Genetically Generated Double-Level Fuzzy Controller with a Fuzzy Adjustment Strategy

Achiche, Sofiane; Wang, Wei; Fan, Zhun; Özkil, Ali Gürcan; Sørensen, Torben; Wang, Jiachuan; Goodman, Erik

Published in:
Genetic and Evolutionary Computation Conference 2007

Publication date:
2007

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Genetically Generated Double-Level Fuzzy Controller with a Fuzzy Adjustment Strategy

Sofiane Achiche
Technical University of Denmark
Nils Koppels Allé, Building 404
2800 Kgs. Lyngby, DK
+45 4525 4166
sac@gst.mek.dtu.dk

Ali Ozkil
Technical University of Denmark
Nils Koppels Allé, Building 404
2800 Kgs. Lyngby, Copenhagen, DK
+45 4525 6271
s053675@student.dtu.dk

Erik Goodman
Michigan State University
East Lansing, 48823
Michigan, USA
+1 517 355 6453
goodman@msu.edu

Wang Wei
Technical University of Denmark
Nils Koppels Allé, Building 404
2800 Kgs. Lyngby, DK
+45 4525 6271
ww@mek.dtu.dk

Zhun Fan
Technical University of Denmark
Nils Koppels Allé, Building 404
2800 Kgs. LyngbyDK
+45 4525 6271
zf@mek.dtu.dk

Torben Sørensen
Technical University of Denmark
Nils Koppels Allé, Building 404
2800 Kgs. Lyngby, Copenhagen, DK
+45 4525 6278
ts@mek.dtu.dk

Jiachuan Wang
Systems Department
United Technologies Research Center
East Hartford, 06128, USA
+1 860 610 7549
WangJ2@utc.com

ABSTRACT
This paper describes the use of a genetic algorithm (GA) in tuning a double-level modular fuzzy logic controller (DLMFLC), which can expand its control working zone to a larger spectrum than a single-level FLC. The first-level FLCs are tuned by a GA so that the input parameters of their membership functions and fuzzy rules are optimized according to their individual working zones. The second-level FLC is then used to adjust contributions of the first-level FLCs to the final output signal of the whole controller, i.e., DLMFLC, so that it can function in a wider spectrum covering all individual working zones of the first-level FLCs. The second-level FLC is again optimized by a GA. An inverted pendulum system (IPS) is used to demonstrate the feasibility of the approach.

Categories and Subject Descriptors
I.2 [Artificial Intelligence]: Problem Solving, Control Methods, and Search – Heuristic methods.

General Terms
Design

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Copyright 2007 ACM  978-1-59593-697-4/07/0007...$5.00.

Keywords
Genetic algorithm, fuzzy logic controller, modularity.

1. INTRODUCTION
Several research and industrial applications concentrated their efforts on providing simple and easy control algorithms to cope with the increasing complexity of the controlled processes/systems [1]. The design method for a controller should enable full flexibility in the modification of the control surface [2]. The systems involved in practice are, in general, complex and time variant, with delays and nonlinearities, and often with poorly defined dynamics. Consequently, conventional control methodologies based on linear system theory have to simplify/linearize the nonlinear systems before they can be used, but without any guarantee of providing good performance. To control nonlinear systems satisfactorily, nonlinear controllers are often developed. The main difficulty in designing nonlinear controllers is the lack of a general structure [3]. In addition, most linear and nonlinear control solutions developed during the last three decades have been based on precise mathematical models of the systems. Most of those systems are difficult/impossible to be described by conventional mathematical relations, hence, these model-based design approaches may not provide satisfactory solutions [4]. This motivates the interest in using FLC; FLCs are based on fuzzy logic theory [5] and employ a mode of approximate reasoning that resembles the decision making process of humans. The behavior of a FLC is easily understood by a human expert, as knowledge is expressed by means of intuitive, linguistic rules. In contrast with traditional linear and nonlinear
control theory, a FLC is not based on a mathematical model and is widely used to solve problems under uncertain and vague environments, with high nonlinearities [6][7]. Since their advent, FLCs have been implemented successfully in a variety of applications such as insurance and robotics [8][9][10]. Unlike neural networks, the first-generation fuzzy systems are not able to learn from data. However, several techniques have been proposed to extract fuzzy rules from training data gathered from observation of the operator control strategy [11][12][13]. As a consequence, an emerging topic in the FLC research community has been to investigate ways to help designers to automatically design FLCs. The majority of the work uses computational intelligence techniques, including neural networks [14][15][16] and genetic algorithms (GA) [17][18][19]. This paper concentrates on the use of a GA. Various approaches have been used to design FLCs using a GA, either concentrating on the fuzzy control rules [20], the membership functions [21], or both [22][23][24]. However, a GA applied to design FLCs has some shortcomings too, mainly because the FLC is built and selected according to the fitness value, which only measures how close the solution is to the desired result under predefined environments, and may therefore decrease its generalization to changing or uncertain environments. FLCs tend to show good control behavior only in ranges near the desired result under predefined environments, and may therefore lose the ability to perform at all in a new environment. 

