Compact Codes.
2. Bicyclic Saturated Hydrocarbons

Milan Randić

Department of Mathematics and Computer Science, Drake University, Des Moines, Iowa 50311, U.S.A. and Ames Laboratory — DOE,* Iowa State University, Ames, Iowa 50011, U.S.A.

Compact codes, recently introduced (ref. 1), have been applied to all known bicyclic saturated hydrocarbons as the first step in a systematic deriving compact codes for more complex ring systems. The codes are compared with IUPAC nomenclature for bicyclic compounds, as well as with WLN (Wiswesser Line Notation). Advantages of the compact codes are outlined, including qualified algebraic manipulations on the codes to derive codes for structurally related compounds.

INTRODUCTION

Since the early attempts to develop suitable chemical nomenclature, which goes back to Beilstein[2] and Baeyer[3], to the present time interest in deriving an »ideal« chemical coding system which can handle past, present and future chemical structures continues. An ideal system should have all the qualities desired by users, which can be summarized by asking that the codes be simple. Reid[4] has explicitly listed a dozen desirable characteristics, such as that codes be unique, reconstactable, brief, can be derived by hand etc. We are considering here general codes, applicable to any structure. If one restricts attention to a special class of compounds, such as benzenoid fused aromatic compounds, one can propose a number of useful alternative codes which may satisfy many of the desired characteristics. For example, several publications[5] discuss reduced graphs for benzenoid systems in which each benzene ring is represented by a single vertex, a notational device first used by Smith[6], rediscovered and developed subsequently particularly by Balaban[7]. Structural codes based on such »dualist« graphs of benzenoid systems have been found to be mathematically equivalent to structural codes representing rotational isomers of alkanes (*i. e.*, paths (self-avoiding walks) on a diamond grid[8] illustrating how two apparently unrelated problems have the same mathematical basis. Moreover, the case illustrates how graph theoretical schemes, despite of the fact that graphs are devoid of any 3-dimensional structure, can neverthelss encode

---

objects (such as rotational isomers) in 3-dimensional space. For an alternative approach see ref. 9, where an interesting connection is pointed out between graphical diagrams and structural codes. Polansky and Rouvray[10], and recently Wenchen and Wenjie[11], elaborated on the topic of encoding spatial arrangements of rings, while Dias[12] produced and arranged systematically various possible structures based on fusion of benzene rings in the periodic table of benzenoid isomers. If one relaxes the conditions and allows fusion of benzene rings in annulenenoid shape, then the diversity of forms increases.[13,14] The codes could be judged on how well they can be *extended* to these novel but closely *related* situations, and they may even suggest novel structures. However, *special* codes will not be suitable for a general structure, which should not diminish their utility.
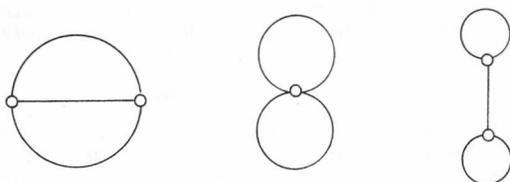

Figure 1. Three possible topologies for bicyclic molecules.
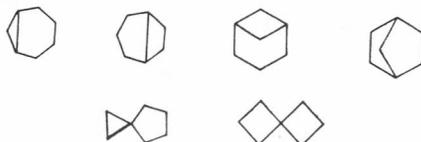

Figure 2. Molecular graphs of norbornane and its bicyclic isomers.

Bicyclic saturated compounds considered in this paper, represent also a very special class of structures. *The codes that we will outline however, are general compact codes applicable to any polycyclic system.* In this case there are three possible topologies (Figure 1); one is associated with three bridges between two branching vertices, another introduces a *spiro* vertex (a tetravalent vertex from which two independent loops originate) and finally in the third two rings are bridged. The last case is excluded from »Ring Systems Handbook« classification of possible rings.[15] For bicyclic systems having three branches Baeyer introduced and adequate and simple nomenclature which is still part of the IUPAC official nomenclature. One simply records the number of carbon atoms in each of the three branches. Hence, norbornane and its isomers (Figure 2) are: bicyclo[2.2.1]heptane; bicyclo-[3.1.1]heptane; bicyclo[3.2.0]heptane; bicyclo[4.1.0]heptane; spiro[3.3]heptane; and spiro[4.2]heptane respectively. When stripped of nonessentials, the above codes become respectively:

(2.2.1); (3.1.1); (3.2.0); (4.1.0); (3.3) and (4,2).

They allow unique reconstruction, when one observes that code (m, n) cannot be confused with a structure having two branches, as such a structure represents a single ring and will have code (m + n + 2).

The problem with simple Baeyer nomenclature is that it cannot be *simply* extended to polycyclic cases. There is ample evidence that already in the case of tricyclic systems, despite rigorously prescribed rules how to label atoms and mark individual branches, frequently incorrect names have been reported. For example, J. Meinwald and J. K. Crandall[16] in a brief appendix on nomenclature listed a few incorrect names found in the literature. K. Conrow[17] developed a computer program to generate Baeyer System names of saturated bridges bicyclic, tricyclic and tetracyclic hydrocarbons and found numerous violations of the IUPAC official rules in the literature. The major cause of difficulties in devising simple notation for polycyclic ring systems is that with an increase in the number of rings the number of distinct topologies also increases. For example, while for bicyclic systems we found only three different topologies, in the case of tricyclic systems the number of distinctive topologies is 12, as found by Tanaka and coworkers[18] using a graph theoretical approach. If one, however, restricts attention to a single topology (single class of structures) then simple codes may be possible, as illustrated in ref. 18 for the class of tricyclic structures to which adamantane belongs.

Our interest is restricted here to bicyclic systems for which we will report compact codes. We will also report IUPAC and WLN codes for the same class of compounds. While IUPAC names are straightforward, we will see that for this rather simple class of compounds already WLN codes could become somewhat lengthy, although not difficult to find. The differences between various coding systems will become even more evident when we extend the comparison to tricyclic systems in the forthcoming publication.[19] Our aim is not to offer criticism of the existing systems, they are brought into focus only to emphasize the advantages of our compact notation! We still have some way to go before our compact codes are extended to the wide class of chemical compounds. This is only the beginning in the development of a *novel* system. One should recollect that it took a large group of chemists and specialists and many years for the present schemes of IUPAC and WLN to establish themselves. There is no doubt that they are far from being perfect, which is reflected in their constant modifications and improvements. In their defense, however, one should be reminded that they have been developed under the pressure of a continuous and rapid growth of the number of chemical compounds and the volume of chemical literature in the real time. Hence, one should not be surprised that they present some conceptual, computational and operational difficulties. Novel systems *ought* to be free of such inconveniences, and we sincerely believe that the compact codes offer a fresh basis for the development of a comprehensive system of chemical codes. The present paper serves as introduction to these novel codes, which, as one can easily verify, have numerous advantages, such as being: unique, brief, simple, derivable by hand, reconstructible, and thus approach the goal of being »ideal«.

