Prime Number Assignment to a Hexagonal Tessellation of a Plane That Generates Canonical Names for Peri-Condensed Polybenzenes

Seymour B. Elk

Elk Technical Associates, 321 Harris Place, New Milford, New Jersey 07646

Received January 18, 1994®

An algorithm, based on the use of prime numbers, is presented that assigns a canonical name first to each hexagon in a designated domain of a hexagonal tessellation of the plane and then a canonical name to any contiguous set of such hexagons. This algorithm is relatively efficient for uniquely cataloging and naming the classes of peri-condensed and corona-condensed polycyclic aromatic hydrocarbons having all rings of size 6, as well as being viable, but inefficient, for the entire class of such compounds, including those that are cata-condensed—for which other algorithms are much more efficient.

1. INTRODUCTION

The formulation of nomenclature systems for that limited class of chemical compounds which may be mathematically represented by the edge fusion of coplanar regular hexagons, referred to by various authors as "polycyclic aromatic compounds (of ring size 6)", "polyhexes", "polybenzenes", etc., has been an ongoing process for many decades. Although one would probably be correct in ascribing the beginning of this endeavor to the seminal work by Patterson[1] in which an attempt is made to orient and number all types of ring systems—of which benzenoid systems were one (albeit important) special subclass, a more focused beginning is the development of specialized coding schemes a half-century later by Henson, Windlinx, and Wiswesser.[2] This, in turn, was followed by very many more mathematicians and chemists creating a large variety of noteworthy nomenclatures. Just a few of these are listed in ref 3. Following up on the various types of nomenclature systems that we[4] have formulated for this class of molecules, we recently applied some principles from number theory in mathematics to present a greatly simplified algorithm for nomenclating the further limited class of cata-condensed[5] rings using base 5 numbers.[6] Now, we herein present a very different number theory application, viz., prime factorization, by which we can canonically name the entire set of polybenzenes. However, because of its general inefficiency, especially in the case of long-chain cata-condensed compounds—which may be more easily nomenclated by other techniques, the use of the method presented herein is recommended only for peri- and corona-condensed polybenzenes.

2. SCHEMA

Begin at any hexagon in a hexagonal tessellation of a plane and consider the 120° sector with that hexagon as the vertex and which extends 60° above and below the horizontal reference line (Figure 1). Assign prime numbers and powers of these prime numbers to each hexagon of the planar sector using the following technique:

(a) Along the horizontal line, assign first the value 2 to the vertex; then, to each successive hexagon along this line assign the square of the previous number. Note that this produces hexagons named 2 in the first hexagon, $2^2 = 4$ in the second hexagon, $2^4 = 16$ in the third hexagon, etc.; i.e., 2^{2^n} is the name of the nth hexagon in this row.

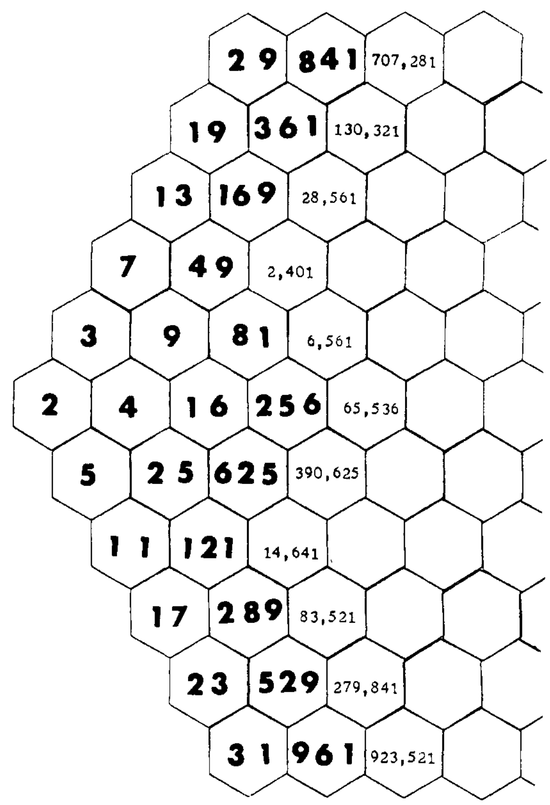

Figure 1.

