[Regular Paper]

A Physicochemical Property Estimation System with Automatic Route Selection

Delin QU^{†1)}, Teng TENG^{†1)}, Masaaki MURAKI^{†2)*}, and Toyohiko HAYAKAWA^{†3)}^{†1)} Dept. of Chemical Engineering, Tsinghua University, Beijing, China 100084^{†2)} Dept. of Industrial Engineering and Management, Tokyo Institute of Technology,
2-12-1 Ookayama, Meguro-ku, Tokyo 152^{†3)} Dept. of Industrial Engineering, Chiba Institute of Technology,
Tsudanuma, Narashino, Chiba 275

(Received January 27, 1994)

In most of the present physicochemical property estimation systems (PPESs), as estimation route, which involves an estimation method and its intermediate estimation methods, is selected by the user. Since many elements such as application domain, accuracy, and availability of data required for intermediate property estimation have their influences on the route selection, it is difficult for the user to comprehensively determine a route considering all these constrained elements. Therefore, an advanced PPES is proposed. It can automatically select a proper estimation route, moreover, it can extend easily to other physicochemical properties adding relevant modules. This PPES is composed of two procedures: routing and sequencing. The former is to choose an estimation route, and it is realized by introducing the priority rule based on the application domain and the module structure composed of control and estimation sections. The control section offers the information about the connection relationship among the estimation methods and the constrained elements, and the estimation section implements the relevant estimation like the module of existing PPESs. The latter (sequencing) is performed by the relevant sequencing module in which the integrative feasibility of the selected route is examined through the detection of a loop, and the calculation order is determined on the basis of mathematical relationship. The average time increase in execution to provide automatic route selection is about 15% of the physicochemical property estimation for typical samples.

1. Introduction

One of the main currents of Computer-Aided Process Engineering (CAPE) is process simulation, and process simulators at present spend more than 80% of their central processing unit (CPU) time in physicochemical property calculations¹⁾. The effectiveness of physicochemical property estimations depends heavily on physicochemical property database systems, and up to the present day, simulators composed of basic Physicochemical Property Database (PCPDB) and of Physicochemical Property Estimation System (PPES) have been developed mostly for the steady-state flowsheet simulation programs^{2),3)}. The problems of basic data storage in the PCPDB have already been solved by logical and physical database design methods^{4),5)}; thus, improving the quality of results and computing performance in the PPES is the remaining problem.

Most of the requested intensive and extensive physicochemical properties can be calculated by several different estimation methods. For instance, there are two methods for computing enthalpy deviation. One is to use the Redlich-Kwong-Soave (RKS) equation, and the other is to rely on the BWR equation of state. It is necessary and important to make a proper choice between the two. Unfortunately, this choice is complicated because each of them requires some intermediate properties which are the results calculated by other estimation methods. In the case of using RKS equation, the acentric factor as an intermediate property needs to be computed by one of the three methods which are Pitzer correlation, Edmister and Lee-Kesler equations. In general, each physicochemical property estimation can be well represented by a routing tree of modules⁶⁾, corresponding to the estimation methods, which involve estimation method, relevant intermediate estimation methods, and calculation order.

In most of the present PPESs, there are two ways

* To whom correspondence should be addressed.

to provide a choice of estimation route for the user. One is to offer options of the relevant estimation methods, and then to construct an estimation route based on the user's choice⁷⁾. For instance, in FLOWTRAN[®] it is possible to choose a method for computing vapor-liquid equilibria. The other is to provide the user with a routing plexus structure, which can specify a particular estimation route^{3), 6), 7)}, and may give flexibility in specifying the route to the user. The estimation method is usually constrained by many elements such as its application domain, substance system, phase state, accuracy and data availability of the relevant intermediate physicochemical properties. It is then necessary for the user to select a method under comprehensive consideration of these elements. Recently, in order to help the user's choice, a prototype for a limited expert system to select *K*-value estimation methods was developed⁹⁾. However, not only the relevant knowledge base and inference mechanism but also an interactive program is necessary to extend it to other properties.