3. GENETIC ALGORITHM

Genetic algorithms are stochastic optimization techniques based on an analogy to the mechanics of biological genetics and imitate the Darwinian survival-of-the-fittest phenomenon. Each individual of a population is a potential FLC. In this research, FLCs are encoded into a genotype before applying four evolutionary operations: reproduction, mutation, evaluation and natural selection, and finally decoded.

3.1 GA Applied to FLC Generation

The key part of a FLC is the Fuzzy Inference System, and it is composed of fuzzy sets and fuzzy rules [28]. One can use some parameters to describe both of them and encode these parameters into a chromosome. There are many ways to encode an FLC into a GA [29]. In this paper, the encoding is based on the scheme proposed in [30], chosen for its simplicity since only few parameters are used to describe the FLC—namely, the number and center values of the input and output membership functions and the linguistic control rules.

3.1.1 Coding of the Membership Functions

In this paper the number of membership functions (MF) is assumed to be an odd number greater than unity and the MFs are distributed symmetrically around a center of zero. Only triangular fuzzy membership functions are considered for the sake of coding simplicity and without loss of generality. The center value is given by:

\[
\phi_{MF} = \text{sgn}(i).c_m \left| \frac{x}{P_n} \right|^2 \quad \text{where} \quad i = -\frac{n+1}{2}, \ldots, 0, \ldots, \frac{n+1}{2}
\]  

(2)

Where \( n \) and \( P_n \) are the number of MFs and the exponent of the power function and \( \text{sgn}(*) \) denotes the SIGN function:

\[
\text{sgn}(x) = \begin{cases} 
1 & \text{if } x \geq 0 \\
-1 & \text{if } x < 0 
\end{cases}
\]  

(3)

Finally, the characteristic function gives the values of \( c_m \):

\[
c_m = \left( \frac{n+1}{2} \right)^{-P_n}
\]

(4)

3.1.2 Coding of the Fuzzy Rule Base

A FLC is constructed using the following linguistic rules:

\[
R_i = \text{if } x_1 \in X_{i1} \text{ and } x_2 \in X_{i2} \text{ then } u \in U_i, i = 1, \ldots, n
\]

(5)

where \( x_i, x_2 \) and \( u \) are input and output variables in the controller; these are described as state error, change-of-state error and a control input, in this paper.
The rules are coded using a combination of a seed line and seed points, which are positioned into a phase plane containing all the possible fuzzy rules one can fire. The locations of the seed points are determined as follow:

\[
x_{si} = L \sin c_i \frac{p_i}{2}
\]
\[
x_{si} = L \tan(\phi_i)
\]
\[i = \frac{n_o + 1}{2}, \cdots, 0, \cdots, \frac{n_o + 1}{2}
\]

where \(\phi_i\) and \(p_i\) are seed angles and are defined as characteristic parameters for the fuzzy linguistic rules. Also \(C_i\) is obtained the same way as \(C_p\) using the number of MFs on the output \(n_o\).

L limits the locations of the seed points. For more details on the coding, see [30].

In Figure 1, the relationship between input and output in the controller can be represented as a function of these design parameters as follows.

\[
u(t) = f(x_1(t), x_2(t)\Psi)
\]

(7)

Where \(f(\Psi)\) denotes a fuzzy function constructed with the design parameter vector \(\Psi\). The parameters of \(\Psi\) are coded into a chromosome using binary coding.

3.1.3 Genotype of FLC

The characteristic parameters of the FLC are encoded into a chromosome as shown in Figure 2. The GA will evaluate this chromosome using the evaluation function to find the best individual and the FLC defined by it. More details are given in [30].