(b) Along the horizontal line immediately above the reference line described in a, i.e., that starts with the hexagon above the hexagons named 2 and 4, assign the next prime number, namely, 3. Then, to each successive hexagon along this line, assign the square of the previous number. Note that this produces hexagons named 3 in the first hexagon, $3^2 = 9$ in the second hexagon, $3^4 = 81$ in the third hexagon, etc.; i.e., 3^{2^n} is the name of the nth hexagon in this row.

(c) Along the horizontal line immediately below the reference line described in a, i.e., that starts with the hexagon below the hexagons named 2 and 4, respectively, assign 5^{2^n} to the nth such hexagon; i.e., 5, 25, 625, 390 625, etc.

(d) Continue this assignment of prime numbers first above and then below the sector just formed so that every hexagon on the defining 120° sector is given the next prime number, p. Then, the adjacent hexagons in a given row are assigned the value p^2, p^4, ..., p^{2^n}.

® Abstract published in *Advance ACS Abstracts*, May 1, 1994.

0095-2338/94/1634-0942$04.50/0

Table 1. Assignment of Hexagons for Integers 2–100

integer	molecule	integer	molecule	integer	molecule
2	benzene	3	benzene	$4 = 2^2$	benzene
5	benzene	$6 = 2 \times 3$	dibenzene	7	benzene
$8 = 2 \times 2^2$	dibenzene	$9 = 3^2$	benzene	$10 = 2 \times 5$	dibenzene
11	benzene	$12 = 2^2 \times 3$	dibenzene	13	benzene
$14 = 2 \times 7$	disjoint	$15 = 3 \times 5$	dibenzene	$16 = 2^4$	benzene
17	benzene	$18 = 2 \times 3^2$	disjoint	19	benzene
$20 = 2^2 \times 5$	dibenzene	$21 = 3 \times 7$	dibenzene	$22 = 2 \times 11$	disjoint
23	benzene	$24 = 2 \times 2^2 \times 3$	tribenzene	$25 = 5^2$	benzene
$26 = 2 \times 13$	disjoint	$27 = 3 \times 3^2$	dibenzene	$28 = 2^2 \times 7$	disjoint
29	benzene	$30 = 2 \times 3 \times 5$	tribenzene	31	benzene
$32 = 2 \times 2^4$	disjoint	$33 = 3 \times 11$	disjoint	$34 = 2 \times 17$	disjoint
$35 = 5 \times 7$	disjoint	$36 = 2^2 \times 3^2$	dibenzene	37	benzene
$38 = 2 \times 19$	disjoint	$39 = 3 \times 13$	disjoint	$40 = 2 \times 2^2 \times 5$	tribenzene
41	benzene	$42 = 2 \times 3 \times 7$	tribenzene	43	benzene
$44 = 2^2 \times 11$	disjoint	$45 = 3^2 \times 5$	disjoint	$46 = 2 \times 23$	disjoint
47	benzene	$48 = 2^4 \times 3$	disjoint	$49 = 7^2$	benzene
$50 = 2 \times 5^2$	disjoint	$51 = 3 \times 17$	disjoint	$52 = 2^2 \times 13$	disjoint
53	benzene	$54 = 2 \times 3 \times 3^2$	tribenzene	$55 = 5 \times 11$	dibenzene
$56 = 2 \times 2^2 \times 7$	disjoint	$57 = 3 \times 19$	disjoint	$58 = 2 \times 29$	disjoint
59	benzene	$60 = 2^2 \times 3 \times 5$	tribenzene	61	benzene
$62 = 2 \times 31$	disjoint	$63 = 3^2 \times 7$	dibenzene	$64 = 2^2 \times 2^4$	dibenzene
$65 = 5 \times 13$	disjoint	$66 = 2 \times 3 \times 11$	disjoint	67	benzene
$68 = 2^2 \times 17$	disjoint	$69 = 3 \times 23$	disjoint	$70 = 2 \times 5 \times 7$	disjoint
71	benzene	$72 = 2 \times 2^2 \times 3^2$	tribenzene	73	benzene
$74 = 2 \times 37$	disjoint	$75 = 3 \times 5^2$	disjoint	$76 = 2^2 \times 19$	disjoint
$77 = 7 \times 11$	disjoint	$78 = 2 \times 3 \times 13$	disjoint	79	benzene
$80 = 2^4 \times 5$	disjoint	$81 = 3^4$	benzene	$82 = 2 \times 41$	disjoint
83	benzene	$84 = 2^2 \times 3 \times 7$	tribenzene	$85 = 5 \times 17$	disjoint
$86 = 2 \times 43$	disjoint	$87 = 3 \times 29$	disjoint	$88 = 2 \times 2^2 \times 11$	disjoint
89	benzene	$90 = 2 \times 3^2 \times 5$	disjoint	$91 = 7 \times 13$	dibenzene
$92 = 2^2 \times 23$	disjoint	$93 = 3 \times 31$	disjoint	$94 = 2 \times 47$	disjoint
$95 = 5 \times 19$	disjoint	$96 = 2 \times 3 \times 2^4$	disjoint	97	benzene
$98 = 2 \times 7^2$	disjoint	$99 = 3^2 \times 11$	disjoint	$100 = 2^2 \times 5^2$	dibenzene

