Extended Adjacency Matrix Indices and Their Applications

Yi-Qiu Yang, Lu Xu,* and Chang-Yu Hu

Applied Spectroscopy Laboratory, Changchun Institute of Applied Chemistry, Academia Sinica, Changchun 130022, Jilin, People's Republic of China

Received March 30, 1994®

In this paper, new topological indices, EAΣ and EAmax, are introduced. They are based on the extended adjacency matrices of molecules, in which the influences of factors of heteroatoms and multiple bonds were considered. The results show that EAΣ and EAmax possess high discriminating power and correlate well with a number of physicochemical properties and biological activities of organic compounds.

INTRODUCTION

The structure of a molecule—geometric and electronic—must contain the features responsible for its physical and chemical properties. The simplest way to represent a molecule's structure is to assign to the structure a number or a set of numbers, termed indices; then, the indices are applied to the study of the correlations with properties ranging from physical to biological. In recent years, one type of these methods seems to hold good promise for the quantitative structure–property relationship (QSPR) and the quantitative structure–activity relationship (QSAR) studies.

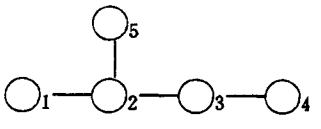
The topological indices are actually graph invariants. They are obtained from the chemical graph (hydrogen-suppressed graph). Uniqueness and correlation are two of the most important requirements for the topological indices. The task of defining an index that could have different, unique, but structurally significant, values for different structures seems to be very difficult. Therefore, more than 100 topological indices are in existence, such as the Wiener index, W,[1] the Hosoya index, Z,[2] the Randić index χ,[3] the Balaban index, J,[4] and the extended Randić index by Kier and Hall.[5] In this paper, the new indices, EA's proposed by us, show lower degeneracy and good performance in correlation.

METHODS

Recently, Randić discussed the strategies for searching optimal molecular descriptors.[6,7] As we know already, most indices to date have been based on two particular matrices: the distance and adjacency matrices. In this paper, the adjacency matrix and its extended form were used to deduce the new topological indices. The entries in the adjacency matrix are symbolized as a_{ij} and are equal to either one or zero, depending respectively on whether or not the vertices are connected:

$$a_{ij} = \begin{cases} 1 & \text{for adjacent vertices} \\ 0 & \text{otherwise} \end{cases}$$

For example the hydrogen-suppressed graph of 2-methylbutane is



and its adjacency matrix takes the form

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

We also introduce the degree vector $\mathbf{V}$, $\mathbf{V} = \{v_i\}$, where v_i

Table 1. Electronegativities of Common Atoms

H	2.1	C	2.5
N	3.0	O	3.5
F	4.0	P	2.1
S	2.5	Cl	3.0
Br	2.8	I	2.5

Table 2. Values of ID, EAΣ, and EAmax of 20 Structures

no.	ID	EAΣ	EAmax
141	18.037 9058	14.798 27	3.337 117
142	24.315 9362	19.659 51	3.337 117
143	15.007 6026	10.461 11	2.936 060
144	19.205 2889	13.375 89	2.653 472
145	19.212 7979	13.119 84	2.672 798
146	19.386 9551	14.105 44	2.829 541
147	26.868 1152	17.446 43	2.829 542
148	7.777 7778	6.000 000	3.000 000
149	17.991 7696	12.000 00	3.000 000
150	20.263 3	13.298 79	2.951 932
151	22.417 2175	14.285 75	2.914 090
152	22.524 8849	14.713 33	2.906 148
153	22.457 3998	15.517 54	3.000 000
154	17.676 2689	12.292 53	3.000 000
155	22.100 2019	15.307 58	2.914 090
156	21.417 0255	14.499 62	2.653 614
157	31.210 7297	21.656 12	2.777 381
158	21.392 4287	15.121 58	2.620 189
159	21.178 7608	16.198 04	3.049 510
160	19.477 3775	12.139 82	3.005 782

Table 3. EA Indices for 18 Graphs

no.	EAΣ	EAmax	no.	EAΣ	EAmax
1	11.25963	2.878785	10	11.65685	3.000000
2	11.49126	2.864037	11	11.73486	2.885716
3	12.21042	2.822746	12	11.73486	2.885716
4	11.55154	2.865975	13	11.32856	2.894902
5	11.55315	2.824196	14	11.32856	2.894902
6	11.57474	2.822746	15	12.47214	3.000000
7	11.42910	2.822746	16	12.47214	3.000000
8	11.64767	2.884355	17	12.00000	3.000000
9	10.68483	2.884355	18	12.00000	3.000000

* Abstract published in *Advance ACS Abstracts*, July 15, 1994.

