New Group Contribution Method for Estimating Properties of Pure Compounds

Leonidas Constantinou and Rafiqul Gani
Engineering Research Center, IVC-SEP, Dept. of Chemical Engineering
Technical University of Denmark, 2800 Lyngby, Denmark

A new group contribution method for the estimation of properties of pure organic compounds is presented. Estimation is performed at two levels: the basic level uses contributions from first-order groups, while the next higher level uses a small set of second-order groups having the first-order groups as building blocks. Thus, the method provides both a first-order approximation (first-order group contributions) and a more accurate second-order prediction (first- and second-order group contributions). This article discusses methods for prediction of normal boiling point, normal melting point, critical pressure, critical temperature, critical volume, standard enthalpy of vaporization at 298 K, standard Gibbs energy, and standard enthalpy of formation at 298 K. The predictions are based exclusively on the molecular structure of the compound, and the method is able to distinguish among isomers. Compared to the currently-used methods, this technique demonstrates significant improvements in accuracy and applicability.

Introduction

In the study of a chemical, biochemical or environmental system, knowledge of a variety of physical and chemical properties of pure compounds and mixtures under specified conditions is required. It is not always possible, however, to find reliable experimental values of these properties for the compounds of interest in the literature, nor is it practical to measure the properties as the need arises.

In process simulation, reliable and accurate property estimation methods play an important role in the solution of various simulation problems where convergence is often traced to failures in the reliable predictions of physical and thermodynamic properties. In design, the screening of alternatives depends on fast, yet accurate, evaluation of properties for a large number of pure compounds and mixtures (Gani et al., 1991; Constantinou et al., 1994a). Therefore, in computer-aided process and product design, simple, efficient, and reliable methods for the estimation of properties of organic compounds from their molecular structure are essential for the analysis and design of products and processes.

For the estimation of physical and thermodynamic prop-

erties of pure compounds, group contribution methods (Joback and Reid, 1984; Lydersen, 1955; Ambrose, 1978, 1980; Klincewicz and Reid, 1984; Reid et al., 1987; Lyman et al., 1990; Horvath, 1992) are the most widely used. In these methods, the property of a compound is estimated as a summation of the contributions of simple first-order groups which can occur in the molecular structure. They provide the important advantage of quick estimates without requiring substantial computational resources. However, many of these methods are of questionable accuracy and utility. For example, their reliability in the estimation of properties essential in separation processes, such as the normal boiling point, has often been questioned (Horvarth, 1992). Furthermore, the estimation of the critical temperature requires experimentally determined values of the normal boiling point which may not always be available. Also, the molecular structure is often oversimplified, making isomers indistinguishable. Finally, the equations employed in these methods often give poor extrapolation results and may yield limiting values inconsistent with theory (Tsonopoulos and Tan, 1993).

To overcome some of the disadvantages of the preceding techniques, several attempts have been reported in the literature. Fedors (1982) has provided a method for the estimation of the critical temperature based exclusively on the molecular

Correspondence concerning this article should be addressed to R. Gani.

structure of the compounds. Although less accurate than other techniques, the method has often been recommended in the literature due to its pure-predictive character (Daubert and Danner, 1989). Among other investigators, Jalowka and Daubert (1986) and Daubert and Bartakovits (1989) have employed another class of group contribution methods involving the type of the functional groups introduced by Benson et al. (1968), that is, particular grouping of atoms in a specified environment of other atoms. Despite the improved accuracy, the employment of Benson-type groups depicts a fairly complicated scheme (Reid et al., 1987). Besides, the expressions of the properties require other determined properties and therefore these techniques have similar restrictions of the earlier approaches. In addition, the determination of group contributions is carried out separately for hydrocarbons and nonhydrocarbons. Kreglenski and Zwolinski (1961), Tsonopoulos (1987), Gray et al. (1989), Teja et al. (1990), and Magoulas and Tassios (1990) have proposed correlating critical properties and normal boiling point to the number of carbon atoms in the molecule for homologous series, such as normal alkanes and alkanols. Although these correlations have been found to be accurate, their range of application has been questioned by Kreglenski and Zwolinski (1961) and Tsonopoulos and Tan (1993).

Constantinou et al. (1993, 1994c) have recently proposed a new additive property estimation method which is based on conjugation operators and applicable to organic compounds. In this method, the molecular structure of a compound is viewed as a hybrid of a number of conjugate forms (alternative formal arrangements of valence electrons), and the property of a compound is a linear combination of these conjugate forms contribution. Constantinou et al. (1993, 1994c) showed that by including more information about the molecular structure of a compound in a systematic way, it is possible to substantially improve the accuracy of a variety of properties of pure compounds, capture the fine differences among isomers, and lead to pure predictive techniques. However, the generation and enumeration of conjugate forms is a nontrivial issue and requires a symbolic computing environment (Prickett et al., 1993). Despite this complexity, the method enriches the area of property estimation in general and provides a basis which can be considered in the development of simpler techniques.

Motivated by the preceding literature review, our efforts have been focused on developing a new property estimation method which overcomes the limitations imposed by the currently used procedures and allows more efficient and reliable analysis of chemical, biochemical, and environmental products and processes. In this article, we propose a new group contribution technique for the estimation of properties of organic compounds. The basic premise is to provide in a simple and methodical way enough information about the molecular structure of a compound so that significantly improved prediction of properties can be made.

In our method, property estimation is done at two levels. The basic level has contributions from first-order functional groups such as those currently applied for the estimation of mixture properties (Fredenslund et al., 1977; Derr and Deal, 1969). The next level has a set of second-order groups which have the first-order groups as building blocks. Thus, the method allows for both first-order estimation (using only first-order groups) and a more accurate second-order approximation (using both first- and second-order groups). The properties are classified into primary properties and secondary properties. Primary properties are estimated exclusively from the molecular structure while secondary properties (not considered in this article) involve primary properties in addition to the molecular structure data. Here, several important primary properties are estimated: normal boiling point, normal melting point, critical pressure, critical temperature, critical volume, standard enthalpy of vaporization at 298 K, standard Gibbs energy, and standard enthalpy of formation at 298 K. Comparisons with existing methods are made for all the above mentioned properties. Through examples, the applicability of the method, the consistency of prediction, and the capability of the method to capture isomer differences are demonstrated.

Development of the New Method

In the proposed method, the molecular structure of a compound is considered to be a collection of two types of groups: simple functional groups, that is, the first-order groups, and functional groups with the first-order groups as building blocks, namely the second-order groups.

First-order groups

For the estimation of properties of pure compounds, most of the currently-used methods employ the same set of first-order groups (Reid et al., 1987; Lyman et al., 1990). Methods for the estimation of properties of mixtures, on the other hand, have mainly used a different set of first-order groups (Fredenslund et al., 1977; Dear and Deal, 1969). Consequently, group contribution-based computational tools need to accommodate two separate molecular-structure descriptions (Gani et al., 1990): one for prediction of properties of pure compounds and another for mixture property estimations. To circumvent this drawback, we propose to use as first-order groups, the set of groups commonly used for the estimation of mixtures properties (Fredenslund et al., 1977; Dear and Deal, 1969). Another advantage of this selection is that a group appearing in an aliphatic ring is considered equivalent to its identical nonring one. Therefore, each group has a single contribution independent of the type of the compound involved (acyclic or cyclic). The first-order groups are presented on Tables 1 and 2. Sample molecular structures which incorporate these groups have been reported in the literature (Reid et al., 1987; Hansen et al., 1991). An important disadvantage of this group definition is that there is no theoretical basis for their identification (Wu and Sandler, 1992). Besides, the groups are not able to capture proximity effects or isomer differences (Wu and Sandler, 1989, 1992; Kehiaian, 1983).

Second-order groups

In the new method, the role of the second-order groups is to provide more structural information about the portions of the molecular structure of a compound where the description through the first-order groups is insufficient. The ultimate goal is to enhance the accuracy, reliability, and the range of applicability of the property estimation, and overcome some of the disadvantages of the first-order groups, such as partial description of proximity effects and distinguish among isomers. Contrary to the case of first-order groups, there can be molecule structures which do not need any second-order groups.

The selection of the first-order groups aimed to achieve

consistency with the best documented mixture property estimation methods. The definition and identification of the second-order groups, on the other hand, has a theoretical basis. Specifically, we propose the principle of conjugation, as introduced by Mavrovouniotis (1990) and Constantinou et al. (1993, 1994c) in the ABC framework, to be the platform for

the selection of the set of second-order groups to be used in the proposed method.

