Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

An Artificial Immune Systems based Predictive Modelling Approach for the Multi-Objective Elicitation of Mamdani Fuzzy Rules

A Special Application to Modelling Alloys

Jun Chen and Mahdi Mahfouf
Dept. of Automatic Control and Systems Engineering
IMMPETUS, the University of Sheffield
Sheffield, U.K.
Jun.chen@sheffield.ac.uk; m.mahfouf@sheffield.ac.uk

***Abstract*—In this paper, a systematic multi-objective Mamdani fuzzy modeling approach is proposed, which can be viewed as an extended version of the previously proposed Singleton fuzzy modeling paradigm. A set of new back-error propagation (BEP) updating formulas are derived so that they can replace the old set developed in the singleton version. With the substitution, the extension to the multi-objective Mamdani Fuzzy Rule-Based Systems (FRBS) is almost endemic. Due to the carefully chosen output membership functions, the inference and the defuzzification methods, a closed form integral can be deducted for the defuzzification method, which ensures the efficiency of the developed Mamdani FRBS. Some important factors, such as the variable length coding scheme and the rule alignment, are also discussed. Experimental results for a real data set from the steel industry suggest that the proposed approach is capable of eliciting not only accurate but also transparent FRBS with good generalization ability.**

***Keywords*—interpretability, multi-objective immune-based optimisation algorihtm, Mamdani fuzzy modelling, variable length coding scheme**

I. INTRODUCTION

The main aim of this paper is to present a systematic multi-objective fuzzy modeling approach which can simultaneously account for the interpretability of the rule-base and its predictive accuracy. Among the two main fuzzy modeling paradigms, viz. TSK [1] and Mamdani [2] FRBS, the latter is chosen as the main focus of this study due to its unique ability of expressing semantic meanings in its consequents. Generally speaking, a FRBS is a static nonlinear mapping between its inputs and outputs [3], which can be formulated as follows:

$$R_i: If\ x_1\ is\ A_i^1\ and\ x_2\ is\ A_i^2, \dots, and\ x_j\ is\ A_i^n\ Then\ y_i = Z_i$$

where, A_i^j is ith linguistic value (fuzzy set) for the jth linguistic variable x_j defined over the universe of discourse $\mho_j$; the function $\mu_{A_i^j}(x_j)$ associated with A_i^j that maps $\mho_j$ to [0, 1] is the corresponding membership function; R_i represents the ith rule in the rule base, and y_i is the output of the ith rule. Typically, Z_i can be the function of the inputs or the linguistic value of the output, which differentiate FRBS into TSK (the former) and Mamdani (the latter) FRBS. Depending on the form of the function, TSK FRBS can be further divided into Singleton ones (Z_i is the zero order function of the inputs) and linear TSK FRBS (Z_i is the linear function of the inputs). In some sense, Singleton FRBS share the basic feature of Mamdani FRBS if one considers the singleton consequents as a special type of fuzzy sets.

In this paper, a systematic multi-objective Mamdani fuzzy modeling approach is proposed and is organized as follows: section II shortly reviews the existing Evolutionary Algorithms (EAs)-based approaches for improving fuzzy model's interpretability; section III introduces the concept of multi-objective optimization (MO) and the developed three-stage modeling procedure; special attentions have been given to the second modeling stage, and the discussions of the variable length coding scheme and the rule alignment which are closely associated with the problem of the so-called 'unordered set of rules' [4] due to the changeable rule-base structure and the blind search process; experimental results for predicting Tensile Strength (TS) of alloy steels are presented in section IV to validate the proposed modeling scheme; finally, conclusions are given in section V.

II. THE REVIEW OF EAS-BASED FUZZY MODELING

Originated from Karr's work [5], the GA approach was initially utilized to adjust the parameters of membership functions, which makes no significant difference from other learning paradigms. The real significance of employing EAs for optimizing FRBS comes from EAs' flexibility in terms of being able to encode and evolve almost every component of FRBS. Such flexibility offers a solution so that one can take into account the interpretability (structure) and the performance of the FRBS in a more coherent way. Broadly speaking, there currently exist two different EA-based streams to tackle the interpretability issues: the first stream is mainly concerned with the linguistic modeling, in which a set of pre-specified fuzzy partitions are given *a priori* by experts or users (grid partition); the task is then to find an optimal FRBS in terms of its compactness and performance [6~8]; rule selection is a common approach that has been adopted in this line of research; the second stream generally takes the approximate

978-1-4244-2794-9/09/$25.00 ©2009 IEEE

fuzzy models as the start point; hence, the task is to improve the model's explanatory ability, which may have been lost during the automatic learning process, through a set of similarity-driven simplification and parameter adjusting operations [9~13].

The difference between the two streams is derived from the difference of the problems that they are facing. In the linguistic modeling stream, the target problems are normally associated with classifications and low-dimensional function approximations; hence, the effect of the 'curse of dimensionality' due to the grid partition and the need for the parameter tuning due to the performance requirement are not serious issues. In the latter case, high-dimensional approximations are often the case; as a result, an approximate FRBS is a better choice to start with due to the accuracy and compactness requirement. Within the second stream, EA-based multi-objective fuzzy modeling has become a recent hotspot for function approximations due to its ability of producing a set of compromised FRBSs [10~13]. However, this is a rather new developing area with several other issues to be addressed. Among which, the following two are the most important: 1) most well-known multi-objective optimization algorithms used in the fuzzy modeling, e.g. NSGA II [14], are originally designed to solve real-valued problems; in order to use such type of real-valued algorithms to simultaneously optimize the rule base structure and the parameters of membership functions, similarity-driven simplifications are normally selected as the mutation operators for the former [10~11], and the heuristic variations (crossover) are proposed for the latter [10~13]; however, the search power of these optimization algorithms relies heavily on their original variation operators; other components of the algorithms are mainly used to advocate diversity and elitism; without using the original variation operators, even if the general framework is kept it is likely that the search capability, in terms of the real-valued (parameter) optimization part, may be compromised; 2) the reason behind the use of the heuristic variation operators for the parameter optimization is that the structure optimization leads to individuals with different sizes, e.g. rule base length, which makes conventional variation operators (e.g. operators that depend on the interaction between individuals) invalid. Hence, new techniques that can cope with the variable length coding and can facilitate the use of the original variations operators are needed.

With the aim of solving high-dimensional approximation problems, this paper falls into the second stream. To address the above two issues, we extend the research in [15~16], which has been shown to be effective for real-valued MO, to a Mamdani fuzzy modeling scenario, and propose a new distance index [17] that is able to cope with the variable-length individuals and unconstraint optimization. The above features plus the newly proposed BEP updating formulas for the Mamdani FRBS are the main contributions of this paper.

III. A Three-Stage Immune Multi-objective Mamdani Fuzzy Modelling Approach

A. The Definition of MO Problems

Many real-world problems are inherently of a multi-objective nature with often conflicting issues. Generally, MO's aim is to minimize/maximize the vector function (1) subject to J inequality and K equality constraints (2):

$$f(x) = [f_1(x), f_2(x), \dots, f_m(x)]^T \quad (1):$$

$$g_j(x) \geq 0 \quad j = 1, \dots J; \quad h_k(x) = 0 \quad k = 1, \dots K \quad (2)$$

where $x = [x_1, x_2, \dots x_n]^T \in \emptyset$ is the vector of decision variables and $\emptyset$ is the feasible region. Instead of a unique solution, a set of trade-off solutions, or the so-called Pareto solutions, is found in the context of MO. The expression of these Pareto solutions in the objective space forms 'Pareto front'. Fig. 1 shows the Pareto front in a biobjective fuzzy modeling scenario where two competing objectives, viz. the predictive error and the rule-base complexity, are minimized simultaneously. The aim is to find a set of 'Pareto FRBSs' as close to the Pareto front as possible. By finding the set of solutions humans can understand the problem in a much greater depth, and finally a single optimal solution to a specific scenario is finally selected and applied. As mentioned in [10], this results in a low human intervention during the modeling process.

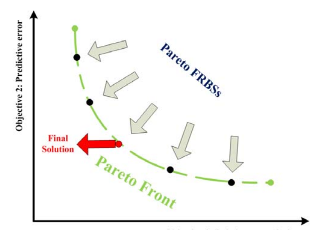

Fig. 1. Pareto front in a biobjective fuzzy modeling case.

B. A Three-stage Immune Multi-objective Mamdani Fuzzy Modeling Approach

Nature-inspired optimization algorithms, such as Genetic Algorithms (GA), have been found to be very promising for MO problems due to the parallel search for a set of solutions. Artificial Immune Systems (AIS) are another instance among such computing paradigms. In [15~16], the authors synergized four immunological models and have proved that a multi-stage immune MO procedure can greatly reduce the computational load of the whole search process. The first stage of the proposed MO procedure emulates the vaccination process, and the second stage is based on the clonal selection and network hypothesis [15]. Based on the immune inspired MO algorithm, in [17], the authors further proposed a three-stage Immune Multi-objective Singleton Fuzzy Modeling (IMOFM_S) approach, in which, the first two stages function exactly the same as the first step of the multi-stage immune MO procedure to extract the so-called 'vaccine model'. Fig. 2 shows the framework of the IMOFM_S, in which Activation calculates the affinity (fitness) for each Antibody (solution) so that an adaptive number of clones can be produced; Affinity maturation mutates the clones so that more search space can be explored; Reselection selects good candidate solutions from the combined parents and clones to provide a selection pressure to effectively drive the candidate solutions towards the Pareto front over many iteration steps; Network suppression is used to regulate the dynamics of the population so that it can adapt to the problem. Although immune algorithms accidentally resemble some characteristics of Genetic Algorithms, a more efficient search could be induced since the population is adaptive.

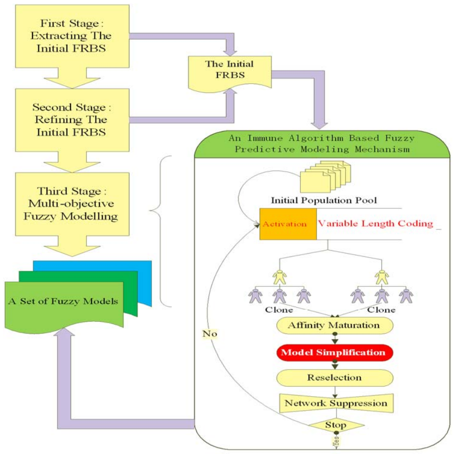

Fig. 2. The framework of the proposed immune based fuzzy predictive modeling methodology.