0095-2338/94/1634-1140$04.50/0

Table 4. Topological Indices and Boiling Points for 149 Alkanes

compounds	W	Z	χ	J	EAΣ	EAmax	bp (°C
ethane	1	2	1.000 000	1.000 000	2.000 000	1.000 000	−88.6
propane	4	3	1.414 214	1.632 993	3.535 534	1.767 767	−42.1
butane	10	5	1.914 214	1.974 745	5.385 165	1.846 291	−0.50
2-methylpropane	9	4	1.732 051	2.323 790	5.773 502	2.886 751	−11.73
pentane	20	8	2.414 214	2.190 610	6.274 918	1.887 459	36.07
2-methylbutane	18	7	2.270 056	2.539 539	7.532 390	2.657 551	27.85
2,2-dimethylpropane	16	5	2.000 000	3.023 716	8.500 001	4.250 000	9.50
hexane	35	13	2.914 214	2.339 092	7.875 304	1.912 646	68.74
2-methylpentane	32	11	2.770 056	2.627 215	8.320 846	2.644 617	60.27
3-methylpentane	31	12	2.808 061	2.754 185	9.089 133	2.441 140	63.28
2,3-dimethylbutane	29	10	2.642 735	2.993 498	9.637 889	2.909 472	57.99
2,2-dimethylbutane	28	9	2.560 660	3.168 490	10.173 18	3.909 893	49.74
heptane	56	21	3.414 214	2.447 473	8.892 971	1.929 549	98.42
2-methylhexane	52	18	3.270 056	2.678 258	9.992 315	2.641 847	90.05
3-methylhexane	50	19	3.308 061	2.831 820	9.929 288	2.428 892	91.85
3-ethylpentane	48	20	3.346 066	2.992 303	9.509 251	2.254 625	93.48
2,4-dimethylpentane	48	15	3.125 898	2.953 223	10.336 42	2.811 188	80.50
2,3-dimethylpentane	46	17	3.180 739	3.144 208	11.169 28	2.786 288	89.78
2,2-dimethylpentane	46	14	3.060 660	3.154 490	10.929 39	3.902 909	79.20
3,3-dimethylpentane	44	16	3.121 320	3.360 435	11.715 59	3.549 463	86.03
2,2,3-trimethylbutane	42	13	2.943 376	3.541 197	12.253 66	3.905 577	80.88
octane	84	34	3.914 214	2.530 061	10.396 22	1.941 614	125.67
2-methylheptane	79	29	3.770 056	2.715 843	10.968 48	2.641 261	117.65
3-methylheptane	76	31	3.808 061	2.872 066	11.564 35	2.425 542	118.93
4-methylheptane	75	30	3.808 061	2.919 613	10.752 93	2.416 806	117.71
3-ethylhexane	72	32	3.846 066	3.074 373	11.045 75	2.246 364	118.54
2,5-dimethylhexane	74	25	3.625 898	2.927 820	12.116 62	2.722 244	109.10
2,4-dimethylhexane	71	26	3.663 903	3.098 828	11.960 89	2.724 521	109.43
2,3-dimethylhexane	70	27	3.680 739	3.180 819	12.014 14	2.780 567	115.61
3-ethyl-2-methylpentane	67	28	3.718 744	3.354 877	11.470 13	2.693 850	115.65
3,4-dimethylhexane	68	29	3.718 744	3.292 478	12.707 51	2.631 516	117.72
2,2-dimethylhexane	71	23	3.560 660	3.111 766	12.621 92	3.902 374	106.84
3,3-dimethylhexane	67	25	3.621 320	3.373 382	12.491 21	3.540 540	111.97
3-ethyl-3-methylpentane	64	28	3.681 981	3.583 212	13.019 51	3.172 473	118.25
2,3,4-trimethylpentane	65	24	3.553 418	3.464 227	13.244 01	2.919 351	113.46
2,2,4-trimethylpentane	66	19	3.416 502	3.388 924	12.934 07	3.914 430	99.23
2,2,3-trimethylpentane	63	22	3.481 381	3.623 281	13.776 70	3.877 758	109.84
2,3,3-trimethylpentane	62	23	3.504 036	3.708 324	13.777 61	3.568 858	114.76
2,2,3,3-trimethylbutane	58	17	3.250 000	4.020 391	14.857 66	4.214 415	106.47
nonane	120	55	4.414 214	2.595 083	11.477 15	1.950 614	150.8
2-methyloctane	114	47	4.270 056	2.746 691	12.504 77	2.641 138	142.8
3-methyloctane	110	50	4.308 061	2.876 623	12.561 53	2.424 473	143.5
4-methyloctane	108	49	4.308 061	2.954 823	12.400 64	2.413 346	142.4
4-ethylheptane	102	51	4.346 066	3.175 341	12.048 32	2.238 559	141.2
3-ethylheptane	104	52	4.346 066	3.092 246	12.103 02	2.243 288	143.0
2,6-dimethylheptane	108	40	4.125 898	2.914 659	13.035 63	2.679 947	135.2
2,5-dimethylheptane	104	43	4.163 903	3.060 821	13.685 03	2.664 994	136.0
2,3-dimethylheptane	102	44	4.180 739	3.155 280	13.647 04	2.779 527	140.5
2,4-dimethylheptane	102	41	4.163 903	3.151 251	12.779 65	2.721 230	133.5
3-ethyl-2-methylhexane	96	45	4.218 744	3.410 085	13.072 26	2.689 473	138.0
4-ethyl-2-methylhexane	98	44	4.201 908	3.307 394	13.155 79	2.680 721	133.8
3,5-dimethylheptane	100	45	4.201 908	3.223 047	13.576 83	2.580 399	136.0
3,4-dimethylheptane	98	46	4.218 744	3.324 760	13.551 02	2.623 622	140.1
3-ethyl-4-methylhexane	94	48	4.256 749	3.499 480	13.045 09	2.512 713	140.4
2,2-dimethylheptane	104	37	4.060 660	3.072 990	13.586 60	3.902 333	132.7
3,3-dimethylheptane	98	41	4.121 321	3.330 074	14.171 24	3.539 675	137.3
4,4-dimethylheptane	96	39	4.121 321	3.431 051	13.262 69	3.531 518	135.2
4,4-diethylpentane	88	48	4.242 640	3.824 684	13.090 17	2.795 085	146.2
3-ethyl-3-methylhexane	92	44	4.181 981	3.617 390	13.840 58	3.161 112	140.6
2,3,5-trimethylhexane	96	37	4.036 582	3.376 601	14.047 27	2.848 745	131.3
2,4-dimethyl-3-ethylpentane	90	39	4.091 423	3.677 616	13.398 14	2.828 666	136.73
2,3,4-trimethylhexane	92	41	4.091 423	3.575 834	14.783 69	2.842 428	139.0
2,2,5-trimethylhexane	98	32	3.916 502	3.280 711	14.748 63	3.903 268	124.0
2,2,4-trimethylhexane	94	33	3.954 507	3.467 262	14.564 33	3.910 517	126.5
2,2,3-trimethylhexane	92	35	3.981 381	3.588 734	14.622 88	3.877 049	131.7
2,2-dimethyl-3-ethylpentane	88	36	4.019 385	3.792 908	14.046 11	3.855 443	133.83
2,4,4-trimethylhexane	92	34	3.977 163	3.576 752	14.502 35	3.561 424	126.5
2,3,3-trimethylhexane	90	36	4.004 036	3.702 086	14.554 98	3.560 822	137.7
3,3,4-trimethylhexane	88	39	4.042 041	3.802 396	15.303 75	3.527 114	140.5
2,3-dimethyl-3-ethylpentane	86	40	4.064 696	3.919 211	15.050 23	3.239 333	141.6
2,2,3,4-tetramethylpentane	86	31	3.854 059	3.877 605	15.849 47	3.882 490	133.0
2,3,3,4-tetramethylpentane	84	33	3.886 752	4.013 737	15.835 85	3.585 167	141.5
2,2,4,4-tetramethylpentane	88	24	3.707 107	3.746 418	15.527 46	4.083 121	122.7
2,2,3,3-tetramethylpentane	82	30	3.810 660	4.144 726	16.375 59	4.054 641	140.3
decane	165	89	4.914 214	2.647 605	12.928 00	1.957 555	174.2
2-methylnonane	158	76	4.770 056	2.773 189	13.563 13	2.641 112	167
3-methylnonane	153	81	4.808 061	2.886 163	14.081 01	2.424 202	167.8