COMPACT CODES

The problem of designing codes for molecular structures is old and still open. The need for coding emerged in diverse applications, the oldest one, of course, being associated with the needs of chemical documentation. More

recently, however, codes became essential for computer processing of complex syntheses.[20] Ultimately, codes are of interest in developing possible structure-property-activity-function relationships.[21] For this purpose structural coding has apparent advantages, although partially structural codes (such as WLN) can to a considerable extent be adapted such research.[22] Despite numerous schemes, the problem may be considered open because no general scheme that satisfies numerous desirable requirements has yet been proposed or accepted. The schemes that are accepted, like the IUPAC Standard Notation, which **has the blessing of an international union** of chemical societies leans heavily on a set of trivial names, which makes it mathematically less attractive, and hence limited when mathematical manipulations of associated codes are considered. Similarly the WLN system, which gradually received wider acceptance, is a hybrid of structural and arbitrary (non-structural) codes. Considerable effort has been made to upgrade WLN and make it available for computer processing.[23] Its prime disadvantage is that it requires trained personnel, the construction rules being very involved and coding becoming quite complex and unwieldy. Lack of criticism of both the IUPAC and WLN is probably due to the *needs* at this point in time. The development of alternative schemes, of which there are numerous illustrations[24] does indicate indirectly some percieved faults or deficiencies of the currently authorized schemes. Hence, enthusiasm expressed for WLN[25] should be viewed with due reservation. Our aim here is not to criticize other schemes but to *confront* them with an alternative — which has yet to be fully developed in the codal and nomenclature system — and let the *users* judge the merits of each. In order to *induce* the use of the novel system we undertook the task of making available codes for several classes of rather common ring structures, starting with bicyclic systems.

We will illustrate the compact codes on an example. Consider norbornane, C_7H_{12}, the IUPAC name (as originally proposed by Baeyer) being: bicyclo-[2.2.1]heptane and its WLN code being: L55 ATJ. Compact codes consist of two parts: the first part lists reduced valencies of atoms which form the *acyclic backbone* of the structure, the second part lists valencies and neighbors for ring forming atoms. Hence, given a molecular structure, like norbornane, we have first to select that acyclic backbone and then consider atoms that make ring closure. The rules for finding the acyclic subgraph are few and simple.[1] In the case of norbornane a single rule suffices: »Remove as few atoms as possible which when erased will produce an acyclic system«. By erasing an atom one automathically also erases the incident bonds. In this case, if we remove one of the branching carbon atoms we have opened the ring system (Figure 3). We obtained a six vertex acyclic graph (tree representing 3-methylpentane). The first step in the construction of the code is to find the compact code for this acyclic portion of the structure. We consider this problem solved, as we adopted here the N-tuple codes of Knop, Müller, Szymanski, Jeričević and Trinajstić[26] to represent any acyclic portion of a structure that arises in our analysis. Knop and coworkers developed one of *the most advanced* codes for acyclic structures: their coding approach *implies* atomic numbering, hence no separate rules for numbering are required. At the same time their codes are compact, the number of entries in the code is n, *i. e.*, given by the number of atoms present (and if desired can be reduced to $n-1$, the number of bonds).[27] The

code is given by a string of digits: A B C D..., where a digit (like A, B) signifies an individual atom and blanks following each digit signify bonds. Terminal atoms are coded by zeros, thus pointing to the end of a path of consecutively connected atoms. In ref. 26 instructions are given how to find the code for acyclic structures like 3-methylpentane: First locate the atom of the highest (graph theoretical) valency. If there are several such atoms, the Rules specify how to select one of many. There is one such atom here carbon 3. The code starts by writing the valency of that atom. Such atom is always a branching atom, while in the case of linear chains it divides the molecule into two parts, instead of three or more. One now examines
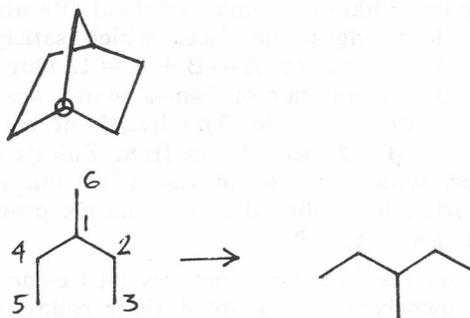

Figure 3. Construction of the compact code for norbornane.

each of the branch separately and then concatenates the codes for the branches. Entries following the leading digit in the code are given by $v - 1$, valency reduced by 1. Thus, for the three branches of 3-methylpentane we have: 10, 10, and 0. The Rule for concatanation states that codes are linked so that the resulting string corresponds to the numerically *largest* possible number. In this case we obtain the string 1 0 1 0 0 to follow the initial digit 3. Hence, the N-tuple code for 3-methylpentane is: 3 1 0 1 0 0. This code is unique, brief, easily derivable reconstructible, *etc.*, *i. e.*, it has all the desirable qualities specified by Ried[4] for molecular code. The *only* limitation is that the N-tuple codes apply to trees, acyclic structures. That, of course, is a serious limitation when one is interested in general graphs, but if one confines interest to trees, the N-tuple codes are ideal! Our Compact Codes can be viewed as an *extension* of the work of Knop, Müller, Jeričević and Trinajstić[26], a *nontrivial* extension to be precise. It is usually either very easy (*i. e.*, trivial) to extend an algorithm valid for acyclic systems to cyclic ones or not possible at all, unless some *novel* concept has been introduced. Our novel concept consists in »opening of rings by erasure of selected vertices«.[1] As outlined in ref. 1, the opening of rings by erasure of bonds alone (leaving all atoms intact) producing spanning trees, does not lead to a useful extension, because the number of spanning trees proliferates very fast. Instead, we erase individual atoms and obtain, what we have called *subspanning* trees. If a single atom is erased, the derived subgraphs are known a Ulam's subgraphs, associated with the famous Ulam's Reconstruction Conjecture.[28] To complete the code one now has to incorporate the erased atom and its connections. We accomplish this in the following way: Write first the valency

of the erased atom. In the case of several atoms, start with those of the lowest valency. Next *add* labels of its neighbours. Recollect that N-tuple code assigns an atomic label to each atom, the number is given by the *position* of the atom in the code. In the case of 3-methylpentane the atomic labels are shown in Figure 3. We see that carbon atoms 3, 5, 6 are terminal, and connect to the excised carbon atom 7. Hence, the full code is 3 1 0 1 0 0 3 3 5 6, in all 10 entries, seven of which designate valencies of atoms, the part 3 1 0 1 0 0 3. Only the first and the last entry here are true valencies, the other are reduced to $v - 1$. The remaining three labels indicate the ring closures. Let A, B and C represent the number of atoms, bonds and cycles. The corresponding customary notation in mathematical literature is respectively: V, E and F for vertices, edges and faces which satisfy the well known Euler's relation: $V - E + F = 1$ or $A - B + C = 1$. Our compact codes do not explicitly record B, the number of bonds, which are implied by blanks following qualified entries in a code. The length of compact code is then given by $A + C + 1$, or $B + 2$ (as follows from Euler's relation). Hence B, the number of edges, which is also indicated by the number of nonzero entries is the upper triangle of the adjacency matrix provides some measure of the complexity of a structure.[29]

