Computer Generation of Baeyer System Names of Saturated Bridged Bicyclic, Tricyclic, and Tetracyclic Hydrocarbons

KENNETH CONROW

Department of Chemistry, Kansas State University, Manhattan, Kansas 66502

A computer program has been written in the Fortran language which accepts structural information for bridged bicyclic, tricyclic, and tetracyclic saturated hydrocarbons and develops their Baeyer system name. Symmetry elements are recognized and identified by the program, and the atoms which serve as the bridgehead atoms in the name are identified in terms of their numbers in the input numbering so that the resulting name can easily be fit to the structure. A list of 29 examples of incorrect Baeyer nomenclature from CA indexes and the Ring Index illustrates the necessity for real care, or the use of a computer, in assigning Baeyer names to these systems. Although extremely limited in its scope, this work represents a first step toward generation of a complete systematic name for organic compounds by computer methods.

We have been engaged in the hydrolytic rearrangement of 2,2',4,4',6,6'-hexamethyl-4,4'-bi-4H-pyran (1) and have isolated some 10 different crystalline products from this reaction under varying conditions of acidity and water concentration. In view of the large number and the complexity of the products of this hydrolytic rearrangement, it was deemed inefficient to attempt a classical structure proof of each individual product. Instead, we chose to prepare a list of probable products of hydrolytic rearrangement with a view to choosing from that list those structures which would best accomodate the observed properties of the isolated products.

Even though a number of limits were set by excluding certain mechanisms and condensations of unlikely occurrence, the list of possible products has grown to well in excess of 1500 structures. These structures arise from approximately 500 different sequences of carbon-carbon bond forming condensation reactions and (as it proved) represent 88 different carbocyclic ring systems. As every chemist who has worked with bridged polycyclic systems knows, detailed comparison of two structures to verify their identity is a time-consuming and unreliable process because of the difficulty of reorienting the structures for detailed comparison. With some 1500 structures under consideration, it was necessary to automate the comparison process in some way so that different sequences of condensations leading to the same product could be identified and the duplicate structures eliminated from the list.

One foolproof approach appeared to be to assign a systematic name to each structure in the list and then to compare the systematic names. Since each structure will generate a unique systematic name, all duplicates should be quickly and unambiguously identified. A problem still remained. As has been pointed out (2), the Baeyer system

nomenclature tends to give ambiguous results for tricyclic and more highly condensed bridged systems unless great care is taken that proper choice of the bridgehead atoms is made. Casual inspection of a structure is as likely to give an incorrect as a correct name, but systematic inspection of all possible combinations of bridgehead atoms for choice of the correct combination is extremely tedious.

It appeared worthwhile to have a computer address the tedious chore of developing systematic names for the ring systems represented in the list of structures. Ideally, development of the complete systematic names, including substituent group and functional group designation, would have permitted the complete comparison of compounds in the list *via* their systematic names. However, the necessity for comparative brevity and simplicity in the program dictated that the substituent groups and functional groups be omitted, and the program finally employed names the hydrocarbon skeleton alone. Final comparisons, for the bipyran work, were done manually.

The procedure of comparing compounds by way of the systematic names of their skeletons is not the most efficient approach. At least two other approaches (3, 4) could have been made, but the possibility of generating organic nomenclature by computer seemed interesting enough to be worth pursuing in its own right, with the bipyran hydrolysis products serving as a source of structures to test the program on. The form of input was initially specialized for the bipyran problem, but has since been generalized so that any system can be input. Only a few restrictions as to the type of system which can be named have remained in the program as a legacy from its original purpose.

To illustrate the difficulties in recognizing the reorientation and numbering necessary to the naming process, Figure 1 gives as A the original orientations and numbering

tricyclo[5.2.1.0^{4,8}]decane (I)   tetracyclo[6.3.1.0^{2,6}.0^{5,11}]dodecane (II)
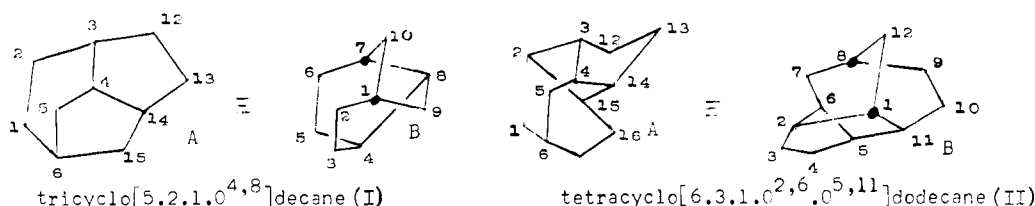
Figure 1. Input numbering, final numbering, and Baeyer names
of typical tricyclic (I) and tetracyclic (II) hydrocarbons.

of two structures as they occur in the structure list. The final structures, drawn and numbered in a way suggested by the name, are given as B.