Theoretical background

In the ABC framework, compounds are represented as hybrids of many conjugate forms. Each conjugate form is an

Table 1. First-Order Groups and Their Contributions for the Physical Properties

Group	t_{c1i}	p_{c1i} (bar$^{-0.5}$)	v_{c1i} (m^3/kmol)	t_{b1i}	t_{m1i}
CH$_3$	1.6781	0.019904	0.07504	0.8894	0.4640
CH$_2$	3.4920	0.010558	0.05576	0.9225	0.9246
CH	4.0330	0.001315	0.03153	0.6033	0.3557
C	4.8823	-0.010404	-0.00034	0.2878	1.6479
CH$_2$=CH	5.0146	0.025014	0.11648	1.7827	1.6472
CH=CH	7.3691	0.017865	0.09541	1.8433	1.6322
CH$_2$=C	6.5081	0.022319	0.09183	1.7117	1.7899
CH=C	8.9582	0.012590	0.07327	1.7957	2.0018
C=C	11.3764	0.002044	0.07618	1.8881	5.1175
CH$_2$=C=CH	9.9318	0.031270	0.14831	3.1243	3.3439
ACH	3.7337	0.007542	0.04215	0.9297	1.4669
AC	14.6409	0.002136	0.03985	1.6254	0.2098
ACCH$_3$	8.2130	0.019360	0.10364	1.9669	1.8635
ACCH$_2$	10.3239	0.012200	0.10099	1.9478	0.4177
ACCH	10.4664	0.002769	0.07120	1.7444	-1.7567
OH	9.7292	0.005148	0.03897	3.2152	3.5979
ACOH	25.9145	-0.007444	0.03162	4.4014	13.7349
CH$_3$CO	13.2896	0.025073	0.13396	3.5668	4.8776
CH$_2$CO	14.6273	0.017841	0.11195	3.8967	5.6622
CHO	10.1986	0.014091	0.08635	2.8526	4.2927
CH$_3$COO	12.5965	0.029020	0.15890	3.6360	4.0823
CH$_2$COO	3.8116	0.021836	0.13649	3.3953	3.5572
HCOO	11.6057	0.013797	0.10565	3.1459	4.2250
CH$_3$O	6.4737	0.020440	0.08746	2.2536	2.9248
CH$_2$O	6.0723	0.015135	0.07286	1.6249	2.0695
CH-O	5.0663	0.009857	0.05865	1.1557	4.0352
FCH$_2$O	9.5059	0.009011	0.06858	2.5892	4.5047
CH$_2$NH$_2$	12.1726	0.012558	0.13128	3.1656	6.7684
CHNH$_2$	10.2075	0.010694	0.07527	2.5983	4.1187
CH$_3$NH	9.8544	0.012589	0.12152	3.1376	4.5341
CH$_2$NH	10.4677	0.010390	0.09956	2.6127	6.0609
CHNH	7.2121	-0.000462	0.09165	1.5780	3.4100
CH$_3$N	7.6924	0.015874	0.12598	2.1647	4.0580
CH$_2$N	5.5172	0.004917	0.06705	1.2171	0.9544
ACNH$_2$	28.7570	0.001120	0.06358	5.4736	10.1031
C$_5$H$_4$N	29.1528	0.029565	0.24831	6.2800	****
C$_5$H$_3$N	27.9464	0.025653	0.17027	5.9234	12.6275
CH$_2$CN	20.3781	0.036133	0.15831	5.0525	4.1859
COOH	23.7593	0.011507	0.10188	5.8337	11.5630
CH$_2$CL	11.0752	0.019789	0.11564	2.9637	3.3376
CHCL	10.8632	0.011360	0.10350	2.6948	2.9933
CCL	11.3959	0.003086	0.07922	2.2073	9.8409
CHCL$_2$	16.3945	0.026808	0.16951	3.9300	5.1638
CCL$_2$	****	****	****	3.5600	****
CCL$_3$	18.5875	0.034935	0.21031	4.5797	10.2337
ACCL	14.1565	0.013135	0.10158	2.6293	2.7336
CH$_2$NO$_2$	24.7369	0.020974	0.16531	5.7619	5.5424
CHNO$_2$	23.2050	0.012241	0.14227	5.0767	4.9738
ACNO$_2$	34.5870	0.015050	0.14258	6.0837	8.4724
CH$_2$SH	13.8058	0.013572	0.10252	3.2914	3.0044
I	17.3947	0.002753	0.10814	3.6650	4.6089
BR	10.5371	-0.001771	0.08281	2.6495	3.7442
CH≡C	7.5433	0.014827	0.09331	2.3678	3.9106
C≡C	11.4501	0.004115	0.07627	2.5645	9.5793
CL-(C=C)	5.4334	0.016004	0.05687	1.7824	1.5598
ACF	2.8977	0.013027	0.05672	0.9442	2.5015
HCON(CH$_2$)$_2$	****	****	****	7.2644	****

(Continued)

Table 1. First-Order Groups and Their Contributions for the Physical Properties (continued)

Group	t_{cli}	p_{cli} (bar$^{-0.5}$)	v_{cli} (m^3/kmol)	t_{bli}	t_{mli}
CF$_3$	2.4778	0.044232	0.11480	1.2880	3.2411
CF$_2$	1.7399	0.012884	0.09519	0.6115	* * * *
CF	3.5192	0.004673	* * * *	1.1739	* * * *
COO	12.1084	0.011294	0.08588	2.6446	3.4448
CCL$_2$F	9.8408	0.035446	0.18212	2,8881	7.4756
HCCLF	* * * *	* * * *	* * * *	2.3086	* * * *
CCLF$_2$	4.8923	0.039004	0.14753	1.9163	2.7523
F (except as above)*	1.5974	0.014434	0.03783	1.0081	1.9623
CONH$_2$	65.1053	0.004266	0.14431	10.3428	31.2786
CONHCH$_3$	* * * *	* * * *	* * * *	* * * *	* * * *
CONHCH$_2$	* * * *	* * * *	* * * *	* * * *	* * * *
CON(CH$_3$)$_2$	36.1403	0.040149	0.25031	7.6904	11.3770
CONCH$_3$CH$_2$	* * * *	* * * *	* * * *	* * * *	* * * *
CON(CH$_2$)$_2$	* * * *	* * * *	* * * *	6.7822	* * * *
C$_2$H$_5$O$_2$	17.9668	0.025435	0.16754	5.5566	* * * *
C$_2$H$_4$O$_2$	* * * *	* * * *	* * * *	5.4248	* * * *
CH$_3$S	14.3969	0.016048	0.13021	3.6796	5.0506
CH$_2$S	17.7916	0.011105	0.11650	3.6763	3.1468
CHS	* * * *	* * * *	* * * *	2.6812	* * * *
C$_4$H$_3$S	* * * *	* * * *	* * * *	5.7093	* * * *
C$_4$H$_2$S	* * * *	* * * *	* * * *	5.8260	* * * *

* For the normal boiling point, the method is not applied to highly partial-fluorinated compounds.

Table 2. First-Order Groups and Their Contributions for the Thermodynamic Properties

| Group	h_{1i} (kJ/mol)	g_{1i} (kJ/mol)	h_{v1i} (kJ/mol)	Group	h_{1i} (kJ/mol)	g_{1i} (kJ/mol)	h_{v1i} (kJ/mol)
CH$_3$	− 45.947	− 8.030	4.116	CH$_2$CL	− 73.568	− 33.373	13.780
CH$_2$	− 20.763	8.231	4.650	CHCL	− 63.795	− 31.502	11.985
CH	− 3.766	19.848	2.771	CCL	− 57.795	− 25.261	9.818
C	17.119	37.977	1.284	CHCL$_2$	− 82.921	− 35.814	19.208
CH$_2$ = CH	53.712	84.926	6.714	CCL$_2$	* * *	* * *	17.574
CH = CH	69.939	92.900	7.370	CCL$_3$	− 107.188	− 53.332	* * *
CH$_2$ = C	64.145	88.402	6.797	ACCL	− 16.752	− 0.596	11.883
CH = C	82.528	93.745	8.178	CH$_2$NO$_2$	− 66.138	17.963	30.644
C = C	104.293	116.613	9.342	CHNO$_2$	− 59.142	18.088	26.277
CH$_2$ = C = CH	197.322	221.308	12.318	ACNO$_2$	− 7.365	60.161	* * *
ACH	11.189	22.533	4.098	CH$_2$SH	− 8.253	16.731	14.931
AC	27.016	30.485	12.552	I	57.546	46.945	14.364
ACCH$_3$	− 19.243	22.505	9.776	BR	1.834	− 1.721	11.423
ACCH$_2$	9.404	41.228	10.185	CH ≡ C	220.803	217.003	7.751
ACCH	27.671	52.948	8.834	C ≡ C	227.368	216.328	11.549
OH	− 181.422	− 158.589	24.529	CL-(C = C)	− 36.097	− 28.148	* * *
ACOH	− 164.609	− 132.097	40.246	ACF	− 161.740	− 144.549	4.877
CH$_3$CO	− 182.329	− 131.366	18.999	HCON(CH$_2$)$_2$	* * *	* * *	* * *
CH$_2$CO	− 164.410	− 132.386	20.041	CF$_3$	− 679.195	− 626.580	8.901
CHO	− 129.158	− 107.858	12.909	CF$_2$	* * *	* * *	1.860
CH$_3$COO	− 389.737	− 318.616	22.709	CF	* * *	* * *	8.901
CH$_2$COO	− 359.258	− 291.188	17.759	COO	− 313.545	− 281.495	* * *
HCOO	− 332.822	− 288.902	* * *	CCL$_2$F	− 258.960	− 209.337	13.322
CH$_3$O	− 163.569	− 105.767	10.919	HCCLF	* * *	* * *	* * *
CH$_2$O	− 151.143	− 101.563	7.478	CCLF$_2$	− 446.835	− 392.975	8.301
CH-O	− 129.488	− 92.099	5.708	F (except as above)	− 223.398	− 212.718	* * *
FCH$_2$O	− 140.313	− 90.883	11.227	CONH$_2$	− 203.188	− 136.742	* * *
CH$_2$NH$_2$	− 15.505	58.085	14.599	CONHCH$_3$	− 67.778	* * *	* * *
CHNH$_2$	3.320	63.051	11.876	CONHCH$_2$	− 182.096	* * *	51.787
CH$_3$NH	5.432	82.471	14.452	CON(CH$_3$)$_2$	− 189.888	− 65.642	* * *
CH$_2$NH	23.101	95.888	14.481	CONCH$_3$CH$_2$	− 46.562	* * *	* * *
CHNH	26.718	85.001	* * *	CON(CH$_2$)$_2$	* * *	* * *	* * *
CH$_3$N	54.929	128.602	6.947	C$_2$H$_5$O$_2$	− 344.125	− 241.373	* * *
CH$_2$N	69.885	132.756	6.918	C$_2$H$_4$O$_2$	* * *	* * *	* * *
ACNH$_2$	20.079	68.861	28.453	CH$_3$S	− 2.084	30.222	16.921
C$_5$H$_4$N	134.062	199.958	31.523	CH$_2$S	18.022	38.346	17.117
C$_5$H$_3$N	139.758	199.288	31.005	CHS	* * *	* * *	13.265
CH$_2$CN	88.298	121.544	23.340	C$_4$H$_3$S	* * *	* * *	27.966
COOH	− 396.242	− 349.439	43.046	C$_4$H$_2$S	* * *	* * *	* * *
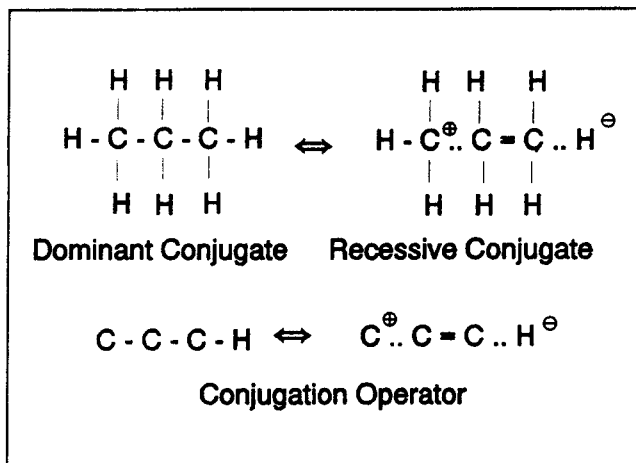
Figure 1. A dominant conjugate, a generated recessive conjugate and the corresponding conjugation operator.

