

A NEURAL NETWORK STUDY ON GLASS TRANSITION TEMPERATURE OF POLYMERS*

Lin-xi Zhang^{a**}, De-lu Zhao^b and You-xing Huang^a

^aDepartment of Physics, Zhejiang University, Hangzhou 310028, China

^bPolymer Physics Laboratory, Center of Molecular Science, Institute of Chemistry, Chinese Academy of Sciences, Beijing 100080, China

Abstract In this paper, an artificial neural network model is adopted to study the glass transition temperature of polymers. In our artificial neural networks, the input nodes are the characteristic ratio C_∞ , the average molecular weight M_e between entanglement points and the molecular weight M_{mon} of repeating unit. The output node is the glass transition temperature T_g , and the number of the hidden layer is 6. We found that the artificial neural network simulations are accurate in predicting the outcome for polymers for which it is not trained. The maximum relative error for predicting of the glass transition temperature is 3.47%, and the overall average error is only 2.27%. Artificial neural networks may provide some new ideas to investigate other properties of the polymers.

Keywords Neural network, Glass transition temperature

INTRODUCTION

The glass transition of polymers is an important phenomenon in polymer physics. Although much experimental and theoretical works has been done in order to explain the physical origin of the glass transition^[1], the theoretical description of this transition is still regarded as an unsolved question^[2]. In the theories, the glass transition is related to the vanishing of the configurational entropy. Although it is possible to understand the empirical Vogel-Fulcher law within the framework of these thermodynamic theories, they cannot account for the detailed dynamical aspects of the glass transition^[3, 4]. One of the final goals for solving the problem of physics of disorder systems in the field of polymers is the explanation of relations between the glass transition temperature of polymers and their chemical structure and the motion mechanism. Some successful relations have been established between the glass transition temperature T_g and the chemical or/and physical factors such as cohesive energy density (CED), hole energy formation, chain stiffness and flex energy^[5-9]. However, correlation of this type emphasizes only one aspect of a rather complicated problem. The results do not help us to understand the nature of dynamic processes at the glass temperature. In fact, the glass transition temperature of a polymer depends on its chemical structure, such as molecular weight, chain dimension, and the interaction between atoms, etc. This paper aims to find the relations between T_g and the chemical structures and high molecular weights. Perhaps the relations are rather complicated, and it is more difficult to obtain the mathematical expression. Hence, an artificial neural network is adopted. Artificial neural networks have been widely applied in chemistry. The advantage of the artificial neural network is that the relations between the input "nodes" and the output "nodes" are found automatically by computer through "training"^[10]. Recently, Darsey *et al.* used an artificial

* This research was financially supported by NSFC (No. 29874012) and the Special Funds for Major State Basic Research Projects (95-12 and G1999064800).

** Corresponding author: Lin-xi Zhang (章林溪), E-mail: zhanglx@mail.hz.zj.cn

Received March 1, 2001; Revised March 29, 2001; Accepted April 1, 2001

neural network to investigate the characteristic ratio and that for the temperature coefficient^[11]. In their work, the overall average error for predicting of the characteristic ratio was 4.82%, and that of the temperature coefficient was 0.89%. In this paper, we will investigate the glass transition temperature by using artificial neural networks, and to find the correlation relation between glass transition temperature and their structures. In the following section, the neural network model is given.

ARTIFICIAL NEURAL NETWORK MODEL

The architecture of the three-layer back-propagation neural network (BPNN) is shown in Fig. 1. The input nodes are fully connected to a middle or hidden layer of nodes, which in turn are fully connected to a set of output nodes. The number of input nodes and output nodes are n and p , respectively, and the number of hidden nodes is m , which depends on the experimental parameters. This neural network algorithm can evolve a set of weights to produce an arbitrary mapping from input to output by presenting pairs of input nodes and their corresponding output nodes.

