Bull. Aust. Math. Soc. **79** (2009), 213–225
doi:10.1017/S0004972708001111

A POLYNOMIAL RING CONSTRUCTION FOR THE CLASSIFICATION OF DATA

A. V. KELAREV ✉, J. L. YEARWOOD and P. W. VAMPLEW

(Received 11 February 2008)

Abstract

Drensky and Lakatos (Lecture Notes in Computer Science, 357 (Springer, Berlin, 1989), pp. 181–188) have established a convenient property of certain ideals in polynomial quotient rings, which can now be used to determine error-correcting capabilities of combined multiple classifiers following a standard approach explained in the well-known monograph by Witten and Frank (*Data Mining: Practical Machine Learning Tools and Techniques* (Elsevier, Amsterdam, 2005)). We strengthen and generalise the result of Drensky and Lakatos by demonstrating that the corresponding nice property remains valid in a much larger variety of constructions and applies to more general types of ideals. Examples show that our theorems do not extend to larger classes of ring constructions and cannot be simplified or generalised.

2000 *Mathematics subject classification*: primary 16S36; secondary 20M25, 94B60.
Keywords and phrases: ring constructions, group rings, classification.

1. Introduction

The classification of data plays one of the central roles in data mining and other applications of mathematical methods, see, for example, [4, 16, 20, 23, 24]. A well-known method of designing efficient multiple classifiers consists in representing them as several individual classifiers combined into one scheme. This method is very effective, and it is often advisable to apply it even in situations where it is possible to build multiple classifiers analysing the data directly, see Witten and Frank [22, Section 7.5]. The main advantage of this method is in the ability of multiple classifiers to correct the errors of individual classifiers. This is why the problem of determining the error-correcting capabilities of the combined multiple classifiers is crucial.

Polynomial quotient rings can be used to introduce additional structure to the class sets of multiple classifiers and generate these sets with small numbers of generators.

The first author was supported by Discovery grant DP0449469 from the Australian Research Council. The second author was supported by a Queen Elizabeth II Fellowship and Discovery grant DP0211866 from the Australian Research Council. The third author was supported by two research grants of the University of Ballarat.

Drensky and Lakatos [7] have established a convenient property of some ideals in polynomial quotient rings, which can now be used to determine error-correcting capabilities of combined multiple classifiers (see Proposition 3.4 in Section 3).

We strengthen and generalise the result of Drensky and Lakatos [7] by demonstrating that the corresponding nice property remains valid in a much larger variety of constructions and applies to more general types of ideals (see Theorems 4.1, 4.2 and 4.4 in Section 4). Examples show that our theorems do not extend to larger classes of ring constructions and cannot be simplified or generalised.

2. Motivation

We are using standard terminology and refer the reader to, for example, [8, 9, 14] for preliminaries on ring constructions and to [20, 22, 24] for background information on classification methods.

Consider the problem of combining several classifiers into a larger multiple classifier. Let p be a prime number, q a power of p, and let $\mathbb{F} = GF(q)$ be the finite field of order q. Suppose that there are $N > 1$ classifiers to be combined and that these classifiers divide their input data into classes by producing outputs $o_1, \ldots, o_N \in \mathbb{F}$ for each input element. Then the sequence $(o_1, \ldots, o_N) \in \mathbb{F}^N$ is called a *class vector* of the combined multiple classifier, and the set of all class vectors is called the *class set*.

Let V be an N-dimensional linear space over $\mathbb{F}$ with a basis $B = \{b_1, \ldots, b_N\}$. The number of nonzero coordinates of v with respect to the basis B is denoted by $\mathrm{wt}(v) = \mathrm{wt}_B(v)$ and is called the *weight*, or *Hamming weight*, of v in the basis B. If it is clear from the context which basis B is being considered, then the weight of v in the basis B is called the *weight* of v.

The *weight* of a class set C is the minimum weight of a nonzero element in C. The *minimum distance* of a class set C is the minimum weight among all weights of nonzero differences between pairs of elements in C. If C is a linear space, then its minimum distance is equal to its weight. For any real number x, denote by $\lfloor x \rfloor$ the *integral part* of x, or the *floor* of x, that is, the largest integer which does not exceed x. It is well known and easy to verify that the number of errors of binary classifiers that the multiple classifier can correct is equal to $\lfloor (d-1)/2 \rfloor$, where d is the minimum distance of the class set of the classifier.

Instead of storing the whole large class set C in computer memory, it is convenient to be able to generate C with one or more generators. To this end we are going to take a polynomial ring and use it to introduce additional structure on the class set of a multiple classifier. This will allow us to multiply the generators with arbitrary elements of $\mathbb{F}^N$ and to take their sums. The structure will enable us to find small generating sets for the classifier and to determine error-correcting capabilities of the whole class set by looking only at its generators.

https://doi.org/10.1017/S0004972708001111 Published online by Cambridge University Press

3. Polynomial generators for class sets

Let $\mathbb{N}_0$ be the set of nonnegative integers, $X = \{x_1, \ldots, x_m\}$ a set of commuting variables, and let $\mathbb{F}[X]$ stand for the ring of polynomials in X over $\mathbb{F}$. Denote by

$$M_X = \{x_1^{a_1} \cdots x_m^{a_m} \mid a_1, \ldots, a_m \in \mathbb{N}_0\}$$

the free commutative monoid generated by X. Choose any subset P of the set $M_X^2 = M_X \times M_X$ of all pairs (u, v), where $u, v \in M_X$. A *binomial ideal* of $\mathbb{F}[X]$ is the ideal

$$I_P = (u - v \mid \text{for all } (u, v) \in P) \tag{3.1}$$

generated by all binomials $u - v$, for all $(u, v) \in P$. The binomials $u - v$, for $u, v \in M_X$, are also sometimes called *pure difference binomials*. We are going to use the polynomial quotient ring $\mathbb{F}[X]/I_P$ to generate multiple classifiers.