Note that, by such an assignment, every prime number or selected powers of a prime number (namely, those of the form p^{2n}) will correspond to a single hexagon. More useful, however, is that every composite number formed using exactly two prime numbers, as well as selected combinations of two powers of a single prime number, will correspond to either naphthalene (dibenzene) or to a disjoint set of two hexagons (for which we could set up a system of biphenyl groups, depending on how far apart these two benzene rings are; in practice, however, this second interpretation of disjoint hexagons is not deemed to be worthwhile). In a similar manner we shall find that for every composite number, by examining its prime factorization we can designate a set of hexagons in the described sector. Furthermore, when these hexagons are contiguous, the product of the numbers in these forming hexagons is a name of the polybenzene that this corresponds to. Table 1 lists the first 99 integers that are either prime or composite (i.e., 2–100).

3. CANONICAL NAME

The criterion for selecting the canonical name from the various different sets of hexagons that correspond to a molecule is simply to select the smallest number that specifies this molecule. Focusing attention on Figure 1, we find the smallest set of hexagons that give the desired molecule. For example, dibenzene is $2 \times 3 = 6$; the three tribenzenes (A-tribenzene = anthracene; B-tribenzene = phenanthrene; and C-tribenzene is a nonviable combination of hexagons in a triangular form) would be named as follows: $A = 2 \times 3 \times 7 = 42$, $B = 3 \times 2 \times 5 = 30$, and $C = 2 \times 3 \times 4 = 24$, respectively. This is shown in the first four lines of Table 2, which assigns names to the various mathematically possible polybenzenes (up through the pentabenzenes), using the synthetic nomenclature presented in ref 4a, as a reference. Notice that completion of the entries for Table 2 is achieved by considering all 12 possible orderings of the hexagons in the plane,[4b] such as is shown in Figure 2 for a representative example, JBA-pentabenzene, and by selecting the smallest of these numbers as the canonical name. Note that, in most instances, all but one or two of the 12 possible names is seen to correspond to