Before reporting on the bicyclic structures let's comment on the N-tuple codes of Knop and coworkers. We adopted their codes for the acyclic part of our compact codes. Hence, one may say that compact codes for acyclic structures are those of Knop and coworkers.[26] They prescribed Rules for finding N-tuple codes, which have been used in a number of applications for search for trees with unusual walks,[30,31] search for endospectral trees.[32] for trees, such as construction and enumeration of all trees of various sizes,[26] search for trees with identical ID numbers,[33,34] generalization to codes for benzenoid structures[35] and application to various problems of benzenoid structures. They adopted the general Rule, that among *alternatives* (when there are several choices for the vertex of the highest valency) one selects the variant that corresponds to the *maximal* numerical (lexical) value for the code. They also decided to represent vertices by $(v - 1)$ rather than v (its formal graph theoretical valency) except for the initial vertex. This choice, which apparently has some computer-advantages can be easily replaced by the choice that all atoms are represented by their *true* valency. Such a change will not change anything in the procedure, one can follow their Rules and at the *end* replace $(v - 1)$ with v for all (but the first) entries. In this case the code for 3-methylpentane instead of N-tuple 3 1 0 1 0 1 we would have M-tuple 3 2 1 2 1 1. For the sake of possible future reference to this alternative we deliberately referred to the latter as M-tuple instead of N-tuple. For practicing chemists the latter appears a better alternative, because the numbers 3,2,1 now refer to tertiary, secondary and primary carbon atoms, respectively. This immediately makes the codes more familiar to organic chemists! We will, however, continue to use N-tuple representation but should one desire a change it is easy to replace 3 1 0 1 0 0 3 3 5 6 with 3 2 1 2 1 1 3 3 5 6. We feel that this kind of changes should be made at a later stage, after the compact codes have been exposed to a wider circle of chemists and have been completed with suggestions for incorporation of heteroatoms, multiple bonds *etc.*

COMPACT CODES FOR BICYCLIC STRUCTURES

As alerady mentioned, there are three distinctive topologies for bicyclic structures. We will first restrict analysis to structures included in »Ring Systems Handbook«, a recent publication of Chemical Abstract Service,[15] which excludes structures in which disjoint rings are linked by a bridge (e. g., structures like bicyclopropyl). We want to contrast our compact codes with available current schemes: IUPAC nomenclature and WLN notation. Strictly speaking, codes and nomenclature are distinct schemes for labelling chemical structures, as emphasized recently by Goodson,[36] but as the size and complexity of compounds increase, the distinction becomes vague, because names for very complex structures will necessarily involve considerable degree of coding. We will group bicyclic compounds into three general classes which differ in their topology and each will be treated separately.

Bicyclo[i.j.k]alkanes

In Tables I and II we listed the compact codes, the IUPAC standard names, and WLN notation for all ring systems listed in the Handbook.[15] Table I contains smaller systems, which we somewhat arbitrarily restricted to rings having eight carbon atoms at the most, while Table II lists systems involving larger rings. The compact codes for structures in Tables I and II are easy to find, because the bicyclic compounds of the form [i.j.k] have two equivalent bridge atoms, and erasure of one of them produces the subspanning tree, the backbone for the *N*-tuple (or *M*-tuple) code. Hence, in this case the vertex to be erased is unique (except for symmetry equivalence, which is immaterial for the construction of the code). In fact, suppose that symmetry is not apparent, then one would have to consider *both* branching carbon atoms as alternative. One would construct the code for both cases and select one that produced the canonical *N*-tuple (*i. e.*, *N*-tuple corresponding to the numerically largest number if the sequence is interpreted as a single number). One would find then in the case of compounds of Tables I and II *that either* selection products the *same N*-tuple. This is tantamount for establishing the presence of some symmetry in the considered graphs. Additional symmetry may be observed by analyzing the code itself. For example the bicyclo[2.2.0]hexane has the compact code: 3 1 0 1 0 3 3 5 6, the subspanning part: 3 1 0 1 0 in this case shows symmetry of the two branches: 1 0 (which are repeated twice), just as the IUPAC code [2.2.0] similarly shows equivalence of two branches. This simple symmetry is also noticeable in WLN code L44 TJ, by the repeating figures 4,4. Virtues of compact codes are not so outstanding in comparison with IUPAC and WLN as seen in Table I, primarily for two reasons: (1) the molecules are simple enough to allow considerable differences in the complexity of the competing codes; and (2) IUPAC codes in particular (and to some degree this is true for WLN notation) for this class are particularly *simple* because they represent *special* coding of the Baeyer design for bicyclic systems. In forthcoming papers on the codes in tricyclic and polycyclic systems the contrast in compact codes and current alternatives will become quite apparent. Some disadvantage of

TABLE I

Compact Codes for Bicyclic Hydrocarbons Having a Bridged Structure in Comparison with IUPAC Notation and WLN Codes