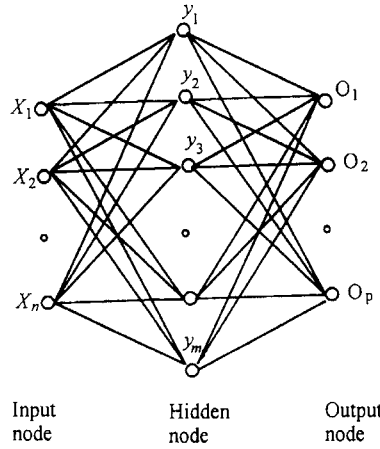


Fig. 1 Architecture of the neural network model used to simulate the glass transition temperature

For an input layer node, the output (Xo) is equal to its input (X):

$$Xo_i^L = X_i^L \quad (i=1, 2, \dots, n; L=1, 2, \dots, S) \quad (1)$$

Here the subscript i indicates the i -th input layer node, and L indicates the L -th trained sample. n is the number of input nodes, and S is the number of trained samples. The outputs from the inputs are weighted and fed to the hidden nodes. The net input of the j -th hidden node in the L -th trained sample is:

$$y_j^L = \sum_{i=1}^n Xo_i^L W_{ij}^L \quad (j=1, 2, \dots, m; L=1, 2, \dots, S) \quad (2)$$

where W_{ij}^L indicates the weight of the connection between the i -th input node and the j -th hidden node in the L -th trained sample, and m is the number of hidden nodes. The outputs of the hidden layer nodes (Yo_k^L) are given after transformation by the sigmoidal function:

$$Yo_k^L = f(y_k^L) \quad (k=1, 2, \dots, m) \quad (3)$$

here $f(x)$ is the sigmoidal function:

$$f(x) = \frac{1}{1 + e^{-x}} \quad (4)$$

The output of the last layer (output layer) nodes is given by the sum after weighted input from the hidden nodes and transformation:

$$O_l^L = f\left(\sum_{j=1}^m Y_{oj}^L V_{jl}^L\right) \quad (l=1, 2, \dots, p; L=1, 2, \dots, S) \quad (5)$$

where V_{jl}^L is the weight connecting the j -th hidden node to the l -th output layer node. After carrying out the procedure for all combinations of input signals, the root-mean-square error (RMSE) is given by:

$$\text{RMSE} = \sqrt{\sum_{\alpha=1}^p \sum_{\beta=1}^S \frac{E_{\alpha\beta}^2}{pS}}, \quad E_{\alpha\beta} = \frac{Y_{\alpha}^{\beta} - O_{\alpha}^{\beta}}{Y_{\alpha}^{\beta}} \quad (6)$$

where p is the number of output layer nodes, and S is the number of trained samples, and Y_{α}^{β} is the target value of β -th trained samples. In this paper, as the output node is the glass transition temperature, only the number of output layer nodes is $p = 1$. Learning is carried out by iteratively updating the connecting weights so as to minimize a measure of the difference between the actual output values of the network and the desired output values. Therefore, it is very important to determine the weights W_{ij}^L and V_{jl}^L .

Each weight, whose initial value is given by the computer randomly, is changed automatically by Eqs. (7), (8) and then the training is repeated until RMSE falls into an acceptable region,

$$\Delta W_{ij}^L(t) = \eta \delta_j^L X_{oi}^L + \mu \Delta W_{ij}^L(t-1) \quad (i=1, 2, \dots, n; j=1, 2, \dots, m; L=1, 2, \dots, S) \quad (7)$$

$$\Delta V_{kl}^L(t) = \eta \xi_l^L Y_{ok}^L + \mu \Delta V_{kl}^L(t-1) \quad (k=1, 2, \dots, m; l=1, 2, \dots, p; L=1, 2, \dots, S) \quad (8)$$

Here $\Delta W_{ij}^L(t)$ is the weight change of the connection between the i -th input node and the j -th hidden node, and $\Delta V_{kl}^L(t)$ is the weight change of the connection between the k -th hidden input node and the l -th output node, and t is the number of iterations of neural computation. η is the learning rate and determines how fast the changes $\Delta W_{ij}^L(t)$ and $\Delta V_{kl}^L(t)$ should be implemented in the iteration cycles. The momentum constant μ prevents sudden changes in the direction in which corrections are made. The errors ξ_l^L and δ_j^L can be expressed as

$$\xi_l^L = (Y_l^L - O_l^L) f'(x)(O_l^L) \quad (l=1, 2, \dots, p) \quad (9)$$

for the output layer l -th node and

$$\delta_j^L = \left(\sum_{\alpha=1}^p \xi_{\alpha}^L V_{\alpha j}^L\right) f'(x)(y_j^L) \quad (j=1, 2, \dots, m) \quad (10)$$

for the hidden layer j -th node. Here $f'(x)$ is the derivative with respect to x of the transformation function, and $f'(x) = (1 - f(x)) \times f(x)$. We let the trained network scan automatically and intensively near the range of experimental value, and the weight factors of the connection W_{ij}^L and V_{kl}^L are almost the same for different samples at the end of training. Therefore, our neural network predicting values can be calculated by Eqs. (1)–(5); here input nodes are the values of testing samples.

