INTERNATIONAL JOURNAL

of

COMPUTERS, COMMUNICATIONS & CONTROL

With Emphasis on the Integration of Three Technologies

IJCCC

Year: 2011 Volume: 6 Number: 3 (September)

Special Issue on

Fuzzy Sets and Systems

Dedicated to the 90th Birthday of Prof. Lotfi A. Zadeh

Guest Editors: Valentina Emilia Balas, Ioan Dzitac and Horia-Nicolai Teodorescu

Agora University Editing House

CCC Publications

www.journal.univagora.ro

International Journal of Computers, Communications & Control

EDITOR IN CHIEF:
Florin-Gheorghe Filip
Member of the Romanian Academy
Romanian Academy, 125, Calea Victoriei
010071 Bucharest-1, Romania, ffilip@acad.ro

ASSOCIATE EDITOR IN CHIEF:
Ioan Dzitac
Aurel Vlaicu University of Arad, Romania
Elena Dragoi, 2, Room 81, 310330 Arad, Romania
ioan.dzitac@uav.ro

MANAGING EDITOR:
Mişu-Jan Manolescu
Agora University, Romania
Piata Tineretului, 8, 410526 Oradea, Romania
rectorat@univagora.ro

EXECUTIVE EDITOR:
Răzvan Andonie
Central Washington University, USA
400 East University Way, Ellensburg, WA 98926, USA
andonie@cwu.edu

TECHNICAL SECRETARY:

Cristian Dziţac
R & D Agora, Romania
rd.agora@univagora.ro

Emma Margareta Văleanu
R & D Agora, Romania
evaleanu@univagora.ro

EDITORIAL ADDRESS:
R&D Agora Ltd. / S.C. Cercetare Dezvoltare Agora S.R.L.
Piaţa Tineretului 8, Oradea, jud. Bihor, Romania, Zip Code 410526
Tel./ Fax: +40 359101032
E-mail: ijccc@univagora.ro, rd.agora@univagora.ro, ccc.journal@gmail.com
Journal website: www.journal.univagora.ro

DATA FOR SUBSCRIBERS
Supplier: Cercetare Dezvoltare Agora Srl (Research & Development Agora Ltd.)
Fiscal code: RO24747462
Headquarter: Oradea, Piata Tineretului Nr.8, Bihor, Romania, Zip code 410526
Bank: MILLENNIUM BANK, Bank address: Piata Unirii, str. Primariei, 2, Oradea, Romania
IBAN Account for EURO: RO73MILB0000000000932235
SWIFT CODE (eq.BIC): MILBROBU

International Journal of Computers, Communications & Control

EDITORIAL BOARD

Boldur E. Bărbat
Lucian Blaga University of Sibiu
Faculty of Engineering, Department of Research
5-7 Ion Raţiu St., 550012, Sibiu, Romania
bbarbat@gmail.com

Pierre Borne
Ecole Centrale de Lille
Cité Scientifique-BP 48
Villeneuve d'Ascq Cedex, F 59651, France
p.borne@ec-lille.fr

Ioan Buciu
University of Oradea
Universitatii, 1, Oradea, Romania
ibuciu@uoradea.ro

Hariton-Nicolae Costin
Faculty of Medical Bioengineering
Univ. of Medicine and Pharmacy, Iaşi
St. Universitatii No.16, 6600 Iaşi, Romania
hcostin@iit.tuiasi.ro

Petre Dini
Cisco
170 West Tasman Drive
San Jose, CA 95134, USA
pdini@cisco.com

Antonio Di Nola
Dept. of Mathematics and Information Sciences
Università degli Studi di Salerno
Salerno, Via Ponte Don Melillo 84084 Fisciano, Italy
dinola@cds.unina.it

Ömer Egecioglu
Department of Computer Science
University of California
Santa Barbara, CA 93106-5110, U.S.A
omer@cs.ucsb.edu

Constantin Gaindric
Institute of Mathematics of
Moldavian Academy of Sciences
Kishinev, 277028, Academiei 5, Moldova
gaindric@math.md

Xiao-Shan Gao
Academy of Mathematics and System Sciences
Academia Sinica
Beijing 100080, China
xgao@mmrc.iss.ac.cn

Kaoru Hirota
Hirota Lab. Dept. C.I. & S.S.
Tokyo Institute of Technology
G3-49,4259 Nagatsuta,Midori-ku,226-8502,Japan
hirota@hrt.dis.titech.ac.jp

George Metakides
University of Patras
University Campus
Patras 26 504, Greece
george@metakides.net

Ştefan I. Nitchi
Department of Economic Informatics
Babes Bolyai University, Cluj-Napoca, Romania
St. T. Mihali, Nr. 58-60, 400591, Cluj-Napoca
nitchi@econ.ubbcluj.ro

Shimon Y. Nof
School of Industrial Engineering
Purdue University
Grissom Hall, West Lafayette, IN 47907, U.S.A.
nof@purdue.edu

Stephan Olariu
Department of Computer Science
Old Dominion University
Norfolk, VA 23529-0162, U.S.A.
olariu@cs.odu.edu

Horea Oros
Dept. of Mathematics and Computer Science
University of Oradea, Romania
St. Universitatii 1, 410087, Oradea, Romania
horos@uoradea.ro

Gheorghe Păun
Institute of Mathematics
of the Romanian Academy
Bucharest, PO Box 1-764, 70700, Romania
gpaun@us.es

Mario de J. Pérez Jiménez
Dept. of CS and Artificial Intelligence
University of Seville, Sevilla,
Avda. Reina Mercedes s/n, 41012, Spain
marper@us.es

Dana Petcu
Computer Science Department
Western University of Timisoara
V.Parvan 4, 300223 Timisoara, Romania
petcu@info.uvt.ro

Radu Popescu-Zeletin
Fraunhofer Institute for Open
Communication Systems
Technical University Berlin, Germany
rpz@cs.tu-berlin.de

Imre J. Rudas
Institute of Intelligent Engineering Systems
Budapest Tech
Budapest, Bécsi út 96/B, H-1034, Hungary
rudas@bmf.hu

Yong Shi
Research Center on Fictitious Economy
& Data Science
Chinese Academy of Sciences
Beijing 100190, China
yshi@gucas.ac.cn
and
College of Information Science & Technology
University of Nebraska at Omaha
Omaha, NE 68182, USA
yshi@unomaha.edu

Athanasios D. Styliadis
Alexander Institute of Technology
Agiou Panteleimona 24, 551 33
Thessaloniki, Greece
styl@it.teithe.gr

Gheorghe Tecuci
Learning Agents Center
George Mason University, USA
University Drive 4440, Fairfax VA 22030-4444
tecuci@gmu.edu

