

Topological Organic Chemistry. 3. Graph Theory, Binary and Decimal Adjacency Matrices, and Topological Indices of Alkanes¹

HARRY P. SCHULTZ*

Department of Chemistry, University of Miami, Coral Gables, Florida 33124

TOR P. SCHULTZ

Forest Products Laboratory and Department of Chemistry, Mississippi State University, Mississippi State, Mississippi 39762

Received August 1, 1990

Adjacency matrices are used as a source of topological indices of alkanes by transformation from the binary to the decimal number system. The data are based upon a canonical numbering system of the graphs.

INTRODUCTION

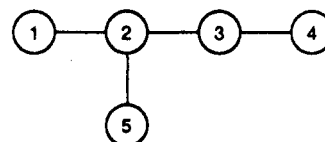
Graph theory serves as a means to numerically characterize chemical structures. Rouvray² has summarized numerous techniques for calculating topological indices, methods which continue to increase in number. This paper extends the study of the utility of structure-descriptive matrices as sources of single-sum, whole numbers attained by established matrix algebraic manipulations, and which may serve as codes and descriptors of alkanes. The values thus derived possess utility for demonstrating quantitative structural-physical property (QSPR) relationships. The topological indices derived from the adjacency (**A**), degree (**v**), and distance (**D**) matrices of alkanes in various combinations such as $v(\mathbf{A} + \mathbf{D})$ and the determinants of $(\mathbf{A} + \mathbf{D})$, both based upon the IUPAC numbering system and described in earlier papers of this series, appeared to be invariant, independent of the sequence of graph vertex numbers. Preliminary experimentation demonstrated, however, that the indices described in this paper and which are based upon binary decimal manipulations of the adjacency matrices varied as the sequence of numbering the vertices of the molecular graphs was altered.

COMPUTATIONS

This circumstance was countered by assembling data using a canonical method of numbering graph vertices based loosely on the Cahn-Ingold-Prelog³ priority rules, and in this paper is referred to as the CIP system of canonical numbering. This canonical mode of numbering graph vertices was quantitatively attained by applying the iterative techniques of Frazer et al.⁴ as adapted and elaborated by Herndon,⁵ who used the adjacency matrix of the graph from which to formulate canonical vertex numbers— $U_{K+1} = U_K + AU_K$, where U_K is a column vector, **A** the adjacency matrix, and U_{K+1} is the new column vector; U_1 has all elements equal to unity. The Frazer method differs only slightly in one step from the Herndon procedure, in that the Frazer process uses one less addition in its last iteration step; the n th Frazer iteration step is the $U_{K+1} = AU_K$ value (column AU_5 of the example), rather than the Herndon $U_{K+1} = U_K + AU_K$ sum (sum of columns U_5 and AU_5 of the example). The number of iterations required to establish apparent priorities among vertices of a graph varied, as displayed in column 8 of Table I, but in no instance did the vertex priorities change as a consequence of utilizing the Herndon iteration method in contrast to the Frazer iteration procedure. Nor did the number of iterations required to establish apparent priorities among the vertices of a graph alter from one iteration system to the other. Column 6 of Table I lists the Frazer n -iteration values and column 7 compares the corresponding Herndon iteration values. The Herndon (H) n -iteration numbers are about 50% larger than the Frazer (F) n -iteration

values. In order to surmount oscillations or inversions among priority data, n iterations, where n is the number of graph vertices, including the initial unity, were calculated for all graphs. In the instance of degenerate (equal) priorities of vertices, Herndon's third rule of numbering was invoked: "If there were two or more correct notations then that number which was lowest at the first point of difference was chosen". In a normal alkane this order of numbering was symmetrically maintained, for example, from the center of the graph on sequential, alternate, flanking vertices out to the terminal points of the graph. The iteration technique is illustrated with 2-methylbutane (7), which like all graphs described in this paper is hydrogen-suppressed with carbon-carbon bond distances of one; observe the degenerate priority values for vertices with CIP numbers three and four.

Structural (Molecular) Graph, IUPAC Numbered.

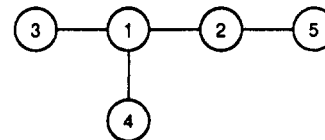


(**A**₂) Matrix, IUPAC Numbered; Five Iterations.

