Fuzzy Modeling Using Genetic Algorithm and Its Application to Biomechanics

Kanta Tachibana

Fuzzy Modeling Using Genetic Algorithm and Its Application to Biomechanics

Kanta Tachibana

Contents

Chapter 1

Introduction

1.1 Background

1.1.1 Knowledge Acquisition from Data

Nowadays, scientists and engineers are well concerned with complex systems. These systems have multi-dimensionality, hierarchical structures, mutual interactions between variables, internal feedback mechanisms, and/or unpredictable dynamics. It is difficult to comprehend such systems only with mechanical/physical models. A paradigm shift from modeling based on first-principle to modeling from data has been underway[1].

The first-principle is a basic approach, on which modern science and engineering has been developed. This approach starts with a basic scientific model, e.g. Newton's laws of mechanics or Maxwell's theory of electromagnetism, and then builds upon them various applications in engineering. In the actual applications, however the underlying first principles are often unknown or the target systems, for example, physical, biological, and/or social systems, are too complex to be mathematically described. Without first-principle models, we have been required to construct models from data. Thanks to the rapid development of computer/semiconductor technologies, data collection and storage from such systems have become easy. Thus, modeling from such readily available data for describing unknown input-output dependencies of target systems is currently attracting much attention.

Common modeling procedure from data can be summarized as

Step 1: Problem Statement and Hypothesis Formulation First, a problem is specified in a particular research field. Next, some hypotheses are formulated to solve the problem. For each of the hypotheses, a set of input variables and output variables are specified. Most data modeling studies are carried out in a specific application domain. Therefore, the modeler must involve experts of the application for meaningful problem statement and for appropriate hypothesis formulation.

Step 2: Experiment Design and Data Collection The next step is designing experiment followed by collecting data. If the modeler can control all of the input variables, the experiment design is easy. If this is not the case, the data are collected from a database. The training data, with which the model will be identified, and the future data, which are used for the prediction, must come from the same population and through the same sampling distribution. Otherwise, the predictive accuracy of the obtained model would be poor.

Step 3: Data Preprocessing This step is executed to remove outliers and to select informative features. Most of outliers are due to errors in the measurement. However, the modeler should be circumspect in removing outliers because some of the outliers may represent input-output dependencies. Anyhow, the outliers largely affet the structure and the parameters of the obtained model. After the outlier removal, a small number of features are selected out from the multidimensional data. Good feature selection enables identification of a concise and accurate model in the next step. The success of the whole procedure is generally due to good preprocessing of data. Involving knowledge and experience of the application experts is indispensable for reasonable preprocessing.

Step 4: Model Identification The hypotheses, formulated in Step 1, are quantified. The structure and the parameters of model are identified using the training data. The main goal of this step is to achieve high predictive accuracy. However, the interpretability of the model should be considered to draw clear conclusion in the next step.

Step 5. Interpretation of the Model and Drawing Conclusions The predictive model is usually utilized for decision making by human. Therefore, an interpretable model, with which the input-output dependencies are clearly shown, needs to be identified. However, the predictive accuracy and the interpretability are contradictory in general, i.e. a predictive accurate model is usually complex but an interpretable model has to be simple.

Throughout the whole procedure, proper utilization of experts' knowledge/experience is necessary for data modeling studies. Modeling methods of input-output dependencies from data have been traditionally explored in such diverse fields as multivariate regression, classification, pattern recognition, knowledge acquisition, and machine learning. Recent interest in learning from data has resulted in the development of biologically motivated methodologies, such as artificial neural networks and fuzzy systems. These methods utilize multiple basis functions for input-output mapping. Each of the basis functions is simple. Combination of these basis functions enables to model complex systems.

The basis functions of fuzzy model are called membership functions. The membership functions granulate continuous values, and has labels. The input-output mapping can be described with fuzzy if-then rules using the labels. This discrete expression is compatible with smooth interpolation between the rules. The output of each fuzzy rule is interpolated with the outputs of neighboring rules by fuzzy inference. Because of high interpretability of fuzzy model that is compatible with the accuracy, this paper is to study fuzzy modeling for the modeling from data.

1.1.2 Fuzzy Modeling

Fuzzy logic was originally motivated by the desire to sidestep the rigidity of the traditional boolean logic, in which any statement was either true or false. Fuzzy logic allows degree of truthfulness that measures to what extent a given object belongs to a fuzzy set. A continuous fuzzy set (linguistic variable) A is specified by the fuzzy membership function $\mu_A(x)$ that gives partial degree of membership of an object x in A.

$$x \in X, \ \mu_A : X \rightarrow [0, 1], \tag{1.1}$$

in contrast to

$$X \rightarrow \{0, 1\} \tag{1.2}$$

in the boolean logic.

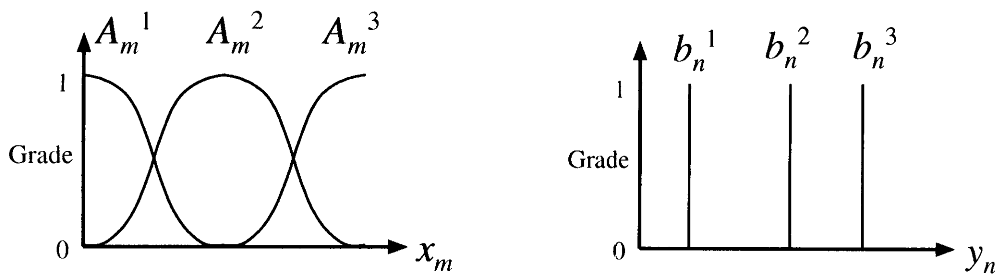
The input-output relationships are described with a set of fuzzy rules using fuzzy sets. The i-th rule, in the case of M dimensional multivariate inputs $\boldsymbol{x}= (x_1, \cdots, x_M)$ and single output y, has a linguistic expression:

$$\text{Rule } i: \text{ IF } (x_1 \text{ is } A_1^i) \text{ AND } (x_2 \text{ is } A_2^i) \cdots \text{ AND } (x_M \text{ is } A_M^i), \text{ THEN } y \text{ is } B^i. \quad (1.3)$$

where $A_m^i(m = 1 \cdots M)$ and B^i are fuzzy sets for the input x_m and the output, respectively. The antecedent part is also expressed as

$$\text{IF } \boldsymbol{x} \text{ is } \boldsymbol{A}^i, \quad (1.4)$$

where $\boldsymbol{A}^i= (A_1^i, A_2^i, \cdots, A_M^i)$. Figure 1.1(a) shows an example of membership functions formed by multiple sigmoid functions. Figure 1.1(b) shows an example with singletons. These fuzzy sets are designed with human's *a priori* knowledge or estimated with a statistical method.



(a) sigmoid functions (b) singletons

Figure 1.1: Membership Functions

The activation value of the fuzzy rule is given by the product of membership grades,

$$\mu_i(\boldsymbol{x}) = \mu_{A_1^i}(x_1) \cdot \mu_{A_2^i}(x_2) \cdots \mu_{A_M^i}(x_M). \quad (1.5)$$

Another inference method utilizes min operation instead of the product operation.

The output fuzzy set in the consequent can be replaced by a linear function. This model, known as Takagi-Sugeno-Kang fuzzy model[2], is given in the following form:

$$\text{Rule } i: \text{ IF } \boldsymbol{x} \text{ is } \boldsymbol{A}^i \text{ THEN } y \text{ is } g(\boldsymbol{x}), \quad (1.6)$$

where $g(\boldsymbol{x})$ is an affine function. Sometimes $g(\boldsymbol{x})$ is a singleton. The inference method with singletons in the consequent is called *simplified fuzzy inference*[3]. The rule is given by

$$\text{Rule } i: \text{ IF } \boldsymbol{x} \text{ is } \boldsymbol{A}^i \text{ THEN } y \text{ is } y_i, \tag{1.7}$$

and the model output $\hat{y}$ is

$$\hat{y} = \frac{\sum_{i=1}^{N_R} \mu_i(\boldsymbol{x})\, y_i}{\sum_{i=1}^{N_R} \mu_i(\boldsymbol{x})}. \tag{1.8}$$

The most distinguishing advantage of fuzzy model is its high interpretability compatible with smooth interpolating capability. The model is interpretable with the discrete expression of rules. The output of each fuzzy rule is interpolated with the outputs of neighboring rules by the above fuzzy inference.

The hierarchical fuzzy modeling method proposed by Nakayama *et al.* [4] uses multiple fuzzy models as sub-models. This hierarchical fuzzy model is applicable to systems with many input variables.

Figure 1.2 shows an example of hierarchical fuzzy model, which consists of three fuzzy models in a two layered hierarchical structure. The figure shows a case where

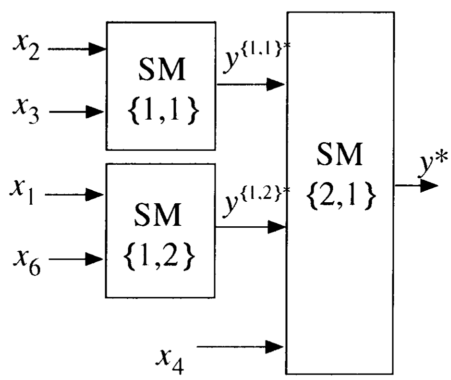


Figure 1.2: A hierarchical fuzzy model

the model has 5 dimensional inputs $\boldsymbol{x} = (x_1, x_2, x_3, x_4, x_6)$ and single output y. In figure 1.2, $y^{\{1,1\}*}$, $y^{\{1,2\}*}$ and y^* are the inferred values of the fuzzy sub-models. In the first layer, the sub-model SM{1,1} is the fuzzy model with the inputs x_2 and x_3. The other fuzzy model with x_1 and x_6 are lined in parallel. The outputs of these models are $y^{\{1,1\}*}$ and $y^{\{1,2\}*}$, respectively. These two fuzzy sub-models in the first layer greatly contribute to the input-output relationships of the system. The input variable x_4 is used by the fuzzy sub-model in the second layer. This input is used for a small adjustment of the model.

The identification of a hierarchical fuzzy model is to carry out the followings:

1. determination of a hierarchical structure of sub-models,

2. selection of input variables of each sub-model from candidate variables,

3. division of input space of each sub-model,

4. determination of fuzzy inference method,

5. tuning of membership functions.

A suitable fuzzy inference method should be adopted referring to the type of input and output variables, the input-output dependencies, and required form of knowledge that will be acquired. Application specific knowledge is helpful to select the suitable inference method.

The fuzzy modeling method was first proposed by Takagi, Sugeno and Kang [2, 5]. Input-output nonlinear relationships were made possible to be described with discrete fuzzy rules by their method. However, this modeling method tends to fall into local minima because of its incremental modeling procedure. This method increases the number of input variables or the number of membership functions one by one to find out a good input space division.

Horikawa *et al.*[6, 7] proposed fuzzy neural network for tuning of membership functions. This method utilizes the back-propagation algorithm to tune fuzzy rules and the membership functions. To obtain an accurate and interpretable fuzzy model, the number of membership functions and coarse allocation of membership functions must be determined before the back-propagation learning.

Karr et al.[8] proposed a combination of fuzzy logic and genetic algorithm (GA). The genetic algorithm was applied to tuning of membership functions of a fuzzy controller from given input-output pairs of data.

Fukuda *et al.*[9] proposed an application of genetic algorithm to identification of hierarchical structure of fuzzy model. This reduces the probability being trapped into local minima.

Matsushita *et al.*[10, 11] applied genetic algorithm to input variable selection of each sub-model and to determination of the hierarchical structure. This method is very effective in the case where the system has a strong nonlinearity and many candidates for the input variables. However, this procedure has the following difficulties:

- Membership functions are allocated evenly on the universe of discourse of each input variable. This often makes the number of membership functions excessive.
- Tree-type division of input space, with which the number of fuzzy rules becomes probably smaller, is not applicable.
- Chromosomes are evaluated under a fitness function that is a weighted sum of the predictive accuracy and a measure of conciseness of model. An appropriate set of weights should be found out by trial and error.

1.1.3 Biomechanics

Human body is a complex mechanical system. Biomechanics, whose origins are mechanics and anatomy, has tried to describe, analyze, and assess humans' motion. Major interests of biomechanics have been focused on humans' locomotion, i.e. how a human moves utilizing his/her body. The fastest and the most popular locomotions, with use of vehicles powered by human body, are cycling on land and rowing on water. The cycling motion is also the most efficient locomotion among those ever found by the creatures on earth.

For biomechanics of the whole body motion, the Newton's laws of mechanics are the *first principle* models. There are three kinds of mechanical models of the human body:

Center of Body Mass Model This model considers only the center of body mass as a particle. It is effective if the relationship between the jump height and the ground reaction force is studied or if the positions of soccer players are discussed, for example. Motions of limbs can not be described by this model.

Link Segment Model This model assumes that the human body consists several (usually, up to 15) rigid segments, linked by the joints. Many position and angle parameters are measurable. Detailed information, such as force, torque and power at each joint, is calculated by inverse dynamics. This model is effective to discuss how the body parts coordinate to complete a task that requires the whole body. This model does not consider effect of co-contraction etc., this neglects force by antagonist muscle, which yields a few percent error.

Musculoskeletal Model This is to model the human body exactly, which has 206 bones and about 700 muscles. Accurate description, analysis and assessment are possible with this model.

One of major reasons of the difficulty of describing humans' motion with mechanical models is the nature of muscle contraction. Hill [12] investigated the force and the speed of muscle contraction. The relationship is inversely proportional, i.e. muscle pulls a light load at a high speed and a heavy load at a low speed. Muscle group has the same nature. As the power produced by muscle is the product of the force and the speed, there exists an optimal speed to produce maximal muscle power.

Human body utilizes leverage to match the angular velocity around joint to the optimal speed. For example, baseball pitchers throw the ball, which weighs only 125g, keeping the elbow joint straight. The straight arm enlarges the distance from the rotating axis to the ball. On the other hand, shot put throwers bend the elbow joint and shorten the distance from the rotating axis to the shot. Another example is a sprinter. A sprinter crouches at the start of 100m race. The velocity of the body mass is zero at the starting position. The crouching start makes the hip nearer to the ground. That makes it easier to produce a large force (i.e. large acceleration) at a low speed. After the start, the sprinter does not need the large acceleration but the high speed, so the hip joint is kept as high as possible.

These examples are simple considerations of a single leverage. Human body consists of parallel and serial combinations of leverages. Exact description of human body with a mechanical model is too difficult. Another nature that muscle group has a spring factor, and a phenomenon that a joint moves automatically when another joint moves are other reasons why the description, analysis and assessment of human body are difficult only with a mechanical model.

Cyclists utilize their lower limbs and rowers utilize the whole body to propel the vehicle and the body itself. These motions are complicated motions. A textbook of sports biomechanics written by Hay[13] mentioned the complexity of the cycling motion. The following is quoted from the textbook:

> Consider the action of one of the cyclist's leg. Here there are at least three simultaneous rotations taking place. First, there is the rotation of his thigh about an axis through his hip joint (which is itself translating). Then there is the rotation of his leg about his knee joint, and finally there is the rotation of his foot about his ankle joint. As can probably be imagined, a study of the combined motions of the elements within such a system can become quite complex.

The above paragraph considers a link-segment model of a lower limb, which consists of three segments. In the case of rowers, they utilize the trunk and the arms as well as the lower limbs. The motion becomes more complex. The data from these kinds of human motions have multi-dimensionality and mutual interaction.

Some researchers [14, 15, 16, 17, 18] have measured kinetic and kinematical variables for analyses of cycling and rowing. However, these studies did not go any further than depicting the measured data. Interpretation of the data has relied mainly on the intuition (or sometimes supertuition) of the researchers and the coaches.

The use of statistical methods for estimating input-output dependencies is effective to comprehend such complex systems. Recently statistical estimating methods like cross validation [19] and discriminant analysis [20] are utilized for biomechanical features of rowers. Artificial neural networks have started to be applied to biomechanics in these years, too. Neural networks are used to describe input-output mapping between training volume, skills and the sports performance [21, 22].

But, any study that applies a fuzzy model to biomechanics has not been found. Since the final goal of the motion analyses in this field is to assist athletes for improving their performances, the fuzzy modeling that extracts interpretable information from the data is promising.

1.2 Purpose of the Study

The purpose of this study is to develop an efficient and effective fuzzy modeling method and break new ground for describing humans' locomotion with fuzzy rules. For this purpose, this dissertation proposes a new identification method for sub-models of hierarchical fuzzy models using the genetic algorithm (GA). The GA determines a combination of input variables. For automatic input space division, this paper proposes two methods. One is based on grid-type division and the other on tree-type division. The proposed grid-type division method allocates membership functions unevenly on the universe of discourse by inserting/deleting membership functions. This method probably decreases the number of membership function without loss of the predictive accuracy. The proposed tree-type division method repeats dividing a subspace into a tree-type structure, and construct a model with a smaller number of fuzzy rules than the grid-type does.

This dissertation proposes a new method for fuzzy modeling utilizing multiple objective genetic algorithm that is able to obtain compromising solutions for conflicting multiple objectives.

The hierarchical fuzzy modeling is effective for identifying systems with strong nonlinearity and many input variables. Biomechanical data of athletes contain nonlinearity with many variables. This thesis applies the proposed method to description of cycling and rowing motions, and reveals nonlinear relationships between variables and the performance.

1.3 Composition of this Dissertation

This dissertation presents solutions to main issues of fuzzy modeling. The issues of fuzzy modeling are input variable selection and input space division of the antecedent part.

The methods of input space division are proposed in chapter 2. The proposed grid-type division method inserts a new membership function and deletes needless membership functions. Numerical experiments show that the proposed method obtains more concise models than the insert-only method and the delete-only method do. The proposed tree-type division method divides the input space incrementally under the criterion of model error. A numerical experiment shows that the conciseness of fuzzy models with the grid-type division obtains a smaller number of fuzzy rules than the tree-type division does for the same predictive accuracy level.

Evaluation criteria of fuzzy models determine the quality of the obtained model. Two encoding methods are proposed for the structure identification of the sub-models utilizing multiple-objective genetic algorithm (MOGA). The MOGA generates wide variety of model candidates with Pareto optimality.

Chapter 3 presents application of hierarchical fuzzy modeling to biomechanical data of cyclists. Human musculoskeletal system is complicated. It has been hard to comprehend such complex motions as cycling with a mechanical model. The relationships between force patterns of pedaling and the output power are analyzed with the identified fuzzy model.

Chapter 4 presents applications of hierarchical fuzzy modeling to biomechanics of rowers. Motion of rowers is more complex than that of cyclists because they utilize the trunk and the arms as well as the legs. In addition, they are not on solid road, but on fluid water. Force patterns of rowing are analyzed in section 4.2. Power and its timing of partial motions of rowers are analyzed in section 4.3. Hierarchical fuzzy modeling method is utilized to clarify the input-output dependencies and to extract advice for the athletes.

Chapter 2

Fuzzy Modeling Using Genetic Algorithm

2.1 Introduction

Fuzzy model[2] is one of the statistical models to estimate input-output dependencies. The most distinctive advantage of fuzzy modeling is its interpretability due to the discrete expression of each of basis functions. The fuzzy inference interpolates smoothly the output of each basis function with those of neighboring basis functions. The discrete expression and the smooth interpolating capability are compatible in the fuzzy model.

The hierarchical fuzzy modeling method using genetic algorithm proposed by Matsushita *et al.*[11] is a good tool to construct interpretable fuzzy models of nonlinear multi-input systems. It is easy from the acquired fuzzy model to extract fuzzy if-then rules that describe the input-output relationships of the target systems. However, as summarized in 1.1.2, this method has the following difficulties: (1) Even allocation of membership functions often results in excessive number of membership functions; (2) The division method of input space is limited only to grid-type division; (3) Weighted sum of evaluated values need trial and error.

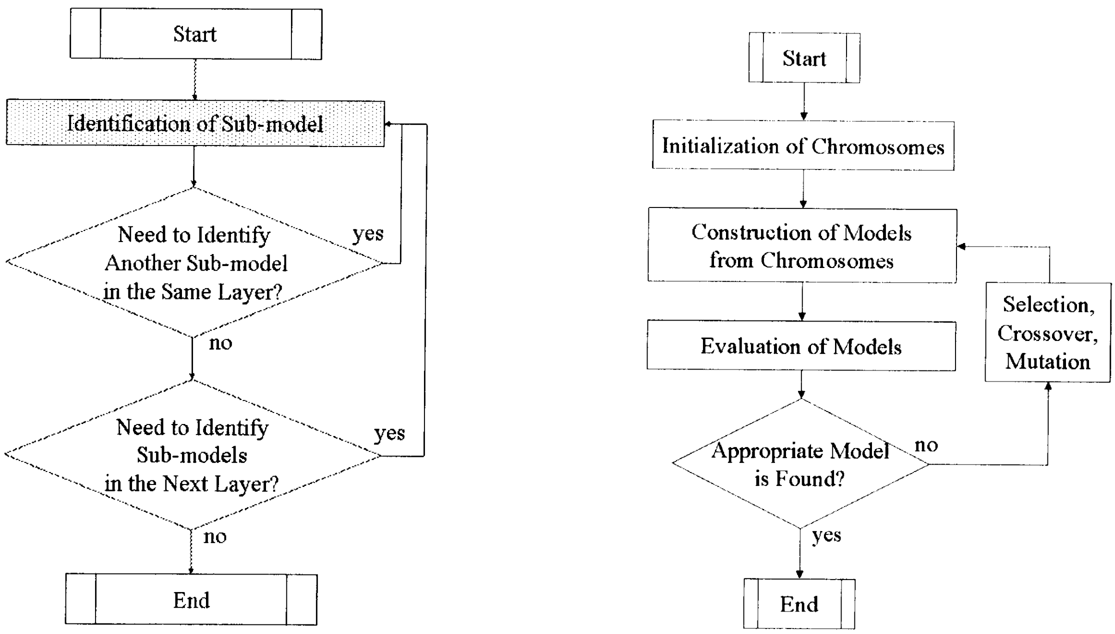
This chapter presents a new hierarchical fuzzy modeling method that solves the above difficulties. For flexible allocation of membership functions and high degree of freedom for input space division, only the combination of input variables is encoded to chromosomes. Two methods are presented in this chapter for division of the input space. One is to divide the input space into grid and the other is to obtain a tree-type division. The proposed grid-type division method can identify fuzzy models with uneven allocation of membership functions. Numerical experiments show (1) the proposed grid-type division method obtains more concise models than the conventional methods do and (2) the tree-type division method obtains more concise models than the grid-type division method does.

To avoid the trial and error for setting the weights on the evaluation functions, multiple-objective genetic algorithm (MOGA) is introduced to the modeling process. Two encoding methods for the MOGA are presented. Numerical experiments show that the application of MOGA with the proposed encoding methods can generate various candidate models on the Pareto front.

2.2 Conventional hierarchical fuzzy modeling method

The main flow of hierarchical fuzzy modeling method in [11] is shown in Figure 2.1(a). Sub-models in a hierarchical fuzzy model are identified one by one using genetic algorithm as shown in Figure 2.1(b). First of all, the first sub-model in the first layer (SM{1-1} in Figure 1.2) is identified. Chromosomes representing combinations of input variables and number of membership functions for each input variables are generated. An example of chromosome is shown in Figure 2.2. Through genetic operations, selection, crossover, mutation, a good sub-model is identified. Next, the second model (SM{1-2}) that compensates the error of the first model is identified in the same layer in the case where two or more candidate variables are remaining, not used as an input of the first sub-model. Some of the remaining variables would be selected. The third, the fourth and more models are identified until the number of remaining variables becomes two or less.

Fuzzy modeling in the next layer is done using the outputs of sub-models in the previous layer. The output of the first model in the previous layer is always used. A combination of this output and some of the candidates is chosen as the inputs of the first model of this new layer. The candidates are the outputs of the other models in the previous layer and the remaining variables that have not been used in the fuzzy modeling of the previous layer or lower layers. The whole hierarchical fuzzy model is evaluated under the predictive accuracy of the model. If the evaluation value is improving, more layers are identified.



(a) Main Flow (b) Identification of Sub-model using Genetic Algorit

Figure 2.1: Flow of Hierarchical Fuzzy Modeling

The hierarchical modeling process in Figure 2.1(a) is summarized as follows:

1. The input-output data is divided into two groups A and B, whose statistical characteristics are nearly the same. The model identified from the data of group A is called model A. The number of layer h is set at 1.

2. A sub-model is identified using gentic algorithm. The flow of this step is shown in Figure 2.1(b). The identified model is denoted by model h-1 and its output is denoted by y^{h-1*}. The model h-k means the k-th model in the h-th layer.

3. If the number of remaining input variables not used in the model h-1 is more than two, another model is identified using genetic algorithm. Some of the remaining variables would be selected as the inputs of model h-2. If more variables remain, model h-3 will then be identified. This modeling is repeated until the number of remaining input variables becomes less than two. The model h-1 will be used for the identification of the models in the succeeding layer. The outputs of the model h-2, -3, $\cdots$ will be the candidates for the input variables of the models in the next layer.

4. Fuzzy modeling in the next layer is done using the outputs of sub-models identified in the previous layer. The output of model h-1 is always used. A combination of this output y^{h-1*} and some of the outputs of models h-2, -3, $\cdots$ as well as the input variables not used in model h-1 is chosen by genetic algorithm. The acquired model is denoted by $(h+1)$-1.

5. The models h-1 and $(h+1)$-1 are compared under the following criterion.

$$\begin{aligned} C &= \sqrt{\sum_{i=1}^{n_A}\left(y_i^A - y_i^{AA}\right)^2 + \sum_{i=1}^{n_B}\left(y_i^B - y_i^{BB}\right)^2} \\ &+ \sqrt{\sum_{i=1}^{n_A}\left(y_i^{AB} - y_i^{AA}\right)^2 + \sum_{i=1}^{n_B}\left(y_i^{BA} - y_i^{BB}\right)^2} \end{aligned} \tag{2.1}$$

where n_A and n_B are the numbers of data groups A, B, respectively; y_i^A and y_i^B are the outputs of data A and data B, respectively; y_i^{AA} and y_i^{BB} are the inferred value of model A with data A and that of model B with data B, respectively; y_i^{AB} and y_i^{BA} are the inferred value of model B with group A and that of model A with group B, respectively. The first term on the right hand side in eq.(2.1) is the precision of the model, and the second term is the criterion for evaluation of the predictive accuracy of the model.

If the $(h+1)$-1 is better than the h-1, the procedure is repeated from 3. If not, the procedure is stopped.

Details of the identification procedure for sub-model in Figure 2.1(b) is given as follows: The number of membership functions for each input variable is coded onto each gene of chromosomes. Figure 2.2 shows an example of chromosome. If the gene

x_1	x_2	x_3	x_4	$\cdots$	x_M
5	0	2	0	$\cdots$	0

Figure 2.2: An example of chromosome

has 2 or a larger number, the corresponding variable is selected as an input. If the gene has 0 or 1, the corresponding variable is not selected. In this example, x_1 and x_3 are selected. Five membership functions are allocated to x_1 and 2 membership functions to x_3.

Fuzzy models are identified based on the chromosomes. Membership functions are allocated evenly on the universe of discourse on each of the selected variables. In the case of the above example, combinations of five membership functions on x_1 and two membership function on x_3 provide ten fuzzy rules. The fuzzy models are evaluated under the accuracy and the conciseness of the model. Selection, crossover and mutation are applied to the chromosomes according to the fitness values. The fuzzy models in the new generation are evaluated. These steps are repeated until the fitness value converges.

The above identification process is summarized as follows:

1. Chromosomes are initialized to have 0 in each gene. The number of chromosomes is n_g. The binary number in each gene is flipped to 1 with the probability of p_i.

2. Chromosomes are evaluated. The chromosome determines a combination of input variables and a number of membership functions for each input variable to be used for the sub-model. In this process, the identification in the consequent part

of the model is done with the data of group A. The performance index F used in this fuzzy modeling process is given by:

$$F = \frac{\sum_{i=1}^{n_B} \left(y_i^{BA} - y_i^{B}\right)^2}{n_B} + kS \tag{2.2}$$

where n_B is the number of the data of group B, y_i^B is the output data of group B, y_i^{BA} is the inferred value of model A with data B, and S is the number of subspaces. The first term evaluates the generality of the identified model, and the second term is a criterion for the conciseness of the model. Coefficient k adjusts the weights on the generality and the conciseness.

3. Individuals are ranked with this performance index. The worst n_w chromosomes are replaced with copies of better chromosomes.

4. Crossover and mutation operations are applied to the population. Crossover is applied to the whole population except for one elite at the rate of p_c. Parents are randomly selected, and one point crossover with randomly selected crossover point is applied. Mutations are applied to each gene of all the chromosomes except for that of the elite chromosome at the rate of p_m.

5. Stop if the performance of the elite chromosome does not improve during m_{end} generations. Otherwise, go to step 2.

This procedure including the flows in Figure 2.1(a) and (b) gives a good solution for determination of hierarchical structure and selection of input variables.

2.3 Proposed method for sub-model identification

This dissertation proposes a new hierarchical fuzzy modeling method using genetic algorithm. In the proposed method, the combination of input variables is determined by genetic algorithm. The chromosomes have binary numbers. An example is shown in Figure 2.3. If the gene has 1, the corresponding variable is used as an input. If the gene has 0, the corresponding variable is not used. This simple coding decreases the number of possible chromosomes and narrows the search space.

With this type of chromosomes, the way to construct models has a freedom of choices. There could be three major types of input space division for fuzzy modeling. They are

x_1	x_2	x_3	x_4	$\cdots$	x_M
1	0	1	0	$\cdots$	1

Figure 2.3: An example of chromosome for the proposed method

grid, tree and cluster types, as illustrated in Figure 2.4. The grid division is the most interpretable for a system with a few input variables. But, in the case where the data set has many input variables, it identifies a model with too many rules. This causes poor generalization (predictive accuracy) and poor interpretability. The tree-type division divides the input space like a decision tree. The number of fuzzy rules becomes smaller than the grid-type division. And the obtained fuzzy rules are interpretable. But, it has a risk of obtaining a sub-optimal division, since it creates the tree structure by incrementally dividing the input space. The cluster-type division probably makes the simplest model with the smallest number of rules. But, the divided subspaces do not cover the whole input space. The identified model sometimes loses predictive accuracy. And, sometimes the normalized membership grades are not convex nor ordinal on an input variable. It is not an interpretable model for some applications.

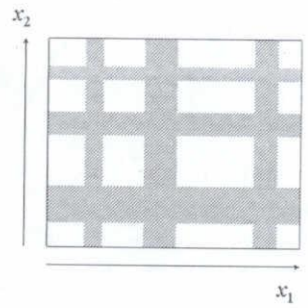
Thus, this paper presents two dividing methods, i.e. grid-type division and tree-type division. This input space division step is carried out in the construction process from chromosomes as shown in Figure 2.5.

2.4 Grid-Type Division of Input Space

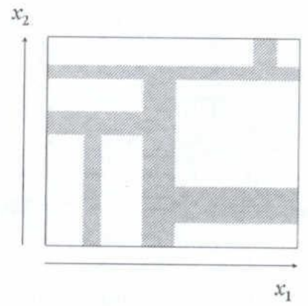
A new grid-type division method is proposed in this section. The proposed method allocates membership functions by inserting a new membership functions at the point with maximal error, and deleting a needless membership function that can be interpolated with the neighboring membership functions. Araki *et al.*[23] proposed an allocation method of membership functions by inserting a new one. The methods in [24] and [25] allocate membership functions by deleting needless ones. However, there has been no method that inserts and deletes membership functions simultaneously for appropriate allocation of membership functions.

2.4.1 Procedure

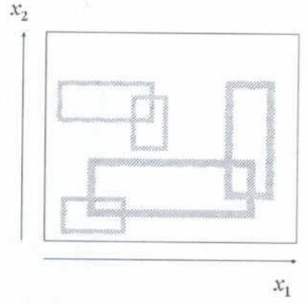
The allocation of membership functions on the universe of discourse for each input variable depends on the nonlinearity of the target system. This allocation is carried out by inserting and deleting membership functions under criteria. In this procedure, two data sets are used. One (data set A) is to estimate the model parameters and the



(a) grid-type division



(b) tree-type division



(c) cluster-type division

Figure 2.4: Three Types of Input Space Division

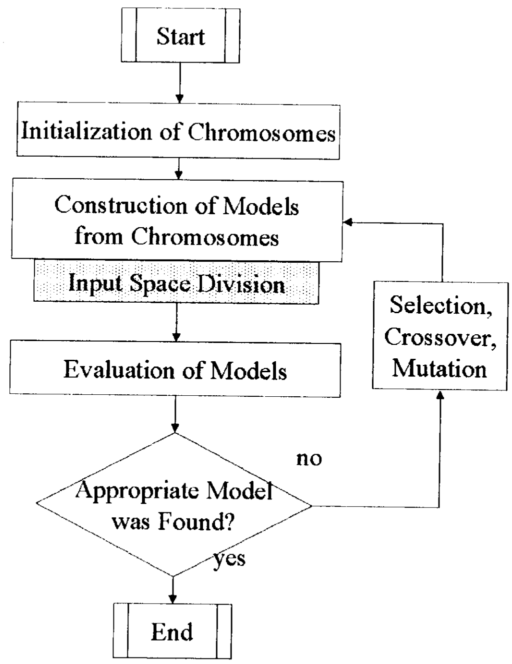


Figure 2.5: Input space division in sub-model identification flow

other set (data set B) is to evaluate the predictive accuracy of the model.

The procedure is as follows.

1. Two membership functions are initially allocated on the universe of discourse for each input variable. Figure 2.6(a) shows an example in this case.

2. The singleton in the consequence of the i-th rule is calculated by:

$$\hat{y}_i = \frac{\sum_s^{N_s} \mu_i(\boldsymbol{x}_s)\, y_s}{\sum_s^{N_s} \mu_i(\boldsymbol{x}_s)}, \tag{2.3}$$

   where N_s is the number of samples. $\boldsymbol{x}_s$ and y_s are the input and output of the s-th sample.

3. The performance of the model is evaluated with a criterion that includes the predictive accuracy and the conciseness of the identified model:

$$F = \frac{\sum_{i=1}^{n_B} \left(y_i^{BA} - y_i^{B}\right)^2}{n_B} + kS \tag{2.4}$$

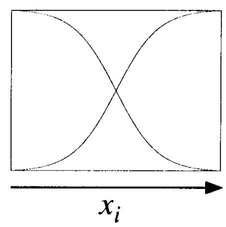
   where n_B is the number of the data of group B, y_i^B is the output data of group B, y_i^{BA} is the inferred value of model A with data B, M is the number of input variables, and S is the number of subspaces. The coefficient k is the weights that balances the generality and the conciseness of the model. The first term evaluates the generality of the identified model, and the second term is a criterion for the conciseness of the model.

4. If the performance index does not improve for n_{end} iterations, stop. If it does, go to step (5).

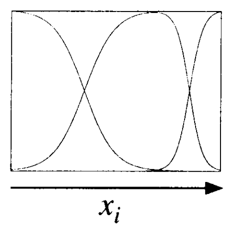
5. A new membership function is inserted in the sub-space where the error is maximum. Among the multiple input variables, the input variable with the highest grade in the sub-space is chosen. Figure 2.6 (b) and (c) show an example how membership functions are inserted into sub-spaces.

6. The singletons in the consequence of the new model are obtained with the same calculation in step 2.

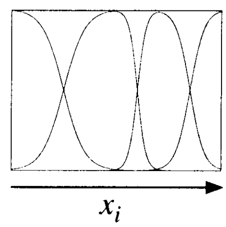
7. Consequent singletons of each membership function for an input variable are examined. If they can be linearly interpolated with neighboring rules, this membership function is deleted.



(a) initial state



(b) one membership function is inserted



(c) two membership functions are inserted

Figure 2.6: Insertion of membership functions

8. Go to step 2.

After this process, a fuzzy model with appropriately allocated membership functions is obtained. The membership functions are unevenly distributed on the universe of discourse, depending on the nonlinearity of the modeling object.

2.4.2 Numerical Experiment

Two numerical experiments were done to demonstrate the effectiveness of the proposed method. The first experiment compared the accuracy and the conciseness of the proposed method to those of a method that only inserts membership functions and to a method that only deletes membership functions. The second experiment compared them to those with even allocation of membership functions.

Experiment 1

An example that had two inputs (x_1, x_2) and an output y was tested.

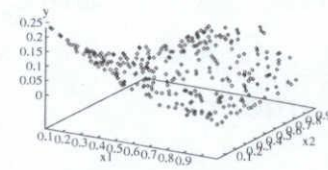
$$y = \max \begin{cases} (x_1(1-x_1)(x_1+x_2))^2 \\ \left(\frac{1}{2}(1-x_1)(1-x_2)\right)^2 \end{cases} \tag{2.5}$$

300 pairs of x_1, x_2 were randomly generated in the range of $0 \leq x_1 \leq 1,\ 0 \leq x_2 \leq 1$.

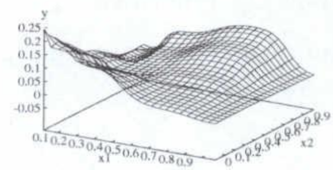
Figure 2.7(a) shows the training data, and (b) shows the input-output mapping obtained by the proposed method. Figure 2.7(c) and (d) show the mappings of the fuzzy models obtained by the insert-only method and by the delete-only method, respectively.

Figure 2.8 shows the obtained membership functions. (a) shows membership functions obtained by the proposed method. (b) shows those by the insert-only method, and (c) shows those by the delete-only method. The proposed method obtained 6 membership functions for x_1 and 7 membership functions for x_2. The insert-only method obtained 6 and 9 each, and the delete-only method obtained 7 and 9, respectively. The numbers of membership functions and the corresponding numbers of fuzzy rules are summarized in Table 2.1.

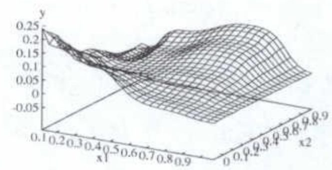
Another data set of randomly generated 300 pairs verified the predictive accuracy of the three methods. Table 2.2 shows the errors for training data and for unknown data. The model obtained by the proposed method had better predictive accuracy with a



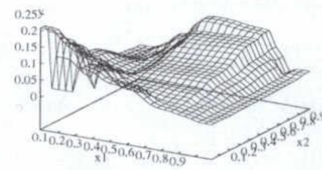
(a) training data



(b) obtained mapping by the proposed method

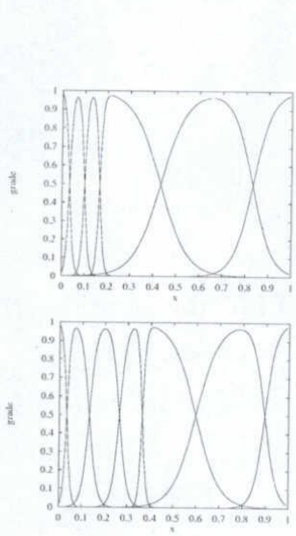


(c) obtained mapping by insert only method

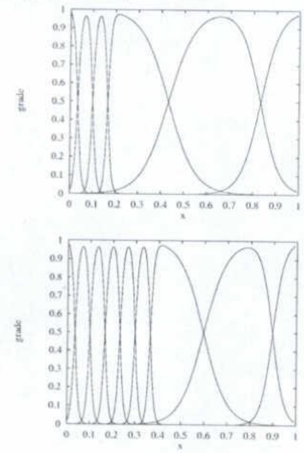


(c) obtained mapping by delete only method

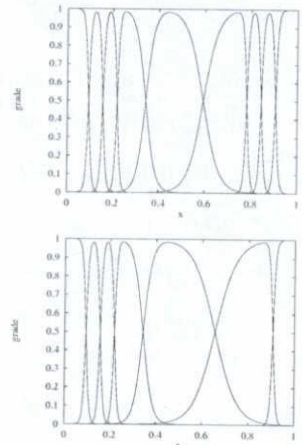
Figure 2.7: Training data and obtained mappings



(a) proposed method



(b) insert-only method



(c) delete-only method

Figure 2.8: Obtained Membership Functions

Table 2.1: Numbers of Membership Functions and Fuzzy Rules

Method	Number of Membership Functions		Number of Fuzzy Rules
	x_1	x_2	
Insert + Delete	6	7	42
Insert Only	6	9	54
Delete Only	7	9	63

Table 2.2: Comparison of Accuracy

	Error for Training Data	Error for Unknown Data
Insert + Delete	0.0250	0.0279
Insert Only	0.0249	0.0287
Delete Only	0.0238	0.0355

smaller number of fuzzy rules. That means more suitable allocations of membership functions was acquired by the proposed method.

Experiment 2

Another numerical experiment was carried out to validate the effectiveness of the proposed identification method of sub-models. The following input-output relationships were tested:

$$\begin{aligned} y &= (-4 + x_0^{0.5} + x_1^{-1})^2 + 5\sin(x_2 + x_3) \\ &+ \exp(1 + x_4 + x_5) \\ &= (-4 + x_{10})^2 + 5\sin(x_{11}) + \exp(1 + x_{12}) \end{aligned} \quad (2.6)$$

$$\begin{aligned} &x_6 = x_0^{0.5}, \quad x_7 = x_1^{-1}, \quad x_8 = x_1^{-2}, \\ &x_9 = x_0^{0.5} x_1^{-1}, \quad x_{10} = x_0^{0.5} + x_1^{-1}, \quad x_{11} = x_2 + x_3, \\ &x_{12} = x_4 + x_5 \end{aligned} \quad (2.7)$$

where x_0 to x_5 were input variables. x_6 to x_{12} were arranged input variables expressed by eq. (2.7). x_{13} was used as a dummy variable, that had no relationship with y. The number of candidates of inputs l_g was 14. The ranges of input variables are shown in Table 2.3. The ranges were decided so that the input variables influenced the output almost equally.

Table 2.3: Range of variable

variable	range
x_0	$\{0, 1, \cdots, 20\}$
x_1	$\{0.5, 0.7, \cdots, 1.5\}$
x_2, x_3	$\{-1.0, -0.9, \cdots, 5.0\}$
x_4	$\{0.0, 0.1, \cdots, 0.5\}$
x_5	$\{-1.5, -1.4, \cdots, 0.5\}$
x_{13}	$\{0, 1, \cdots, 17\}$

Twenty sets of eighty pairs of input-output data were generated. Every two sets of them were used for A and B groups. The input-output data were normalized within the range $[0, 1)$ for the fuzzy modeling. The parameters of GA were set as follows: the number of chromosomes $n_g = 20$, the probability of crossover $p_c = 0.75$, the probability of mutation $p_m = 0.1$, the number of generations for the stopping condition $m_{end} = 10$ and the number of iterations for the stopping condition of learning $n_{end} = 15$.

Table 2.4 shows the results of 10 experiments with 10 different pairs of sets of data. Criterion C was calculated by (2.1) with the output of sub-model h-1. h is the number of layers. s is the number of sub-models. v is the total number of input variables of all sub-models. m means the total number of membership functions.

Table 2.4: Models obtained with the proposed method

data set #	C in (2.1)	h	s	v	m
0, 1	0.775	2	4	7	35
2, 3	0.748	2	3	7	35
4, 5	1.001	3	5	10	44
6, 7	0.706	2	2	6	21
8, 9	0.885	2	2	4	19
10, 11	0.788	3	5	12	51
12, 13	0.879	2	1	3	19
14, 15	1.021	2	3	8	32
16, 17	1.113	2	2	6	30
18, 19	0.996	3	6	16	54
average	0.891	2.3	3.3	7.9	34.0
SD	0.137	0.5	1.6	3.9	12.6

Experiments with the conventional hierarchical fuzzy modeling method[26] which allocated membership functions evenly were also done. Table 2.5 shows the results.

Comparing these results, the values of C were a little bit larger (worse) with the proposed method. But, the number of layers h, the number of sub-models s, the number of variables v and the number of membership functions m were smaller. Figure 2.9 shows comparison of C and m. Boxes represent the average values and the intervals represent standard deviations. Statistical tests tell that the difference of h and the difference of m are significant ($p < 0.05$) and those of the others are not significant.

Table 2.5: Models obtained with the conventional method which evenly allocate membership functions

data set #	C in (2.1)	h	s	v	m
0, 1	0.930	3	4	11	50
2, 3	0.619	4	6	16	80
4, 5	0.704	4	5	10	53
6, 7	0.480	5	5	16	74
8, 9	0.985	2	3	9	37
10, 11	0.681	5	10	26	100
12, 13	0.814	2	3	7	40
14, 15	1.097	4	4	9	41
16, 17	0.861	5	7	16	62
18, 19	0.947	2	3	6	31
average	0.811	3.6	5.0	12.6	56.8
SD	0.189	1.6	2.2	6.0	22.0

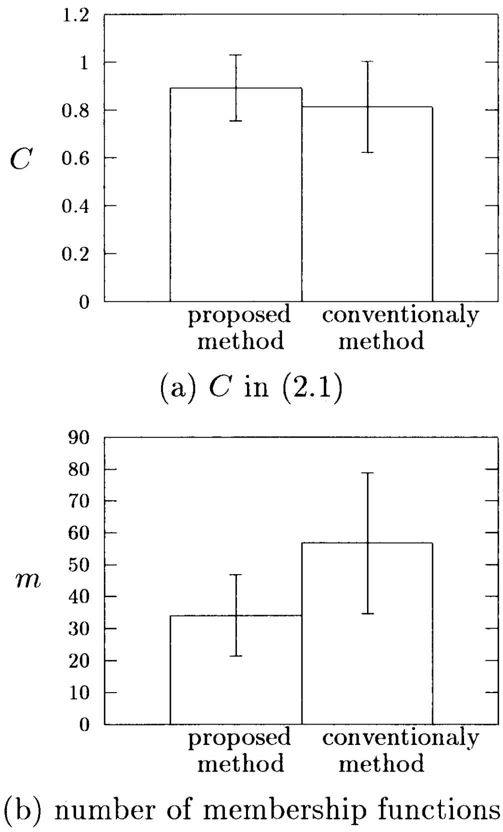


Figure 2.9: Comparisons of performances

Figure 2.10 shows the identified model with the data set 0 and 1. Numbers in the parenthesis correspond to the number of membership functions for the input variables. The number of input variable was 7 and the total number of membership functions was 35. The allocations of membership functions are shown in Figure 2.11.

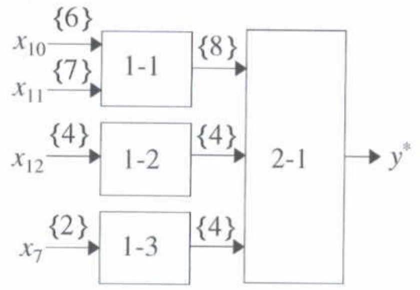


Figure 2.10: Identified model for the data set 0 and 1

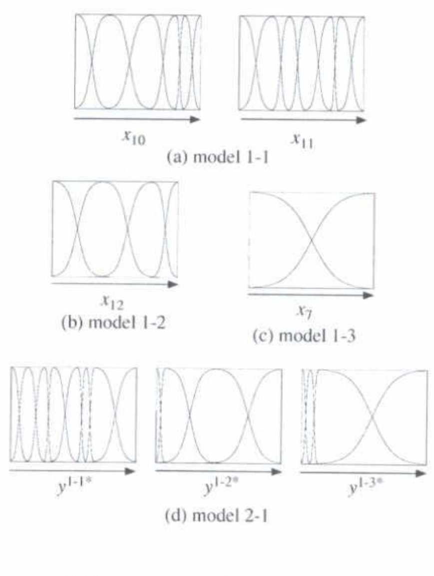


Figure 2.11: Allocation of membership functions

Figure 2.12 shows the identified model with evenly allocated membership functions for the data set 0 and 1. The number of input variables was 11 and the total number of membership functions was 50. Both of them were larger than those of the model obtained by the proposed method.

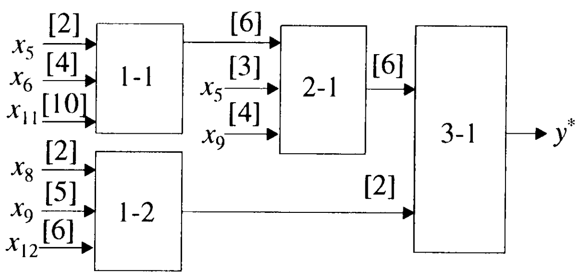


Figure 2.12: Identified model with evenly allocated membership functions

2.5 Tree-Type Division of Input Space

The other division method of input space is based on the tree-type division. The input space is divided incrementally to obtain a tree structure. The input-output relationships of data are used to determine the dividing points. The subspace with the largest variance of data output is divided so that the model error becomes the smallest. The tree-type division obtains a model with a smaller number of fuzzy rules than the grid-type division does without loss of predictive accuracy.

2.5.1 Procedure

The division of input space determines the predictive accuracy and the interpretability of the fuzzy model. This section presents a tree division using the statistical t test. This test is to examine whether the average value of a group of data is significantly different from that of the other group. This method uses the p value of the t test. We apply the t test to the stopping condition of the following division algorithm.

1. The input space is not divided initially, i.e. the number of subspaces is 1. The number of fuzzy rules N_R is also 1, initially. The subspace, which covers the whole input space, is denoted by S_1.

2. The data are divided into N_R data sets. If the input of a data is in the r-th subspace, this data belongs to the r-th data set. One of subspaces is going to be divided into two new subspaces. The selected subspace for the division S_{DIV} is the one with the maximum variation of the output. This subspace is decided as:

$$S_{DIV} = \arg \max_{r=1,\cdots,N_R} (n_r - 1)V(S_r) \tag{2.8}$$

   where $V(S_r)$ and n_r are the variance of outputs and the number of data in the r-th data set, respectively.

   The subspace S_{DIV} is divided into new two subspaces, S_{new1} and S_{new2}. The dividing point x_d is decided as:

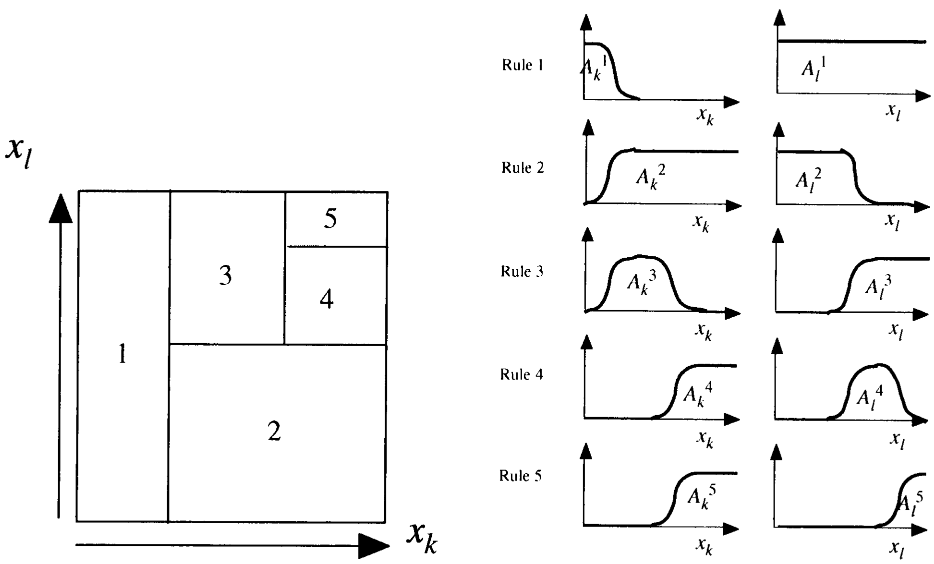
$$\min_{i=1,\cdots,I} \min_{xd_i \in S_{DIV}} \{(n_{new1} - 1)V(S_{new1}(i, xd_i)) + (n_{new2} - 1)V(S_{new2}(i, xd_i))\} \tag{2.9}$$

where $S_{new1}(i, xd_i)$ and $S_{new2}(i, xd_i)$ denote the data sets generated by dividing S_{DIV} at xd_i, which is the dividing point on the i-th axis. The number of rules N_R is increased by 1, i.e. $N_R := N_R + 1$.

3. The division of the data set is evaluated by t test. If the p value is less than the pre-set significance level, step 2 is repeated to divide another subspace. We consider that if the p value of a division is less than a pre-specified level, the average output values in the two divided subspaces are significantly different from each other.

4. Fuzzy sets that correspond to the subspaces are generated. The consequent singletons of all N_R rules are tuned by the back propagation learning.

The division process is repeated until the p value becomes greater than the significance level or the number of rules becomes greater than a pre-decided number. Figure 2.13 shows an example of tree-type division of the input space on 2 inputs (x_k and x_l) and the corresponding membership functions.



(a) Uneven division (b) Membership functions

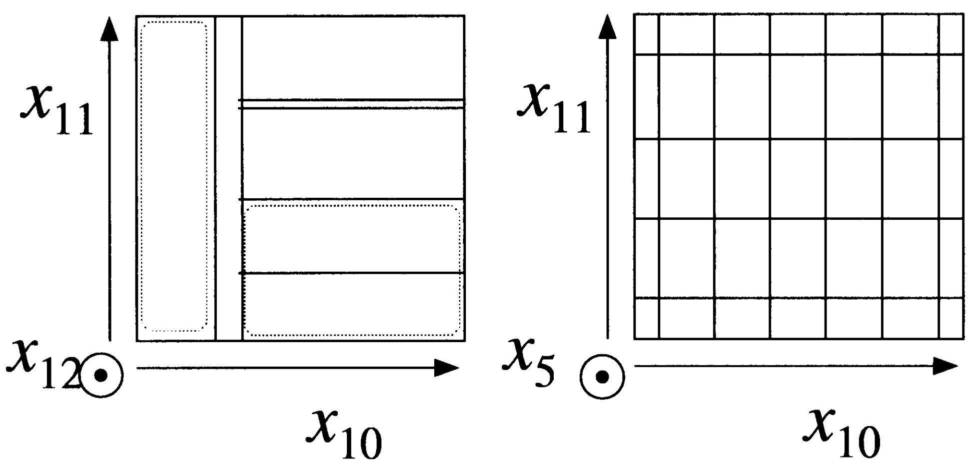
Figure 2.13: Obtained division and membership functions

2.5.2 Numerical Experiment

A numerical experiment was done to compare the performances of the grid-type and the tree-type divisions. The target nonlinear system was the same as that used by the Experiment 2 in sub-section 2.4.2.

Sixteen sets of three hundred pairs of input-output data were generated. The input-output data were normalized within the range [0, 1] for the fuzzy modeling. Two of the sixteen sets were used for A and B groups in a modeling experiment. Eight experiments were done.

Figure 2.14(a) shows an example of the input division of sub-model 1-1 obtained by the tree-type division method in one of the experiments. x_{10}, x_{11} and x_{12} were selected for the inputs. The input space was divided into 9 subspaces, and the number of fuzzy rules (NOR) was 9. The root mean square error (RMSE) for the test data was 0.0944. Figure 2.14(b) shows an example of the input division of sub-model 1-1 identified by the conventional method in [26]. The numbers of membership functions were 2 for x_5, 7 for x_{10} and 5 for x_{11}. The NOR was 70. The RMSE was 0.1066.



(a) Tree-type division (b) Grid-type division

Figure 2.14: Division of input space

The numbers of rules (NORs) and the root mean square errors (RMSEs) are summarized in Table 2.6. The table shows (Average) ± (Standard deviation) of the eight experiments in the case where the coefficient k in eq.(2.4) was set at 10^{-3}. The tree-type

Table 2.6: Comparison of the models 1-1

	tree-type division	grid-type division
NOR	6.1±1.4	54.6±16.9
RMSE	0.098±0.012	0.093±0.016

division method could identify much more concise models than the grid-type division method did without degrading their generality.

2.6 Multiple Objective Genetic Algorithm

This section discusses the step: Evaluation of Models shown with shade in the Figure 2.15. Numerical experiments investigate the effect of evaluation function to the

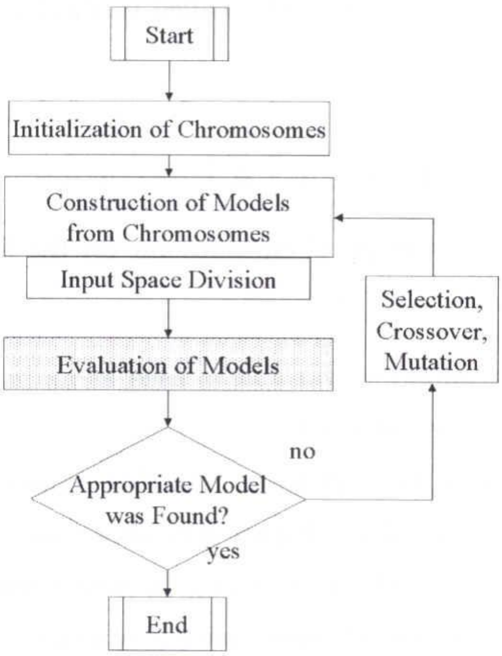


Figure 2.15: Evaluation of models in sub-model identification flow

conciseness and the accuracy of the obtained fuzzy models.

Multiple-objective genetic algorithm[27, 28] optimizes the conflicting multiple objectives. This algorithm generates various compromising solutions. In this dissertation, the multiple-objective genetic algorithm is applied to the fuzzy modeling to obtain various models having conciseness and predictive accuracy different from each other. The application of multiple-objective genetic algorithm makes the modeler free from the trial and error that has been needed to set the weights on the evaluation function appropriately. This section proposes two methods of coding for the new fuzzy modeling. The first method of coding encodes only the combination of input variables. The second method encodes both the combination of input variables and the allowable number of fuzzy rules. A numerical experiment compares the models obtained by the two coding methods.

2.6.1 Procedure of Multiple Objective Genetic Algorithm

The first method encodes the combination of input variables. Figure 2.16 shows an example of the chromosome. This method is called method 1 in this section.

x_1	x_2	x_3	x_4	$\cdots$	x_M
1	0	1	0	$\cdots$	1

Figure 2.16: An example of chromosome for the first method

If a gene has 1, the corresponding variable is used as an input variable of the sub-model. If it is 0, the variable is not used.

The other method encodes the maximum allowable number of fuzzy rules as well as the combination of input variables. Figure 2.17 shows an example of chromosome. The number of fuzzy rules is encoded in n_1 to n_K by binary numbers. In this case, the stopping condition for the division of input space is not only the pre-set p value. The input space determined by the chromosome is divided up to the number given by the chromosome.

x_1	x_2	$\cdots$	x_M	n_1	$\cdots$	n_K
1	0	$\cdots$	0	1	$\cdots$	1

Figure 2.17: An example of chromosome for the second method

$\boldsymbol{g}(p) = (g_1(p), g_2(p), \cdots, g_L(p))$ denotes the p-th chromosome in the population. Each gene has a binary number, 0 or 1, as shown in the example. The length of each chromosome L is the same as M in the first method and $M+K$ in the second method. The population, that is the number of chromosomes, is denoted by P.

The evolution of individuals is carried out by the following procedure:

1. Chromosomes are initialized to have 0 in each gene. The binary number in each gene is flipped to 1 with the probability of pr_{ini}.

$$\forall p = 1, \cdots, P, \ \forall n = 1, \cdots, L, \ g_n(p) = \begin{cases} 1, & \text{with the probability of } pr_{ini} \\ 0, & \text{otherwise} \end{cases}$$

2. Chromosomes are evaluated. Chromosome $\boldsymbol{g}(p)$ decides a combination of input variables and a number of fuzzy rules. A model θ^p is identified using only the selected input variables and the output variable. The division process of the input space described in sub-section 2.5 is carried out.

   The fitness vector $\boldsymbol{f}(p) = (f_1(p), f_2(p), f_3(p))$ is used for the evaluation of the chromosome. The elements in the fitness vector evaluate estimation ability of the sub-model, the number of rules, and the number of input variables, respectively. $f_1(p)$ is Akaike Information Criterion (AIC)[29] given by:

$$f_1 = N_D \log\left(\frac{E}{N_D}\right) + 2DOF \tag{2.10}$$

   where N_D is the number of data, E is the total error and DOF is the degree of freedom of the model. DOF is given by:

$$DOF = 2N_R - 1 \tag{2.11}$$

   for this uneven division of input space, where N_R is the number of rules. The AIC is a good criterion for the estimation capability of sub-models. The other criteria $f_2(p)$ and $f_3(p)$ are N_R:the number of rules and N_V:the number of selected input variables of the model θ^p, respectively.

$$f_2 = N_R \tag{2.12}$$

$$f_3 = N_V \tag{2.13}$$

   The smaller these values are, the better the sub-model is. These values, especially f_1 and f_2, f_1 and f_3, are the search for candidates that are in trade-off with each other.

The rank of chromosome $\boldsymbol{g}(p)$ in the population is given by

$$\text{rank}(p) := 1 + n(p) \tag{2.14}$$

where $n(p)$ is the number of chromosomes that are 'superior' to the chromosome $\boldsymbol{g}^p$. The 'superiority' is defined that $\boldsymbol{g}(q)$ is *superior* to $\boldsymbol{g}(p)$ in the case where

$$(\forall i = 1,2,3)\ f_i(q) \leq f_i(p) \quad \wedge \quad (\exists i = 1,2,3)\ f_i(q) < f_i(p). \tag{2.15}$$

3. If rank(p) is 1, the chromosome is an elite. Elite chromosomes are reserved for the next generation.

4. The crossover operation is applied to the whole population except for the elites at the rate of pr_{cr}. Parents are selected according to the total fitness value of chromosomes. The total fitness value of chromosome $\boldsymbol{g}(p)$ is given by 1/rank(p).

   The mutation operation is applied to each gene of all the chromosomes except for that of the elite chromosomes at the rate of pr_{mut}.

5. Stop if the number of elite chromosomes becomes larger than a pre-set number or generations becomes larger than a pre-decided limit number. Otherwise, go to step 2.

Multiple elite models are obtained by the above processes. A human designer should select a desirable model out of the obtained elite models. If the designer regards that the conciseness (interpretability) of the model is more important than the predictive accuracy, he/she would choose a model with a smaller number of rules.

2.6.2 Numerical Experiments

Numerical experiments using a data set generated from a nonlinear equation are done. In this section, two coding methods introduced in sub-section 2.6.1 are examined. The results show that a new coding that decides the number of division of input space of each sub-model is effective for generating various models on the Pareto front.

We used the following nonlinear equation as the target system:

$$y = \sin(\pi x_1) + \cos(\pi(x_1 + x_2)) + \cos(\pi x_1 x_3) + \sin(\pi(x_1 - x_4)). \tag{2.16}$$

The input x_5 that had no relationship with y was also used. The value of each input was set at $\{0,0.5,1\}$. ($3^5 =$) 243 data were generated from this equation.

The first fuzzy sub-model in the first layer $\mathrm{FM}^{1,1}$ was identified. Figure 2.18 shows the AIC vs. the numbers of rules of all the possible combinations smaller than four input variables.

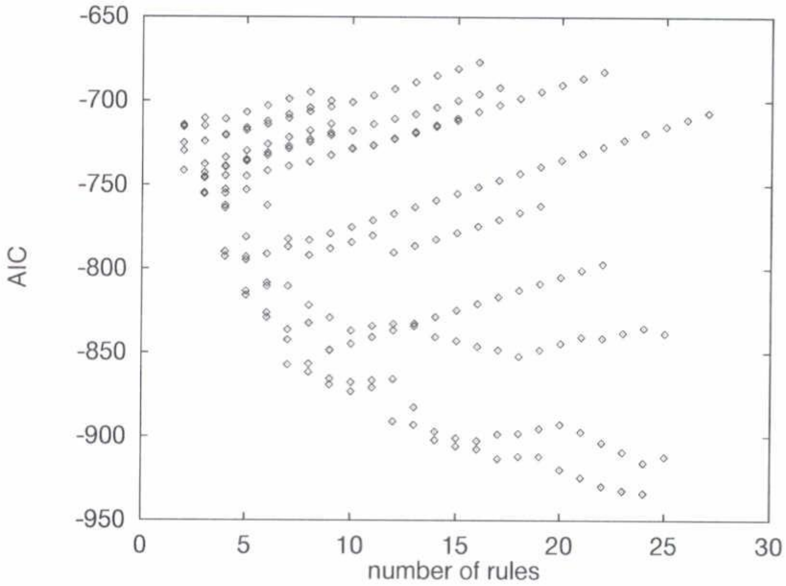


Figure 2.18: All possible solutions with smaller than 4 inputs

Figure 2.19 shows an example of elite models obtained by method 1. Seven solutions were obtained. Figure 2.20 shows another example of candidate models obtained by method 2. Fifteen solutions were obtained. These models generated by method 2 were spread widely over the Pareto front. The significant difference in these results was due to the difference in the coding method. Method 2 let the number of rules N_R be determined by genetic algorithm. So the method 2 could generate various candidates.

Figure 2.21 and Figure 2.22 show two examples of the obtained divisions of input space. In both cases, x_1 and x_2 were selected as input variables. The obtained fuzzy rules are shown in Table 2.7 and Table 2.8, respectively. The fuzzy rules can be expressed by linguistic labels as shown in the tables.

Table 2.9 shows another set of fuzzy rules obtained by method 2. There were 21 fuzzy rules. These rules had 3 input variables. This model had less amount of error but the

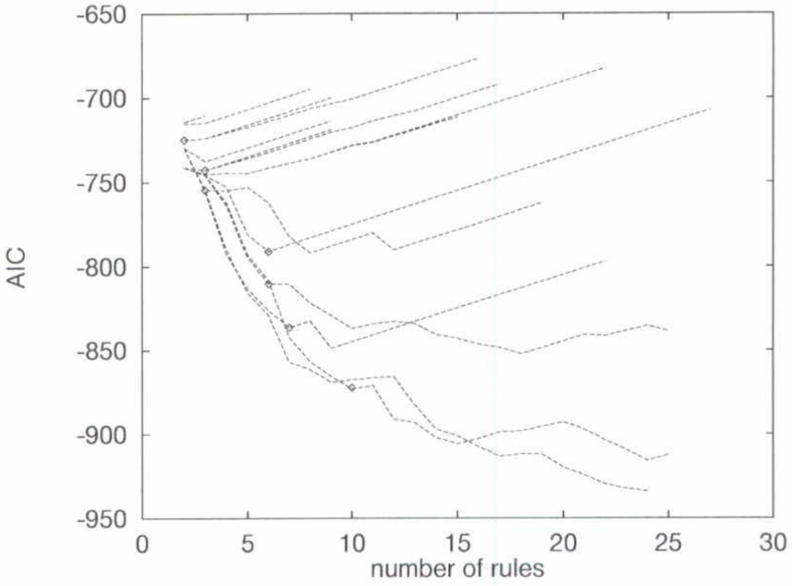


Figure 2.19: Models obtained by method 1

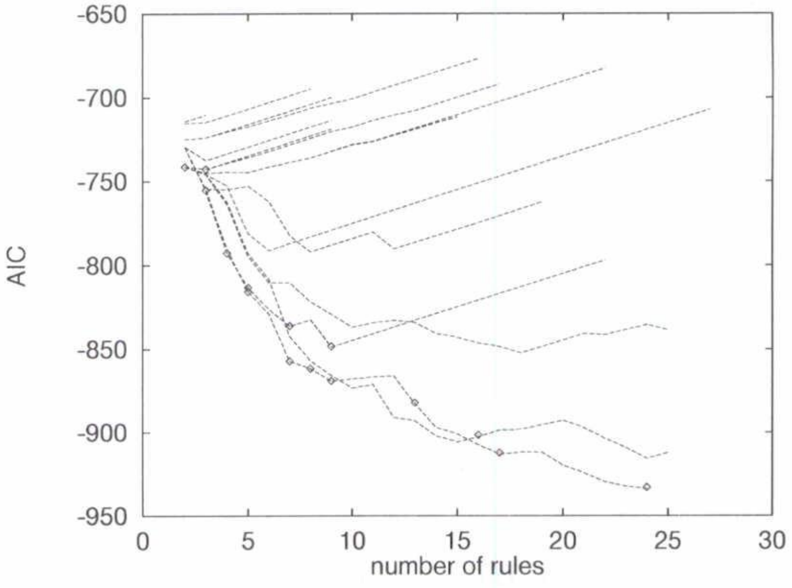


Figure 2.20: Models obtained by method 2

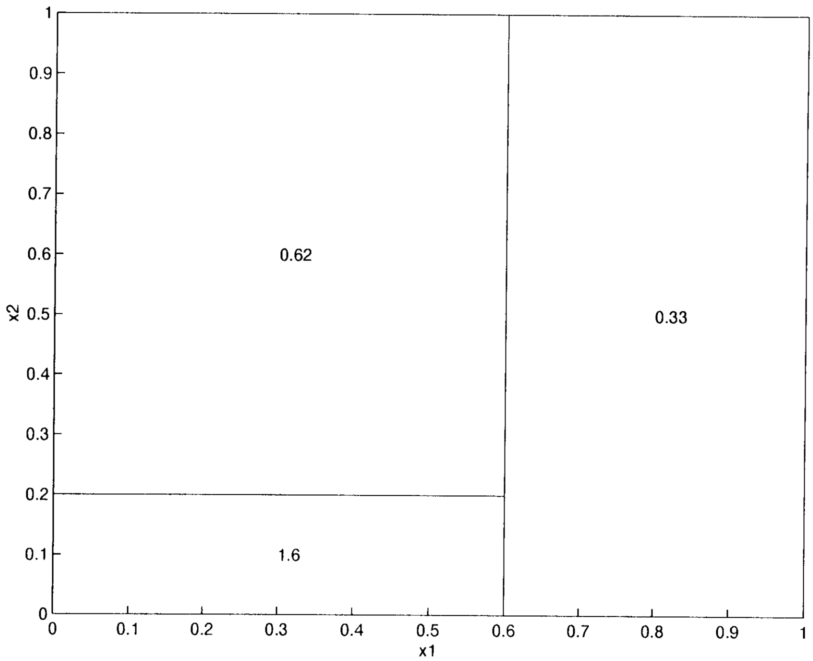


Figure 2.21: Obtained division of input space with 3 rules

Table 2.7: Obtained rule set with 3 fuzzy rules

IF	x_1 is Not Large and	x_2 is Small	THEN $y = 1.6$
IF	x_1 is Not Large and	x_2 is Not Small	THEN $y = 0.62$
IF	x_1 is Large		THEN $y = 0.33$

Table 2.8: Obtained rule set with 7 fuzzy rules

IF	x_1 is Not Large	and	x_2 is Small	THEN $y = 1.6$
IF	x_1 is Medium	and	x_2 is Large	THEN $y = 1.6$
IF	x_1 is Large	and	x_2 is Large	THEN $y = 1.3$
IF	x_1 is Medium	and	x_2 is Medium	THEN $y = 0.57$
IF	x_1 is Large	and	x_2 is Medium	THEN $y = 0.33$
IF	x_1 is Small	and	x_2 is Not Small	THEN $y = 0.16$
IF	x_1 is Large	and	x_2 is Small	THEN $y = -0.67$

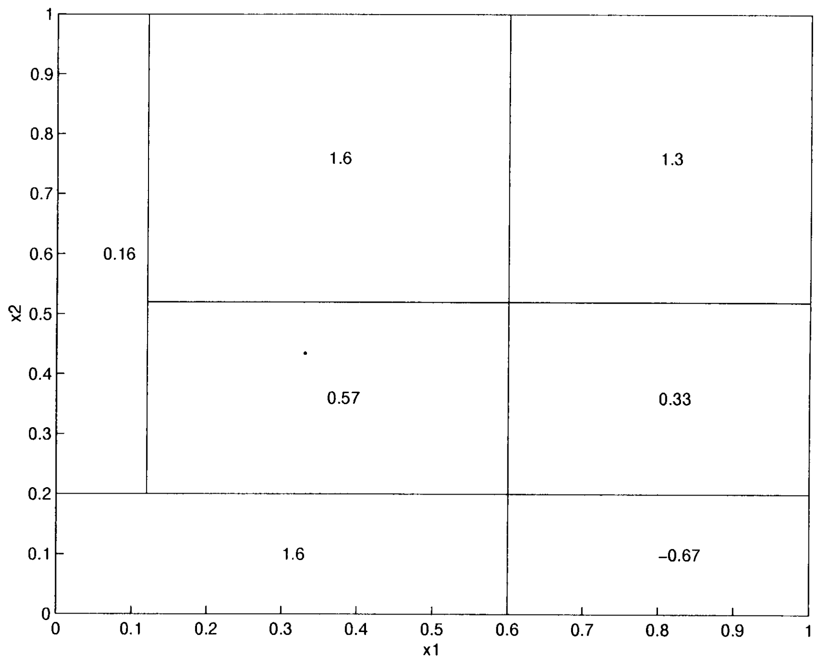


Figure 2.22: Obtained division of input space with 7 rules

number of rules were large. Human users can choose one of the solutions according to their own purposes.

Table 2.9: Obtained rule set with 21 fuzzy rules

IF	x_1 is Medium	and	x_2 is Small	and	x_3 is Small	THEN $y = 0.91$
IF	x_1 is Medium	and	x_2 is Not Small	and	x_3 is Small	THEN $y = 0.81$
IF	x_1 is Large	and	x_2 is Medium	and	x_3 is Large	THEN $y = 0.80$
IF	x_1 is Medium	and	x_2 is Large	and	x_3 is Medium	THEN $y = 0.71$
IF	x_1 is Medium	and	x_2 is Small	and	x_3 is Medium	THEN $y = 0.71$
IF	x_1 is Small	and	x_2 is Small	and	x_3 is Not Small	THEN $y = 0.70$
IF	x_1 is Small	and	x_2 is Not Large	and	x_3 is Small	THEN $y = 0.68$
IF	x_1 is Large	and	x_2 is Large	and	x_3 is Small	THEN $y = 0.60$
IF	x_1 is Large	and	x_2 is Medium	and	x_3 is Medium	THEN $y = 0.60$
IF	x_1 is Large	and	x_2 is Large	and	x_3 is Large	THEN $y = 0.60$
IF	x_1 is Medium	and	x_2 is Large	and	x_3 is Large	THEN $y = 0.51$
IF	x_1 is Medium	and	x_2 is Small	and	x_3 is Large	THEN $y = 0.51$
IF	x_1 is Small	and	x_2 is Medium	and	x_3 is Not Small	THEN $y = 0.50$
IF	x_1 is Medium	and	x_2 is Medium	and	x_3 is Not Small	THEN $y = 0.41$
IF	x_1 is Large	and	x_2 is Medium	and	x_3 is Small	THEN $y = 0.40$
IF	x_1 is Large	and	x_2 is Small	and	x_3 is Medium	THEN $y = 0.40$
IF	x_1 is Large	and	x_2 is Medium	and	x_3 is Large	THEN $y = 0.40$
IF	x_1 is Small	and	x_2 is Large	and	x_3 is Small	THEN $y = 0.40$
IF	x_1 is Small	and	x_2 is Medium	and	x_3 is Not Small	THEN $y = 0.30$
IF	x_1 is Large	and	x_2 is Small	and	x_3 is Large	THEN $y = 0.20$
IF	x_1 is Large	and	x_2 is Small	and	x_3 is Small	THEN $y = 0.20$

2.7 Conclusion

This chapter proposed new hierarchical fuzzy modeling methods using genetic algorithm. In the proposed method, the combination of input variables is determined by genetic algorithm. Two division methods of input space were presented, one was based on the grid-type division and the other was on the tree-type division. Numerical experiments showed (1) the proposed grid-type division method obtained more concise models than the conventional methods did and (2) the tree-type division method obtained more concise models than the grid-type division method did.

Multiple-objective genetic algorithm was applied to the hierarchical fuzzy modeling to obtain various models having conciseness and predictive accuracy different from each other. Two methods of coding were proposed for the fuzzy modeling. One method encoded only the combination of input variables. The other method encoded the combination of input variables and the allowable number of fuzzy rules. A numerical experiment showed the latter method was able to obtain more various models, spread widely on the Pareto front.

Chapter 3

Application to Biomechanics of Cyclists

3.1 Introduction

Cyclists utilize the lower limbs to propel the body and the vehicle. This motion is complicated. When muscle contraction generates power, the angle of joints changes at the same time. The change of joint angle alters the distances between the fulcrum, the point of action and the point of a lever where the force is applied. That changes muscle contraction speed. The shape and the inertia of each segment as well as the time-series patterns of the power supplied by the muscle contraction affect the bicycle speed. The motion is hard to comprehend with a mechanical model.

Some scientists measured the force exerted on the pedals and discussed good patterns of pedaling force. But, such data are difficult to explain explicitly. Fuzzy modeling has not been applied to these data before.

This chapter presents a biomechanical analysis of cycling by the fuzzy modeling method developed in the previous chapter. The relationships between the force patterns on the pedal and the output power of cyclists are described with fuzzy rules. The hierarchical fuzzy modeling selects out the most influencing variables among the absolute force, the angle of the pedal and the force direction in each quarter of the pedaling cycle. Feedback to the athletes to improve their output power is also obtained, utilizing the activation values of fuzzy rules for each pedaling data.

3.2 Force Pattern Analysis of Cycling

Pedaling data of fourteen international level cyclists [14] are analyzed to extract pedaling know-how. All cyclists were measured while they were pedaling at their maximal effort. Six of the 14 cyclists were also measured while they were pedaling at 90% of their maximal outputs. The number of data was 20. The data set consisted of pedal angle θ, normal and tangential forces to the pedal as functions of crank angle ranged from 0 to 360 degrees. Two variables, absolute value of the force exerted on the pedal $|F|$, and angle of the force direction to the pedal θ_F were calculated. The variables $|F|$, θ_F and θ were averaged in every 90 degrees of the crank angle. Three variables in 4 ranges, totally 12 variables, were obtained as the inputs. The output variable y was average output power [watt].

Figure 3.1 depicts force patterns of 5 cyclists. Cyclist A had the weakest output

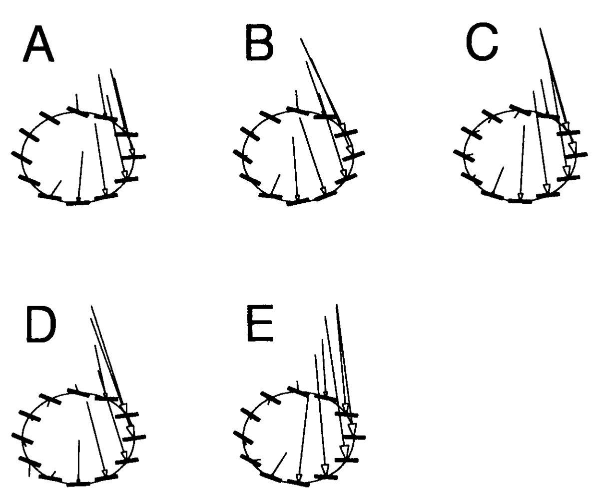


Figure 3.1: force patterns of five cyclists

among the 14 cyclists. Cyclist B was better than A, and cyclist C was better than B and so on. Cyclist E was the best of the 14.

Common interests by cyclists with this kind of data were: the factor that affected the output most and :the point that they should concentrate on to improve their outputs.

3.2.1 Hierarchical fuzzy model of cyclists' data

The hierarchical fuzzy modeling with tree-type division presented in section 2.5 and the coding method 2 presented in section 2.6.1 was utilized to analyze the data. Figure 3.2 depicts the final structure of the obtained hierarchical model. The sub-models SM {1-1} and SM {1-2} were identified. The outputs of the obtained two sub-models were added in the second layer.

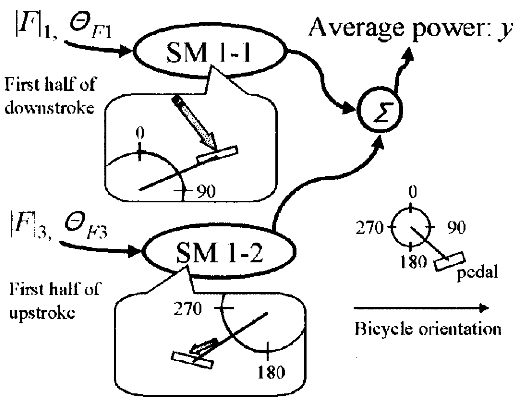
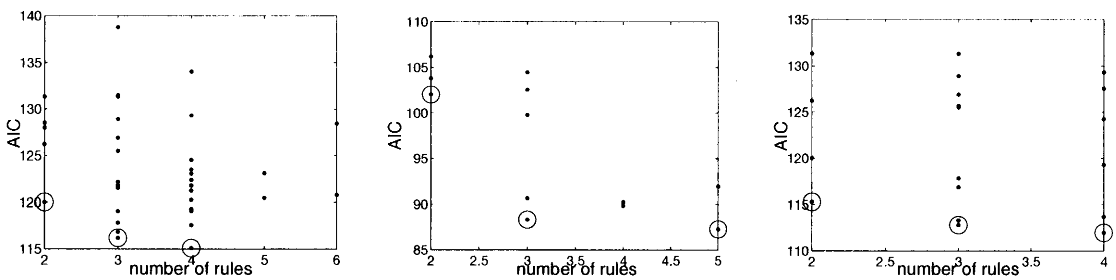


Figure 3.2: Obtained hierarchical model

Figure 3.3(a) shows the solutions for the first sub-model in the first layer searched by the multiple objective genetic algorithm. The solutions, which were on the Pareto front in the final generation, are indicated by circles. The author selected the model



(a) 1st model in 1st layer (b) 2nd model in 1st layer (c) 1st model in 2nd layer

Figure 3.3: Solutions searched by MOGA for each sub-model

with 4 rules and 2 inputs out of the compromising solutions on the Pareto front as the sub-model SM {1,1}. Table 3.1 shows the fuzzy rules of this sub-model. The two inputs were $|F|_1$: the force strength; and θ_{F1}: the force direction; in the first half of downstroke. Linguistic labels and ranges in which the grades are greater than 0.5 are shown. The consequent singletons y_1 are also shown. The universe of discourse of the input is shown in the bottom row. The obtained rules are interesting. Though the

Table 3.1: Fuzzy rules in sub-model 1-1

No.	$\|F\|_1$	θ_{F1}	y_1
1	Large (315 ∼)	Large (−78 ∼)	253
2	Medium (274 ∼ 315)	—	209
3	Large (315 ∼)	Small (∼ −78)	207
4	Small (∼ 274)	—	184
	176 ∼ 371	−83 ∼ −48	

force $|F|_1$ of rule 3 was stronger than that of rule 2, the application of the force to a wrong direction, i.e. small θ_{F1} limited the output.

Further information would not be obtained from this quite a small number of data unless hierarchical modeling was utilized.

The second sub-model SM {1,2} was identified using the error of sub-model 1-1 as its output. Figure 3.3(b) shows the solutions for sub-model 1-2. I chose the model with 3 rules and 2 inputs. Table 3.2 shows the rules. The strength $|F|_3$ and the force

Table 3.2: Fuzzy rules in sub-model 1-2

No.	$\|F\|_3$	θ_{F3}	y_2
1	Large (97 ∼)	Large (−137 ∼)	+16
2	—	Small (∼ −137)	+6
3	Small (∼ 97)	Large (−137 ∼)	−15
	42 ∼ 114	−244 ∼ −97	

direction θ_{F3} in the first half of upstroke were selected as the inputs of the compensative sub-model. These rules were also interesting. If the force direction was small (nearly parallel to the pedal surface), the strength did not contribute to the output. If it is large, the error of SM {1,1} depends on $|F|_3$.

For the first sub-model in the second layer SM {2,1}, the outputs of sub-models in the first layer were also the candidates of its inputs. Figure 3.3(c) shows the solutions for sub-model 2-1. We chose the model with 4 rules and 2 inputs. Table 3.3 shows the rules. The rules told that the compensation by the sub-model 1-2 was needed only for the cyclists in intermediate levels.

Table 3.3: Fuzzy rules in sub-model 2-1

No.	y_1	y_2	y
1	Large (229 $\sim$)	—	253
2	Medium (196 $\sim$ 229)	Large (-15 $\sim$)	212
3	Medium (196 $\sim$ 229)	Small ($\sim$ -15)	188
4	Small ($\sim$ 196)	—	184
	183 $\sim$ 266	-18 $\sim$ $+12$	

But, the model did not have less error than the sum of the output of sub-model 1-1 and that of sub-model 1-2. So, in the second layer, a sum unit was employed instead of the fuzzy model SM {2,1}.

The rooted mean squared errors of the hierarchical fuzzy model was 7.1 watts that was 3.3 percents of the average output, 211 watts. The model had 12 degrees of freedom in total. A linear regression with 13 degrees of freedom gave a smaller error. The rooted mean squared errors with the linear regression was 4.8 watts. However, more detailed information were extracted from the fuzzy model as discussed in the next subsection.

3.2.2 Advice Extraction from Fuzzy Rules

The fuzzy model gave us more information about the pedaling technique. The strength and the force direction in the first half of downstroke were important. And those in the first half of upstroke were also important. This information was hard to be extracted from the linear regression. The correlation coefficients of input variables to the output were 0.78 at the maximum. The most related input variable to the output was $|F|_2$: strength in the second half of downstroke. The second maximal coefficient was 0.60 of the strength in the first half of downstroke $|F|_1$. The correlation coefficients of the force directions in the halves of downstroke θ_{F1} and θ_{F2} were -0.08 and 0.47, respectively. The correlation coefficients of the force strength $|F|_3$ and direction θ_{F3} in the first half of upstroke, which were selected by the fuzzy model, were 0.52 and 0.20, respectively.

The fuzzy model found out the combination effects of the force strength and the direction, which the linear regression could not reveal. Furthermore, the fuzzy model was able to give detailed information as follows: Table 3.4 shows the normalized activation values of four fuzzy rules in SM {1,1} and three fuzzy rules in SM {1,2} in the case of five cyclists A-E. The activation values are normalized so that the sum of the all activation values of a sub-model is 100%.

Table 3.4: Activation values of fuzzy rules of five cyclists

	rule no.(consequent singleton)							output	
	1 (253)	2 (209)	3 (207)	4 (184)	1 (+16)	2 (+6)	3 (-15)	inferred	actual
A	0	0	0	100	0	51	49	176.2	163.0
B	29	61	8	2	3	33	64	203.5	196.4
C	11	0	74	16	70	9	20	222.7	230.5
D	61	37	1	1	0	100	0	244.2	248.2
E	97	0	1	2	63	28	9	278.5	279.6

Cyclist A matched the 4th rule in sub-model 1-1 and it matched the 2nd and the 3rd rules in sub-model 1-2 half and half. Referring to table 3.1, the major cause of his relatively low performance was explained by the small force in the first half of downstroke. And referring to table 3.2, the minor cause was explained by the application of force to a wrong direction in the first half of upstroke.

According to the second row in table 3.4, table 3.1 and table 3.2, the pedaling of cyclist B was characterized by the medium force in the first half of downstroke and the small force in the first half of upstroke. He should care his upstroke more than downstroke.

The third line in table 3.4 and the rule tables tell that the pedaling of cyclist C was characterized by the strong but wrong directional force in the first half of downstroke and the large force in the first half of upstroke. He should mind the force direction in his downstroke.

Cyclist D had acquired proper force direction in the first half of downstroke, because the 1st and the 2nd rules of sub-model 1-1 had higher activation values. He should improve strength in downstroke and correct the direction in upstroke.

These pieces of advice were extractable because the acquired fuzzy models had the interpretability.

3.3 Conclusion

Pedaling patterns of U.S. national team cyclists were analyzed using hierarchical fuzzy modeling. The most influencing variables were the absolute force and the force direction in the first half of the downstroke. These variables were the inputs of the first model. The absolute force and the force direction in the first half of upstroke affected the output power as the compensation of the first model. These variables had mutual interactions and saturations. The obtained fuzzy rules clarified the nonlinear relationships.

Feedback to the athletes was obtained from activation values of the fuzzy rules for each pedaling data. Some cyclists should improve force in the first half of downstroke. Other should concentrate on the force direction rather than increasing the amount. Other should correct the technique in upstroke.

Chapter 4

Applications to Biomechanics of Rowers

4.1 Introduction

Rowers utilize the trunk and the arms as well as the legs to propel the body and the vehicle. The whole body motions of the rowers become more complex than those of cyclists. Furthermore, the center of body mass is moved forward and backward on the boat. The movement makes the boat surge and pitch. That causes additional complexity.

Conventionally, the forces exerted on the oars have been measured and analyzed. But, any of the studies did not go further than plotting the measured patterns of good rowers and comparing them to those of other rowers. The interpretation of the force patterns was mainly dependent on the intuition of the researchers.

The first application of the fuzzy modeling to biomechanics of rowers, presented in section 4.2, investigates the relationships between the force pattern and the average power applied to the oar. Theoretically, the boat speed is proportional to the cube square of the average power. According to preceding studies and comments of coaches, maximal force is influencing, of course, and the timing and the speed of stroke affect the power as well. This study investigates maximal force, wasted time and angular velocities of beginning, middle and ending parts of the stroke as factors that may affect the average power.

Rowers apply power mainly by three partial motions, extension of the legs, swing of the trunk and pulling of the arms. Coaches know empirically that the coordination of the partial motions is important for high performance. But, the coordination has not been studies for on-water rowing or sculling. Rowers are able to improve the performance more easily if the partial motion that should be corrected is specified clearly. The second application of the fuzzy modeling investigates the relationships between the partial motions and the rowing performance in section 4.3.

4.2 Force Pattern Analysis of Rowing

Rowers and coaches intend to progress the speed of their boat. Good skill must be pursued to realize an excellent performance. The boat speed is theoretically proportional to the cube root of the power supplied by rowers via oars. To enhance the power, the relationship between the force patterns and the output power should be explicitly clarified.

Smith and Spinks [20] used discriminant analysis to classify the force patterns of novice, good and elite rowers. This method distinguished each class with linear connection of input biomechanical factors. I noticed the necessity of utilizing a nonlinear modeling method to describe the relationships more accurately. Fuzzy modeling identifies nonlinear input-output dependencies with high interpretability.

4.2.1 Data Collection and Preprocessing

The forces exerted on oars and the angles of oars were measured while women national team rowers were rowing 100m runs. Using the measured data, I examined the relationships between the boat speed and the power. Then, we identified the relationships between the power and the skill parameters in a form of a fuzzy model.

Subjects and measurement:

Six women rowers participated. Two of them rowed a double scull as a crew. They rowed 4 runs of 100 meters. They were directed to row at 20, 24, 26 and 28 strokes per minutes in the four runs,respectively. Other 4 crews carried out the same runs. During these runs, the forces exerted on oars were measured by strain gauges, and the angles were measured by goniometers. The sampling frequency was 50Hz. The data were low-pass filtered with the cut-off frequency of 12.5Hz. An example of measured force and angle data is shown in figure 4.1.

Boat speed dependence on power

We calculated power at each time step, by multiplying force $F_i(t)$, angular velocity $\omega_i(t)$ and outboard length of the oar l_i. The angular velocity was the difference of the angle data ($\omega_i(t) = (\theta_i(t) - \theta_i(t-1))/(\delta t)$, where δt is the sampling time).

$$P_i(t) = l_i F_i(t) \omega_i(t) \tag{4.1}$$

where the index i denotes each oar. A double scull has four oars, so $i = 1, \cdots, 4$. The average power supplied by the four oars is:

$$P = \frac{\sum_{t=1}^{T} \sum_{i=1}^{4} P_i(t)}{T} \tag{4.2}$$

where T is the final time step of the 100 m run. The boat speed is:

$$v = \frac{100}{T\,\delta t}. \tag{4.3}$$

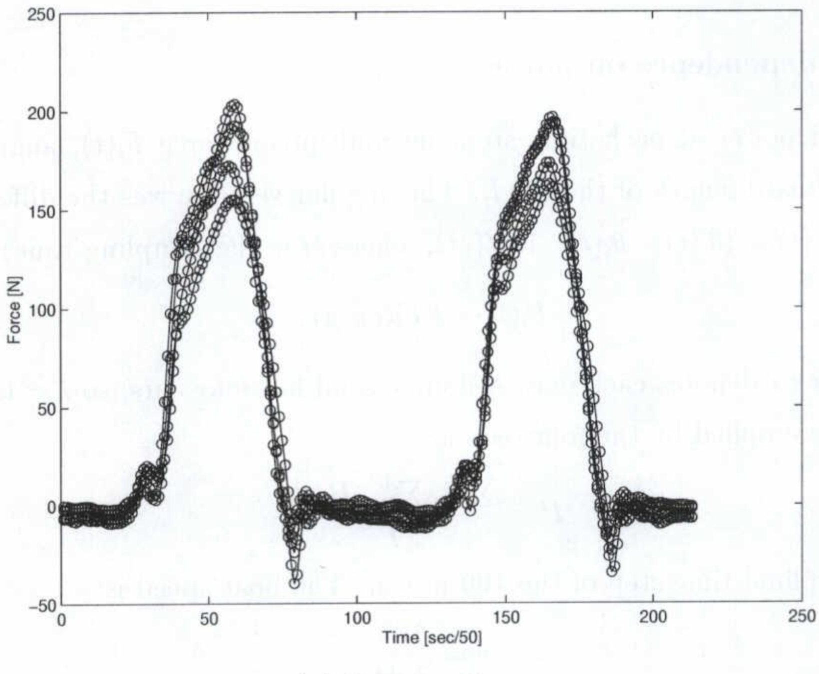
It is a well-known theory that the resistance force of a rigid body moving in a fluid is proportional to the square of the velocity. So, the power is proportional to the cube of the velocity.

Circles in figure 4.2 show the relationship between the boat speed and the average power. The curve in figure 4.2 shows the theoretical curve, the cube root of the power. The main reason of the deviations was due to the disturbance of the wind.

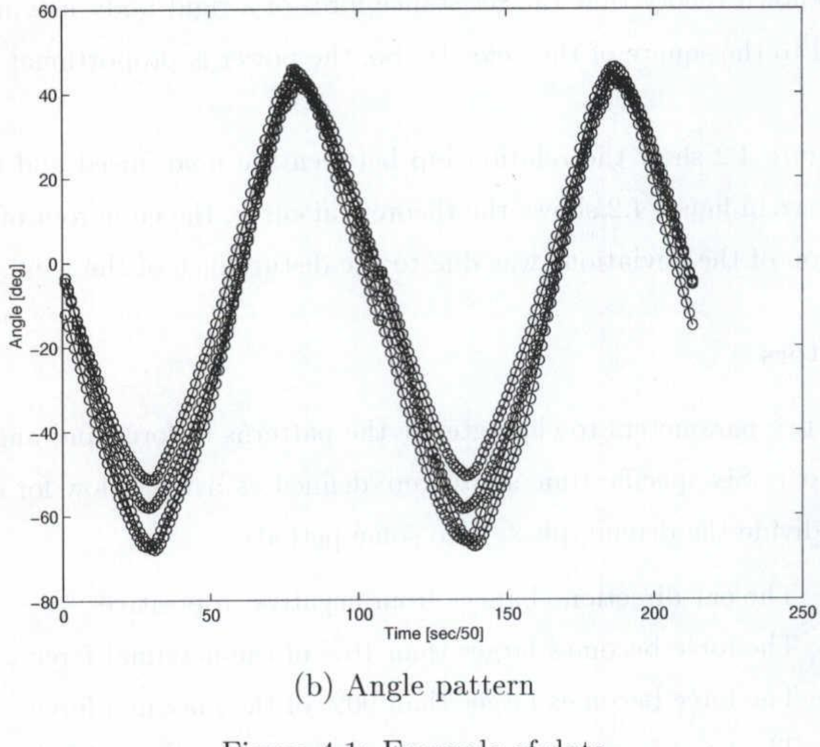
Skill parameters

We calculated five parameters to characterize the patterns of force and angle for each stroke of each oar. Six specific time steps were defined as listed below for each stroke of each oar to divide the driving phase into some periods.

- t_1: The oar direction changes from negative to positive
- t_2: The force becomes larger than 10% of the maximal force
- t_3: The force becomes larger than 90% of the maximal force
- t_4: The force becomes smaller than 90% of the maximal force
- t_5: The force becomes smaller than 10% of the maximal force
- t_6: The oar direction changes from positive to negative



(a) Force pattern



(b) Angle pattern

Figure 4.1: Example of data

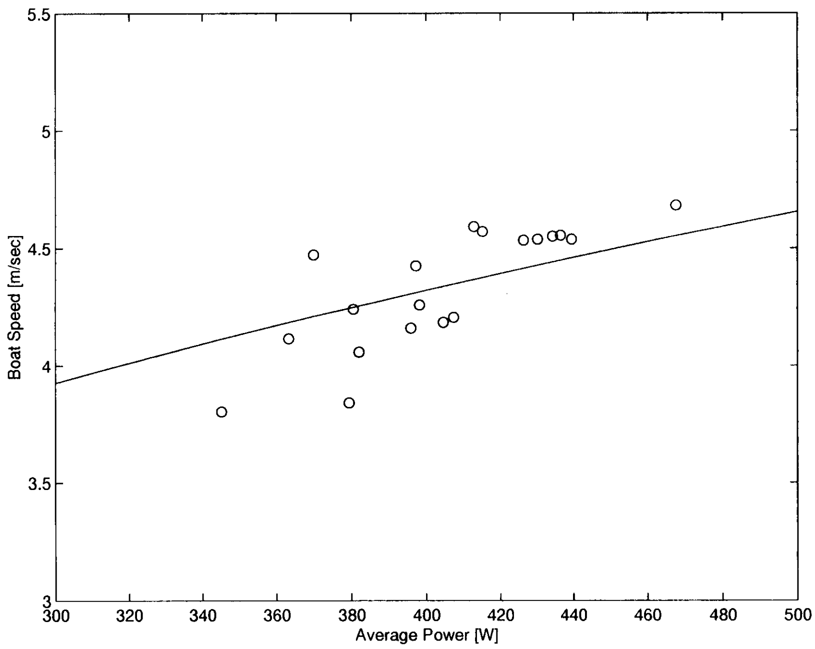


Figure 4.2: Boat speed dependence on power

The followings are the parameters of rower's skill:

- Maximal force: $F_m = \max_{t=t_1}^{t_6} F(t)$
- Wasted time: $t_w = (t_2 - t_1) + (t_6 - t_5)$
- Angular velocity in the beginning of effective driving: $\omega_B = \frac{\theta(t_3)-\theta(t_2)}{(t_3-t_2)\delta t}$
- Angular velocity in the middle of effective driving: $\omega_M = \frac{\theta(t_4)-\theta(t_3)}{(t_4-t_3)\delta t}$
- Angular velocity at the end of effective driving: $\omega_E = \frac{\theta(t_5)-\theta(t_4)}{(t_5-t_4)\delta t}$

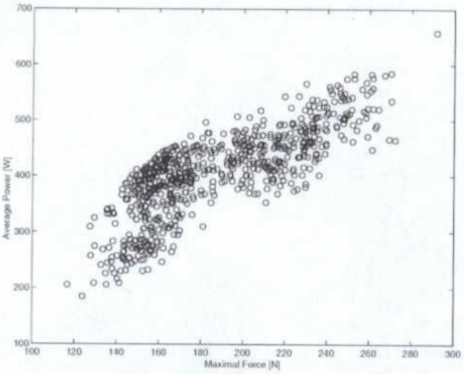
A rower can improve the performance more easily when she concentrates on one of these parameters than when she tries to improve all of them at the same time. The relationships between these parameters and the output power are shown in figure 4.3.

4.2.2 Fuzzy Models and Advice Extraction

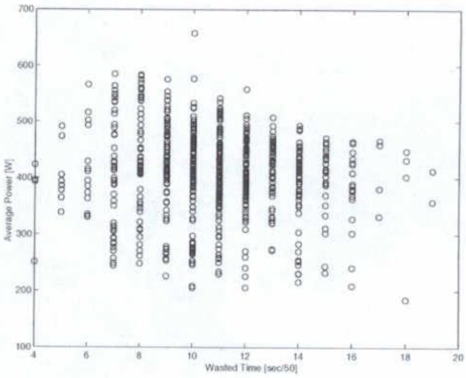
The hierarchical fuzzy modeling method with tree-type division presented in section 2.5 and the coding method 2 presented in section 2.6 was applied to the data in the previous sub-section. The obtained fuzzy rules are shown in Table 4.1. They are listed and numbered in order of the output power. The relationships between the skill parameters and the power were described with 5-inputs 1-output rules. The modeling process stopped identifying the fuzzy model in the first layer. The rooted mean squared error of the fuzzy model was 26W, which was 6.4% of the average output. Effective advice is extractable from this table. For example, Rule 8 has a difference in FM from rule 9. So "push strongly" is effective advice for a rower who can row about 570 watts in power. If we look at rule 6 and 8, the difference is in t_w. "Reduce wasteful motion" is very good advice for a rower who outputs about 470 watts. There is a difference in ω_M between rule 7 and rule 8. "Speed up in the middle" is effective for a rower outputting about 530 watts.

4.3 Motion Analysis of Rowing

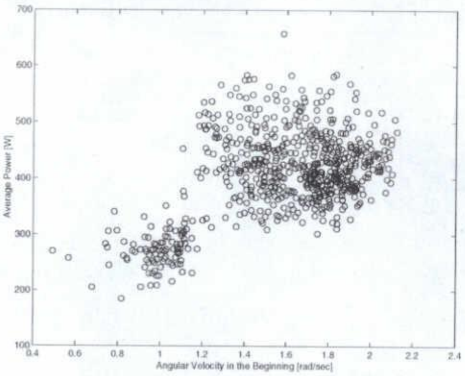
The rowing performance depends on biomechanical factors and energetic factors. For lightweight class rowers, biomechanical factors are more important because the regulation of body weight limits energetic factors. To analyze biomechanical factors of



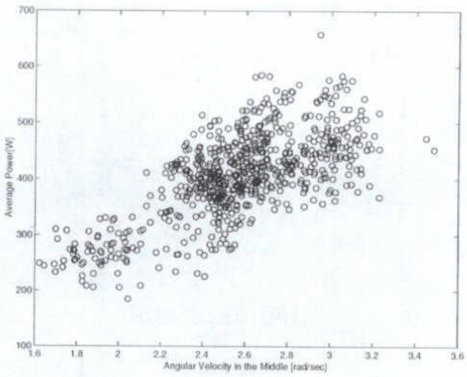
(a) Maximal force



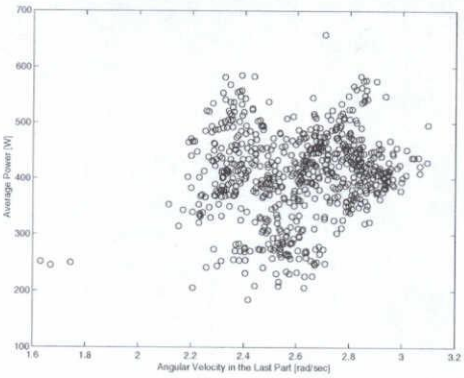
(b) Wasted time



(c) Angular velocity in the beginning



(d) Angular velocity in the middle



(e) Angular velocity at the end

Figure 4.3: Relationships between skill parameters and average power

Table 4.1: Obtained linguistic rules with the fuzzy model

Rule Number	If part					Then part
	F_m	t_w	ω_B	ω_M	ω_E	P
1	Small		Small			253
2	Small		Large		Small	344
3	Very Small		Large		Large	362
4	Large		Large	Small		412
5	Medium Small		Large		Large	436
6	Large	Large		Large		469
7	Large		Small	Small		528
8	Medium Large	Small		Large		576
9	Very Large	Small		Large		617

rowing, propulsive power has been studied [30, 31, 32]. The propulsive power is the final result of the contributions of the whole body parts. Partial motions, such as leg extension, trunk swing and arm pull, are easier to sense and control for a rower.

However, there are two hurdles to give rowers effective suggestions associated with the partial motions. One is that it is hard to measure rower's muscle power directly during on-water rowing. The other is that, by only showing the graphs of partial motions, it is difficult for athletes and coaches to understand which motion is suitable for high performance and which one should be corrected.

Against the first hurdle, this study utilizes inverse dynamics [33] to calculate power of the partial motions. Against the second one, we utilize the fuzzy modeling, which give the athletes and the coach explicit knowledge on the relationships between the motions and the performance.

This section presents a combination of the inverse dynamics and the fuzzy modeling for biomechanical feedback to rowers. When the subjects row 100m by single scull, forces exerted on the oars, angles of the oars in the horizontal plane and the acceleration of the boat are measured. Joint positions of the rower are videotaped and digitized. The inverse dynamics is applied to the data. It outputs time series patterns of joint power. Average power and the representative timing of the three partial motions are calculated as parameters of motion of the rower. Efficiency is also calculated as one of two performance indices. The other index is the boat speed, which is measured by an electromagnetic speedometer.

The six parameters are the inputs of the fuzzy model and the performance indices are the outputs. Fuzzy modeling identifies nonlinear input-output relationships. It is easy for a human to extract linguistic advice from the fuzzy rules. The advice is fed back to the rower. The rower refers to the feedback and remembers the original motion at the same time, then creates a new motion.

4.3.1 Data Collection and Preprocessing

Subjects and Measurement

The subjects were eight rowers from the Japanese national team and seven rowers from a university rowing team. All of them were male and lightweight class. When each of them rowed 100m runs a few times by single scull, forces on oar handles f_{OAR}, oar angles in the horizontal plane θ_{OAR} and the acceleration of the boat were measured. Knee, hip, waist, shoulder, elbow and handle positions were videotaped, digitized and projected to the sagittal plane. The sampling frequency was set at 15 Hz. An electromagnetic speedometer measured the average speed of the boat.

Inverse Dynamics

A two-dimensional link segment model with six body segments, i.e. shank, thigh, pelvis, trunk, upper arm and forearm, was constructed. The coordinate system had x-axis to the boat direction and y-axis to the vertical direction with a reference point on the still water. Figure 4.4 illustrates the link segment model.

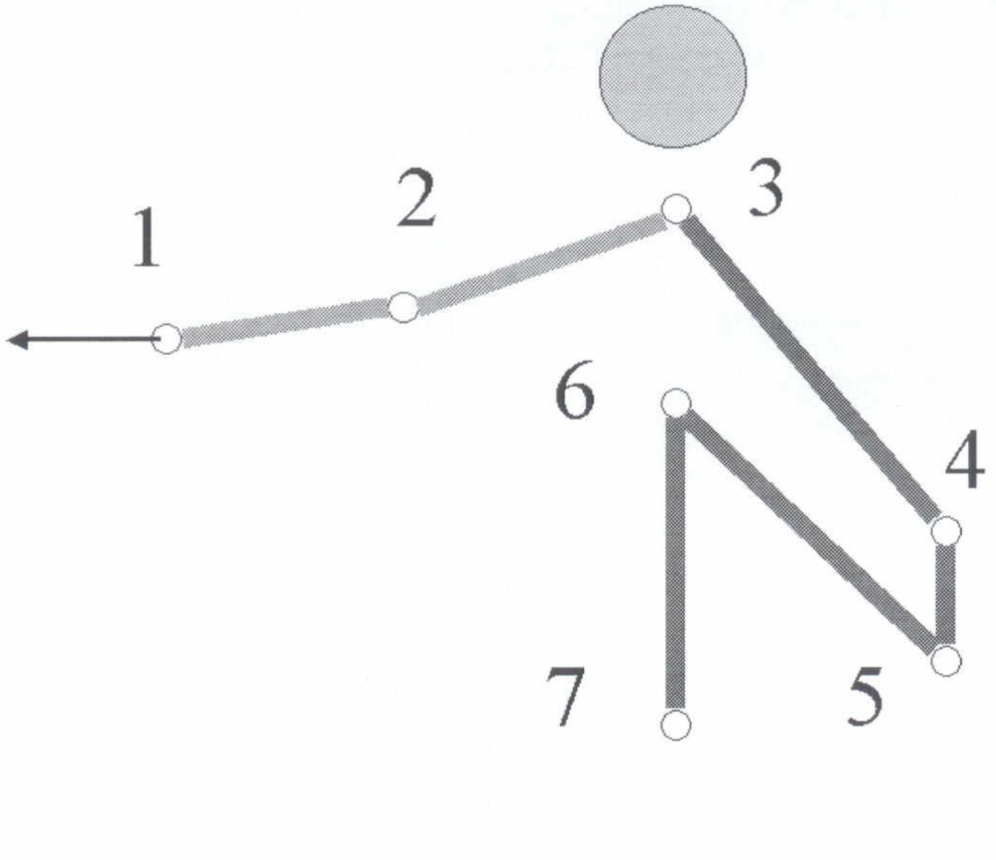


Figure 4.4: Link segment model in the sagittal plane

Joint forces were calculated from the elbow to the foot:

$$\boldsymbol{f}_{i+1} = \boldsymbol{f}_i + m_i(\boldsymbol{g} - \ddot{\boldsymbol{x}}_i) \tag{4.4}$$

where $\boldsymbol{g} = [0, \ -g]^T$ was the gravity acceleration, m_i and $\boldsymbol{x}_i = [x_i, y_i]^T$ were the mass and the position of i-th segment, respectively, double dots denote the second derivative. The hand force was given from the measured variables,

$$\boldsymbol{f}_1 = [f_{OAR} \sin\theta_{OAR}, 0]^T. \tag{4.5}$$

The vertical force on the hip from the seat f_{SEATy} was estimated by,

$$f_{SEATy} = \frac{(\boldsymbol{r}'_{HAND} \times \boldsymbol{f}_1 + \boldsymbol{r}'_{COG} \times M\boldsymbol{g})}{x'_{SEAT}} \tag{4.6}$$

where $\boldsymbol{r}'_{HAND}$ and $\boldsymbol{r}'_{COG}$ were the relative displacement to the foot of the hand and the center of gravity, respectively, M was the body mass, x'_{SEAT} was the relative displacement of the seat. The force was added in the equation of pelvis motion. The joint torque T_i was calculated by:

$$T_{i+1} = T_i + \boldsymbol{a}_i \times \boldsymbol{f}_i + \boldsymbol{b}_i \times (-\boldsymbol{f}_{i+1}) - I_i\ddot{\theta}_i \tag{4.7}$$

where $\boldsymbol{a}_i$ and $\boldsymbol{b}_i$ were the relative displacements of the segment ends to the segment center of mass, I_i and $\ddot{\theta}_i$ were the inertial moment and the angular acceleration, respectively, in the sagittal plane. The joint torque power P_i was calculated by:

$$P_i = T_i \cdot (\dot{\theta}_i - \dot{\theta}_{i-1}). \tag{4.8}$$

The efficiency η was

$$\eta = \frac{\text{(output work)}}{\text{(output work)} + \text{(internal consumption work)}} \tag{4.9}$$

where the output work and the internal consumption work were:

$$\text{(output work)} = \sum_t \{\dot{\boldsymbol{f}}_7 \cdot \dot{\boldsymbol{x}}_7 - \dot{\boldsymbol{f}}_1 \cdot \dot{\boldsymbol{x}}_1\}, \tag{4.10}$$

and

$$\text{(internal consumption work)} = \sum_t \sum_{i=2}^{7} P_i, \tag{4.11}$$

respectively.

Time-series patterns of leg extension power $P_{LE}(t)$, trunk swing power $P_{TS}(t)$, and arm pull power $P_{AP}(t)$ were given by $P_7(t) + P_6(t)$, $P_5(t) + P_4(t)$, and $P_3(t) + P_2(t)$, respectively. Totally 43 runs of the 15 rowers were analyzed. The power patterns of two rowers are shown in Figure 4.5. Both of them applied leg extension in the first half, trunk swing in the middle, and arm pull in the second half of the driving phase. The power patterns of each run were parameterized to average value and representative timing during the driving phase as shown with asterisks in the graphs.

4.3.2 Fuzzy Models and Advice Extraction

Fuzzy models were identified for the boat speed and the efficiency separately by applying the hierarchical fuzzy modeling method with tree-type division presented in section 2.5 and the coding method 2 presented in section 2.6.

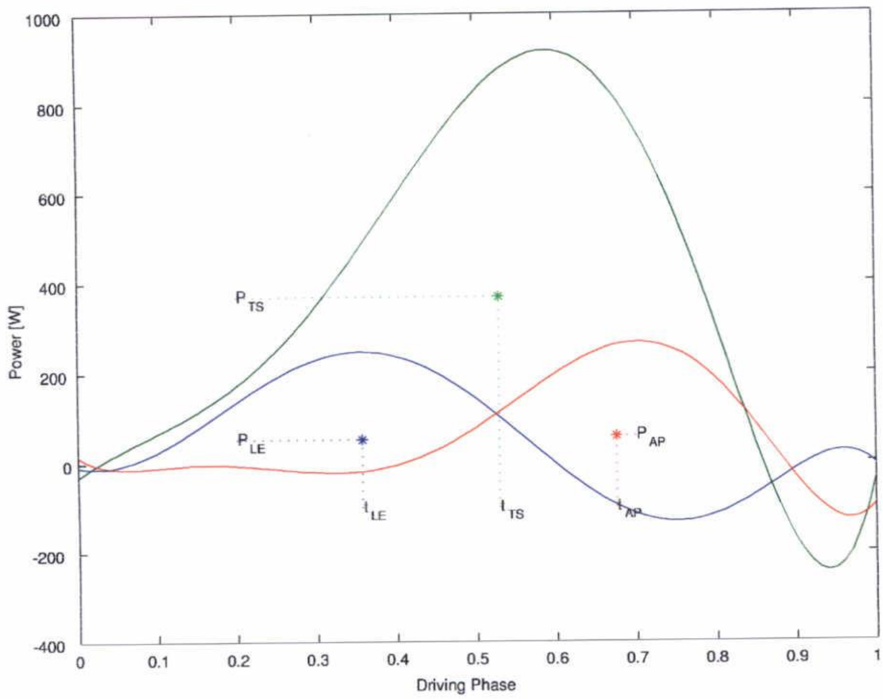
Fuzzy model of boat speed

Three fuzzy rules were obtained for the boat speed. The rules are shown in table 4.2. The hierarchical fuzzy modeling identified only the first model in the first layer.

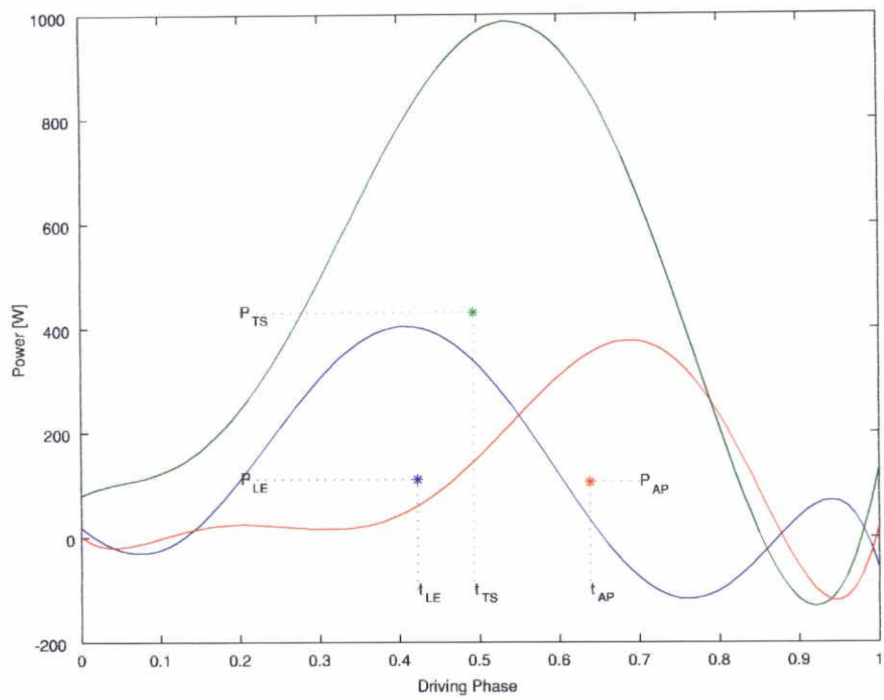
Table 4.2: Fuzzy rules of boat speed v, and activations of rules for rowers A and B

Antecedent part	Consequent part	Rower A	Rower B
IF P_{TS} is Strong AND t_{LE} is Early,	THEN $v = 4.53$	55	4
IF P_{TS} is Strong AND t_{LE} is Late,	THEN $v = 4.08$	44	96
IF P_{TS} is Weak,	THEN $v = 3.61$	1	0

The trunk swing power affected the boat speed. The timing of leg extension affected v in the case where P_{TS} was strong. The third and the fourth columns in table 4.2 show the activation values of rower A and those of rower B, respectively. Figure 4.6 shows the input-output relationships. Green dots in the top graphs show P_{TS}, t_{LE} and v of rower A. Red dots show those of rower B. The left bottom graph shows the membership functions of "P_{TS} is Strong" and "P_{TS} is Weak", respectively. Both A and B had so large P_{TS}, more than 350 W, that the grade of "P_{TS} is Strong" was almost 100%, and that of "P_{TS} is Weak" was almost zero. The right bottom graph shows



(a) Rower A



(b) Rower B

Figure 4.5: Power patterns and parameters of rowers A and B

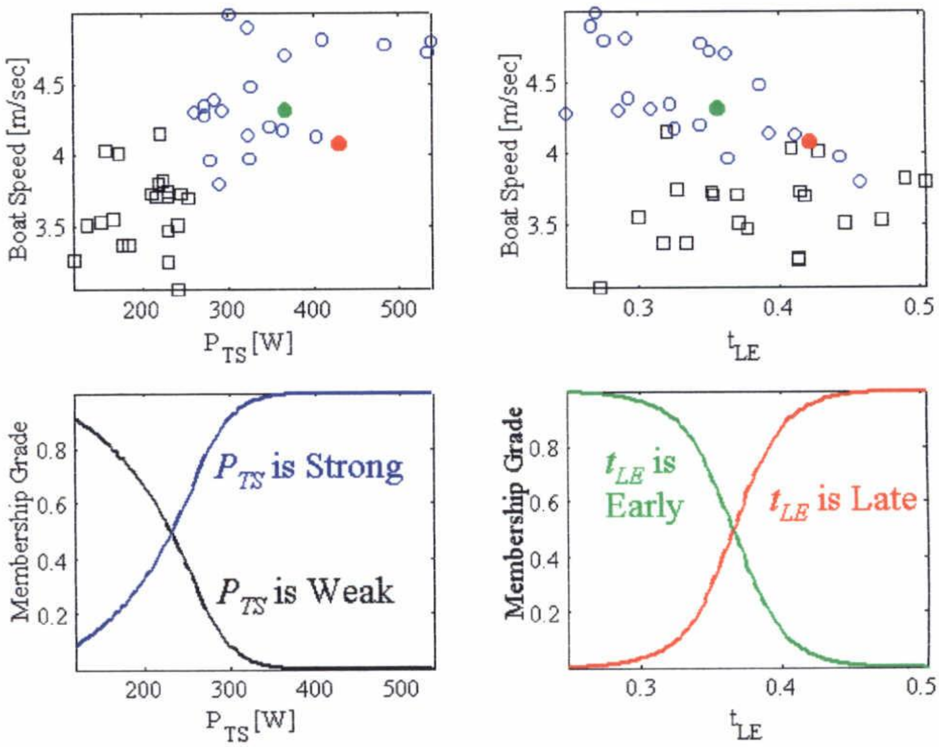


Figure 4.6: Relationships between P_{TS}, t_{LE} and v (top graphs). Membership grades on each input variable (bottom graphs).

the membership functions of "t_{LE} is Early" and "t_{LE} is Late", respectively. The t_{LE} of rower B was so late (large) that the grade of "t_{LE} is Late" was nearly 100%. The t_{LE} of rower A was around 0.35, so the two rules were activated about 55 and 44%, respectively.

The inference error of this model was 0.27 m/sec in rooted mean squared error.

Fuzzy model of efficiency

Another fuzzy model with three fuzzy rules was constructed for the efficiency. Table 4.3 shows the obtained rules. Efficient rowers applied small power of arm pull. Rowers with large arm power and with large leg power were inefficient. The P_{AP} and the P_{LE}

Table 4.3: Fuzzy rules of Efficiency η, and activations of rules for rowers A and B

Antecedent part	Consequent part	Rower A	Rower B
IF P_{AP} is Weak,	THEN $\eta = 0.63$	78	13
IF P_{AP} is Strong AND P_{LE} is Weak,	THEN $\eta = 0.59$	20	14
IF P_{AP} is Strong AND P_{LE} is Strong,	THEN $\eta = 0.51$	2	73

of rower A were Weak as shown by green dots in Figure 4.7. On the other hand, rower B applied Strong P_{AP} and Strong P_{LE}, both were more than 100 W. His efficiency was relatively low.

The inference error of this model was 0.04.

Advice for rowers

Making t_{LE} a little earlier will be effective for rower A to enhance the boat speed. A suggestion not to emphasize on P_{AP} will be useful for better efficiency. For rower B, making t_{LE} earlier will be effective rather than let him try to enlarge P_{LE} and P_{AP}.

4.4 Conclusion

This chapter first analyzed the relationships between the force patterns and the output power. The average power during the driving phase was important to enhance the

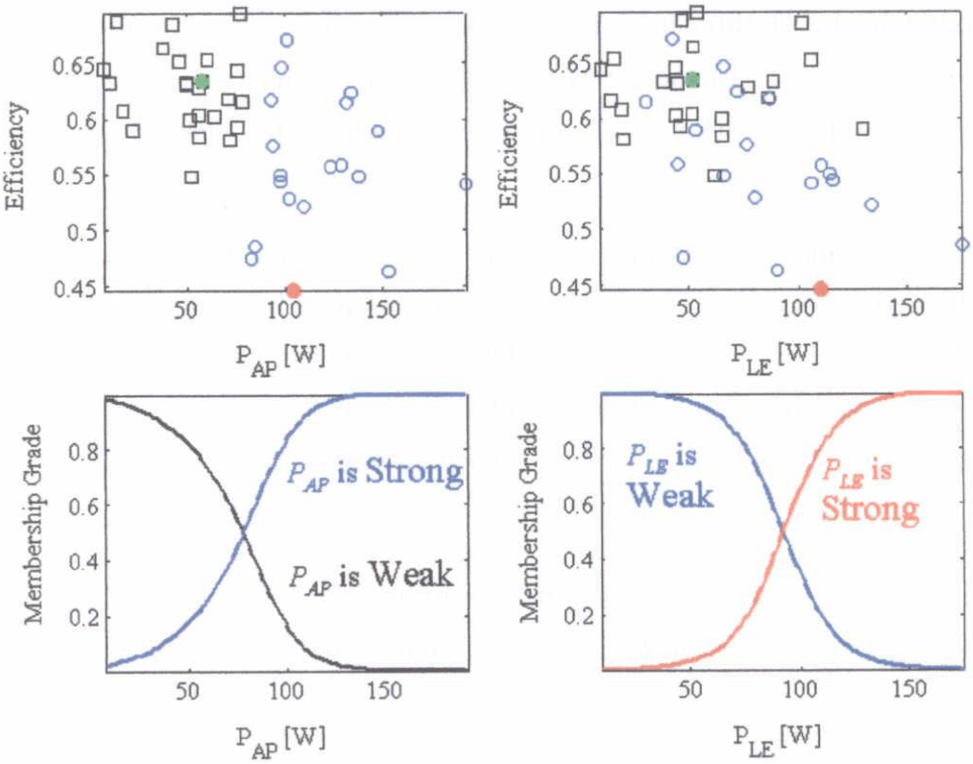


Figure 4.7: Relationships between P_{AP}, P_{LE} and η (top graphs). Membership grades on each input variable (bottom graphs).

boat speed. The average power depended on rower's skill nonlinearly. Fuzzy modeling revealed the nonlinear relationships, and effective advice for rowers in accordance with their performance level were extracted from the obtained model. The maximal force largely affected the output power of women national team rowers. The other factors that affected the output power were in combination with the maximal force.

This chapter then presented a combination of inverse dynamics and fuzzy modeling. The fuzzy modeling clarified that the boat speed was affected by the trunk swing power and by the timing of leg extension. It also clarified that strong arm pulling and strong leg extension decreased the efficiency. Pieces of advice were obtained for each run, referring to the fuzzy rules and their activations.

Chapter 5

Conclusion

5.1 Summary of the Dissertation

Nowadays, scientists and engineers are concerned with systems that are too complex to describe with a basic physical model in many fields. In the absence of the basic model, it is important to estimate input-output dependencies, buried in the data collected from such complex systems, with a statistical method. Fuzzy modeling is one of the statistical methods to estimate input-output dependencies. The most distinctive advantage of fuzzy modeling is its interpretability because each basis function has interpretable discrete expression. The fuzzy inference interpolates smoothly the output of each basis function with those of neighboring basis functions. The discrete expression and the smooth interpolating capability are compatible.

The hierarchical fuzzy modeling proposed by Matsushita *et al.* was effective for strongly nonlinear systems with many input variables. The main issues of the hierarchical fuzzy modeling were input variable selection, input space division. This method searched the combination of input variables and number of membership functions at the same time. However, the even allocation of membership functions on the universe of discourse of each input variable made the number of membership functions larger than needed. The tree type division wa not applicable to this model. The modeler needed trial and error in setting appropriate weights on evaluating functions to obtain a suitable model.

This dissertation proposed new hierarchical fuzzy modeling methods using genetic algorithm in chapter 2. In the proposed method, the combination of input variables was determined by genetic algorithm. Two division methods of input space were presented, one was on the grid-type division and the other was based on the tree-type division. Numerical experiments showed that (1) the proposed grid-type division method obtained more concise models than the conventional methods did and (2) the tree-type division method obtained more concise models than the grid-type division method did.

Multiple-objective genetic algorithm was applied to the hierarchical fuzzy modeling to obtain various models having conciseness and predictive accuracy different from each other. The application of multiple-objective genetic algorithm made the modeler free from the trial and error that had been needed to set the weights on the evaluating functions appropriately. This dissertation proposed two methods of coding for MOGA. The first method encoded only the combination of input variables. The second method encoded the combination of input variables and the allowable number of fuzzy rules. A numerical experiment showed the second method was able to obtain more various models, spread widely on the Pareto front.

Skills of athletes were focused in chapters 3 and 4. Human motions such as cycling and rowing are typical examples of systems that are difficult to be described mathematically with mechanical model. Conventionally, simple models, like center of body mass models, were utilized for analyses of human body motions. Even if attentions had been payed to more complicated models, the interpretation of measured data relied on the intuition of researchers and that of coaches because the motions were hard to describe exactly.

Against the data collected from such a complex system like cycling and rowing motion, the modeling methods from data were effective. Estimation methods of input-output dependencies were not widely used for human motion analyses, except for statistical models and neural networks. Any study that utilized fuzzy models, which had high interpretability, have not been found, in spite of the final goal of the skill analyses of athletes has been the decision making by humans.

Chapter 3 presented application of fuzzy modeling to biomechanics of cyclists. The relationships between the force patterns on the pedal and the output power of cyclists were analyzed. The obtained model clarified the important factors for the pedaling technique. Pieces of advice were able to be extracted from the fuzzy model.

Chapter 4 presented applications of fuzzy modeling to biomechanics of rowers. The hierarchical fuzzy modeling method was utilized to clarify the relationships between the force patterns on the oar and the output power of rowers in section 4.2. The obtained model revealed the nonlinear input-output relationships. The dependencies of rowing performance on partial motions were analyzed in section 4.4. The obtained models explicitly showed the mutual interaction between the input variables. In addition, pieces of advice to improve the performance (output variable) were extractable for each data point, referring to the obtained fuzzy rules and the activations of fuzzy rules.

These applications clearly showed advantages of application of fuzzy modeling to human motion analyses.

5.2 Future Outlook of this Study

This dissertation presented new fuzzy modeling methods and its application to biomechanics. The distinct feature of fuzzy modeling is its interpretability originated from its discrete expression of basis functions that is compatible with the smooth interpolating capability. The hierarchical fuzzy modeling method probably obtains an interpretable model even if the system is very complex.

The human body is hard to describe with a mechanical model. An estimating method of input-output dependencies must be used to analyze the data collected from the human body.

Nowadays, computer simulation enables to visualize complex mechanical models. Computer simulation has also been utilized in biomechanics field.

Figure 5.1 envisions the future research project. The fuzzy modeling and the computer simulation work together. The fuzzy modeling finds out the input-output dependencies from the data. And the feedback can be obtained from the fuzzy model. A new set of parameters of motion, such as joint torque patterns is generated, referring to the feedback. The new set of parameters is inputted to computer simulation that visualizes the motion and calculates the performance of the simulated motion. This combiantion is expected to make a breakthrough for the analysis of biomechanics.

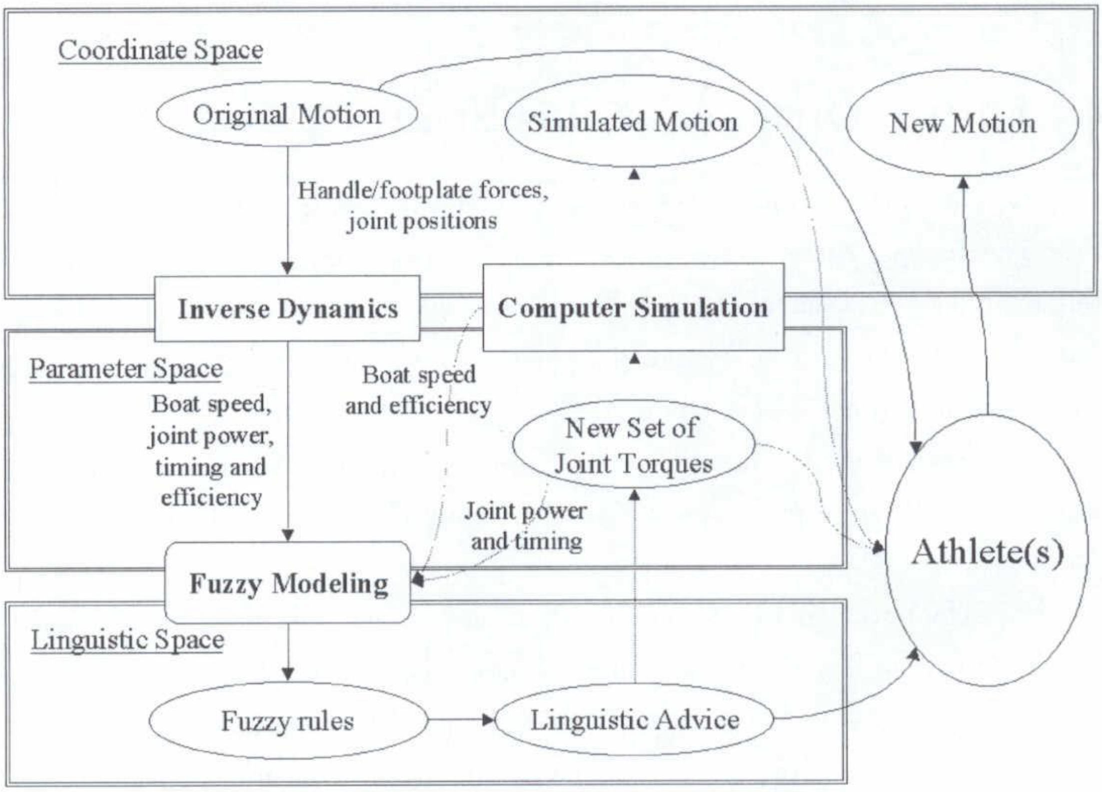


Figure 5.1: Combination of fuzzy modeling and computer simulation

Bibliography

[1] V. Cherkassky, and F. Mulier, "Learning from Data", Wiley Inter-Science (1998)

[2] T. Takagi and M. Sugeno, "Fuzzy Identification of Systems and Its Applications to Modeling and Control", IEEE Transactions on Systems, Man and Cybernetilcs, Vol.SMC-15, pp.116–132 (1985)

[3] M. Mizumoto "Fuzzy Reasoning", Journal of Japan Society of Fuzzy Theory and Systems, Vol.4, No.2, pp.254–256 (1992)

[4] S. Nakayama, T. Furuhashi, Y. Uchikawa, "A Proposal of Hierarchical Fuzzy Modeling Method", Journal of Japan Society of Fuzzy Theory and Systems, vol.5, no.5, pp.1155–1168 (1993).

[5] G. Kang and M. Sugeno, "Fuzzy Modeling", Trans. of the Society of Instrument and Control Engineers, Vol.23, No.6, pp.650–652 (1987).

[6] S. Horikawa, T. Furuhashi, S. Okuma, Y. Uchikawa, "A Fuzzy Controller Using a Neural Network ans its Capability to Learn Expert's Control Rules", Proc. of Int'l Conf. on Fuzzy Logic and Neural Networks (IIZUKA'90), pp.103–106 (1990)

[7] S. Horikawa, T. Furuhashi, Y. Uchikawa, "On Fuzzy Modeling Using Fuzzy Neural Networks with the Back-Propagation Algorithm", IEEE Trans. on Neural Networks Vol.3,5 pp.801–806 (1992)

[8] C.L. Karr, L. Freeman, D. Meredith, "Improved Fuzzy Process Control of Spacecraft Autonomous Rendezvous Using a Genetic Algorithm", SPIE Conf. on Intelligent Control and Adaptive Systems, pp.274–283 (1989)

[9] T. Fukuda, H. Ishigami, T. Shibata, and F. Arai, "Structure Optimization of Fuzzy Neural Network by Genetic Algorithm", Proc. of IFSA'93, pp. 964–967 (1993)

[10] S. Matsushita, A. Kuromiya, M. Yamaoka, T. Furuhashi, Y. Uchikawa, "A Hierarchical Fuzzy Modeling Method Using Multiple Submodels", Journal of Japan Society of Fuzzy Theory and Systems, Vol.8, 2, pp.488–498 (1996)

[11] S. Matsushita,A. Kuromiya, M. Yamaoka, T. Furuhashi and Y. Uchikawa, "Determination of Antecedent Structure for Fuzzy Modeling Using Genetic Algorithm", Proc. of International Conference on Evolutionary Computation (ICEC'96), pp.235–238 (1996)

[12] A.V. Hill, "Heat of Shortening and the Dynamic Constraints of Muscle", Proc. Roy. Soc., B126, pp.136–195 (1938)

[13] J.G. Hay "The Biomechanics of Sports Techniques – Fourth Edition –", Printice Hall, (1993)

[14] S. A. Kautz, M. E. Feltner, E. F. Coyle, A. M. Bayler, "The Pedaling Technique of Elite Endurance Cyclist: Changes with Increasing Workload at Constant Cadence", International Journal of Sport Biomechanics, 7, pp. 29–53 (1991), and http://isb.ri.ccf.org/data/kautz/

[15] A. Millward, "A Study of the Forces Exerted by an Oarsman and the Effect on Boat Speed", Journal of Sports Sciences. (5), 2, pp.93–103 (1987)

[16] J. Janiak, R. Stupnicki, "Static Muscle Force in Athletes Practicing Rowing", Biology of Sport, (10), 1, pp.29–34 (1993)

[17] U. Hartmann, A. Mader, K. Wasser, I. Klauer, "Peak Force, Velocity, and Power During Five and Ten Maximal Rowing Ergometer Strokes by World Class Female and Male Rowers. ", International Journal of Sports Medicine, (14), Suppl. 1, pp.42–45 (1993)

[18] H. Gerber, E. Strussi, "Hydraulic Force and Flow at the Oar-Blade in Rowing a Skiff", Abstracts of the International Society of Biomechanics, XIVth Congress, Paris, 4-8 July, 1993, pp.466–467 (1993)

[19] R.L. Jensen, G.M. Kline, "The Resampling Cross-Validation Technique in Exercise Science: Modelling Rowing Power", Medicine & Science in Sports & Exercise, (26), 7, pp.929-933 (1994)

[20] R.M. Smith, W.L. Spinks, "Discriminant Analysis of Biomechanical Differences Between Novice, Good and Elite Rowers", Journal of Sports Science. 13, (5), pp.377-385 (1995)

[21] K.D. Maier, P. Meier, H. Wagner, R. Blickhan, "Neural Network Modeling in Sport Biomechanics Based on the Example of Shot-Put Flight", Proceedings of XVIII International Symposium on Biomechanics in Sports, pp.550-553 (2000)

[22] A. Hohmann, J. Edelmann-Nusser, and B. Henneberg, "Modeling and Prognosis of Competitive Performances in Elite Swimmig", Proceedings of Swim Sessions, XIX International Symposium on Biomechanics in Sports, pp.100-104 (2001)

[23] S. Araki, H. Nomura, I. Hayashi, N. Wakami, "A Fuzzy Modeling with Iterative Generation Mechanism of Fuzzy Inference Rules", Journal of Japan Society of Fuzzy Theory and Systems, Vol.4, No.4, pp.722-732 (1992)

[24] Y. Ishizuka, H. Miyajima, S. Fukumoto, "Fuzzy Modeling Using a Neural Network Deleting the Rules in Learning" , Proc. of the 9th Fuzzy System Symposium, pp.865-868 (1993)

[25] S. Horikawa, T. Furuhashi, Y. Uchikawa, "Fuzzy Modeling Using a Fuzzy Neural Network (IV)", Proc. of the 10th Fuzzy System Symposium, pp.697-700 (1994)

[26] T. Furuhashi, S. Matsushita, H. Tsutsui, "Evolutionary Fuzzy Modeling Using Fuzzy Neural Networks and Genetic Algorithm", proc. of 1997 IEEE Int'l Conf. on Evolutionary Computation (ICEC'97), pp.623-627 (1997).

[27] J. D. Schaffer, "Multiple Objective Optimization with Vector Evaluated Genetic Algorithms", Proceedings of First International Conference on Genetic Algorithms, pp.93-100 (1985)

[28] C. M. Fonseca and P. J. Fleming, "Genetic Algorithms for Multiobjective Optimization: Formulation, Discussion and Generalization", Proceedings of the Fifth International Conference on Genetic Algorithms, pp.416-423 (1993)

[29] H. Akaike, "Information theory and an extention of the maximum likelihood principle", 2nd Inter. Symp. on Information Theory, pp.267–281 (1973)

[30] W.S. Erdmann, "Diagnostic Procedures in Rowing", Proc. of the 18th International Symposium of Biomechanics in Sports, Vol. II, p.951 (2000)

[31] V. Kleshnev, "Power in Rowing", Proc. of the 18th International Symposium of Biomechanics in Sports, Vol. II, pp.662–666 (2000)

[32] D.P. Buck, R.M. Smith, P.J. Sinclar, "Peak Ergometer Handle and Foot-Stretcher Force on Concept II and Rowperfect Rowing Ergometers", Proc. of the 18th International Symposium of Biomechanics in Sports, Vol. II, pp.622–625 (2000)

[33] D.A. Winter, "Biomechanics and Motor Control of Human Movement — Second Edition —", Wiley Inter-Science (1990)

Acknowledgment

I am heartily grateful to many people who contributed to this dissertation in various ways.

I would like to thank Professor Yasuhito Suenaga, Graduate School of Engineering, Nagoya University, for his valuable comments and encouragements. I would like to thank Professor Masaaki Sugihara and Professor Yukio Kaneda, Graduate School of Engineering, Nagoya University, for valuable suggestions and comments. I am especially indebted to Professor Takeshi Furuhashi, School of Engineering, Mie University, for his continuous advice and encouragements. I would like to acknowledge the late Professor Yoshiki Uchikawa, Associate Professor Akio Ishiguro, Graduate School of Engineering, Nagoya University, and Associate Professor Tetsuji Kodama, School of Engineering, Mie University, for their help.

I would like to thank Professor Tetsuo Fukunaga and Associate Professor Yasuo Kawakami, Graduate School of Arts and Sciences, University of Tokyo, for valuable comments.

I thank my colleagues; Akinobu Fujii, Takeshi Inoue, Toshihiro Suzuki, Arao Funase, Kenji Nakahira and Seiji Tokura, for their help. I also thank all members of the complex system engineering laboratory.

Finally, I would like to thank my family for any kinds of their help, support and encouragement.