Horia-Nicolai Teodorescu
Faculty of Electronics and Telecommunications
Technical University "Gh. Asachi" Iasi
Iasi, Bd. Carol I 11, 700506, Romania
hteodor@etc.tuiasi.ro

Dan Tufiş
Research Institute for Artificial Intelligence
of the Romanian Academy
Bucharest, "13 Septembrie" 13, 050711, Romania
tufis@racai.ro

Lotfi A. Zadeh
Professor,
Graduate School,
Director,
Berkeley Initiative in Soft Computing (BISC)
Computer Science Division
Department of Electrical Engineering
& Computer Sciences
University of California Berkeley,
Berkeley, CA 94720-1776, USA
zadeh@eecs.berkeley.edu

Guest Editors of this Special Issue:

Valentina Emilia Balas
Aurel Vlaicu University of Arad, Romania
Elena Dragoi, 2, 310330 Arad, Romania
balas@drbalas.ro

Ioan Dzitac
Aurel Vlaicu University of Arad, Romania
Elena Dragoi, 2, 310330 Arad, Romania
ioan.dzitac@uav.ro

Horia-Nicolai Teodorescu
Technical University "Gh. Asachi" Iasi
Iasi, Bd. Carol I 11, 700506, Romania
hteodor@etc.tuiasi.ro

International Journal of Computers, Communications & Control

Short Description of IJCCC

Title of journal: International Journal of Computers, Communications & Control
Acronym: IJCCC
International Standard Serial Number: ISSN 1841-9836, E-ISSN 1841-9844
Publisher: CCC Publications - Agora University
Starting year of IJCCC: 2006
Founders of IJCCC: Ioan Dzitac, Florin Gheorghe Filip and Mişu-Jan Manolescu
Logo:

Number of issues/year: IJCCC has 4 issues/odd year (March, June, September, December) and 5 issues/even year (March, September, June, November, December). Every even year IJCCC will publish a supplementary issue with selected papers from the International Conference on Computers, Communications and Control.

Coverage:

- Beginning with Vol. 1 (2006), Supplementary issue: S, IJCCC is covered by Thomson Reuters - SCI Expanded and is indexed in ISI Web of Science.
- Journal Citation Reports/Science Edition 2010:
  - Impact factor = 0.650
- Beginning with Vol. 2 (2007), No.1, IJCCC is covered in EBSCO.
- Beginning with Vol. 3 (2008), No.1, IJCCC, is covered in Scopus.

Scope: IJCCC is directed to the international communities of scientific researchers in universities, research units and industry. IJCCC publishes original and recent scientific contributions in the following fields: Computing & Computational Mathematics; Information Technology & Communications; Computer-based Control.

Unique features distinguishing IJCCC: To differentiate from other similar journals, the editorial policy of IJCCC encourages especially the publishing of scientific papers that focus on the convergence of the 3 "C" (Computing, Communication, Control).

Policy: The articles submitted to IJCCC must be original and previously unpublished in other journals. The submissions will be revised independently by at least two reviewers and will be published only after completion of the editorial workflow.

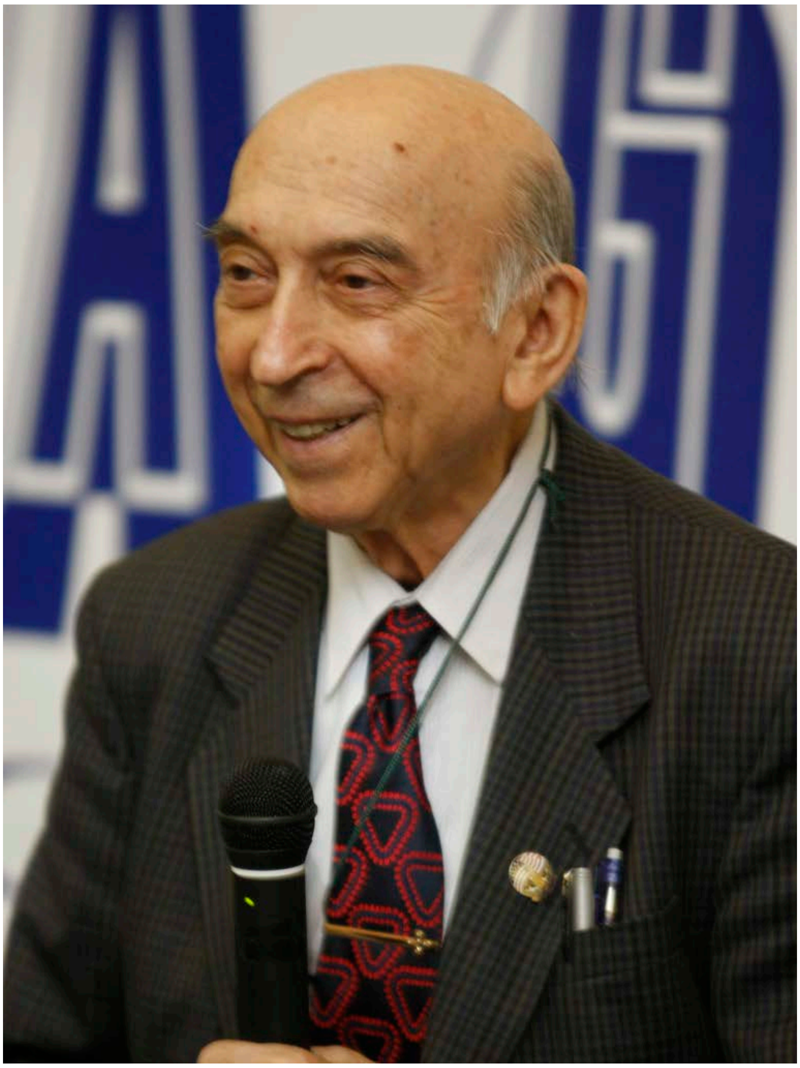
Copyright © 2006-2011 by CCC Publications

International Journal of Computers, Communications & Control

Special Issue on

Fuzzy Sets and Systems

Dedicated to the 90th Birthday of Prof. Lotfi A. Zadeh



Lotfi A. Zadeh

(B. February 4, 1921)

Contents

Int. J. of Computers, Communications & Control, ISSN 1841-9836, E-ISSN 1841-9844
Vol. VI (2011), No. 3 (September), pp. 385-386

Foreword

L.A. Zadeh

Lotfi A. Zadeh
Professor in the Graduate School
Director, Berkeley Initiative in Soft Computing (BISC)
729 Soda Hall 1776 Computer Science Division
Department of Electrical Engineering and Computer Sciences
University of California Berkeley, CA 94720-1776
E-mail: zadeh@eecs.berkeley.edu
http://www.cs.berkeley.edu/ zadeh/

I feel honored by the dedication of the Special Issue of IJCCC to me. I should like to express my deep appreciation to the distinguished Co-Editors and my good friends, Professors Balas, Dzitac and Teodorescu, and to distinguished contributors, for honoring me. The subjects which are addressed in the Special Issue are on the frontiers of fuzzy logic.

