

A Computer-Based Methodology for Matching Polymer Structures with Required Properties

GEORGE C. DERRINGER and RICHARD L. MARKHAM, *Battelle-Columbus Laboratories, Columbus, Ohio 43201-2963*

Synopsis

A methodology is presented in which a computer generates polymer structures and estimates their properties using empirical relationships. An optimization method is then employed to select the candidate polymers which best comply with a set of property specifications.

INTRODUCTION

A common problem in polymer science is that of finding a polymer which can meet a number of property constraints. For example, a polymer might be desired exhibiting high modulus, low density, low dielectric strength, etc. When only a few property constraints exist, a suitable candidate can often be found from textbook property tabulations or individual experience. However, when several property constraints must be simultaneously met or when constraints are difficult to meet with existing polymers, the only recourse may be a costly review of the literature and/or a polymer synthesis/evaluation study. The purpose of this paper is to illustrate how viable polymer candidates may be found by computer calculation of approximate polymer properties using empirical equations from the literature. The proposed methodology is illustrated with an example.

PREDICTING POLYMER PROPERTIES FROM CHEMICAL STRUCTURES

Recent work by Van Krevelen,¹ McGinniss,² and others has demonstrated that many polymer properties can be estimated by semiempirical relationships involving only the chemical structure of the polymer. In these approaches, the basic structural unit is divided into smaller groups of atoms. A knowledge of the contribution of each of the groups to a given polymer property is then used to predict the property of a polymer composed of the given structural repeat unit. For example, in one of Van Krevelen's numerous examples, the density of poly(ethylene terephthalate) was calculated based solely on group contributions of $-\text{C}_4\text{H}_6-$, $-\text{COO}-$, and $-\text{CH}_2-$, which are the building blocks of the basic structural unit.

Predicting Polymer Structure to Meet Required Properties

Van Krevelen's book¹ alone contains empirical equations for many of the most important properties of a wide variety of polymers. It should therefore be possible to find candidate polymers to meet a set of property constraints

simply by carrying out the necessary calculations. Since the number of calculations is large this would ideally be done on a computer. A flow chart illustrating one possible scheme is shown in Figure 1. Here the first step is the selection of the number of groups in the polymer repeat unit. Then from a table of chemical groups such as $-\text{CH}_2-$, $-\text{COO}-$, $-\text{CHCl}-$, etc., the appropriate number of groups is selected either randomly or according to a systematic selection procedure. In step 3 the properties of interest are calculated from group contribution data stored in the computer. In step 4 each property is checked against the corresponding constraint. If all constraints are met, the repeat unit is printed along with calculated properties. If all constraints are not met the process is repeated from step 1 or 2.

Example of Polymer Structure Prediction

A computer program was written to find viable polymer structures for meeting constraints involving density, water absorption, and glass transition temperature T_g . The groups comprising the repeating structural unit were limited to $-\text{CH}_2-$, $-\text{CO}-$, $-\text{COO}-$, $-\text{O}-$, $-\text{CONH}-$, $-\text{CHOH}-$, and $-\text{CHCl}-$. The contributions of these groups to the three properties of interest as well as the empirical equations relating them to the chosen properties were those given by Van Krevelen.¹ The group contributions were used to calculate water absorption, W , glass transition temperature, T_g , and density, D , according to the following equations developed by Van Krevelen:¹

$$W = \frac{18 \sum H_i}{\sum M_i} \text{ g H}_2\text{O/g polymer} \quad (1)$$

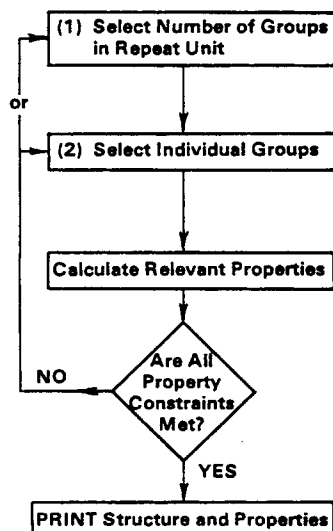


Fig. 1. Flow chart for proposed methodology.

$$T_g = \frac{\sum Y_{gi}}{\sum M_i} - 273^\circ\text{C} \quad (2)$$

$$D = \frac{\sum M_i}{\sum V_{gi}} \text{ g/cm}^3 \quad (3)$$

Where H_i is the molar water content, M_i the gram molecular weight, Y_{gi} the molar glass transition function and V_{gi} the group contribution to molar volume. Group contributions for H_i , M_i , Y_{gi} , and V_{gi} are shown in Table I for the seven groups used in the example. (Van Krevelen also takes a more sophisticated approach which takes into account the interaction among the groups in the structural unit as well as the effects of crystallinity. In the program used for this example the simpler approach was used. For more detail please consult Ref. 1). These were stored as part of the computer program. In the computer program the number of groups comprising the structural unit could either be preselected or selected on a random basis for each iteration. Results from a search using the latter method are given in Table II. In this example, 300 structural units were evaluated using random selection of both the number of groups per structural unit and the groups themselves. Of these only the 12 shown in Table II met the required specifications.