To this end let us now review an alternative representation for the quotient ring $\mathbb{F}[X]/I_P$. Let M be a monoid. The *monoid algebra* $\mathbb{F}[M]$ is the $\mathbb{F}$-algebra spanned by the elements of M with multiplication defined by the distributive law and the multiplication of M. Note that the polynomial ring $\mathbb{F}[X]$ coincides with the monoid algebra $\mathbb{F}[M_X]$. If M is a group, then $\mathbb{F}[M]$ is called a *group algebra*. These constructions were considered, for example, in [2, 3, 5, 10–12, 17].

Denote by ϱ_P the congruence generated in M_X by all pairs (u, v), for all $(u, v) \in P$. Then the quotient ring $\mathbb{F}[X]/I_P$ is isomorphic to the monoid algebra $\mathbb{F}[M_X/\varrho_P]$, see [14]. We identify the polynomial quotient ring $\mathbb{F}[X]/I_P$ and the monoid algebra $\mathbb{F}[M_X/\varrho_P]$ so that

$$\mathbb{F}[X]/I_P = \mathbb{F}[M_X/\varrho_P]. \tag{3.2}$$

Hence, the dimension of the quotient ring $\mathbb{F}[X]/I_P$ over $\mathbb{F}$ is equal to the cardinality $|M|$ of the quotient monoid $M = M_X/\varrho_P$.

Further, we assume that the dimension of the quotient ring is equal to the number of classifiers being combined. This means that $|M| = N$ and so $M = \{m_1, \ldots, m_N\}$. In order to use two operations of the quotient ring $\mathbb{F}[X]/I_P = \mathbb{F}[M]$ for $\mathbb{F}^N$, let us identify $\mathbb{F}[M]$ with $\mathbb{F}^N$ by identifying every element

$$r = \sum_{i=1}^{N} r_i m_i \in \mathbb{F}[M] \tag{3.3}$$

with the sequence

$$(r_1, \ldots, r_N) \in \mathbb{F}^N. \tag{3.4}$$

This makes sense since the standard addition of vectors is defined on the linear space $\mathbb{F}^N$ componentwise, and so it coincides with the definition of addition in the monoid algebra $\mathbb{F}[M]$.

We can now use two operations to generate classifiers. An element $r \in \mathbb{F}^N$ is said to be *generated* by the elements $g_1, \ldots, g_k \in \mathbb{F}^N$ if it belongs to the ideal generated by these elements, that is, if it is equal to a sum of multiples of these generators.

https://doi.org/10.1017/S0004972708001111 Published online by Cambridge University Press

Accordingly, the whole class set C of a multiple classifier is said to be *generated* by the elements $g_1, \ldots, g_k$ in $\mathbb{F}^N$ if C coincides with the ideal generated by these elements, that is, if C is equal to the set of all sums of multiples of these generators,

$$\begin{aligned} C &= C(g_1, \ldots, g_k) \\ &= \left\{ \sum_{i=1}^{m_1} r_{1i} g_1 + \cdots + \sum_{i=1}^{m_k} r_{ki} g_k \,\middle|\, \text{where } r_{ji} \in \mathbb{F}^N \right\}. \end{aligned} \tag{3.5}$$

In this case the notation $C = C(g_1, \ldots, g_k)$ is used when it is necessary to indicate the generators explicitly.

Fix a set $P \subseteq M_X^2$ and consider the ring $F[X]/I_P$. Let $D \subseteq \mathbb{N}_0^m$. Denote by U_D the set of polynomials

$$u_d = \prod_{i=1}^{m} (x_i^2 - x_i)^{d_i} \in \mathbb{F}[X], \tag{3.6}$$

for all $d = (d_1, \ldots, d_m) \in D$. As is customary, it is assumed that all zero powers $(x_i^2 - x_i)^0$ are equal to the identity element 1 of $\mathbb{F}[X]/I_P$. Let I_D be the ideal generated in the polynomial quotient ring $\mathbb{F}[X]/I_P$ by the set U_D.

DEFINITION 3.1. A class set $C \subseteq \mathbb{F}^N$ will be called a *Drensky class set* if $C = I_D$ for some $D \subseteq \mathbb{N}_0^m$.

Drensky class sets were considered in [7] using a different terminology.

DEFINITION 3.2. A set $U \subseteq \mathbb{F}^N$ will be called a *visible generating set*, or a *visible set of generators*, if the weight of the class set $C = C(U)$ generated by U in $\mathbb{F}^N$ is equal to the minimum of the weights of the generators $u \in U$.

This concept is analogous to the notion of a visible basis introduced in [21], see also [6].

DEFINITION 3.3. We say that a Drensky class set I_D with $D \subseteq \mathbb{N}_0^m$ is *visible* if its standard generating set U_D is visible.

A convenient method for finding the minimal distances of some Drensky class sets has been obtained in [7, Proposition 1.2]. An *elementary abelian p-group* is a group isomorphic to a direct product $\mathbb{Z}_p^k$, where $\mathbb{Z}_p$ stands for the cyclic group of order p and k is a nonnegative integer.