Although, such a framework was initially proposed for the singleton fuzzy modeling, it is not restricted to such an implementation. This paper extends the framework to its Mamdani version by replacing the singleton consequents with bell-shape membership functions. A set of new BEP updating formulas are thus developed to account for such changes, which forms the baseline for substituting the second modeling stage in IMOFM_S. It is worth mentioning that in the Mamdani implementation the merging operation takes place not only in the premises but also in the consequents. In following space, the three-stage Immune Multi-objective Mamdani Fuzzy Modeling (IMOFM_M) approach is introduced with the special emphasis on the second modeling stage. Special attentions are also given to the variable length coding scheme and its concomitant problem, viz. 'unordered set of rules', because it is somehow overlooked in the past research. Interested readers are referred to [15~17] for more details regarding the immune inspired MO algorithm and the first and the third modeling stages.

1) First stage: An Evolutionary Based Clustering Algorithm-G3Kmeans

Clustering is incorporated into fuzzy modeling especially when numerical data reflects a high dimensionality mapping between input and output spaces. The purpose of clustering is to extract the relationship between independent system variables so that the initial fuzzy structure with only a conservative number of rules can be obtained. In [17], an evolutionary based clustering algorithm-G3Kmeans, representing the combination of G3PCX [18] and K-means is proposed. The purpose of such a combination is to utilize the global search capability of GA to find a set of cluster centers so that a within-cluster-distance criterion is minimized. By doing so, the sensitivity to the initial settings is avoided in the first place, and more importantly a good global fuzzy partition is extracted, which can ease the optimization in the second modeling stage. Interested readers about the detailed steps included in G3Kmeans are referred to [17]. As will be seen in the next part, the Gaussian membership function is used for the inputs in FRBS. In such a case, the identified cluster centers C in the input dimensions correspond directly to the centers of Gaussian membership functions. The spread of the Gaussian membership function is obtained by first calculating the U matrix as follows:

$$U(i,m)=\left(\sum_{l=1}^{k}\frac{\|X_m-C_i\|}{\|X_m-C_l\|}\right)^{-1} \quad (3)$$

Where, $C_1, C_2, \ldots C_k$ are k cluster centers, and U specifies the degrees of data points belonging to each cluster center. Spread σ_i^j is then calculated as follows:

$$\sigma_i^j=\max\left(\sqrt{\frac{-\left(x_m^j-C_i^j\right)^2}{2\cdot\log\left(U(i,m)\right)}}\right) \quad m=1,2,\ldots t \quad (4)$$

where, j indicates the dimension of the spread for the *ith* cluster, t is the total number of data points. With centers and spreads obtained from the clustering algorithm, the Gaussian membership function can be specified as follows:

$$\mu_{A_i^j}\left(x^j\right)=\exp\left(-\frac{1}{2}\cdot\left(\frac{x^j-c_i^j}{\sigma_i^j}\right)^2\right) \quad (5)$$

It is worth mentioning that G3Kmeans is operated on the product space of the inputs and output. Hence, after the first stage, a FRBS with the pre-specified number of rules is extracted from the numerical data with the input membership functions defined by (5). Instead of using Gaussian membership functions, bell-shape membership functions are used for the consequents as described below:

$$\mu_{B_i}(y)=\frac{1}{1+\left(\frac{y-c_i^y}{\sigma_i^y}\right)^2} \quad (6)$$

where, $\mu_{B_i}(y)$ is the ith output membership function; c_i^y is the ith cluster center on the output dimension, and σ_i^y is calculated using eq. (4) on the output dimension. The reason behind the use of bell-shape membership functions for the consequents is that a closed form integral can be deducted in the defuzzification step, which makes the defuzzification step differentiable and thus facilitates the deduction of a set of BEP parameter updating laws for fine tuning Mamdani FRBS.

2) Second stage: Refining the Initial Model with a Constriant BEP

The initial fuzzy model extracted from the first stage is not optimal from two perspectives: 1) the structure of FRBS is not optimal as far as the interpretability is concerned; 2) the membership function parameters need to be tuned further. A constrained back BEP algorithm is thus proposed to first improve the accuracy of the initial FRBS so that a 'vaccine model' can be obtained for the next operation in the multi-objective optimization stage. As mentioned in [13], if the initial population can be constructed using some heuristics, e.g. an optimized FRBS in terms of its predictive performance, then many generations of evolutionary search can be saved. The 'vaccine model' constructed by the first two stages acts similarly to these heuristics.

The original Mamdani FRBS [2] is based on the so-called 'sup-star compositional rule of inference' (eq. (7) ~ (9)) and the overall implied fuzzy set (eq. (9)) [3] defined as follows:

$$\mu_{\hat{B}_i}(y) = \mu_i(X) * \mu_{B_i}(y) \qquad (7)$$

$$\mu_i(X) = \mu_{A_i^1}(x^1) * \mu_{A_i^2}(x^2) * \ldots * \mu_{A_i^j}(x^j) \qquad j = 1 \ldots n \quad (8)$$

$$\mu_{\hat{B}}(y) = \mu_{\hat{B}_1}(y) \oplus \mu_{\hat{B}_2}(y) \oplus \cdots \oplus \mu_{\hat{B}_i}(y), \;\; i = 1 \ldots k \quad (9)$$

where, the 'sup' corresponds to the $\oplus$ operation, and the 'star' corresponds to *. A special instance of the 'sup-star', which uses maximum for $\oplus$ and minimum for *, was adopted in the original Mamdani implementation, and center of average defuzzification was applied on the overall implied fuzzy set in order to derive a crisp output, which leads to two problems as mentioned in [3]: 1) the overall implied fuzzy set $\hat{B}$ is itself difficult to compute, and 2) the defuzzification techniques based on the overall implied fuzzy set are also difficult to compute. More importantly, if an analytical solution cannot be deducted from the defuzzification step the BEP technique cannot be utilized. Hence, in this paper, center of gravity defuzzfication is applied on the implied fuzzy set (eq. (7)). Instead of using minimum, product is used for *. If the Gaussian membership functions are used for the premises and bell-shape membership functions for the consequents, a Mamdani FRBS can be formulated as follows:

$$y^{crisp} = \frac{\sum_{i=1}^{k} b_i \cdot \int_y \mu_{\hat{B}_i}(y)\,dy}{\sum_{i=1}^{k} \int_y \mu_{\hat{B}_i}(y)\,dy} = \frac{\sum_{i=1}^{k} b_i \cdot \mu_i(X) \cdot \int_y \mu_{B_i}(y)\,dy}{\sum_{i=1}^{k} \mu_i(X) \cdot \int_y \mu_{B_i}(y)\,dy} \stackrel{\text{def}}{=} y^{crisp}(X|\theta) \qquad (10)$$

where, b_i is the center of area of the membership function $\mu_{B_i}(y)$ and is the peak if such a membership function is symmetric. $\int_y \mu_{\hat{B}_i}(y)\,dy$ denotes the area under $\mu_{\hat{B}_i}(y)$ over the output interval $y\colon[y_L, y_U]$ and $\int_y \mu_{B_i}(y)\,dy$ can be calculated as below:

$$\int_y \mu_{B_i}(y)\,dy = \sigma_i^y\left[arctan\left(\frac{y_U - b_i}{\sigma_i^y}\right) - arctan\left(\frac{y_L - b_i}{\sigma_i^y}\right)\right] \stackrel{\text{def}}{=} g\left(b_i, \sigma_i^y\right) \qquad (11)$$

$\theta = \left(b_i, \sigma_i^y, c_i^j, \sigma_i^j\right)$ is the parameter vector subject to the minimization of the mean square error. Using the BEP technique, a set of parameter update laws are derived in eqs. (12) ~ (16) for the Mamdani FRBS. If one compares the update laws in this paper with those developed in [17], one will find that the width of the output membership function is also included in eqs. (12) ~ (16). Since there are no constraints on updating these parameters, during the course of the optimization, centres are likely to be placed outside the boundaries. Although this does not affect the ultimate accuracy of FRBS, it does cause confusion for the users when assigning linguistic labels, and more importantly it may violate the search space defined in the next stage. Hence, in this work, a constraint handling scheme is added, which checks the boundary violation for centres during each iteration step and drive any violated centres back to the boundaries. The step size λ and the gain of momentum term β are all set to 0.035 in this work without any loss of generality.

3) *Third stage: Multi-objective Mamdani Fuzzy Modeling*

a) *Forming the objective functions*

Two conflicting objective functions are formulated with the first focusing on the prediction accuracy and the second on the structure simplification as follows:

$$Objective\ 1{:}\, RMSE = \sqrt{\frac{\sum_{m=1}^{t}\left(y_{prediction\_m} - y_{real\_m}\right)^2}{t}} \qquad (17)$$

$$Objective\ 2{:}\, Complexity = Nrule + Nset + RL$$

where, $y_{prediction\_m}$ and y_{real_m} are predicted and real outputs respectively; *Nrule* is the number of fuzzy rules in FRBS; *Nset* is the total number of fuzzy sets; *RL* is the summation of the rule length of each rule.

b) *Forming the initial population pool*

The vaccine model elicited from the first two stages is used to seed the initial population pool, in which a set of initial

$$b_i(t+1) = b_i(t) - \lambda_1 \cdot \varepsilon_m(t) \cdot \frac{\mu_{i(t)}(X_m) \cdot \left[g\left(b_{i(t)}, \sigma_{i(t)}^y\right) + b_{i(t)} \cdot g'\left(b_{i(t)}\right) - g'\left(b_{i(t)}\right) \cdot y^{crisp}\left(X_m|\theta(t)\right)\right]}{\sum_{i=1}^{k} \mu_{i(t)}(X_m) \cdot g\left(b_{i(t)}, \sigma_{i(t)}^y\right)} + \beta_1 \cdot \Delta b_i(t-1) \qquad (12)$$

$$\sigma_i^y(t+1) = \sigma_i^y(t) - \lambda_2 \cdot \varepsilon_m(t) \cdot \frac{\mu_{i(t)}(X_m) \cdot g'\left(\sigma_{i(t)}^y\right) \cdot \left[b_{i(t)} - y^{crisp}\left(X_m|\theta(t)\right)\right]}{\sum_{i=1}^{k} \mu_{i(t)}(X_m) \cdot g\left(b_{i(t)}, \sigma_{i(t)}^y\right)} + \beta_2 \cdot \Delta\sigma_i^y(t-1) \qquad (13)$$

$$c_i^j(t+1) = c_i^j(t) - \lambda_3 \cdot \varepsilon_m(t) \cdot \frac{g\left(b_{i(t)}, \sigma_{i(t)}^y\right) \cdot \left[b_{i(t)} - y^{crisp}\left(X_m|\theta(t)\right)\right]}{\sum_{i=1}^{k} \mu_{i(t)}(X_m) \cdot g\left(b_{i(t)}, \sigma_{i(t)}^y\right)} \cdot \mu_{i(t)}(X_m) \cdot \left[\frac{x_m^j - c_{i(t)}^j}{(\sigma_{i(t)}^j)^2}\right] + \beta_3 \cdot \Delta c_i^j(t-1) \qquad (14)$$

