Evolution of Empirical Models for Metallurgical Process Systems

Greeff, D.J.

Chemical Engineering Dept.
University of Stellenbosch
Private Bag X1, Matieland, 7602,
Stellenbosch, South Africa
djgreeff@ing.sun.ac.za

Aldrich, C.[#]

Chemical Engineering Dept.
University of Stellenbosch
Private Bag X1, Matieland, 7602,
Stellenbosch, South Africa
ca1@maties.sun.ac.za

[#]Author to whom all correspondence should be addressed

In order to maintain cost-effective operation of process plants, it is often necessary to adjust plants operating conditions, e.g. through recalibration of instrumentation, and adjustment of plant units to accommodate changes in crude oil feedstocks in petrochemical plants, variation in ore feed in mineral processing plants, etc. Virtually all of these adjustments are based on some model or representation of the process, without which adjustment could be highly inefficient. These process models have the added advantage that they can be analyzed to improve understanding of the relevant physical phenomena, while training of process operators to control complex systems or emergency situations can also be facilitated considerably. Since many process systems are not understood sufficiently well to be represented by a first-principles approach, they have to be modelled empirically. Moreover, with inadequate knowledge of the behaviour of the process, considerable time can be spent on a trial-and error determination of a suitable model structure to which parameters can be fitted. This usually requires several assumptions to be made with regard to the variables and the model, e.g. the distribution of the data (usually assumed to be multivariate normal), the independence of the variables, etc. Real world problems seldom meet all these criteria.

Modelling of complex systems can be expedited by the use of genetic programming techniques, which automatically evolve suitable model structures. One of the main difficulties of model development with genetic programming is the proliferation of complex tree-like structures that are difficult to interpret, and which tend to compromise the evolution of generalized models. Several approaches can be followed to curtail tree size explosion or bloating. For example, a penalty factor can be added to the fitness function, which is proportional to the size of the tree. In this study genetic programming was used to evolve empirical models for metallurgical processes. The process data were randomly divided in training and test sets, in order to cross-validate the models.

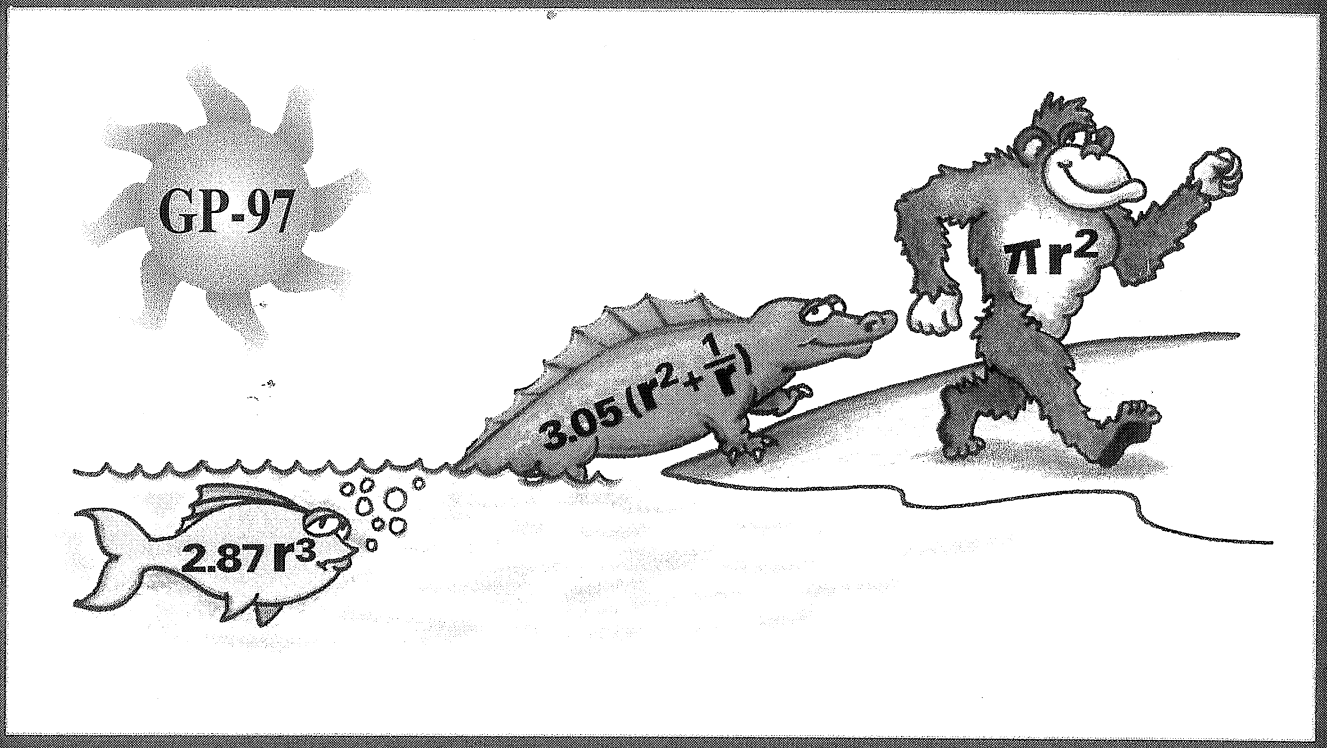
The chief obstacle that had to be overcome with regard to the genetic programming algorithm, concerned the evolution of numerical values for the model parameters. These values were evolved through sub-trees, which were in principle able to produce any rational number desired. However, in practice the trees tended to grow disproportionately large (and their structures difficult to interpret), while runs were more likely to fail to converge.

It was found to be more effective to separate the parameter search from the variable search or model structure evolution, and to determine the parameters by means of standard methods, such as gradient descent techniques, simulated annealing or genetic algorithms.

With these modifications, the genetic programming scheme provided an efficient means for eliminating the trail-and-error procedures of identification of appropriate model structures, and yielded results as least as accurate as those obtained by other least square approaches. The models also compared favourably with representations formed by artificial neural networks.

CONFERENCE PROCEEDINGS

Genetic Programming 1997



Proceedings of the Second Annual Conference
July 13-16, 1997, Stanford University

Edited by
John R. Koza
Kalyanmoy Deb
Marco Dorigo
David B. Fogel
Max Garzon
Hitoshi Iba
Rick L. Riolo