The sequence of steps in generating the Baeyer system name (5) from the structural representation may be expressed operationally as follows:

1. Choose the largest possible ring as the main ring.
2. Choose the largest possible sequence of remaining atoms as the main bridge. If there are two candidates with an equal number of atoms, choose that one which divides the main ring most nearly symmetrically.
3. Designate subsidiary bridges by the number of atoms in them and, in superscripts, the points of attachment of the ends of these subsidiary bridges. If subsidiary bridges are located unsymmetrically with respect to the main ring and the main bridge, the choice of name is such that these superscripts have the smallest possible value.

This sequence of rules permits the unambiguous choice of a single pair of tertiary atoms, called main bridgehead atoms (marked as dots in B, Figure 1), which serve as the endpoints in numbering the system. The bridgehead atoms are connected by three main bridges, which are unbroken chains of atoms which proceed from one main bridgehead to the other. Although it is not explicit, it appears implicit in the rules that there must be three continuous, nonoverlapping bridges between the pair of bridgehead atoms which serve as the basis of a Baeyer name. Each bridge may contain subsidiary bridgehead atoms which are the termini of subsidiary bridges. Although not in the IUPAC rules, the term "segment" will be used herein to designate that portion of a bridge which extends from one bridgehead atom to an adjacent one, regardless of whether the bridgehead atoms are principal or subsidiary. A bridge might be composed of one, two, or more segments.

Once the proper bridgehead atoms are chosen, the number of atoms in the bridges connecting the main bridgehead atoms are given in brackets in descending order followed by the subsidiary bridges which bear superscripts denoting the positions of their attachment to the main bridges. Numbering of the ring system proceeds by starting at one bridgehead atom, proceeding along the longer branch of the main ring to the other bridgehead atom, back to the first bridgehead atom along the shorter branch of the main ring, and along the main bridge to the second bridgehead atom. Any atoms in subsidiary bridges are numbered last.

The steps which lead to the systematic name in the computer follow this same sequence. Every possible com-

bination of bridgehead atoms is examined. The relative acceptabilities of each alternative are examined at each level in the above sequence and the more satisfactory alternatives are retained. In this way, the remaining steps in the development of the name are applied only to the satisfactory alternatives, resulting in more efficient use of machine time. If only one alternative survives the sorting process, the appropriate information for that alternative is converted into the systematic name. In the event that the hydrocarbon skeleton is sufficiently symmetrical that two or more equally acceptable numberings exist, the program causes a symmetry message to be printed and a set of bridgehead atoms is identified for every acceptable numbering.

Structural information is handled in the computer in the form of a connection matrix (6). The digit "1" is entered in a 20 × 20 matrix (a 20-atom skeleton is the maximum the program handles) in the ij and the ji positions if atoms i and j are bonded in the ring system, all other entries in the matrix being zero. Structural information is input in two categories. First are given those atoms which are consecutively connected in the system as it is arbitrarily numbered (1-7 and 11-17 in the bipyran work). Then are given the extra connections which exist (the 4-14 bond and the bonds established by condensation reactions). The sum of the digits in row i of the connection matrix identifies carbon atom i as primary, secondary, tertiary, or quaternary.

Before the naming of the skeleton can begin, substituents, which can cause confusion in subsequent operations, must be removed. This is done by symbolically deleting each row (and column) of the connection matrix whose sum is 1, since these atoms are the termini of carbon chain substituents. The atom attached to the one discarded becomes terminal if it is part of the substituent and is discarded in its turn. When all substituent atoms are discarded, there remain only the secondary and tertiary atoms of the polycyclic system. The number of atoms remaining gives the root name of the system.

(When quaternary ring atoms are identified, a message indicating this fact is printed and no further processing of the name is done. Polycyclic systems containing such quaternary bridgehead atoms are rare and the inability of the program to handle these systems is not a serious shortcoming.)

This device for eliminating substituent groups succeeds for all alkyl groups, but fails for cycloalkyl groups because these groups have no primary atoms to initiate the discard process. Recognition of cycloalkyl groups becomes possible at a later stage after segments have been identified,

because a cycloalkyl group is a segment whose first and last atoms are identical. A message identifying these cases is printed, and no further development of the name is done.

The bridgehead atoms are identified by their tertiary nature, and a list of these atoms is accumulated, each one identified by its row in the connection matrix which corresponds to its initial numbering in the input system. The number of bridgehead atoms (2,4, or 6, respectively) identifies the system as bicyclic, tricyclic, or tetracyclic. Extension of the program to the naming of acyclic, monocyclic, and pentacyclic or more complex systems was not attempted.