A ₂	U ₁	AU ₁	U ₂	AU ₂	U ₃	AU ₃	U ₄	AU ₄	U ₅	AU ₅	CIP no.
0 1 0 0 0	1	1	2	4	6	11	17	33	50	92	3
1 0 1 0 1	1	3	4	7	11	21	33	59	92	172	1
0 1 0 1 0	1	2	3	6	9	16	25	47	72	131	2
0 0 1 0 0	1	1	2	3	5	9	14	25	39	72	5
0 1 0 0 0	1	1	2	4	6	11	17	33	50	92	4

$\Sigma =$ 303 559
 $n =$ iteration index (F) 559
 $n =$ iteration index (H) 862

Structural (Molecular) Graph, CIP Numbered.



(**A**₂) Matrix, CIP Numbered.

A ₂
0 1 1 1 0
1 0 0 0 1
1 0 0 0 0
1 0 0 0 0
0 1 0 0 0

Table I. Topological Indices Derived from the Binary and Decimal Adjacency Matrices of Representative C₁–C₁₅ Alkanes

								(-1)ΔH _f ^o (g) (kJ/mol)			
no.	compound	(A ₁₀)	(v × A ₁₀)	(A ₁₀ × D)	n - (A ₂)	n - (A ₂)	iterations to priority definition	(v × A ₁₀) index		n-iterations	n-iterations
		index	index	index	iterations index (F)	iterations index (H)		obsd ^a	calcd ^b	index (F), calcd ^c	index (H), calcd ^d
1	methane	0	0	0				74.45			
2	ethane	3	3	3	4	8	1	83.45	60.0	74.9	78.0
3	propane	11	14	30	24	41	1	104.67	100.4	103.2	104.1
4	butane	27	42	132	110	178	1	125.66	129.3	127.3	127.5
5	2-methylpropane	31	45	141	132	208	1	134.19	131.1	130.2	130.0
6	pentane	59	106	437	448	707	2	146.77	153.6	149.5	149.5
7	2-methylbutane	71	116	500	559	862	2	153.68	156.0	153.0	152.7
8	2,2-dimethylpropane	79	124	508	728	1 093	1	167.95	157.8	157.2	156.5
9	hexane	123	234	1 181	1 712	2 668	2	167.03	174.5	170.7	170.8
10	3-methylpentane	151	274	1 463	2 216	3 364	2	172.1	178.6	174.8	174.5
11	2-methylpentane	151	252	1 538	2 068	3 158	3	174.8	176.4	173.7	173.5
12	2,2-dimethylbutane	175	298	1 596	3 168	4 704	2	186.1	180.8	180.5	179.8
13	2,3-dimethylbutane	159	285	1 497	2 592	3 848	1	178.3	179.6	177.3	176.6
14	heptane	251	490	3 464	5 099	8 486	3	187.7	193.9	188.0	189.2
15	2-methylhexane	311	524	4 407	7 714	11 723	3	194.6	195.6	194.6	194.4
16	3-methylhexane	311	586	4 032	8 317	12 540	3	191.3	198.6	195.7	195.5
17	2,2-dimethylpentane	353	632	4 353	11 040	16 313	5	205.9	200.6	200.2	199.7
18	3-ethylpentane	311	622	3 844	8 019	12 028	1	189.3	200.1	195.2	194.8
19	2,2,3-trimethylbutane	383	697	4 475	16 855	24 589	2	204.5	203.1	206.9	206.2
20	2,3-dimethylpentane	335	642	4 112	10 293	15 309	4	198.0	201.0	199.1	198.7
21	3,3-dimethylpentane	367	678	4 353	11 992	17 611	2	201.2	202.4	201.5	200.9
22	2,4-dimethylpentane	335	525	4 549	9 112	13 666	3	201.7	195.7	197.2	196.8
23	octane	507	1 002	8 970	22 304	34 222	3	208.7	212.7	211.4	211.5
24	4-methylheptane	631	1 110	10 413	29 610	44 458	2	212.0	215.4	215.8	215.7
25	3-methylheptane	625	1 190	10 808	28 921	43 515	8	212.5	217.2	215.5	215.3
26	2,2-dimethylhexane	737	1 272	12 239	39 222	57 918	4	224.6	218.9	220.3	219.9
27	2,2,3,3-tetramethylbutane	831	1 596	11 766	61 544	88 270	1	225.5	224.9	227.4	226.6
28	3-ethyl-2-methylpentane	687	1 450	10 377	38 996	57 560	2	211.0	222.4	220.2	219.8
29	3,3-dimethylhexane	739	1 426	11 207	44 234	64 798	3	220.0	222.0	222.2	221.7
30	3,4-dimethylhexane	687	1 386	10 656	38 210	56 450	2	212.8	221.2	219.9	219.5
31	2,4-dimethylhexane	547	1 236	10 507	38 404	56 696	4	212.6	218.2	219.9	219.6
32	2,5-dimethylhexane	687	1 053	12 672	31 500	47 258	3	222.5	214.0	216.8	216.7
33	3-ethyl-3-methylpentane	751	1 502	10 692	48 106	70 088	2	214.9	223.3	223.5	223.0
34	nonane	1 019	2 026	22 407	76 760	117 263	4	228.7	231.2	230.9	231.2
35	decane	2 043	4 074	54 783	261 800	398 746	4	249.7	249.5	250.3	250.7
36	undecane	4 091	8 170	131 186	880 984	1 338 779	5	270.9	267.8	269.5	270.1
37	dodecane	8 187	16 362	309 984	2 932 410	4 448 336	5	289.7	286.1	288.5	289.3
38	tridecane	16 379	32 746	722 249	9 672 836	14 365 889	6	311.5	304.3	307.4	308.0
39	tetradecane	32 763	65 514	1 666 887	31 664 680	47 911 604	6	332.1	322.5	326.2	327.2
40	pentadecane	65 531	131 050	3 807 756	104 067 884	157 337 449	7	352.8	340.8	345.0	346.2
41	2,3-dimethyl-6-ethyl-5-isopropyloctane ^e	94 567	233 850	3 842 363	430 246 510	620 543 361	11		356.0	367.5	368.1
42	2,6-dimethyl-5-ethyl-3-isopropyloctane ^e	93 567	233 210	3 952 864	431 430 641	621 616 202	15		355.9	367.5	368.2