PROPOSITION 3.4 (Drensky and Lakatos [7]). *Let $\mathbb{F}$ be a finite field, and let P be a subset of M_X^2 such that M_X/ϱ_P is an elementary abelian p-group. Then every Drensky class set I_D in the polynomial quotient ring $\mathbb{F}[X]/I_P$ is visible.*

4. Main results

Our new theorems generalise Proposition 3.4 and give us an efficient method for finding the error-correcting capabilities of the corresponding multiple classifiers.

https://doi.org/10.1017/S0004972708001111 Published online by Cambridge University Press

Indeed, when a class set has a visible generating set, then it is very easy to determine its weight and the number of errors it can correct. Recall that a commutative semigroup is called a *semilattice* if it entirely consists of idempotents (see [6] for a recent related result).

THEOREM 4.1. *Let $\mathbb{F}$ be a finite field, and let P be a subset of M_X^2 such that M_X/ϱ_P is a subsemigroup of a direct product of a semilattice, an elementary abelian 2-group and an elementary abelian p-group. Then every Drensky class set I_D in the polynomial quotient ring $\mathbb{F}[X]/I_P$ is visible.*

In the case of $\mathrm{char}(\mathbb{F}) = 2$, it turns out to be possible to prove even more.

THEOREM 4.2. *Let $\mathbb{F}$ be a finite field with $\mathrm{char}(\mathbb{F}) = 2$, and let P be a subset of M_X^2 containing all pairs (x^3, x), for all $x \in X$. Then every Drensky class set I_D in the polynomial quotient ring $\mathbb{F}[X]/I_P$ is visible.*

In addition, we show that more general types of generating sets are also visible.

DEFINITION 4.3. A class set C in $R = \mathbb{F}[X]/I_P$ will be called a *binomial class set* if $C = C(b_1, \ldots, b_k)$, where

$$b_i = \prod_{j=1}^{k_i} (w_{i,j}^2 - w_{i,j})^{d_{i,j}}, \tag{4.1}$$

for some $w_{i,j} \in M_X$, $d_{i,j} \in \mathbb{N}_0$.

Obviously, every Drensky class set is a binomial class set.

THEOREM 4.4. *Let P be a subset of M_X^2 containing all pairs (x^{p+1}, x), for all $x \in X$. Then the following conditions are equivalent.*

(i) *Every Drensky class set in $R = \mathbb{F}[X]/I_P$ is visible.*
(ii) *Every binomial class set in R is visible.*

REMARK 4.5. Theorem 4.1 generalises Proposition 3.4 in any characteristic. In the case of characteristic $\mathrm{char}(\mathbb{F}) = 2$, Theorem 4.2 generalises Proposition 3.4 and Theorem 4.1, see the beginning of proof of Theorem 4.1 in Section 6.

Examples given in Section 5 show that our theorems cannot be simplified or generalised. In particular, in the case of $\mathrm{char}(\mathbb{F}) > 2$, it is impossible to generalise Theorem 4.1 to an analogous version of Theorem 4.2.

5. Examples

Our first example demonstrates that Theorem 4.1 and Proposition 3.4 cannot be extended to groups which are not p-groups.

EXAMPLE 5.1. Let p' be a prime such that $2 < p' \neq p$, and let $X = \{x\}$, $P = \{(x, x^{p'+1})\}$. Then $\mathbb{F}[X]/I_P$ is isomorphic to the group algebra $\mathbb{F}[G]$ of the cyclic group $G = \mathbb{Z}_{p'}$ of order p'. Consider the ideal J generated in $F[G]$ by $g = (1-x)^{p'-1}$. Since $p' \neq p$, it follows that $\mathrm{wt}(g) = p'$ in $\mathbb{F}[G]$. However, every commutative $\mathbb{F}$-algebra satisfies the identity $(y+z)^p = y^p + z^p$ for all y, z. Choose k such that $p^k > p' - 1$. Then we obtain $(1-x)^{p^k} = 1 - x^{p^k}$, and so $\mathrm{wt}((1-x)^{p^k})^I = 2$ in $\mathbb{F}[G]$. Therefore, $\mathrm{wt}(J) = 2 < \mathrm{wt}(g)$, because $(1-x)^{pp'} \in J$.

The next two examples show that Theorem 4.1 cannot be generalised to p-groups which are not elementary abelian groups.

EXAMPLE 5.2. Let $p = 2$, $\mathbb{F} = \mathbb{F}_2 = GF(2)$, $X = \{x\}$, and let $P = \{(x, x^{3p+1})\}$. Then $\mathbb{F}_2[X]/I_P$ is isomorphic to the group algebra $\mathbb{F}_2[G]$ of the cyclic group $G = \mathbb{Z}_{p^3}$ of order p^3. Consider the ideal J generated in $\mathbb{F}_2[G]$ by $g = (1-x)^3$. Clearly, $g = 1 + x + x^2 + x^3 \in \mathbb{F}_2[G]$ and so $\mathrm{wt}(g) = 4$. However, we see that $\mathrm{wt}((1-x)g) = 2$. Therefore, $\mathrm{wt}(J) = 2 < \mathrm{wt}(g)$. This demonstrates that, in the case of $p = 2$, Theorem 4.1 does not generalise to p-groups which are not elementary abelian.