The next task is to identify the segments which connect the bridgehead atoms. Three segments lead from each bridgehead atom, so there are 6, 12, and 18 segments, respectively, to be identified in bicyclic, tricyclic, and tetracyclic cases. (Each segment is listed twice, once in each direction.) Suppose atoms j and n are bridgehead atoms connected by the chain of atoms, j, k, l, m, and n. There will appear a "1" in the kth column of the jth row of the connection matrix. If row k is now examined, it will be found to contain "1" in positions kj and kl. Since the jk connection was already known, it is ignored and row l is next examined, where it is found that the next atom in sequence is m. This procedure is followed until the bridgehead atom n is encountered, and the segment is recognized as complete. Three items are stored as a result of this identification of a segment: The bridgehead atom at the start (j), the number of atoms in the segment (5), and the bridgehead atom encountered at the end of the segment (n). This same procedure is followed three times for each bridgehead atom so that information about all 6, 12, or 18 segments is stored. This operation completes the use of the connection matrix. Any alternative mode of storage of structural information could be used to produce this collection of segments, but from this point on it should not be necessary to change the program to accomodate changes in the form of the input structural information. The stored information at this stage is as given in Figure 2. It is at this point that cycloalkyl substituents are recognized and aborted.

Before the Baeyer name can be constructed, the complete bridges between every combination of bridgehead atoms must be formed by combination of the segments already listed. Only after that can the characteristics of the complete bridges be compared for selection of the correct name. The complexity of this procedure is a steep function of the number of rings in the hydrocarbon system. A bicyclic system offers no difficulty because only two bridgehead atoms exist and these two necessarily serve as the basis for construction of the name. The segments

already listed constitute the complete bridges and are stored in a standard form without further processing.

In tri- and tetracyclic systems, several topologically different ways in which the segments may be combined into the complete systems may be envisaged (7). This observation enables one to systematize the conversion of bridge segments into permissible complete bridges. The topologically distinct possibilities are given schematically in Figure 3.

In the tricyclic case, it is necessary that at least one bridge consist of a single segment. The two cases differ in regard to the other bridges. If, as in case a, two one-segment bridges connect a given pair of bridgehead atoms, then the third bridge is a three-segment bridge which involves another pair of bridgehead atoms. The tricyclic system is such that one or the other pair of bridgehead atoms *must* serve as the basis for the Baeyer name to avoid overlapping of the bridges. Attention is immediately focused on these two pairs of bridgehead atoms and both alternative long bridges based on each pair of bridgehead atoms are stored in the standard form to be described below.

In the more symmetrical case (case b), no immediate selection of certain pairs of bridgehead atoms on the basis of lack of symmetry is possible, and every combination of bridges must be discovered and listed. In this case, every bridgehead may be reached from a given bridgehead by one one-segment and two two-segment bridges. Since, coincidentally, there are as many combinations of bridgehead atoms (12) as there are segments (12), it is convenient to base the search on the combination of bridgehead atoms represented in each segment in turn. Using segment J, the one-segment bridge, as a starting place, the segment J1 is found since it is one of the other segments starting from bridgehead 1. Segment K is the only segment which connects the end of segment J1 to bridgehead atom 2 (the end of segment J). Similarly, the bridge J2-M is constructed by connecting the end of the third segment from bridgehead 1 to that segment which reaches from it to' bridgehead 2. Finally the segment MM which connects the two subsidiary bridgehead atoms is found by search. Appropriate information about these bridges is stored in a standard form to be given below.

In the tetracyclic case, a number of different arrangements of the bridge segments into over-all cyclic systems may be envisaged. Three of these are analogous to case a of the tricyclic systems in that there is at least one pair of segments which begin and end at the same pair of bridgehead atoms. Since the requirement of nonoverlapping bridges places special restrictions on the choice of bridgehead atoms which would require special handling, and since these systems do not occur in the bipyran work, the program was arranged to recognize and abort these

3	3	3	4	4	4	6	6	6	14	14	14	0	0	0	0	0	0
4	2	4	2	3	2	4	3	3	2	4	3	0	0	0	0	0	0
6	4	14	3	6	14	3	4	14	4	3	6	0	0	0	0	0	0

2	2	2	3	3	3	4	4	4	6	6	6	14	14	14	15	15	15
3	2	2	2	2	4	2	3	2	3	3	4	2	4	2	2	2	4
6	3	15	2	4	14	3	6	14	2	4	15	4	3	15	2	14	6