The purpose of this study is to develop an advanced PPES which can automatically estimate physicochemical properties with reasonable accuracy and can easily extend to other physicochemical property estimations. Therefore, it is necessary for the above-mentioned PPES to realize the following two functions in consideration of the ease of such extension. The first is to choose a proper estimation method and to constitute intermediate estimation methods automatically. The second is to examine the integrative feasibility of this choice and to determine the calculation order. We propose the PPES which is composed of two procedures (routing and sequencing) corresponding to these two functions. Furthermore, in order to improve the flexibility in the extension to other physicochemical properties, we adopt a modular structure for each estimation method, which has control and estimation sections. The control section offers only the information about the connection relationship between the estimation method and its intermediate estimation methods and the constrained elements. The estimation section corresponds to the module of the existing PPESs, and its information is used to implement relevant property estimations.

Using these two procedures and the control section of each module, the proposed PPES can automatically determine a proper estimation route of required accuracy and easily extend itself to other physicochemical properties. The average time increase in automatic estimation route selection is about 15% of property calculations for

typical samples. The effectiveness of this proposed PPES is illustrated with examples.

2. Basic Structure of PPES

Process simulation requires thermodynamic data (such as vapor/liquid equilibria and enthalpies) and transport properties (such as viscosity, conductivity, and surface tension) in large quantities^{10), 11)}. These properties usually cannot be stored in the PCPDB directly; they need to be calculated by physicochemical property estimation methods with the basic data stored in the PCPDB. Sometimes several methods are available to estimate the same physicochemical property, and each method has its own route characteristic. The concept of the routing tree⁶⁾ is to indicate the relationship between estimation objects and methods visibly. For example, a routing tree for the *K*-value estimation is shown in Fig. 1, where the estimation objects represented by circles take "logical or operations," the estimation methods represented by rectangles take "logical and operations," and the lines connecting the estimation objects and methods represent estimation routes, which show the alternatives to be selected in the PPES.

Up to now, some of the PPESs adopt PLEXUS routing structure^{6), 7)} that provides all of the theoretically feasible estimation routes among which the user needs to select a proper one. However, it is difficult for the user to comprehensively consider all the constrained elements, so that the selected estimation route may not be always feasible. Furthermore, the selected route may fail in practical calculation because the data required for some intermediate estimation methods or basic data are not available. In order to select an integrally feasible estimation route from the alternatives such as shown in Fig. 1, two procedures of routing and sequencing are indispensable. The former is to choose a proper estimation method with its intermediate estimation methods among many alternatives under the consideration of the connecting relationship between these methods and the constrained elements mentioned above automatically, and the latter (sequencing) is to examine the integrative feasibility of the aforementioned choice based on the availability of the data and to determine the calculation order. For implementing them, the basic structure of the proposed PPES is shown in Fig. 2. Routing and sequencing are represented by the encircled dash lines separately. Details of the procedures are described in the following sections, and therefore only an outline of the

Fig. 1 Routing Tree for Estimating K-Value

Fig. 2 The Process of Automatical Route Choosing

proposed PPES is explained here.

Since in routing, it is difficult to complete this kind of choice with once-through operation, the strategy of repeating the following two steps is

proposed. In the first step, the priority rule based on the application domain which is the key information for determining the local feasibility is adopted, and the estimation method with the highest local feasibility is chosen. In the second step, using the information stored in the control section of the individual module, the substance system and phase state are examined and then the accuracy required is judged for the selected method. If this method is not feasible in accordance with these previous conditions, return to the first step and choose an alternative with a secondary highest local feasibility. The estimation method can be determined by the joint action of the priority rule and the control section, and the estimation route can be determined through the chain-branching provided by the control sections of the relevant estimation modules.