EXAMPLE 5.3. Let $p > 2$, $X = \{x\}$, and let $P = \{(x, x^{2p+1})\}$. Now $\mathbb{F}[X]/I_P$ is isomorphic to the group algebra $\mathbb{F}[G]$ of the cyclic group $G = \mathbb{Z}_{p^2}$ of order p^2. This time we look at the ideal J generated in $\mathbb{F}[G]$ by $g = (1-x)^2 = 1 - 2x + x^2 \in \mathbb{F}[G]$. Here $\mathrm{wt}(g) = 3$, because $p \neq 2$. However, $(1-x)^p = 1 - x^p$ in $\mathbb{F}[G]$. Therefore, $\mathrm{wt}((1-x)^p) = 2$ and $\mathrm{wt}(J) = 2 < \mathrm{wt}(g)$. Thus, in the case where $p > 2$, Theorem 4.1 cannot be generalised to p-groups which are not elementary abelian.

Our next example shows that Theorem 4.1 cannot be generalised to monoids which are unions of p-groups but are not contained in a direct product of an elementary abelian 2-group, an elementary abelian p-group and a semilattice.

EXAMPLE 5.4. Let $X = \{x_1, x_2, x_3\}$, and let

$$P = \{(x_i, x_i^{p+1}) \mid i = 1, 2, 3\} \cup \{(x_2^p, x_3^p), (x_1^p x_2, x_1)\}, (x_1^p x_3, x_1)\}.$$

Then M_P is a union of two elementary abelian groups: $\langle x_2, x_3\rangle \cong \mathbb{Z}_p \times \mathbb{Z}_p$ and $\langle x_1\rangle \cong \mathbb{Z}_p$. Consider the ideal J generated in $R = \mathbb{F}[X]/I_P$ by $g = (1-x_2)(1-x_3)$. We have $\mathrm{wt}(g) = 4$. However, $x_1 g = 1 + 2x_1 + x_1^2$ in R. If $p = 2$, then $\mathrm{wt}(x_1 g) = 2$. On the other hand, if $p > 2$, then $\mathrm{wt}(x_1 g) = 3$. In any case we obtain $\mathrm{wt}(J) < \mathrm{wt}(g)$.

The following example demonstrates that Theorem 4.1 does not generalise to monoids which are not unions of groups.

EXAMPLE 5.5. Let $X = \{x_1, x_2\}$, and let

$$P = \{(x_1^4, x_1^5), (x_2^4, x_2^5), (x_1^4, x_2^4), (x_1^3 x_1, x_1^4), (x_1^2 x_1^2, x_1^4), (x_1 x_2^3, x_1^4)\}.$$

https://doi.org/10.1017/S0004972708001111 Published online by Cambridge University Press

Consider the ideal J generated in $R = F[X]/I_P$ by $g = (x_1 - x_1^2)(x_2 - x_2^2)$. We have $\mathrm{wt}(g) = 4$. However, $x_1 g = x_1^2 x_2 - x_1^4$ in R. Hence, $\mathrm{wt}(x_1 g) = 2 < \mathrm{wt}(g)$. Therefore, $\mathrm{wt}(J) < \mathrm{wt}(g)$.

The next example shows that, for a set P satisfying the hypothesis of Theorem 4.2, there may exist another set Q such that $I_P = I_Q$, but Q does not contain all pairs (x^{p+1}, x), for all $x \in X$.

EXAMPLE 5.6. Let $m = 1$, $X = \{x\}$, $P = \{(x, x^{p+1})\}$ and $Q = \{(x, x^{2p+1}), (x^{p+1}, x^{2p+1})\}$. Then it is clear that $I_P = I_Q$ and so condition (ii) of Theorem 4.1 is satisfied. However, P does not contain the pair (x^{p+1}, x).

The following example shows that the analogue of Theorem 4.2 is not valid for the case of $\mathrm{char}(\mathbb{F}) > 2$. In other words, Theorem 4.1 cannot be generalised to include the case of all unions of p-groups.

EXAMPLE 5.7. Let $p > 2$, $m = 3$, $X = \{x_1, x_2, x_3\}$, $P = \{(x_1, x_1^3), (x_2, x_2^3), (x_3, x_3^3), (x_1 x_3, x_3^2), (x_2 x_3, x_3^2)\}$, and let $g = \{(x_1 - x_1^2)(x_2 - x_2^2)\}$. Then M_X/ϱ_P is a union of 2-groups $\langle x_1, x_2\rangle$ and $\langle x_3\rangle$. Consider the ideal J generated by g in $\mathbb{F}[X]/I_P$. Obviously, $\mathrm{wt}(g) = 4$. Since $x_3 g = 2x_3^2 - 2x_3$, we obtain $x_3 g = 2$ in $\mathbb{F}[X]/I_P$. Therefore, $\mathrm{wt}(J) < \mathrm{wt}(g)$.

6. Proofs

LEMMA 6.1. *Let P be an arbitrary subset of M_X^2, $R = \mathbb{F}[X]/I_P$, and let $b \in R$ be a polynomial of the form (4.1), that is,*

$$b = \prod_{j=1}^{k} (w_j^2 - w_j)^{d_j}, \tag{6.1}$$

for some $w_j \in M_X$, $d_j \in \mathbb{N}_0$. Then b can be represented in the form

$$b = \sum_{i=1}^{\ell} r_i \prod_{j=1}^{k} (x_{i,j}^2 - x_{i,j})^{d_j}, \tag{6.2}$$

for some $r_i \in R$, $x_{i,j} \in X$, $d_i \in \mathbb{N}_0$.

Representation (6.2) is equivalent to saying that b belongs to the Drensky class set generated by the products $\prod_{j=1}^{k} (x_{i,j}^2 - x_{i,j})^{d_j}$, for $i = 1, \ldots, \ell$.