			Ring Class
C_4H_6	bicyclo(1.1.0)butane 2 0 0 3 1 2 3	L 33 TJ	(3, 3)
C_5H_8	bicyclo(2.1.0)pentane 2 1 0 0 3 1 3 4	L 34 TJ	(3, 4)
C_6H_{10}	bicyclo(3.1.0)hexane 2 1 1 0 0 3 1 4 5	L 35 TJ	3, 5)
C_7H_{12}	bicyclo(4.1.0)heptane 2 1 1 1 0 0 3 1 5 6	L 36 TJ	(3, 6)
C_8H_{14}	bicyclo(5.1.0)octane 2 1 1 1 1 0 0 3 1 6 7	L 37 TJ	(3, 7)
C_6H_{10}	bicyclo(2.2.0)hexane 2 1 1 0 0 3 2 4 5	L 44 TJ	(4, 4)
C_5H_8	bicyclo(1.1.1)pentane 3 0 0 0 3 2 3 4	L 44 ATJ	
C_7H_{12}	bicyclo(3.2.0)heptane 2 1 1 0 0 3 1 4 6	L 45 TJ	(4, 5)
C_6H_{10}	bicyclo(2.1.1)hexane 3 1 0 0 0 3 3 4 5	L 45 ATJ	
C_8H_{14}	bicyclo(4.2.0)octane 2 1 1 1 0 1 0 3 1 5 7	L 46 TJ	(4, 6)
C_7H_{12}	bicyclo(3.1.1)heptane 3 1 1 0 0 0 3 4 5 6	L 46 ATJ	
C_7H_{12}	bicyclo(2.2.1)heptane 3 1 0 1 0 0 3 3 5 6	L 55 ATJ	(5, 5)
C_8H_{16}	bicyclo(3.2.1)octane 3 1 1 0 1 0 0 3 4 6 7	L 56 ATJ	(5, 6)
C_9H_{16}	bicyclo(3.3.1)nonane 3 1 1 0 1 1 0 0 3 4 7 8	L 66 ATJ	(6, 6)
C_8H_{14}	bicyclo(2.2.2)octane 3 1 0 1 0 1 0 3 3 5 7	L 66 A BTJ	
$C_{10}H_{18}$	bicyclo(4.3.1)decane 3 1 1 0 1 1 0 0 3 4 7 8	L 67 ATJ	(6, 7)
C_9H_{16}	bicyclo(3.2.2)nonane 3 1 1 0 1 0 1 0 3 4 6 8	L 67 A BTJ	
$C_{13}H_{24}$	bicyclo(5.5.1)tridecane 3 1 1 1 1 0 1 1 1 1 0 0 3 6 11 12	L 88 ATJ	(8, 8)
$C_{12}H_{22}$	bicyclo(4.4.2)dodecane 3 1 1 1 0 1 1 1 0 1 0 3 5 9 11	L 88 A BTJ	
$C_{11}H_{20}$	bicyclo(3.3.3)undecane 3 1 1 0 1 1 0 1 1 0 3 4 7 10	L 88 A B CTJ	

WLN notation can be, however, seen already on bicyclic systems in Table II, in which ring sizes are increased. One can observe that as ring sizes increase, specifically with an increase in the number of *even* rings, the number of possible underlying graphs increases. Thus, ring types (3, 3); (3, 4); (3, 5); (3, 6); (3, 7); *etc.* have a single possibility, the ring types (4, 4); (4, 5); (4, 6); (4, 7); *etc.* have *two* possible forms, and the same is again true for ring types (5, 5); (5, 6); (5, 7); (5, 8); *etc.* With the increase of one of the rings to size 6 we have an additional ring type, as seen by examining Table I for (6, 6); (6, 7); (6, 8); *etc.* and (7, 7); (7, 8) *etc.* Still larger rings introduce new ring forms: for ring types (8, m) and (9, m), m = 8, 9, 10 ... there are four possible graphs (Figure 4), but not all of them are to be found in the Ring

TABLE II

Compact Codes for Bicyclic Compounds Having Larger Rings Illustrating Some
Limitations of WLN Codes When Bridges of Greater Length Occur

Class:

(8, 9)   IUPAC name / WLN notation / Compact code

$C_{13}H_{24}$ bicyclo(5.4.2)tridecane
                L 89 A B TJ
                3 1 1 1 1 0 1 1 1 0 1 0 3 6 10 12

$C_{12}H_{22}$ bicyclo(4.3.3)dodecane
                L 89 A B C TJ

(9, 9)

$C_{15}H_{26}$ bicyclo(6.6.1)pentadecane
                L 99 ATJ
                3 1 1 1 1 1 0 1 1 1 1 0 0 3 7 13 14

$C_{14}H_{22}$ bicyclo(5.5.5)tetradecane
                L 99 A BTJ
                3 1 1 1 1 0 1 1 1 0 1 0 3 6 11 13

(10, 10)

$C_{16}H_{30}$ bicyclo(6.6.2)hexadecane
                L-10--10- A BTJ
                3 1 1 1 1 1 0 1 1 1 1 0 1 0 3 7 13 15

$C_{14}H_{26}$ bicyclo(4.4.4)tetradecane
                L-10--10- A B C DTJ
                3 1 1 1 0 1 1 1 0 1 1 1 0 3 5 9 13

. . .

(22, 22)

$C_{32}H_{62}$ bicyclo(10.10.10)dotriacontane
                L-22--22- A B C D E F G H I JTJ
                3 1^9 0 1^9 0 1^9 0 3 11 20 30

$1^9 = 1 1 1 1 1 1 1 1 1$

Systems Handbook, presumably those not included are currently experimentally unavailable. Still larger ring sizes have additional possibilities, but only a fraction of them are listed in the Handbook, apparently due to limited availability of most of the possible structures. In Table II we list *all* those reported in the Handbook together with their respective compact codes, IUPAC names and WLN codes. As one can see, these larger cases show some (perhaps minor) limitations of the WLN codes: the necessity to enlist intermediate bonds in the bridge by use of letters A. B. C. D. E, ... Already some of the cases in Table II have used up almost half of the alphabet (!), and clearly the things will get worse with future reports on still larger structures. Because the IUPAC and WLN systems operate in real time, if current expressions become cumbersome and unwieldy one can expect some modifications to be introduced. An advantage of compact codes is that, regardles of future needs, the codes can be applied and will not produce awkward expressions.
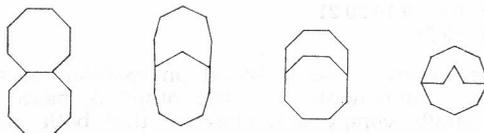

Figure 4. All possible bicyclic structures for class (8, 8) showing 1, 2, 3, and 4 common bonds for the two eight membered rings.

Spiro (m. n) alkanes

In deriving the compact codes for spiro compounds we have to decide between the two alternatives: (a) to erase *one* vertex and obtain *disjoint* fragments; and (b) to erase *two* vertices and have one *connected* subspanning tree. Which one should be chosen? We included in Table III *both* alternatives for the compact codes and tentatively selected the second option, but again