The Foreword gives me an opportunity to share with the readers of the Journal my recent thoughts regarding a subject which I have been pondering about for many years - fuzzy logic and natural languages. The first step toward linking fuzzy logic and natural languages was my 1973 paper," Outline of a New Approach to the Analysis of Complex Systems and Decision Processes." Two key concepts were introduced in that paper. First, the concept of a linguistic variable - a variable which takes words as values; and second, the concept of a fuzzy if- then rule - a rule in which the antecedent and consequent involve linguistic variables. Today, close to forty years later, these concepts are widely used in most applications of fuzzy logic.

The second step was my 1978 paper, "PRUF - a Meaning Representation Language for Natural Languages." This paper laid the foundation for a series of papers in the eighties in which a fairly complete theory of fuzzy - logic-based semantics of natural languages was developed. My theory did not attract many followers either within the fuzzy logic community or within the linguistics and philosophy of languages communities. There is a reason. The fuzzy logic community is largely a community of engineers, computer scientists and mathematicians - a community which has always shied away from semantics of natural languages. Symmetrically, the linguistics and philosophy of languages communities have shied away from fuzzy logic.

In the early nineties, a thought that began to crystallize in my mind was that in most of the applications of fuzzy logic linguistic concepts play an important, if not very visible role. It is this thought that motivated the concept of Computing with Words (CW or CWW), introduced in my 1996 paper "Fuzzy Logic = Computing with Words." In essence, Computing with Words is a system of computation in which the objects of computation are words, phrases and propositions drawn from a natural language. The same can be said about Natural Language Processing (NLP.) In fact, CW and NLP have little in common and have altogether different agendas.

In large measure, CW is concerned with solution of computational problems which are stated in a natural language. Simple example. Given: Probably John is tall. What is the probability that John is short? What is the probability that John is very short? What is the probability that John is not very tall? A less simple example. Given: Usually Robert leaves office at about 5 pm. Typically it takes Robert about an hour to get home from work. What is the probability that Robert is home at 6:l5 pm.? What should be noted is that CW is the only system of computation which has the capability to deal with problems of this kind. The problem-solving capability of CW rests on two key ideas. First, employment of so-called restriction-based semantics (RS) for translation of a natural language into a mathematical language in which the concept of a

Copyright © 2006-2011 by CCC Publications

restriction plays a pivotal role; and second, employment of a calculus of restrictions - a calculus which is centered on the Extension Principle of fuzzy logic.

What is thought-provoking is that neither traditional mathematics nor standard probability theory has the capability to deal with computational problems which are stated in a natural language. Not having this capability, it is traditional to dismiss such problems as ill-posed. In this perspective, perhaps the most remarkable contribution of CW is that it opens the door to empowering of mathematics with a fascinating capability - the capability to construct mathematical solutions of computational problems which are stated in a natural language. The basic importance of this capability derives from the fact that much of human knowledge, and especially world knowledge, is described in natural language.

In conclusion, only recently did I begin to realize that the formalism of CW suggests a new and challenging direction in mathematics - mathematical solution of computational problems which are stated in a natural language. For mathematics, this is an unexplored territory.

Lotfi A. Zadeh
Berkeley, CA
August 13, 2011

Int. J. of Computers, Communications & Control, ISSN 1841-9836, E-ISSN 1841-9844
Vol. VI (2011), No. 3 (September), pp. 387-402

Reconfigurable Takagi-Sugeno Fuzzy Logic Control for a Class of Nonlinear System considering Communication Time Delays on Peripheral Elements

H. Benítez-Pérez, F. Cárdenas-Flores, F. García-Nocetti

Hector Benítez-Pérez, F. Cárdenas-Flores, Fabian García-Nocetti
Universidad Nacional Autónoma de México
Departamento de Ingeniería de Sistemas Computacionales y Automatización
Apdo. Postal 20-726., Admón. No. 20
Del. A. Obregón, México D. F., CP. 01000, México.
E-mail: hector@uxdea4.iimas.unam.mx

Abstract: Nowadays the study of faults and their consequences becomes an issue into highly safety critical computer network systems. How to bound the effects of a fault and how to tackle them into a dynamic system is still an open field. In here an approach to tackle this problem is presented. The use of Takagi-Sugeno fuzzy logic structure is given in order to accomplish two challenges, the presence of local faults and the respective time delays within a real-time distributed system. This approach is pursued as reconfigurable strategy according to communication time delays within a real-time distributed system. This approach is pursued as reconfigurable strategy according to communication delays.
Keywords: fuzzy logic, network control.

1 Introduction

The emergence of smart sensor and actuator technology removes the need for centralized control with feedback loops to dumb peripheral actuators replacing it with a databus connection [1]. This gives an autonomous actuator installation [2] as well as local control, self-calibration, health monitoring and reconfiguration availabilities. Several strategies for managing time delay within control laws have been studied for different research groups. For instance [3] proposes the use of a time delay scheme integrated to a reconfigurable control strategy based upon a stochastic methodology. On the other hand, [4] proposes a reconfiguration strategy based upon a performance measure from a parameter estimation fault diagnosis procedure. Another strategy has been proposed by [5] where time delays are used as uncertainties, which modify pole placement of a robust control law. In [6] presents an interesting view of fault tolerant control approach related to time delay coupling. Reconfigurable control has been studied from the point of view of structural modification since fault appearance as presented by [7] it performs a combined modification of system structure and dynamical systems as studied by [8–10]. Present approach takes time delays due to communication as deterministic measured variables. As well as actuator fault presence by modification of B matrix in order to propose a Takagi-Sugeno fuzzy control with two conditions loose of local peripheral elements and the related time delays. In here fuzzy logic control (FLC) law [11] views time delays as a result of deterministic reconfigurable communications based upon scheduling algorithm. These time delays are a structural consequence determined by the insertion of new elements within communication channels due to fault appearance. In fact, fault presence is taken into account as the lost of the related peripheral element, specifically, sensor or actuator elements. The result of fault effects over the peripheral element is the lost of the related measurement, in this case, the structure of the plant is modified in terms of the lost of specific peripheral elements. As an obvious consequence control structure is modified.

