Expert Systems with Applications 38 (2011) 15014–15019

Contents lists available at ScienceDirect

Expert Systems with Applications

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

Application of genetic programming for modelling of material characteristics

Leo Gusel *, Miran Brezocnik

University of Maribor, Faculty of Mechanical Engineering, Smetanova 17, 2000 Maribor, Slovenia

ARTICLE INFO

Keywords:
Evolutionary computation
Genetic programming
Modelling
Metal forming
Material characteristics

ABSTRACT

Genetic programming, which is one of the most general evolutionary computation methods, was used in this paper for the modelling of tensile strength and electrical conductivity in cold formed material. No assumptions about the form and size of expressions were made in advance, but they were left to the self organization and intelligence of evolutionary process. Genetic programming does this by genetically breeding a population of computer programs using the principles of Darwinian's natural selection and biologically inspired operations. In our research, copper alloy was cold formed by drawing using different process parameters and then tensile strengths and electrical conductivity (dependent variables) of the specimens were determined. The values of independent variables (effective strain, coefficient of friction) influence the value of the dependent variables. Many different genetic models for both dependent variables were developed by genetic programming. The accuracies of the best models were proved by a testing data set. Also, comparison between the genetic and regression models is presented in the paper. The research showed that very accurate genetic models can be obtained by the proposed method.

1. Introduction

For high quality and full functionality of the formed product, the properties of the material, which the future product will be made of, have to be determined as precisely as possible. Several modelling methods for predicting dependent output variables have been developed to reduce the costs of the experiments and computer computations. In most conventional deterministic modelling methods, such as regression analysis, a prediction model is determined in advance. Traditional methods can be expensive and often result in sub-optimal solution. Because of the pre-specified size and shape of the model, the latter is often incapable of capturing complex relationships between influencing parameters. It is very important that the independent input variables influence on the dependent output variables and, consequently, on the product quality has been already examined in the early stages of a metal forming process.

Evolutionary computation (*EC*) is generating considerable interest for solving real engineering problems. They are proving robust in delivering global optimal solutions and helping to resolve those limitations encountered in traditional methods. *EC* harnesses the power of natural selection to turn computers into optimization tools. This is very applicable to different problems in the manufacturing industry (Koza, 1992; Paskowicz, 2009). One of most important *EC* methods is genetic programming (*GP*) which is, similarly to a genetic algorithm, an evolutionary computation method for imitating biological evolution of living organisms. Several researches have been carried out using a neural network or genetic algorithms for modelling, thus forming process parameters (Fakhrzad & Khademi Zare, 2009; Ganguly, Datta, & Chakraborti, 2007; Odugava, Tiwari, & Roy, 2005; Özel & Karpat, 2005; Pierrevall, Caux, Paris, & Viguier, 2003; Tanguy, Besson, Piques, & Pineau, 2005; Zadeh, Darvizeh, Jamali, & Moeini, 2005), but only a few dealing with much more general genetic programming method (Baykasoglu, Güllü, Çanakçi, & Özbakir, 2008; Brezocnik, Kovacic, & Gusel, 2005; Chang, Kwang, & Kim, 2005; Dimitriu, Bhadeshia, Fillon, & Poloni, 2009). In the *GP* method, the structure subject for adaptation is the population of hierarchically-organized computer programs. The *GP* method is most often used for complex system modelling, but it can also be effectively used for the modelling of a relatively simple system, such as the systems described in our paper.

This paper describes an evolutionary computation method approach for the modelling of tensile strength and electrical conductivity of formed copper alloy. Experimental data obtained during the cold drawing processes under different conditions serves as an environment which, during simulated evolution, models for the tensile strength and electrical conductivity have to be adapted to. Different values for effective strains and coefficients of friction were used as independent input variables, while tensile strength and electrical conductivity were dependent output variables. *GP* method was used for the evolutionary computation of the models

* Corresponding author. Tel.: +386 2 2207851; fax: +386 2 2207990.
E-mail address: leog@uni-mb.si (L. Gusel).

0957-4174/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2011.05.045

for prediction of both dependent variables, on the basis of a training data set.

2. Method used

Genetic programming is evolutionary computation methods in which the structures subject to adaptation are those hierarchically organized computer programs whose size and form dynamically change during simulated evolution.

The space for solutions in the *GP* method is the huge space of all possible computer programs consisting of components describing the problem area studied. Possible solutions in genetic programming are all those possible computer programs that can be composed in a recursive manner from a set of function genes *F* and a set of terminal genes *T*. Function genes *F* are arithmetical functions, Boolean functions, relation functions, etc., while terminal genes are numerical constants, logical constants, variables, etc. (Ashlock, 2006; Koza, 1992).