TABLE III

Compact Codes for Spiro-Bicyclic Compounds*

C_5H_8	Spiropentane	L 3XTJ A-& AL3XTJ
	1 0 1 0 4 1 2 3 4	
	2 0 0 2 1 2 2 1 3	
C_6H_{10}	Spirohexane	L 4XTJ A-& AL3XTJ
	2 0 0 1 0 4 2 3 4 5	
	2 1 0 0 2 1 3 2 1 4	
C_7H_{12}	Spiro(2.4)heptane	L 5XTJ A-& AL3XTJ
	2 1 0 0 1 0 4 3 4 5 6	
	2 1 1 0 0 2 1 4 2 1 5	
C_8H_{14}	Spiro(2.5)octane	L 6XTJ A-& AL3XTJ
	2 1 1 0 1 0 4 4 5 6 7	
	2 1 1 1 0 0 2 1 5 2 1 6	
C_9H_{16}	Spiro(2.6)nonane	L 7XTJ A-& ALXTJ
	2 1 1 1 0 0 1 0 4 5 6 7 8	
	2 1 1 1 1 0 0 2 1 6 2 1 7	
C_7H_{12}	Spiro(4.4)heptane	L 4XTJ A-& AL4XTJ
	2 0 0 2 0 0 4 2 3 5 6	
	2 1 1 0 0 2 2 4 2 2 5	
C_8H_{14}	Spiro(3.4)octane	L 5XTJ A-& AL4XTJ
	2 1 0 0 2 0 0 4 3 4 6 7	
	2 1 1 1 0 0 2 2 5 2 2 6	
C_9H_{16}	Spiro(3.5)nonane	L 6XTJ A-& AL4XTJ
	2 1 1 0 0 2 0 0 4 4 5 7 8	
	2 1 1 1 1 0 0 2 2 6 2 2 7	
C_9H_{16}	Spiro(4.4)nonane	L 5XTJ A-& AL5XTJ
	2 1 0 0 2 1 0 0 4 3 4 7 8	
	2 1 1 1 1 0 0 2 3 6 2 3 7	
$C_{10}H_{18}$	Spiro(4.5)decane	L 6XTJ A-& AL5XTJ
	2 1 1 0 0 2 1 0 0 4 4 5 8 9	
	2 1 1 1 1 1 0 0 2 3 7 2 3 8	
$C_{11}H_{20}$	Spiro(5.5)undecane	L 6XTJ A-& AL6XTJ
	2 1 1 0 0 2 1 1 0 0 4 4 5 9 10	
	2 1 1 1 1 1 1 0 0 2 4 8 2 4 9	
$C_{12}H_{22}$	Spiro(5.6)dodecane	L 7XTJ A-& AL6XTJ
	2 1 1 1 0 0 2 1 1 0 0 4 5 6 10 11	
	2 1 1 1 1 1 1 1 0 0 2 4 9 2 4 10	
$C_{16}H_{30}$	Spiro(7.8)hexadecane	L 9XTJ A-& AL8TJ
	2 1 1 1 1 1 0 0 2 1 1 1 0 0 4 7 8 14 15	
	2 1 1 1 1 1 1 1 1 1 1 0 0 2 6 13 2 6 14	
$C_{22}H_{42}$	Spiro(10.11)docosane	L-12-XTJ A-& AL-11-XTJ
	2 1^8 0 0 2 1^7 0 0 4 9 10 20 21	
	2 1^{17} 2 9 19 2 9 20	

* Two alternatives are shown, one is based on excision of a single (spiro) atom which produced two components and the other is based on excision of two vertices leaving a single component. Observe that both alternatives give codes of the same length. Observe also the somewhat cumbersome form of WLN codes, the length of which does not reflect the relative simplicity of the spirocompounds.

TABLE IV

Compact Codes for the Selection of Bicyclic Systems Having Exocyclic Bridges (i. e., Bridges Connecting Disjoint Rings), not Included in the Ring Systems Handbook

C_6H_{10}		2 1 0 0 2 1 4 2 2 3
C_7H_{12}		2 1 1 0 0 2 1 5 2 3 4
		2 1 1 0 0 2 1 5 2 2 4
C_8H_{14}		2 1 1 1 0 0 2 1 6 2 4 5
		2 1 1 1 0 0 2 1 6 2 3 5
		2 1 1 1 0 0 2 1 6 2 2 5
		2 1 1 1 0 0 2 2 6 2 3 5
C_9H_{16}		2 1 1 1 1 0 0 2 1 7 2 5 6
		2 1 1 1 1 0 0 2 1 7 2 4 6
		2 1 1 1 1 0 0 2 1 7 2 3 6
		2 1 1 1 1 0 0 2 1 7 2 2 6
		2 1 1 1 1 0 0 2 2 7 2 4 6
		2 1 1 1 1 0 0 2 2 7 2 3 6
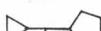
the final decision should be made once a sufficiently large body of compounds is examined. The prime criterion should be the simplicity of deriving the compact codes, and because bicyclic compounds are not complex enough they do not offer sufficient insight for taking the proper decision. We notice in passing that the codes derived by excision of the spiro carbon atom produce numerically *smaller* codes, which is one of the criteria for selecting the subspanning tree discussed in ref. 1. One should also recognize that for some molecules the opening of rings will necessarily introduce disjoint fragments (*e. g.,* radialenes)
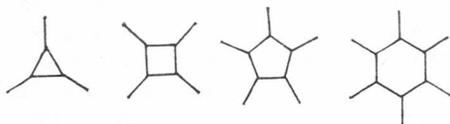
Hence, there is no compelling argument not to base selection on choice (a), but the decision is best delayed until we examine a number of spirocompounds having many rings.[37]

Spirocompounds have simple IUPAC names, while WLN appear somewhat lengthy, doubling the number of entries in the code as compared to codes of other bicyclic structures. Here we see some redundancies in the

WLN codes as compared to the compact codes. Let's consider the simplest case: spiropentane or what should have been spiro[2.2]pentane were the IUPAC names truly internally consistent. One can argue that [2.2] is redundant in the name spiro[2.2]pentane, but then if one is to be internally consistent then [1.1.0] is also redundant in the name of bicyclo[1.1.0]butane. The name bicyclobutane alone uniquely determines the form for the carbon skeleton of the molecule bicyclo[1.1.0]butane! Both alternative codes derived by either erasure of one carbon (spirocarbon) atom or two carbon atoms have the same number digits in their codes. The code can be considered as *additive*, since various branches or disjoint fragments are added, specifically concatenated. The codes end with list of ring closures. In the case of WLN codes we also see an additive process, as reflected by sign & but symbols encode larger building blocks, here fragments such as rings, rather than atoms. This is an important feature of numerous codes (including IUPAC standard names), as chemists prefer to see rings, fusion, bridges, functional groups, rather than individual atoms, which have a lower priority in conception of larger and more complex molecules. The »cost« of using larger fragments is a possible »overlapping« of smaller portions of the structure, atoms in particular. WLN constructs the codes by addition of two strings, one for each ring of a spirocompound: in case of spiropentane we have L3XTJ A-8 AL3XTJ. Observe the use of special symbols *and* use of spacings. Because the above labels describe rings rather than atoms we can expect further complications if the central atom is not carbon, but, say, silicon. Our compact codes are ready for inclusion of heteroatoms. Moreover, there is no need of special rules to define the codes for heteroatoms, one simply positions the symbol for heteroatoms at the corresponding site in the code, which depends on which carbon atom is to be replaced. If 2 0 0 2 1 2 2 1 3 represents spiropentane and if X stands for a heteroatom, then one can construct codes like: 2 X 0 0 2 1 2 2 1 3 and 2 0 X 0 2 1 2 2 1 3 to correspond to heterocompounds:
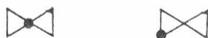
The heterobond is automatically deduced from the code by reconstruction.

Bridged bicyclic compounds

In Table IV we give the compact codes for the selection of bicyclic hydrocarbons having bridged cyclic regions. Here one has to excise two vertices to open the two disjoint rings. The codes for the bicyclic bridges systems are structurally very similar to the compact codes of spiroalkanes based on the erasure of two vertices. One can even derive one set of codes from the other by introducing the corresponding bridge enlargement operator B(+1) which is illustrated by the codes of spiropentane and derived bicyclopropyl: B(+1){2 0 0 2 1 2 2 1 3} = 2 1 0 0 2 1 4 2 2 3. The effect of B(+1) becomes more evident if one rewrites the spiropentane code as 2 0 0 2 1 3 2 1 2, then one has to add 1 in the N-tuple portion of the code and increase all labels representing the ring closure, except label 1. Observe also another regularity in Table IV: all isomers have the same N-tuple part of the code; they differ in the label part of the codes.

CODE MANIPULATIONS

Comparison of compact codes associated with molecules of same underlying topology but involving rings of different size shows simple regularity illustrated Table V. These regularities allow one to write down the form

TABLE V

Regularities in Codes of Structurally Related Molecules, Allowing to Write a General Code as well as Operate on Codes of Smaller Molecules to Derive the Codes of Larger Ring Structures

2 0 0 3 1 2 3	2 0 0 2 1 2 2 1 3
2 1 0 0 3 1 3 4	2 1 0 0 2 1 3 2 1 4
2 1 1 0 0 3 1 4 5	2 1 1 0 0 2 1 4 2 1 5
2 1 1 1 0 0 3 1 5 6	2 1 1 1 0 0 2 1 5 2 1 6

$$2\,1^{n-4}\,0\,0\,3\,1\,n{-}2\,n{-}1 \qquad\qquad 2\,1^{n-4}\,0\,0\,2\,1\,n{-}2\,2\,1\,n{-}1$$

$$R(+1)\,(2\,1^{n-4}\,0\,0\,3\,1\,n{-}2\,n{-}1) = 2\,1^{n-3}\,0\,0\,3\,1\,n{-}1\,n$$

$$R(+1)\,(2\,1^{n-4}\,0\,0\,2\,1\,n{-}2\,2\,1\,n{-}1) = 2\,1^{n-3}\,0\,0\,2\,1\,n{-}1\,2\,1\,n$$

of the code for a general member of the class, regardless its size. In the case of bicyclo[m.1.0]alkanes we see that an *increase* in the size of the ring from m to m+1 is reflected in its compact code by addition of an entry 1 and the increase of the terminal labels t-2 and t-1 by 1. Hence, we can introduce an operator R(+1) (ring enlargement) defined by:

$$R(+1)\{2\,1\,1\ldots0\,0\,3\,1\,t{-}2\,t{-}1\} = 2\,1\,1\,1\ldots0\,0\,t{-}1\,t$$

which relates the compact codes of structures differing in size of a single ring. In the case of spirocompounds the ring enlargement operator is defined by:

$$R(+1)\{2\,1\,1\ldots0\,0\,2\,1\,t{-}2\,2\,1\,t{-}1\} = 2\,1\,1\,1\ldots0\,0\,2\,1\,t{-}1\,2\,1\,t.$$

Other related operations can be similarly defined, like simultaneous enlargement of two rings, or increase of one ring and simultaneous decrease of the other. Some such operations may become of interest in special applications, such as substructure search, and we intend to examine the subject more thoroughly at a later date.

CONCLUDING REMARKS

The compact codes illustrate the coding system which is simple to implement and does not require any prior experience. The bicyclic structures considered here are simple enough and do not justly illustrate the simplicity of our scheme, as compared to others, because almost *any* scheme devised will be simple in the case of bicyclic structures and codes will be brief. In subsequent papers of this series we will see how compact codes surpass other schemes in simplicity and efficiency of coding. It is very likely that the present compact codes are the shortest possible codes not using special symbols and not being confined to a special class of compounds. But being brief, one may argue, is today a less compelling quality in view of the advances in computer facilities than some 15—20 years ago. This argument was raised as one that was used to suggest that the WLN notation may not be so important in comparison with the full list of adjacency relation-

ships.[38] While it is true that brief is not as essential today as it was in the recent past because of the increased capacities of computers, it is still a desirable quality. Most chemical graphs are represented by sparse matrices, which are »wasteful« of memory. On the other hand, frequent use of codes, in searches of data, would accumulate »waste«, and no argument will sustain the criticism of necessity for short codes, *providing that such codes are simple and easy to use,* particularly easy to use by everyone. It is likely that the considerable opposition to numerous coding systems has been provoked by the systems themselves — being too complex for use by chemists, except a few specialists or specially trained people.[39] The compact codes will, no doubt, defuse such opposition — and are likely to mushroom and replace most systems — by virtue of their pronounced simplicity.

