match no. 17 pp. 255-268 1985

FORMULATION OF A CANONICAL NOMENCLATURE FOR POLYBENZENES USING TRIANGULAR-SHAPED HEXAGONAL TESSELLATION ENVELOPES

Seymour B. Elk
Mathematics Consultant
321 Harris Place
New Milford, New Jersey 07646

(received: October 1984)

ABSTRACT: Rather than tessellating the entire plane with hexagons, as had previously been a basis for creating various nomenclature schemes for the class of polybenzenes, a grouping of minimum sized triangularly shaped envelopes covering only the relevant portion of the plane is formulated. In this manner, the set of coverings that must be examined for a given configuration is limited to, at most, twelve orientations of the given configuration. A simple ordering is then created for this set which yields a unique canonical name for any polybenzene. Similarly, starting from any canonical name, the model that produced this name can be formulated.

In Reference 1, an "almost intuitive" nomenclature scheme for clusters of benzene rings was introduced, but not fully developed. Instead, the emphasis was placed on the formulation of synthesis and analysis algorithms -- which are more closely related to the I.U.P.A.C. nomenclature[2]. Refinement of the techniques used in these synthesis and analysis algorithms are carried out in Reference 3 by turning away from geometrical considerations (tessellations) and relying on more analysis-based (graph theoretical) principles. Now, however, instead

of discarding the use of tessellations (as was the direction chosen in Reference 3), by a change in emphasis in the cluster nomenclature from considerations of symmetry to the superimposition of an appropriate covering set, an efficient, pragmatic, general nomenclature algorithm is presented that is readily adapted to the computer.

Adhering to the convention of drawing benzene rings with a set of vertical sides[4], any benzene cluster can be drawn in twelve or less different orientations (exactly twelve when no symmetry axis or point exists). Consequently, an orientation-based nomenclature algorithm will require the formulation of, at most, twelve preliminary names, followed by the selection of a canonical name from this set. For example, consider the pentabenzene shown in Figure 1 (designated by fbb-pentabenzene in the analytic nomenclature of Reference 1). In order to form all twelve orientations, begin with the configuration shown and rotate this figure 60^{o}, then a second 60^{o}, etc. until making a complete circle. This will produce six different orientations. The second six are formed by reflecting each of the orientations formed through a vertical line. Figure 2 illustrates the twelve orientations of the compound represented in Figure 1.

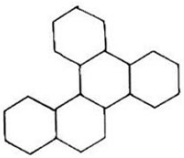


FIGURE 1

Next, for each of these orientations, draw a horizontal "base" line through the "bottom" hexagon(s). Using this base, draw in the smallest "triangle" of hexagons that completely covers the given figure. Such "triangular-shaped hexagonal tessellation envelopes" are shown in Figure 3. Note that in three pairs (i.e., six) of these tessellation envelopes, a triangle of side length 4 is sufficient; namely, in a, c and e. Label these using n-tuplets formed by including a "1" for the presence of a ring in the given orientation and a "0" otherwise. This yields a set of ordered binary prefixes:

FIGURE 2

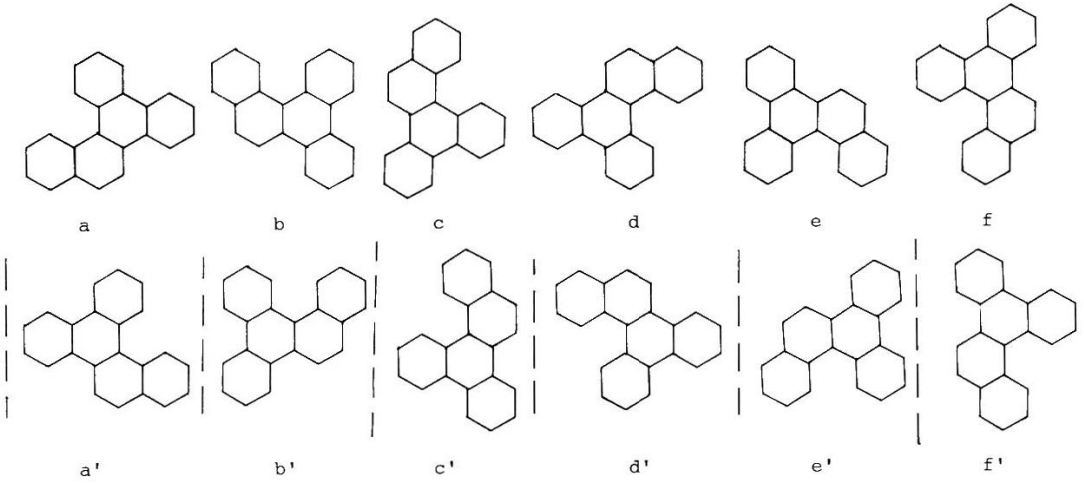


FIGURE 3

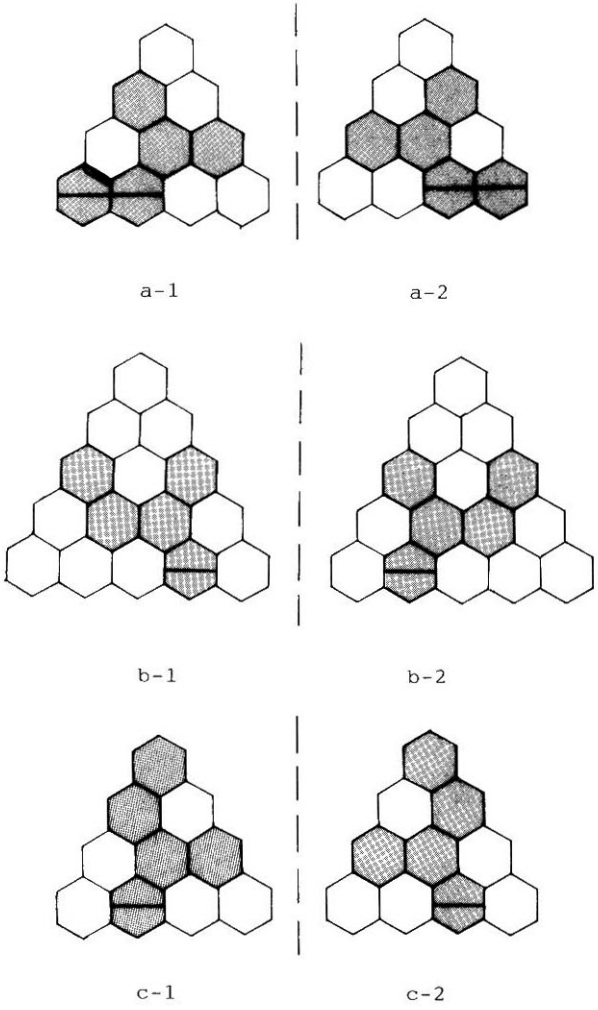
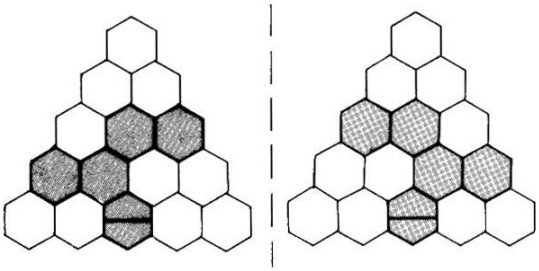
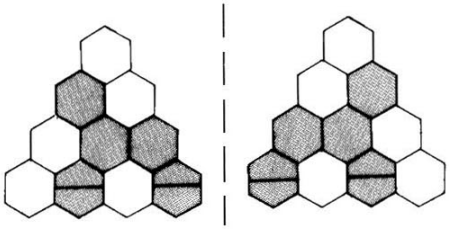


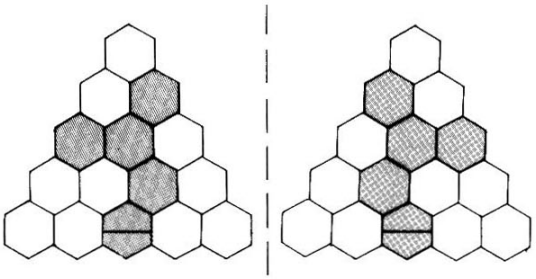
FIGURE 3 (continued)



d-1 d-2



e-1 e-2



f-1 f-2

a-1 = 0,10,011,1100 a-2 = 0,01,110,0011
c-1 = 1,10,011,0100 c-2 = 1,01,110,0010
e-1 = 0,10,011,0101 e-2 = 0,01,110,1010

which, for convenience, may be expressed as decimal integers:

a-1 = 0,2,3,12 a-2 = 0,1,6,3
c-1 = 1,2,3,4 c-2 = 1,1,6,2
e-1 = 0,2,3,5 e-2 = 0,1,6,10

Finally, selecting the smallest prefix (a-2 for this example) yields the suggested name: 0163-pentabenzocluster -- where the commas may be deleted when all of the digits are less than 10.

The general algorithm is presented in Table 1, and its application to naming all of the mathematically possible tri- and tetra-benzenes is given in Tables 2 and 3. Also, application to larger compounds (including their corresponding I.U.P.A.C. nomenclature[2]) is given in Table 4; thereby illustrating the simplicity of the proposed name vs. the corresponding presently-accepted name.