The initial population is obtained by the creation of random computer programs consisting of available function genes from set *F* and available terminal genes from set *T*. Each program represents a random point in the searching space. The creation of an initial population is a blind random search for solutions within the huge space of possible solutions. The next step is the calculation of individual's adaptation to the environment (i.e., calculation of fitness for each computer program). Fitness is a guideline for modifying those structures undergoing adaptation. Computer programs change in *GP*, in particular during genetic operations regarding reproduction and crossover. The reproduction operation gives a higher probability of selection to more successful organisms. They are copied unchanged into the next generation. The crossover operation ensures the exchange of genetic material between computer programs. The mutation operation increases the genetic diversity of a population.

After finishing the first cycle, which includes creation of the initial population, calculation of fitness for each individual of the population, and genetic modification of the contents of the computer programs and formation of a new population, an iterative repetition of fitness calculation and genetic modification follows. After a certain number of generations the computer programs are usually much better adapted to the environment. The definition of the environment depends on the problem dealt with. The evolution is terminated when the termination criterion is fulfilled. This can be a prescribed number of generations or sufficient quality of the solution. Since evolution is a non-deterministic process, it does not end with a successful solution after each run (i.e., civilization). In order to obtain a successful solution, the problem must be processed over several independent runs. The number of runs required for the satisfactory solution depends on the difficulty of the problem.

3. Experimental work

The aim of the experimental work was to determine the influence of the effective strain ε_e and coefficient of friction μ during cold drawing on the change of tensile strength and electrical conductivity of cold drawn copper alloy CuCrZr. This special copper alloy has high electrical and thermal conductivity, with excellent mechanical and physical properties at elevated temperatures.

Copper alloy rods were deformed by cold drawing under different conditions. The drawing speed was 20 m/min and the angle of drawing die was $\delta = 28°$. Copper alloy rods were drawn from an initial diameter of $D = 20$ mm to six different diameters (i.e. six different effective strains). Three different lubricants with different coefficients of friction ($\mu = 0.07$, $\mu = 0.11$ and $\mu = 0.16$) were used for the drawing process. In order to evaluate the tensile strength, standard specimens for tensile tests were prepared from locations in the middles of the drawn rods. In this way we obtained 18 different experimental specimens.

The tensile strengths of all specimens were determined by providing three tensile tests for each specimen in order to provide reliable results. Electrical conductivity of drawn rods was measured by Sigmatest measurement equipment at frequency $f = 120$ kHZ and temperature $T = 20$ °C. Three measurements of electrical conductivity for each specimen were carried out. The results (average values) for tensile strength and for electrical conductivity are presented in Table 1. These experimental data serve as an environment which, during simulated evolution, models for tensile strength and electrical conductivity have to adapt.

4. Genetic programming modelling of tensile strength and electrical conductivity

In the *GP* method the initial random population $P(t)$ consists of randomly generated organisms which are, in fact, mathematical models. The variable t represents the generation time. Each organism in the initial population consists of the available function genes *F* and terminal genes *T*. Terminal genes are in fact independent variables: strain and coefficient of friction. Random floating-point numbers within the range $[-10, 10]$ are added to the set of terminals to increase the genetic diversities of the organisms. Function genes *F* are basic arithmetical, exponential and cosine functions.

4.1. Evolutionary parameters

The absolute deviation $R(i, t)$ of individual model i (organism) in generation time t for the *GP* approach, was introduced as the standard raw fitness measurement (Ashlock, 2006):

$$R(i,t) = \sum_{j=1}^{n} |E(j) - P(i,j)|, \tag{1}$$

where $E(j)$ is the experimental value for measurement j, $P(i, j)$ is the predicted value returned by the individual model i for measurement j, and n is the maximum number of measurements. The aim of the optimization task is to find those models that Eq. (1) would give as having as low an absolute deviation as possible. However, because it is unnecessary that the smallest values of the above equation also means the smallest percentage deviation of this model, the average absolute percentage deviation of all measurements for individual model i was defined as (Koza, 1992):

$$\Delta(i) = \frac{R(i,t)}{|E(j)|n} \cdot 100\% \tag{2}$$

Eq. (2) was not used as the fitness measurement for evaluating population, but only to find the best organism in population after completing the run.

In the *GP* method, reproduction, crossover, and mutation operations were used for altering the population $P(t)$. Evaluation and altering of the population $P(t)$ were repeated until termination condition had been fulfilled. The termination condition was the prescribed maximum number of generation to be run. Reproduction, crossover, and mutation were used as genetic operations. For example, Fig. 1 shows the operation of the crossover. Two randomly selected parts of two parental organisms (in boldface) are interchanged. Thus two offspring are created.

The evolutionary processes were controlled by the following evolutionary parameters: