

# MB GP IN MODELLING AND PREDICTION

Carlos Oliver-Morales  
DISCA-IIMAS-UNAM  
Circuito Escolar, Cd. Universitaria  
04510 Mexico City, MEXICO

Katya Rodríguez Vázquez  
DISCA-IIMAS-UNAM  
Circuito Escolar, Cd. Universitaria  
04510 Mexico City, MEXICO

**Abstract.** *The paper describes a hybrid approach for dynamic system modelling. This proposal is mainly based on a Least Squares algorithm and a Multi-Branch Genetic Programming (MB-GP) encoding. Having multiple branches representing an individual allows us to get simpler mathematical expressions, and therefore, reduces the computational evaluation time.*

## 1 INTRODUCTION

The encoding proposed in this work is illustrated in Figure 1. It is composed of an  $n$  number of branches,  $n+1$  coefficients (one for each branch and the constant term) and the addition function. The branches are mathematical expressions representing function terms and encoding as traditional GP structures; however, the maximum depth of these branches are much more lower than the one of an approach using traditional GP encoding. Coefficients of each branch are estimated by means of a Least Squares algorithm. The addition function has the aim of adding the product of each branch with its associated coefficient in order to construct the model.

Crossover operator was specifically defined for the MB-GP encoding. Crossing over two parent individuals consists of selecting randomly a branch in each parent and swapping selected branches. Mutation consists of randomly selecting a branch, eliminating the selected branch, and finally, substituting it for a new branch created randomly from primitive functions.

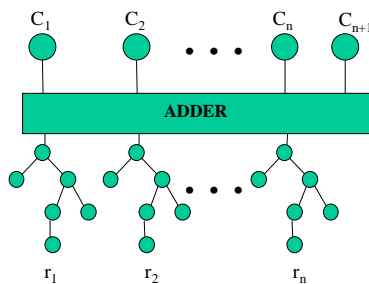


Figure 1. MB-GP encoding.

## 2 PRELIMINARY RESULTS

The data used in this work (local behaviour of temperature in a large period of time) was measured near Mexico City, at Texcoco Lake. Temperature ( $T$ ), relative humidity ( $H$ ), solar radiation ( $R$ ) and wind speed ( $V$ ) and direction ( $D$ ) were recorded. The time interval was 15 minutes.

Based on gathered data, 15% of information corresponding to 324 observations was used for modelling. Experiments were carried out depending on each encoding (MB-GP and Koza-style GP). Cost function was predictive error-based metric. For each experiment, 20 runs were evaluated in order to provide relevant statistical information. The model, which exhibited the best performance from 20 runs, is shown and used to predict future observations (testing data). Based on Koza-style GP (Koza, 1992), the following expression emerged:

$$(+ (divd T_{T-1} (- (+ (- (* V_{T-2} R_{T-2}) (- H_{T-1} R_{T-3})) (divd (divd (exp V_{T-2}) (exp D_{T-3})) (* (- TT-1 D_{T-2}) (divd 0.54707 D_{T-3})))) (cos (sin (sin (sin V_{T-2})))))) (+ T_{T-1} (divd (exp (divd -0.77074 D_{T-1})) (- (+ (exp (exp V_{T-3})) (cos R_{T-2})) (cos (cos (* D_{T-2} T_{T-2}))))))$$

In this case, translating previous expression into a mathematical function turns into a difficult task due to complexity in structure. It is also clear that this expression corresponds to a complex function.

In the case of MB-GP, the following expression was generated,

$$(model T_{T-1} R_{T-1} (* T_{T-2} R_{T-1}) R_{T-3} (* R_{T-3} T_{T-2}) 0.95716 0.021437 -0.00087839 -0.020309 0.00090912 0.02055)$$

Equivalent to,

$$T(t) = 0.0255 + 0.9572 T(t-1) + 0.0214 R(t-1) - 0.0203 R(t-3) - 0.0008 T(t-2)R(t-1) + 0.0009 T(t-2)R(t-3)$$

## 3 CONCLUSIONS AND FUTURE WORK

An alternative representation in GP for dynamic system modelling and prediction was presented. This MB-GP approach has used small values of GP parameters but these preliminary results showed this encoding could produce simple functions and reduce the search space without penalising complex solutions. MB-GP showed also to be consistent in all experiments.

## ACKNOWLEDGEMENTS

Authors gratefully acknowledge the financial support of CONACyT under the project J34900-A. They also thank Dr. Rojano who provided the data.

## REFERENCES

KOZA, J.R. (1992) *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. MIT Press.