idealized structure with integer-order localized bonds and integer charges on atoms. The purely covalent conjugate form is the dominant conjugate and the ionic forms are the recessive conjugates, which can be obtained from the dominant form by a rearrangement of electron pairs. A conjugation operator defines a particular pattern of electron rearrangement. When applied to the dominant conjugate, an operator yields an entire class of recessive conjugates. Conjugation operators are represented by a distinct subchain with two or three bonds, such as $C - C - C - H$ and $O = C - C$. Figure 1 presents a dominant conjugate, a generated recessive conjugate form, and the corresponding conjugation operator (Constantinou et al., 1993). In the framework, the property of a compound is estimated by determining and combining properties of its conjugate forms. Properties of conjugate forms are estimated through conjugation operators (a fixed contribution is assigned to each operator (through regression) reflecting the contribution of the corresponding class of conjugate forms). The basic property in the ABC framework is the standard enthalpy of formation at 298 K, because for the estimation of this property, the contribution of the conjugate forms can be expressed in terms of their physical significance rather than adjustable parameters (Constantinou et al., 1993). Thus, it is possible to distinguish those classes of conjugate forms (the most important conjugate forms) which exert the strongest influence not only on the estimation of the enthalpy of formation but also on other properties to be estimated as well. The conjugation operators associated with the most important conjugate forms are assigned the largest contributions (contributions of all the operators are given by Constantinou (1993)). In this way, it is possible to identify the classes of conjugate forms with the highest conjugation activity by inspecting the contributions of the corresponding operators. The group identification focuses on the operators which correspond to the important conjugate forms, that is, the operators with significantly higher contribution than others.

Group identification

In the new method, we take the following as our principles for the identification of the second-order groups:

(1) The structure of a second-order group should incorporate the distinct subchain of at least one important conjugation operator (due to structure similarity of the operators, second-order groups can involve the subchains of more than one operator); for example, the CH_3COCH_2 second-order group incorporates the $O = C - C$, along with the $O = C - C - H$ and $C - C - C - H$ operators.

(2) The structure of a functional second-order group should have adjacent first-order groups as building blocks and it should be as small as possible; for example, the CH_3COCH_2 has the CH_3CO, and CH_2 as building blocks.

(3) Second-order groups based on common operator(s) should be equally treated in the method; for example the $CH_n = CH_m - CH_p = CH_k$ $[n,k,p,m\epsilon(0,2)]$ represent the second-order groups obtained by the $C = C - C = C$ operator.

(4) The performance of second-order groups is independent of the molecule in which the groups occur, satisfying a fundamental group-contribution principle.

Tables 3 and 4 give lists of second-order groups that have been defined for the new method. Sample assignments, along with the occurrences of the groups, are provided in Table 4. Examples of how the second-order groups are used (if they do exist in a molecular structure) are given in Appendix A.

Another systematic procedure for identification of groups based on theoretical principles is the work of Wu and Sandler (1992), where the identification of groups is based on *ab initio* orbital calculations, rather than conjugation. Nevertheless, common groups can be found in the set proposed by Wu and Sandler (1992) and in the set proposed in this work.

Proposed model

Let C_i be the contribution of the first-order group type-i which occurs N_i times and D_j be the contribution of the second-order group type-j with M_j occurrences in a compound. With $f(X)$ a simple function of the property X, the property-estimation model takes the form of Eq. 1.

$$f(X) = \sum_i N_i C_i + W \sum_j M_j D_j \qquad (1)$$

The constant W is assigned to unity in the second-level estimation (second-order approximation), where both first and second-order group contributions are involved, and 0 in the basic level (first-order approximation), where only the contributions of first-order groups are employed. The selection of $f(X)$ has been based on the following factors: (1) The function has to achieve additivity in the contributions C_i and D_j; (2) it has to demonstrate the best possible fit of the experimental data; and (3) the expressions should be able to provide sufficient extrapolating behavior and therefore wide range of applicability. A large number of expressions have been tested for each property (including those found in Constantinou et al., 1993; Reid et al., 1987; Horvath, 1992). The selected functions of $f(X)$ for all the properties are presented in Table 5. The symbols t_{c1i}, p_{c1i}, v_{c1i}, t_{m1i}, t_{b1i}, g_{1i}, h_{v1i}, h_{1i} represent the contributions (C_i) of the first-order group of type-i for the corresponding properties. Similarly, t_{c2j}, p_{c2j}, v_{c2j}, t_{m2j}, t_{b2j}, g_{2j}, h_{v2j}, h_{2j} are the contributions (D_j) of the second-order group of type-j. The t_{c0}, p_{c1}, p_{c2}, v_{c0}, t_{m0}, t_{b0}, g_0, h_{v0}, and h_{f0} are additional adjustable parameters of the expressions (universal constants).

The determination of the adjustable parameters of the model,

Table 3. Second-Order Groups and Their Contributions for the Physical Properties

Groups	t_{c2j}	p_{c2j} $(bar^{-0.5})$	v_{c2j} $(m^3/kmol)$	t_{b2j}	t_{m2j}
$(CH_3)_2CH$	−0.5334	0.000488	0.00400	−0.1157	0.0381
$(CH_3)_3C$	−0.5143	0.001410	0.00572	−0.0489	−0.2355
$CH(CH_3)CH(CH_3)$	1.0699	−0.001849	−0.00398	0.1798	0.4401
$CH(CH_3)C(CH_3)_2$	1.9886	−0.005198	−0.01081	0.3189	−0.4923
$C(CH_3)_2C(CH_3)_2$	5.8254	−0.013230	−0.02300	0.7273	6.0650
3 membered ring*	−2.3305	0.003714	−0.00014	0.4745	1.3772
4 membered ring*	−1.2978	0.001171	−0.00851	0.3563	* * * *
5 membered ring*	−0.6785	0.000424	−0.00866	0.1919	0.6824
6 membered ring*	0.8479	0.002257	0.01636	0.1957	1.5656
7 membered ring*	3.6714	−0.009799	−0.02700	0.3489	6.9709
$CH_n=CH_m\text{-}CH_p=CH_k$ $k,n,m,p \in (0,2)$	0.4402	0.004186	−0.00781	0.1589	1.9913
$CH_3\text{-}CH_m=CH_n$ $m,n \in (0,2)$	0.0167	−0.000183	−0.00098	0.0668	0.2476
$CH_2\text{-}CH_m=CH_n$ $m,n \in (0,2)$	−0.5231	0.003538	0.00281	−0.1406	−0.5870
$CH\text{-}CH_m=CH_n$ or $C\text{-}CH_m=CH_n$ $m,n \in (0,2)$	−0.3850	0.005675	0.00826	−0.0900	−0.2361
Alicyclic side chain $C_{cyclic}C_m$ $m>1$	2.1160	−0.002546	−0.01755	0.0511	−2.8298
CH_3CH_3	2.0427	0.005175	0.00227	0.6884	1.4880
CHCHO or CCHO	−1.5826	0.003659	−0.00664	−0.1074	2.0547
CH_3COCH_2	0.2996	0.001474	−0.00510	0.0224	−0.2951
CH_3COCH or CH_3COC	0.5018	−0.002303	−0.00122	0.0920	−0.2986
$C_{cyclic}(=O)$	2.9571	0.003818	−0.01966	0.5580	0.7143
ACCHO	1.1696	−0.002481	0.00664	0.0735	−0.6697
CHCOOH or CCOOH	−1.7493	0.004920	0.00559	−0.1552	−3.1034
ACCOOH	6.1279	0.000344	−0.00415	0.7801	28.4324
CH_3COOCH or CH_3COOC	−1.3406	0.000659	−0.00293	−0.2383	0.4838
$COCH_2COO$ or COCHCOO or COCCOO	2.5413	0.001067	−0.00591	0.4456	0.0127
CO-O-CO	−2.7617	−0.004877	−0.00144	−0.1977	−2.3598
ACCOO	−3.4235	−0.000541	0.02605	0.0835	−2.0198
CHOH	−2.8035	−0.004393	−0.00777	−0.5385	−0.5480
COH	−3.5442	0.000178	0.01511	−0.6331	0.3189
$CH_m(OH)CH_n(OH)$ $m,n \in (0,2)$	5.4941	0.005052	0.00397	1.4108	0.9124
$CH_{m\,cyclic}\text{-}OH$ $m \in (0,1)$	0.3233	0.006917	−0.02297	−0.0690	9.5209
$CH_m(OH)CH_n(NH_p)$ $m,n,p \in (0,3)$	5.4864	0.001408	0.00433	1.0682	2.7826
$CH_m(NH_2)CH_n(NH_2)$ $m,n \in (0,2)$	2.0699	0.002148	0.00580	0.4247	2.5114
$CH_{m\,cyclic}\text{-}NH_p\text{-}CH_{n\,cyclic}$ $m,n,p \in (0,2)$	2.1345	−0.005947	−0.01380	0.2499	1.0729
$CH_m\text{-}O\text{-}CH_n=CH_p$ $n,n,p \in (0,2)$	1.0159	−0.000878	0.00297	0.1134	0.2476
$AC\text{-}O\text{-}CH_m$ $m \in (0,3)$	−5.3307	−0.002249	−0.00045	−0.2596	0.1175
$CH_{m\,cyclic}\text{-}S\text{-}CH_{n\,cyclic}$ $m,n \in (0,2)$	4.4847	* * * *	* * * *	0.4408	−0.2914
$CH_m=CH_n\text{-}F$ $m,n \in (0,2)$	−0.4996	0.000319	−0.00596	−0.1168	−0.0514
$CH_m=CH_n\text{-}Br$ $m,n \in (0,2)$	−1.9334	−0.004305	0.00507	−0.3201	−1.6425
$CH_m=CH_n\text{-}I$ $m,n \in (0,2)$	* * * *	* * * *	* * * *	−0.4453	* * * *
ACBr	−2.2974	0.009027	−0.00832	−0.6776	2.5832
ACI	2.8907	0.008247	−0.00341	−0.3678	−1.5511
$CH_m(NH_2)\text{-}COOH$ $m \in (0,2)$	* * * *	* * * *	* * * *	* * * *	* * * *

* Corrections for stress-strained effects are treated similar to second-order groups.

that is, C_js and D_js, has been divided into a two-step regression analysis:

(1) Regression is carried out to determine the contribution of the first-order groups (that is, values for C_js) while W is set to zero (that is, second-order groups are not considered in this step);

(2) Using the estimated values of C_js, the second-order groups are activated ($W = 1$) and the contributions of the second-order groups (D_js) are estimated through regression. This two-step regression procedure enables the contribution of first-order groups to be independent of that of the second-order groups. Furthermore, the contribution of second-order groups serves as a correction to the first-order approximation.

Selection of experimental data

A common problem in the literature on experimental thermodynamic data is that a portion of the data are actually estimated. For example, Constantinou (1993) mentions that an extensive literature search revealed that a portion of the data reported in Reid et al. (1987) were actually estimated using group-contributions and other additive methods (during the data compilation, nonexperimental sources may not have been screened out). Therefore, one must be very cautious in the use of these data for regression. The reason is that minor differences in the data, if they are the result of a systematic bias, can adversely influence the parameters to be determined, even when they do not adversely affect the standard deviation of the regression. In this work, a systematic search has been made to satisfy the need of providing reliable and consistent experimental data for the regression step. The DIPPR databank (Daubert and Danner, 1989) has been the main source for experimental values of the critical properties, normal boiling point, normal melting point, standard enthalpy of formation, and standard Gibbs energy at 298 K. The reason was that the availability of data sources and quality codes in DIPPR has permitted us to judge whether specific data are reliable experimental or not. Similarly, the experimental data of Nikitin et al. (1994) of the critical properties of heavy alkanes and values of the compilation of Constantinou (1993), not included in DIPPR, have also been used. The experimental data for the standard enthalpy of vaporization at 298 K were obtained from Cox and Pilcher (1970), a descriptive thermodynamic data source.

Regression

After the selection of data, a least-square analysis has been carried out to determine the contributions of first- and second-order groups (adjustable parameters) for the eight primary

Table 4. Second-Order Groups and Their Contributions for the Thermodynamic Properties along with Sample Assignments

Groups	h_{2j} (kJ/mol)	g_{2j} (kJ/mol)	h_{v2j} (kJ/mol)	Sample Group Assignment (Occurrences)
$(CH_3)_2CH$	−0.860	0.297	0.292	2-Methylpentane (1)
$(CH_3)_3C$	−1.338	−0.399	−0.720	2,2-Dimethylpentane (1), 2,2,4,4-Tetramethylpentane (2)
$CH(CH_3)CH(CH_3)$	6.771	6.342	0.868	2,3-Dimethylpentane (1), 2,3,4-Tetramethylpentane (2)
$CH(CH_3)C(CH_3)_2$	7.205	7.466	1.027	2,2,3 Dimethylpentane (1), 2,2,3,4,4-Pentamethylpentane (2)
$C(CH_3)_2C(CH_3)_2$	14.271	16.224	2.426	2,2,3,3 Dimethylpentane (1), 2,2,3,4,4 hexamethylpentane (2)
3 membered ring*	104.800	94.564	* * * *	cyclopropane (1)
4 membered ring*	99.455	92.573	* * * *	cyclobutane (1)
5 membered ring*	13.782	5.733	−0.568	cyclopentane (1), EthylCyclopentane (1)
6 membered ring*	−9.660	−8.180	−0.905	cyclohexane (1), Methylcyclohexane (1)
7 membered ring*	15.465	20.597	−0.847	cycloheptane (1), EthylCycloheptane (1)
$CH_n = CH_m\text{-}CH_p = CH_k$ $k,n,m,p \in (0,2)$	−8.392	−5.505	2.057	1,3 butadiene (1)
$CH_3\text{-}CH_m = CH_n$ $m,n \in (0,2)$	0.474	0.950	−0.073	butene-2 (2), 2-Methyl-Butene-2 (3)
$CH_2\text{-}CH_m = CH_n$ $m,n \in (0,2)$	1.472	0.699	−0.369	1,4 Pentadiene (2)
$CH\text{-}CH_m = CH_n$ or $C\text{-}CH_m = CH_n$ $m,n \in (0,2)$	4.504	1.013	0.345	4-methylpentene-2 (1)
Alicyclic side-chain $C_{cyclic}C_m$ $m > 1$	1.252	1.041	−0.114	Ethylcyclohexane (1), Propylcycloheptane (1)
CH_3CH_3	−2.792	−1.062	* * * *	Ethane (only)
CHCHO or CCHO	−2.092	−1.359	0.207	2-methyl Butyraldehyde (1)
CH_3COCH_2	0.975	0.075	−0.668	2-pentanone (1)
CH_3COCH or CH_3COC	4.753	* * * *	0.071	3-methyl-pentanone-2 (1)
$C_{cyclic} = O$	14.145	23.539	0.744	Cyclohexanone (1)
ACCHO	−3.173	−2.602	−3.410	Benzaldehyde (1)
CHCOOH or CCOOH	1.279	2.149	* * * *	2-methyl butanoic acid (1)

(continued)

Table 4. Second-Order Groups and Their Contributions of Thermodynamic Properties along with Sample Assignments
(Continued)

