

GA optimized FRBS for travel systems

J.H. Ricketts^a and D. Olaru^b

^a *Queensland Climate Change Centre of Excellence, Ecosciences Precinct, 41 Boggo Rd, Dutton Park, Queensland.*

^b *Management and Organisations, Faculty of Business, University of Western Australia, 35 Stirling Hwy, Crawley, WA (Email: doina.olaru@uwa.edu.au.)*

Abstract: This paper reports further on work presented at MODSIM09 (Ricketts 2009), which itself extended modeling of behavioural rules for activity scheduling in the transport domain as previously published (Olaru, Smith 2005). We again report on the use of genetic algorithms to tune a modified Mamdani fuzzy rule based system (FRBS), this time focusing on the optimization of the rule set composition.

A Mamdani fuzzy knowledge base system is a fuzzy logic rule based system (FRBS) initially proposed by (Mamdani 1974) as a fuzzy logic controller. One uses a combination of fuzzification, fuzzy inference and defuzzification together with a knowledge base comprising database of fuzzy sets and a rule-base of fuzzy rules. This version substitutes fuzzy selection for defuzzification and, after training, outputs predicted travel schedule decisions given a coding of an individual's situation. The tuning of such a system is an open problem.

To tune the system, two genetic algorithms were applied. One, the rule-base GA (*rb-GA*), taking the database as fixed, attempts to firstly maximize the classification rate and secondarily minimize the size of the rule base. The other, the fuzzy set GA (*fs-GA*) attempts to pre-tune the partitioning of the fuzzy sets for the *rb-GA* using an information entropy-like measure as a heuristic. The *fs-GA* was the focus of (Ricketts 2009), and hence the *rb-GA* is the focus of this report.

The population consists of 100 chromosomes, each of which is of variable length of up to 100 genes. Each gene is a pair, representing an antecedent rule and a consequent. Crossover is an asymmetric single-point variable-length operator. Several mutation operators are defined, point mutation, delete, extend, inversion. Selection is elitist, using a ranking system. Maturation consists of the translation of a chromosome into a full FRBS, and fitness is the proportion of correct predictions, further weighted according to a minimal defining length criterion to first prefer accuracy and then compactness.

The effect on the convergence, classification rates and rule-set compactness of four variations was investigated.

- The type of fuzzy rule composition/combination on the classification and compaction of the FRBS rule set; whether the “*single-best*” operator of (Olaru, Smith 2005), or a new operator we call “*first-best*”.
- The degree of disruption within the GA on the convergence properties of the GA,
- The effects of a specific selection strategy that we call *heuristic mate selection* on the convergence properties of the GA.
- The effects of pre-tuning the FRBS partitioning parameters. This is already reported in (Ricketts 2009).

Heuristic mate selection was not found to be effective. *Single-best* composition was found to give better classification rates, but *first-best* composition produced much more compact rule sets with better entropic qualities. The degree of disruption in the GA had an effect in that it converged on better classification rates at the cost of larger rule sets. Pre-tuning, as previously reported, led to more compact rule set and better classification rates.

Keywords: Mamdani, Fuzzy Rule Base System (FRBS), Genetic Algorithm (GA), Trip Scheduling, Information Entropy

1. INTRODUCTION

1.1. The Trip Scheduling Problem

People are required to make trip-scheduling decisions in their daily lives. They deal with a rich set of uncertainties – changes of venue, priority, and time of activities – choices in mode, route, and time of trips – information horizons – delays and changes in current activities, and in trips. Populations both evolve and contain, a rich set of applicable strategies. As previously stated in (Ricketts 2009), Olaru and Smith (Olaru, Smith 2005) published a study based on transport decisions in hypothetical situations made by 126 respondents in a follow-up survey after a larger survey of students and academics in Bucharest in November 1998. They attempted to obtain a set of travel decision rules. Olaru and Smith showed that a Mamdani FRBS (Cordón 2001) allowed them to “examine a broader and richer field than traditional methods” and “represents the real system in a form close to human perception”. Their system had 56 rules and achieved 82-87% successful classification rates. This problem has been further analysed, concentrating on the rule sets, the formal grammar of the fuzzy rules, and the composition rules. The problem is now cast as a learning/classification challenge, where the inputs are coded as real numbered tuples, and outputs of the FRBS are discrete values representing predicted responses of individuals to their situation.

1.2. Fuzzy Rule Based Systems

Fuzzy rule based systems (FRBS) use a fuzzy logic (Zadeh 1965) inference engine to process a knowledge base encoded as a combination of a database of facts and a rule-base of inferential rules. The Mamdani FRBS was proposed initially proposed as a fuzzy logic controller (Mamdani 1974). This research compares two inferencing/composition and defuzzification approaches emphasizing two facets: *competition* (*first-best* maximises matching/the degree of association of the pattern to the class) and *collaboration* (as (Cordón 2001) indicated, *single-best* considers classification by other active rules (p.96)). The latter operator was used in Olaru and Smith.

1.3. Genetic Algorithms

Genetic algorithms (GAs), first described in (Holland 1975), are useful for optimisation problems including both the database tuning and rule-base structure of FRBSs (Cordón 2001, Maniadakis, Surmann 1999). They make use of objective functions, implicitly utilising an error space, and thus have to deal with the topology and curvature of that space. We examine four distinct areas in the application of GA to optimising the FRBS with respect to classification capability and rule complexity of a modified Mamdani FRBS, namely:

1. The effects of the type of fuzzy rule composition/combination.
2. The effects of pre-tuning the FRBS partitioning parameters.
3. The effects of varying the degree of disruption within the GA on the convergence properties of the GA, leading secondarily to variation of classification capability, and parsimony of the FRBS.
4. The effects of a specific selection strategy that we call heuristic mate selection on the convergence properties of the GA.

Disruption in a GA has sometimes been characterized as a problem, defined in terms of the schema theorem as the probability of a schema surviving intact from generation to generation (Falkenauer 1996). It has also been shown to improve convergence in difficult topographies (Kuo, Hwang 1996). Amongst reasons for non-convergence are, stalling due to loss of population diversity, undirected search due to loss of schemas, and the topography of the error space. Thus, one aspect of a GA worth exploring is the degree of disruption. Heuristic mate selection was conceived as a different approach to the same problem, and was implemented as a part of the selection process whereby a selected breeder chromosome had the chance to “reject” a potential mate.

2. THE MODIFIED MAMDANI SYSTEM

This version of the Mamdani FRBS substitutes selection for defuzzification since the application is not a control one but a classification one with disjoint classes. The modified FRBS is fed tuples of real values. The selection interface takes the fuzzy partition from the inference system, and outputs a discrete value, as illustrated in Figure 1. At least one rule of n rules fires for each decision situation, y , and each firing produces a fuzzy set membership $\mu_{B'_n}(y)$.

The rules are *composed* using two forms:

- a *first-best* rule (similar to what Cordon calls aggregation by means of the maximum t-conorm)

$$\mu_{B'}(y) = \max\{\mu_{B'_1}(y), \dots, \mu_{B'_n}(y)\} \tag{1}$$

- *single-best* (sum activations), similar to what Olaru and Smith (2005) used in their paper.

$$\mu_{B'}(y) = \sum \mu_{B'_n}(y) \tag{2}$$

Single-best composition tends to act to give the system some of the properties of a neural network, i.e. the behaviour of the rule base as a whole depends on the interaction between specific rules. *First-best* presumes that at most one rule dominates for any one situation and is more closely related to Boolean logic. In both cases, to break a tie, the first encountered outcome is selected.

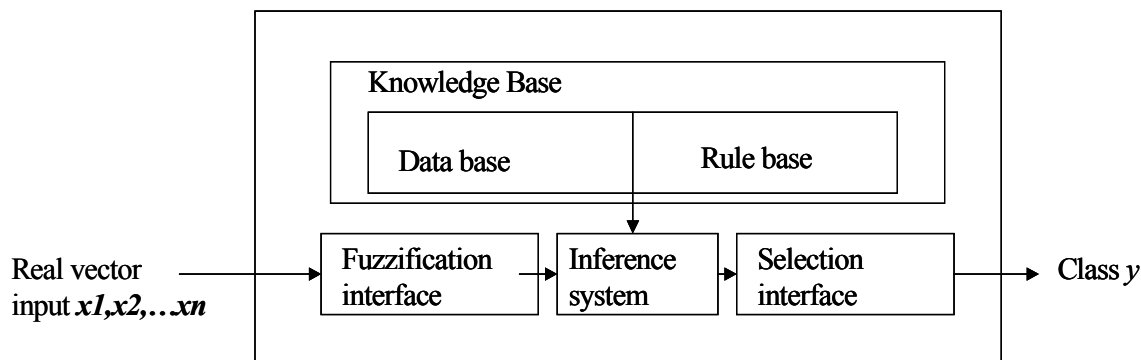


Figure 1. The modified Mamdani FRBS with Selection replacing de-fuzzification. (From (Ricketts 2009), after(Cordón 2001))

3. THE GENETIC ALGORITHM

Two different Genetic Algorithms (GAs) were used in this study for different purposes. The “rule-base” GA (*rb-GA*) and a “tuning” GA (*fs-GA*). The *fs-GA* was used to provide alternative parameters for the fuzzy membership functions (FMFs) in the FRBS as a contrast to those deduced by (Olaru, Smith 2005). The details of this are given in (Ricketts 2009), noting here again that the *fs-GA* differs from other work in that it is a pre-tuning step using a heuristic, executed before the rules are established. The *rb-GA*, the major subject of this report, was used to optimise the FRBS with respect to both the number of rules and the maximum classification rate. The rest of this section describes the *rb-GA*.

3.1. Rule Structure

Each rule consists of an implication with monotone monomials in the antecedent, and a single literal consequent. Note that for each adjective there is a default `frb_X_none` defined (where ‘X’ is replaced by ‘f’, ‘ts’ etc), and simply means “do not consider”, and has the effect of generalising the rule – equivalently, of shortening it. E.g. “IF flexibility=flexible AND time_saving=small AND time_of_day= night AND duration_next_activity=short THEN action=remove”. I.e. if the schedule is flexible and the time saving is small and the time of day is night and the duration of the next activity is short, then remove the next item from the schedule. The list of antecedents and consequents is shown in Figure 2.

3.2. Chromosome

Given the definitions in Figure 2, there is a limited number (768) of different possible antecedent rule sets each of which will have a single consequent. Thus, each gene is a pair of integers, the antecedent part and the consequent part. We will refer to this form as being *integer-pair* form, and its translation as *implication* form. Each individual chromosome consists of a variable length chromosome of these genes. Since the rule count deduced by (Olaru, Smith 2005) was 56, the maximum chromosome was limited to 100 genes.

3.3. Population and Breeding

The population consists of 100 chromosomes. Additionally the population maintains a fitness value for each chromosome. Breeding consists of the application of two phases of selection and the application of various

operators to the population. After application of the selection policy the population consists of parents and their offspring in equal numbers. The fittest parents are preserved unaltered but the majority of the population is subjected to a bout of mutation using a random selection of mutation operators.

Antecedent conditions and their adjectives

- frb_flexibility = (frb_f_none, frb_f_rigid, frb_f_not_rigid, frb_f_flexible);
- frb_time_saving = (frb_ts_none, frb_ts_neg_large, frb_ts_neg_small, frb_ts_small, frb_ts_medium, frb_ts_large);
- frb_time_of_day = (frb_tod_none, frb_tod_night, frb_tod_early_am, frb_tod_am, frb_tod_noon, frb_tod_early_pm, frb_tod_later_pm, frb_tod_evening);
- frb_duration_next_activity = (frb_dna_none, frb_dna_short, frb_dna_medium, frb_dna_long);

Consequents

- frb_action = (frb_a_none, frb_a_do_nothing, frb_a_change_time, frb_a_change_duration, frb_a_location, frb_a_remove, frb_a_new, frb_a_change_transport);

Figure 2. The four antecedent conditions, each of which is represented by a partition of fuzzy sets. Each of the fuzzy sets except the * *none* represents the degree of membership applicable to one “linguistic adjective”. The consequents are by contrast not a linear partition, but disjoint fuzzy sets.

3.4. Selection and Mate Selection

Selection consists of the selection of a breeding pool (based on fitness), and the separate phase of mate selection or pairing. A form of heuristic mate selection was introduced in an attempt to preserve differences between parents, selecting partners randomly until either five attempts had been made or a partner with a different length or different fitness score. This is a rather weak heuristic, but serves to prevent incest. Selection is *elitist* ranking selection, in which individuals are first ranked by fitness and then working from fittest to least fit, the first 50 individuals mate with selected members of the old-population. Offspring of each mating replace the least fit individuals, and the size of the old population is reduced accordingly.

3.5. Crossover

Crossover in this case was an asymmetric single point crossover where a random crossover point was computed for each parent separately and two new strings were composed from the head of one and the tail of the other. During appending of the tail parts each gene is eliminated if its antecedent is already part of this new chromosome to prevent contradiction. As a result of this policy the two offspring are generally of different lengths.

3.6. Mutation

No mutator may create a chromosome that has duplicate antecedent parts in order to avoid contradiction.

- *Point Mutation*: A randomly selected locus in the antecedent part *or* the consequent of one gene is incremented or decremented. If a duplicate antecedent is generated then the antecedent is replaced at random.
- * *Delete*: Delete a sequence of genes from a chromosome. No more than one quarter of genes is deleted.
- * *Shuffle*: Up to eight pairs of genes chosen at random are swapped within a chromosome.
- * *Branch*: Extends the chromosome with a rule generated by a point mutation of a randomly selected existing gene.
- *Inversion*: A sequence of genes is simply inverted in place.
- *Extend*: Extends the chromosome with a randomly generated gene.

Shuffle and Inversion are two similar strategies for regrouping genes so that crossover has a chance to test different building blocks. * denotes operators used only for runs defined as *disruptive*.

3.7. Decoding, Evaluation and Fitness

Each chromosome is interpreted as a fuzzy rule set and tested for its ability to correctly partition the entire set of test cases by executing it in a Mamdani FRBS, against the set of examples. Fitness of an individual f_{ind} , is

the proportion of correct classifications $N_{correct}$, made against the test population N_{pop} , weighted by reduced rule count $N_{max} - N_{ind}$. This gives a ranking based mostly on strength of prediction and secondarily to compactness of representation.

$$f_{ind} = \frac{1}{N_{pop}} \left(\frac{(N_{max} - N_{ind})}{N_{max}} + N_{correct} \right) \quad (3)$$

3.8. Pre-tuning

The pre-tuning step (Ricketts 2009) was tested in this work using a single execution of the *fs-GA*. As such the Mamdani FRBS was tested with two different data-bases. The ‘‘Olaru’’ sets are as reported in (Olaru, Smith 2005) and the ‘‘tuned’’ sets are those derived from separate runs of the *fs-GA*.

4. EXPERIMENTS AND RESULTS

The *rb-GA* was run with the following variations. Each run was halted after 10,000 generations.

1. Two types of fuzzy combination were used within the Mamdani machine: (a) *single-best*, and (b) *first-best*.
2. Two different sets of genetic operator were used within the GA: (a) a simple set using crossover, mutation plus single gene extension and single gene shortening, and inversion; (b) disruptive GA with all those operators, plus deletion, shuffling and branching.
3. Additionally all the above were run: (a) with a heuristic mate selection operator and (b) without.
4. Two different approaches to fuzzy set tuning were tested: (a) the original sets published by Olaru (replicated four times), and (b) the outcome of four different tunings using the *fs-GA* (above).

This gave 64 separate runs of the GA.

For each run we record the number of rules in the rule-base, the proportion classified correctly (both of these quantities being part of the fitness criterion). From the last two a mean entropy per rule can be computed, using a variant of standard Shannon entropy, $H = -\sum \text{Pr}(x) \ln(\text{Pr}(x))$ where $\text{Pr}(x)$ is in this case the probability of the Mamdani FRBS correctly classifying a randomly selected antecedent. Since the GA is elitist, the generation number in which the final elite solution first appeared was also recorded.

4.1. Statistical Methods

Using the R package (<http://www.r-project.org/>), we used ANOVA with orthogonal contrasts to determine which variation/factors or combinations of factors were significant. It was immediately apparent that heuristic mating had no effect on mean rule counts and on proportion classified. The three factors included in the ANOVAs reported here were *combination*, *disruption* and *tuning*. These were tested for their effect on rule count and on proportion classified.

4.2. Analysis

Proportion Classified ANOVA of Proportion Classified showed that degree of disruption of the GA is significant and both fuzzy combination type and the tuning of the fuzzy sets were highly significant, with highest mean classification rate being shown by a disruptive GA running against tuned fuzzy sets and with Single-best composition. The highest rate overall rate of 86% was attained with this combination with 81 rules but classification rates of 84%, with 49 rules were also achieved this way. Using *first-best* combination classification rates were generally lower although with much lower rule counts. One instance run with a disruptive GA, first-best combination and tuned fuzzy sets gave 84% classification with 47 rules, however 79% was attained with just 33 rules.

Rule counts ANOVA of Rule Counts showed that both the degree of disruption of the GA and fuzzy combination type are highly significant and that they interact strongly. Mean rule counts were much lower for *first-best* combination, ranging from 20 to 47 with a median of 30, as compared to *single-best* with a range of 33 to 89 with a median of 53. The disruptive GA showed a strong tendency to push up rule counts, especially in combination with *single-best* combination.

Mean Rule Entropy This measure is a function of both rule count and proportion classified. Due to space constraints this is reported without further analysis. The degree of disruption of the GA and fuzzy

combination method are highly significant, fuzzy set tuning is significant and there is some interaction between combination and tuning. Tuning reduced rule counts and increased mean rule entropy, disruption operated vice-versa.

Termination time/generation ANOVA showed that the degree of disruption of the GA and fuzzy combination method are highly significant with the disruptive GA terminating much later on average, especially in combination with *single-best* combination, and at the cost of higher rule counts.

Table 1: Means (over a total of 64 runs) of Proportion Classified, Rule Count, Termination Generation, and Mean Entropy per Rule, for each of eight variants after heuristic mate selection was pooled.

Run Parameters	Code	Proportion Classified	Rule Count	Termination Generation	Mean Rule Entropy
GA Disruptive, Combination FirstBest, Sets Olaru	DFO	0.7309	28.38	5277	0.0211
GA Disruptive, Combination FirstBest, Sets Tuned	DFT	0.7602	33.38	6033	0.0173
GA Disruptive, Combination SingleBest, Sets olaru	DSO	0.7626	66.75	9393	0.0084
GA Disruptive, Combination SingleBest, Sets Tuned	DST	0.8060	57.13	9334	0.0093
GA Simple, Combination FirstBest, Sets Olaru	SFO	0.7112	26.38	3306	0.0234
GA Simple, Combination FirstBest, Sets Tuned	SFT	0.7357	30.13	3990	0.0194
GA Simple, Combination SingleBest, Sets Olaru	SSO	0.7429	39.38	7471	0.0146
GA Simple, Combination SingleBest, Sets Tuned	SST	0.7753	44.00	6464	0.0124

5. DISCUSSION

5.1. Combination procedure

If the aim is to generate easily analysed rules then *first-best* seems a good choice. The rules interact in an easily understood and interpreted way. A rule has utility only if its activation is maximal for some situation.

Single-best however seems to give better classification performance but at the cost of an increased number of rules, and possibly at the cost of increased complexity in the topology of the objective function surface. The latter is shown by the tendency of the GA to stall (show epistasis), given the number of runs which gave high rule counts and low classification rates. It first sums activations, and then selects the maximally activated set. Since there is a summing step, multiple rules may contribute to the classification of a single situation. This has several effects, (a) the FRBS acquires a slightly neural net like property as knowledge is distributed within the rule set, (b) the complexity of the problem the GA must solve is increased by the interaction between rules. Rules are more complex in their effects and so are harder to understand. The explanatory power of the rule set, as far as a human analyst is concerned, is reduced.

5.2. Disruption

Disruption was intended to prevent premature termination of the GA by providing more variation in the population. Operators which extend the chromosome length are intended to provide extra sites for mutation to operate on. With *first-best* combination, this is relatively benign since there is no interaction between rules, and extra rules have no effect on classification unless they are improvements. The only other effect would be to dilute the effect of mutations which may slow the progress of the GA marginally. With *single-best* combination, randomly introduced rules have immediate and unpredictable effects. Since rules interact, introducing a number of random rules can increase the activation of fuzzy sets and change the classification of the system, not because strong rules predominate, but because a number of weak rules can dominate a single strong one.

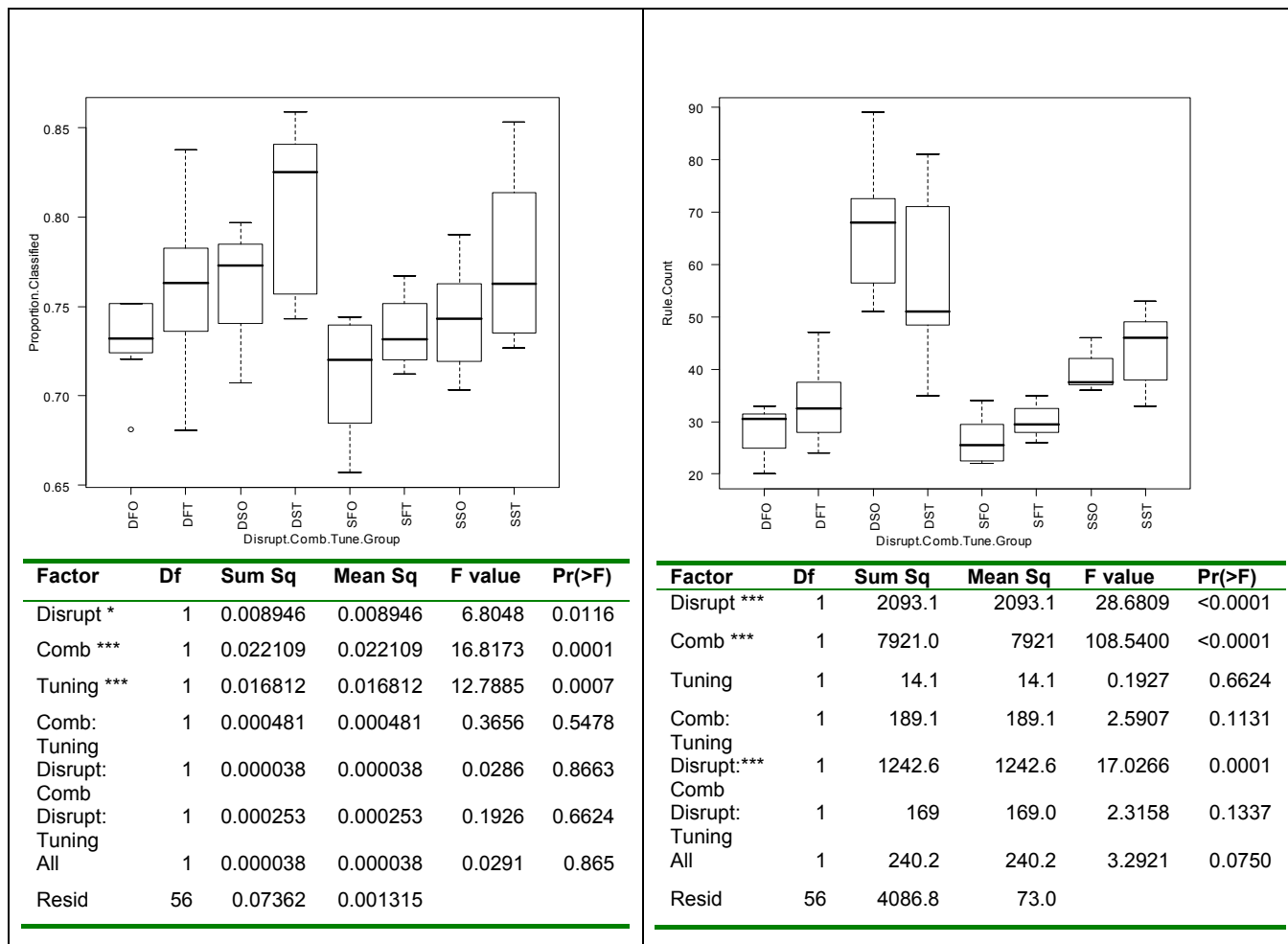


Figure 2: Box and whisker plots, and ANOVA tables showing the effect of modifying disruption, the fuzzy combination and pre-tuning. The codes on the y-axes are from the second column of Table 1.

REFERENCES

Cordón, O. 2001, *Genetic fuzzy systems: evolutionary tuning and learning of fuzzy knowledge bases*, World Scientific Pub Co Inc.

Falkenauer, E. 1996, "A hybrid grouping genetic algorithm for bin packing", *Journal of Heuristics*, vol. 2, no. 1, pp. 5-30.

Holland, J. 1975, *Adaptation in natural and artificial systems*, Addison-Wesley Publishing Company, Reading, Massachusetts.

Kuo, T. & Hwang, S.Y. 1996, "A genetic algorithm with disruptive selection", *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 26, no. 2, pp. 299-307.

Mamdani, E.H. 1974, "Application of fuzzy algorithms for control of simple dynamic plant", *Proc.Iee*, vol. 121, no. 12, pp. 1585-1588.

Maniadakis, M. & Surmann, H. 1999, "A genetic algorithm for structural and parametrical tuning of fuzzy systems", *European Symposium on Intelligent Techniques, ESIT99*, June 3, 1999.

Olaru, D. & Smith, B. 2005, "Modelling behavioural rules for daily activity scheduling using fuzzy logic", *Transportation*, vol. 32, no. 4, pp. 423-441.

Ricketts, J.H. 2009, "Tuning a modified Mamdani fuzzy rule base system with a genetic algorithm for travel decisions", 18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation, Cairns, Australia, 13-17 July 2009, pp. 768-774.

Zadeh, L.A. 1965, "Fuzzy sets*", *Information and control*, vol. 8, no. 3, pp. 338-353.