Table 4 (Continued)

compounds	W	Z	χ	J	EAΣ	EAmax	bp (°C
4-methylnonane	150	79	4.808 061	2.968 015	13.390 22	2.412 358	165.7
5-methylnonane	149	80	4.808 061	2.998 419	14.037 90	2.409 867	165.1
4-isopropylheptane	138	81	4.846 066	3.295 082	12.967 97	2.231 204	162.0
4-ethyloctane	141	83	4.846 066	3.205 535	13.560 17	2.235 614	163.64
3-ethyloctane	145	84	4.846 066	3.086 901	13.574 78	2.242 122	166.0
2,7-dimethyloctane	151	65	4.625 898	2.909 472	14.615 88	2.659 508	159.87
2,6-dimethyloctane	146	69	4.663 903	3.033 297	14.633 00	2.646 812	158.54
2,3-dimethyloctane	143	71	4.680 739	3.129 602	14.645 22	2.779 339	164.31
2,5-dimethyloctane	143	68	4.663 903	3.124 402	14.522 39	2.663 637	156.8
2,4-dimethyloctane	142	67	4.663 903	3.160 036	14.431 02	2.720 603	153.0
4-isopropylheptane	131	72	4.718 744	3.499 857	14.091 77	2.685 158	160.0
4-ethyl-2-methylheptane	134	70	4.701 908	3.390 786	14.119 81	2.679 002	160.0
3-ethyl-2-methylheptane	134	73	4.718 744	3.397 790	14.089 72	2.688 596	166.0
5-ethyl-2-methylheptane	138	72	4.701 908	3.255 530	14.185 72	2.650 472	159.7
3,6-dimethyloctane	141	74	4.701 908	3.168 174	15.255 24	2.508 182	160.0
3,4-dimethyloctane	137	75	4.718 744	3.308 840	15.184 55	2.621 908	166.0
3,5-dimethyloctane	138	71	4.701 908	3.268 555	14.398 18	2.573 861	160.0
4,5-dimethyloctane	135	73	4.718 744	3.375 929	14.394 87	2.615 580	162.1
4-ethyl-3-methylheptane	129	77	4.756 750	3.563 688	14.625 80	2.506 718	167.0
5-ethyl-3-methylheptane	133	76	4.739 913	3.412 256	14.731 24	2.496 523	158.3
3-ethyl-4-methylheptane	130	76	4.756 749	3.529 931	13.877 44	2.502 991	167.0
3,4-diethylhexane	125	80	4.794 754	3.698 220	13.992 08	2.371 022	162.0
2,2-dimethyloctane	146	60	4.560 660	3.043 758	15.131 77	3.902 330	155.0
3,3-dimethyloctane	138	66	4.621 321	3.276 962	15.142 73	3.539 593	161.2
4,4-dimethyloctane	134	64	4.621 321	3.417 512	14.945 76	3.530 638	157.5
3,3-diethylhexane	121	76	4.742 640	3.874 775	14.584 04	2.781 714	166.3
4-ethyl-4-methylheptane	126	69	4.681 981	3.690 295	14.946 70	3.149 601	167.0
3-ethyl-3-methylheptane	129	72	4.681 981	3.575 505	15.489 28	3.159 644	163.8
2,3,6-trimethylheptane	136	61	4.536 582	3.301 403	15.767 26	2.797 082	155.7
2,4,6-trimethylheptane	135	56	4.519 745	3.337 430	14.804 87	2.787 556	144.8
3-isopropyl-2-methylhexane	124	63	4.591 423	3.728 003	15.084 90	2.824 479	163.0
2,5-dimethyl-3-ethylhexane	127	62	4.574 586	3.603 336	15.189 79	2.772 741	157.0
2,3,5-trimethylheptane	131	64	4.574 586	3.461 674	15.662 41	2.808 731	157.0
2,4,5-trimethylheptane	130	63	4.574 586	3.502 718	15.583 71	2.768 814	157.0
2,3,4-trimethylheptane	128	65	4.591 423	3.583 327	15.626 95	2.839 443	163.0
2,4-dimethyl-3-ethylhexane	122	67	4.629 428	3.797 908	14.983 69	2.748 271	164.0
2,3-dimethyl-4-ethylhexane	123	68	4.629 428	3.756 108	15.125 35	2.797 341	164.0