^aClark, T.; McKervey, M. A. In *Comprehensive Organic Chemistry*; Stoddart, J. F., Ed.; Pergamon: Oxford, 1979; Vol. 1, p 69. The values for compounds 38, 39, and 40 are listed as estimated by Stull, D. R.; Westrum, E. F., Jr.; Sinke, E. F. *The Chemical Thermodynamics of Organic Compounds*; Wiley: New York, 1969. ^bCalculated by using $(-1)\Delta H_f^o(g)$ kJ/mol = $31.084 + 60.512 (\log \times [v \times A_{10}])$ index). ^cCalculated by using $(-1)\Delta H_f^o(g)$ kJ/mol = $52.948 + 36.4276 (\log [n - (A_2)])$ F index). ^dCalculated by using: $(-1)\Delta H_f^o(g)$ kJ/mol = $44.749 + 36.7775 (\log [n - (A_2)])$ H index). ^eSuggested by Hansen and Jurs.⁶

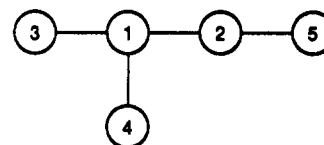
Inspection and intuition confirmed the CIP canonical numbers assigned to all graph points of all compounds except those of compounds 17, 22, 25, 29, 31, 41, and 42. In each of these instances, an interior secondary vertex adjacent to or flanked by quaternary or tertiary vertices sometimes possessed a higher priority number than expected; in some instances higher even than possessed by a quaternary or tertiary vertex. These circumstances, however unexpected and atypical, created no computational problems but seemed to designate such rare kinds of secondary vertices as possessing most unusual priorities. Examples of matrix manipulations pictured subsequently in this paper use only the CIP canonical numbering system.

The use of the adjacency matrix as the basis from which to calculate a molecular topological index was attained by applying its definition as it reads in Hansen and Jurs's review.⁶ That the adjacency matrix is a binary system is affirmed by Trinajstić⁷ and Randić,⁸ who utilized a binary linear representation of the adjacency matrix, combining separate vectors, top to bottom. Kier and Hall⁹ illustrated the binary and

decimal equivalents of an adjacency matrix.

A specific procedure which subjects the adjacency matrix of 2-methylbutane (7) to binary decimal transformations for the derivation of topological indices is illustrated.