Figure 2. The lists of segments generated during
computer treatment of compounds I and II.
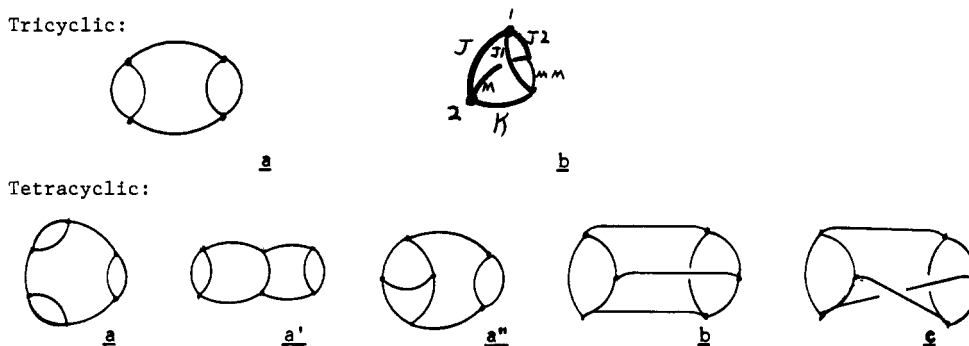
Tricyclic:

Tetracyclic:

Figure 3. Possible combinations of segments to form tricyclic and tetracyclic skeletons.

systems when encountered. This is probably the most severe of the restrictions to the program's applicability.

The remaining arrangment of bridgehead atoms (b) is so highly symmetrical that every combination of bridgehead atoms must be systematically listed with the complete bridges between them, and then the lists compared for the choice of the best name. This arrangement is topologically a triangular prism: Every bridgehead atom is in a topologically identical situation so that a general description of the bridges between arbitrary combinations of bridgehead atoms may be made. (A compound which is topologically a twisted prism [c] gives absurd names when given to the program. Fortunately these compounds are not chemically feasible unless one bridge is very long, so this restriction on the program will not prove severe.)

As illustrated in Figure 4, diametrically opposite bridgehead atoms are connected by bridges of three, two, and two segments. Bridgehead atoms on a common edge of a quadrangular face are connected by bridges of three, three, and one segments. Bridgehead atoms on the same triangular face are connected by bridges of four, two, and one segments. The prism is oriented in such a way that segment J is along a quadrangular face of the prism. (This is recognized by the fact that there is no single segment connecting the end of J to the end of J1 or J2.) The remaining segments are linked up by searching in the same general way as done for the tricyclic case. Once the segments are identified in terms of their positions on the triangular prism, it is a simple matter to write down the characteristics of the five sets of complete bridges which connect the first bridgehead to each of the five other bridgehead atoms in the system. This process is repeated for each of the six bridgehead atoms, thus generating 30 lists of bridges corresponding to each of the 30 possible names for a given tetracyclic system. This information is stored in a standard form. These operations complete the use of the lists of bridge segments.

The standard form of storage of bridge information is a 21-row column vector. Bicyclic, tricyclic, and tetra-

cyclic systems generate 2, 12, and 30 of these vectors, respectively. The first 12 rows in each vector (three bridges of up to four segments) are reserved for the lengths of the segments which comprise each bridge. Rows 13 and 14 give the first and second main bridgehead atom as numbered by its position in the input connection matrix. Rows 15 to 17 and 18 to 20 are used to store information about subsidiary bridges. Entries 15 and 18 give the length of the subsidiary bridge and the next two entries are two-digit codes giving, first, the bridges bridged by the subsidiary bridge, and second, the subsidiary bridgeheads in the main bridges connected by the subsidiary bridge. Position 21 is reserved for a code indicating an ambiguity in the assignment of superscripts. At this stage the stored information is as given in Figure 5.

The next task is to arrange each set of bridges in order of decreasing size. Although the bridges have already been stored in order of decreasing number of segments, it is necessary to calculate their lengths in terms of the number of atoms because this may give a different order. In calculating bridge length, proper care must be taken that subsidiary bridgehead atoms are not twice counted as they would be if the segment lengths were simply added. If an inversion in the order of bridges is required to get them in descending order, this is done and appropriate reassignment of code numbers in rows 16 and 19 of the column vector is made so as to keep track of which main bridges the subsidiary bridges are located on. If two main bridges have the same number of atoms, they are ordered so that the first superscript on the subsidiary bridge will be as small as possible. In tetracyclic cases, the subsidiary bridges are also arranged in order of decreasing numbers of atoms. In cases where the number of atoms in the bridges does not determine the order of their representation in tetracyclic systems, the order of representation of the subsidiary bridges and of the main bridges interact in a complex way. Rather than try to program this detailed interaction, it was merely arranged that a message that the assignment of super-
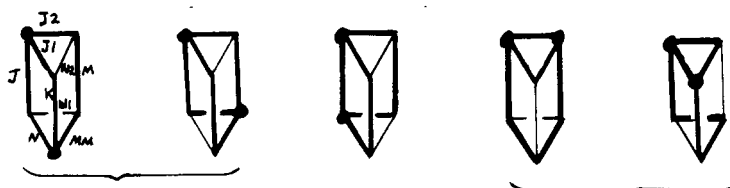

Figure 4. The combinations of segments connecting arbitrary pairs of bridgehead atoms (dotted) are indicated in heavy lines on the triangular prism.

