Jordan algebras arising from intermolecular recombination

Murray R. Bremner
Department of Mathematics and Statistics
University of Saskatchewan
McLean Hall (Room 142), 106 Wiggins Road
Saskatoon, SK, S7N 5E6, Canada
bremner@math.usask.ca

10 August 2005

Abstract

We use computer algebra to show that a linearization of the operation of intermolecular recombination from theoretical genetics satisfies a nonassociative polynomial identity of degree 4 which implies the Jordan identity. We use the representation theory of the symmetric group to decompose this new identity into its irreducible components. We show that this new identity implies all the identities of degree ≤ 6 satisfied by intermolecular recombination.

Introduction

In the theory of DNA computing many operations on formal languages are studied which model the processes of molecular genetics. Computer scientists approach these operations through the theory of monoids. Linearization of these operations permits an approach through the theory of nonassociative algebras (Bremner [2]). In particular the operation of intermolecular recombination (Landweber and Kari [6], operation (2), page 9) has the form

$$uxv + u'xv' \Longrightarrow uxv' + u'xv, \tag{1}$$

where u, u', v, v', x are words over some alphabet S. This notation indicates the replacement of two strings with a common substring by two other strings. (For the general theory of DNA computing see Păun [9].) Operation (1) can be interpreted as a bilinear product on a free partially commutative associative algebra. We show that the resulting nonassociative operation is commutative and satisfies a polynomial identity of degree 4 which implies the Jordan identity. We use the representation theory of the symmetric group to decompose this new identity into its irreducible components. We show further that this identity implies all the identities of degree ≤ 6 satisfied by intermolecular recombination. Our proofs are computational: we use Maple 9.5 (especially the packages `LinearAlgebra` and `LinearAlgebra[Modular]`) running on a Sun Blade 1000 with 512 megabytes of RAM.

Jordan algebras first appeared in mathematical genetics in the work of Holgate [4] on population genetics; the present paper establishes a connection between Jordan algebras and theoretical molecular genetics. (For the general theory of Jordan algebras see McCrimmon [7].)

1 Preliminaries

Let $S = \{x_1, \ldots, x_m\}$ and $T = \{y_1, \ldots, y_n\}$ be two finite nonempty sets (the alphabets); we may assume that $S = T$ but this is not necessary. Let $\mathbb{F}$ be a field and let A be the free partially commutative associative algebra on $S \cup T$ over $\mathbb{F}$: that is, A is the associative algebra generated by $S \cup T$ subject only to the relations $x_i y_j = y_j x_i$ for $i = 1, \ldots, m$ and $j = 1, \ldots, n$. Let $M(S)$ be the free monoid on S; that is, the set of all finite words $a = a_1 \cdots a_\ell$ where $a_k \in S$ for $k = 1, \ldots, \ell$, with concatenation as the associative binary operation. (We allow $\ell = 0$ giving the empty word 1.) Since any basis monomial for A can be written (using partial commutativity) in the form $a'a''$ with $a' \in M(S)$, $a'' \in M(T)$, we see that the algebra A has a basis over $\mathbb{F}$ consisting of the ordered pairs $(a', a'') \in M(S) \times M(T)$. With respect to this basis the natural associative operation on A takes the form $(a', a'')(b', b'') = (a'b', a''b'')$. We introduce a new operation $\triangleright$ (splicing) on A defined on basis elements by

$$(a', a'') \triangleright (b', b'') = (a', b''), \tag{2}$$

and extended bilinearly to all of A. It is easy to check that this operation is again associative. We now define the operation $\circ$ (intermolecular recombination) to be the Jordan product obtained from (2) defined on basis elements by

$$(a', a'') \circ (b', b'') = (a', a'') \triangleright (b', b'') + (b', b'') \triangleright (a', a'') = (a', b'') + (b', a''),$$

and extended bilinearly. This is the nonassociative linearization of operation (1). The substring x is now superfluous; equivalently, we take x to be the empty word. (Taking x to be nonempty is related to the theory of mutations of algebras; see Elduque and Myung [3].) From this definition and the fact that splicing is associative it follows immediately that intermolecular recombination is commutative and satisfies the Jordan identity. However we will see that it satisfies a much stronger identity.