3,4,5-trimethylheptane	125	70	4.629 427	3.685 406	16.322 97	2.729 738	164.0
2,2,6-trimethylheptane	139	51	4.416 502	3.205 456	15.651 22	3.902 401	148.2
2,2,5-trimethylheptane	134	55	4.454 507	3.355 508	16.315 85	3.902 960	148.0
2,2,3-trimethylheptane	130	57	4.481 381	3.518 426	16.255 24	3.876 995	158.0
2,2,4-trimethylheptane	131	52	4.454 507	3.469 465	15.381 43	3.910 422	147.7
2,2-dimethyl-3-ethylhexane	122	58	4.519 386	3.808 926	15.663 29	3.854 864	159.0
2,2-dimethyl-4-ethylhexane	126	56	4.492 512	3.630 845	15.783 09	3.907 570	147.0
2,5,5-trimethylheptane	131	57	4.477 163	3.464 727	16.296 56	3.541 770	152.8
2,3,3-trimethylheptane	127	59	4.504 036	3.633 392	16.234 34	3.560 056	160.1
2,4,4-trimethylheptane	127	53	4.477 163	3.625 600	15.272 46	3.552 846	153.0
3,3,5-trimethylheptane	126	59	4.515 168	3.641 863	16.129 13	3.553 431	155.68
3,3,4-trimethylheptane	123	62	4.542 041	3.778 372	16.149 45	3.525 855	164.0
3,4,4-trimethylheptane	122	61	4.542 041	3.823 180	16.080 77	3.518 440	164.0
3,3-dimethyl-4-ethylhexane	118	64	4.580 046	3.971 137	15.585 07	3.494 951	165.0
3,3-diethyl-2-methylhexane	114	68	4.625 356	4.153 513	14.992 00	2.951 106	174.0
2,3-dimethyl-3-ethylhexane	119	63	4.564 696	3.943 570	15.875 25	3.230 296	169.0
2,4-dimethyl-4-ethylhexane	122	60	4.537 823	3.802 578	15.866 17	3.206 202	158.0
3,4-dimethyl-3-ethylhexane	117	68	4.602 701	4.020 494	16.582 03	3.172 749	170.0
2,4-dimethyl-3-isopropylpentane	117	54	4.464 102	3.983 500	15.278 07	2.924 988	157.04
2,3,4,5-tetramethylhexane	121	58	4.464 102	3.813 995	16.860 020	2.924 877	161.0
2,2,4,5-tetramethylhexane	124	47	4.327 186	3.684 242	16.651 36	3.911 134	148.2
2,2,3,5-tetramethylhexane	123	48	4.337 223	3.734 823	16.656 44	3.878 260	148.4
2,2,4-trimethyl-3-ethylpentane	115	50	4.392 064	4.072 918	15.964 57	3.859 398	155.3
2,2,3,4-tetramethylhexane	118	53	4.392 064	3.941 813	17.389 66	3.878 834	154.9
2,2,3,5-tetramethylhexane	120	49	4.359 879	3.865 559	16.566 71	3.579 268	153.0
2,3,4,4-tetramethylhexane	116	55	4.414 720	4.034 119	17.377 24	3.539 185	162.2
2,3,3,4-tetramethylhexane	115	56	4.424 757	4.089 289	17.362 61	3.547 729	164.59
2,3,4-trimethyl-3-ethylpentane	112	57	4.447 412	4.228 989	17.074 37	3.287 522	169.44
2,2,5,5-tetramethylhexane	127	41	4.207 107	3.563 001	17.381 50	3.954 520	137.46
2,2,4,4-tetramethylhexane	119	43	4.267 767	3.887 595	17.098 26	3.976 226	153.3
2,2,3,3-tetramethylhexane	115	47	4.310 660	4.101 784	17.153 43	4.052 231	158.0
2,2,3-trimethyl-3-ethylpentane	110	52	4.371 320	4.328 342	17.638 52	3.949 285	168.0
3,3,4,4-tetramethylhexane	111	53	4.371 320	4.281 757	17.894 99	3.821 292	170.5
2,2,3,4,4-pentamethylpentane	111	40	4.154 701	4.231 133	18.454 19	4.020 850	159.29
2,2,3,3,4-pentamethylpentane	108	43	4.193 376	4.403 818	18.432 52	4.050 615	166.05