Copyright © 2006-2011 by CCC Publications

Fuzzy logic control engulf both undesirable situations by defining the time delays during fault appearance as well as the related lost of the faulty peripheral element. The goal of this paper is to define a strategy for control reconfiguration based upon communication system reconfiguration. Some considerations need to be stated in order to define this approach. First, faults are strictly local in peripheral elements, faults are tackled by just eliminating the faulty element. In fact, faults are catastrophic and local. Time delays are bounded and restrictive to scheduling algorithms. Global stability can be reached using Takagi-Sugeno fuzzy logic approach. Fuzzy logic has been a very powerfull tool to systematic modelling an design complex nolinear systems as presented by Zadeh in [23–26].

This paper is placed as a strategy for reconfigurable systems as shown in Figure 1. In fact, this paper is focused into reconfigurable control law due to the presence of local faults and time delays as consequences. Time delays are measurable and bounded according a real-time scheduling algorithm. In this case scheduling algorithm is the well known earliest deadline first algorithm (EDF). According to Figure 1, structural reconfiguration takes place as result of EDF performance and related user request. This action provokes control law modification. How this is modified is the scope of this paper by using Takagi-Sugeno approach.

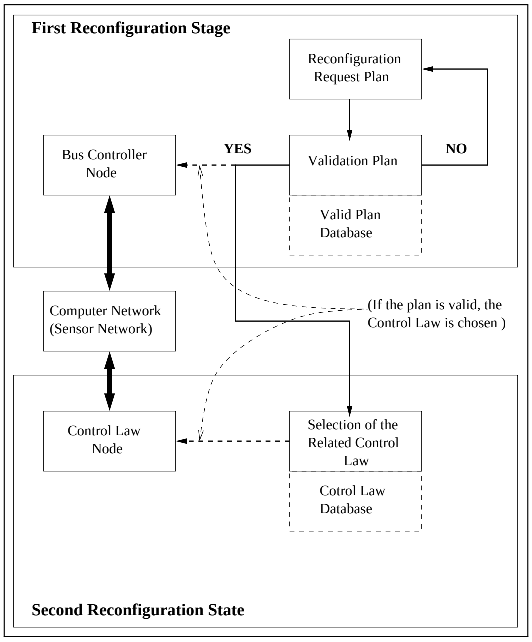


Figure 1: General structure of Reconfigurable System over a Computer Network

The map of this paper is next, first section is current introduction, second section is scheduling approach, third section is plant design, fourth section is control approach fifth section presents a case study and finally some concluding remarks are given.

2 Scheduling aproach

The communication network plays a key role in order to define the behaviour of the dynamic system in terms of time variance from communications and processing although it presents a nonlinear behaviour. In order to understand such a nonlinear behaviour, time delays are incorporated by the use of real-time system theory that allows time delays to be bounded even in the case of causal modifications due to external effects. Several algorithms can be pursued such as

Rate Monotonic (RM), Deadline Monotonic or Earliest Deadline First (EDF) [15–17]. The use of last algorithm is pursued in here due to flexibility of task reorganisation during online performance. requires several characteristics from each task such as deadlines, consumption times and priorities. For instance, consider three tasks with next characteristics (Table 1 and Figure 2) first section under EDF algorithm if a task changes its deadline at Δt it would have a higher priority than those tasks already defined (Table 2).

Table 1: Task used to exemplified EDF algorithm.

Task number	Consumption Time (C)	Periodic Time (P)	Deadline (D)	Priority
1	C_1	P_1	D_1	P_{r_2}
2	C_2	P_2	D_2	P_{r_3}
3	C_3	P_3	D_3	P_{r_1}

From Figure 2 there are two scenarios, in the first task 3 has the smallest slack time (t_{s3}) therefore it has the highest priority P_{r_1}, thereafter, task 1 has next priority (P_{r_2}) and task 2 has the lowest priority P_{r_3} (Table 1).

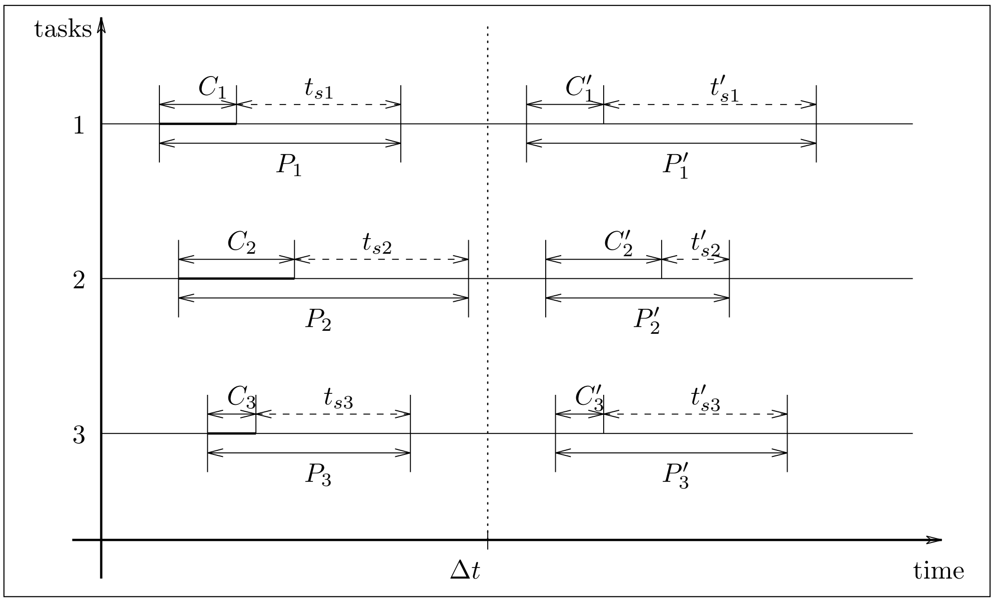


Figure 2: Time Graph Related to Table 1.

Second scenario presents a different priority conformation according to slack times modifications. Thats it for the case of deadline modification as display in Figure 2 priorities are modified as shown in Table 2 where task 2 has the smallest slack time (t_{s2}) therefore it has the highest priority P_{r_1}, task 3 has next priority and the task 1 has the lowest priority P_{r_3}. Therefore the task that is going to be executed first is the one who has the shortest deadline in comparison to the rest of alive tasks in a particular time. Since this algorithm is performed, two conditions are stated upon this approach. Control law must be executed after all sensors have performed their task and before any actuator takes an action over the plant, from these restriction system performance has two main scenarios fault present and fault free situations. These scenarios are

Table 2: New priority order after at reorganisation.

Task number	Consumption Time (C)	Periodic Time (P)	Deadline (D)	Priority
1	C_1	P_1'	D_1'	P_{r3}
2	C_2	P_2'	D_2'	P_{r_1}
3	C_3	P_3'	D_3'	P_{r_2}

exposed in Figure 3, it shows a common configuration considering local time delays.