We assume that the base field $\mathbb{F}$ has characteristic 0; this implies that every polynomial identity over $\mathbb{F}$ is equivalent to a set of homogeneous multilinear identities (Zhevlakov [11], Chapter 1). From now on we omit the symbol $\circ$ and write intermolecular recombination simply as juxtaposition:

$$ab = (a', a'')(b', b'') = (a', b'') + (b', a''). \tag{3}$$

	$(ab)c$	$(ac)b$	$(bc)a$
(a', b'')	0	1	1
(a', c'')	1	0	1
(b', a'')	0	1	1
(b', c'')	1	1	0
(c', a'')	1	0	1
(c', b'')	1	1	0

Table 1: The expansion matrix in degree 3

The polynomial identities we study are elements of the nullspace of the expansion matrix. For degree n, this matrix has columns labelled by the commutative nonassociative monomials in the variables a_k for $i = 1, \ldots, n$ where $a_k = (a'_k, a''_k)$, and rows labelled by the ordered pairs (a'_k, a''_ℓ) with $1 \leq k \neq \ell \leq n$.

Proposition 1. *The expansion matrix in degree n has size $n(n-1) \times (2n-3)!!$ where for odd m the symbol $m!!$ is the product of all odd integers from 1 to m.*

Proof. The number of distinct multilinear commutative nonassociative monomials of degree n equals $(2n-3)!!$ (see Bremner [1], Proposition 1, page 80 for the anticommutative case; since we consider multilinear monomials, the same proof works in the commutative case). Since each variable a_k for $k = 1, \ldots, n$ represents an ordered pair (a'_k, a''_k) we see that $n^2 - n$ distinct ordered pairs can arise as terms in the expansion of each monomial (all ordered pairs (a'_k, a'_ℓ) with $k \neq \ell$). The (i, j) entry of the expansion matrix contains the coefficient of the i-th ordered pair in the expansion of the j-th nonassociative monomial. $\square$

2 Identities of degree 3

Lemma 2. *Every identity of degree ≤ 3 satisfied by intermolecular recombination follows from the commutative identity $ab = ba$.*

Proof. For a commutative nonassociative operation there are three inequivalent multilinear monomials in degree 3: $(ab)c$, $(ac)b$, $(bc)a$. Since we consider multilinear identities it suffices to replace the variables a, b, c by basis elements in $M(S) \times M(T)$. So we write $a = (a', a'')$, $b = (b', b'')$, $c = (c', c'')$. When we evaluate any of the three monomials each term in the result is one of the six basis elements (a', b''), (a', c''), (b', a''), (b', c''), (c', a''), (c', b''). We can therefore store the expansions of the monomials in the columns of the 6×3 matrix in Table 1. Any identity of degree 3 for intermolecular recombination which is not a consequence of commutativity must lie in the nullspace of this matrix. The matrix has rank 3, and hence nullspace $\{0\}$. $\square$

$$\begin{pmatrix}
0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 2 & 2 & 0 & 1 & 1 \\
0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 2 & 2 & 0 & 1 & 1 & 0 & 1 \\
1 & 0 & 1 & 0 & 0 & 0 & 2 & 2 & 0 & 1 & 0 & 1 & 1 & 1 & 0 \\
0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 2 & 2 & 0 & 1 & 1 \\
0 & 1 & 0 & 1 & 2 & 2 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 \\
1 & 0 & 2 & 2 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\
0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 2 & 2 & 0 & 1 & 1 & 0 & 1 \\
0 & 1 & 0 & 1 & 2 & 2 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 \\
2 & 2 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\
1 & 0 & 1 & 0 & 0 & 0 & 2 & 2 & 0 & 1 & 0 & 1 & 1 & 1 & 0 \\
1 & 0 & 2 & 2 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\
2 & 2 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1
\end{pmatrix}$$

Table 2: The expansion matrix in degree 4

3 Identities of degree 4

3.1 Computational linear algebra

Theorem 3. *Every identity of degree ≤ 4 satisfied by intermolecular recombination follows from the commutative identity and the identity*

$$(ab)(cd) = ((ab)d)c + ((ac)b)d + ((bc)a)d - 2((ab)c)d. \tag{4}$$

This identity is symmetric in a, b.