In a similar manner, given any appropriate sequence of numbers, the structure can be generated (provided that the name yields a connected set of rings). For example, consider a random sequence, such as 7,4,8. First, express this in binary form: 111,100,1000. Next add the appropriate number of zero rows and leading zeroes: 0,00,111,0100,01000. This is the representation shown in Figure 4, where the number of rings is determined by counting the number of binary ones; i.e., a pentabenzocluster.

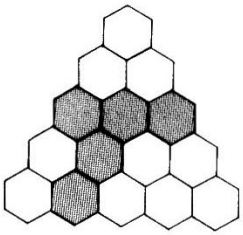


FIGURE 4

An important observation to be made at this point is that this structure was drawn without regard to whether this is, in fact, the canonical name, or whether one of the other eleven orientations should have been used. For this particular compound, a rotation of -120°, followed by a vertical re-

flection yields the canonical name 0,0,1,14,4-pentabenzocluster (Figure 5). Note that the zeroes may be omitted: 1,14,4-pentabenzocluster.

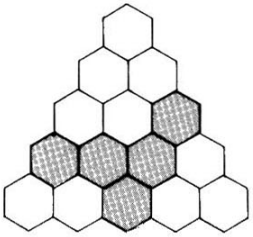


FIGURE 5

A simplification in usage that may be employed on an interim basis; for example , for research, rather than in a final report, is to use the orientation exactly as drawn and to then draw in the base line wherever appropriate with an arrowhead to indicate whether reflected (clockwise) or not. This is illustrated in Figure 6, and is comparable to starting with a standardized curve, such as a sine wave, and then drawing in the axes and scales later (Figure 7), rather than the more traditional method of replotting the curve.

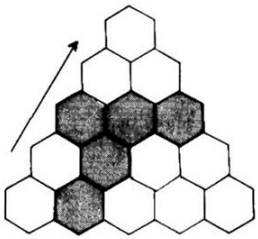


FIGURE 6

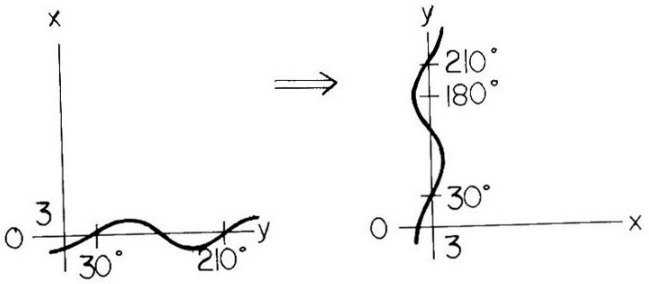


FIGURE 7

A second type of simplification, of a more fundamental nature, occurs whenever the given sequence is all even; namely, there exists a smaller triangular-shaped hexagonal tessellation envelope than the one selected. For example, consider 6,4,8. Converting to binary and adding the implied zeroes yields 0,00,110,0100,01000 -- which is illustrated in Figure 8. Notice that the right hand boundary of hexagons in the triangular-shaped tessellation envelope is all empty; and should be omitted. Now the resulting triangular-shaped hexagonal tessellation envelope has a side length of 4, and the sequence prefix: 0,11,010,0100, which, in decimal form, is 0324 -- exactly one-half the starting sequence.

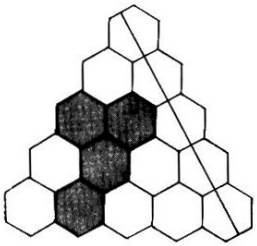


FIGURE 8

Mathematically, it is not surprising that, in the above process, there is no significance to leading zeroes or zero apical rows. This is a consequence that, for example:

3, 03, 003, ..., 00000003, etc.

all represent the same quantity: 3. For this reason, all such zeroes in the name may be omitted without ambiguity. In the regeneration of the picture from the name, therefore, it should be kept in mind that the n-th row must be represented by a number less than 2^n. This is equivalent to saying that the bottom row in the name for isoviolanthrene must contain at least 7 hexagons (See Table 4). Also, at least four rows of all zeroes will be required.

TABLE 1

ALGORITHM

1. Draw each possible orientation (maximum of 12).
2. For each drawing, using the base line, enclose the given figure in the smallest triangular-shaped hexagonal tessellation envelope.
3. If there exist different size smallest triangular-shaped hexagonal tessellation envelopes for the different orientations, retain only those of minimum size.
4. In each of the remaining envelopes, label each hexagon by assigning either a "1" or a "0", depending if that hexagon is a part of the given figure or not.
5. Starting from the apex of the triangular-shaped hexagonal tessellation envelope, form an n-tuple for each orientation. This n-tuple may be expressed in either binary, or decimal, form.
6. Choose the numerically lowest n-tuple as the prefix of the canonical name. The remainder of the name is the appropriate Greek prefix for the number of rings and the word "benzocluster".