Table 2. Assignment of Canonical Names to Polybenzenes

2	benzene
$2 \times 3 = 6$	dibenzene
$2 \times 3 \times 4 = 24$	C-tribenzene
$3 \times 2 \times 5 = 30$	B-tribenzene
$2 \times 3 \times 7 = 42$	A-tribenzene
$2 \times 3 \times 4 \times 5 = 120$	CB-tetrabenzene
$4 \times 2 \times 3 \times 7 = 168$	CA-tetrabenzene
$5 \times 2 \times 3 \times 7 = 210$	BA-tetrabenzene
$5 \times 2 \times 3 \times 9 = 270$	BB-tetrabenzene
$7 \times 3 \times 4 \times 5 = 420$	DB-tetrabenzene
$2 \times 3 \times 7 \times 13 = 546$	AA-tetrabenzene
$3 \times 4 \times 5 \times 16 = 960$	FB-tetrabenzene
$5 \times 2 \times 4 \times 3 \times 7 = 840$	DBA-pentabenzene
$2 \times 3 \times 5 \times 4 \times 9 = 1\,080$	CCA-pentabenzene
$5 \times 2 \times 3 \times 9 \times 7 = 1\,890$	CBA-pentabenzene
$2 \times 4 \times 16 \times 3 \times 5 = 1\,920$	ICA-pentabenzene
$2 \times 4 \times 3 \times 7 \times 13 = 2\,184$	CAA-pentabenzene
$11 \times 5 \times 2 \times 3 \times 7 = 2\,310$	FBA-pentabenzene
$2 \times 4 \times 16 \times 3 \times 7 = 2\,688$	DCA-pentabenzene
$5 \times 2 \times 3 \times 7 \times 13 = 2\,730$	BAA-pentabenzene
$11 \times 5 \times 2 \times 3 \times 9 = 2\,970$	EBA-pentabenzene
$2 \times 4 \times 5 \times 9 \times 11 = 3\,960$	JBA-pentabenzene
$5 \times 2 \times 3 \times 9 \times 16 = 4\,320$	BBB-pentabenzene
$11 \times 5 \times 4 \times 3 \times 7 = 4\,620$	FDB-pentabenzene
$2 \times 3 \times 9 \times 7 \times 13 = 4\,914$	DAA-pentabenzene
$5 \times 4 \times 3 \times 7 \times 13 = 5\,460$	GBA-pentabenzene
$2 \times 4 \times 16 \times 9 \times 5 = 5\,760$	JCA-pentabenzene
$2 \times 4 \times 9 \times 7 \times 13 = 6\,552$	DBB-pentabenzene
$5 \times 2 \times 3 \times 7 \times 49 = 10\,290$	BBA-pentabenzene
$2 \times 3 \times 7 \times 13 \times 19 = 10\,374$	AAA-pentabenzene
$2 \times 4 \times 7 \times 9 \times 25 = 12\,600$	FBB-pentabenzene
$7 \times 9 \times 4 \times 5 \times 11 = 13\,860$	LBA-pentabenzene
$2 \times 4 \times 5 \times 9 \times 81 = 29\,160$	KBA-pentabenzene
$2 \times 4 \times 9 \times 49 \times 25 = 88\,200$	IBA-pentabenzene