Proof. By Lemma 2 it suffices to consider identities of degree 4. By Proposition 1 the expansion matrix has size 12×15 which we create using Maple's `Matrix` command. The row labels of this matrix are the 12 ordered pairs:

$$\begin{array}{cccccc}
(a', b''), & (a', c''), & (a', d''), & (b', a''), & (b', c''), & (b', d''), \\
(c', a''), & (c', b''), & (c', d''), & (d', a''), & (d', b''), & (d', c'').
\end{array}$$

The column labels of this matrix are the 15 nonassociative monomials:

$$\begin{array}{llllll}
((ab)c)d, & ((ab)d)c, & ((ac)b)d, & ((ac)d)b, & ((ad)b)c, & ((ad)c)b, \\
((bc)a)d, & ((bc)d)a, & ((bd)a)c, & ((bd)c)a, & ((cd)a)b, & ((cd)b)a, \\
(ab)(cd), & (ac)(bd), & (ad)(bc). & & &
\end{array} \tag{5}$$

We compute the expansions of the monomials and store the results to get the expansion matrix displayed in Table 2. We now use the Maple procedure `ReducedRowEchelonForm` to compute the row canonical form of the expansion matrix which is displayed in Table 3. (We omit the zero rows at the bottom of the matrix.) This matrix has rank 6, and so its nullspace has dimension 9. We use Maple's `NullSpace` procedure to find a basis for the nullspace which consists of the rows $N_1, \ldots, N_9$ of the matrix N displayed in Table 4.

$$
\begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 1 & 0 & -3 & -3 & -5 & 0 & -3 & -2 & 1 & -2 \\
0 & 1 & 0 & 0 & 0 & -1 & 0 & 3 & 3 & 5 & 1 & 4 & 1 & 0 & 3 \\
0 & 0 & 1 & 0 & 0 & -1 & 0 & -1 & 1 & 0 & -2 & -2 & 1 & -2 & 0 \\
0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 2 & 2 & 0 & 1 & 1 \\
0 & 0 & 0 & 0 & 1 & 1 & 0 & -2 & -1 & -3 & -1 & -3 & 0 & 0 & -2 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 3 & 1 & 3 & 1 & 3 & 1 & 1 & 1
\end{pmatrix}
$$

Table 3: The row canonical form of the expansion matrix in degree 4

$$
\begin{pmatrix}
-1 & 1 & 1 & -1 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
3 & -3 & 1 & -1 & 2 & 0 & -3 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
3 & -3 & -1 & 0 & 1 & 0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
5 & -5 & 0 & -1 & 3 & 0 & -3 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & 2 & -2 & 1 & 0 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
3 & -4 & 2 & -2 & 3 & 0 & -3 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
2 & -1 & -1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
-1 & 0 & 2 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
2 & -3 & 0 & -1 & 2 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}
$$

Table 4: A basis for the nullspace of the expansion matrix in degree 4

For each basis vector in the nullspace, we apply all 24 permutations of a, b, c, d to the corresponding identity, straighten the terms to obtain monomials in the standard form (5), and store the results in a matrix of size 24×15. The row space of this matrix is the S_4-submodule generated by the basis vector. For the basis vectors listed in Table 4, the corresponding dimensions are 3, 6, 6, 6, 6, 6, 9, 9, 9. For row 7, which corresponds to identity (4), the matrix we obtain is displayed in Table 5. Using Maple's `Rank` command, we find that this matrix has rank 9, and so the row space of this matrix equals the nullspace of the expansion matrix. □

Identity (4) can be proved directly using the formulas

$$
\begin{aligned}
(ab)c &= ac + bc, \\
((ab)c)d &= ac + bd + 2cd, \\
(ab)(cd) &= ac + ad + bc + bd,
\end{aligned}
$$

which can be easily derived using definition (3). This however does not prove that identity (4) implies *all* the identities of degree ≤ 4 satisfied by intermolecular recombination.

