Computational Materials Science 118 (2016) 56–65

Contents lists available at ScienceDirect

Computational Materials Science

journal homepage: www.elsevier.com/locate/commatsci

Evolution of glass forming ability indicator by genetic programming

Manwendra K. Tripathi [a,b], Subhas Ganguly [a,*], Partha Dey [c], P.P. Chattopadhyay [b]

[a] National Institute of Technology Raipur, Raipur 492010, India
[b] Indian Institute of Engineering Science and Technology, Shibpur, Howrah 711103, India
[c] Department of Mechanical Engineering, Academy of Technology, Adisaptagram, Hooghly 712121, India

ARTICLE INFO

Article history:
Received 17 November 2015
Received in revised form 22 February 2016
Accepted 23 February 2016

Keywords:
Glass forming ability
GFA criteria
Bulk metallic glass
Genetic programming

ABSTRACT

A symbolic regression technique has been employed to evolve the functional relationship among the characteristic transformation temperatures, *viz.* glass transition temperature (T_g), onset crystallization temperature (T_x) and offset temperature of melting (T_l) concerning glass forming ability (GFA) of bulk metallic glasses (BMGs). The critical diameters (D_{max}) of 410 reported BMGs, along with their T_g, T_x and T_l values, forms the training data, for a genetic programming based computer code which attempts to evolve an expression leading to high correlation with D_{max} as the target variable. Another set of recently reported 184 BMGs data, is used to assess the performance of the evolved expression. The evolved expression shows significantly improved correlations with critical diameter D_{max}, for training data, test data and training and test data considered together. The same also compares well with the high correlation GFA indicators reported earlier in the literature.

1. Introduction

The critical diameter (D_{max}), *i.e.*, the maximum possible thickness of the metallic glass that can be produced from an supercooled liquid, has emerged as a practical consideration to compare the relative potential of glass formation in bulk metallic glasses (BMGs). The value of D_{max} is in inversely proportion to the critical cooling rate (R_c) which is conceptually regarded as a good quantification of its glass forming ability. The experimental determination of D_{max} is relatively easier than R_c. However, multiple experiments are still unavoidable to determine D_{max} for each alloy composition which is eventually dependent on the fabrication method [1,2].

The experimental determination of characteristic transformation temperatures of the metallic glasses is relatively easy using differential scanning calorimetry (DSC) or differential thermal analyser (DTA), as compared to other physical and thermal property measurement such as viscosity of a melt, in the course of estimation of glass forming ability [2,3]. This fact has led to the development of expressions representing glass forming ability in terms of characteristic temperatures. A single experimental synthesis of an alloy followed by characteristic temperature measurements using DSC/DTA is sufficient for estimating the maximum possible diameter of glassy structure of the alloy. Nevertheless, such expressions in terms of characteristic temperatures should be well correlated with D_{max} and be consistent with the physical metallurgy principals. Therefore, developing GFA expression in terms of characteristics temperatures, having high correlation with critical diameters of BMG alloys, can significantly reduce the experimental effort to determine glass forming ability of a composition with reasonable accuracy.

Inoue proposed to measure GFA on the basis of stability of the metallic glass against crystallization while heating, quantified as $\Delta T_x (= T_x - T_g)$ [4]. It implies that a good glass former should have either high onset temperature of crystallization T_x or low T_g. It may be noted that the first GFA expression proposed by Turnbull ($T_{rg} = T_g/T_l$) suggests that T_g should be high for good GFA, which is essentially contradicted by the Inoue's proposition. The correlation coefficients with the experimentally measured D_{max} (and $\log R_c$) values were in better agreement with ΔT_x compared to T_{rg}, which favoured ΔT_x to be accepted as a better GFA indicator.

Subsequently, several other indicators have been proposed on the basis of similar empirical considerations of glass formation in terms of T_g, T_x and T_l like γ, γ_m, δ, new-β, ω, etc. [2,5–21]. Despite of substantial efforts devoted in the development of reliable GFA expressions for past few decades, the frequently used expressions like γ and ω parameters have been recently reported to fail to represent the glass forming ability of the BMGs [15,22–24]. Attempts have been made to develop expressions like Φ [10] and θ [25] by fitting in suitable indices to pre-fixed expression in terms of the characteristics temperatures to satisfy the requirement of high correlation with critical diameter, D_{max}. However, with synthesis of new compositions, these expressions also have shown poor correlations with D_{max}, since the fitting indices were dependent on the *training data*.

* Corresponding author.

http://dx.doi.org/10.1016/j.commatsci.2016.02.037
0927-0256/© 2016 Elsevier B.V. All rights reserved.

It comes by no surprise that while majority of the propositions regarding GFA expressions are based on thermodynamics and kinetics based considerations. Recently, Chattopadhyay et al. have classified the reported glass forming ability expressions into four categories on the basis of characteristic transformation temperatures, thermodynamic parameters, topology based parameters and phase transformation kinetics based parameters [26]. The authors have shown that kinetic viscosity based approach to estimate glass forming ability correlates well for few compositions (45 in number), with the critical cooling rate calculated from the TTT diagram, unlike other three proposed classes. However, this criterion is not statistically tested over large number of alloys since the kinetic data have not been reported. Similar approach has been considered by Johnson [27] and Kozmidis-Petrović [28] where the availability of kinetic data limits its practical application. Therefore, GFA expression in terms of characteristics temperatures with high correlation with D_{max} remains a pragmatic solution to estimate glass forming ability of a composition.