Groups	h_{2j} (kJ/mol)	g_{2j} (kJ/mol)	h_{v2j} (kJ/mol)	Sample Group Assignment (Occurrences)
ACCOOH	12.245	10.715	8.502	Benzoic acid (1)
CH₃COOCH or CH₃COOC	− 7.807	− 6.208	− 3.345	Isopropyl Acetate (1)
COCH₂COO or COCHCOO or COCCOO	37.462	29.181	* * * *	Ethylacetoacetate (1)
CO-O-CO	− 16.097	− 11.809	1.517	Propanoic anhydride (1)
ACCOO	− 9.874	− 7.415	* * * *	Benzoic Acid ethyl ester (1)
CHOH	− 3.887	− 6.770	− 1.398	2-butanol (1)
COH	− 24.125	− 20.770	0.320	2-methyl butanol-2 (1)
CH$_m$(OH)CH$_n$(OH) $m,n \in (0,2)$	0.366	3.805	− 3.661	1,2,3 propanetriol (2)
CH$_{m\,cyclic}$-OH $m \in (0,1)$	− 16.333	− 5.487	4.626	Cyclopentanol (1)
CH$_m$(OH)CH$_n$(NH$_p$) $m,n,p \in (0,2)$	− 2.992	− 1.600	* * * *	1-amino-2-butanol (1), 1-Hydroxy-N-Methyl-Butylamine (1)
CH$_m$(NH₂)CH$_n$(NH₂) $m,n \in (0,2)$	2.855	1.858	* * * *	1,2-propanediamine (1)
CH$_{m\,cyclic}$-NH$_p$-CH$_{n\,cyclic}$ $m,n,p \in (0,2)$	0.351	8.846	2.311	Pyrrolidine (1)
CH$_m$-O-CH$_n$=CH$_p$ $n,n,p \in (0,2)$	− 8.644	− 13.167	* * * *	Ethyl-vinyl Ether (1)
AC-O-CH$_m$ $m \in (0.3)$	1.532	− 0.654	* * * *	Ethyl Phenyl Ether (1)
CH$_{m\,cyclic}$-S-CH$_{n\,cyclic}$ $m,n \in (0,2)$	− 0.329	− 2.091	0.972	Tetrahydro thiophene (1)
CH$_m$=CH$_n$-F $m,n \in (0,2)$	* * * *	* * * *	* * * *	1-Fluoro propene-1 (1)
CH$_m$=CH$_n$-Br $m,n \in (0,2)$	11.989	12.373	* * * *	1-Bromo propene-1 (1)
CH$_m$=CH$_n$-I $m,n \in (0,2)$	* * * *	* * * *	* * * *	1-Iodo propene-1 (1)
ACBr	12.285	14.161	− 7.488	Bromotoluene (1)
ACI	11.207	12.530	− 4.864	Iodotoluene (1)
CH$_m$(NH₂)-COOH $m \in (0,2)$	11.740	* * * *	* * * *	2-amino hexanoic acid (1)

* Corrections for stress-strained effects in molecules are treated similar to second-order groups.

properties listed earlier. The optimization algorithm used for the regression is the modified-Levenberg approach and the objective function was to minimize the sum of squares of the differences between the experimental and the predicted values

Table 5. Selected Function for Each Property

Property (X)	Lefthand Side of Eq. 1 [Function $f(X)$]	Righthand side of Eq. 1 (Group-Contribution Expression)
Critical Temperature (T_c)	$\exp(T_c/t_{c0})$	$\Sigma_i N_i t_{c1i} + \Sigma_j M_j t_{c2j}$
Critical Pressure (P_c)	$(P_c - p_{c1})^{-0.5} - p_{c2}$	$\Sigma_i N_i p_{c1i} + \Sigma_j M_j p_{c2j}$
Critical Volume (V_c)	$V_c - v_{c0}$	$\Sigma_i N_i v_{c1i} + \Sigma_j M_j v_{c2j}$
Melting Point (T_m)	$\exp(T_m/t_{m0})$	$\Sigma_i N_i t_{m1i} + \Sigma_j M_j t_{m2j}$
Normal Boiling Point (T_b)	$\exp(T_b/t_{b0})$	$\Sigma_i N_i t_{b1i} + \Sigma_j M_j t_{b2j}$
Standard Gibbs Energy at 298 K (ΔG_f)	$\Delta G_f - g_0$	$\Sigma_i N_i g_{1i} + \Sigma_j M_j g_{2j}$
Standard Enthalpy of Vaporization at 298 K (ΔH_v)	$\Delta H_v - h_{v0}$	$\Sigma_i N_i h_{v1i} + \Sigma_j M_j h_{v2j}$
Standard Enthalpy of Formation at 298 K (ΔH_f)	$\Delta H_f - h_{f0}$	$\Sigma_i N_i h_{1i} + \Sigma_j M_j h_{2j}$

of the functions $f(X)$. This method has been used in all the regressions reported in this article.

A list of the contributions of the first-order groups for the studied properties is presented in Tables 1 and 2. In all the subsequent tables, four asterisks (****) denotes group-contributions which have not been estimated because reliable experimental values of the compounds involving the groups could not be found in the open literature. The additional adjustable parameters determined in the first step of the regression (universal constants) are presented in Table 6.

With the first-order group contribution known, attention was switched to the determination of the contributions of the second-order groups, that is, the second-step of the regression. Tables 3 and 4 report the contributions of these groups for all the properties.

Reliability of the model

The reliability of the expressions determined through regression has been tested for each property through a least-squares regression in which a randomly chosen subset of the experimental data has been excluded (Constantinou et al., 1993). Once the results are obtained, the mean-square residual J, defined as:

Table 6. Values of the Additional Adjustable Parameters

Adjustable Parameter (Universal Constants)	Value
t_{c0}	181.128 K
p_{c1}	1.3705 bar
p_{c2}	0.100220 bar$^{-0.5}$
v_{c0}	-0.004350 m^3/kmol
t_{b0}	204.359 K
t_{m0}	102.425 K
g_0	-14.828 kJ/mol
h_{v0}	6.829 kJ/mol
h_{f0}	10.835 kJ/mol

$$J = \sqrt{\frac{1}{n} \Sigma (X_i - Y_i)^2} \qquad (2)$$

is calculated. In Eq. 2, n is the number of data excluded from the regression, X_i is the property of the compound i estimated by the initial regression (full data set), and Y_i is the property of the same compound estimated by the regression in which the n randomly selected data were excluded from the data set (including the compound currently studied). For all the properties, the residual is much smaller than the estimation errors reported in Table 6, confirming the reliability of the method.

Table 7 gives for each property the standard deviation, the average error, and the average absolute percentage error for the first- and second-order approximations. The number of experimental data used in the regressions is also reported. As mentioned in preceding sections and indicated by the reported statistics, the contribution of the second-order groups serve as a correction to the first-order approximation.

Results and Discussion

Comparisons with existing methods

The property predictions with the new methods after the second-level estimation are compared with several currently available group contribution based methods: the methods of Lydersen (1955), Ambrose (1978, 1980), Klincewicz and Reid (1984), and Joback and Reid (1984) for the estimation of critical properties; Joback and Reid (1984) for the estimation of the normal boiling point and normal melting point; Benson (1968) and Joback and Reid (1984) for the estimation of standard enthalpies and Gibbs energies of formation. The comparisons of the physical properties are reported in Tables 8 and 9. Table 10 presents the comparisons for thermodynamic properties. The statistics of the preceding methods have been obtained from the reported literature (Joback and Reid, 1984; Reid et al., 1987). It is worth mentioning that it would have been a fair comparison if the proposed model had been compared with other pure-predictive models with the same range of applicability and using a common set of reliable experimental data. However, such models cannot be found in the open literature. Also, a comparison of models with the same data set has difficulties since the models do not have the same range of application.

For the critical temperature, comparisons of the predicted property with the proposed method and those of Klincewicz and Reid (1984) and Lydersen (1955) show that significantly better results are obtained with the new method. Joback and Reid (1984) exhibits the same accuracy as this work. Also, the slightly better accuracy of the method of Ambrose (1978, 1980) when compared to the proposed method is nonappreciable since the accuracy of both methods is within the limits of experimental errors. Finally, compared with another purely predictive method (Fedors, 1982), the proposed technique performs much better. Attention should be focused on the fact that with the exception of the method of Fedors (1982), all the existing methods for the estimation of the critical temperature require reliable experimental value of the normal boiling point (a property which may not always be available). Therefore, the reliability and applicability of the estimations is restricted to those compounds where accurate experimentally-determined value of the normal boiling point is available (Joback and Reid, 1984). On the contrary, the new method can reliably estimate the critical temperature of any type of compound exclusively from the molecular structure. Such estimations are valuable for biochemical and environmentally-related processes which often involve large and complex-structured compounds. Furthermore, reliable value of critical temperature is required for the estimation of other properties, such as vapor pressure (Reid et al., 1987).

For critical pressure, an important property for the corresponding state principles, the estimates with the new method

Table 7. Comparison of the First- and Second-Order Approximations of the Method

Property X	Data Points	Standard Deviation		AAE		AAPE	
		1st-order	2nd-order	1st-order	2nd-order	1st-order	2nd-order
T_c	285	12.96 K	6.98 K	9.12 K	4.85 K	1.62%	0.85%
P_c	269	2.34 bar	2.02 bar	1.43 bar	1.13 bar	3.72%	2.89%
V_c	251	0.012 m^3/kmol	0.010 m^3/kmol	0.008 m^3/kmol	0.006 m^3/kmol	2.30%	1.79%
T_b	392	10.48 K	7.70 K	7.71 K	5.35 K	2.04%	1.42%
T_m	312	22.51 K	18.28 K	17.39 K	14.03 K	8.90%	7.23%
ΔH_v	225	2.20 kJ/mol	1.83 kJ/mol	1.40 kJ/mol	1.11 kJ/mol	3.22%	2.57%
ΔH_f	373	8.28 kJ/mol	6.12 kJ/mol	5.45 kJ/mol	3.71 kJ/mol	—	—
ΔG_f	333	7.25 kJ/mol	5.23 kJ/mol	4.78 kJ/mol	3.24 kJ/mol	—	—

$$\text{Standard Deviation} = \sqrt{\frac{\Sigma (X_{est} - X_{exp})^2}{N}}; \quad \text{AAE} = \frac{1}{N}\Sigma |X_{est} - X_{exp}|; \quad \text{AAPE} = \frac{1}{N}\Sigma \frac{|X_{est} - X_{exp}|}{X_{exp}} 100\%$$

where N is the number of experimental data, X_{est} is the estimated value of the property X, and X_{exp} the experimental value of the property X.