4. DOUBLE-LEVEL MODULAR FLC

FLCs designed automatically by GAs or other techniques generally exhibit a drop in performances with an increase in the working zone. The working zone (also called the universe of discourse) is defined as the interval between the minimum and maximum values in which the control actions occur. This drop in efficiency is due to the GA’s learning paradigm, which generally selects the desired solution by evaluating the fitness function. A fitness function typically measures how closely a FLC’s output fits the data in working zone used in the training set, which is generally smaller than the overall working zone (smaller interval and/or fewer values or even one single value). The difficulty this causes will be explained further below.

3.2 Evaluation Function

The evaluation function is called by the GA to compute the fitness of a set of parameters. The parameters are passed to the evaluation function, which processes them and returns a value corresponding to how well the controller defined by those parameters performed the task at hand. In this paper, the evaluation function returns the root mean square error between the desired and actual (computed) positions of the pendulum. This function first extracts the relevant parameters from the chromosome passed in. A SIMULINK model is then called, from which a record of the error in the pole angle throughout the duration of the simulation is obtained. The square of the error is multiplied by a time weight and the sum of this time-weighted square error is inverted to give a fitness value. If the pole angle should at any stage of the simulation saturate—i.e., reach ±90°, then the simulation is stopped immediately so that time is not wasted modeling controllers that fail to balance the pole. The MatLab source code for this evaluation function can be found in [33].

4. DOUBLE-LEVEL MODULAR FLC

FLCs designed automatically by GAs or other techniques generally exhibit a drop in performances with an increase in the working zone. The working zone (also called the universe of discourse) is defined as the interval between the minimum and maximum values in which the control actions occur. This drop in efficiency is due to the GA’s learning paradigm, which generally selects the desired solution by evaluating the fitness function. A fitness function typically measures how closely a FLC’s output fits the data in working zone used in the training set, which is generally smaller than the overall working zone (smaller interval and/or fewer values or even one single value). The difficulty this causes will be explained further below.

In Figure 3, \(g_1\) and \(g_2\) denote the minimum and maximum values of the FLC’s working zone, respectively. The horizontal axis is the whole working zone. For the rest of the paper, the interval \(G\) (\(g_1, g_2\)) will refer to the working zone of the FLC and \(ts\) will refer to the training sets. \(G\) and \(ts\) are linked by the following expression:

\[ts \in G \text{ or } ts = G\]

(8)

There is no established formula that can describe explicitly this link. But it is important to note that the performance quality depends on the selection of \(ts\), since it determines the fitness function evaluation.

When the \(ts\) working zone is smaller than \(G\), the near-optimal FLC may not work properly for the whole \(G\), as shown in Figure 4. But increasing the \(ts\) working zone would not only increase learning time, but also causes the GA to get stuck in local optima because of larger/more complex search space, yielding an FLC
that is not optimum, and therefore cannot solve the problem either efficiently or effectively.

One way to solve this problem is to divide \( G \) into smaller working zones, choose a \( ts \) from each (can be a single value), and construct an FLC that works well for each. A method to integrate all FLCs is then needed so that the combination covers the entire working zone \( G \). However, how to divide \( G \) properly so that each smaller working zone can have the right (producing optimum FLC) \( ts \) is still an open problem.

Another similar way is to randomly select values from \( G \) and use them as \( ts \), and then to identify an appropriate working zone for each \( ts \) by testing. With this method, the working zone of the system is controlled by several FLCs. This approach will be referred to as the switch mode in this paper, in which the concatenation of the FLCs’ smaller working zones gives the total working zone of the system. In the switch mode approach, the control switches to a different FLC depending on where the value of the external input is located in \( G \). Figure 5 shows an example in which all the FLCs have the same external input, but as their working zones are different, a switch chooses the FLC to use to control the target (FLC2 output in Figure 6).

However, the switch mode approach can still fail quite frequently. One reason for this is that a smoother transition from one FLC to another is needed to control the dynamic systems successfully. An improvement of the switch mode approach is to use a linear combination of the outputs of the FLCs to control the system; such an approach has generated some good results, as described in [26].

Both approaches described above are difficult to realize properly, since the working zone must be divided into several smaller ones, and the learning by GA must be done for each, which is time consuming. The question would be: is it possible to reduce the number of smaller working zones and/or choose a finite and limited number of values from the system’s working zone as \( ts \) in order to get an FLC that works well across its entire working zone? To achieve this, a double-level modular FLC (DLMFLC) is proposed in this paper.