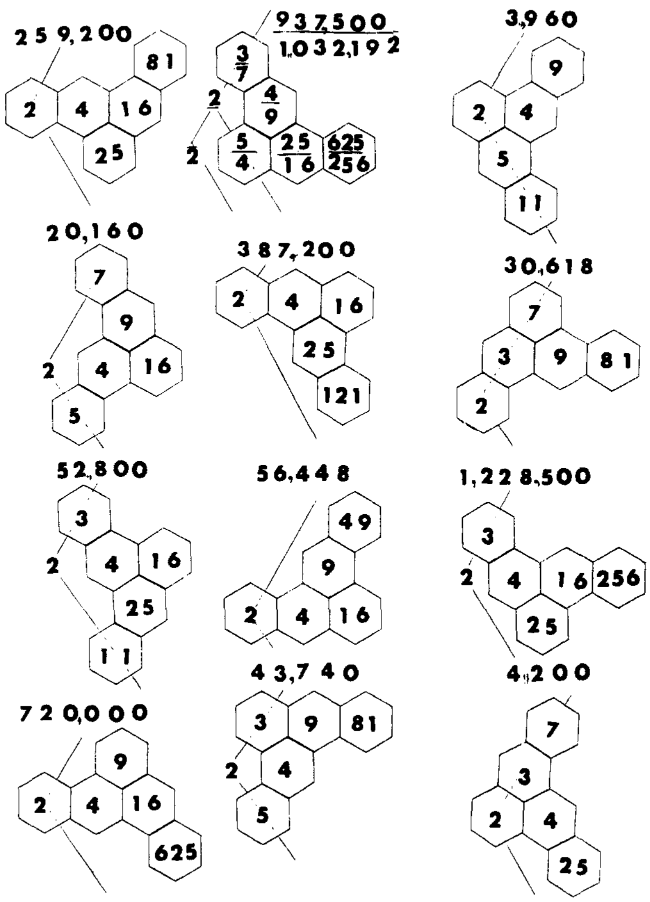

Figure 2.

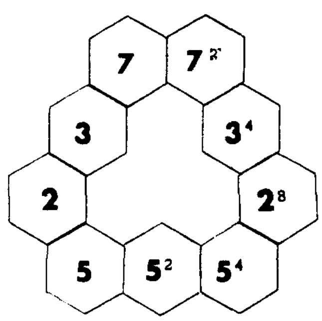

Figure 3.

a much larger number and so can be eliminated by a cursory inspection. Also note that in some instances two or more possible locations of the reference vertex (to be designated "2") may have to be examined. This is shown in the only one of the 12 possible orientations when it was not *immediately* evident which hexagon to choose as the vertex. This is illustrated for the orientation shown in row 1, column 2 of Figure 2 by listing one set of numbers above a dashed line and the other below.

4. USE OF PRIME FACTORS

4.1. Analyzing a Given Name. Starting now from the prime factorization of a given number, we can determine the number of hexagons in the system by breaking down each power of a prime into the sum of powers of 2. For example, consider $4320 = 2^5 \times 3^3 \times 5$. First, notice that $2^5 = 2 \times 16$ and $3^3 = 3 \times 9$. Thus, we have five hexagons named: 2, 3, 5, 9, 16, which from Table 2 we see is the canonical form for BBB-pentabenzene. As a second example, let us consider $3\,333\,960\,000\,000 = 2^9 \times 3^5 \times 5^7 \times 7^3$. Again, grouping the various powers of the prime as products of terms of the form p^{2^n}, since $2^9 = 2^8 \times 2$, $3^5 = 3^4 \times 3$, $5^7 = 5^4 \times 5^2 \times 5$, and $7^3 = 7^2 \times 7$, we thus have 9 hexagons, which, when drawn on a plane, are seen to be a corona-condensed array (Figure 3).

Table 3. Product of n Members of Defining Set

		$n = 3$		
$2 \times 3 \times 4 = 24$	$2 \times 3 \times 5 = 30$	$2 \times 3 \times 7 = 42$	$2 \times 3 \times 9 = 54$	...
	$2 \times 4 \times 5 = 40$	$2 \times 4 \times 7 = 56$	$2 \times 4 \times 9 = 72$	...
		$2 \times 5 \times 7 = 70$	$2 \times 5 \times 9 = 90$	...
			$2 \times 7 \times 9 = 126$	...
	$3 \times 4 \times 5 = 60$	$3 \times 4 \times 7 = 84$	$3 \times 4 \times 9 = 108$	...
		$3 \times 5 \times 7 = 105$	$3 \times 5 \times 9 = 135$	...
			$3 \times 7 \times 9 = 189$	...
		$4 \times 5 \times 7 = 140$	$4 \times 5 \times 9 = 180$	...
			$4 \times 7 \times 9 = 252$	...
			$5 \times 7 \times 9 = 315$	...
		$n = 4$		
$2 \times 3 \times 4 \times 5 = 120$	$2 \times 3 \times 4 \times 7 = 168$	$2 \times 3 \times 4 \times 9 = 216$		...
	$2 \times 3 \times 5 \times 7 = 210$	$2 \times 3 \times 5 \times 9 = 270$		...
		$2 \times 3 \times 7 \times 9 = 378$		...
	$2 \times 4 \times 5 \times 7 = 280$	$2 \times 4 \times 5 \times 9 = 360$		...
		$2 \times 4 \times 7 \times 9 = 504$		...
		$2 \times 5 \times 7 \times 9 = 630$		...
	$3 \times 4 \times 5 \times 7 = 420$	$3 \times 4 \times 5 \times 9 = 540$		...
		$3 \times 4 \times 7 \times 9 = 756$		...
		$3 \times 5 \times 7 \times 9 = 945$		...
		$4 \times 5 \times 7 \times 9 = 1260$		...
		$n = 5$		
$2 \times 3 \times 4 \times 5 \times 7 = 840$	$2 \times 3 \times 4 \times 5 \times 9 = 1080$			
$2 \times 3 \times 4 \times 7 \times 9 = 1512$	$2 \times 3 \times 5 \times 7 \times 9 = 1890$			
$2 \times 4 \times 5 \times 7 \times 9 = 2520$	$3 \times 4 \times 5 \times 7 \times 9 = 3780$			...
		$n = 6$		
$2 \times 3 \times 4 \times 5 \times 7 \times 9 = 7560$	$2 \times 3 \times 4 \times 5 \times 7 \times 11 = 9240$			...