Structural (Molecular) Graph, CIP Numbered.



Binary Adjacency (A₂) Matrix.

$$A_2 = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

Decimal Adjacency (A₁₀) Matrix.

$$A_2 = \begin{bmatrix} 14 \\ 17 \\ 16 \\ 16 \\ 8 \end{bmatrix} = \{14 \ 17 \ 16 \ 16 \ 8\}$$

Generation of Decimal Numbers of the (A_2) Matrix.

$$\Sigma = 2m^{n-1} + 2m^{n-2} + \dots 2m^0$$

Sum of Elements, Decimal Adjacency Vector.

$$(A_{10}) = \{14 \ 17 \ 16 \ 16 \ 8\} = 71 = \text{index}$$

Sum of Elements, Matrix Product of Degree Vector, and Decimal Adjacency Vector.

$$\begin{aligned} (v \times A_{10}) &= (v) \times (A_{10}) \\ &= [3 \ 2 \ 1 \ 1 \ 1] \times \{14 \ 17 \ 16 \ 16 \ 8\} \\ &= [42 \ 34 \ 16 \ 16 \ 8] = 116 = \text{index} \end{aligned}$$

Sum of Elements, Matrix Product of Decimal Adjacency Vector, and Distance Matrix.

$$\begin{aligned} (A_{10} \times D) &= (A_{10}) \times (D) \\ &= [14 \ 17 \ 16 \ 16 \ 8] \times \begin{bmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 2 & 2 & 1 \\ 1 & 2 & 0 & 2 & 3 \\ 1 & 2 & 2 & 0 & 3 \\ 2 & 1 & 3 & 3 & 0 \end{bmatrix} \\ &= [65 \ 86 \ 104 \ 104 \ 141] = 500 = \text{index} \end{aligned}$$

The binary adjacency matrix (A_2) is transformed into its decimal adjacency column vector (A_{10}), which to save space may also be represented as a row vector within braces rather than a column vector within brackets. The generation of decimal numbers from the binary notation of the adjacency matrix is accomplished for each row by utilizing the depicted expression from right to left, where n is the number of elements in each row of the binary adjacency matrix and m is either 0 or 1, an element of the binary adjacency matrix itself. The manipulation can also be applied to the binary adjacency matrix's columns; summing from bottom to top yields a row vector of the same elements. Summing of the elements of the decimal adjacency vector yields a molecular topological index.

The decimal adjacency column vector (postmultiplier) is conformable with the degree vector (premultiplier) for matrix multiplication; summation of the product vector elements yields the $(v \times A_{10})$ topological index.

In a parallel fashion, the decimal adjacency row vector (premultiplier) is conformable for multiplication with the distance matrix (postmultiplier) to give the $(A_{10} \times D)$ molecular topological index.

The (A_{10}) , $(v \times A_{10})$, and $(A_{10} \times D)$ indices of Table I all increase as the apparent complexities of the molecules increase within a family of isomers. Bertz¹⁰ regards this circumstance as one of the minimum criteria for topological indices.

As work progressed on the various topological indices derived from the decimal adjacency vectors, it became apparent that the sums of the iterative priority numbers derived from the binary adjacency matrices for canonical numbering of the various graphs also had the potential of serving as molecular topological indices, for all were single-sum, apparently unique, and constituted a monotonic series of numbers. The comparisons between iterative indices had perforce to be based upon sums derived after executing an equivalent (n) number (where n is the number of vertices in a graph) of iterations for each graph—a cumbersome goal to attain. Indices based upon n -iterative sums possessed the advantage of invariancy,

not exhibited by the prior (A_{10}) , $(v \times A_{10})$, or $(A_{10} \times D)$ indices described above. Additionally, the n -iterations indices also adhered to the Bertz minimum criteria for topological indices, in that the parent alkanes of a series of structural isomers had the lowest n -iterative index sums, with values increasing as the structural isomers increased in apparent intricacy.

RESULTS

Column 3 of Table I compares the results of computing the (A_{10}) vector index numbers of representative alkanes. The data confirm that the transformation of the base-2 adjacency (A_2) matrix into its base-10 adjacency (A_{10}) vector, followed by summation of the vector elements, gives the potential to calculate a topological index number for each alkane. Such (A_{10}) index numbers, however, were not unique: for as seen in Table I, compounds **10** and **11** (151); **15**, **16**, and **18** (311); **20** and **22** (335); and **28**, **30**, and **32** (687) all possessed replicate index values.