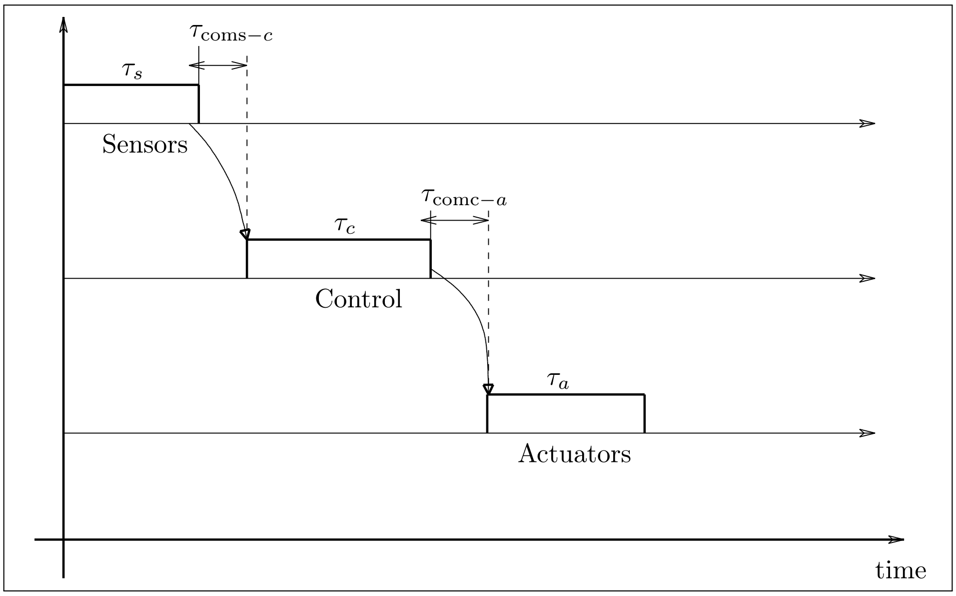


Figure 3: Related time delays from control configuration.

Related time delays are defined in terms of scheduling algorithm behaviour. Therefore time delays are stated as:

delay sensor to controller $\tau_{sc} = \tau_s + \tau_{cons-c}$

delay of the controller τ_c

delay controller actuator $\tau_{ca} = \tau_a + \tau_{comc-a}$

where τ_{coms-c} and τ_{comc-a} are the source of variations due to communication variation between elements.

3 Plant Approach

The proposed dynamic plant is based upon the following structure:

$$\begin{aligned} x(k+1) &= a^p x(k) + B^p u(k) \\ y &= c^p x(k), \end{aligned} \tag{1}$$

where $a^p \in \Re^{n \times n}$, $c^p \in \Re^{n \times 1}$ and $B^p \in \Re^{n \times 1}$ are matrices related to the plant. $x(k)$, $u(k)$ and $y(k)$ are the states, inputs and outputs respectively. Specially B^p is stated as:

$$B^p = \sum_{i=1}^{N} \rho_i B_i \sum_{j=1}^{M} \int_{\tau_j^i}^{\tau_{j-1}^i} e^{-a^p(t-\tau)} d\tau, \tag{2}$$

where $\rho_i \in \{0,1\}$ and $\sum_{i=1}^{N} \rho_i = 1$ taking into account that N are the total number of possible faults and M are the involved time delays from each fault. Current communication time delays are expressed as t_{j-1}^i and t_j^i. Considering that $\sum_{j=1}^{M} \tau_j^i \leq T$, where T is the sampling period and depends of the faults scenarios, then B_i is integrated as:

$$B_i = \begin{bmatrix} b_1 \\ b_2 \\ 0_i \\ \vdots \end{bmatrix} \rightarrow, i \text{ fault element},$$

where $b_i \rightarrow b_n$ are the elements conformed at the input of the plant (such as actuators) and 0_i is the lost element due to local fault where B^p represents only one scenario following Equation 2. Therefore B_i^p defined as:

$$B_i^p = B_i \sum_{j=1}^{M} \int_{t_j^i}^{t_{j-1}^i} e^{a^p(t-\tau)} d\tau, \tag{3}$$

considers local faults and related time delays. For simplicity purposes B_i^p is used in order to depict local linear plants.

From this representation fuzzy plant is defined as follows, taking into each time delay and fault cases:

$$r_1 \textbf{ IF } x_1 \text{ is } A_{1i} \text{ and } x_2 \text{ is } A_{2i} \text{ and } \ldots \text{ and } x_\ell \text{ is } A_{\ell i} \textbf{ THEN } a_i^p x(k) + B_i^p u(k), \tag{4}$$

where $x_1 \ldots x_\ell$ are current state measures, ℓ is the number of states, $i = 1, \ldots, N$ is one of the fuzzy rules, N is the number of the rules which is equal to the number of possible faults and A_{ij} are the related membership functions which are gaussians defined as:

$$A_{ij}(x_i) = exp\left(-\frac{(x_i - c_{ij})^2}{\sigma_{ij}^2}\right) \tag{5}$$

where c_{ij} and σ_{ij} are constants to be tuned. Then it can be write h_i as follows:

$$h_i = \prod_{j=1}^{\ell} A_{ij}(x_j) \tag{6}$$

The final representation of plant as integrated system is based upon center of area defuzzification is expressed as:

$$x^p(k+1) = \frac{\sum_{i=1}^{N} h_i(a_i^p x(k) + B_i^p u(k))}{\sum_{i=1}^{N} h_i} \tag{7}$$

It is important to remember Equation 2 in order to pursue system response. From this representation of a global nonlinear system, the integration of several linear systems, it is necessary to define global stability as a result of this fuzzy system. This review is given in following section considering fuzzy logic control approach. The result of this system representation allows the integration of nonlinear stages and transitions to basically a group of linear plants.

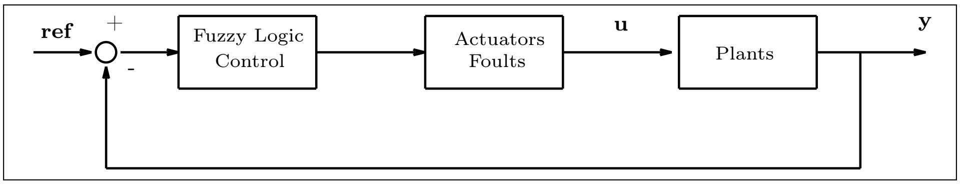


Figure 4: Plant Configuration using Fuzzy Logic Control.