PROOF. We proceed by induction on the maximum degree d_m of all monomials $w_j \in M_X$, $j = 1, \ldots, k$. The induction basis, where $d_m = 1$ and all $w_j \in X$, is trivial, because then b itself is of the form (6.2). Further, we assume that $d_m > 1$ and that the assertion has been proved for smaller values of d_m.

https://doi.org/10.1017/S0004972708001111 Published online by Cambridge University Press

Consider any $1 \leq j \leq k$. If $\deg(w_j) < d_m$, then the induction assumption allows us to express $(w_j^2 - w_j)^{d_j}$ as

$$(w_j^2 - w_j)^{d_j} = \sum_{i=1}^{m} a_i \prod_{j=1}^{k} (x_{i,j}^2 - x_{i,j})^{d_j}, \tag{6.3}$$

for some $a_i \in R$. On the other hand, if $\deg(w_j) = d_m$, then there exist $u, v \in M_X$ such that $w_j = uv$ and $\deg(u), \deg(v) < d_m$. Then we can represent each of the elements $u^2 - u$ and $v^2 - v$ in the form (6.3) and substitute these representations into the equality

$$(u^2v^2 - uv) = u^2(v^2 - v) + v(u^2 - u).$$

This demonstrates that $(w_j^2 - w_j)^{d_j}$ can be expressed in the form (6.3) in this case again.

If we substitute all expressions (6.3) for all w_j into (6.1) and apply the distributive law, then a representation (6.2) for b follows. This completes the proof. □

PROOF OF THEOREM 4.4. The implication (ii) ⇒ (i) is trivial, because every Drensky class set is a binomial class set. Let us prove the reverse implication.

(i) ⇒ (ii): Suppose that condition (i) holds. Choose any binomial class set $C = C(B)$, generated by a set $B = \{b_1, \ldots, b_k\}$, where all of the b_i satisfy (4.1). Lemma 6.1 implies that each polynomial b_i can be represented in the form (6.2). Since $b_i \neq 0$ and P contains all pairs (x^{p+1}, x), we see that the monogenic subsemigroup generated by each $w_{i,j}$ is a cyclic group of order p. The same is also true of every $x_{i,j}$ occurring in (6.3) and in the resulting expression (6.2) for b_i. It follows that b_i belongs to a Drensky class set with standard generator polynomials having the same weights as b_i. Since the Drensky class set is visible, its weight is equal to the minimum weight of these generators. Hence, it follows that the weight of C is also equal to the minimum of the weights of the generating elements b_i. This completes our proof. □

Let S be a semigroup. An $\mathbb{F}$-algebra R is said to be *S-graded*, if $R = \bigoplus_{s \in S} R_s$ is a direct sum of $\mathbb{F}$-modules R_s and $R_s R_t \subseteq R_{st}$, for all $s, t \in S$ (see [14] and [13]). The $\mathbb{F}$-modules R_s are called the *homogeneous components* of the grading. Let $R = \bigoplus_{s \in S} R_s$ be an S-graded ring. An element of R is said to be *S-homogeneous*, or *homogeneous*, if it belongs to the union $\bigcup_{s \in S} R_s$ of the homogeneous components. An ideal I of R is said to be *homogeneous*, or *S-homogeneous*, if it is equal to the sum

$$I = \bigoplus_{s \in S} I \cap R_s. \tag{6.4}$$

LEMMA 6.2. *Let S be a semigroup, $\bigoplus_{s \in S} A \cap R_s$ a finite-dimensional S-graded $\mathbb{F}$-algebra, and let B be a basis of R regarded as a linear space over $\mathbb{F}$. Suppose that B consists entirely of homogeneous elements, and I is a homogeneous ideal of R. Then every nonzero element of minimum weight in I is homogeneous, and in particular the weight of I is equal to the minimum weight of a nonzero homogeneous element in I.*

https://doi.org/10.1017/S0004972708001111 Published online by Cambridge University Press

PROOF. Denote the elements of B by $b_1, \ldots, b_k$, so that $B = \{b_1, \ldots, b_k\}$. Choose a nonzero element r with minimum weight in I. Look at the expression $r = r_1 b_1 + \cdots + r_k b_k$. Since $r \neq 0$, there are nonzero coefficients r_i in this expression. Without loss of generality, we may assume that $r_1 \neq 0$. Since B consists of homogeneous elements, for each $i = 1, \ldots, k$, there exists $s_i \in S$ such that $b_i \in R_{s_i}$. We can reorder the vectors in the basis and collect all basis vectors, which belong to R_{s_1}, in the beginning of the basis. Then we may assume that $s_1 = \cdots = s_\ell$, for some $1 \leq \ell \leq k$, and that $s_i \neq s_1$ for all $i = \ell + 1, \ldots, k$. It follows that the s_1-component r_{s_1} of r is equal to

$$r_{s_1} = \sum_{i=1}^{\ell} r_i b_i.$$

Since I is homogeneous, (6.4) implies that $r_{s_1} \in I$. By the minimality of $\mathrm{wt}(r)$, we obtain $r = r_{s_1}$. Thus, r is a homogeneous element, as required. □

LEMMA 6.3. *Let $G = G_2 \times G_p$ be a direct product of an elementary abelian 2-group G_2 and an elementary abelian p-group G_p. Then every Drensky class set I_D, $D \subseteq \mathbb{N}_0^m$, in the group algebra $\mathbb{F}[G]$ is visible.*

PROOF. If $p = 2$, then $G_2 \times G_p$ is an elementary abelian p-group, and our lemma coincides with the assertion of Proposition 3.4. Further, we assume that $p > 2$.