$$\sigma_i^j(t+1) = \sigma_i^j(t) - \lambda_4 \cdot \varepsilon_m(t) \cdot \frac{g\left(b_{i(t)}, \sigma_{i(t)}^y\right) \cdot \left[b_{i(t)} - y^{crisp}\left(X_m|\theta(t)\right)\right]}{\sum_{i=1}^{k} \mu_{i(t)}(X_m) \cdot g\left(b_{i(t)}, \sigma_{i(t)}^y\right)} \cdot \mu_{i(t)}(X_m) \cdot \left[\frac{(x_m^j - c_{i(t)}^j)^2}{(\sigma_{i(t)}^j)^3}\right] + \beta_4 \cdot \Delta\sigma_i^j(t-1) \qquad (15)$$

$$\varepsilon_m \triangleq y^{crisp}(X_m|\theta) - y_m$$

$$where; \qquad g'(b_i) \triangleq g'\left(b_i, \sigma_i^y\right)|_{b_i} = \frac{1}{1+\left(\frac{y_L - b_i}{\sigma_i^y}\right)^2} - \frac{1}{1+\left(\frac{y_U - b_i}{\sigma_i^y}\right)^2}$$

$$g'\left(\sigma_i^y\right) \triangleq g'\left(b_i, \sigma_i^y\right)|_{\sigma_i^y} = \frac{1}{\sigma_i^y} \cdot \left[g\left(b_i, \sigma_i^y\right) + \frac{y_L - b_i}{1+\left(\frac{y_L - b_i}{\sigma_i^y}\right)^2} - \frac{y_U - b_i}{1+\left(\frac{y_U - b_i}{\sigma_i^y}\right)^2}\right] \qquad (16)$$

FRBSs will be randomly generated around the original vaccine model using the following equations:

$$\begin{aligned} {C_{initial}}_i^j &= \alpha \cdot range^j \cdot randn + {C_{vaccine}}_i^j \\ {\sigma_{initial}}_i^j &= \beta \cdot randn + {\sigma_{vaccine}}_i^j \\ {C_{initial}}_i^y &= \alpha \cdot range^y \cdot randn + {C_{vaccine}}_i^y \\ {\sigma_{initial}}_i^y &= \beta \cdot randn + {\sigma_{vaccine}}_i^y \end{aligned} \quad (18)$$

$$range = \min(|C_{vaccine} - U_{limit}|, |C_{vaccine} - L_{limit}|) \quad (19)$$

where, *randn* is a random number within [0, 1]. $range$ defines the minimum interval between the center and its corresponding upper U_{limit} and lower L_{limit} limits of input (or the output) variable, whichever is smaller. The inclusion of $range$ is to ensure that the newly generated centers are most likely within the inputs' (or the output's) domains. Any violation of the domains will be corrected by dragging those centers (or consequents) back to the upper or lower limits, whichever is closest. α and β are the user specified parameters which define how much different the newly generated FRBSs are from the original vaccine one in order to maintain a certain diversity in the initial population, and is set to 0.2 and 0.1 in this work with loss of generality.

c) Variable length coding scheme

Encoding scheme plays a vital role in all kinds of EA-based optimization. As far as the multi-objective fuzzy modeling is concerned, different encoding schemes have been proposed and can be broadly divided into two categories: 1) encoding based on the global data base (linguistic term set); 2) encoding based on the effective rule parameters. The former is mainly found within the linguistic modeling stream [6], [8]; While the latter is mainly found in the approximate modeling stream due to the lack of the global data base in the first place [10~12]. [7] and [13] represent the variants of the first encoding scheme, in which encoding comprises the structure coding and the parameter (data base) coding. The structure coding controls the 'on-and-off' of the genes in the parameter coding.

The drawback of using the first encoding scheme and its variants is that it suffers the 'curse of dimensionality'. In such a case, the length of the chromosome grows exponentially with the increased dimensions, which causes the difficulty for the EAs as far as the search capability is concerned. A typical problem associated with the variants is illustrated in Fig. 6 (a). Since most heuristic search methods rely on the interaction between individuals in the phenotypic space, which is the major thrust directing the search mechanism, an ineffective real-valued optimization may be induced because some active parameter genes (grey ones) may interact with the inactive ones (blank ones). Conversely, if only the effective rule parameters are included in the coding, a variable length coding scheme is inevitable. In [19], one of the first attempts of this kind for designing fuzzy controllers has been proposed. Similar coding scheme can be found in [10~12]. Such a variable length coding scheme, which only encodes effective rules, is also employed in this work to account for the efficiency of the search and the curse of dimensionality. Fig. 6 (b) and (c) give examples of how to encode FRBSs with the different number of rules.

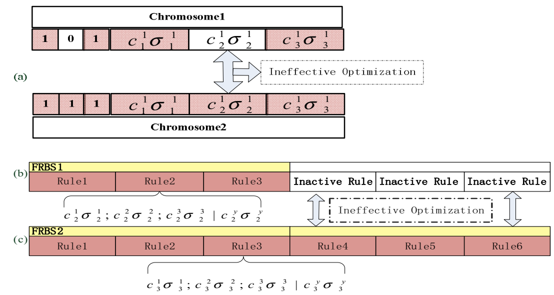

Fig. 6. (a) Ineffective optimization caused by the interaction of inactive gene (blank one) and active gene (colored ones); (b) and (c) variable length coding scheme for a three-rule FRBS and a six-rule FRBS.

Given the variable length coding scheme and the unconstrained optimization, a concomitant effect of the so-called 'unordered sets of rules' [4] may occur, which may also affect the variation operators based on the interaction of individuals. The effect of this is equivalent to combining the mother's gene for good vision and father's gene for curly hair [4]. Hence, an alignment procedure is required to align the closest rules from different FRBSs in order to have a meaningful interaction over the 'unordered sets of rules'. Although this problem has been realized and solved early-on during the development of the binary GA-based fuzzy controller, it was somehow overlooked in the development of the real-valued GA-based fuzzy models. Hence, in this paper a new distance index is proposed to calculate the affinity for the immune algorithm in the activation step. This will facilitate the use of the original effective search operator, viz. affinity maturation. The basic idea is to find the distance of the closest rules in different FRBSs rather than the distance of corresponding rules. The mathematical description of the idea is as follows:

$$dist(R_j, R_k) = \frac{\sum_{i1=1}^{k1} \sum_{l=1}^{rl} \left(R_j^{i1}(l) - R_k^{Ci1}(l)\right) + \sum_{i2=1}^{k2} \sum_{l=1}^{rl} \left(R_k^{i2}(l) - R_j^{Ci2}(l)\right)}{rl \cdot (k1 + k2)} \quad (20)$$

Where, R_j, R_k are two FRBS with $k1$ and $k2$ rules; rl is the length of the rule; R_k^{Ci1} (R_j^{Ci2}) represent the closest rule in R_k (R_j) with respect to the *i1*th (*i2*th) rule in R_j (R_k). The above distance index is used to replace the one in the immune inspired MO algorithm for calculating the affinity (refer to [15~16] for more details about the affinity calculation).

d) Model Simplification

A model simplification step is added into the immune MO algorithm with the aim of removing the redundancy both in rules and in fuzzy sets so that one can pursue the FRBS structure optimization along with the accuracy at the same time. On the rule level: 1) one of insignificant rules (rules that contribute the least to any prediction error increase when not include these rules) is removed for each cloned FRBS at each iteration step unless the rule base reaches the fewest rules designated by the user; 2) one of singleton rules [12] (rules whose comprising fuzzy sets are similar to singleton set) is removed for each cloned FRBS at each iteration step; 3) the most similar pair of rules based on the Similarity of Rule Premise (SRP) [20] are merged for each cloned FRBS at each iteration step. On the fuzzy sets level: 1) one fuzzy set that is the most similar to the universal fuzzy set is labeled as 'Don't care' for each clone at each iteration step; 2) two most similar fuzzy sets from the inputs and output dimensions are merged to form a single fuzzy set for each cloned FRBS at each iteration step based on the similarity measure $S(A_i^j, A_l^j)$ [20]. A set of thresholds which control the various similarity measures are specified by the users. Interested readers are referred to [17] for the detailed steps and user specified parameters involved in the model simplification process. In the following experiment, all the user specified parameters are kept the same as those in [17]. However, during the experiments, we found that the thresholds are not critical parameters due to the following two reasons: 1) only one fuzzy rule or two fuzzy sets are removed or merged at each iteration step; 2) elitism is adopted to record any non-dominated solution found at each iteration step.

IV. Experimental Studies

To validate the proposed modeling framework, it is applied to the modeling of Tensile Strength (TS) of alloy steels. 3760 TS data are used, in which 75% of the data are used for training and the remaining data are used for testing. Another 12 more recent samples are used as the unseen data set to validate the generalisation properties of the model. The TS data includes 15 inputs and one output. The inputs consists of the weight percentages for the chemical compositions, the test depth, the size of the specimen and the site where it has been produced, the cooling medium, the quenching and tempering temperatures. The output is the tensile strength itself. The number of rules used to extract initial FRBS using G3Kmeans is 12, and the initial population is set to 7. Fig. 7 shows the Pareto front after 2000 iteration, in which a set of 47 non-dominated FRBS have been elicited.

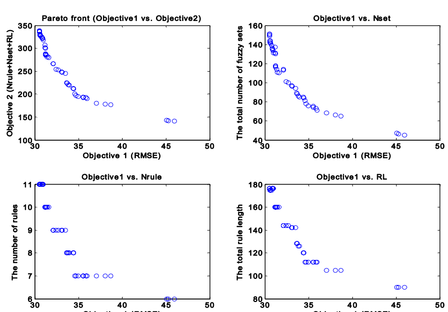

Fig. 7. The Pareto fronts obtained using the proposed three-stage procedure for Example 2: (1) Objective1 vs. Objective2; (b) Objective1 vs. Nsets; (c) Objective1 vs. Nrule; (d) Objective 1 vs. The total rule length.

Table 1 summarized the results obtained using IMOFM_S and IMOFM_M, and compared these results with those in [21]. Due to the constraints on space, only a few obtained 'Pareto' FRBS are presented without any loss of generality. As shown in Table 1, the predictive performance of the initial Mamdani FRBS in the first modelling stage is slightly worse than that of the singleton FRBS and the results presented in [21]. However, using the proposed BEP updating formulas, the accuracy of such an inaccurate Mamdani FRBS has been improved greatly in the second modelling stage. More importantly, a much better generalisation ability has been seen for the refined Mamdani FRBS on the 12 unseen data. After the third modelling stage, a set of 'Pareto' FRBSs are obtained using the proposed IMOFM_M. As one can see, even with a 7-rule simplified FRBS, the predictive performances on the training and testing data sets are better than those presented in [21] using a 12-rule FRBS. The refined singleton FRBS performed badly on the validation data when using a 12-rule FRBS. However, its generalisation ability is much improved when its structure redundancies have been removed. Those redundancies are responsible for the overfitting of the training data, which may lead to a bad generalisation on unseen situations. Since the singleton FRBS is a special type of TSK model, good generalisation ability using fewer rule is expected, e.g. an 8-rule FRBS in IMOFM_S. With a few more rules, Mamdani FRBS represents a competitive generalizer, e.g. a 10-rule FRBS in IMOFM_S.

TABLE I. Modeling Results for Tensil Strength

Modeling Methods	First Stage (clustering algorithm)		Second Stage (single objective refining)		
	Training (RMSE)	Testing (RMSE)	Training (RMSE)	Testing (RMSE)	Validation (RMSE)
[21]	100.54	108.26	37.45	43.07	-
IMOFM_S	113.54	112.32	30.93	35.65	53.61
IMOFM_M	120.43	123.44	31.21	35.49	**37.23**
Third Stage (multi-objective fuzzy modeling)					
Modeling Methods	**No. of rules**	**No. of Fuzzy sets in inputs**		**Modeling performance**	
				Training (RMSE)	Testing/ Validation
[21] Pareto FRBS1	12	Inputs: [9 11 10 12 8 10 8 9 10 10 6 11 10 10 10 10] Output: 10		37.45	43.07/-
Pareto FRBS2	9	Inputs: [97 8 7 5 6 4 6 8 8 2 6 7 8 7], Output: 9		42.82	43.90/-
IMOFM_S Pareto FRBS1	10	Inputs: [4 7 8 8 4 7 3 8 7 7 3 4 4 7 7], Outputs: 10		32.38	34.82/41.01
Pareto FRBS2	8	Inputs: [2 4 4 7 3 3 3 5 4 5 2 2 3 6 6], Output: 8		36.43	37.63/**33.00**
Pareto FRBS3	7	Inputs: [3 4 4 4 1 3 3 4 3 4 1 1 2 6 5], Output: 7		42.91	43.87/46.34
IMOFM_M Pareto FRBS1	10	Inputs: [8 9 10 10 6 10 6 9 9 7 4 7 6 10 9], Output: 10		31.21	35.32/**35.65**
Pareto FRBS2	7	Inputs: [5 7 7 7 2 4 3 6 6 6 2 3 1 7 7], Output: 5		34.70	36.44/**37.23**
Pareto FRBS3	6	Inputs: [2 2 2 5 2 2 1 4 3 3 0 2 1 2 4], Output: 5		45.83	44.30/49.87

Fig. 8 shows the predictive performance of the 7-rule simplified FRBS. Fig. 9 shows the simplified fuzzy sets of in a few selected inputs and output compared to the refined 12-rule FRBS. A much improved interpretability has been achieved so that a set of linguistic terms can be associated with the fuzzy sets in each input and output. Fig. 10 shows a snapshot of the approximate Pareto fronts at 10, 100, 500, 800, 1000 and 2000 iteration respectively. As one can see from this figure, the evolution starts from the refined FRBS and expands the Pareto front during the course of the optimization. The variable length coding and the new distance index play very important roles in expanding such Pareto fronts and in the fine-tuning of the parameters of the evolved simpler FRBS. The MO search

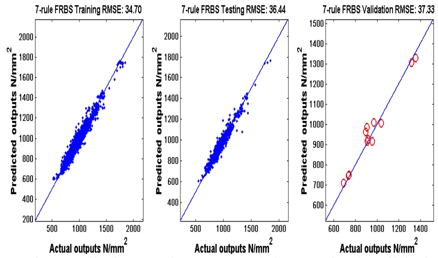

Fig. 8. The predictive performance of the 7-rule simplified FRBS (left to right: training, testing and validation).

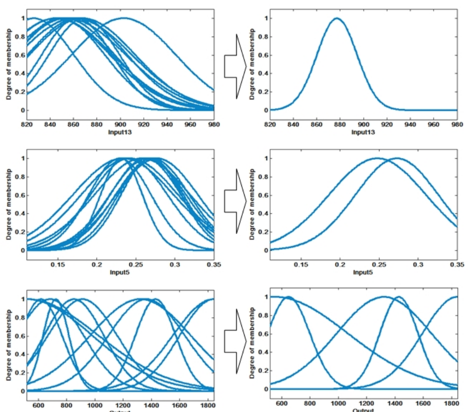

Fig. 9. The fuzzy sets of input 13, 5 and output: (right) the refined 12-rule model; (left) the optimized 7-rule model.

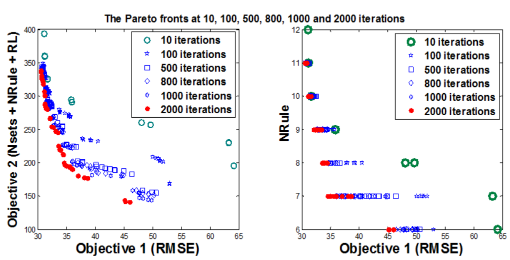

Fig. 10. The Pareto fronts at 10, 100, 500, 800, 1000 and 2000 iteration steps.

process is efficient since after 800 iterations it has already approached very closely to the approximate Pareto front.

V. Conclusions

In this paper, a systematic multi-objective Mamdani fuzzy modelling approach is proposed. By using bell-shape membership functions for the consequents and merging fuzzy set not only in the inputs but also in the output, the resulted consequents are more interpretable compared to the singleton ones. The parameter updating formulas developed in the second modelling stage using BEP technique could be used separately in order to improve the predictive performance of Mamdani FRBS. It is also worth mentioning that the proposed three-stage modelling framework is fairly general since one can easily replace any one or all of the modelling stages with their own developed single objective and multi-objective optimization algorithms. The mentioned 'unordered sets of rules' is a common problem associated with all EA-based fuzzy modelling approaches. Hence, special cautions should be taken when one intends to devise a variation operator based on the interaction of individuals.

References

[1] T. Takagi, M. Sugeno, "Fuzzy Identification of Systems and Its Application to Modeling and Control", *IEEE Trans. Syst., Man, Cybern.*, vol. 15, pp. 116-132, 1985.

[2] E. H. Mamdani, "Applications of Fuzzy Algorithm for Control a Simple Dynamic Plant", in *Proceedings of Proceedings of IEEE*, vol. 121 (12), pp. 1585-1588, 1974.

[3] K. M. Passino, S. Yurkovich, *Fuzzy Control*, Addison-Wesley, 1998.

[4] L. Magdalena, 'Crossing Unordered Sets of Rules in Evolutionary Fuzzy Controllers', International Journal of Intelligent Systems, vol. 13 (10/11), pp. 993-1010, 1998.

[5] C. L. Karr, "Genetic Algorithms for Fuzzy Controllers", *AI Expert*, vol. 6 (2), pp. 26-33, 1991.

[6] H. Ishibuchi, T. Yamamoto, "Fuzzy Rule Selection by Multi-Objective Genetic Local Search Algorithms and Rule Evaluation Measures in Data Mining", *Fuzzy Sets and Systems*, vol. 141, pp. 59-88, 2004.

[7] R. Alcal□, M. J. Gacto, F. Herrera, "A Multi-Objective Genetic Algorithm for Tuning and Rule Selection to Obtain Accurate and Compact Linguistic Fuzzy Rule-Based Systems", *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 15 (5), pp. 539-557, 2007.

[8] M. Cococcioni, P. Ducange, B. Lazzerini, F. Marcelloni, "A Pareto-based Multi-objective Evolutionary Approach to the Identification of Mamdani Fuzzy Systems", Soft Computing, vol. 11, pp. 1013-1031, 2007.

[9] M. Setnes, R. Babuška, U. Kaymak, H. Lemke, "Similarity Measures in Fuzzy Rule Base Simplification", *IEEE Transactions on Systems, Man, and Cybernetics-Part B*, vol. 28 (3), pp. 376-386, 1998.

[10] F.Jiménez, A. F. Gómez-Skarmeta, H. Roubos, R. Babuška, "Accurate, Transparent, and Compact Fuzzy Models for Function Approximation and Dynamic Modeling through Multi-Objective Evolutionary Optimization", in *E. Zitzler et al. (Eds.): EMO 2001, LNCS* 1993, pp. 653-667, 2001.

[11] F.Jiménez, G. Sánchez, A. F. Gómez-Skarmeta, H. Roubos, R. Babuška, "Fuzzy Modeling with Multi-Objective Neuro-Evolutionary Algorithms", *IEEE International Conference on Systems, Man, and Cybernetics*, vol. 3, 2002.

[12] H. L. Wang, S. Kwong, Y. C. Jin, W. Wei, K. F. Man. "Multi-objective Hierarchical Genetic Algorithm for Interpretable Fuzzy Rule-based Knowledge Extraction", *Fuzzy Sets and Systems*, Vol. 149 (1), pp. 149-186, 2005.

[13] J. González *et al.*, "Improving the Accuracy While Preserving the Interpretability of Fuzzy Function Approximators by Means of Multi-objective Evolutionary Algorithms", *International Journal of Approximate Reasoning*, 2006.

[14] K. Deb, Multi-Objective Optimization using Evolutionary Algorithms, Wiley, Chichester, U.K., 2001.

[15] J. Chen, M. Mahfouf, "A Population Adaptive Based Immune Algorithm for Solving Multi-Objective Optimization Problems", *ICARIS 2006, LNCS* 4163, pp. 280-293, 2006

[16] J. Chen, M. Mahfouf, "Artificial Immune Systems as a Bio-inspired Optimization Technique and Its Engineering Applications", in *Artificial Immune Systems and Natural Computing: Applying Complex Adaptive Technologies,* H. W. Mo (eds.), 2008.

[17] J. Chen, M. Mahfouf, "An Immune Algorithm Based Fuzzy Predictive Modeling Mechanism using Variable Length Coding and Multi-objective Optimization Allied to Engineering Materials Processing", in *proceedings of the 2008 IEEE International Conference on Granule Computation (GrC 2008),* pp.26-28, 2008.

[18] K. Deb, A. Anand, D. Joshi, "A Computationally Efficient Evolutionary Algorithm for Real-Parameter Optimization", *Evolutionary Computation*, vol. 10 (4), MIT Press, pp. 371-395, 2002.

[19] M. G. Cooper, J. J. Vidal, "Genetic Design of Fuzzy Controllers: The Cart and Jointed-Pole Problem", *in Proceedings of the Third IEEE Conference on Fuzzy Systems*, vol. 2, pp. 1332-1337, 1994.

[20] Y. Jin, "On generating FC^3 Fuzzy Rule Systems From Data Using Evolution Strategies", *IEEE Trans. On SMC*, vol. 29 (6), pp. 829-845, 1999.

[21] Q. Zhang, M. Mahfouf, "Fuzzy Predictive Modeling Using Hierarchical Clustering and Multi-Objective Optimisation for Mechanical Properties of Alloy Steels", in *Proceedings of the 12th IFAC Symposium on Automation in Mining, Mineral and Metal Processing*, 2007.