Figure 5. The collection of column vectors describing the possible bridges as listed before ordering during computer treatment of compounds I and II.

scripts may be incorrect be printed out. A "1" in row 21 of the column vector indicates this circumstance. Every effort was made to have the program correctly calculate the superscripts even in ambiguous cases; in the five examples so far encountered, it has done so. The information stored at this stage of the work is presented in Figure 6.

The column vectors corresponding to each choice of bridgehead atoms are in the order which corresponds to the name which would result if that bridgehead combination were used. It remains to choose among all the possibilities that one which gives the best name. The size of the largest ring is found by adding the lengths of the segments which compose the two longest bridges connecting the chosen bridgehead atoms, proper allowance again being made to prevent twice counting subsidiary bridgehead atoms. The length of the main bridge is found by adding the segments composing the third of the three bridges. The difference in the lengths of the first and second main bridges is determined. These three items permit the discard of most of the alternative candidates for the Baeyer name. The candidates with the largest main ring, the longest main bridge, and the main bridge most symmetrically dividing the main ring are retained at the end of this stage in the program. The retained vectors at this stage are shown in Figure 7.

It remains to convert the codes which specify the points of attachment of the subsidiary bridges into numbers which specify the atoms as correctly numbered in the Baeyer system. This is a matter of summing the appropriate bridge segments, proper allowance being made to prevent double counting and cognizance being taken of the convention for direction of numbering along each bridge. The lengths of subsidiary bridges are reduced by two to account for the fact that their termini are already specified (Figure 8). When the superscripts designating the ends of the subsidiary bridges have been determined, that candidate for the name which has the smaller superscripts at the first difference, if any, is kept (Figure 9).

A knowledge of the symmetry of a bridged system, besides its intrinsic interest in subtle cases, is a necessity for the correct choice among two or more alternative numberings in the case of substituted or functional derivatives. The foregoing process of eliminating unsatisfactory candidates for the name reveals one kind of symmetry if two or more candidates for the name remain after all the discards have been made. This occurs if there is an axis of symmetry which rotates one bridgehead atom into the other (or, alternatively, one pair of bridgehead atoms into another equivalent pair). Thus, if more than one column vectors remain, a symmetry message is printed and a message identifying each equivalent pair of bridgehead atoms is printed. A different, more restricted, symmetry element exists when a plane of symmetry contains the bridgehead atoms and one bridge, but reflects the other two bridges into one another. This condition is detected by pairwise comparison of the segments of the
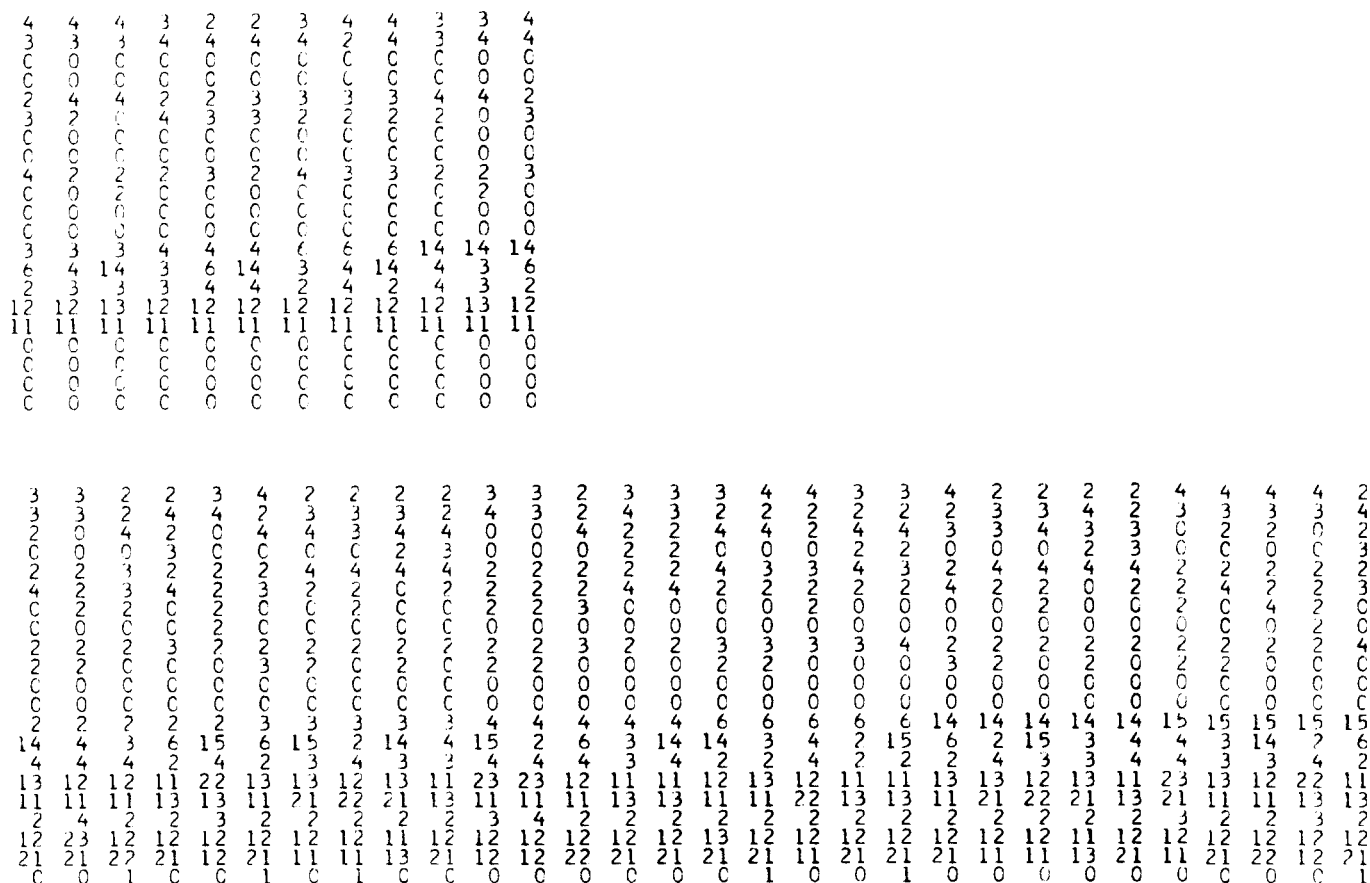
```
4  4  4  3  2  2  3  4  4  3  3  4
3  3  3  4  4  4  4  2  4  3  4  4
0  0  0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0  0  0  0  0
2  4  4  2  2  3  3  2  3  4  4  2
3  2  4  4  3  3  2  2  2  2  0  3
0  0  0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0  0  0  0  0
4  2  2  2  3  2  4  3  3  2  2  3
0  0  0  0  0  0  0  0  0  0  2  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
3  3  3  4  4  4  6  6  6  14 14 14
6  4  14 3  6  14 3  4  14 4  3  6
2  3  3  3  4  4  2  4  2  4  3  2
12 12 13 12 12 12 12 12 12 12 13 12
11 11 11 11 11 11 11 11 11 11 11 11
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  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0  0  0  0  0
```