There has been at least two attempts [10,25] in developing the GFA expressions using statistical considerations. Multivariate statistical analysis has been attempted for feature extraction and classification of bulk metallic glasses by Tripathi et al. [29]. While statistical regression may be seen as the embryonic form for dealing with huge amount of data, giving meaning to it, and helping scientists to arrive at theories accurately. The advent and subsequent advances of the computing power and associated storage capacity of computers have laid open a new avenue for researchers to explore *i.e.,* data analytics or informatics. Materials informatics, being a fast developing and highly successful paradigm for dealing with quantitative structure-attribute relationships [30–33]takes advantage of this huge search potential in finding appropriate structure that would best fit into a given physical system, and provide plausible physical explanation to the obtained expression.

In view of the above, the present study attempts to evolve an expression using the transformation temperatures having high correlation with critical diameter D_{max} by employing symbolic regression. It is imperative to mention here that symbolic regression is a very versatile technique capable of forming any arbitrary expression by combining a set of operators and given variables into a flexible form with no *a priori* assumptions regarding its structure. A set of 410 BMG data, compiled by Long et al. [12], has been undertaken to evolve the developed expression(s), referred as training data. Additional 184 BMG data have been compiled after publication of Long et al. [12] and the coefficient of correlations have been calculated to evaluate the performance of evolved expression(s).

In context of compositional design, evolutionary computational approaches have already been explored in the field of BMG composition design problem. One of the pioneer work on evolutionary computation applied to BMG composition design problem was done by Dulikravich et al. [34], where the composition evolution has been formulated in a multi-objective optimization sense considering simultaneously maximizing T_g, T_l and T_g/T_l and minimizing density of the designed alloys. In this work, the evolutionary tool of *genetic programming* has been used to evolve subsequent generations of flexible expressions representing GFA (constructed from the symbols of addition, subtraction, multiplication and division, operated on the characteristic transformation temperatures T_g, T_x and T_l) which gradually improve over the generations to produce the best possible expressions of a GFA indicator.

2. Methodology

Since the efforts made so far in developing a GFA parameter is constructed by the terms T_g, T_x and T_l using the operators addition (+), subtraction (−), multiplication (×) and division (/) following thermodynamic and kinetic interpretation of the operators {+, −, ×, /} in the expression and the constituent factors. Here, it is proposed to search for an expression that can be assembled using T_g, T_x and T_l and the operators {+, −, ×, /} which eventually describe the experimentally observed D_{max} data for a wide spectrum of alloy composition. In this context, the genetic programming, an evolutionary search tool, has been exercised. The genetic programming attempts to generate a number of trial GFA expressions and from those the best GFA expression (G_p) has been evolved.

A code was developed in Fortran 90/95 for solving the symbolic regression problem using *Genetic Programming* [35,36] and applied to evolution of a GFA expression in terms of T_g, T_x and T_l. Trees grown in the GP code were initialized with the ramped half-and-half approach, where half of the trees grew to their full depth, and the rest of them terminated prematurely before being fully grown. At every generation, a pair of parents are selected from the entire previous generation following the roulette wheel selection procedure, on which the size-fair crossover [37] was applied to generate offspring. The size-fair crossover restricts the crossover points of the two parents to have the same depth-thus controlling bloat [38,39]. Additional care is taken so that not more than 25% of the sub-trees being interchanged among the parents were single leaves – thus ensuring that a substantial amount of genetic material is in fact interchanged [35]. As mutation is believed to play a nominal role in genetic programming [40], point mutation was used that arbitrarily changes a function node selected from the tree at random. The formulation of the evolution methodology has been presented in Fig. 1.

While bloat was completely prevented through the strict restrictions of the size-fair crossover, allowing no tree to exceed the maximum depth, the user could reset the max-depth value to get larger and more complicated trees in next run. The issue of rogue trees was intrinsically left unsolved. However, characterizing lower fitness values, rogue trees participated very little in parental regeneration, and died out eventually from the overall survival of the fittest principle.

In line with the philosophy of evolutionary optimization – where it is not necessary to locate the exact/best solution to a problem in practice, but it is sufficient to find an acceptable solution that serves the requirements to a satisfactory extent – the GP code used in the present article was developed by us may lag a little behind the state-of-the-art in GP, but gives an acceptable solution to the development of GFA expression with just three variables. Focus was kept at simplicity rather than accuracy, and in that way a glass forming indicator substantially accurate than the reported ones were arrived at, which could be explained based on thermodynamic and kinetic considerations.

In genetic programming, the evolving structures, *i.e.* GFA indicator expressions expressed in terms of T_g, T_x and T_l in the present context, are the hierarchically organized mathematical expressions (called *computer programs*), represented as tree-like structures as shown Fig. 1, with size and form dynamically changing during the process of *simulated evolution*. In the process of evolution, it attempts to maximize the correlation of the expression with the D_{max} value for 410 BMGs reported reference [12] which is introduced to the present code as a data file containing T_g, T_x and T_l together with D_{max}.

The population of individuals, a set of GFA indicator expressions, is initialized by creating programs through the combination of the set of functions F = $\{+, -, *, /\}$ and the set of terminal T = $\{1, T_g, T_x, T_l\}$. Each program represents a trial solution means GFA indicator expression that initially searches randomly for a possible candidate of GFA indicator expressions. The search is guided by the fitness of a GFA indicator expression, *i.e.*, correlation of the expression with the D_{max}, which helps to select individuals *i.e.*, GFA indicator expressions. A computer program in genetic programming, formulated as described above, evolves through several genetic operators