1$.

Denote by $H_{\min}(C) = \{x \in C \; : \; \operatorname{wt}(x) = \operatorname{wt}(C)\}$ the set of all nonzero elements of minimal weight in C. For $\lambda \in \Lambda$, we define $H_c(\lambda) = \left\{ \sum_{j=1}^n (e; i_j, \lambda) \middle| \; n \geq 1, \; i_j \in I \right\}$ and

$$H_c = \cup_{\lambda \in \Lambda} H_c(\lambda).$$

Finally, for $i \in I$, let us also set $H_r(i) = \left\{ \sum_{j=1}^n (e; i, \lambda_j) \middle| \; n \geq 1, \; \lambda_j \in \Lambda \right\}$ and

$$H_r = \cup_{j \in I} H_r(j).$$

Since $\operatorname{wt}(C) > 1$, we get $C \cap (T \setminus \{\theta\}) = \emptyset$. Consider several possible cases.

Case 1: $H_{\min}(C) \cap H_r \neq \emptyset$. Choose y in $H_{\min}(C) \cap H_r$. Then there exists $i \in I$ such that $y \in H_r(i)$. Put

$$L(y) = \{\lambda \in L \; : \; \operatorname{supp}(y) \cap T_{*\lambda} \neq \emptyset\},$$
$$\Lambda(y) = \{\lambda \in \Lambda \setminus L \; : \; \operatorname{supp}(y) \cap T_{*\lambda} \neq \emptyset\}.$$

Suppose that, for some $s \in T$, the product ys is nonzero. Then $\operatorname{wt}(ys) = \operatorname{wt}(y)$. Since $ys \in C$, we get $ys \in H_{\min(C)}$. By (1), $H_r s \subseteq T \setminus \{\theta\}$. Hence $ys \in T \setminus \{\theta\}$. This contradicts $H_{\min}(C) \cap (T \setminus \{\theta\}) = \emptyset$ and shows that $yT = 0$.

For each $j \in I$, it follows from $yT = 0$ and the definition of L that the number of elements $\lambda \in \Lambda(y)$, satisfying $p_{\lambda j} \neq \theta$, is even. Hence $|\Lambda(y)| \leq |L_E|$. Obviously, $|L(y)| \leq |L|$, too. Therefore $\operatorname{wt}(y) = |L(y)| + |\Lambda(y)| \leq |L| + |L_E| = M_r$. Hence we get $\operatorname{wt}(C) = \operatorname{wt}(y) \leq M_r$.

By the choice of C, condition (ii) implies that $\operatorname{wt}(C) \geq M_r$, and so $\operatorname{wt}(y) = M_r$. It follows that $L(y) = L$ and $|\Lambda(y)| = |L_E|$. The definition of $\mathcal{P}_r$ and $\mathcal{G}_r$ show that $y \in \mathcal{G}_r$. Thus, in this case C contains an element of $\mathcal{G}_r$.

Case 2: $H_{\min}(C) \cap H_c \neq \emptyset$. This case is dual to Case 1, and we omit the proof.

Case 3: $H_{\min}(C) \cap H_r = H_{\min}(C) \cap H_c = \emptyset$. Choose any element y in $H_{\min}(C)$.

For any $s \in T$, it is clear that $sy, ys \in H_{\min}(C) \cup \{0\}$. Besides, it is easily seen that if $sy \neq 0$, then $sy \in H_r$. Similarly, if $ys \neq 0$, then $ys \in H_c$. Hence the hypothesis of Case 3 implies that $sy = ys = 0$. Therefore $Ty = yT = 0$, and so $y \in \mathrm{Ann}\,(F_0[T])$.

Let us introduce the following subsets X, Y, V, Z of $\mathrm{supp}(y)$:

$$
\begin{aligned}
X &= \mathrm{supp}(y) \cap (G \times (I \setminus R) \times L), \\
Y &= \mathrm{supp}(y) \cap (G \times R \times (\Lambda \setminus L)), \\
V &= \mathrm{supp}(y) \cap (G \times (I \setminus R) \times (\Lambda \setminus L)), \\
Z &= \mathrm{supp}(y) \cap T_{RL}.
\end{aligned}
$$

We are going to show that $X \in \mathcal{P}_{cz}$, $Y \in \mathcal{P}_{rz}$, $V \in \mathcal{P}_{ez}$ and $Z = T_{RL}$.

