A Nonunique Path Connectivity Matrix

Donald J. Polton*

Department of Computer Science, University of Hull, Hull HU6 7RX, England

Received March 5, 1992

A connectivity matrix has been developed that is based on structural components that are not single atoms but linear fragments (paths). The path connectivity matrix (PCM) preserves the structure as specified in the input source and is, therefore, not unique since the same structure may be described by different nomenclature systems in different manners. It is readily obtained by computer processing, and preferred names can be computed from the PCM. It has been used in the study of nomenclature systems in general[1] and is of particular benefit as an aid to generation and validation of names.

INTRODUCTION

The Concise Connection Table (CCT) of Rayner[2] gives a full description of a chemical structure in a minimum space by preserving whole chains and rings rather than breaking them down into individual atoms, as is usual in a connection table. The CCT describes chains and rings one at a time in a four-column table which contains information on their sizes, connectivity points, and the nature of atoms and bonds, etc. It is derived from IUPAC nomenclature,[3] and each structure has a unique CCT. In polycyclic structures, it preserves the orientation necessary for deriving the correct IUPAC name.

For the study of different representations of chemical structures by computer, a connectivity matrix has been developed which may take different forms for a particular structure according to the input source used. By strict control of the computer algorithm used for its derivation, it may, however, be standardized for each individual method of structural representation. This matrix, like the CCT, is also based on chains, or paths, of atoms both in cyclic and acyclic structures. It is referred to here as the path connectivity matrix (PCM). It is readily obtained by computer processing and preserves the structure as described in the input source.

The PCM was created for the study of nomenclature systems which have, over the years, been put forward as possible alternatives to conventional IUPAC nomenclature. For such a purpose, it is preferable to keep the structural fragments as far as possible in string form rather than to break them down into individual atoms and then reunite them. The work involves creating a matrix from the name and a name from the matrix and, thus, covers generation, validation, and interconversion of names. A full account of the procedures used and the results are reported elsewhere.[1] This account is intended to cover the idea of the PCM only, there being no other connectivity matrix using strings rather than single atoms, apart from the CCT of Rayner,[2] reported in the general literature. It must also be emphasized that connectivity matrices are not confined to the chemical world but have important applications, for example, in geography.

The structure of the PCM for acyclic and cyclic structural components is considered separately. This is followed by a description of the manner in which heteroatoms, unsaturation, and aromaticity are recorded.

GENERAL CONSTRUCTION OF THE PCM

Although referred to as a matrix, the entire record of the stored structure consists of a number of sections. The main section, the matrix, shows how fragments of the structure are joined together, by which nodes in the case of acyclic structures, and how they are cyclized and bridged in the case of ring structures. The lengths of the fragments are defined prior to the matrix section. The size of the matrix, i.e., the number of fragments of which it is composed, precedes the chain length section, and the whole begins with information on the source used to produce it. Nodal variations, or heterocyclic atoms, and edge variations, or information on bond types, follow as a suffix to the main matrix section. For multicomponent systems, the structure may be broken down into separate PCM sections with data on interconnections between the components. To summarize, the general construction of a single component PCM is as follows:

```
                    Source type
                  source name/code
components...
                  number of fragments
                   fragment lengths
                  fragment connections
node variations...
          identifier and count of node variations
               node type (heterocyclic atom)
                node position in the matrix
edge variations...
          identifier and count of edge variations
                     variation type
                    location in the matrix
```

For the purposes of display, these data need not be given in full, and strict order is not maintained. In the examples which follow, the sources and fragment counters are omitted. Fragment lengths are not displayed before the matrix, but each length is shown alongside the row in which the information on that fragment occurs, with the indicator 'L'. Suffix information follows the matrix in abbreviated form, and the counters are not displayed.

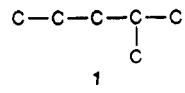
PCM FOR ACYCLIC STRUCTURES

The matrix section of the PCM for acyclic structures consists of the tabulated link positions of the n chains of which the structure is composed, where the maximum value of n permitted is the number of terminal positions minus 1.