4 Control Approach

From the representation of the plant as fuzzy system [18] it is considered to develop the control law as a group of bounded local linear control laws related to each local linear system. The structure of each fuzzy rule is:

$$r_1 \textbf{ IF } x_1 \text{ is } A^c_{1i} \text{ and } x_2 \text{ is } A^c_{2i} \text{ and } \ldots \text{ and } x_\ell \text{ is } A^c_{\ell i} \textbf{ THEN } u(k) = -g_i x(k), \tag{8}$$

where $i = 1, \ldots, N$, N is the number of fuzzy rules which is the number of faults to be represented, $x_1, \ldots, x_\ell$ are current states of the plant, A^c_{ij} are the gaussians members functions like:

$$A^c_{ij} = exp\left(-\frac{(x_i - c^c_{ij})^2}{(\sigma^c_{ij})^2}\right) \tag{9}$$

where c^c_{ij} and σ^c_{ij} are constants to be tuned. Furthermore g_j represents the related control gain. Similar to fuzzy system plant, fuzzy control representation is integrated as:

$$w_i = \prod_{j=1}^{\ell} A^c_{ij}(x_j) \tag{10}$$

and

$$u(k) = \frac{\sum_{j=1}^{N} w_i(g_i x(k))}{\sum_{i=1}^{N} w_i} \tag{11}$$

The configuration of FLC is integrated to the already explored plant where final representation is given as closed loop system of linear feedback plant as shown in Figure 4.

$$\begin{aligned} x(k+1) &= \frac{\sum_{i=1,j=1}^{N} h_i w_j((a_i - c_i g_j B^p_i)x(k) + B^p_i g_j \mathbf{ref})}{\sum_{i=1,j=1}^{N} h_i w_j} \\ y &= c_i u_i = c_i g_j B_j \end{aligned} \tag{12}$$

where **ref** is the reference to be followed by controller and the variables i and j are used due to fuzzy rules interconnections as the representation of different linear plants and theirs controllers. From this representation stability needs to be stated as the following Lyapunov function:

$$V(x(k)) = x^T(k)Px(k) \tag{13}$$

and

$$\Delta V(x(k)) = v(x(k+1)) - v(x(k)), \tag{14}$$

where

$$V(k+1) = x(k+1)^T P x(k+1)$$
$$= \left(\frac{\sum_{i=1,j=1}^{N} h_i w_j ((a_i - c_i g_j B_j^p) x(k+1) + B_j^p g_j \mathbf{ref})}{\sum_{i=1,j=1}^{N} h_i w_j} \right)^T P \left(\frac{\sum_{i=1,j=1}^{N} h_i w_j ((a_i c_i g_j B_j^p) x(k+1) + B_j^p g_j \mathbf{ref})}{\sum_{i=1,j=1}^{N} h_i w_j} \right). \quad (15)$$

Therefore

$$\Delta V(x(k)) = \left(\frac{\sum_{i=1,j=1}^{N} h_i w_j ((a_i - c_i g_j B_j^p) x(k+1) + B_j^p g_j \mathbf{ref})}{\sum_{i=1,j=1}^{N} h_i w_j} \right)^T P \left(\frac{\sum_{i=1,j=1}^{N} h_i w_j ((a_i c_i g_j B_j^p) x(k+1) + B_j^p g_j \mathbf{ref})}{\sum_{i=1,j=1}^{N} h_i w_j} \right) - X^T P X, \quad (16)$$

and

$$\Delta V(x(k)) = \frac{\sum_{i=1,j=1}^{N} \left((h_i w_j)^2 x^T \begin{bmatrix} -(a_i^T P c_i g_j B_j^P) + a_i^T P (B_j^P g_i \mathbf{ref}) - \\ (c_i g_j B_j^P)^T P a_i - \\ (c_i g_j B_j^P) P (B_j^P g_j \mathbf{ref}) + \\ (g_i B_j^P \mathbf{ref})^T P a_i - \\ (g_j B_j^P \mathbf{ref})^T P (c_i g_j B_j^P) + \\ a_i^T P a_i + (c_i g_j B_j^P) P (c_i g_j B_j^P) + \\ B_j^P g_j \mathbf{ref} P (g_j B_j^P \mathbf{ref}) \end{bmatrix} x - xPx \right)}{\sum_{i=1,j=1}^{N} (h_i w_j)^2}, \quad (17)$$

then by considering $\mathbf{ref}$=0

$$\Delta V(x(k)) = \frac{\sum_{i=1,j=1}^{N} \left((h_i w_j)^2 x^T \begin{bmatrix} -(a_i^T P c_i g_j B_j^P) - (c_i g_j B_j^p)^T a_i \\ + a_i^T P a_i + (c_i g_j B_j^P) P (c_i g_j B_j^P) \end{bmatrix} x - xPx \right)}{\sum_{i=1,j=1}^{N} (h_i w_j)^2}. \quad (18)$$

It is important to remember that:

$$\Delta V(x(k)) \leq 0, \quad (19)$$

and

$$\begin{bmatrix} -(a_i^T P c_i g_j B_j^P) - (c_i g_j B_j^p)^T a_i \\ + a_i^T P a_i + (c_i g_j B_j^P) P (c_i g_j B_j^P) \end{bmatrix} < 0, \quad (20)$$

where $P > 0$, $|| \bullet ||$ is the Euclidean norm and it is possible to define:

$$g_j > ||B_j^P||. \quad (21)$$

This condition has to be given for every single time delay and local fault appearance. Furthermore the stability and the convergence of states should be assured by the adequate selection of matrices g_i (Equation 8) and the related parameters from both fuzzy systems. In this case recommendable procedure to follow is multi-objective optimisation in order to define those suitable values.

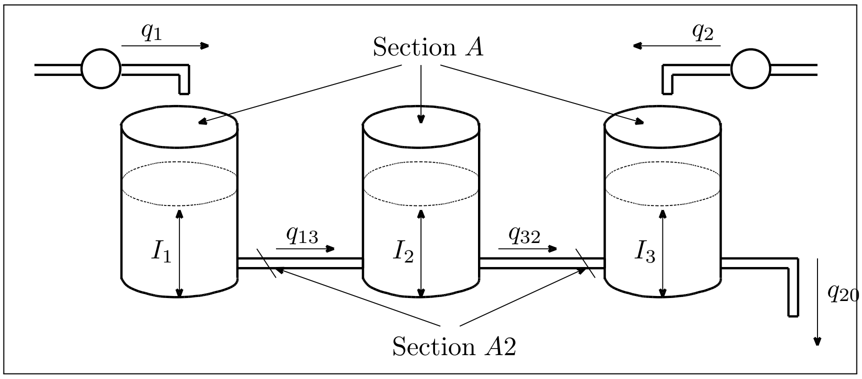


Figure 5: Three tanks representation.