The definitions of the sets R, L show that $g_Z \in \mathrm{Ann}\,(F_0[T])$. Therefore it follows from $y \in \mathrm{Ann}\,(F_0[T])$ that $g_X, g_Y \in \mathrm{Ann}\,(F_0[T])$. Hence $g_V \in \mathrm{Ann}\,(F_0[T])$, too.

Lemma 3 implies that $X \in Z_c$, because $g_X \in \mathrm{Ann}\,(F_0[T])$. Hence it follows from the definition of Z_c that, for each $\lambda \in L$, $\mu \in \Lambda$, the number $|\{i \in X^{(e\lambda)} : p_{\mu i} \neq \theta\}|$ is even. Notice that this is the same condition, which occurs for X in the definition of the set $\mathcal{P}_{cz}$. Besides, Lemma 3 yields us that $Y \in Z_r$. It follows that, for each $i \in R$, $j \in I$, the number $|\{\lambda \in Y^{(ei)} : p_{\lambda j} \neq \theta\}|$ is even. Notice that this is the same condition which occurs for Y in the definition of the set $\mathcal{P}_{rz}$.

Lemma 3 also shows that $V \in Z_c \cap Z_r$. Hence it follows that, for all $j, k \in I$, $\mu, \nu \in \Lambda$, the numbers $|\{(e; i, \nu) \in V : p_{\mu i} \neq \theta\}|$ and $|\{(e; j, \lambda) \in V : p_{\lambda k} \neq \theta\}$ are even. These are the same conditions which occur for V in the definition of the set $\mathcal{P}_{ez}$.

Pick any $\lambda \in L$. The set $X^{(e\lambda)}$ is a subset of $I \setminus R$. As we have established, the number of $i \in X^{(e\lambda)}$, such that $p_{\mu i} \neq \theta$, is even, for each $\mu \in \Lambda$. This means that the set $X^{(e\lambda)}$ satisfies the property which occurs in the definition of the set R_E. The maximality of $|R_E|$ shows that $|X^{(e\lambda)}| \leq |R_E|$. Therefore

$$|X| \leq |R_E| \cdot |L|. \tag{4}$$

For each $i \in R$, a dual argument demonstrates that $|Y^{(ei)}| \leq |L_E|$ and

$$|Y^{(ei)}| \leq |L_E| \cdot |R|. \tag{5}$$

We have already established that, for each $j, k \in I$, $\mu, \nu \in \Lambda$, the number of $(e; i, \nu) \in V$, satisfying $p_{\mu i} \neq \theta$, is even and the number of $(e; j, \lambda) \in V$, satisfying $p_{\lambda k} \neq \theta$, is even. These are precisely the numbers which occur in the definition of the set D_z. It follows from the maximality of $|D_z|$ that

$$|V| \leq |D_z|. \tag{6}$$

Since $Z \subseteq T_{RL}$, the following inequality is obvious:

$$|Z| \leq |R| \times |L|. \tag{7}$$

Now we see that the maximality of the weight of C, condition (iii) and inequalities (4), (5), (6) and (7) imply that all of these inequalities turn into exact equalities: $|X^{(e\lambda)}| = |R_E|$, $|Y^{(ei)}| = |L_E|$,

$|V| = |D_z|$ and $Z = R \times L$. It follows that $X \in \mathcal{P}_{cz}$, $Y \in \mathcal{P}_{rz}$ and $V \in \mathcal{P}_{ez}$. Therefore $y = g_Z + g_X + g_Y + g_V \in \mathcal{G}_z$ in this case.

Thus we see that in all cases C always contains an element of minimal weight, which belongs to the union $\mathcal{G}_c$, $\mathcal{G}_r$ and $\mathcal{G}_z$, as claimed.

If we combine the last conclusion with Lemma 5 and conditions (i), (ii) and (iii), then it follows that $\text{wt}(C) \leq W$. On the other hand, the maximality of $\text{wt}(C)$ and conditions (i), (ii) and (iii) also yield us that $\text{wt}(C) \geq W$. Therefore $\text{wt}(C) = W$, which completes the proof. $\square$

4   Acknowledgements

References

[1] Auinger, K., Hall, T.E., Reilly, N.R. and S. Zhang, *Congruences on the lattice of pseudovarieties of finite semigroups*, Internat. J. Algebra Comput. **7**(1997)(4), 433–455.