Further experiments were tried to ascertain whether the (A_{10}) vector possessed any other utility as a source of unique topological indices. Each degree vector was multiplied by the corresponding (A_{10}) vector, followed by summation of the elements of the vector product, as outlined above under Computations. As recorded in column 4 of Table I, the consequent $(v \times A_{10})$ index values were different, and continually increased in value through the series, as well as through each family of isomers. There appeared to be a regular progression of $2n + 22$ units in the $(v \times A_{10})$ series of straight-chain alkanes (where n is the index number of the prior normal alkane), except for the first three members (**2**, **3**, and **4**) of the alkane series. The regular progression of the $(v \times A_{10})$ index numbers of the normal alkanes afforded a possible tool with which to quantify relative intricacies among isomers of different carbon atom content. For example, the index ratios of pentane/2-methylbutane ($106/116 = 0.913$) and hexane/2-methylpentane ($234/252 = 0.929$) indicated that by the standards of these $(v \times A_{10})$ indices the two branched-chain isomers were of similar structure. In sharp contrast, if the above ratios are compared to the heptane/2,2,3-trimethylbutane ($490/697 = 0.703$) ratio, it is seen that of the branched-chain isomers compared above, the latter is much more intricate.

As shown in Table I, column 5, the $(A_{10} \times D)$ indices also continually increase in value through the series, as well as through each family of isomers, with one duplicative set of values—those of **17** and **21** (4353).

As seen in columns 6 and 7 of Table I, the n -iterative Frazer and Herndon indices continually increase in value through the series, as well as through each family of isomers; all appear to be unique, single-sum, and members of a monotonic sequence of values. If the n -iterative indices are used as a basis for the comparisons of molecular complexities, the isomers of heptane (**14**) become more intricate in the increasing order **14** < **15** < **18** < **16** < **22** < **20** < **17** < **21** < **19**, with 2,2,3-trimethylbutane (**19**) being the most complex molecule of the heptane isomers. Although intuitively reasonable, the above sequence differs somewhat from the indices listed in columns 3, 4, and 5 of Table I, as well as differing from other modes of deriving molecular topological indices which are described in earlier papers of this series.

EXAMPLE APPLICATION

The utility of the unique topological indices—the $(v \times A_{10})$ indices, which depended upon the use of the canonical CIP numbering system, and the invariant n -iterative indices—was tested with the use of heats of formation data. Column 9 of Table I presents the observed heats of formation data for all the compounds listed; columns 10, 11, and 12 record the values

for heats of formation derived from utilization of ($v \times A_{10}$) indices and Frazer and Herndon n -iterative indices, respectively. Methane excluded, the ($v \times A_{10}$) index values displayed a linear relationship described by the equation:

$$(-1)\Delta H_f^\circ(g) \text{ kJ/mol} = 31.084 + 60.512 (\log [v \times A_{10}] \text{ index})$$

[with an r^2 value of 0.983 and F test equal to 2136. The Frazer n -iterative (F) index values displayed a linear relationship described by the equation:

$$(-1)\Delta H_f^\circ(g) \text{ kJ/mol} = 52.948 + 36.4276 (\log [n - (A_2)] \text{ F index})$$

[with an r^2 value of 0.992 and F test equal to 4869, methane excluded. The best correlation by a slight margin was exhibited by the Herndon n -iterative (H) index values with a linear relationship described by the equation:

$$(-1)\Delta H_f^\circ(g) \text{ kJ/mol} = 44.749 + 36.7775 (\log [n - (A_2)] \text{ H index})$$

[with an r^2 value of 0.993 and F test equal to 5341, methane excluded. The above relatively simple one-parameter modes of estimating heats of formation compare reasonably well with the seven structural parameters utilized by Kalb et al.¹¹ to approximate such alkane heats of formation to within ± 1.5 kJ/mol.