Note: The properties of a polymer with the structural unit predicted by this methodology are the result of contributions of each group within this unit. The order of these groups in the polymer chain can be varied if necessary to devise a feasible synthetic scheme from synthesizable monomers. In other words, the properties predicted for $-(\text{CH}_2-\text{CH}_2-\text{CHOH}-\text{CHCl})-$ are the same as those predicted for $-(\text{CH}_2-\text{CHOH}-\text{CH}_2-\text{CHCl})-$. Multiples of the structural unit may also be rearranged to identify a feasible polymer. For example, the properties predicted for $(-\text{CHCl}-\text{CHCl}-\text{CH}_2-)$ are also expected in the polymer $[(\text{CH}_2-\text{CHCl})_{2n}(\text{CHCl}-\text{CHCl})_n]$.

RANKING OF CANDIDATE MATERIALS

In some situations, the number of candidates which are identified may be quite large. It is therefore desirable to have a means of ranking the selected polymers. The desirability function approach developed by Har-

TABLE I
Group Contributions Used in Example

Group	Y_{gi}	V_{gi}	H_i	M_i
$-\text{CH}_2-$	2700	15.85	3.3×10^{-5}	14
$-\text{CO}-$	27,000	13.4	0.11	28
$-\text{COO}-$	8000	23.0	0.075	44
$-\text{O}-$	4000	10.0	0.02	16
$-\text{CONH}-$	12,000	24.9	0.75	43
$-\text{CHOH}-$	13,000	19.15	0.75	30
$-\text{CHCl}-$	20,000	29.35	0.015	48.5

TABLE II
Computer Search for Viable Polymer Structures

Specifications			
Density	1-1.5 g/cm ³		
Water absorption	0-0.18 g(H ₂ O)/g (polymer)		
<i>T_g</i>	25°C		
Results			
Structural units	g(H ₂ O)/g (polymer)	<i>T_g</i> (°C)	Density (g/cm ³)
—(CH ₂ —CHCl)—	0.004	90	1.38
—(CH ₂ —COO—CH ₂)—	0.019	79	1.32
—(CO—CH ₂)—	0.047	433	1.44
—(CO—CH ₂ —O—CH ₂)—	0.032	232	1.31
—(CHCl—CHCl—CH ₂)—	0.005	112	1.49
—(CH ₂ —COO—CH ₂ —CO)—	0.033	251	1.47
—(CHCl—CH ₂ —CHOH—O)—	0.130	93	1.46
—(CO—CH ₂ —CH ₂)—	0.035	305	1.24
—(CH ₂ —CO—CH ₂ —CONH)—	0.156	175	1.42
—CO—O—CH ₂)—	0.040	308	1.48
—(CHCl—CH ₂ —O)—	0.008	67	1.42
—(CH ₂ —CHOH—CHCl)—	0.149	113	1.44

rington³ and modified by Derringer and Suich⁴ provides such a means. This approach involves the following steps:

1. For each predicted property the property level *Y* is transformed to a 0-1 desirability, *d_i*, scale. On this scale, 0, corresponds to a property level that makes the polymer useless for the application being considered. A desirability of 1.0, on the other hand, corresponds to a property level such that no other property level would make the polymer more useful for the application. For intermediate values, the higher the *d*, the more desirable is the corresponding property value.

2. The individual desirabilities are combined into a composite desirability *D* by taking the geometric mean of the individual desirabilities as follows:

$$D = (d_1 d_2 d_3 \dots d_k)^{1/k} \quad (4)$$

where *k* is the number of properties. It will be noted that *D* reflects the overall optimality of the polymer of interest. Like the individual desirabilities, *D* also ranges between 0 and 1 with the same interpretations of these values. The presence of a single zero-valued *d* will result in *D* also being zero. This reflects most optimization strategies where one property which is unsatisfactory will render the material useless.

It is obvious that once a *D* is computed for each identified polymer, it will serve as a ranking index. Higher *D* values correspond to more promising materials than lower *D* values.

Types of Constraints

Property constraints are either one-sided or two-sided. For example, if a property must be greater than some minimum value or less than some maximum value, the constraint is one-sided. If, however, a property value

must fall between a minimum and maximum value the constraint is said to be two-sided. Derringer and Suich⁴ presented equations for relating property levels to desirability levels for both one- and two-sided constraints.

For the case of a one-sided lower constraint the desirability equation is

$$d_i = \begin{cases} 0 & \hat{Y}_i \leq Y_{i^*} \\ \left[\frac{\hat{Y}_i - Y_{i^*}}{Y_i^* - Y_{i^*}} \right]^r & Y_{i^*} \leq \hat{Y}_i \leq Y_i^* \\ 1 & \hat{Y}_i \geq Y_i^* \end{cases} \quad (5)$$

Here \hat{Y}_i represents the predicted property value and Y_{i^*} is the minimum acceptable level of property Y_i . Values below this value will have a desirability of zero. Y_i^* is the highest value of Y_i which will translate into improved utility. The desirability at Y_i^* equals 1.0 and does not increase as Y_i increases above Y_i^* . The value of r is adjustable and determines the rate of increase of d_i with Y_i between Y_{i^*} and Y_i^* . This equation can be used for a property to be minimized by simply maximizing the negative value of the property.

For the case of a two-sided constraint the desirability equation is

$$d_i = \begin{cases} \left[\frac{\hat{Y}_i - Y_{i^*}}{c_i - Y_{i^*}} \right]^s & Y_{i^*} \leq \hat{Y}_i \leq c_i \\ \left[\frac{\hat{Y}_i - Y_i^*}{c_i - Y_i^*} \right]^t & c_i \leq \hat{Y}_i \leq Y_i^* \\ 0 & \hat{Y}_i < Y_{i^*} \text{ or } \hat{Y}_i > Y_i^* \end{cases} \quad (6)$$

Here Y_{i^*} and Y_i^* are, respectively, the lower and upper constraints on property Y_i below and above which the desirability will equal zero. Parameter c_i is the most desirable level of the property Y_i and corresponds to a desirability of 1.0. Note that c_i need not be midway between Y_{i^*} and Y_i^* . The rate of decrease of d_i above and below c_i is determined by adjustable parameters t and s . Values of Y_{i^*} , Y_i^* , c_i , r , t , and s in eqs. (5) and (6) are specified by the user and will reflect his judgments about the optimization problem at hand.

Example Using Desirability Ranking

The specifications from the earlier example were expressed as desirabilities by using eqs. (5) and (6). Table III provides the values used for the parameters in the equations. Plots of these properties and their corresponding desirabilities are presented in Figures 2-4 for water absorption, T_g , and density, respectively. Figure 2 represents a one-sided constraint, and Figures 3 and 4 represent two-sided constraints. Notice that there is considerable flexibility in the desirability functions. For example, in Figure 3 the most desirable value of T_g is 200°C. Below 200°C, however, the desirability function falls much more slowly than it does above 200°C. This reflects a situation where a value below 200°C is more desirable than a value the same number of degrees above 200°C. In Figure 4 on the other hand, density values greater than 1.3 were more acceptable than values equally distant but below 1.3.

TABLE III
Desirability Equation Parameters Employed for Desirability Ranking Using
Eqms. (5) and (6)

Property	Y_i	Y_i^*	Desirability Parameter Values Used				Eq
			r	c	s	t	
-Water absorption	- 0.18	0	3	—	—	—	(5)
T_g	25	365	—	200	0.3	3	(6)
Density	1.0	1.5	—	1.3	3	1	(6)

Table IV is the result of the same search which resulted in Table II. In this case, however, desirabilities were calculated for each property corresponding to Figures 2-4, and the resulting candidate polymers were ranked in order of decreasing composite desirability, D . Only the most promising six candidates are presented.