RESULTS AND DISCUSSION

The glass transition phenomenon is at present a major problem of the physics of disordered system. The glass

transition temperature is a key polymer parameter that controls melt viscoelasticity, solid mechanical behavior and adhesion of polymers, and depends mainly on the polymer structure. In order to study those structure factors that the glass temperature depends on, we should consider the experimental background of the presented correlation of T_g values of different polymers. First, DSC and analogously dynamic-mechanical (DMS) and dielectric relaxation (DS) spectroscopies show that T_g dependence has a saturation effect in the region $M > M_e$, where M_e is the average molecular weight of chain parts between successive entanglements^[12, 13]. Secondly, small angle neutron scattering (SANS) measurements of polymers in the bulk state prove that the configuration of chains in the elastic and glassy state of amorphous polymers is close to the chain configuration in solution under θ -condition^[14, 15]. Therefore, we think that the glass transition temperature may depend on the characteristic ratio C_∞ and the average molecular weight M_e . At the mean time, the molecular weight M_{mon} of the repeating unit may also be considered. Thus in our calculation, the input nodes are C_∞ , M_e and M_{mon} , and the number of the input layer n is 3 (see Fig. 1). The output node is the glass transition temperature and the number of the output layer p is equal to 1. The values of C_∞ , M_e and M_{mon} are taken from the literature^[16, 17]. Although DSC, DMS and DS show that T_g dependence has a saturation effect in the region $M > M_e$ without including the region $M \leq M_e$, we can also predict the T_g in the region $M \leq M_e$ from the neural network. The reason is that the relationship between the input “nodes” and the output “nodes” can be found automatically through “training”. We also know that the experimental values of C_∞ , M_e and M_{mon} may be different. If the training number of polymers is large, the results are credible. The initial architecture of the neural network is only an initial guess; it must be modified after performing the calculations. The learning rate η and the momentum constant μ should be selected carefully. In the training process, if the value of the learning rate and the momentum constant are too small, RMSE would show large fluctuation, and could not reach a global minimum quickly. In our training, $\eta = 0.1$ and $\mu = 0.9$.

The selection of the number of hidden layer nodes is important in the artificial neural network. In order to obtain the best neural network structure, several artificial neural network systems with different number of hidden nodes were tested. Here the testing polymer of output node was polyisobutylene, the number of polymer samples (excluding polyisobutylene) was 17, and RMSE = 0.2. The results are shown in Fig. 2. In Fig. 2, the relative error Δ is equal to $|T_g(\text{experiment}) - T_g(\text{predicted})|/T_g(\text{experiment})$, and m is the number of hidden layer nodes. In Fig. 2, we find that there is a minimum value for the relative error. For example, $\Delta = 8.91\%$ for $m = 3$ and $\Delta = 13.4\%$ for $m = 10$. From Fig. 2, we know that $m = 6$ is the best fit with our system, and it is also true for the other testing polymer chains. Therefore, we choose the 3-6-1 structure as the optimized system.

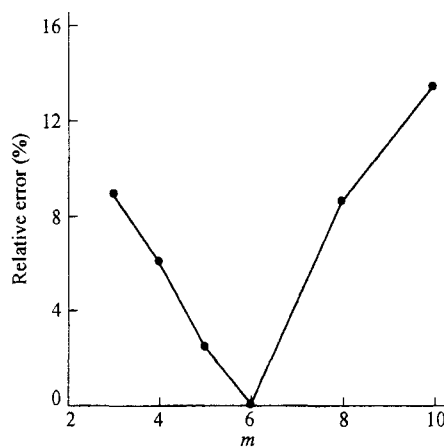


Fig 2 A plot of relative error versus m (the number of hidden nodes) for polyisobutylene with RMSE = 0.2

Table 1 lists the polymers used for training and testing purposes. We calculate a total of 20 polymers including 17 for training and 3 for testing.