is the degree of the ith vertex, and the **V** of 2-methylbutane is

$$\mathbf{V} = \{1,3,2,1,1\}$$

The extended adjacency matrix (EA matrix) is represented as follows:

$$\mathrm{EA} = \{g_{ij}\}$$

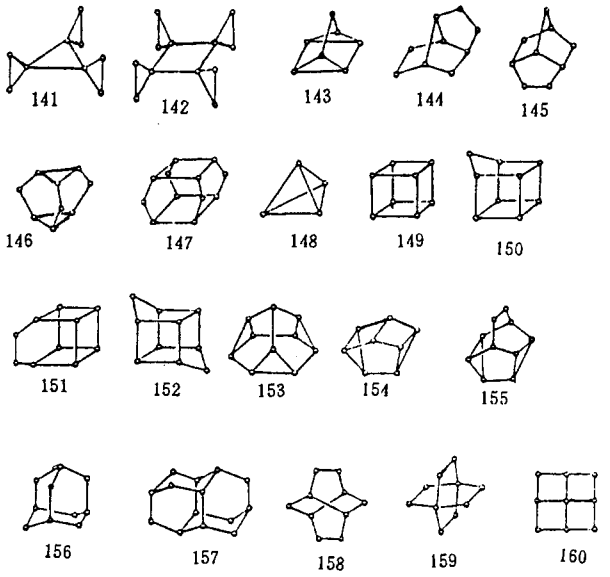


Figure 1. Skeletons of organic compounds showing different ring structures including unusual ring systems.

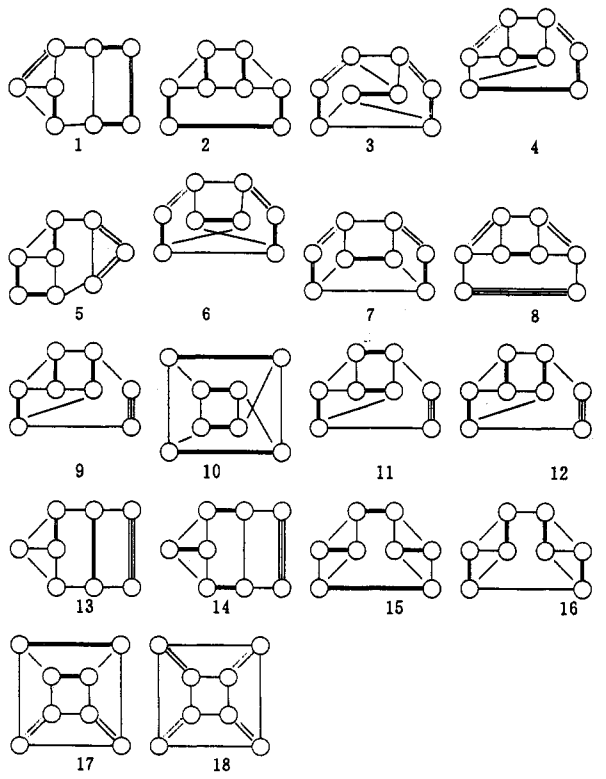


Figure 2. Skeletons of partial cyclic graphs with eight vertices and the degree of each vertex being 4.

whereas

$$g_{ij} = a_{ij}\frac{v_i/v_j + v_j/v_i}{2} \tag{1}$$

where g_{ij} is an element of the EA matrix and a_{ij} denotes an element of matrix **A**. So, the extended adjacency matrix of 2-methylbutane is

$$\mathrm{EA} = \begin{bmatrix} 0 & 1.67 & 0 & 0 & 0 \\ 1.67 & 0 & 1.08 & 1.67 & 0 \\ 0 & 1.08 & 0 & 0 & 1.25 \\ 0 & 0 & 1.25 & 0 & 0 \\ 0 & 1.67 & 0 & 0 & 0 \end{bmatrix}$$

Then, the eigenvalues of the EA matrix can be calculated, and the new topological indices—EAΣ and EAmax—will be constructed. EAΣ is the sum of the absolute eigenvalues of the EA matrix, and EAmax is the maximum of the absolute eigenvalues of the EA matrix. EAΣ and EAmax are together called the EA indices. We also take 2-methylbutane as an example for which the eigenvalues of its EA matrix are

$$X_1 = -2.657\,551 \quad X_2 = 2.657\,550 \quad X_3 = 3.318\,116 \times 10^{-8}$$
$$X_4 = 1.108\,644 \quad X_5 = -1.108\,644$$

EAΣ and EAmax are calculated respectively on the basis of the definitions

$$\mathrm{EA}\Sigma = |X_1| + |X_2| + |X_3| + |X_4| + |X_5| = 7.532390$$

$$\mathrm{EAmax} = \max(|X_i|) = 2.657551$$

According to the preceding discussion, the hydrogen-suppressed graph is a simple graph, and neither heteroatoms nor multiple bonds were considered. Thus, the indices derived from these graphs could not be expected to possessing high selectivity; for example, 2-methylbutane, 2-butanol, and 3-methyl-1-butene have the same chemical graphs. Of course, the EA indices are the same. In order to increase the discriminating power of the new indices, two factors, heteroatoms and multiple bonds, embraced within species should be subsequently taken into account.