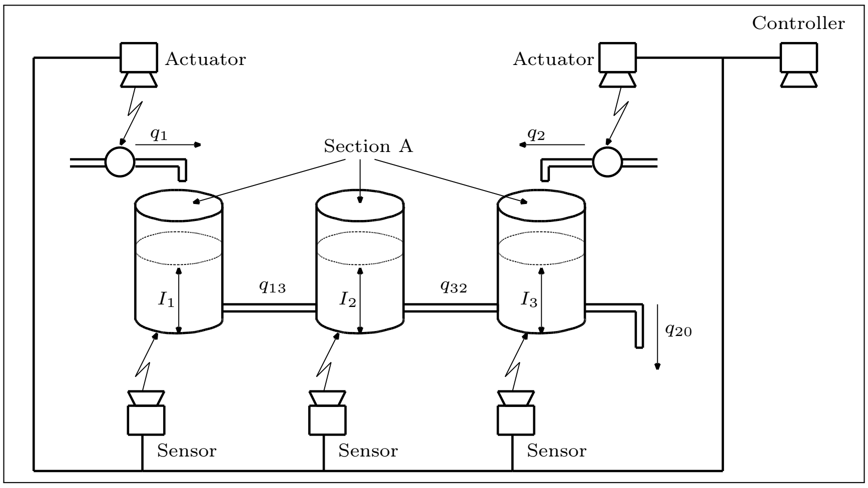


Figure 6: Three Tanks Representation based upon Computer Network.

5 Case of Study

Case study is based upon three tanks approach following Figure 5. This is composed by three tanks with identical cross section A. The tanks are coupled by two pipes with cross section of $A2$. Two pumps are driven by DC motors who supply two inflows q_1 and q_2. Liquid level in each tank is measured and reported as I_1, I_2 and I_3. The computer network integration is shown in Figure 6. This system has a sampling period of 80 ms and nominal communication time delay of 20 ms. Common representation of this system is stated in Eq. 22 as state space representation [19]. This representation takes into account the inherent non-linearities of the model represented as:

$$x(k+1) = \begin{bmatrix} -\dfrac{q_{l3}}{A} \\ \dfrac{q_{32}x - q_{20}x}{A} \\ \dfrac{q_{l3}x - q_{32}x}{A} \end{bmatrix} + \begin{bmatrix} \frac{1}{A} & 0 \\ 0 & \frac{1}{A} \\ 0 & 0 \end{bmatrix} U, \tag{22}$$

where $x = [l_1\ l_2\ l_3]^T$, $y = [l_1\ l_2]^T$ and $u = [q_1\ q_2]^T$ and

$$\begin{aligned} q_{ij} &= \mu_{ij} A * \text{signum}(l_i - l_j)\sqrt{2g|l_i - l_j|} \\ q_{20} &= \mu_{20} A \sqrt{2gl_2}. \end{aligned} \tag{23}$$

Here μ_{ij} are the outflow coefficients. In terms of fuzzy representation, three rules are given due the possible number of faults which are two due to just two actuators are available.

$$r_i \textbf{ IF } x_1 \text{ is } A_{11}^C \text{ and } x_2 \text{ is } X_{21}^c \text{ and } x_3 \text{ is } A_{31}^c \textbf{ THEN } a_1^p x(k) + B_1^p u(k),$$
$$r_2 \textbf{ IF } x_1 \text{ is } A_{12}^c \text{ and } x_2 \text{ is } A_{22}^c \text{ and } x_3 \text{ is } A_{32}^c \textbf{ THEN } a_2^p x(k) + B_2^p u(k),$$
$$r_3 \textbf{ IF } x_1 \text{ is } A_{13}^c \text{ and } x_2 \text{ is } A_{23}^c \text{ and } x_3 \text{ is } A_{33}^c \textbf{ THEN } a_3^p x(k) + B_3^p u(k),$$

In this case B_1^p, B_2^p and B_3^p are defined as:

$$\begin{aligned} B_1^p &= B_1 \left(\int_{0.0}^{0.03} e^{-a_1^p(t-\tau)} d\tau + \int_{0.03}^{0.08} e^{-a_1^p(t-\tau)} d\tau \right), \\ B_2^p &= B_2 \left(\int_{0.0}^{0.04} e^{-a_2^p(t-\tau)} d\tau + \int_{0.04}^{0.08} e^{-a_2^p(t-\tau)} d\tau \right), \\ B_3^p &= B_3 \left(\int_{0.0}^{0.05} e^{-a_3^p(t-\tau)} d\tau + \int_{0.05}^{0.08} e^{-a_3^p(t-\tau)} d\tau \right). \end{aligned} \tag{24}$$

These three plant representations are modified just as B matrix where the reconfiguration is performed. In here two integrals are presented due to important time delays of related faults. For instance, B_2^p has a time delay of 0.04 and B_3^p has a time delay of 0.05 seconds. In the same way control laws are represented, where three control laws are presented.

$$r_i \textbf{ IF } x_1 \text{ is } A_{11}^C \text{ and } x_2 \text{ is } X_{21}^c \text{ and } x_3 \text{ is } A_{31}^c \textbf{ THEN } u(k) = -g_1 x(k),$$
$$r_2 \textbf{ IF } x_1 \text{ is } A_{12}^c \text{ and } x_2 \text{ is } A_{22}^c \text{ and } x_3 \text{ is } A_{32}^c \textbf{ THEN } u(k) = -g_2 x(k),$$
$$r_3 \textbf{ IF } x_1 \text{ is } A_{13}^c \text{ and } x_2 \text{ is } A_{23}^c \text{ and } x_3 \text{ is } A_{33}^c \textbf{ THEN } u(k) = -g_3 x(k).$$

Based upon this equation, final closed loop equation is:

$$x(k+1) = \frac{\sum_{i=1,j=1}^{3} h_i w_j((a_i - c_i g_j B_j^k)x(k) + B_j^k g_j \text{ref})}{\sum_{i=1,j=1}^{3} h_i w_j}. \tag{25}$$

Since this representation is given and based upon stability proposal in last section, optimisation toolbox from MATLAB [20] is used to define current values of g matrices.

6 Implementing Approach

Following the review of case study, this is implemented over a computer network simulation [8,21,22] based upon True-Time strategy. This simulation consists of a CSMA/CA CAN network integrated over ten nodes, where the typical time diagram is presented in Figure 7.

Inconsistencies appear during communication and consumption times as well as the jitter play an important role.