Figure 6. The collection of column vectors after each has been put in order of decreasing bridge length.

bridges; if two bridges are identical, a symmetry message is printed. Thus, a symmetry message with a single bridgehead choice is indicative of a plane of symmetry through the bridgehead atoms, and a symmetry message with more than one bridgehead message is indicative of an axis of symmetry. Two symmetry messages indicate both symmetry elements.

To construct the actual name, the numerals in brackets have to be determined. To accomplish this, the lengths of the segments composing the three bridges are again added, taking account of the fact that the bridgehead atoms are not included in these numerals. The number of atoms in the subsidiary bridges and the superscripts on them are already available in rows 16 to 20 of the column vector and are printed out directly. The number of rings and the number of atoms in the polycyclic system are available from earlier calculations and are incorporated into the systematic name in the final output (Figure 10).

The approach of the present program cannot easily be extended to pentacyclic and more complex systems, because the present program is dependent upon there being a limited number of topologically different patterns in which the four or six bridgehead atoms of tricyclic and tetracyclic systems can be arranged. The eight bridge-
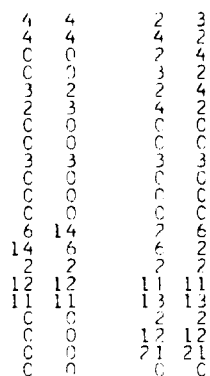
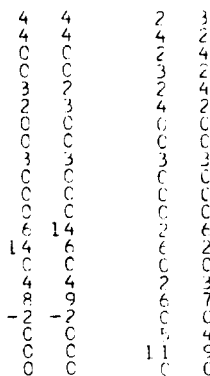Figure 7. The column vectors remaining after discard on the basis of the first two rules.