There exist positive integers m_1 and m_2 such that $G_2 \cong \mathbb{Z}_2^{m_1}$ and $G_p \cong \mathbb{Z}_p^{m_2}$. Put $X_1 = \{x_1, \ldots, x_{m_1}\}$, $X_2 = \{x_{m_1+1}, \ldots, x_{m_1+m_2}\}$ and $X = X_1 \cup X_2$. Then $\mathbb{F}[G] \cong \mathbb{F}[X]/I_P$, where

$$P = \{(1, x_i^2) \mid i \in [1, m_1]\} \cup \{(1, x_i^p) \mid i \in [m_1 + 1, m_1 + m_2]\}.$$

Fix any i such that $1 \leq i \leq m_1$. It is known that every Drensky class set is visible in the group algebra $\mathbb{F}[x_i]/(1 - x_i^p) \cong \mathbb{F}[\mathbb{Z}_p]$; and this fact was used in the proof of Proposition 3.4 in [7]. Now we claim that, in a similar fashion, for $p \neq 2$, every Drensky class set is also visible in the group algebra $R_i = \mathbb{F}[x_i]/(1 - x_i^2) \cong \mathbb{F}[\mathbb{Z}_2]$.

Indeed, let us first consider the ideal J generated by $g = 1 - x_i$ in R_i. It is easily seen that J is equal to the *augmentation ideal* of R_i, that is, the set

$$\left\{ \sum_{s \in \mathbb{Z}_2} r_s s \;\middle|\; \sum_{s \in \mathbb{Z}_2} r_s = 0 \right\}.$$

Therefore, $\mathrm{wt}(J) = 2 = \mathrm{wt}(g)$, and so J is visible.

Further, consider the element g^d, for a positive integer d. Easy induction on d shows that $g^d = 2^{d-1} g$. Since $p > 2$, it follows that g^d generates the same ideal J as g and $\mathrm{wt}(g^d) = 2$. Therefore, $\mathrm{wt}(J) = \mathrm{wt}(g^d)$ again. Thus, every Drensky class set in the group algebra R_i coincides with J and is visible.

Keeping this fact in mind it is routine to verify that all steps of the proof of Proposition 3.4 given in [7] remain valid in our more general situation. It follows that the exact analogue of Proposition 3.4 holds for every direct product of an elementary abelian 2-group and an elementary abelian p-group. This completes the proof. □

https://doi.org/10.1017/S0004972708001111 Published online by Cambridge University Press

REMARK 6.4. An alternative proof of Lemma 6.3 follows from the main theorem of [21, Section 2], which uses the notion of a visible basis of a vector space. Indeed, it is easily seen that every visible generating set of an ideal generates (with respect to multiplication) a visible basis of the ideal regarded as a vector space.

PROOF OF THEOREM 4.2. Let $\mathbb{F}$ be a finite field with $\mathrm{char}(\mathbb{F}) = 2$. Since M_X is commutative and P contains all pairs (x^3, x), for all $x \in X$, it follows that the monoid $M = M_X/\varrho_P$ satisfies the identity $x^3 = x$, for all $x \in M$. Therefore, the monogenic subsemigroup $\langle x \rangle$ is isomorphic to the cyclic group $\mathbb{Z}_3$, for each $x \in M$. Hence, M is a union of cyclic groups isomorphic to $\mathbb{Z}_3$.

Let Y be a semilattice. A semigroup S is said to be a Y*-semilattice* of subsemigroups S_y, $y \in Y$, if $S = \bigcup_{y \in Y} S_y$ is a disjoint union of the S_y, and $S_x S_y \subseteq S_{xy}$ for all $x, y \in Y$.

Denote by Y the subsemigroup generated in M by all elements x^2 for all $x \in X$. For any $y \in Y$, put

$$G_y = \{x \in M \mid x^2 = y\}.$$

It is straightforward to verify that Y is a semilattice, every G_y is an elementary abelian 2-group, and $M = \bigcup_{y \in Y} G_y$ is a semilattice of groups G_y. This fact is well known and is recorded, for example, as [15, Proposition 2.1]. Hence, it follows that $\mathbb{F}[M] = \bigoplus_{y \in Y} \mathbb{F}[G_y]$ is a Y-graded ring.

For $y \in Y$, denote by e_y the identity of the elementary abelian 2-group G_y. It is convenient to keep in mind the fact that every semilattice is a partially ordered set with respect to the natural order $\leq$ defined by the rule $x \leq y \Leftrightarrow xy = x$.

Choose an arbitrary subset D of $\mathbb{N}_0^2$, and consider the Drensky class set I_D in $R = \mathbb{F}[X]/I_P$. We claim that the weight of I_D is equal to the minimum of the weights of the generators in the set U_D defined by (3.6).

Take a nonzero element r of minimal weight in I_D. It follows from (3.5) that $r \in I_D = C(U_D)$ can be represented in the form

$$r = \sum_{d \in D} r_d u_d, \tag{6.5}$$

where $r_d \in \mathbb{F}[M]$. Here $r_d = \sum_{y \in Y} r_{d,y}$, where $r_{d,y} = (r_d)_y \in \mathbb{F}[G_y]$ for each $y \in Y$. Therefore,

$$r = \sum_{d \in D} \sum_{y \in Y} r_{d,y} u_d, \tag{6.6}$$

where $r_{d,y} \in \mathbb{F}[G_y]$.