A DLMFLC is a further improvement of the method presented in [26]. The basic idea is to use a second-level FLC to adjust the contribution of the first-level FLCs to the output of the integrated controller. The first-level FLCs would be the modules of the control system. These FLCs are called modular FLCs; they can work in their smaller working zones but not necessarily well in the whole system’s work zone. Having several FLCs, how can we integrate them together into a single FLC? How can this second level be constructed? For each input that falls into the working zone of a first-level FLC, the output of this particular FLC module should contribute more to the output of the overall controller. However, this does not necessarily mean that other FLC modules should not contribute at all to the output of the overall controller. They will simply contribute to a lesser extent. The key here is to decide proper ‘weights’ of all FLC modules so that their overall contributions can control the system successfully across the entire working zone.

The method in [26] utilizes a linear combination of the outputs of the FLC modules, in which the ‘weights’ of the FLC modules are fixed for all inputs. In this paper, a second-level FLC is used to integrate the modular FLCs so that the ‘weights’ are adjusted by the second-level FLC according to different inputs. The basic scheme is presented in Figure 7.

The existing FLCs constitute the first-level FLCs, which are FLC1, FLC2, ..., FLCn in Figure 7. A second-level FLC, which resides above the first-level FLCs (Figure 7), is used to integrate outputs of the first-level FLCs.
\[ \text{OUT} = \sum_{i=1}^{n} f_i \times \text{out}_i \]  \hspace{1cm} (11)

5. Case Study
In order to test and validate the control approach proposed in Section 4, the Inverted Pendulum system (IPS) control problem is used, since it is considered a typical testbed for control experiments.

5.1 Inverted Pendulum System (IPS)
As we chose the IPS as our testbed, it is necessary to introduce some basic variables and components.

The variables shown in Figure 8 and 9 are as follows:
- M, the mass of the cart,
- m, the mass of the pole,
- l, the distance from the pivot to the center of mass of the pole,
- I, the moment of inertia of the pole around the pivot,
- x, the cart’s position,
- \( \theta \), the angle the center of the pole makes with the vertical,
- F, the horizontal control force applied to the cart,
- g, gravity.

![Figure 8: IPS Model.](image)

![Figure 9: IPS--Pole in Isolation](image)

Taking the second derivative of the position of the pole’s center of gravity, one obtains:
\[ \ddot{x}_G = \ddot{x} - l \dot{\theta} \cos \theta + l \dot{\theta}^2 \sin \theta \]  \hspace{1cm} (12)

Newton’s second law of motion yields:
\[ F = M \ddot{x} + m \ddot{x}_G \]  \hspace{1cm} (13)

Combining equations 12 and 13 yields:
\[ F = \ddot{x}(M + m) + ml \dot{\theta}^2 \sin \theta - ml \dot{\theta} \cos \theta \]  \hspace{1cm} (14)

Considering that, for a rigid body, the sum of the moments about a fixed point, P, is equal to the moment of inertia of the body about P multiplied by the angular acceleration plus the product of the mass of the body, its linear acceleration and the perpendicular distance between point P and the vector representing the acceleration yields:
\[ M_p = I_p \dot{\alpha} + m \ddot{a} \]  \hspace{1cm} (15)

Recombination, insertion and rearrangement of these equations yields (for a uniform rod of mass m and length L=2\times1):
\[ \ddot{x} = \frac{4}{3} \frac{F + mg \cos \theta \sin \theta - \frac{4}{3} l \dot{\theta}^2 \sin \theta}{(M + m) - m \cos^2 \theta} \]  \hspace{1cm} (16)

\[ \ddot{\theta} = \frac{(M + m)g \sin \theta + \cos \theta F - ml \dot{\theta}^2 \cos \theta \sin \theta}{\frac{4}{3} (M + m)l - ml \cos^2 \theta} \]  \hspace{1cm} (17)

More details on the inverted-pendulum can be found in [33].

5.2 SIMULINK model of the IPS
Using the equations shown in Section 5.1, the SIMULINK model shown in Figure 10 is developed. In the model, the input is the control action. The angle and angular velocity from the previous step are combined to compute the cart’s acceleration and the pole’s angular acceleration, respectively. The speed and position of the cart are gotten by integrating the angle and the angular velocity. The pole’s angle is also limited to remaining in the range of \([-\pi/2, \pi/2]\).

![Figure 10: SIMULINK model of the IPS](image)

For the sake of uniformity of the parameters during the test, the values of three variables are fixed:
- M=1kg
- m=0.1kg
- l=0.5m

Gravity is fixed at g = 9.8m/s².