In sequencing, two functions are necessary based on the requirements of the PPES. The first is to examine the integrative feasibility by detecting whether or not there exists a loop in the selected estimation route, because the lack of data required in estimation of some intermediate properties produces a loop in the route. The second is to determine the calculation order based on the mathematical relationship between the selected estimation method and its intermediate estimation methods. If a loop is detected, return to routing in order to find an alternative route; otherwise, the calculation order is determined by the sequencing module which is an independent and general program. This sequencing module is called by the control section of the relevant estimation module. As a result, the estimation route obtained is integrally feasible. If there is no alternative without a loop in the route, man-machine interactive manner is applied to continuing this estimation. In this PPES, the advanced modular structure composed of control and estimation sections is adopted to improve the flexibility in extension^{9)~12)}, and to ensure a reasonable route selection with integrative feasibility. Based on this modular design, an estimation route for the required property can be automatically assembled.

3. Routing

Routing is to automatically choose an estimation route with reasonable accuracy. It differs from the procedure of first collecting all feasible routes and then identifying the optimal one on the user's choice^{6),13),14)}. Furthermore, many new estimation methods have been proposed along with the development of chemical thermodynamic models; then it is necessary for routing to consider

Table 1 The Priority for Estimating Activity Coefficients in the Liquid Phase

Estimation method	Application domain	Priority order (from high to low)
UNIFAC	High nonideal mixtures including strongly polar or hydrogen bonded component, polymers or electrolytes; be applicable to liquid-liquid (L-L) and vapor-liquid systems (V-L).	1
ASOG	High nonideal mixtures including strongly polar or hydrogen bonded component, electrolytes; be applicable to L-L, V-L systems, temperature range 303-423 K.	2
UNIQUAC	For binary and multicomponent mixtures containing a variety of nonelectrolyte components; be applicable to L-L, V-L systems.	3
NRTL	For binary mixtures excluding electrolyte and polymers, be applicable to L-L, V-L binary systems.	4
WILSON	Be not applicable to L-L system; a modified form of WILSON's equation can be successfully represent equilibria in polymer solution.	5
VAN LAAR	Be only applicable to V-L system for binary mixture excluding electrolyte, polymers and strongly hydrogen bonded component; the temperature dependence of A and B may be neglected with little loss in accuracy.	6
MARGULAS	Be only applicable to V-L system and interactive parameters A and B are functions of temperature. And the relation should be used only for isothermal data; for binary mixtures excluding electrolyte, polymers, and strongly hydrogen bonded components.	7

not only the constrained elements for choosing a proper method but also the structure of the PPES for extending the user's choice. In routing, we propose a priority rule for estimation methods and a control section for each module.

3. 1. Priority Rule

The application domain is to be the key information of local feasibility among other constrained elements such as substance system, phase state, accuracy and data availability of relevant intermediate physicochemical properties. Based on this domain, the priority rule is established for each property estimation method of both chemical species and their mixtures. According to this rule, we give a higher priority to the estimation method with a wider application domain because of its statistical height of local feasibility. Priority is then arranged in the order of decreasing application domain. The estimation methods for physicochemical properties are divided into two classes. One is to use physicochemical property estimation equations, and the other is to rely on group contribution methods. For instance, there are several methods for estimating activity coefficients in the liquid phase, and the relevant priority order is determined according to the previous principle as shown in **Table 1**, where the group contribution methods such as UNIFAC and ASOG possess high priority. According to this approach, the individual priority order is established for each physicochemical property estimation method. Then based on this order, the estimation method can be chosen tentatively in routing.

3. 2. Control Section

The priority rule only provides the basis of local feasibility for choosing an estimation method.

However, in order to obtain a reasonable estimation method and route, there are still two requirements. The first is to consider the alternative constrained elements which may affect the estimation method selection. The second is to identify the connecting relationship between the estimation method and its intermediate estimation methods which may affect the estimation route. The implementation of these two requirements depends on the information related to the estimation method, and then we propose an advanced modular structure that not only provides this information but also improves flexibility in extension. Each estimation method is designed to a separate module pre-processed as an executable version, which is composed of control and estimation sections. The control section provides the following information.

1) Applicable substance system and accuracy.

2) Suitable phase state.