Table 1. Polymers used in training and testing

Polyisobutylene	Isotactic polystyrene
Polydimethylsiloxane	Polycaprolactam
Polyoxymethylene	Poly(hexamethylene adipamide)
Polyoxyethylene	Poly(ethylene isophthalate)
Polyethylene	Polyacrylonitrile
Poly(vinyl chloride)	Poy(<i>n</i> -decyl methacrylate)
Poly(<i>n</i> -hexyl methacrylate)	Poly(<i>n</i> -octyl methacrylate)
Poly(methyl methacrylate)	Poly(methyl acrylate)
Poly(ethyl methacrylate)	Poly(vinyl acetate)
Atactic polystyrene	Polycarbonate of bisphenol-A

Table 2 presents the results for three runs in which three polymers were chosen from the 20 polymers listed in Table 1 to test our training. Each run represents a different set of three polymers. In Table 2, we find that the relative error ranges from 1.13% to 3.47%, and the overall average error is only 2.27%, which is acceptable. This means that if C_∞ , M_e and M_{mon} can be known accurately, the predictions of the glass temperature using the neural network are accepted.

Table 2. Results of the predicted glass transition temperature for selected polymers

Polymer	Predicted T_g (K)	Experimental T_g (K)	Relative error in %
Run A:			
Polyisobutylene	205.3	203	1.13
Atactic polystyrene	365.0	373	2.14
Polydimethylsiloxane	155.2	150	3.47
Run B:			
Polyisobutylene	205.7	203	1.15
Poly(hexamethylene adipamide)	332.9	323	3.07
Polyoxymethylene	185.2	190	2.53
Run C:			
Poly(vinyl chloride)	366.1	360	1.69
Polycaprolactam	315.3	323	2.38
Poly(methyl acrylate)	292.3	301	2.89

This work demonstrates the feasibility of using neural network to model the glass transition temperature of polymeric systems. Although the results are not perfect yet, they are very satisfactory. The maximum error for predicting the glass transition temperature is only 3.47%. The advantage of using neural network instead of the more traditional techniques is that once you properly train your network, it requires no more than one second or less of computer time to obtain results on other polymers for which the network has not been trained. From our calculations, we know that the glass transition temperature depends on the polymer structure, mainly on C_∞ , M_e and M_{mon} . This investigation may provide some new methods for studying the properties of polymers, and such work is already in progress.

Acknowledgements We thank Profs. Zhi-tong Cao and Ping Zhu for their useful discussions.

REFERENCES

1 Gotze, W., "Liquids, Freezing and the Glass Transition", ed. by Hansen J.P., Levesque D. and Zinnjustin, J., North-Holland, Amsterdam, 1990

2 Zallen, R., "The Physics of Amorphous Solids", Wiley, New York, 1983, p.63

3 Vogel, H., Phys. Z, 1921, 22: 645

- 4 Fulcher, G.S., *J. Am. Ceram. Soc.*, 1925, 8: 339
- 5 Lee, W.A. and Sewell, J.H., *J. Appl. Polym. Sci.*, 1968, 12: 1397
- 6 Wrasidlo, W., *Adv. Polym. Sci.*, 1974, 13: 2
- 7 Privalko, V.P. and Lipatov, Y.S., *J. Macro. Sci. Phys.*, 1974, B9: 551
- 8 Gibbs, J.A. and DiMarzio, E.A., *J. Chem. Phys.*, 1958, 28: 373
- 9 Linxi, Z., Agen, X., Jianmin, X. and Delu, Z., *Chinese J. Polym. Sci.*, 1999, 17: 543
- 10 Zurada, J.M., "Introduction to Artificial Neural Systems", Weit Publishing Company, New York, 1992, p.32
- 11 Darsey, J.A., Soman, A.G. and Noid, D.W., *Makrom. Chem. Theory Simul.*, 1993, 2: 711
- 12 Boese, D. and Kremer, F., *Macromolecules*, 1990, 23: 829
- 13 Turnet, D.T., *Polymer*, 1978, 19: 789
- 14 Fischer, E.W. and Dettenmaier, M., *J. Non-Cryst. Solids*, 1978, 31: 181
- 15 de Gennes, P.G., "Scaling Concepts in Polymer Physics", Cornell University Press, Ithaca, New York, 1979, p.107
- 16 Lee, W.A. and Rutherford, R.A., "Glass Transition Temperature of Polymers", in "Polymer Handbook", ed. by Brandrup, J. and Immergut, E.H., Wiley, New York, 1975, p.236
- 17 Wu, S., *Polym. Engin. Sci.*, 1992, 32: 823