REFERENCES

1. M. Randić, *J. Chem. Inf. and Comput. Sci.* (in press).
2. See: R. Luckenbach, *J. Chem. Inf. Comput. Sci.* **21** (1981) 82.
3. A. Baeyer, *Ber.* **38** (1900) 3771.
4. R. C. Read, *J. Chem. Inf. Comput. Sci.* **23** (1983) 135.
5. N. Trinajstić, Z. Jeričević, J. V. Knop, W. R. Muller, and K. Szymanski, *Pure and Appl. Chem.* **55** (1983) 379. P. Dzonova-Jerman--Blazic and N. Trinajstić, *Computers and Chemistry* **6** (1982) 121. (In this paper different reduced graph are considered in which alternate carbon atoms are only retained). J. V. Knop, K. Szymanski, Z. Jeričević, and N. Trinajstić, *Int. J. Quant. Chem.* **23** (1983) 713. W. C. Herndon and H. Hosoya, *Tetrahedron* **40** (1984) 3987. T. Yamaguchi, M. Suzuki, and H. Hosoya, *Nat. Sci. Report Ochanomizu Univ.* **26** (1975) 39. I. Gutman and S. El-Basil, *Z. Naturforsch.* **9a** (1983) 276. D. Bonchev and A. T. Balaban, *J. Chem. Inf. Comput. Sci.* **21** (1981) 223. S. El-Basil, *Croat. Chem. Acta* (in press).
6. F. T. Smith, *J. Chem. Phys.* **34** (1961) 793. Smith introduced a simplified notation to indicate individual molecules in a table: »A dot or the intersection of two lines represent a benzene hexagon, a line represents the junction of two hexagons through a common edge« a footnote to the table is all that is said. Balaban and Harary (vide infra) have rediscovered and were first to use the above representation of benzenoid molecules as reduced molecular graphs, which thus become mathematical objects.
7. A. T. Balaban and F. Harary, *Tetrahedron* **24** (1968) 2505.
8. A. T. Balaban, *Rev. Roum. Chim.* **22** (1977) 45.
9. M. Randić, *Int. J. Quant. Chem. Symp.*
10. O. E. Polansky and D. H. Rouvray, *MATCH* **2** (1976) 63.
11. H. Wenchen and H. Wenjie, *Theor. Chim. Acta* (in press).
12. J. R. Dias, *J. Chem. Inf. Comput. Sci.* **24** (1984) 124. J. R. Dias, *Can. J. Chem.* **62** (1984) 2914. J. R. Dias, *J. Chem. Inf. Comput. Sci.* **22** (1982) 15. J. R. Dias, *MATCH (Math. chem.* **14** (1983) 83. J. R. Dias, *Nouv. J. Chim.* **9** (1985) 125.
13. G. Ege and H. Vogler, *Theor. Chim. Acta* **26** (1972) 55. G. Ege and H. Fisher, *Tetrahedron*, **23** (1967) 149. U. E. Meissner, A. Gensler, and H. A. Staab, *Tetrahedron Lett.* (1977) 3. U. E. Meissner, A. Gensler, and H. A. Staab, *Angew. Chem.* **88** (1976) 374. U. Meissner, B. Meissner, and H. A. Staab, *Angew. Chem. Int. Ed.* **12** (1973) 916. W. Jenny and R. Paioni, *Chimia* **23** (1969) 41.
14. M. Randić, B. M. Grimarc, and N. Trinajstić, (in preparation).
15. *Ring Systems Handbook,* 1984 Edition, publication of Chemical Abstract Service.
16. J. Meinwald and J. K. Crandall, *J. Amer. Chem. Soc.* **88** (1966) 1301.
17. K. Conrow, *J. Chem. Docum.* **6** (1966) 206.

18. N. Tanaka, T. Iizuka, and T. Kan, *Chem. Lett., (Japan)* (1974) 539.

19. M. Randić, work in progress (part 3 of this series).

20. Besides the few review articles we listed for several laboratories presently active in the field, a number of their earlier publications ending with a more recent, one so that interested reader can backtrack from the latest publication any other of the selected laboratories: J. Lederberg, Dendral-64. *A System for Computer Construction, Enumeration and Notation of Organic Molecules as Tree Structures,* (NASA CR 57029), (1964). J. Lederberg, *Proc. Natl. Acad. Sci.* **53** (1965) 134. R. E. Carhart, D. H. Smith, H. Brown, and C. Djerassi, *J. Amer. Chem. Soc.* **97** (1975) 5755. J. C. Wenger and D. H. Smith, *J. Chem. Inf. Comput. Sci.* **22** (1982) 29. C. L. Stevens, K. G. Taylor, M. E. Munk, W. S. Marshall, K. Noll, G. D. Shah, L. G. Shah, and K. Uzu, *J. Med. Chem.* **8** (1965) 1. M. E. Munk, C. S. Sodano, R. L. McLean, and T. D. Haskell, *J. Amer. Chem. Soc.* **89** (1967) 4158. D. B. Nelson, M. E. Munk, K. B. Gash, and D. L. Herald, Jr., *J. Org. Chem.* **34** (1969) 3800. M. E. Munk, C. A. Shelly, H. B. Woodrull, and M. O. Trulson, *Fresenius Zeit. Anal. Chem.* **313** (1982) 473. E. J. Corey and W. T. Wipke, *Science,* **166** (1969) 178. E. J. Corey, *Quart. Rev., (London)* **25** (1971) 455. E. J. Corej and G. A. Peterson, *J. Amer. Chem. Soc.* **94** (1972) 460. J. B. Hendrickson, *J. Amer. Chem. Soc.* **93** (1971) 6847. J. B. Hendrickson and G. Toczko, *J. Chem. Inf. Comput. Sci.* **23** (1983) 171. V. V. Serov, M. E. Elyashberg, and L. A. Gribov, *J. Mol. Struct.* **31** (1976) 381. J. Gasteiger, P. D. Gillespie, D. Marquarding, and I. Ugi, *Topics Curr. Chem.* **40** (1974) 1. J. Friedrich and I. Ugi, *MATCH* **6** (1979) 201. A. J. Thakkar, *Topics Curr. Chem.* **39** (1973) 3. G. Klopman and M. McGonigal, *J. Chem. Inf. Comput. Sci.* **21** (1981) 48. Y. Kudo and S. Sasaki, *J. Chem. Inf. Comput. Sci.* **16** (1976) 43. T. Yamasaki, H. Abe, Y. Kudo, and S. Sasaki, *CHEMICS,* in *Computer-Assisted Structure Elucidation,* (D. H. Smith, ed.), *Amer. Chem. Soc. Symposium Series* **54** (1977) 108. S. Sasaki, H. Abe, Y. Hirota, Y. Ishida, Y. Kudo, S. Ochiai, K. Saito, and T. Yamasaki, *J. Chem. Inf. Comput. Sci.* **18** (1978) 211. P. Picchiottino, G. Georgoulis, G. Sicouri, A. Panaye, and J. E. Dubois, *J. Chem. Inf. Comput. Sci.* **24** (1984) 241. M. F. Lynch and P. Willett, *J. Chem. Inf. Comput. Sci.* **18** (1978) 154. S. Zhu and J. Zhang, *J. Chem. Inf. Comput. Sci.* **22** (1982) 34. M. Uchino, *J. Chem. Inf. Comput. Sci.* **20** (1980) 116. M. Uchino, *J. Chem. Inf. Comput. Sci.* **20** (1980) 121. M. Uchino, *J. Chem. Inf. Comput. Sci.* **20** (1980) 124. M. Uchino, *J. Chem. Inf. Comput. Sci.* **22** (1980) 201.

21. I. Motoc, *Topics Curr. Chem.* **114** (1983) 93. N. A. B. Gray, C. W. Crandell, J. G. Nourse, D. H. Smith, M. L. Dageforde, and C. Djerassi, *J. Org. Chem.* **46** (1981) 703. N. A. B. Gray, R. E. Carhart, A. Lavnchy, D. H. Smith, T. Varkony, B. G. Buchanan, W. C. White, and L. Creary, *Anal. Chem.* **52** (1980) 1095. M. F. Delaney and P. C. Uden, *Anal. Chem.* **51** (1979) 1242. L. Hodes, *J. Chem. Inf. Comput. Sci.* **21** (1981) 128. M. Razinger, J. Zupan, M. Penca, and B. Barlič, *J. Chem. Inf. Comput. Sci.* **20** (1980) 158. J. Zupan, D. Hadži, M. Penca, M. Razinger, and B. Barlič, *Anal. Chem. Acta* **122** (1980) 103.