Lemma 6.2 implies that r is Y-homogeneous, and so $r = r_v$ for some $v \in Y$. Obviously, every generator u_d belongs to the ring $\mathbb{F}[G_{y_d}]$ for some $y_d \in Y$. Therefore, (6.6) can be rewritten as

$$r = r_v = \sum_{y y_d = v} r_{d,y} u_d. \tag{6.7}$$

https://doi.org/10.1017/S0004972708001111 Published online by Cambridge University Press

We may assume that all summands in (6.7) are nonzero and similar terms have been combined.

Let us consider any term $r_{d,y}u_d$ in (6.7). By (3.6), $u_d = \prod_{i=1}^{m}(x_i^2 - x_i)^{d_i} \in \mathbb{F}[G_{y_d}]$. Let $e = e_{y_d}$ be the identity of G_{y_d}. Then we get $u_d = eu_d = \prod_{i=1}^{m}((ex_i)^2 - ex_i))^{d_i}$. Since M is a union of 2-groups, $\mathrm{char}(\mathbb{F}) = 2$ and $u_d \neq 0$, we see that $d_i \in \{0, 1\}$ for all i. Therefore,

$$u_d = eu_d = \prod_{i=1}^{m}((ex_i)^2 - ex_i). \tag{6.8}$$

It follows from the definition of $R = \mathbb{F}[X]/I_P$ that $ex_i \in M_X$ for all i. Hence, u_d generates a binomial class set $C(u_d)$ in $\mathbb{F}[G_{y_d}]$.

It follows from (6.8) that there exists a subgroup H_{y_d} of G_{y_d} such that $u_d = \sum_{h\in H_y} h$. Condition $yy_d = v$ in (6.7) implies that $e_y e_{y_d} = e_v$. Therefore, we can rewrite the term $r_{d,y}u_d$ as follows

$$\begin{aligned} r_{d,y}u_d &= (e_v r_{d,y})(e_v u_d) \\ &= (e_v r_{d,y})\bigg(e_v \sum_{h\in H_y} h\bigg). \end{aligned} \tag{6.9}$$

Obviously, $e_v H_y$ is a subgroup of G_v. Lagrange's theorem implies that $|e_v H_y|$ divides $|G_v|$, and so it is a power of two. Likewise, $|e_v H_y|$ divides $|H_y|$. Hence, $|H_y|/|e_v H_y|$ is a power of two. It is straightforward to verify that

$$e_v \sum_{h\in H_y} h = \frac{|H_y|}{|e_v H_y|} \sum_{h\in e_v H_y} h. \tag{6.10}$$

Since $\mathrm{char}(\mathbb{F}) = 2$, we see that $e_v u_d \neq 0$ implies $|H_y| = |e_v H_y|$. It follows that every nonzero element $e_v u_d$ in (6.9) has weight equal to $\mathrm{wt}(u_d)$, and is a binomial generator of the form (4.1). Hence, all of the $e_v u_d$ generate a binomial class set C_B in $\mathbb{F}[G_v]$. Proposition 3.4 and Theorem 4.4 show that C_B is visible. Hence, it follows that $\mathrm{wt}(r)$ is not less than the minimum weight $\mathrm{wt}(u_d)$ for some $d \in D$. Therefore, $\mathrm{wt}(r) = \mathrm{wt}(u_d)$. This completes our proof. □

Proof of Theorem 4.1. First, consider the case where $p = 2$. Then $M = M_X/\varrho_P$ is a subsemigroup of a product of a semilattice and an elementary abelian 2-group. Hence, M satisfies the identity $x^3 = x$, for all $x \in M$. Therefore, there exists $Q \subseteq M_X^2$ such that $I_P = I_Q$ and Q contains all pairs (x^3, x), for all $x \in X$. Thus, the hypotheses of Theorem 4.2 are satisfied for the set Q. Theorem 4.2 implies that every Drensky class set in $\mathbb{F}[X]/I_Q$ is visible. Since $\mathbb{F}[X]/I_P = \mathbb{F}[X]/I_Q$, we see that Theorem 4.1 holds in this case.

Further, we assume that $p > 2$. Let Y be the subsemigroup generated in $M = M_X/\varrho_P$ by the set E of all elements $e_i = x_i^2$, $i = 1, \ldots, m_1$, and all elements $e_i = x_i^p$, $i = m_1, \ldots, m$. Let z be the product of all elements in Y. Denote by G_2 the

https://doi.org/10.1017/S0004972708001111 Published online by Cambridge University Press

multiplicative subgroup generated in M by all elements $zx_1, \ldots, zx_{m_1}$. Let G_p be the multiplicative subgroup generated in M by all elements $zx_{m_1+1}, \ldots, zx_m$. Given that M is isomorphic to a subsemigroup of a direct product of a semilattice, an elementary abelian 2-group, and an elementary abelian p-group, a tedious but routine verification shows that Y is a semilattice with zero z, G_2 is an elementary abelian 2-group, G_p is an elementary abelian p-group, and M is isomorphic to a subsemigroup of $Y \times G_2 \times G_p$. Therefore, $R = \mathbb{F}[X]/I_P$ is isomorphic to a subring of the monoid algebra $\mathbb{F}[Y \times G_2 \times G_p]$.

Choose a subset D of $\mathbb{N}_0^2$ and consider the Drensky class set I_D. We claim that the weight of I_D is equal to $\mathrm{wt}(U_D)$, that is, the minimum of the weights of the generators in the set U_D. Obviously, it is enough to prove the inequality $\mathrm{wt}(I_D) \geq \mathrm{wt}(U_D)$.