Table 8. Comparison of the Accuracy between Widely-Used Existing Methods and the Proposed for the Critical Properties

Method	T_c		P_c		V_c	
	AAE (K)	AAPE %	AAE (10^5 Pa)	AAPE %	AAE (10^{-6} m^3/mol)	AAPE %
Ambrose (1978, 1980)	4.3	0.7	1.8	4.6	8.5	2.8
Fedors* (1982)	—	5.0	—	—	—	3.15
Joback and Reid (1984)	4.8	0.8	2.1	5.2	7.5	2.3
Klincewicz and Reid (1984)	7.5	1.3	3.0	7.8	8.9	2.9
Lydersen (1955)	8.1	1.4	3.3	8.9	10.0	3.1
Proposed*	4.85	0.85	1.13	2.89	6.00	1.79

* Pure predictive methods.

Table 9. Comparison of the Accuracy Between Widely-Used Existing Methods and the Proposed for the Normal Boiling Point and Normal Melting Point

Method	T_b		T_m	
	AAE (K)	AAPE %	AAE (K)	AAPE %
Joback and Reid (1984)	12.9	3.6	22.6	11.2
Proposed	5.35	1.42	14.03	7.23

are drastically more accurate than all the others compared in this work. Also, for the estimation of the critical volume, as indicated in Table 10, the estimated values derived by the new method are clearly the most accurate.

In the case of the normal boiling point, the currently available group contribution methods (Joback and Reid, 1984) and correlations (Kreglenski and Zwolinski, 1961; Tsonopoulos, 1987) have been of limited accuracy or utility. Thus, the lack of fast and reliable prediction methods for the normal boiling point has been a major obstacle to the analysis and design of vapor-liquid equilibrium processes. The comparisons in Table 9 show the reliability and improvement in accuracy achieved in this method with respect to the only available method with wide range of applicability (a large number of correlations for the estimation of the normal boiling point are applicable to a specific homologous series, such as normal alkanes). For the normal melting point, group contribution methods, in general, have difficulties in providing a reliable estimation. The reason is that melting point highly depends upon intermolecular interaction and molecular symmetry (Reid et al., 1987; Horvath, 1992). Having this in mind, our method aimed to improve the accuracy of available methods and the comparisons are reported in Table 9.

Table 10 presents the comparisons for the thermodynamic properties. For the standard enthalpy of formation and the standard Gibbs energy, the proposed method performs decisively better than widely-used methods, even sophisticated ones (Benson et al., 1968). For the standard enthalpy of vaporization at 298 K, no other group contribution techniques could be

found in the literature (the methods reported in the literature estimate the standard enthalpy of vaporization at the normal boiling point). Note that with the above thermodynamic properties known at a given temperature (298 K in this case), simple correlations can be employed to determine these properties at other temperatures as well (Reid et al., 1987).

Internal consistency

Another issue on which little attention has been placed in the currently-used group contribution methods is the internal consistency of the limiting values (the values of the properties in a homologous series as the number of carbons tends to an infinite value) of estimated properties. Let us consider as an example the limiting behavior of critical density (inverse of the critical volume) and critical pressure of normal straight chained alkanes. For this purpose, the expression of the critical density and critical pressure (as reported in Table 5) have been reformulated to be functions of the number of carbons. A normal alkane with k number of carbons can be described by two first-order groups, that is, $-CH_3$ with 2 occurrences, and $-CH_2-$, with k-2 occurrences. Furthermore, the molecular weight of the alkane can be expressed as $14k+2$ (kg/kmol). Therefore, the critical density and pressure are given by the expression 3a and 3b, respectively (the critical density is estimated in kg/m^3, and the critical pressure in bar).

$$d_c = \frac{1}{V_c} = \frac{14k+2}{0.03421 + 0.05576k} \tag{3a}$$

$$P_c = 1.371 + \frac{1}{(0.11891 + 0.010558k)^2} \tag{3b}$$

As k tends to an infinite value, the limiting value for the critical density is 251 kg/m^3 and that of the critical pressure 1.371 bar, that is, both limits are nonzero values. Applying the same procedure, the limiting values of the currently-used group contributions have been obtained and reported in Table 11. Con-

Table 10. Comparison of the Accuracy Between Widely-Used Existing Methods and the Proposed for Thermodynamic Properties

Method	ΔH_f		ΔH_v		ΔG_f	
	AAE (kJ/mol)	AAPE %	AAE (kJ/mol)	AAPE %	AAE (kJ/mol)	AAPE %
Benson et al. (1968)	4.2	—	—	—	—	—
Joback and Reid (1984)	8.4	—	—	—	8.4	—
Proposed	3.71	—	1.11	2.57	3.24	—

Table 11. Limiting Values of the Critical Density vs. the Critical Pressure

Property Estimation Method	Limiting Value for dc (kg/m³)	Limiting Value for Pc (bar)
Ambrose (1978, 1980)	254.1	0
Joback and Reid (1984)	250.0	0
Klincewicz and Reid (1984)	253.2	0
Lydersen (1955)	254.5	0
Proposed Method	251.0	1.371

Table 12. Group Identification of Isomeric Dimethylhexanes

	2,3 Dimethylhexane	2,4 Dimethylhexane	2,5 Dimethylhexane
First-Order Groups			
CH₃	4	4	4
CH₂	2	2	2
CH	2	2	2
Second-Order Groups			
CH(CH₃)CH(CH₃)	1	—	—
(CH₃)₂CH	—	1	2

trary to the new method, the methods of Joback and Reid (1984), Lydersen (1955), Ambrose (1980), and Klincewicz and Reid (1984) assign finite limiting value to the critical density but zero limiting value to the critical pressure. The nonzero limiting values of both properties in our method seem to be more reasonable and in agreement with theoretical foundation (Tsonopoulos and Tan, 1993). Also, the limiting value of the critical pressure turns out to be based on the experimental data used in the regression. For instance, Tsonopoulos and Tan (1993) report two different limiting values for the critical pressure, 0.05 and 2.86 bar, by using two different data sets. Teja et al. (1990), on the other hand, give as limiting value 8.42 bar based on the regression of a selected experimental data set. Therefore, the limiting value reported in our method may be subject to changes in the future if reliable experimental data for heavy compounds suggest the need.

A second example of consistency, reliability, and wide range of applicability of the method is the study of the critical temperature and normal boiling point of normal alkanes up to one hundred carbon number. As shown in Figure 2, and in agreement with basic physical principles, throughout the homologous series, the normal boiling point smoothly approaches the critical temperature as the carbon number increases. At the same time, the ratio T_b/T_c is positive and less than unity, whereas widely-used methods, such as the method of Lydersen

(1955), assign negative values to T_b/T_c for normal alkanes with the number of carbons more than 77 (Tsonopoulos, 1987). Detailed and extensive analysis of the extrapolating behavior, together with comparisons, is beyond the scope of this work. (Current work is involved with this analysis.)

Properties of isomers

Another important characteristic of the proposed method is the ability to distinguish between isomers. Consider for example the estimation of the normal boiling point and critical temperature of isomeric dimethyl hexanes. Table 12 presents the first-order and second-order groups of these compounds. Tables 13 and 14 show the estimations through the method as opposed to experimental data. Application of any of the existing group contribution methods (even the most complicated ones) as well as the first level of the proposed method, would obtain the same values of the properties of these compounds because they consist of the same first-order groups. However, in the second level of the estimation in this technique, different second-order groups are incorporated. Therefore, the correction to the first-order approximation is achieved in such a manner that the estimated values exhibit differences similar to the experimental measurements. Note that these modest differences are important in various chemical engineering applications. For example, in parallel chemical reactions, a small difference in the properties of two compounds (such as enthalpy of formation) can decisively affect the selectivity of the process (Constantinou et al., 1993).
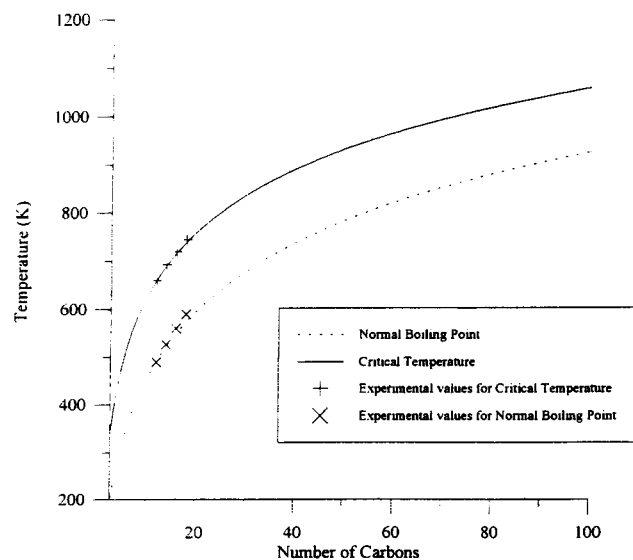

Figure 2. The normal boiling point vs. the critical temperature for normal alkanes (experimental values of heavy alkanes as reported in DIPPR are also presented).