[2] Cazaran, J., Kelarev, A.V., Quinn, S.J. and D. Vertigan, *An algorithm for computing the minimum distances of extensions of BCH codes embedded in semigroup rings*, Semigroup Forum **73**(2006), 317–329.

[3] Davey, B.A., Jackson, M., Maróti, M. and R.N. McKenzie, *Principal and syntactic congruences in congruence-distributive and congruence-permutable varieties*, J. Aust. Math. Soc. **85**(2008)(1), 59–74.

[4] Easdown, D. and W.D. Munn, *Trace functions on inverse semigroup algebras*, Bull. Austral. Math. Soc. **52**(1995)(3), 359–372.

[5] Hall, T.E., *The radical of the algebra of any finite semigroup over any field*, J. Austral. Math. Soc. Ser. A **11**(1970), 350–352.

[6] Hall, T.E., *Biprefix codes, inverse semigroups and syntactic monoids of injective automata*, Theoretical Computer Science **32**(1984)(1-2), 201–213.

[7] Hall, T.E., *Amalgamation for inverse and generalized inverse semigroups*, Trans. Amer. Math. Soc. **310**(1988), 313–323.

[8] Hall, T.E. and T. Imaoka, *Representations and amalgamation of generalized inverse *-semigroups*, Semigroup Forum **58**(1999), 126–141.

[9] Howie, J.M., "Fundamentals of Semigroup Theory", Clarendon Press, Oxford, 1995.

[10] Jackson, M. and M. Volkov, *Undecidable problems for completely 0-simple semigroups*, J. Pure Applied Algebra **213**(10)(2009), 1961-1978.

[11] Kelarev, A.V., "Ring Constructions and Applications", World Scientific, River Edge, 2002.

[12] Kelarev, A.V., "Graph Algebras and Automata", Marcel Dekker, New York, 2003.

[13] Kelarev, A.V., Göbel, R., Rangaswamy, K.M., Schultz, P. and C. Vinsonhaler, "Abelian Groups Rings and Modules", Contemporary Mathematics **273**, American Mathematical Society, New York, 2001.

[14] Kelarev, A.V. and D.S. Passman, *A description of incidence rings of group automata*, Contemporary Mathematics **456**(2008), 27–33.

[15] Kelarev, A.V. and C.E. Praeger, *On transitive Cayley graphs of groups and semigroups*, European J. Combinatorics **24**(2003)(1), 59–72.

[16] Kelarev, A., Ryan, J. and J. Yearwood, *Cayley graphs as classifiers for data mining: The influence of asymmetries*, Discrete Mathematics **309**(2009)(17), 5360–5369.

[17] Kelarev, A.V. and O.V. Sokratova, *On congruences of automata defined by directed graphs*, Theoretical Computer Science **301**(2003), 31–43.

[18] Kelarev, A.V., Yearwood, J.L. and M.A. Mammadov, *A formula for multiple classifiers in data mining based on Brandt semigroups*, Semigroup Forum **78**(2009)(2), 293–309.

[19] Kelarev, A.V., Watters, P. and J.L. Yearwood, *Rees matrix construction for clustering of data*, J. Aust. Math. Soc. **87**(2009), 377–393.

[20] Kelarev, A.V., Yearwood, J.L. and P.W. Vamplew, *A polynomial ring construction for classification of data*, Bull. Aust. Math. Soc. **79**(2009), 213–225.

[21] Okniński, J., "Semigroup Algebras", Marcel Dekker, New York, 1991.

[22] Witten, I.H. and E. Frank, "Data Mining: Practical Machine Learning Tools and Techniques", Elsevier/Morgan Kaufman, Amsterdam, 2005.

[23] Yearwood, J.L. and M.A. Mammadov, "Classification Technologies: Optimization Approaches to Short Text Categorization", Idea Group Inc., 2010.

[24] Yearwood, J., Webb, D., Vamplew, P., Ma, L., Ofoghi, B. and A. Kelarev, *Applying clustering and ensemble clustering approaches to phishing profiling*, Eighth Australasian Data Mining Conference, AusDM 2009, (Melbourne, Australia, December 2009), Conferences in Research and Practice in Information Technology **101**(2009), 25–34.