3) Requested intermediate properties list for this property estimation.

The estimation section is used for calculating property values and it corresponds to the typical module of the existing PPESs. The control section adopts the structural programming and it is independent of the control sections of other estimation methods. Accuracy is related to the application substance system and phase state, and in this study it is divided into three classes (good, better, and fair) according to the up-to-date published¹⁰. Considering the priority rule and the information provided by the control section of the individual module, an estimation method not only with the highest local feasibility but also with required accuracy is determined. Because the determined method usually requires several

Table 2 List of Estimation Files for General Compounds

File name	Function	Number of estimation methods
m11	normal boiling point	2
m12	normal melting point	1
m13	critical temperature	5
m14	critical pressure	3
m15	critical volume	3
m16	critical compressibility factor	2
m17	acentric factor	3
m18	latent heat of vaporization at T_b	6
m19	vapor pressure	7
m20	heat of vaporization for pure component	7
m21	heat capacity of liquid	3
m22	molecular volume of liquid	2
m23	compressibility factor	3
m24	fugacity coefficient	4
m25	activity coefficient	7
m26	viscosity for vapor	12
m27	viscosity for liquid	11
m28	enthalpy deviation	2
m29	entropy deviation	2
m30	heat capacity for vapor at constant pressure	1
m31	standard state fugacity	1
m32	thermal conductivity for vapor	12
m33	thermal conductivity for liquid	11
m34	diffusion coefficient for vapor	3
m35	diffusion coefficient for liquid	7
m36	surface tension	11
m37	K -value	2

intermediate physicochemical properties, their estimation methods can be also determined by the same way. Consequently, an estimation route is defined. In addition, different estimation methods for the same estimation object are put into one estimation file for convenience of choosing, and a list of all estimation files is shown in **Table 2**.

4. Sequencing

There are several estimation routes available for the same physicochemical property, but they may not be feasible in practice because they may have loops. This looping is caused by two reasons. One is the complicated relationship between the required physicochemical property and its intermediate properties. The other is the nonexistence of data required for some intermediate physicochemical properties. In order to clarify this source, the process for estimating intermediate properties is termed a descendant, while the process that requires the descendant is termed a precursor. If the route is not assembled properly, it may lead to the possibility that the source data of the descendant are just the objective data of its precursor. As a result, a loop appears in the estimation route causing some calculation failures in the estimation. For instance, there is a route for estimating the acentric factor, as shown by bold line in **Fig. 3**. Lee-Kesler equation(2) is chosen as the estimation method in routing, and such intermediate properties as critical pressure P_c ,

Fig. 3 Routing Tree for Estimating Acentric Factor

critical temperature T_c , and enthalpy of vaporization H_v are required among which P_c and T_c are known as the basic data in the PCPDB, and only H_v needs to be estimated. According to the routing procedure for H_v , Pitzer correlation method is chosen in which the acentric factor, which is just the objective property, is required. Apparently, this route is not feasible integrally because there is a loop in the route. Therefore, it is not sufficient for an efficient estimation system to have only the route choosing the function. Identifying the loop

and determining the integrally feasible route are also important functions.

These two functions can be implemented by the proposed sequencing procedure, which is distinguished from the general conception of ordering. This procedure examines whether or not there is a loop in the estimation route. If a loop is detected, it indicates that the estimation route is infeasible; otherwise, the route is reasonable. For the infeasible case, it is necessary to return to the routing procedure for choosing an alternative route. Generally, searching for such alternative route can be carried out in the order of decreasing the estimation level. Actually, sequencing is to examine the integrative feasibility of the estimation route selected in routing. If there is no proper alternative route without a loop, the automatic manner has to return to the man-machine interactive manner to avoid looping in which the user is required to input necessary information for continuing the estimation. However, for the feasible case, sequencing can arrange the calculation order in which the estimation is performed.