Table 13. Critical Temperature of Isomeric Dimethylhexanes

Compound	Exp. Value (K)	First-Order Approx. (K)	Second-Order Approx. (K)
2,3-Dimethylhexane	563.40	557.91	566.60
2,4-Dimethylhexane	553.50	557.91	553.41
2,5-Dimethylhexane	550.00	557.91	548.81

Table 14. Normal Boiling Point of Isomeric Dimethylhexanes

Compound	Exp. Value (K)	First-Order Approx. (K)	Second-Order Approx. (K)
2,3-Dimethylhexane	388.76	385.93	391.41
2,4-Dimethylhexane	382.58	385.93	382.32
2,5-Dimethylhexane	380.26	385.93	378.64

Flexibility of the proposed model-application to existing group contribution approaches

The key element of the proposed model is that it allows the determination of the contributions of the second-order groups independently of that of first-order groups. Therefore, the model is applicable to existing group contribution approaches as well: the proposed second-order groups can be incorporated in existing group contributions where, in reality, only a first-level estimation was done. The values of their contributions can be carried out through regression analysis based on the existing first-order group contributions and experimental values (second-level estimation). The main goal is the enhancement of the accuracy of the methods by employing the corrective behavior of second-order groups and the distinguish among isomers as well. The idea can be implemented in both pure compound and mixture property estimation methods. Current work is devoted to the application of the preceding principles in widely-used methods for the estimation of activity coefficients. For example, in UNIFAC the estimation of the liquid-phase activity coefficients is the sum of two terms: the combinatorial and the residual (Fredenslund et al., 1977). The estimation of the former term involves pure-compound properties, that is, the Van der Waals Volume and Surface Area. With the contribution of second-order groups incorporated in the estimation of these properties, the accuracy of the combinatorial term can be improved enabling the method to capture the isomer differences and part of the proximity effects.

Conclusion

The employment of two sets of functional groups, one for a first-order approximation (basic level) and another one for a more accurate prediction (second-level) in a simple and widely-applicable model has led to a new group contribution method for the estimation of important physical and thermodynamic properties of pure compounds. With the proposed technique, inherent limitations of widely-used group contributions have been tackled. Distinction of isomers has been achieved and drastically better accuracy has been fulfilled through a pure molecular-based approach. Motivated by these excellent results, we are currently extending the method for the estimation of other properties of pure compounds, such as acentric factor, viscosity, liquid molar volume, vapor pressure, and solubility parameters. The performance of the model for properties of polymers will also be demonstrated. Challenging issues, such as the capture of isotope differences in simple engineering-oriented methods by introducing higher-level group contributions, are under investigation as well.

The successful applicability of the method, along with the elements that it introduces, highlights the significance of enriched and updated group contribution principles in providing simple yet accurate tools in the analysis of chemical, biochemical, and environmentally-related products and processes.

Acknowledgment

Notation

dc = critical density
ΔH_f = standard enthalpy of formation at 298 K
ΔH_v = standard enthalpy of vaporization at 298 K
g_0 = additional adjustable parameter used in estimation of the Gibbs energy of formation
g_{1i} = contribution of the type-i first-order group for estimation of the Gibbs energy of formation
g_{2j} = contribution of the type-j second-order group for estimation of the Gibbs energy of formation
ΔG_v = standard Gibbs energy of formation at 298 K
h_{f0} = additional adjustable parameter used in estimation of the enthalpy of formation
h_{1i} = contribution of the type-i first-order group for estimation of the enthalpy of formation
h_{2j} = contribution of the type-j second-order group for estimation of the enthalpy of formation
h_{v0} = additional adjustable parameter used in estimation of the enthalpy of vaporization
h_{v1i} = contribution of the type-i first-order group for estimation of the enthalpy of vaporization
h_{v2j} = contribution of the type-j second-order group for estimation of the enthalpy of vaporization
k = number of carbons in a normal alkane
M_j = number of occurrences of the type-j second-order group in a compound
N_i = number of occurrences of the type-i first-order group in a compound
p_{c1} = first additional adjustable parameter used in estimation of the critical pressure
p_{c2} = second additional adjustable parameter used in estimation of the critical pressure
p_{c1i} = contribution of the type-i first-order group for estimation of critical pressure
p_{c2j} = contribution of the type-j second-order group for estimation of critical pressure
P_c = critical pressure
t_{b0} = additional adjustable parameter used in estimation of the critical pressure
t_{b1i} = contribution of the type-i first-order group for estimation of the normal boiling point
t_{b2j} = contribution of the type-j second-order group for estimation of the normal boiling point
t_{c0} = additional adjustable parameter used in estimation of the critical temperature
t_{c1i} = contribution of the type-i first-order group for estimation of the critical temperature
t_{c2j} = contribution of the type-j second-order group for estimation of the critical temperature
t_{m0} = additional adjustable parameter used in estimation of the normal melting point
t_{m1i} = contribution of the type-i first-order group for estimation of the normal melting point
t_{m2j} = contribution of the type-j second-order group for estimation of the normal melting point
T_b = normal boiling point
T_c = critical temperature
T_m = normal melting point
v_{c0} = additional adjustable parameter used in estimation of the critical volume
v_{c1i} = contribution of the type-i first-order group for estimation of the critical volume
v_{c2j} = contribution of the type-j second-order group for estimation of the critical volume
V_c = critical volume
W = constant in the model with values zero or unity

Literature Cited

Ambrose, D., "Correlation and Estimation of Vapor-Liquid Critical Properties: I. Critical Temperatures of Organic Compounds," *NPL Rep. Chem.* 92, Nat. Physical Lab., Teddington, UK (1978).
Ambrose, D., "Vapor-Liquid Critical Properties: II. Critical Pressure

and Critical Volume," *NPL Rep. Chem. 107*, Nat. Physical Lab., Teddington, UK (1980).

Benson, S. W., *Thermochemical Kinetics*, Wiley, New York (1968).

Bondi, A., *Physical Properties of Molecular Crystals, Liquids and Glasses*, Wiley, New York (1968).

Congilio, L., E. Rauzy, and C. Berro, "Representation and Prediction of Thermophysical Properties of Heavy Hydrocarbons," *Fluid Phase Equilibr.*, **87**, 53 (1993).

Constantinou, L., R. Gani, A. Fredenslund, D. T. Wu, and J. A. Klein, "Computer-Aided Product Design: Problem Formulation and Application," *Proc. PSE Conf.*, **2**, 345, Korea (1994a).

Constantinou, L., R. Gani, and J. P. O'Connell, "Prediction of Primary and Secondary Properties through a New Group-Contribution Model," in preparation (1994b).

Constantinou, L., S. E. Prickett, and M. L. Mavrovouniotis, "Estimation of Properties of Acyclic Organic Compounds Using Conjugation Operators," *Ind. Eng. Chem. Res.*, **39**, 395 (1994c).

Constantinou, L., R. Gani, and J. P. O'Connell, "Estimation of the Acentric Factor and the Liquid Molar Volume at 298 K through a New Group Contribution Method," *Fluid Phase Equilib.*, in press (1994d).

Constantinou, L., and R. Gani, "A New Group Contribution Method for the Estimation of Properties of Pure Hydrocarbons," IVC-SEP 9319, Institut for Kemiteknik, The Technical Univ. of Denmark (1993).

Constantinou, L., "Estimation of Properties of Acyclic Organic Compounds through Conjugation," PhD Diss., Univ. of Maryland, College Park (1993).

Constantinou, L., S. E. Prickett, and M. L. Mavrovouniotis, "Estimation of Thermodynamic and Physical Properties of Acyclic Hydrocarbons Using the ABC Approach and Conjugation Operators," *Ind. Eng. Chem. Res.*, **32**(8), 1734 (1993).

Cox, J. D., and G. Pilcher, *Thermochemistry of Organic and Organometallic Compounds*, Academic Press, New York (1970).

Daubert, T. E., and R. P. Danner, *Physical and Thermodynamic Properties of Pure Compounds: Data Compilation*, Hemisphere, New York (1989).

Daubert, T. E., and R. Bartakovits, "Prediction of Critical Temperature and Pressure of Organic Compounds by Group Contribution," *Ind. Eng. Chem. Res.*, **28**, 638 (1989).

Derr, E. L., and C. H. Deal, Jr., "Analytical Solutions of Groups," *Inst. Chem. Eng. Symp. Ser. (London)*, **3**, 40 (1969).

Fedors, R. F., "A Method to Estimate Critical Volumes," *AIChE J.*, **25**(1), 202 (1979).

Fedors, R. F., "A Relationship between Chemical Structure and the Critical Temperature," *Chem. Eng. Commun.*, **16**, 149 (1982).

Fredenslund, Aa., J. Gmehling, and P. Rasmussen, *Vapor Liquid Equilibria using INIFAC*, Elsevier Scientific, Amsterdam (1977).

Fredenslund, Aa., and P. Rasmussen, "Correlation of Pure Component Gibbs Energy Using UNIFAC Group Contribution," *AIChE J.*, **25**(1), 203 (1979).

Gani, R., B. Nielsen, and Aa. Fredenslund, "A Group Contribution Approach to Computer-Aided Molecular Design," *AIChE J.*, **37**(9), 1318 (1991).

Gani, R., J. Perregaard, and H. Johansen, "Simulation Strategies for Design and Analysis of Complex Chemical Processes," *Trans. IChemE*, **68**, 407 (1990).

Gani, R., and J. P. O'Connell, "A Knowledge Based System for the Selection of Thermodynamic Models," *Comp. and Chem. Eng.*, **13**(4–5), 397 (1989).

Gray, R. D. Jr., J. L. Heidman, R. D. Springer, and C. Tsonopoulos, "Characterization and Property Prediction for Heavy Petroleum and Synthesis Liquids," *Fluid Phase Equilibr.*, **53**, 355 (1989).