1. Heteroatoms. The electronegativity is the quantitative property of an atom, which is the power of an atom in a molecule to attract electrons to itself and is concerned with atoms in molecules. Therefore, we introduced the electronegativities in our topological indices so as to reflect the influences of heteroatoms. Table 1 shows the electronegativities of common atoms.

If a compound contains heteroatoms, the degree vector and the elements of the EA matrix will be changed as follows:

$$v_i^h = v_i \mathrm{ELC}_i \tag{2}$$

$$g_{ii}^h = \mathrm{ELC}_i \tag{3}$$

where ELC_i is the electronegativity of atom i; v_i^h is used instead of v_i and g_{ii}^h is the new g_{ii}.

2. Multiple Bonds. The further extension is to enable improved differentiation to better characterize structures containing double and triple bonds.

$$v_i^b = v_i^h, \quad b_{ij} = 1$$

$$v_i^b = v_i^h + {}^1/_2, \quad v_j^b = v_j^h + {}^1/_2, \quad b_{ij} = 2$$

$$v_i^b = v_i^h + {}^1/_3, \quad v_j^b = v_j^h + {}^1/_3, \quad b_{ij} = 3$$

where b_{ij} is the bond type to connect atoms i and j and v_i^b is

Table 5. Results of Relationships of 149 Structures

indices	eq	R	F	S
W	$\mathrm{bp/°C} = 149.99 + 0.14W - 1.19W^{1/2}$	0.9844	2281.71	12.08
Z	$\mathrm{bp/°C} = 138.23 + 0.55Z - 2.42Z^{1/2}$	0.9710	1202.20	16.81
χ	$\mathrm{bp/°C} = -5.13 - 3.14\chi + 8.65\chi^2$	0.9879	2951.82	10.00
J	$\mathrm{bp/°C} = 78.59 - 36.91J + 15.51J^2$	0.9101	352.29	27.81
EA	$\mathrm{bp/°C} = 134.79 + 3.45\mathrm{EA}\Sigma - 10.43\mathrm{EAmax}$	0.9914	4178.34	9.46

Table 6. Two Topological Indices and Boiling Points of 37 Alcohols

no.	compound	χ	$^1\chi^v$	EAΣ	EAmax	bp (°C)
1	methanol	1.000 000	0.447 214	6.000 000	4.169 124	64.7
2	ethanol	1.414 214	1.023 335	8.500 000	4.466 509	78.3
3	1-propanol	1.914 214	1.523 335	11.000 00	4.484 931	97.2
4	2-propanol	1.732 051	1.412 899	11.209 88	5.358 974	82.3
5	1-butanol	2.414 213	2.023 335	13.500 00	4.494 513	117.7
6	2-butanol	2.270 056	1.950 904	13.500 00	5.149 825	99.6
7	2-methyl-1-propanol	2.270 56	1.879 177	13.748 93	5.169 988	107.9
8	2-methyl-2-propanol	2.000 000	1.723 607	16.399 06	6.616 247	82.4
9	1-pentanol	2.914 213	2.523 335	16.000 00	4.500 160	137.8
10	2-pentanol	2.770 055	2.450 904	16.000 00	5.138 161	119.0
11	3-pentaol	2.808 060	2.488 909	16.000 00	4.957 134	115.3
12	2-methyl-1-butanol	2.808 060	2.417 181	16.000 00	4.963 617	128.7
13	3-methyl-1-butanol	2.770 056	2.379 177	16.275 47	5.147 513	131.2
14	2-methyl-2-butanol	2.560 660	2.284 267	18.161 19	6.279 202	102.0
15	3-methyl-2-butanol	2.642 734	2.323 583	16.536 68	5.391 488	111.5
16	2,2-dimethyl-1-propanol	2.560 660	2.169 781	18.799 45	6.409 247	113.1
17	1-hexanol	3.414 213	3.023 335	18.500 00	4.503 749	157.7
18	2-hexanol	3.270 055	2.950 904	18.500 00	5.135 648	139.9
19	3-hexanol	3.808 060	2.988 909	18.500 00	4.946 299	135.4
20	2-methyl-1-pentanol	3.808 060	2.917 181	18.500 00	4.952 079	148.0
21	3-methyl-1-pentanol	3.808 060	2.917 181	18.500 00	4.935 991	152.4
22	4-methyl-1-pentanol	3.270 055	2.879 176	18.780 81	5.142 475	151.8
23	2-methyl-2-pentanol	3.060 660	2.784 266	20.643 11	6.271 849	121.4
24	3-methyl-2-pentanol	3.180 739	2.861 588	18.699 99	5.271 415	134.2
25	4-methyl-2-pentanol	3.125 897	2.806 746	18.927 43	5.299 961	131.7
26	2-methyl-3-pentanol	3.180 739	2.861 588	18.856 79	5.276 77	126.5
27	3-methyl-3-pentanol	3.121 320	2.844 927	19.876 70	5.927 895	122.4
28	2-ethyl-1-butanol	3.346 065	2.955 186	18.500 00	4.790 382	146.5
29	2,2-dimethyl-1-butanol	3.121 320	2.730 441	20.570 58	6.049 779	136.8
30	2,3-dimethyl-1-butanol	3.180 739	2.789 860	19.047 25	5.290 365	149.0
31	3,3-dimethyl-1-butanol	3.060 660	2.669 781	21.304 26	6.402 860	143.0
32	2,3-dimethyl-2-butanol	2.943 376	2.666 983	20.695 65	6.281 723	118.6
33	3,3-dimethyl-2-butanol	2.943 376	2.624 224	21.257 00	6.395 639	120.0
34	1-heptanol	3.914 213	3.523 335	21.000 00	4.506 150	176.3
35	1-octanol	4.414 214	4.023 335	23.500 00	4.507 813	195.2
36	1-nonanol	4.914 214	4.523 335	26.000 00	4.508 997	213.1
37	1-decanol	5.414 214	5.023 335	28.500 00	4.509 859	230.2