Corollary 4. *Every commutative nonassociative algebra satisfying identity (4) is a Jordan algebra.*

Proof. Setting $a = b = c$ in (4) gives the Jordan identity $-(a^2 d)a + a^2(da)$. □

$$
\begin{pmatrix}
2 & -1 & -1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
-1 & 2 & 0 & 0 & -1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 1 & 0 & 0 \\
-1 & 0 & 2 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & -1 & 2 & 0 & -1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 \\
0 & -1 & 0 & 0 & 2 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & -1 & -1 & 2 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 1 \\
2 & -1 & -1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
-1 & 2 & 0 & 0 & -1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 1 & 0 & 0 \\
-1 & 0 & -1 & 0 & 0 & 0 & 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & -1 & 2 & 0 & -1 & 0 & -1 & 0 & 0 & 1 \\
0 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 2 & -1 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & 2 & 0 & -1 & 0 & 1 & 0 \\
-1 & 0 & 2 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & -1 & 2 & 0 & -1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 \\
-1 & 0 & -1 & 0 & 0 & 0 & 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & -1 & 2 & 0 & -1 & 0 & -1 & 0 & 0 & 1 \\
0 & 0 & 0 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 2 & -1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & -1 & -1 & 2 & 1 & 0 & 0 \\
0 & -1 & 0 & 0 & 2 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & -1 & -1 & 2 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 1 \\
0 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 2 & -1 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & 2 & 0 & -1 & 0 & 1 & 0 \\
0 & 0 & 0 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 2 & -1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & -1 & -1 & 2 & 1 & 0 & 0
\end{pmatrix}
$$

Table 5: Row space of this matrix is the submodule generated by identity (4)

Question: Is every algebra satisfying identity (4) a special Jordan algebra? (In other words, is it isomorphic to a subspace of an associative algebra A which is closed under the Jordan product $ab + ba$?)

3.2 Representations of the symmetric group

The proof of Theorem 3 shows that identity (4) generates a submodule R of dimension 9 of the S_4-module of all possible multilinear commutative nonassociative polynomials of degree 4. A similar calculation shows that the multilinear form of the Jordan identity generates a submodule of dimension 4. It is in this sense (comparing dimensions) that identity (4) is much stronger than the Jordan identity. We can make precise the difference between identity (4) and the Jordan identity by applying the representation theory of the symmetric group S_4. We use the Maple command `combinat[character](4)` to compute the character table of S_4 which is displayed in Table 6. The rows are labelled by the partitions of 4 which identify the irreducible representations of S_4; the columns are labelled by representatives of the conjugacy classes in S_4 (expressed as products of disjoint cycles). If we know the character of a representation of S_4 (the

	$()$	(ab)	$(ab)(cd)$	(abc)	$(abcd)$
[4]	1	1	1	1	1
[31]	3	1	-1	0	-1
[22]	2	0	2	-1	0
[211]	3	-1	-1	0	1
[1111]	1	-1	1	1	-1

Table 6: The character table of the symmetric group S_4

trace of a representative of each conjugacy class), then we can use the character table to decompose the representation into its irreducible components. (For the representation theory of the symmetric group see James and Kerber [5].)

Proposition 5. *The 9-dimensional representation $r\colon S_4 \to GL(R)$ generated by identity (4) decomposes as*

$$R = [4] \oplus [31] \oplus [22] \oplus [211],$$

the direct sum of one copy of each irreducible representation for each of the first four partitions of 4.