22. G. W. Adamson and D. Bawden, *J. Chem. Inf. Comput. Sci.* **15** (1975) 215. D. Elkins, and C. Hansch, *J. Chem. Docum.* **14** (1974) 65. M. J. Romanec, *J. Chem. Docum.* **14** (1974) 49. M. Randić, and C. L. Wilkins, *J. Chem. Inf. Comput. Sci.* V. V. Avidon, I. A. Pomarantsev, V. E. Golender, and A. B. Rozenblit, *J. Chem. Inf. Comput. Sci.* **22** (1982) 207.

23. E. Hyde, F. W. Matthews, L. H. Thomson, and W. J. Wiswesser, *J. Chem. Docum.* **7** (1967) 200. M. F. Lynch, *J. Chem. Docum.* **8** (1968) 130. C. M. Bowman, F. A. Landee, N. W. Lee, and M. H. Reslock, *J. Chem. Docum.* **8** (1968) 133. C. D. Farrell, A. R. Chauvenet, and D. A. Koniver, *J. Chem. Docum.* **11** (1971) 52. G. A. Miller, *J. Chem. Docum.* **12** (1972) 60. S. R. Heller and D. A. Koniver, *J. Chem. Docum.* **12** (1972) 55. C. E. Granito, S. Roberts, and G. W. Gibson, *J. Chem. Docum.* **12** (1972) 190. A. Zamora and T. Ebe, *J. Chem. Inf. Comput. Sci.* **16** (1976)

342 M. RANDIĆ

36. M. Osinga, and A. A. V. Stuart, *J. Chem. Inf. Comput. Sci.* **18** (1978) 26.

24. J. A. Silk, *J. Chem. Docum.* D. Lefkowitz, *J. Chem. Docum.* **7** (1967) 186. Y. Kudo and S. Sasaki, *J. Chem. Docum.* **14** (1974) 200. C. Lin, *J. Chem. Inf. Comput. Sci.* **18** (1978) 41. N. Lozach, A. L. Goodson, and W. H. Powell, *Angew. Chem. Int. Ed.* **18** (1979) 887. T. Nakayama and Y. Fujiwara, *J. Chem. Inf. Comput. Sci.* **20** (1980) 23. A. L. Goodson, *J. Chem. Inf. Comput. Sci.* **20** (1980) 172. A. L. Goodson, *Croat. Chem. Acta* **56** (1983) 315. W. C. Herndon, in: *Chemical Applications of Topology and Graph Theory*, (R. B. King, editor), Elsevier, Amsterdam. H. Abe, Y. Kudo, T. Yamasaki, K. Tanaka, M. Sasaki, and S. Sasaki, *J. Chem. Inf. Comput. Sci.* **24** (1984) 212. H. Abe, H. Hayasaka, Y. Miyashita, and S. Sasaki, *J. Chem. Inf. Comput. Sci.* **24** (1984) 216. H. Abe, T. Okuyama, I. Fujiwara, and S. Sasaki, *J. Chem. Inf. Comput. Sci.* **24** (1984) 220.

25. See: M. D. Rosenberg: *Introduction to the Symposium on the Uses and Applications of the Wiswesser Line Notation Today* (presented on Aug. 27, 1980 during the 180th National Meeting of the American Chemical Society, Las Vegas, Nevada and subsequent articles in *J. Chem. Inf. Comput. Sci.* **22** (1982) p. p. 87—113, written by W. J. Wiswesser; M. D. Rosenberg, M. Z. Debardeleben, and J. F. Debardeleben; W. A. Warr; D. R. Eakin; L. E. Fritts and M. M. Schwind; and T. M. Johns and M. Clare; for a less enthusiastic article about future of WLN which lists a number of limitations of the system, presented at the same Symposium, see **ref. 37.**

26. J. V. Knop, W. R. Müller, Z. Jeričević, and N. Trinajstić, *J. Chem. Inf. Comput. Sci.* **21** (1981) 91. N. Trinajstić, Z. Jeričević, J. V. Knop, W. R. Müller, and K. Szymanski, *Pure and Appl. Chem.* **55** (1983) 379.

27. The last zero in the N-tuple code can be deleted, in which case the number of entries is that of the number of bonds.

28. F. Harary, *Graph Theory*, Addison-Wesley, Reading, Mass. (1969).

29. The concept of »complexity of a graph« has yet to be defined. It is likely that there will be several alternative definitions, and that (like with the concept of similarity among structures) one will have to *specify* the concept with respect to a particular aspect. One such specification is complexity as measured by the number of bonds, another with respect to the number of paths, *etc.*

30. J. V. Knop, W. R. Müller, K. Szymanski, M. Randić, and N. Trinajstić, *Croat. Chem. Acta* **57** (1983) 409.

31. »Unusual walks« have been introduced in: M. Randić, W. L. Woodworth, and A. Graovac, *Int. J. Quant. Chem.* **24** (1983) 435.

32. J. V. Knop, W. R. Müller, K. Szymanski, N. Trinajstić, A. F. Kleiner, and M. Randić, *J. Math. Phys.* (in press).

33. K. Szymanski, W. R. Müller, J. V. Knop, and N. Trinajstić, *J. Chem. Inf. Comput. Sci.* (submitted).

34. Molecular ID numbers were introduced in: M. Randić, *J. Chem. Inf. Comput. Sci.* **24** (1984) 164.

35. J. V. Knop, K. Szymanski, and N. Trinajstić, *MATCH* **16** (1984) 103 and references cited therein.

36. A. L. Goodson, *J. Chem. Inf. Comput. Sci.* **20** (1980) 167.

37. M. Randić, vork in progress.

38. V. B. Bond, C. M. Bowman, L. C. Davison, P. F. Roush, and L. F. Young, *J. Chem. Inf. Comput. Sci.* **22** (1982) 103.

39. Very few coding systems appear simple so that the work can be delegated to administrative staff, which is an important asset, as pointed out by L. Quadrelli, V. Bareggi, and S. Spiga, *J. Chem. Inf. Comput. Sci.* **18** (1978) 37 their linear representation of chemical structures satisfying this ciriterion.

SAŽETAK

Kompaktni kodovi. 2. Biciklički zasićeni ugljikovodici

M. Randić

Kompaktni kodovi primijenjeni su na zasićene bicikličke ugljikovodike, kao prvi korak u nastojanju da se sustavno razviju zbijeni kodovi za složene prstenaste sustave. Ti su kodovi uspoređeni s IUPAC-ovom nomenklaturom za bicikličke spojeve, kao i s Wiswesserovom linearnom notacijom. Navedene su prednosti kompaktnih kodova, uključivši algebarske manipulacije kodovima radi izdvajanja kodova za strukturno srodne spojeve.