Figure 8. The column vectors after conversion of subsidiary bridge codes into superscript numbers.
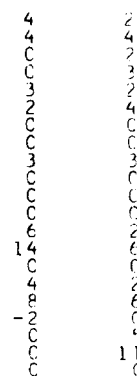
Figure 9. The column vectors after final discard on the basis of the third rule.

```
BRIDGEHEAD 1 IS  6, BRIDGEHEAD 2 IS 14.
 414 1 6 31215 6 C C C C TRICYCLO( 5. 2. 1. 0( 4, 8)) 10ANE.


BRIDGEHEAD 1 IS  2, BRIDGEHEAD 2 IS  6.
 414 1 6 31217 615 2 C C TETRACYCLO( 6. 3. 1. 0( 2, 6). 0( 5,11)) 12ANE.
```

Figure 10. The computer output for compounds I and II, identifying the bridgehead atoms by their original numbering, identifying the system (in terms of its input form), and giving the Baeyer name. Since the examples lack symmetry or ambiguity in assignment of superscripts, these messages do not appear.

Table I. Incorrect and Corrected Baeyer Names from the Literature

Literature Reference[a]	Literature Name(Incorrect)	Computer Name
CA 62,2781s	tricyclo[2.1.1.0^{5,6}]hexane	tricyclo[3.1.0.0^{2,6}]hexane
CA 62,2669s	tetracyclo[3.1.0.0^{2,4}.0^{3,6}]hexane	tetracyclo[2.2.0.0^{2,6}.0^{3,5}]hexane
CA 62,2780s / RRI 8212	tricyclo[2.2.1.0^{2,7}]heptane	tricyclo[3.2.0.0^{2,7}]heptane
CA 62,2780s	tricyclo[3.1.1.0^{6,7}]heptane	tricyclo[4.1.0.0^{2,7}]heptane
CA 61,2595s	tetracyclo[2.2.1.0^{2,6}.0^{3,5}]heptane	tetracyclo[3.2.0.0^{2,7}.0^{4,6}]heptane
CA VI,11959s / RRI 8248	tricyclo[2.2.1.1^{2,7}]octane	tricyclo[3.3.0.0^{2,7}]octane
RRI 2135	8-thia-2,5,7-triaza...	2-thia-1,3,6-triaza...
RRI 2095	tricyclo[2.2.2.0^{2,6}]octane	tricyclo[3.2.1.0^{2,7}]octane
CA VI,8229s / RRI 8247	6-oxatricyclo[3.1.1.1^{3,7}]octane	4-oxatricyclo[3.2.1.0^{3,6}]octane
RRI 2055	6-oxatricyclo[3.2.1.0^{2,8}]octane	6-oxatricyclo[3.3.0.0^{2,8}]octane
CA 60,277s	7-azatricyclo[3.2.1.0^{2,6}]octane	8-azatricyclo[3.3.0.0^{2,7}]octane
CA VI,8229s / RRI 8251	3-oxatricyclo[3.2.1.1^{6,8}]nonane	4-oxatricyclo[4.3.0.0^{2,8}]nonane
RRI 8252	4-oxa...	3-oxa...
CA VI,8202s / RRI 2458*	8-oxa-9-azatricyclo[3.2.1.1^{3,6}]nonane	2-oxa-6-azatricyclo[3.3.1.0^{3,7}]nonane
RRI 2266*	tricyclo[3.2.1.1^{3,8}]nonane	tricyclo[4.2.1.0^{3,7}]nonane
CA 60,277s	6-aza...	4-aza...
CA V,8867s / RRI 2253*	6-oxa...	4-oxa...
RRI 2224*	6-oxa-7-thia...	4-oxa-5-thia...
CA 61,1873s / *CA* VI,8229s / RRI 10153	6-oxatricyclo[3.2.2.0^{2,8}]nonane	6-oxatricyclo[3.3.1.0^{2,8}]nonane
CA 61,1873s	2-oxatricyclo[3.2.2.0^{4,7}]nonane	7-oxatricyclo[4.3.0.0^{3,9}]nonane
RRI 2264*	tricyclo[3.3.1.0^{2,6}]nonane	tricyclo[4.3.0.0^{3,7}]nonane
CA 60,1848s	2-oxatricyclo[3.3.1.0^{4,8}]nonane	4-oxatricyclo[4.3.0.0^{3,7}]nonane
CA 62,2669s	tetracyclo[3.2.1.1^{3,8}.0^{2,4}]nonane	tetracyclo[4.3.0.0^{2,4}.0^{3,7}]nonane
RRI 2418*	2,5,9-trioxatricyclo[4.2.2.0^{3,10}]decane	5,9,10-trioxatricyclo[5.2.1.0^{4,5}]decane
CA V,8867s	4-oxatricyclo[4.3.1.0^{2,7}]decane	4-oxatricyclo[4.4.0.0^{2,5}]decane
CA VI,11515s	tetracyclo[3.3.1.1^{3,9}.0^{2,4}]decane	tetracyclo[4.4.0.0^{2,4}.0^{3,6}]decane
CA 57,2563	tricyclo[3.3.1.2^{2,8}]undecane	tricyclo[5.2.2.0^{4,8}]undecane
RRI 2854* / *CA* V,4090s	1,7-diazatricyclo[3.3.1.2^{2,9}]undecane	8,10-diazatricyclo[5.4.0.0^{4,8}]undecane
CA V,4090s	3,9-diazatricyclo[3.3.1.2^{3,9}]undecane	1,8-diazatricyclo[5.3.1.0^{3,8}]undecane
RRI 2577 / *CA* 62,2781s	tricyclo[4.2.1.2^{7,9}]undecane	tricyclo[5.4.0.0^{2,9}]undecane
RRI 3642	4-oxa-1-azatricyclo[4.3.1.1^{2,7}]undecane	5-oxa-1-azatricyclo[5.4.0.0^{3,9}]undecane
CA VI,1237s	3-azatricyclo[5.3.0.1^{2,6}]undecane	10-azatricyclo[5.3.1.0^{2,6}]undecane
CA 60,277s	1-azatricyclo[6.2.1.0^{4,11}]undecane	4-azatricyclo[5.3.1.0^{4,11}]undecane
CA 61,899s / RRI 4059	3,8-dioxatetracyclo[4.2.1.1^{2,5}.1^{9,11}]undecane	3,11-dioxatetracyclo[6.3.0.0^{2,6}.0^{5,9}]undecane
RRI 10652	3-oxa-8-aza...	3-oxa-11-aza...