5.3 Construction of the first-level FLCs
The number of membership functions (MFs) is arbitrarily restricted to the integers from 3 to 9, inclusively, as most FLCs are within this range. The number of MFs is limited to odd integers because of the way the FLC is encoded in the GA. In the tests reported in this paper, we selected the number of MFs to be 5.
5.3.1 Running the GA
The parameters for running the GA to evolve the FLCs in both second and first levels are:
- Population size: 50
- Selection mechanism: Stoc.Rem (as in Stochastic Remainder Sampling, a MatLab script added to the GA tool box of MatLab [33])
- Crossover: single-point crossover at 0.5
- Mutation: single-bit, at 0.05/bit

Three different disturbances—250N, 500N and 1000N—were used to create the training sets for evolving the first-level FLCs. Then the working zone for each evolved FLC module was obtained by testing the FLC against different external disturbances ranging from 0N to 1200N, with a 1N increment.

5.4 Second-level FLC
Figure 12 shows the basic scheme of the DLMFLC. The new blocks on the left are the modular FLC, also called the second-level FLC. There are some differences in the process of getting the second-level FLC, as compared to finding the first-level ones. The disturbance is different—it lasts 0.01s—but the same three disturbances, which are 250N, 500N and 1000N, are used, and the simulation starts from 0s, 2s and 4s, respectively. The FLC’s structure is also different; while the first-level FLC’s inputs are error and change ratio of error and the output is a force to reach (control) the target (the pole of IPS), the second-level FLC is a one-input, three-output system. The input is the value of the disturbance and the output is the gain (a real value with no actual physical meaning) for every modular FLC (3 outputs).

The parameters for encoding the second-level FLC into the GA are therefore also different. The number of MFs is set to 3 and the output scaling is 0-1 (gain). The parameters for running the GA are the same as those used to evolve the modular FLCs.

In order to test the repeatability of the experiment, the whole process was repeated 5 times. A different modular FLC was selected from each of the 5 experiments. Furthermore, an additional challenge was imposed on the second-level FLC: the sum of the outputs of the overall system should not exceed 500N, which is a way of reproducing the torque limit on a motor that would stabilize the inverted pendulum; however, the modular FLCs’ outputs have no such limitations (i.e., the limitation is imposed by the gains evolved for the second-level FLC).

5.5 Results
In order to test the efficiency of using a second-level FLC, it is compared to the results obtained by training the FLC with the first level only (gains/parameters of the first level FLC are decided by GA).

Figure 14 and 15 show the working zones for each FLC. A reading of ‘1’ on the y-axis means that for the disturbance force shown on the x-axis, the FLC can control and stabilize the IPS successfully. A reading of ‘0’ indicates otherwise.
It is noteworthy that the FLC obtained for the 250N disturbance has a very limited control range, and examination of the ones obtained for 500N and 1000N shows that there are many intervals where the FLC cannot control the IPS. Hence, using a first-level FLC alone does not give satisfactory results when testing the controller with different values than the ones used for the training. After the second-level FLC is linked to the first-level ones, comprising the DLMFLC, one can see in Figure 16 that the controller is able to work properly for the whole range of disturbance forces.

Figure 17 shows the necessary time to stabilize the inverted pendulum using a maximal force of 500N, while disturbing the system with different forces; one can conclude that the response time is satisfactory in all cases studied.

6. CONCLUSION

In order to get a FLC that performs efficiently across the whole working zone of a controlled system using a genetic algorithm (GA), a double-level modular FLC (DLMFLC) was proposed and developed successfully in this paper. The approach was tested with an IPS controller design problem.

The simulation results show that the DLMFLC obtained works well across the entire desired working zone and needs only small training sets during the search process of the GA.

More issues must still to be investigated for the DLMFLC. For example, one should investigate the relationship between first-level FLCs and the second-level FLC, so as to guide how to choose the training sets for the first-level FLCs that make it easier for the second-level FLC to evolve a better overall controller—i.e., DLMFLC.

In the test presented in this paper, the second-level FLC is a SIMO system. In this particular case study of IPS controller design, external disturbance is the only input for the second-level FLC. But in practice, several external parameters can influence a system. Should one add more variables as inputs to the second-level FLC, and how? All these questions merit further study to make the DLMFLC usable in more general applications.

7. ACKNOWLEDGMENTS

Financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC) under Post-Doctoral Grant BP-328508-2006 is gratefully acknowledged.

REFERENCES