In **1** there are three terminal positions, and the number of fragments in the PCM is therefore 2.

* Address correspondence to this address: 'Rosefarm', Denne Manor Lane, Shottenden, Canterbury, Kent CT4 8JJ, England.

0095-2338/93/1633-0095$04.00/0



1

The two chains of **1** may be taken as those of lengths 5 and 1, or those of lengths 3 and 3. The correct name in IUPAC nomenclature is 2-methylpentane, i.e., a chain of length 5 joined at its 2-position to a chain of length 1. The PCM for this arrangement is

	chain 1	chain 2	length
chain 1	0	2	L.5
chain 2	1	0	L.1

with the lengths displayed as indicated above, alongside the chains containing the connectivity information, and prefixed by 'L.'. (Note: Only in this first example of a PCM are the row and column headers shown.)

If the structure is incorrectly named 4-methylpentane it gives rise to the PCM

0 4	L.5
1 0	L.1

which, however, is a valid PCM for the structure. This PCM when reconverted to an IUPAC name will show the original input name to have been wrong.

The structure could also be incorrectly named as 2-propylpropane, giving the PCM

0 2	L.3
1 0	L.3

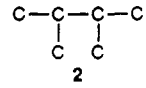
which is also a valid PCM and is convertible to the correct name.

This particular structure named by other methods of nomenclature gives rise to the same PCM as is given by the IUPAC name. Some of these are

Lozac'h nodal[4-6]	$[5.1^2]$hexane
new Dyson[7]	pentan,monan-2
HIRN system[8]	hexa$[5.^2 1]$carbane

In the PCM for acyclic structures, the *n,n* positions in the matrix are always 0. A nonzero here implies that the chain is linked to itself, i.e., there is cyclization.

It was stated above that the maximum value of the number of chains in the PCM is the number of terminal positions minus 1. This maximum count is arrived at by taking the first chain as a through chain, terminal to terminal position, and adding all subsequent chains to intermediate positions of this main chain as the acyclic system is built up. It may be equally feasible to take the second and maybe further chains also as through chains, provided there is a direct connection between the new chain and an earlier named chain. This is illustrated by **2**, which in IUPAC nomenclature is named 2,3-dimethylbutane, a three-fragment system of lengths 4, 1 and 1:



2

The PCM derived from this IUPAC name is

0 2 3	L.4
1 0 0	L.1
1 0 0	L.1

The structure could also be looked upon as consisting of two chains of length 3, joined at their 2-positions. This will lead to the valid PCM

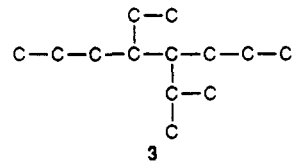
0 2	L.3
2 0	L.3

which would be obtained from the incorrect IUPAC name 2-isopropylpropane.

Those nomenclature systems which depend on the longest chain will all, of course, give rise to the three-membered PCM. Some of these are

Lozac'h	$[4.1^2 1^3]$hexane
new Dyson	tetran,monan-2,3
HIRN system	$[4.^2 1^3 1]$carbane

As a final example, **3** shows a more complex branched acycle and the PCMs derived from it:



3

The full IUPAC name for this, 4-ethyl-5-(1-methylethyl)-octane, gives the PCM

0 4 5 0	L.8
1 0 0 0	L.2
1 0 0 1	L.2
0 0 1 0	L.1

The methyl branch is on the second ethyl chain because the software analyzes unparenthesized groups first. The ethyl group occurs first in the name because of its alphabetical seniority over 'methylethyl'.

The new Dyson name for **3** is octan,(2-trian)4,dian-5, and the PCM from this is

0 5 4	L.8
1 0 0	L.2
2 0 0	L.3

in which there are only three components. The eight-membered chain has two substituents, one of which is a three-membered chain linked by its 2-position.

In the nodal nomenclature of Lozac'h, **3** is $[8.2^4 2^5 1^9]$-tridecane, and the corresponding PCM is

0 4 5 0	L.8
1 0 0 1	L.2
1 0 0 0	L.2
0 1 0 0	L.1

Here the branched chain is described first, so that the PCM differs from that produced by the IUPAC name in the location of the methyl group.

All of the above PCMs are valid for this structure. Depending on the software used to produce them, further forms of the PCM might be produced. This illustrates the necessity for the PCM to be nonunique. To derive a unique PCM would require lengthy computer procedures, and it would be easier to produce an atom connection table. A unique PCM would destroy its own use as a connection table which maintains all the characteristics of the input used to produce it.

Before going on to describe the PCM for cyclic structures, it should be noted that the matrix section of the PCM for an unbranched chain consists of '0' only. It may be looked upon

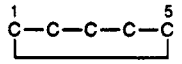
as a one-component matrix. For example, the PCM for *n*-pentane is

0 L.5

PCM FOR CYCLIC STRUCTURES

The path connectivity matrix for a cyclic system is based on one selected ring of the structure. A monocycle is a one-component system; but for polycycles, all additional rings are considered as bridges, first of all across the initial ring and then across nodes in the polycycle as it is being constructed.

Initial Cyclization. The initial cyclization is shown by an entry in the [1,1]-position of the matrix. As already mentioned, this implies that the chain is joined to itself. To depict cyclization of the first chain, it is necessary to show that its ends are joined together. In the case of a five-membered chain, the 1- and the 5-position must be seen in the PCM to be linked:

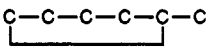


It is necessary, therefore, to add another dimension to the array to allow for both link positions to be quoted:

1^5 L.5

In this particular case of a monocycle, the *x* and *y* dimensions are of 1 unit only, and the two locants necessary to show cyclization are 1 in the [1,1,1]- and 5 in the [1,1,2]-positions.

It is also possible by this means to show partial cyclization of a chain. A six-membered chain joined across its 1- and 5-positions only is a five-membered ring with a substituent chain of length 1 and might be depicted in the same way:



1^5 L.6

The use of an additional dimension to show the two positions of a chain is avoided by placing a limitation on chain length in the system as a whole. Taking this limitation as 99, the two locants can be united by combining them in the form

100 × (locant 1) + (locant 2)

This gives a PCM for the cyclized five-membered chain as

105 L.5

and for the partial cyclized six-membered chain as

105 L.6

If it were considered necessary to include chains of length greater than 99, the multiplication factor could be increased to 1000. Alternatively, real numbers may be used, the cyclization in the above examples being shown as 1.05 and 1.06, or if the maximum allowed chain length is 999, 1.005 and 1.006.

Partial cyclization is not used in the PCM. It is only permissible to show full cyclization of the initial ring, otherwise the PCM ceases to be an instrument for the designation of purely acyclic or cyclic structures. It has been found advantageous to abbreviate full cyclization by using only a single digit in the [1,1]-position. If full cyclization only is allowed, it is only necessary to enter a nonzero in the [1,1]-position to distinguish a cyclic PCM from an acyclic PCM. The number 1 alone is used, and the PCM for a cyclized five-membered chain is

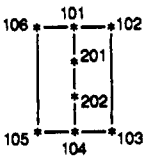
1 L.5

The advantage of this usage will be seen later when the PCM for aromatic rings is described.

Bicyclic Systems. Bicyclic structures are considered as bridged monocycles, that is, only the initial ring of the PCM is shown as a cyclized chain. The additional ring is shown as a bridging chain without indication of cyclization. Bridges have two locants, which are entered into the PCM in the format just described. With the limit of 99 on the chain length these are entered as

100 × (bridge position 1) + (bridge position 2)

The two bridge nodes in the main chain are shown in the same way. Taking as an example **4**, which is a six-membered chain cyclized and bridged by a two-membered chain, the PCM is as shown

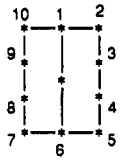


4

1 104 L.6
102 0 L.2

In this PCM, the initial chain is shown to be cyclized by the presence of 1 in the [1,1]-position and is bridged from its 1- to its 4-position (nodes 101 and 104 in the diagram) by a two-membered chain, this being shown by the entry 104 in the chain 1 row. The bridge connections of chain 2 are shown as node 102 in the chain 2 (length 2) row. The individual chain numbers and node locants are easily retrieved from these PCM entries by taking the value in the PCM entry DIV 100 and MOD 100.

In the case just considered, only the one PCM is possible. The largest ring is six-membered, and there is no ring of smaller size. Structure **5** shows a 10-membered ring bridged by a one-membered chain.



5

The PCM for this as a bridged ten-membered ring is

1 106 L.10
101 0 L. 1

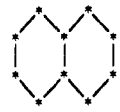
There is another way of describing this structure, as reflected in different forms of nomenclature. It may be considered as two seven-membered rings fused across three common nodes. In this case, one of the rings must be taken as the initial ring, and since there are no other constraints, it does not matter which. The remaining four nodes of the other ring are then taken as a bridge, across the initial ring or cyclized chain, and the PCM becomes

1 103 L.7
104 0 L.4

Both of these are valid PCMs and, in fact, are arrived at from different input sources. The first corresponds to the

nodal system of Lozac'h, and the second corresponds to the Dyson system.

Bridges without Nodes. Many, in fact most, common bicyclic structures are depicted as having two rings fused across one bond only. A simple case is naphthalene, the hydrogenated form known as decalin being shown as **6**.



6

This may be considered as two six-membered rings fused by a common bond or as a 10-membered ring bridged by an empty chain. As before, different naming systems do consider the structure in these ways. The first, which is the method of the new Dyson nomenclature, gives a PCM similar to the above examples by taking the additional nodes of the second ring as a bridge across the 1- and 2-positions of the initial cyclized six-membered chain

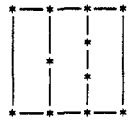
1	102	L.6
104	0	L.4

In the second consideration, all 10 nodes of the system are contained in the initial cyclized chain, and the bridge is of length 0. The matrix still consists of two components, and the bridge is entered as of length 0, i.e., it is a 'null chain'. The PCM, derived from the Lozac'h name bicyclo[010.0^{1,6}]decane, is

1	106	L.10
0	0	L. 0

It might be suggested that this 'null chain' may be omitted. For bicycles this may be true, but with polycycles it may not necessarily be the last chain quoted, and its ring connections will be in the corresponding column. For this reason the null chain is always included in the matrix, even in simple cases such as **6**.

Polycycles. In a polycycle, the bridges may link nodes which are part of the same PCM chain. If so, the PCM will be similar to the examples for bicycles given above. **7** is a tricycle composed of an initial cyclized chain across which there are two bridges.



7

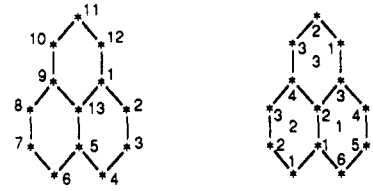
Taking this as a doubly bridged eight-membered ring, the PCM from the Lozac'h name tricyclo[08.2^{1,4}1^{5,8}]undecane, in which longer bridges are quoted first, is

1	104	508	L.8
102	0	0	L.2
101	0	0	L.1

However, based on the smallest set of smallest rings, it is a fused system of three rings of sizes 7, 6, and 5. The initial ring in the new Dyson system is the largest of these, and the other two are represented by two-node bridges so that the PCM is

1	104	507	L.7
102	0	0	L.2
102	0	0	L.2

The case sometimes arises where a bridge links two nodes of a partially built polycycle which belong to different chains of the PCM. It is then necessary to quote the two connecting nodes of the bridged chains separately, against the two chains to which they are connected, and the $100x + y$ system does not apply. In **8**, which may be considered as before in more than one way, the second bridge links the initial cyclized chain and the first bridge:



8

This structure is considered in the Lozac'h system as based on the outer 12-membered ring numbered as shown. This is bridged across its 1- and 5-positions by a chain of length 1, which becomes node 13. A third bridge, a null chain links this new node 13 with node 9 of cyclized chain number 1, and the PCM is

1	105	9	L.12
101	0	1	L. 1
0	0	0	L. 0

The same structure is considered in the Dyson system as consisting of three six-membered rings, numbered in the diagram in the center of the rings. These appear in the PCM as an initial cyclized six-membered chain bridged across its 1 and 2 nodes by a four-membered chain, with the third ring as a three-membered bridge from the 3-position of cyclized chain 1 to the 4-position of chain 2. The enumeration of each chain is shown in the diagram. The PCM for this is

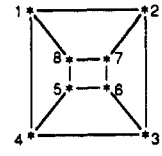
1	102	3	L.6
104	0	4	L.4
1	3	0	L.3

Another enumeration, dependent on the initial ring being numbered the other way round, gives the bridge locants in column 3 the values 6 and 7. This is a valid PCM and would arise from incorrect input.

By way of example, some more complex polycycles are now given with their PCMs.

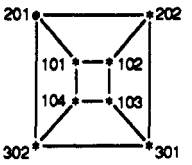
9 is cubane and is shown with two enumerations. It may be taken as an eight-membered ring with four bridges shown as null chains or as five four-membered rings, the first of which is a cyclized chain, the next two represented by two-membered bridges, and the last two by null chains. In this case, the two null chains connect different chains of the PCM, and the connecting bridge entries are single digit only.

The first enumeration is obtained from Lozac'h input; the second is obtained by Dyson input.



9 with Lozac'h enumeration

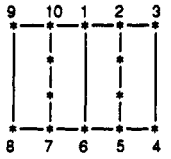
1	104	207	306	508	L.8
0	0	0	0	0	L.0
0	0	0	0	0	L.0
0	0	0	0	0	L.0
0	0	0	0	0	L.0



9 with Dyson enumeration

1	102	304	0	0	L.4
102	0	0	1	2	L.2
102	0	0	2	1	L.2
0	0	0	0	0	L.0
0	0	0	0	0	L.0

10 is an example showing the nonterminal position of a null chain in the PCM derived from the author's development of nodal nomenclature.[9]



10

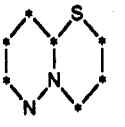
1	106	205	710	L.10
0	0	0	0	L. 0
102	0	0	0	L. 2
102	0	0	0	L. 2

Here the shorter bridges are quoted first, beginning with the null chain across the 1- and 6-positions of the main ring.

NODAL VARIATIONS: HETEROCYCLES

Variation in the character of a node is shown as a suffix to the PCM. If there are heteroatoms present, then a title character, Z, is followed by the total number of nodal variations, i.e., the number of heteroatoms. There then follows a description of the nature of each node by its chemical symbol and location. The node location is expressed in a similar way to bridge location in the matrix, that is, as $100x + y$ where x is the chain number, and y is the number of the node in that chain. Thus, a nitrogen heterocycle in the 3-position of chain 1 is quoted as N103. The heterocyclic symbols are not arranged in any fixed order, their sequence depends entirely on the input source. Some naming systems put elements in symbol order; others put them in some form of alphabetical order. For example, a nitrogen heterocycle might be named by the use of the prefix aza- and when it would appear before boron with the suffix bora-, although the chemical symbol order is the reverse: B, N. Fortunately, the most commonly occurring heterocycles are nitrogen, oxygen, and sulfur, and the prefix order aza-oxa-thia coincides with the symbol order N, O, S. Other systems base element order on the periodic table, which may give a completely different arrangement. The Lozac'h system arranges them in decreasing group number and increasing atomic weight, so that the order is O-S-N, oxa-thia-aza, O and S being in group VI and N in group V of the periodic table.