[a] The "s" in the *CA* references refers to the subject index pages. The Roman numeral *CA* references are to the Collective Indexes. The RRI references are to the serial numbers of the Ring Index, Second Edition (1960), and its supplements (1963 and 1964). The * indicates that the name is noted in RRI as being in some way unsatisfactory.

head atoms of a pentacyclic system can be arranged in a large number of topologically different ways (7), so a large number of distinct cases would have to be considered. A nomenclature program based on the present approach would have to distinguish between them and contain a number of different sections specialized to each topology. Each compound will give 56 column vectors to be processed. A nomenclature program for these systems and more complex ones does not appear likely to repay the labor involved in writing it.

RESULTS

Although the program is quite lengthy and was written in FORTRAN, a language not perfectly adapted to a program which is more logical and nonrepetitive than arithmetic and repetitive, it proved fairly efficient. About 500 polycyclic systems were named in 2.58 hours, including compiling time, on an IBM 1410 computer. A listing of the program will be sent to persons requesting it.

A compilation of cases from the literature in which the Baeyer names have been incorrectly given is made in Table I, which contains 29 different basic skeletons. It is unlikely that this table represents an exhaustive list of such errors, since *Chemical Abstracts* subject indexes and the Ring Index were scanned rather cursorily for compounds whose names had a wrong look; in addition the limitations of the present program (no pentacyclic systems, no quaternary carbon atoms, no two one-segment bridges between bridgehead atoms in tetracyclic systems)

prevented the exhaustive computer treatment of even all those names with a wrong look. In the column giving computer names, the oxa-azaprefixes have been supplied and the partially digitized form of the computer output has been converted to standard nomenclature. Note that by far the most common error in the assignment of names has been failure to include the maximum number of atoms in the main ring of the compound.

LITERATURE CITED

(1) Conrow, K., Radlick, P., *J. Org. Chem.* 26, 2260 (1960).
(2) Meinwald, J., Crandall, J. K., *J. Am. Chem. Soc.* 88, 1300 (1966).
(3) Morgan, H. L., *J. Chem. Doc.* 5, 107 (1965).
(4) Spialter, L., *ibid.,* 4, 261, 269 (1964).
(5) See IUPAC rules A-31 and A-32, *J. Am. Chem. Soc.* 82, 5557 (1960).
(6) Spialter, L., *ibid.,* 85, 2012 (1963).
(7) Lederberg, J., *Proc. Natl. Acad. Sci. U.S.* 53, (1), 134 (1965).

The Application of Computers to the Retrieval of Selective Information Regarding the Anticancer Activity of Coordination Compounds*

STANLEY KIRSCHNER, STANLEY H. KRAVITZ, and JOHN MACK
Wayne State University, Detroit, Michigan 48202

During a study of the anticancer activity of complex inorganic compounds, a need arose for information concerning the anticancer and other physiological activity of complex inorganic compounds already tested for such activity. Additional information about these complexes, including structure and physical properties, was also desired, but all of this information proved to be difficult to locate. The primary problem is that the various compilations of data on anticancer activity (1, 2) make little

or no attempt to separate complex inorganic compounds from other types of compounds for which data are given. Therefore, it is necessary to examine data on all the compounds which have been tested in order to determine whether a substance is a complex compound, a simple salt, or some other type of material. Then, for compounds which are found to be complexes, the desired physical, chemical, structural, and physiological information must be extracted and stored in such a way as to be readily retrievable in a useful form. This paper provides one solution to the problem.