4.2. Simplifying the Synthesis Process. As well as analyzing a given number in terms of its primes, we can also greatly simplify the synthesis process (rather than using the process shown in generating Table 1), by considering the set of composite numbers formed using n different primes or powers of primes of the form p^{2^n}. Note that the members of this set are

$$\{2, 3, 5, 7, 11, ..., 4, 9, 25, 49, ..., 16, 81, ..., 256, ...\}$$

or, more conveniently in sequential order

$$\{2, 3, 4, 5, 7, 9, 11, 13, 16, 17, 19, 23, 25, 29, ...\}$$

Consequently, the smallest number and thus the canonical name for (mono)benzene is the smallest prime: 2. Similarly, for $n = 2$, since again there is one unique dibenzene, we assign as the canonical name the product of the two smallest members of the above named set: $2 \times 3 = 6$. Continuing to higher numbers of hexagons, we tabulate the various possible products of three members of this set, four members of this set, etc. This is shown in Table 3.

Comparing now Tables 2 and 3 for $n = 3$, we see that the first three entries (and three of the four lowest numbers entered) correspond to the canonical names of the three mathematically possible tribenzenes and that all of the other numbers in this table correspond either to less favorable placing of these three contiguous hexagons that comprise one of the mathematically viable tribenzenes or else to a set of hexagons that are not contiguous. Unfortunately, such a perspective—examining just the lowest numbers obtained—is insufficient for larger molecules. For example, examining the tetrabenzenes, the three lowest numbers do correspond to canonical

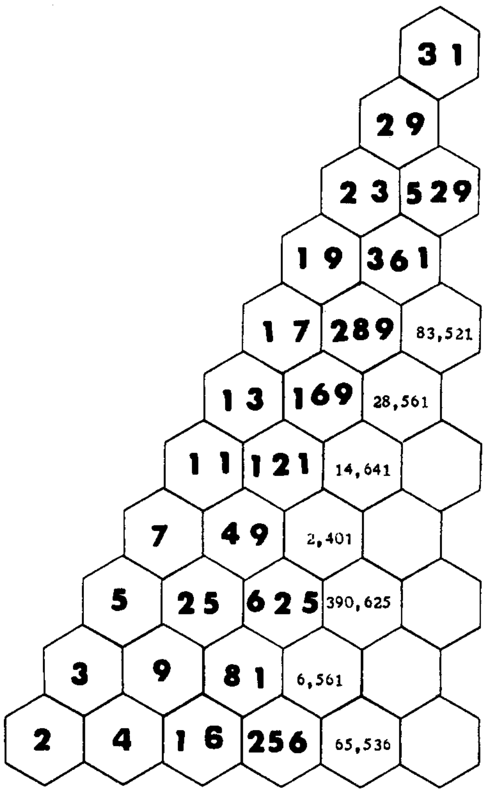

Figure 4.

names of tribenzenes; however, of the fifteen products formed by using just the single digit entries (i.e., the first six members of the set), only five are canonical names, eight are redundancies for hexagons that are not optimally placed, and two are not contiguous. To obtain the other two combinations (AA- and FB-tetrabenzene), we would have to extend this table through the eighth and ninth entries of the above-described set of integers, respectively (i.e., 13 and 16), with the concomitant creation of many more redundancies or noncontiguous combinations before achieving the total listing.