TABLE 2

TRIBENZENES

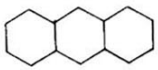
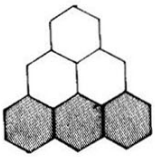
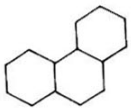
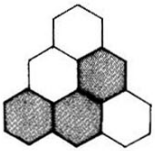
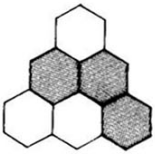
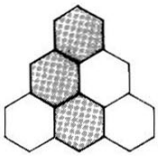
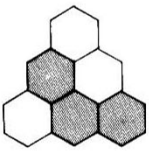
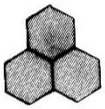
Structure	Anal. Name	Tessellation Envelope	Name Prefix Binary	Dec.
	a-		0,00,111	007
	b-		0,01,110	016
			1,01,010	112
			0,11,001	031
			0,11,100	034
			1,10,010	122

TABLE 2 (continued)

	b-		0,10,011	023
	c-		1,11	13

Notes: (1) Because the simplest orientation gives the canonical name for both a- and c-tribenzene, only one tessellation envelope has to be examined.

(2) Because of symmetry considerations, only six of the twelve b-tribenzenes have to be examined.

TABLE 3

TETRABENZENES

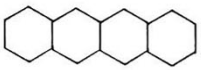
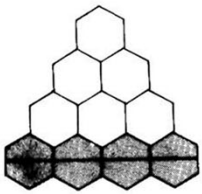
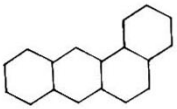
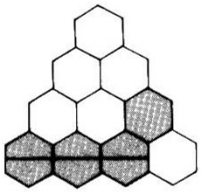
Structure	Anal. Name	Tessellation Envelope	Name Prefix Binary	Dec.
	aa-		0, 00, 000, 1111	0,0, 0,15
	ba-		0, 00, 001, 1110	0,0, 1,14

TABLE 3 (continued)

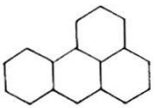
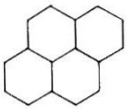
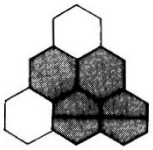
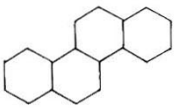
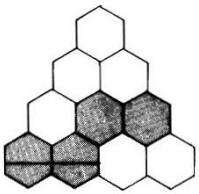
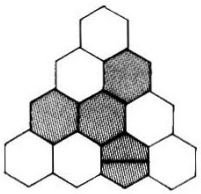
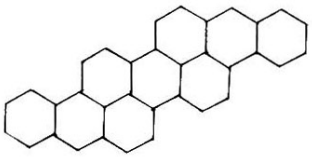
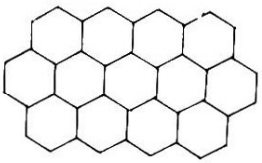
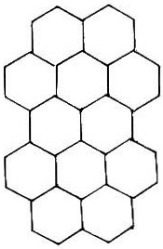
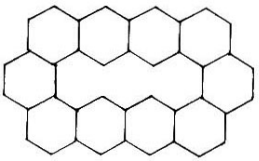
	ca-		0, 01, 111	017
	db-		0, 11, 101	035
	eb-		0, 11, 011	033
	bb-		0, 00, 011, 1100	0,0, 3,12
	fb-		0, 01, 110, 0010	0162

TABLE 4

REPRESENTATIVE COMPOUNDS

Structure	Common Name	IUPAC Name	Cluster Name
	(Dehydro-) Isoviolanthrene	(Dehydro-) Dinaphtho (1,2,3-cd: 1',2',3'-lm) perylene	7,28,112-Nonabenzocluster
	Circumanthrene	Phenanthro (3,4,5,6-bcdef)ovalene	15,30,31-Trideca-benzocluster
	Circobiphenyl	Naphth(2',1',8',7':4,10,5)-anthra-(1,9,8-abcd)-coronene	3,15,30,24-Dodecabenzocluster
	---	15,21:16,20 Dimetheno-naphtho-(1-2a)-phenanthro-(3,4-l)-naphthacene	15,17,30-Decabenzocluster

REFERENCES

1. Elk, Seymour B., "A Nomenclature for Regular Tessellations and its Application to Polycyclic Aromatic Hydrocarbons", MATCH - Communications in Mathematical Chemistry, 8, 121-158, 1980.

2. "Nomenclature of Organic Chemistry: Section A", IUPAC, Pergamon Press, Oxford, England, 1979.

3. Elk, Seymour B., "Refinement of Systematic Nomenclature for Polybenzenes and its Extension to Systems of General Arenes", MATCH - Communications in Mathematical Chemistry, 13, 239-269, 1982.

4. Patterson, Austin M., "Proposed International Rules for Numbering Organic Ring Systems", J.Am.Chem.Soc., 47, 543-561. 1925.