the new v_i. Similar to eq 1, we have

$$g_{ij} = a_{ij}\frac{v_i^b/v_j^b + v_j^b/v_i^b}{2} \qquad (4)$$

and the indices are expected to reflect the influences of heteroatoms or/and multiple bonds will be generated from the modified EA matrix. For instance, the indices EAΣ and EAmax to compound 2-butanol are 13.500 00 and 5.149 825, respectively.

EXAMINATION OF UNIQUENESS

The purpose of topological indices is to classify structures and to serve for structure–property correlations. Nonuniqueness is not necessarily a disadvantage when correlation study is the prime target, because there are compounds with similar or the same properties which requires a similar or a same index. However, interest continues in trying to devise an index that would be unique. In this section we will exam the selectivity of the new indices introduced in this study.

For examining the selectivities of the EA indices, over 20 000 structures and graphs have been detected. These structures and graphs include the following families.

(1) Acyclic alkane molecular graphs up to $n = 16$ ($n > 16$, not to be detected), the total number of alkane isomers being 18 030, can be differentiated without degeneracy by using the EA indices.

(2) The structures and graphs selected by Randić[8] for examining the uniqueness of ID numbers were detected, containing monocyclic graphs up to $n = 8$ (112 cases), bicyclic graphs up to $n = 7$ (79 cases), all graphs on five vertices, sesquiterpenes (30 cases), and miscellaneous, etc. The EA indices show a remarkable ability to discriminate among the structures and graphs. As an example, the EA index values and the corresponding skeletons for graphs 141–160 as numbered by Randić are shown in Table 2 and Figure 1, respectively. Evidently, all of these structures can be discriminated.

(3) For more rigorously determining the uniqueness of EA indices, the cyclic graphs having $n = 8$ vertices, the degree of each vertex being 4, offer novel comparisons. All of the 204 structures were enumerated with the program ISOM,[9] developed for elucidation of structures of organic compounds in our laboratory. The EA indices also show a high selectivity, because most of the very closely related structures can be discriminated by the EA indices, although there are some pair counter examples having the same index values. As an example, Figure 2 shows the partial cyclic graphs of these structures, and Table 3 gives the corresponding values of EAΣ and EAmax. In which, 11 and 12, 13 and 14, 15 and 16, and 17 and 18 are the degeneracy pairs.

Note, that if two structures do not have the same EAΣ and EAmax index values simultaneously, these two structures are considered to be discriminated in this study.

APPLICATIONS TO CORRELATION

Topological indices developed for the purpose of obtaining correlations with physicochemical properties and biological activities of chemical substances have been applied for a very extensive range. The current major applications include bibliographical species classification, physicochemical parameter evaluation, and pharmaceutical drug design. In this section the employed results will be given for alkanes, alcohols, and barbiturates.

Table 7. Results of Relationships for 37 Alcohols

indices	eq	R	F	S
$^1\chi^v$	bp/°C = 29.64 + 88.83$^1\chi^v$	0.9556		10.29
$^1\chi\,^1\chi^v$	bp/°C = −34.00 + 168.10$^1\chi$ − 127.98$^1\chi^v$	0.9923	1088.86	4.39
EA	bp/°C = 132.14 + 7.70EAΣ − 26.36EAmax	0.9838	511.90	6.35

Table 8. log *P* and the EA Indices for Barbiturates with Structure I

no.	R1	R2	R3	EAΣ	EAmax	log *P*
1	methyl	ethyl	methyl	42.442 200	5.588 789	1.15
2	ethyl	ethyl	methyl	44.295 792	5.355 969	1.65
3	propyl	ethyl	methyl	46.774 139	5.349 703	2.15
4	isopropyl	ethyl	methyl	47.172 211	5.443 005	1.95
5	methyl	methyl	ethyl	42.610 088	5.647 092	1.15
6	ethyl	methyl	ethyl	44.549 950	5.415 510	1.65
7	propyl	methyl	ethyl	47.032 730	5.406 142	2.15
8	isopropyl	methyl	ethyl	46.949 421	5.412 628	1.95
9	methyl	propyl	methyl	44.936 668	5.587 068	1.65
10	ethyl	propyl	methyl	46.786 049	5.354 176	2.15
11	methyl	isopropyl	methyl	45.041 950	5.610 556	1.45
12	methyl	butyl	methyl	47.435 841	5.586 836	2.15
13	ethyl	butyl	methyl	49.283 989	5.353 878	2.65
14	ethyl	ethyl	propyl	49.262 020	5.350 319	2.65