Hansen, H. K., P. Rasmussen, Aa. Fredenslund, M. Schiller, and J. Gmehling, "Vapor-Liquid Equilibria by UNIFAC Group Contribution. 5. Revision and Extension," *Ind. Eng. Chem. Res.*, **30**, 2352 (1991).

Horvath, A. L., *Molecular Design*, Elsevier, Amsterdam (1992).

Jalowka, J. W., and T. E. Daubert, "Group Contribution Method to Predict Critical Temperature and Pressure of Hydrocarbons," *Ind. Eng. Process Des. Dev.*, **25**, 139 (1986).

Joback, K. G., and R. C. Reid, "Estimation of Pure-Component

Properties from Group Contributions," *Chem. Eng. Commun.*, **57**, 233 (1983).

Kehiaian, H. V., "Group Contribution Methods for Liquid Mixtures: A Critical Review," *Fluid Phase Equilibr.*, **13**, 243 (1984).

Klincewicz, K. M., and R. C. Reid, "Estimation of Critical Properties with Group Contribution Methods," *AIChE J.*, **30**(1), 137 (1984).

Kontogeorgis, G., I. F. Smirlis, V. I. Harismiadis, Aa. Fredenslund, and D. P. Tassios, "Vapor Pressure and Saturated Liquid Volume Predictions with Cubic Equations of State for Medium and High Molecular Weight Compounds," IVC-SEP 9329, Institut for Kemiteknik, The Technical Univ. of Denmark (1993).

Kreglewski, A., and B. J. Zwolinski, "A New Relation for Physical Properties of n-Alkanes and n-Alkyl Compounds," *J. Phys. Chem.*, **65**, 1050 (1961).

Kudchadker, A. P., and B. J. Zwolinski, "Vapor Pressures and Boiling Points of Normal Alkanes, C21 to C100," *J. Chem. Eng. Data*, **11**(2), 253 (1966).

Leibovici, C. F., "A Consistent Procedure for the Estimation of Properties Associated to Lumped Systems," *Fluid Phase Equilibr.*, **87**, 189 (1993).

Lydersen, A. L., "Estimation of Critical Properties of Organic Compounds," *Coll. Eng. Univ. Wisconsin, Engineering Experimental Station Rept. 3*, Madison, WI (1955).

Lyman, W. J., W. F. Reehl, and D. H. Rosenblatt, *Handbook of Chemical Property Estimation Methods*, Amer. Chem. Soc., Washington, DC (1990).

Magoulas, K., and D. P. Tassios, "Thermophysical Properties of n-Alkanes from C1 to C20 and their Prediction to Higher Ones," *Fluid Phase Equilibr.*, **56**, 119 (1990).

Mavrovouniotis, M. L., "Estimation of Properties from Conjugate Forms of Molecular Structures: The ABC Approach," *Ind. Eng. Chem. Res.*, **32**, 1734 (1990).

McCann, D. W., and R. P. Danner, "Prediction of Second Virial Coefficients of Organic Compounds by a Group Contribution Method," *Ind. Eng. Chem. Process Des. Dev.*, **23**, 523 (1984).

Nikitin, E. D., P. A. Pavlov, and N. V. Bessonova, "Critical Constants of n-Alkanes from 17 to 24 Carbon Atoms," *J. Chem. Thermodynamics*, **26**, 178 (1994).

O'Connell, J. P., personal communication (1994).

Prickett, S. E., L. Constantinou, and M. L. Mavrovouniotis, "Computational Identification of Conjugate Paths for Estimation of Properties of Organic Compounds," *Molec. Simulation*, **11**(2–4), 205 (1993).

Reid, R. C., J. M. Prausnitz, and B. E. Poling, *The Properties of Gases and Liquids*, 4th ed., McGraw-Hill, New York (1987).

Reid, R. C., J. M. Prausnitz, and T. K. Sherwood, *The Properties of Gases and Liquids*, 3rd ed., McGraw-Hill, New York (1977).

Siepmann, J. I., S. Karaborni, and B. Smit, "Simulating the Critical Behavior of Complex Fluids," *Nature*, **365**, 330 (1993).

Somayajulu, G. R., "Estimation Procedures for Critical Constants," *J. Chem. Eng. Data*, **34**, 106 (1989).

Teja, A. S., R. J. Lee, D. J. Rosenthal, and M. Anselme, "Correlation of the Critical Properties of Alkanes and Alkanols," *Fluid Phase Equilibr.*, **56**, 153 (1990).

Tsonopoulos, C., "Critical Constants of Normal Alkanes from Methane to Polyethelene," *AIChE J.*, **33**(12), 2080 (1987).

Tsonopoulos, C., and Z. Tan, "The Critical Constants of Normal Alkanes from Methane to Polyethelene. II. Application of the Flory Theory," *Fluid Phase Equilibr.*, **83**, 127 (1993).

Voulgaris, M. S., S. Stamatakis, K. Magoulas, and D. P. Tassios, "Prediction of Physical Properties of non-Polar Compounds, Petroleum and Coal Liquid Fractions," *Fluid Phase Equilibr.*, **64**, 73 (1991).

Wu, S. E., and S. I. Sandler, "Proximity Effects on the Prediction of the UNIFAC Model: I. Ethers," *AIChE J.*, **35**(1), 168 (1989).

Wu, S. E., and S. I. Sandler, "Use of *ab Initio* Quantum Mechanics Calculations in Group Contribution Methods: 1. Theory and the Basis for Group Identifications," *Ind. Eng. Chem. Res.*, **30**, 881 (1991).

Wu, S. E., and S. I. Sandler, "Use of *ab Initio* Quantum Mechanics Calculations in Group Contribution Methods: 2. Test of New Groups in UNIFAC," *Ind. Eng. Chem. Res.*, **30**, 889 (1991).

Table 15a. Application Examples

(1) $CH_3C(OH)(CH_3)CH_2CH(OH)CH_3$			(2) $CH_3CH(CH_3)CH_2COOH$		
First-Order Groups	Occurrences	Contribution	First-Order Groups	Occurrences	Contribution
CH_3	3	0.8894×3	CH_3	2	0.8894×2
CH_2	1	0.9225×1	CH_2	1	0.9225×1
CH	1	0.6033×1	CH	1	0.6033×1
C	1	0.2878×1	$COOH$	1	5.8337×1
OH	2	3.2152×2			$\Sigma N_i t_{b1i} = 9.1383$
		$\Sigma N_i t_{b1i} = 10.9122$			

Second-Order Groups	Occurrences		Second-Order Groups	Occurrences	Contribution
COH	1	-0.6331×1	$(CH_3)_2CH$	1	-0.1157×1
$CHOH$	1	-0.5385×1			$\Sigma M_j t_{b2j} = -0.1157$
		$\Sigma M_j t_{b2j} = -1.1716$			

Table 15b. Application Examples

(1) $CH_2 = C(CH_3)CH = CH_2$			(2) $4(ACH)(ACCH_3)(ACCH_2)(CH_3)$		
First-Order Groups	Occurrences	Contribution	First-Order Groups	Occurrences	Contribution
$CH_2 = C$	1	88.402×1	CH_3	1	-8.030×1
$CH_2 = CH$	1	84.926×1	$ACCH_2$	1	41.228×1
CH_3	1	-8.030×1	$ACCH_3$	1	22.505×5
		$\Sigma N_i g_{b1i} = 165.298$	ACH	4	22.533×4
					$\Sigma N_i g_{b1i} = 145.835$

Second-Order Groups	Occurrences				
$CH_n = CH_m - CH_p = CH_k$	1	-5.050×1	No second-order groups are involved		
		$\Sigma M_j g_{2j} = -5.050$			

Table 16. First- and Second-Order Estimations of the Compounds of Tables 15a and 15b

	First-Order Approx.	Second-Order Approx.	Exp. Value (% Error)	Existing Methods (Reid et al., 1987)
(1)	$T_b = 204.359 \ln (10.9122)$ $= 488.39$ K	$T_b = 204.359 \ln (10.9122 - 1.1716)$ $= 465.18$ K	470.65 K (1.2%)	517.37 (9.9%)
(2)	$T_b = 204.359 \ln (9.1383)$ $= 452.14$ K	$T_b = 204.359 \ln (9.1383 - 0.1157)$ $= 449.53$ K	448.25 K (0.3%)	458.87 (2.4%)
(3)	$\Delta G_f = -14.828 + (165.298)$ $= 150.472$ kJ/mol	$\Delta G_f = -14.828 + (165.298 - 5.505)$ $= 144.965$ kJ/mol	145.90 kJ/mol (0.6%)	158.35 kJ/mol (8.5%)
(4)	$\Delta G_f = -14.828 + (145.835)$ $= 131.007$ kJ/mol	It does not exist	131.20 kJ/mol (0.1%)	127.68 kJ/mol (2.7%)

Appendix A

To illustrate the proposed technique and demonstrate its simplicity and accuracy, we provide the estimation of the normal boiling point of 2-Methyl-2,4-pentanediol (1), 3-methyl butanoic acid (2), as well as the Gibbs Energy of Formation at 298 K of 2-Methyl-1,3-Butadiene (3), and 1-Methyl-2-Ethyl Benzene (4). The experimental data and estimations of existing widely-used methods are also reported (see Tables 15a, 15b and 16).