Proof. We apply each conjugacy class representative to the polynomial identity corresponding to each basis vector of the nullspace of the expansion matrix (the rows $N_1, \ldots, N_9$ of the matrix N in Table 4) and use the Maple procedure `LinearSolve` to express the result as a linear combination of the basis vectors. That is, for each conjugacy class representative $\pi \in S_4$ we compute the matrix $r(\pi) \in GL(R)$ defined by the equation

$$r(\pi)(N_j) = \sum_{i=1}^{9} r(\pi)_{ij} N_i.$$

For representative (), the identity permutation, we obtain the 9×9 identity matrix. For the other four representatives we obtain the matrices displayed in Tables 7–10. The list of traces is $[9, 1, 1, 0, 1]$. Since the characters form a basis for the space of all class functions on S_4, we can determine the decomposition by using `LinearSolve` again to express this vector as a linear combination of the rows of the character table. $\square$

For the (multilinear form of the) Jordan identity a similar computation gives a 4-dimensional representation of S_4 with decomposition $[4]\oplus[31]$. The first irreducible summand corresponds to the (multilinear form of the) power associative identity

$$(a^2 a)a = (a^2)^2, \tag{6}$$

and the second irreducible summand corresponds to the identity

$$a(b(cd)) - b(a(cd)) = (a(bc))d - (b(ac))d + c(a(bd)) - c(b(ad)), \tag{7}$$

$$\begin{pmatrix}
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
-1 & -1 & 0 & -1 & -2 & -2 & 0 & -1 & -1 \\
-1 & 2 & 1 & 3 & 1 & 3 & 0 & 0 & 2 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0
\end{pmatrix}$$

Table 7: Representing matrix for (ab)

$$\begin{pmatrix}
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & -1 & 0 & 2 & 2 & -1 & 2 & 0 \\
-1 & -1 & 0 & -1 & -2 & -2 & 0 & -1 & -1 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}$$

Table 8: Representing matrix for $(ab)(cd)$

which states that the commutator of two multiplication operators is a derivation. In addition to these two identities, intermolecular recombination satisfies two other identities corresponding to the irreducible summands $[22] \oplus [211]$ of Proposition 5.

Proposition 6. *The summands* $[22]$ *and* $[211]$ *of Proposition 5 are generated respectively by the identities*

$$\begin{aligned}
((ab)c)d + ((ab)d)c - ((ad)b)c - ((ad)c)b - ((bc)a)d - ((bc)d)a \\
+ ((cd)a)b + ((cd)b)a + 2(ab)(cd) - 2(ad)(bc) = 0, \quad (8) \\
((ab)c)d - ((ab)d)c - ((ac)b)d + ((ac)d)b + ((ad)b)c - ((ad)c)b = 0. \quad (9)
\end{aligned}$$

Proof. To determine the remaining two generators we use the projection operators onto the isotypic components of a representation (Serre [10], Theorem 8, page 21):

$$P_i = \frac{n_i}{24} \sum_{\pi \in S_4} \chi_i(\pi) r(\pi). \qquad (10)$$

Here n_i is the dimension of the i-th irreducible representation of S_4 (from the first column of the character table) and χ_i is the character of that representation (row i of the character table); P_i is the projection onto the corresponding isotypic component (the sum of all subrepresentations of R isomorphic to that

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & -3 & -3 & -5 & -1 & -4 & -1 & 0 & -3 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 2 & 1 & 3 & 1 & 3 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

Table 9: Representing matrix for (abc)

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ -1 & 3 & 3 & 5 & 0 & 3 & 2 & -1 & 2 \\ -1 & -1 & 0 & -1 & -2 & -2 & 0 & -1 & -1 \\ 1 & 1 & -1 & 0 & 2 & 2 & -1 & 2 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -3 & -1 & -3 & -1 & -3 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

Table 10: Representing matrix for $(abcd)$

representation). To evaluate this formula, we first compute the matrix $r(\pi)$ representing each permutation as we did for the conjugacy class representatives in the proof of Proposition 5. We write a procedure to compute the disjoint cycle factorization of each permutation, and then we use the Maple commands `permgroup` and `group[areconjugate]` to determine the conjugacy class of each permutation. We then use Maple's `MatrixAdd` procedure to evaluate the sum (10). For $i = 3, 4$ we obtain the matrices in Tables 11 and 12.