Sequencing procedure is implemented by the sequencing module, which is an independent and general program. The control section of the relevant estimation module can call upon the sequencing module to implement sequencing, and the sequencing module works based on the following information:

- (1) estimation method, objective physicochemical properties and other required intermediate physicochemical properties (the first level property estimation).
- (2) estimation methods for intermediate physicochemical properties and required intermediate physicochemical properties (the second level property estimation) and so on.
- (3) mathematical relation between these physicochemical properties.

5. Implementing the Estimation

In order to illustrate the whole process, estimating the vapor-liquid equilibrium ratio K_i of acetone(1)-*n*-hexane(2) system at 55°C and $x_1=0.1846^{15}$, is taken as an example referring to **Fig. 1**. Based on the priority rule, the priority of "activity coefficient method" is located in a higher position than that of "state equation method." According to the information provided by the control section of the K -value estimation module, the former is determined; then its intermediate properties (liquid-phase activity coefficient γ_i , standard state fugacity f_i^s , and vapor-phase fugacity

coefficient ϕ_i^v), which are termed second level estimation, would be required. UNIFAC is determined for estimating γ_i because of its priority based on **Table 1**. Then, the feasibility of UNIFAC for the specified binary (acetone-*n*-hexane) system should be examined. According to the control section of the UNIFAC module, both the substance system (acetone-*n*-hexane) and the phase state (vapor-liquid equilibrium) are suited to the domain of its estimation equation. Moreover, it possesses high accuracy under these conditions, and UNIFAC is then determined for calculating activity coefficient γ_i . In accordance with the way mentioned above, vapor-phase fugacity coefficient ϕ_i^v is estimated with the second order virial equation of state¹⁶. However, standard state fugacity f_i^s has only one thermodynamic relation as shown in **Fig. 1**. Then, its intermediate physicochemical properties labeled as the third level estimation includes vapor pressure p_i^s of pure component i at temperature T , fugacity coefficient ϕ_i^s at p_i^s and T , and liquid molar volume V_i . According to the same procedure, ϕ_i^s can be estimated with the same method of ϕ_i^v and p_i^s and V_i that can be estimated with Antoine and Guun-Yamada methods, respectively. Routing is then determined and represented by the thick bold line in **Fig. 1**. After routing, using the sequencing procedure no loop is detected, and the estimation sequence is determined as V_i , ϕ_i^s and $p_i^s-f_i^s$, ϕ_i^v and γ_i-K_i . Finally, K_1 and K_2 are estimated as 2.5973 and 0.6384, respectively, and the value measured for K_1 is 2.5785 at total pressure of 780.9 mmHg¹⁵, which is close to the estimated one.

For the example shown in **Fig. 3**, the control section of module "Lee-Kesler(2)" calls the sequencing module. Because the physicochemical property required for the second level estimation (Pitzer correlation method) is the objective property of the first level estimation, a loop is detected. Then, an alternative routing needs to be searched for again. Based on the routing procedure, the estimation method of Lee-Kesler(1) is chosen instead of Lee-Kesler(2), and the loop can be avoided. Therefore, the estimation sequence can be determined as $T_b-\omega$, in which T_b is first estimated with Lydersen method, and then the estimation is implemented with Lee-Kesler(1). The combined estimation route is shown with dashlines in **Fig. 3**.

The PPES has been developed in Fortran 77 and C languages under the UNIX operating system, which efficiently implements the automatic route selection and improves the flexibility in extension. With the cooperation of the man-machine interac-

tive manner, the proposed PPES can realize all required property estimations. Twenty six typical physicochemical properties shown in **Table 2** were tested, and the average overhead (the time spent in routing and sequencing procedures), which is associated with the automatic route selection, was 15% of the execution time, and it is superior to the 27% of others^{5)~7)}. Through the efficient interface design, the user cannot only use the application programs but also link them to the PPES.