Table 9. log *P* and the EA Indices for Barbiturates with Structure II

no.	R1	R2	EAΣ	EAmax	log *P*
15	methyl	1-methyl, 1-propenyl	40.068 581	5.633 902	0.65
16	ethyl	1-methyl, 1-propenyl	41.963 631	5.396 735	1.15
17	propyl	1-methyl, 1-propenyl	44.444 069	5.390 218	1.65
18	butyl	1-methyl, 1-propenyl	46.940 262	5.388 160	2.15
19	methyl	1-methylvinyl	37.686 401	5.675 070	0.15
20	ethyl	1-methylvinyl	39.664 139	5.447 943	0.65
21	propyl	1-methylvinyl	42.148 720	5.441 714	1.15
22	butyl	1-methylvinyl	44.646 042	5.440 750	1.65
23	isobutyl	1-methylvinyl	44.718 639	5.461 265	1.45
24	amyl	1-methylvinyl	47.145 580	5.440 601	2.15
25	isoamyl	1-methylvinyl	47.212 662	5.459 260	1.95

1. Alkanes. It is of interest to test methods on data for alkanes because good data are generally available for complete isomer sets. The first 150 alkanes (i.e., up to 10 carbon atoms) have been used by Mihalic et al.[10] for a comparison of performances of the 10 distance indices and 2 connectivity indices in the structure–boiling point correlations. In this paper, the same set of compounds except methane, i.e., 149 alkanes, were employed, and EA indices of all of the 149 alkanes were calculated and are listed in Table 4. For comparisons, the index values calculated by other schemes, *W*,[1] *Z*,[2] χ,[3] and *J*[4] are also given in Table 4.

Table 5 shows a summary of the correlation coefficients, *r*, standard deviations, *s*, and *F*-test values by all of the indices. Evidently, the best result has been achieved by using EA indices.

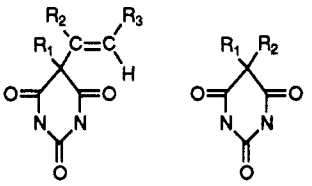
2. Alcohols. Because the most interesting structures possessing activity are rather complex molecules with multiple bonds and/or heteroatoms, it is quite important that topological indices are able to correlate these kinds of molecules. Alcohol contains one heteroatom, oxygen, whose ELC is 3.5 (see Table 1).

Similarly, for comparisons, besides boiling points of 37 alcohols and their EA indices, χ and $^1\chi^v$ [5] are also given in Table 6. The results of the correlation analysis are shown in Table 7.

It is clear that the better result can be obtained with EA indices, though the results obtained using χ and $^1\chi^v$ is slightly superior to ours.

3. Barbiturates. Barbiturates were thought to be nonspecific narcotic agents principally because log *P* (*P* = partition coefficient in octanol–water) correlates very well with their biological potency.[11] Other studies[12] showed a dependence of the action of barbiturates upon chemical structure. Therefore, it was of interest to carry out a correlation analysis of log *P* and topological parameters. Correlations between the EA indices and log *P* for barbiturates have been revealed by this investigation.

The EA indices and log *P* for 25 barbiturate acid derivatives with structure **I** and structure **II** are list in Table 8 and Table 9, respectively.



I II

The results of the correlation analysis between the EA indices and log *P* values of these compounds are

$$\log P = -2.8302 + 0.1900\text{EA}\Sigma - 0.7395\text{EAmax}$$

$$R = 0.9910 \quad F = 602.8965 \quad S = 0.0861 \quad n = 25$$

ACKNOWLEDGMENT

The financial support of the Natural Science Foundation of China is gratefully acknowledged.

REFERENCES AND NOTES

(1) Wiener, H. Structural Determination of Paraffin Boiling Points. *J. Am. Chem. Soc.* **1947**, *69*, 17–20.
(2) Hosoya, H. Topological Index. A Newly Proposed Quantity Characterizing the Topological Nature of Structural Isomers of Saturated Hydrocarbons. *Bull. Chem. Soc. Jan.* **1971**, *44*, 2332–2339.
(3) Randić, M. On Characterization of Molecular Branching. *J. Am. Chem. Soc.* **1975**, *97*, 6609–6615.
(4) Balaban, A. T. Topological Indices Based on Topological Distances in Molecular Graphs. *Pure Appl. Chem.* **1983**, *55*, 199–206.
(5) Kier, I. B.; Hall, L. H. *Molecular Connectivity in Chemistry and Drug Research*; Academic: New York, 1976.
(6) Randić, M.; Trinajstic, N. In search for graph invariants of chemical interest. *J. Mol. Struct. (THEOCHEM)* **1993**, *300*, 551–571.
(7) Randić, M.; Trinajstic, N. Viewpoint 4-Comparative structure-property studies: the connectivity basis. *J. Mol. Struct. (THEOCHEM)* **1993**, *284*, 209–221.
(8) Randić, M. On Molecular Identification Numbers. *J. Chem. Inf. Comput. Sci.* **1984**, *24*, 164–175.
(9) Hu, C.-Y.; Xu, L. Studies on Expert System for the Elucidation of the Structures of Organic Compounds-Structural Generator of ESESOC-II. *Sci. China*, in press.
(10) Mihalic, Z.; Nikolic, S.; Trinajstic, N. Comparative Study of Molecular Descriptors Derived from the Distance Matrix. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 28–37.
(11) Hansch, C.; Anderson, S. M. Structure–Activity Relation in Barbiturates and Its Similarity to That in Other Narcotics. *J. Med. Chem.* **1967**, *10*, 745–753.
(12) Bask, S. C.; Monsrud, L. J.; Rosen, M. E.; Frane, C. M.; Macnuson, V. R. A Comparative Study of Lipophilicity and Topology Indexes in Biological Correlation. *Acta Pharm. Jugosl.* **1986**, *36*, 81–95.