A basis for each isotypic component is the column space of each matrix. By Proposition 5 we know that each isotypic component is in fact irreducible, so the column spaces will have dimensions 2 and 3. We use Maple's `ColumnSpace` procedure to compute a basis for each isotypic component, and then we use `MatrixVectorMultiply` to compute $N^T v$ for each basis vector v, which changes basis from the 9 rows of matrix N to the 15 nonassociative monomials (5). For $i = 3$, the irreducible summand [22] is generated by identity (8), which alternates in the pairs a, c and b, d. For $i = 4$, the irreducible summand [211] is generated by identity (9), which is an alternating sum over b, c, d. $\square$

A classification of irreducible identities of degree 4 for commutative nonassociative algebras is given by Osborn [8].

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & -2 & -2 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & -2 & -2 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & -2 & 4 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & -2 & 4 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 4 & -2 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 4 & -2 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 8 & -4 & -4 \\ 0 & 0 & 0 & 0 & 0 & 0 & -4 & 8 & -4 \\ 0 & 0 & 0 & 0 & 0 & 0 & -4 & -4 & 8 \end{pmatrix}$$

Table 11: Projection matrix for isotypic component [22]

$$\begin{pmatrix} 8 & -8 & -8 & -16 & 0 & -8 & -4 & 4 & -8 \\ 0 & 16 & 8 & 16 & 8 & 16 & 4 & 4 & 8 \\ 0 & 8 & 8 & 8 & 0 & 8 & 4 & 0 & 4 \\ 0 & -8 & -8 & -8 & 0 & -8 & -4 & 0 & -4 \\ 0 & 8 & 0 & 8 & 8 & 8 & 0 & 4 & 4 \\ 0 & -8 & 0 & -8 & -8 & -8 & 0 & -4 & -4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Table 12: Projection matrix for isotypic component [211]

4 Identities of degree 5

Theorem 7. *Every identity of degree ≤ 5 satisfied by intermolecular recombination follows from commutativity and identity (4).*

Proof. We follow the same algorithm as in the proof of Theorem 3; that result also shows that it suffices to consider identities of degree 5. By Proposition 1 the expansion matrix in degree 5 has size 20×105. We compute that this matrix has rank 10, and so there are 95 linearly independent identities satisfied by intermolecular recombination in degree 5. We will show that they are all consequences of identity (4). We write identity (4) in the form

$$I(a,b,c,d) = 2((ab)c)d - ((bc)a)d - ((ac)b)d - ((ab)d)c + (ab)(cd). \tag{11}$$

There are five distinct ways to lift this identity to degree 5:

$$I(ae,b,c,d), \quad I(a,be,c,d), \quad I(a,b,ce,d), \quad I(a,b,c,de), \quad I(a,b,c,d)e. \tag{12}$$

(Since $I(a,b,c,d) = I(b,a,c,d)$ we could simplify the computations a little by omitting the second lifting.) We create a matrix of size 225×105, consisting of a block of size 105×105 on top of a block of size 120×105, and initialize it to zero. We apply all 120 permutations of a,b,c,d,e to the first lifted identity

$I(ae, b, c, d)$ and store the results in the lower block. We then compute the row canonical form of the matrix; the lower block is now zero. We repeat this fill and reduce process with the other four lifted identities, preserving at each stage the relations in the upper block. After all five lifted identities have been processed, the matrix has rank 95, which shows that the space of lifted identities equals the space of all identities in degree 5. $\square$

5 Identities of degree 6

To study identities of degree 6 we will use linear algebra over a finite field as implemented in the Maple package `LinearAlgebra[Modular]`. This is necessary to keep control over the size of the entries when computing the row canonical form of a large matrix. This approach is justified by the following result.