5. OTHER POTENTIAL ALGORITHSM

One other comment of note is the use of different algorithms to assign the sequence of prime numbers to the hexagons: Two important ones to examine are as follows.

(1) Let us start with only a 60° sector and assign sequential primes to the next higher row, rather than alternating odd numbered ($p^{2^{2n-1}}$) vs even numbered ($p^{2^{2n}}$) primes about the reference line, as was selected. This produces a different name for the various hexagons (Figure 4); however, in the case of most, if not all, symmetries, the resultant set of numbers formed for most of the polybenzenes are larger. This alternate assignment is thus eschewed; despite that for the much smaller set of molecules having long chains containing fairly long straight segments, it usually produces somewhat lower canonical names.

(2) Let us assign just the prime numbers to the various hexagons (rather than any powers of primes). One such assignment scheme that emphasizes compactness is achieved by numbering sequential hexagons first along the principal direction, then above, and next below. In other words, 2 and 3 lie on the horizontal line, then 5 is assigned above and 7 below this line creating a 2 × 2 lattice; next 11 extends the principal line and 13 and 17 are one row above and below, while 19 and 23 are two rows above and below this principal line, respectively. This completes the 3 × 3 lattice. Primes

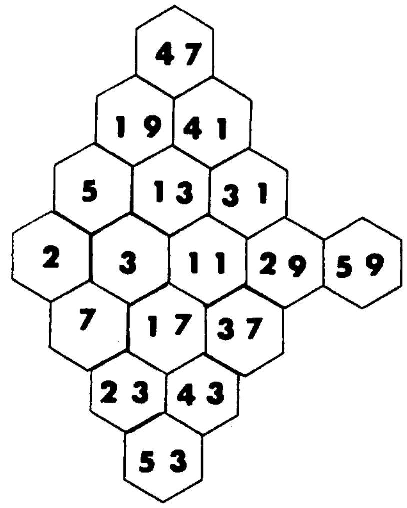

Figure 5.

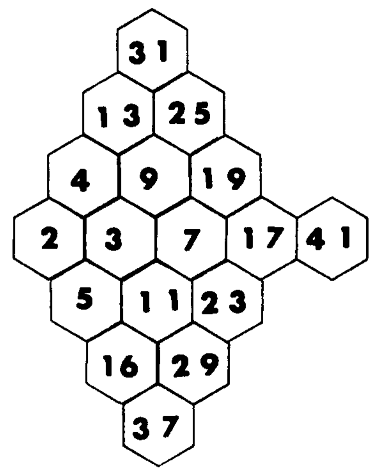

Figure 6.

named 29–53 completes the 4 × 4 lattice, etc. This is shown in Figure 5. Note that a slightly smaller set would be created if, instead of using only primes, we also used numbers of the form p^{2^2}; i.e., the successive hexagons are numbered using the above named sequence (Figure 6). Notice, however, that the saving is minimal and the implementation is tedious. Furthermore, a sequence involving only prime numbers to the first power is much more tedious to decode than the previously described sequence, which usually has a liberal number of powers of 2, 3, and 5, thereby making the final breakdown in prime factors much easier to both accomplish and, once the prime factors are determined, to interpret in terms of which hexagons are used. Consequently, we believe that, in a similar manner to (1), the disadvantages of this alternate assignment again outweighs the single advantage that it also usually produces somewhat lower canonical names than the method described in detail earlier in this report.