6. Conclusion

An advanced PPES, which includes routing and sequencing procedures with the modules composed of control and estimation sections, is proposed for automatically choosing a reasonable estimation route and easily extending it to other physicochemical properties. In this PPES, two techniques, which establish the priority rule and construct the control section in each module, are developed for routing. Sequencing is to examining the integrative feasibility of the estimation route selected in routing by detecting a loop and to determine the calculation order, and this is implemented by the sequencing module called by the relevant control section. By the design of this advanced PPES, high flexibility in extension can be realized, and the creation of estimation routes for specific properties can be promoted with

modules in the program library. The effectiveness of this structure is illustrated by examples.

References

- 1) Winter, P., *CEP*, **88**, (2), 76 (1992).
- 2) Donaldson, R. W., *CEP*, **81**, (9), 14 (1985).
- 3) Vecchiotti, A., Leone, H., Melli, T., Montagna, J., Cerro, R., *Comput. Chem. Eng.*, **11**, 519 (1987).
- 4) Encarnacao, J., "Computer Aided Design," Springer-Verlag, Berlin, Heidelberg, New York (1982), p. 227.
- 5) Qu, D-L., Zhao, K-Y., Muraki, M., Hayakawa, T., *Sekiyu Gakkaishi*, **36**, (5), 347 (1993).
- 6) Seider, W. D., Evans, L. B., *Ind. Eng. Chem., Process Des. Dev.*, **18**, 292 (1979).
- 7) Evans, L. B., Seider, W. D., Joseph, B., *AIChE J.*, **23**, 658 (1977).
- 8) Seader, J. D., Seider, W. D., Pauls, A. C., "Flowtran Simulation, An Introduction," 2nd ed., CACHE: Ann Arbor, (1977).
- 9) Banares, A. R., Westerberg, A. W., Rychener, M. D., *Comput. Chem. Eng.*, **9**, 127 (1985).
- 10) Reid, R. C., Prausnitz, J. M., Poling, B., "The Properties of Gases and Liquids," McGraw-Hill, New York (1987).
- 11) Richard, S. H. Mah., "Foundations of Computer-Aided Chemical Process Design," Vol. I., Engineering Foundation, New York (1981).
- 12) Evans, L. B., Seider, W. D., *CEP*, **72**, (6), 80 (1976).
- 13) Meadows, E. L., *CEP*, **61**, (5), 93 (1965).
- 14) Norris, R. C., *CEP*, **61**, (5), 96 (1965).
- 15) Kojima, K., Tochigi, K., "ASOG and UNIFAC," Chemical Engineering, Tokyo (1987).
- 16) Hayden, J. G., O'Connell, J. P., *IEC Proc. Des. Dev.*, **14**, 209 (1975).

要 旨

ルート選択機能を持つ物性推算システム

曲 徳林^{†1)}, 膳 藤^{†1)}, 村木正昭^{†2)}, 早川豊彦^{†3)}

^{†1)} 清華大学化学工学科, 100084 中国北京市清華園

^{†2)} 東京工業大学経営システム工学科, 152 東京都目黒区大岡山 2-12-1

^{†3)} 千葉工業大学経営工学科, 275 千葉県習志野市津田沼 2-17-1

多くの物性推算システム (PPES) では, 物性推算ルートはユーザーにより選択されている。しかし, 物性推算ルートの選択は, 適用範囲, 精度, 計算途中で必要となるデータの利用の可否などの多くの要素に影響され, ユーザーがこれらの要素を総合的に考慮してルートを決めることは困難である。そこで, 本研究では合理的な推算ルートを自動的に選択でき, かつモジュールの追加により他の物性推算に拡張が容易な二段階 (ルーティング, シーケンシング) からなる PPES を提案した。第一段階

(ルーティング) では, 適用範囲に基づく優先ルールと制御と推算部分からなるモジュール構造を利用して推算ルートを選択する。前者は推算方法間とルート選択に影響する要素に関する情報を持ち, 後者は従来の物性推算におけるモジュールに相当する。第二段階 (シーケンシング) では選択された推算ルートの総合的な実行可能性と推算手順を確定する。自動ルート選択に要する計算時間は平均して約 15% の増加であった。

.....

Keywords

Process simulation, Property estimation, Routing, Sequencing, Priority rule, Control section