DISCUSSION OF CANDIDATE POLYMERS FROM DESIRABILITY SEARCH

The first candidate in Table IV, $(-\text{CH}_2-\text{COO}-\text{CH}_2-)_n$, is poly(β -propiolactone), which can be prepared from β -propiolactone. It has been prepared but is not a commercially available material. The second, $(-\text{CH}_2-\text{CHCl}-)_n$, is, of course, poly(vinyl chloride), which is readily avail-

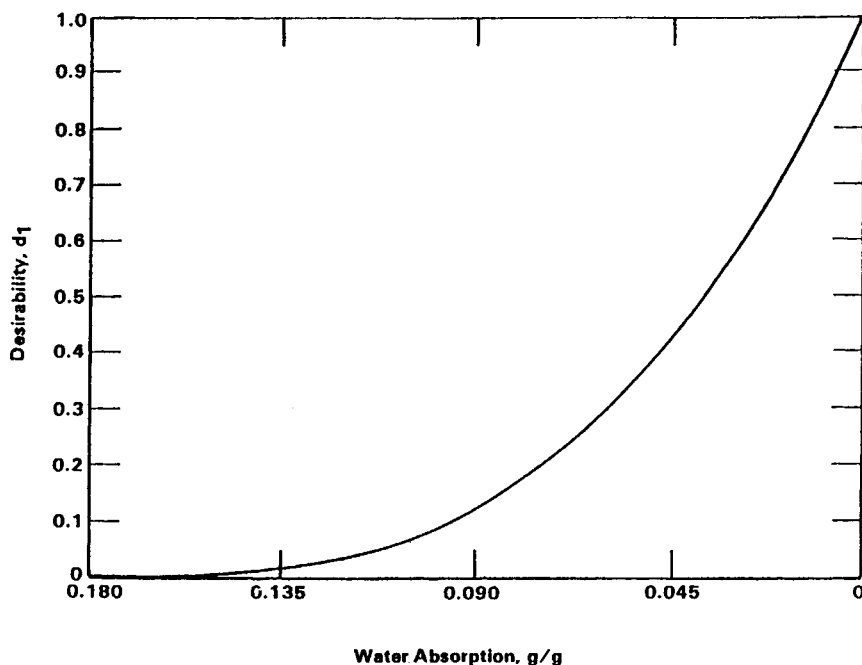


Fig. 2. Desirability function for water absorption in example.

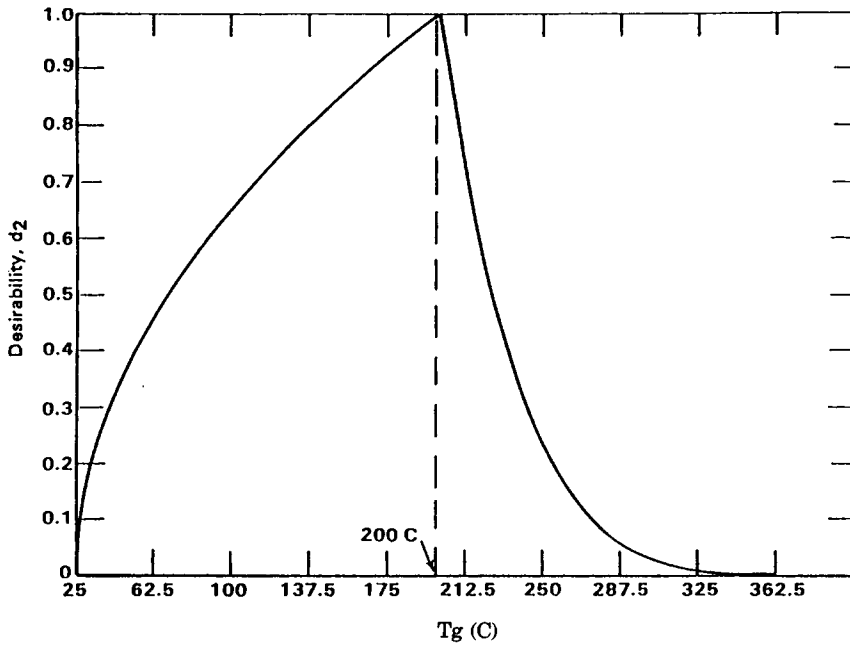


Fig. 3. Desirability function for T_g in example.