As an example, **1** is given with its PCM and hetero suffix. The PCMs for the basic ring system of this structure from Lozac'h and Dyson names shown earlier are repeated here with the heterocyclic suffix, the node variation identifier being 'Z'



11

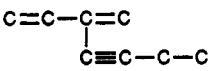
Lozac'h	Dyson
1 106 L.10	1 102 L.6
0 0 L. 0	104 0 L.4
Z3 S102 N106 N107	Z3 N101 N106 S204

In further examples, for the purposes of display, the heterocyclic indicator Z and the count will be omitted.

EDGE VARIATIONS: UNSATURATION AND OTHER BOND TYPES

If there are any double or triple bonds in a structure, these are shown as a suffix which follows any nodal variations already described.

The count of edge variations is indicated after the identifier 'U'. The character of the variation is indicated by a symbol, which is followed by the location of the unsaturated bond, both positions being quoted. For double bonds the symbol 'be' is used; for triple bonds the symbol 'by' is used. **12** is an acyclic structure containing double and triple bonds, with its PCM based on the IUPAC name.



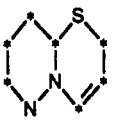
12

0	3	L.7
1	0	L.1

U3 be101102 be103201 by104105

For further display purposes, the title symbol 'U' and the count will be omitted.

An example with both nodal and edge variations is **13**.



13

Lozac'h

1	106	L.10
0	0	L. 0
S 102	N106	N107

be104105

It is of course not necessary to show the second locant of unsaturation if this is sequential to the first, other than for the sake of clarity.

AROMATIC RINGS AND PART AROMATIC RING SYSTEMS

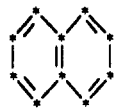
The entry in the [1,1]-position of the matrix shows whether the structure is acyclic or cyclic, by '0' or '1'. Any nonzero in the [1,1]-position could be used to describe a cyclic PCM. Double bonds in a ring system are shown by a suffix, and structures so described are nonaromatic.

The name of a ring structure shows whether or not it is aromatic, perhaps with indicated hydrogen. Aromatic ring codes in the Dyson systems begin with 'B' or use 'phen'. Lozac'h

names may end in 'arene'. IUPAC names are composed of defined trivial components. If a structure described as aromatic contains atoms which are not doubly bonded, these are shown as H substituents.

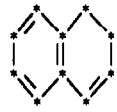
The method for dealing with aromatic structures in the PCM is to indicate this by putting '2' in the [1,1]-position. Because this is nonzero, a cyclic structure is still indicated. There will be no double bond suffix of course, but there may be a hydrogen locant suffix. This '2' is used as a classification symbol.

As an example, the PCMs for **14** (naphthalene) and **15** (1,2-dihydronaphthalene), entered as Dyson names, are shown



14

```
 2 102      L.6
104   0      L.4
```

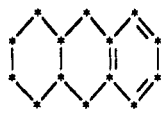


15

```
 2 102      L.6
104   0      L.4
H103 H104
```

A structure may be largely nonaromatic but contain one (or more) aromatic rings. If the aromatic ring bonds are described as double bonds in the name, these will be put in the suffix as they are numbered. In the case of Dyson names, a benzene ring component in a polycyclic structure is shown by 'b' after the ring fusion locant(s). In this case, all six atoms are stored as aromatic nodes, under the header 'ba' with their positions. For conversion of the PCM back into a name or code, the presence of all six atoms of a ring in the aromatic or double bond list shows an aromatic fused ring.

The hydrogenated anthracene, **16**, is coded using the locant series 1b4, the 'b' showing the location of the fused benzene ring to be across the 1-2-positions of the initial ring.

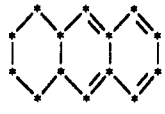


16

The PCM for this will be

```
  1 102 405      L.6
104   0   0      L.4
104   0   0      L.4
ba201 202 203 204 101 102
```

In the hydrogenated anthracene **17**, the unsaturation is shown as double bonds and not aromatization.