CONCLUSION

Transformation of the base-2 adjacency matrix of an alkane to the base-10 vector enabled a series of matrix computations

to be performed that led to several kinds of topological indices for each alkane. One of these, the index calculated by summing the vector elements of the product of the degree vector and the base-10 adjacency vector, gave an index that appeared to be unique, single-sum, and with other indices formed a monotonic series. The above indices were not invariant; hence, a canonical numbering system was used for each graph. There was additionally derived from the iterative procedures that served as the source of the canonical graph numbers a series of invariant molecular topological indices, which appeared to be unique, single-sum, and constituted monotonic series of values.

REFERENCES AND NOTES

- (1) Paper 2 of this series: Schultz, H. P.; Schultz, E. B.; Schultz, T. P. *J. Chem. Inf. Comput. Sci.* **1990**, *30*, 27-29.
- (2) Rouvray, D. H. In *Chemical Applications of Topology and Graph Theory*; King, R. B., Ed.; Elsevier: Amsterdam, 1983; pp 159-177.
- (3) Cahn, R. S.; Ingold, C.; Prelog, V. *Angew. Chem., Int. Ed. Engl.* **1965**, *385-415*.
- (4) Frazer, R. A.; Duncan, W. J.; Collar, A. R. *Elementary Matrices*; Macmillan: New York, 1946; p 142.
- (5) Herndon, W. C. In *Chemical Applications of Topology and Graph Theory*; King, R. B., Ed.; Elsevier: Amsterdam, 1983; pp 231-242.
- (6) Hansen, P. J.; Jurs, P. C. *J. Chem. Educ.* **1988**, *65*, 574-580.
- (7) Trinajstić, N. *Chemical Graph Theory*; CRC: Boca Raton, FL, 1983; Vol. 1, pp 31-41.
- (8) Randić, M. *J. Chem. Phys.* **1974**, *60*, 3920-3928; **1975**, *62*, 309-310. Randić, M. *J. Chem. Inf. Comput. Sci.* **1977**, *17*, 171-180.
- (9) Kier, L. B.; Hall, L. H. *Molecular Connectivity in Chemistry and Drug Research*; Academic: New York, 1976; p 32.
- (10) Bertz, S. H. In *Chemical Applications of Topology and Graph Theory*; King, R. B., Ed.; Elsevier: Amsterdam, 1983; p 208.
- (11) Kalb, A. J.; Chung, A. L. H.; Allen, T. L. *J. Am. Chem. Soc.* **1966**, *88*, 2938-2942.

Chemical Abstracts Online: A Study of the Quality of Controlled Terms

SABINE MARTIN and GÜNTER BERGERHOFF*

Institute for Inorganic Chemistry, University of Bonn, Gerhard-Domagk-Strasse 1, D-5300 Bonn 1, FRG

Received August 27, 1990

Controlled vocabulary is said to be the key to facts covered in the vast amount of literature. An analysis of a certain fraction of CAS controlled vocabulary from the CA File shows a number of errors and inconsistencies. This leads to conclusions for the retrieval process, and more use of the basic index is recommended until the controlled vocabulary has been revised.

INTRODUCTION

CAS ONLINE, being the largest database in chemistry worldwide, demands a high degree of familiarity with its rules of indexing. This leads to unsatisfactory results, especially when used by the bench chemist. As the chemist himself is most familiar with his research subject, the classical documentalist expects that the scientist himself would search the literature for the solution to his problems. Therefore, modern documentation systems should be conceived such that their undeniable advantages in rapidity and flexibility are not offset by overly complicated user instructions.

The search for chemical compounds is a problem of general importance in chemical databases. It has been dealt with in various ways. Whether it is the introduction of registry numbers (the concise and unambiguous representation of substances in most cases), the description of structural characteristics by means of easily-machine-readable topological

codes, or the possibility of graphical entry of the structures in question—the effort always aims to make the search as easy and successful as possible (see STN Express, MOLKICK).

But CAS ONLINE users have to cope with difficulties in fact-searching. The problems may be of a formal nature (typographic errors, parameters for headings, length of terms, use of special characters, admissibility of abbreviations) or of a logical nature (homonyms, synonyms, combination of terms, etc.). Right from the beginning CA has provided a controlled vocabulary (CV) for assistance especially in this type of search. The CV applies also to the database version. However, the CV used from the 9th Collective Index (9CI) on, beginning with 1972, differs somewhat from that for the 8CI (1967-1971) and differs clearly from that in the CA File.

The controlled vocabulary is said to be the key to the solution of the problems cited above. Looking for possible improvements we began with an examination of the CV in the