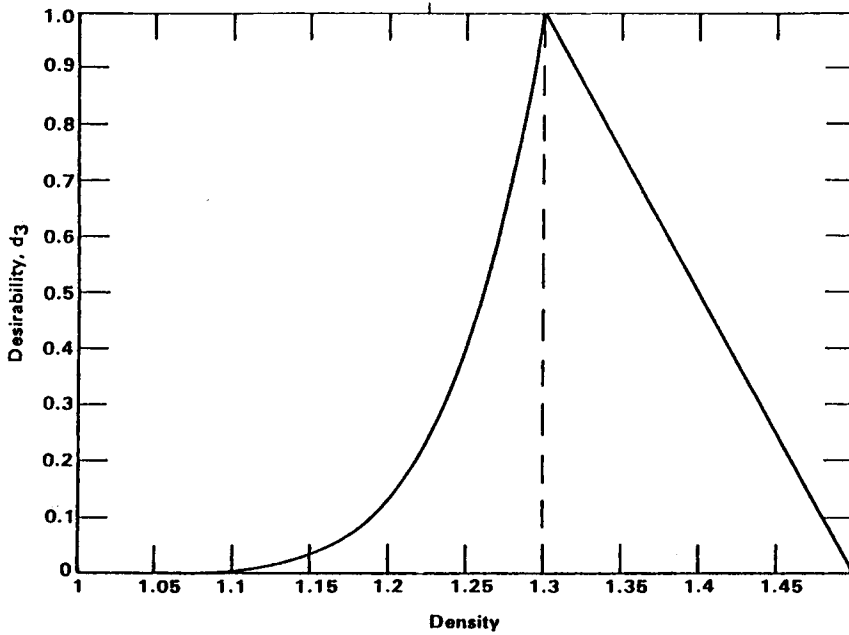


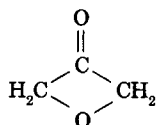
Fig. 4. Desirability function for density in example.

TABLE IV
Prioritized Listing of Candidate Polymers From Desirability Search

Desirability search ^a				
Structural unit	Water absorption (g/g)	T_g (°C)	Density (g/cm)	Composite desirability
—(CH ₂ —COO—CH ₂)—	0.019(0.715)	79(0.555)	1.32(0.900)	0.709
—(CH ₂ —CHCl)—	0.004(0.934)	90(0.610)	1.38(0.585)	0.692
—(CO—CH ₂ —O—CH ₂)—	0.032(0.556)	232(0.418)	1.31(0.950)	0.604
—(CHCl—CH ₂ —O)—	0.008(0.873)	67(0.490)	1.42(0.400)	0.555
—(CHCl—CHCl—CH ₂)—	0.005(0.919)	112(0.705)	1.49(0.050)	0.319
—(CH ₂ —COO—CH ₂ —CO)—	0.033(0.545)	251(0.229)	1.47(0.150)	0.266

^a Values in parentheses represent calculated desirabilities.

available and inexpensive. The third candidate, (—CO—CH₂—O—CH₂)_n, in poly(3-keto oxetane). This could be made from 3-keto oxetane,



but it is expected that it would be difficult to prepare on a commercial basis. The fourth candidate, (—CH—Cl—CH₂—O—)_n, chlorinated poly(ethylene oxide), could possibly be made by chlorinating poly(ethylene oxide). It is not known, however, whether or not this is feasible on a commercial scale. The fifth candidate, (—CHCl—CHCl—CH₂—)_n, is a form of chlorinated poly(vinyl chloride), although synthesis may not be simple. Finally, the last candidate, (—CH₂—COO—CH₂—CO—), is somewhat exotic. The cyclic monomer will provide the polymer if it can be polymerized by a ring-opening reaction.

The above example was carried out primarily to illustrate the potential of the proposed methodology. The next step would typically be a more thorough literature search to determine if those identified candidates which are not familiar polymer materials have ever been synthesized, and if so, the potential for commercialization. It is quite interesting that this relatively simple example identified some potential polymers which may be truly novel. An expanded data base with more properties and polymer repeat units appears to hold considerable promise.

DISCUSSION

The major advantage of the proposed methodology is the ability to determine, prior to experimentation, the types of polymer structures which are most likely to meet the given specifications. In this way much fruitless experimental effort can be eliminated. A good polymer synthesis chemist can inspect the output to determine (1) which polymers already exist, (2) which represent energetically impossible combinations of groups, and (3)

which are not available but can be made. Perhaps the last category will be most fruitful.

The purpose of this paper was simply to introduce the methodology, and the example given was, therefore, rather simple. Obvious extensions include the addition of more structural unit groups as well as more properties to the data base.

This paper is based on a paper presented at the Symposium on Properties of Polymers and Correlation with Chemical Structure in the Division of Organic Coatings and Plastics, American Chemical Society Meeting, March 1982, Las Vegas.

References

1. D. W. Van Krevelen, *Properties of Polymers*, 2nd ed., Elsevier, Amsterdam, 1976.
2. V. D. McGinniss, *Org. Coat. Plast. Chem. Prepr.* **39**, 529-524 (1978).
3. E. C. Harrington, Jr., *Ind. Quality Control*, **21**, 494 (1965).
4. G. C. Derringer and R. Suich, *In Quality Technol.* **12**(4), 214 (1980).

Received December 17, 1983

Accepted March 5, 1985