6. PRUNING THE SET OF NUMBERS THAT MUST BE EXAMINED

As a final remark, we note that there are many ways of pruning the set of integers that may be examined, such as noting that every combination must contain at least one prime number, rather than all the integers being higher powers of the primes; also, at least one of the n integers chosen must be from the set of $p_1 = 2$ thru p_n; i.e., in the case for $n = 3$ in Table 3, we did not need to tabulate any combinations of three numbers whose lowest member was larger than 5.

Furthermore, in all cases examined so far, at least two prime numbers from this set seem to be included. However, at this time, we have not formulated a systematic pruning algorithm and are not certain whether such an algorithm is feasible.

REFERENCES AND NOTES

(1) Patterson, A. M. Proposed International Rules for Numbering Organic Ring Systems. *J. Am. Chem. Soc.* **1925**, *47*, 543–61.

(2) Henson, R. A.; Windlinx, K. J.; Wiswesser, W. J. Lowest Order Computer-Oriented "Ring Index" Diagrams—Verifying Correct Orientation of Fused Hexagonal Ring Systems. *Comput. Biomed. Res.* **1975**, *8*, 53–65.

(3) (a) Balasubramanian, K.; Kaufman, J. J.; Koski, W. S.; Balaban, A. T. Graph Theoretical Characterization and Computer Generation of Certain Carcinogenic Benzenoid Hydrocarbons and Identification of Bay Regions. *J. Comput. Chem.* **1980**, *1*, 149–157. (b) He, W.; He, W. Generation and Enumeration of Planar Polycyclic Aromatic Hydrocarbons. *Tetrahedron* **1986**, *42*, 5291–9. (c) Herndon, W. C.; Bruce, A. J. Perimeter Codes for Benzenoid Aromatic Hydrocarbons. *Stud. Phys. Theor. Chem. (Graph Theor. Topol. Chem.)* **1987**, *51*, 491–513. (d) Randic, M.; Nikolic, S.; Trinajstic, N. Compact Molecular Codes for Polycyclic Systems. *THEOCHEM* **1988**, *42*, 213–28. (e) Müller, W. R.; Szymznski, K.; Knop, J. V.; Nikolic, S.; Trinajstic, N. On Enumeration and Generation of Polyhex Hydrocarbons. *J. Comput. Chem.* **1990**, *11*, 223–35. (f) John, P. E. Two Codes for Hexagonal Systems. *MATCH-Commun. Math. Chem.* **1992**, *28*, 209–18.

(4) (a) Elk, S. B. A Nomenclature for Regular Tessellations and Its Application to Polycyclic Aromatic Hydrocarbons. *MATCH-Commun. Math. Chem.* **1980**, *8*, 121–58. (b) Elk, S. B. Formulation of a Canonical Nomenclature for Polybenzenes Using Triangular-Shaped Hexagonal Tessellation Envelopes. *MATCH-Commun. Math. Chem.* **1985**, *17*, 255–68. (c) Elk, S. B. Canonical Orderings of Hexagons That: (1) Nearly Meet the Intent of Patterson's Nomenclature Rules and (2) Orient a Molecule for Purposes of Assigning a IUPAC Name. *Polycyclic Aromat. Compounds* **1990**, *1*, 109–21. (d) Elk, S. B. A Canonical Ordering of Polybenzenes and Polymantanes Using a Prime Factorization Technique. *J. Math. Chem.* **1990**, *4*, 55–68.

(5) (a) Balaban, A. T.; Harary, F. Chemical Graphs 5. Enumeration and Proposed Nomenclature Cata-condensed Polycyclic Aromatic Hydrocarbons. *Tetrahedron* **1968** *24*, 2505–14. (b) Balaban, A. T. Chemical Graphs 7. Proposed Nomenclature of Branched Cata-condensed Benzenoid Polycyclic Hydrocarbons. *Tetrahedron* **1969**, *25*, 2949–56.

(6) Elk, S. B. A Simplified Algorithm Using Base 5 Numbers To Assign Canonical Names to Cata-Condensed Polybenzenes. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 637–40.