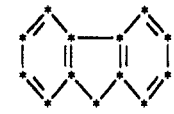
17

The PCM for this will be

```
  1 102 405      L.6
104   0   0      L.4
104   0   0      L.4
be101106 102103 201202 203204
```

The double bonds are represented as a 'be' list and not as 'ba' nodes because they are coded using the double bond symbol 'E' and not the aromatization symbol 'b'. Examination of the double bond node positions will show that one ring has all six nodes doubly bonded, nodes 201–204 and 101–102, but as this leaves two isolated doubly bonded nodes 103 and 106, this ring is not given the aromatic 'b' code.

Fully aromatic structures may of course have indicated hydrogen. **18** is an example, showing the hydrogen as a suffix to the PCM.



18

```
 2 106 813      L.13
 0   0   0      L. 0
 0   0   0      L. 0
H 107
```

MULTICOMPONENT STRUCTURES

Although work on the PCM has been restricted to acyclic and cyclic structures only, some thought has been put into how the PCM might be extended to cover multicomponent structures. There are two ways in which this could be done. Either the components may be mixed in the same PCM, or they may be arranged as a series of PCMs. With the former, it becomes necessary to have a procedure to break down the PCM when it is read, and each chain will have to be labeled with its component number. With the latter, the components are kept apart with component connectivity data between them. The latter separate component method is considered the better of the two.

Multicomponent structures consisting of cyclic components only could be expressed as a single unit PCM. For example biphenyl Ph–Ph may be

```
 2 1      L.6
 1 2      L.6
```

Here the second ring as well as the first is shown to be cyclized by the presence of the nonzero in the [2,2]-position, and the two rings are linked by their 1-positions.

Using the separate component method the PCM for this would consist of two sections with suffix information to show the connections:

```
1  (unit in component 1)
     2    L.2
c1 101 (connection 1 on 101)
1  (unit in component 2)
     2    L.6
c1 101 (connection 1 on 101)
```

ACKNOWLEDGMENT

I acknowledge the contribution of Dr. G. H. Kirby of the Department of Computer Science, Hull University, under

whose supervision the studies, of which this paper forms a part, were carried out.

REFERENCES AND NOTES

(1) This work is part of a postgraduate research study carried out at Hull University (Polton, Ph.D., 1991), which is to be published in revised form: Polton, D. J. *Chemical Structures and the Computer*; Research Studies Press: Taunton, Somerset, England, in press.
(2) Rayner, J. D. A Concise Connection Table Based on Systematic Nomenclatural Terms. *J. Chem. Inf. Comput. Sci.* **1985**, *25*, 108–111.
(3) International Union of Pure and Applied Chemistry. *Nomenclature of Organic Chemistry, Sections A–F and H*; Pergamon: Oxford, 1979.
(4) Lozac'h, N.; Goodson, A. L.; Powell, W. H. Nodal Nomenclature—General Principles *Angew. Chem., Int. Ed. Engl.* **1979**, *18*, 887–899.
(5) Lozac'h, N.; Goodson, A. L. Nodal nomenclature II—Specific Nomenclature for Parent Hydrides, Free Radicals, Ions and Substituents. *Angew. Chem., Int. Ed. Engl.* **1984**, *23*, 33–46.
(6) Lozac'h, N. Principles for the Continuing Development of Organic Nomenclature. *J. Chem. Inf. Comput. Sci.* **1985**, *25*, 180–185.
(7) Dyson, G. M. Some New Concepts in Organic Chemical Nomenclature. Unpublished.
(8) Hirayama, K. *The HIRN System, Nomenclature of Organic Chemistry*; Maruzen: Tokyo; Springer-Verlag: Berlin, 1984.
(9) Polton, D. J. A New Method of Nodal Numbering for Acyclic and Cyclic Structures. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 430–436.