Lemma 8. *The decompositions of any integral representation of S_n over the field $\mathbb{Q}$ of rational numbers and over the field $\mathbb{F}_p$ with p elements where $p > n$ have the same multiplicities of irreducible components.*

Proof. By the representation theory of the symmetric group S_n, we know that the group algebra $\mathbb{F}S_n$ is semisimple for any field $\mathbb{F}$ of characteristic 0 or $p > n$. Furthermore, since $\mathbb{Q}$ is a splitting field for S_n (that is, all the characters take values in $\mathbb{Q}$), for the field $\mathbb{F}_p$ with $p > n$ we have the algebra isomorphism

$$\mathbb{Z}S_n \otimes_{\mathbb{Z}} \mathbb{F}_p \cong \mathbb{F}_p S_n.$$

This guarantees that the decomposition of any integral representation of S_n will be the same over $\mathbb{Q}$ and over $\mathbb{F}_p$ for any $p > n$. $\square$

Theorem 9. *Every identity of degree ≤ 6 satisfied by intermolecular recombination follows from commutativity and identity (4).*

Proof. By Theorem 7 it suffices to consider identities of degree 6. By Proposition 1 the expansion matrix in degree 6 has size 30×945. As in the proof of Theorems 3 and 7, but now using modular arithmetic with $p = 101$, we find that this matrix has rank 15, and so there are 930 linearly independent identities satisfied by intermolecular recombination in degree 6. We will show that they are all consequences of identity (4). Each of the five liftings of $I(a, b, c, d)$ to degree 5 can be lifted to degree 6 in six different ways. After processing these 30 lifted identities in degree 6 we obtain a matrix of rank 930, which shows that the space of lifted identities equals the space of all identities in degree 6. $\square$

Acknowledgements

I thank I. McQuillan for bringing reference [6] to my attention, S. Sehgal for information about representations of the symmetric group, and NSERC (the Natural Sciences and Engineering Research Council of Canada) for financial support through a Discovery Grant.

References

[1] M. Bremner, *Varieties of anticommutative n-ary algebras*, Journal of Algebra 191 (1997) 76–88. MR1444489 (99f:17002)

[2] M. Bremner, *DNA computing, insertion of words and left-symmetric algebras*, pages 229–242 of the Proceedings of the Maple Conference 2005 (July 17-20, 2005), edited by Ilias S. Kotsireas, published by Waterloo Maple Inc., 2005, ISBN 1-894511-85-9.

[3] A. Elduque and H. C. Myung, *Mutations of alternative algebras*, Kluwer, Dordrecht, 1994, ISBN 0-7923-2735-7. MR1282493 (95d:17033)

[4] P. Holgate, *Jordan algebras arising in population genetics*, Proceedings of the Edinburgh Mathematical Society (2) 15 (1966/1967) 291–294. MR0223420 (36 #6468)

[5] G. James and A. Kerber, *The representation theory of the symmetric group*, Addison-Wesley, Reading, 1981, ISBN 0-201-13515-9. MR0644144 (83k:20003)

[6] Laura Landweber and Lila Kari, *The evolution of cellular computing: nature's solution to a computational problem*, Biosystems 52 (1999) 3–13.

[7] K. McCrimmon, *A taste of Jordan algebras*, Springer-Verlag, New York, 2004, ISBN 0-387-95447-3. MR2014924 (2004i:17001)

[8] J. M. Osborn, *Identities of non-associative algebras*, Canadian Journal of Mathematics 17 (1965) 78–92. MR0179221 (31 #3470)

[9] G. Păun, G. Rozenberg and A. Salomaa, *DNA computing: new computing paradigms*, Springer-Verlag, Berlin, 1998, ISBN 3-540-64196-3. MR1724525 (2000m:68063)

[10] J.-P. Serre, *Linear representations of finite groups*, Springer-Verlag, New York-Heidelberg, 1977, ISBN 0-387-90190-6. MR0450380 (56 #8675)

[11] K. A. Zhevlakov, A. M. Slinko, I. P. Shestakov and A. I. Shirshov, *Rings that are nearly associative*, Academic Press, New York-London, 1982, ISBN 0-12-779850-1. MR0668355 (83i:17001)