Take a nonzero element r of minimal weight in I_D. It follows from (3.5) that

$$r = \sum_{d \in D} r_d u_d, \tag{6.11}$$

where $r_d \in \mathbb{F}[M]$. Lemma 6.2 implies that r is Y-homogeneous, and so $r = r_v$ for some $v \in Y$.

It is easily seen that zR is an ideal of R isomorphic to the group algebra $\mathbb{F}[G]$, where $G = zM_X \cong G_2 \times G_p$. This and (6.11) imply that r has the same weight as the element

$$zr = \sum_{d \in D} (zr_d)(zu_d), \tag{6.12}$$

which belongs to the Drensky class set generated in the group algebra $\mathbb{F}[G]$ by the elements zu_d, for $d \in D$. Since $\mathrm{wt}(zu_d) = \mathrm{wt}(u_d)$, for all d, it follows from Lemma 6.3 that the weight of r is not less than the minimum of the weights of u_d, for $d \in D$. This completes our proof, because $\mathrm{wt}(r) = \mathrm{wt}(I_D)$. □

In conclusion let us note that formulas for the maximum number of errors, which can be corrected by multiple classifiers and clusterers defined by ideals and one-sided ideals in the algebras of Brandt semigroups and Rees matrix semigroups have been obtained in [18] and [19], respectively.

References

[1] R. Alfaro and A. V. Kelarev, 'Recent results on ring constructions for error-correcting codes', *Contemp. Math.* **376** (2005), 1–12.

[2] ———, 'On cyclic codes in incidence rings', *Studia Sci. Math. Hungarica* **43**(1) (2006), 69–77.

[3] I. M. Araújo, A. V. Kelarev and A. Solomon, 'An algorithm for commutative semigroup algebras which are principal ideal rings with identity', *Comm. Algebra* **32**(4) (2004), 1237–1254.

[4] A. M. Bagirov, A. M. Rubinov and J. Yearwood, 'A global optimization approach to classification', *Optim. Eng.* **3** (2002), 129–155.

[5] J. Cazaran and A. V. Kelarev, 'Generators and weights of polynomial codes', *Arch. Math. (Basel)* **69** (1997), 479–486.

[6] J. Cazaran, A. V. Kelarev, S. J. Quinn 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.

[7] V. Drensky and P. Lakatos, *Monomial Ideals, Group Algebras and Error-correcting Codes*, Lecture Notes in Computer Science, 357 (Springer, Berlin, 1989), pp. 181–188.

[8] V. S. Drensky, *Free Algebras and PI-Algebras* (Springer, Singapore, 2000).

[9] V. S. Drensky, A. Giambruno and S. K. Sehgal, *Methods in Ring Theory* (Marcel Dekker, New York, 1998).

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

[11] J. East, 'Cellularity of inverse semigroup algebras', submitted.

[12] T. E. Hall, 'The radical of the algebra of any finite semigroup over any field', *J. Aust. Math. Soc. Ser. A* **11** (1970), 350–352.

[13] A. V. Kelarev, 'On classical Krull dimension of group-graded rings', *Bull. Aust. Math. Soc.* **55** (1997), 255–259.

[14] ———, *Ring Constructions and Applications* (World Scientific, River Edge, NJ, 2002).

[15] ———, *Graph Algebras and Automata* (Marcel Dekker, New York, 2003).

[16] A. Kelarev, B. Kang and D. Steane, *Clustering Algorithms for ITS Sequence Data with Alignment Metrics*, Lecture Notes in Artificial Intelligence, 4304 (Springer, Berlin, 2006), pp. 1027–1031.

[17] A. V. Kelarev and D. S. Passman, 'A description of incidence rings of group automata', *Contemp. Math.* **456** (2008), 27–33.

[18] A. V. Kelarev, J. L. Yearwood and M. A. Mammadov, 'A formula for multiple classifiers in data mining based on Brandt semigroups', *Semigroup Forum*, to appear, DOI:10.1007/s00233-008-9098-9.

[19] A. V. Kelarev, J. L. Yearwood and P. Watters, 'Rees matrix constructions for data mining', submitted.

[20] G. Luger, *Artificial Intelligence. Structures and Strategies for Complex Problem Solving*, 5th edn (Addison-Wesley, Reading, MA, 2005).

[21] H. N. Ward, 'Visible codes', *Arch. Math. (Basel)* **54** (1990), 307–312.

[22] I. H. Witten and E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques* (Elsevier, Amsterdam, 2005).

[23] J. L. Yearwood, A. M. Bagirov and A. V. Kelarev, 'Optimization methods and the k-committees algorithm for clustering of sequence data', submitted.

[24] J. L. Yearwood and M. Mammadov, *Classification Technologies: Optimization Approaches to Short Text Categorization* (Idea Group Incorporated, Hershey, PA, 2007).

A. V. KELAREV, School of Information Technology and Mathematical Sciences, University of Ballarat, PO Box 663, Ballarat, Victoria 3353, Australia
e-mail: a.kelarev@ballarat.edu.au

J. L. YEARWOOD, School of Information Technology and Mathematical Sciences, University of Ballarat, PO Box 663, Ballarat, Victoria 3353, Australia
e-mail: j.yearwood@ballarat.edu.au

P. W. VAMPLEW, School of Information Technology and Mathematical Sciences, University of Ballarat, PO Box 663, Ballarat, Victoria 3353, Australia
e-mail: p.vamplew@ballarat.edu.au