Table 3 shows the consumptions, periods and variations of the involved nodes where sensors and actuators are organised according to EDF algorithm. Moreover, Table 4 shows the related time delays according to Figure 6. Based upon these conditions, scheduling modifications are performed. This scheduling modification affects performance response of peripheral elements where B_i^p (Equation 3) and c_{ij} (Equation 5) matrices are altered as where time delays are defined as follows (Table 5):

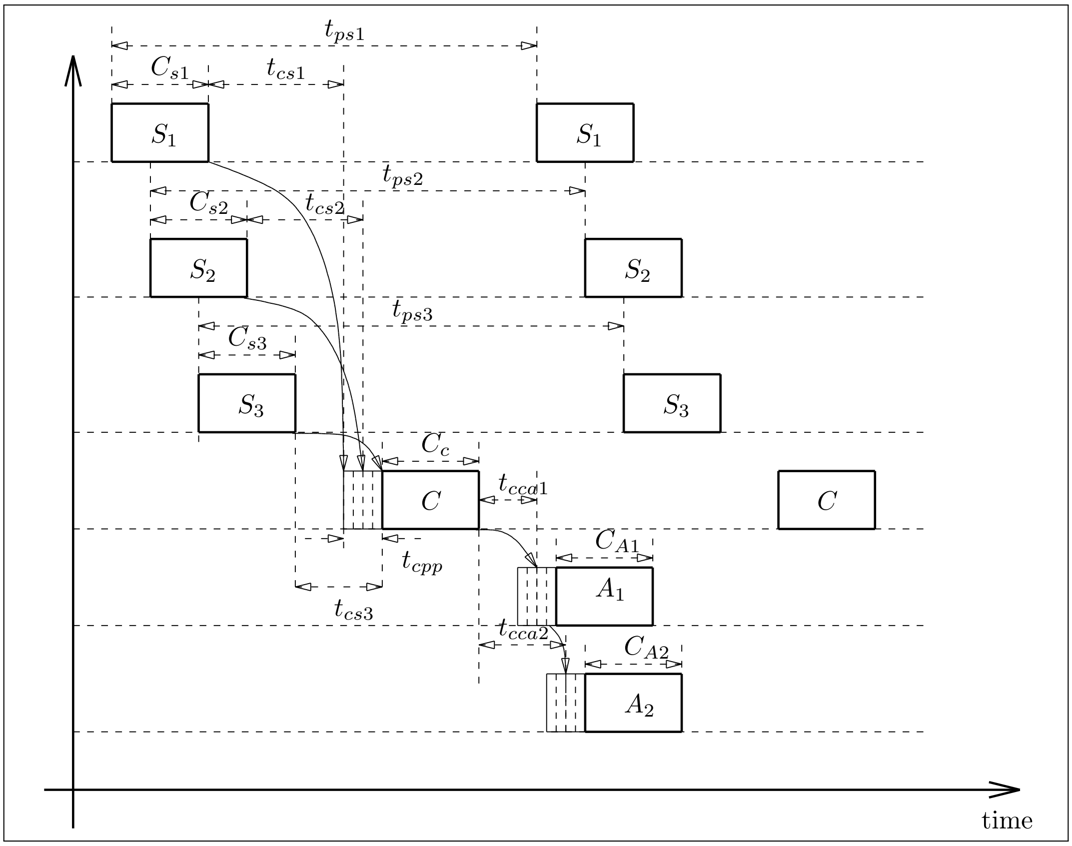


Figure 7: Time Diagram Representation based upon Computer Network.

Table 3: Consumption times and periods from tasks. (in seconds)

Component	Consumption time	Variation %	Time deadlines	Variation %
S_1	$C_{s1} = 0.03$	6-8	0.08	2-3
S_2	$C_{s2} = 0.03$	6-8	0.08	2-3
S_3	$C_{s3} = 0.03$	6-8	0.08	2-3
C	$C_c = 0.05$	7-9	0.08	2-3
A_1	$C_{A1} = 0.045$	5-7	0.08	2-3
A_2	$C_{A2} = 0.045$	5-7	0.08	2-3

Table 4: Communication Times needed (in seconds)

Communication time	Spent time
t_{cs1}	0.02
t_{cs2}	0.02
t_{cs3}	0.02
t_{cca1}	0.02
t_{cca2}	0.02
jitter	0.015

$$\tau_1^1 = 0, \tau_2^1 = t_{cca1} + \text{jitter} + \text{variation, and } \tau_3^1 = 0.08$$
$$\tau_1^2 = 0, \tau_2^2 = t_{cca2} + \text{jitter} + \text{variation, and } \tau_3^2 = 0.08$$

Therefore control reconfiguration becomes necessary in order to keep certain response level. For this case study three different control laws are proposed as shown in Table 6. The type of membership functions are gaussians normally distributed over the rank of each state as shown in Figure 8. In the case that any state is gone further its effected are reduced through the related control law.

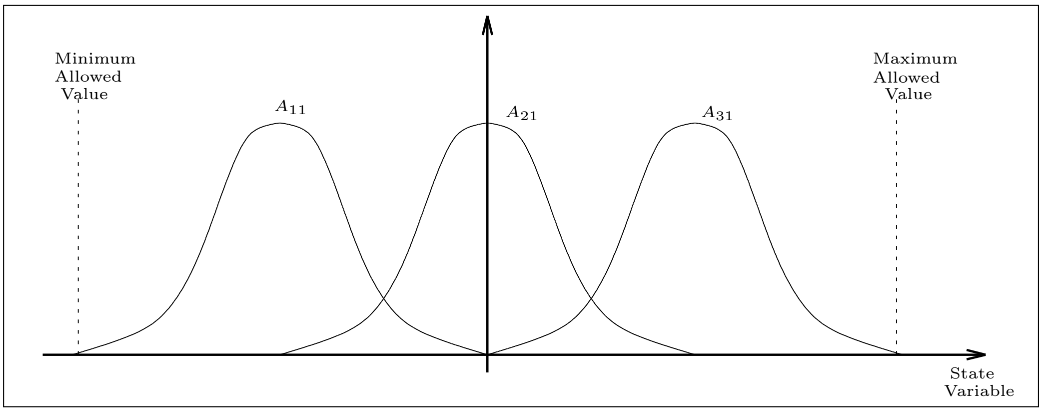


Figure 8: State Variable.

Table 5: Fuzzy logic rules considering the related plant.

Rule	Linearised Plant
if x_1 is A_{11} and x_2 is A_{12} and x_3 and A_{13}	$x(k+1) = a_1^p x(k) + B_1^p u$
if x_1 is A_{21} and x_2 is A_{22} and x_3 and A_{23}	$x(k+1) = a_2^p x(k) + B_2^p u$
if x_1 is A_{31} and x_2 is A_{32} and x_3 and A_{33}	$x(k+1) = a_3^p x(k) + B_3^p u$

Table 6: Fuzzy logic rules considering the related control law.

Rule	Control Law
if x_1 is A_{11} and x_2 is A_{12} and x_3 and A_{13}	$U = g_1X$
if x_1 is A_{21} and x_2 is A_{22} and x_3 and A_{23}	$U = g_2X$
if x_1 is A_{31} and x_2 is A_{32} and x_3 and A_{33}	$U = g_3X$

7 Results

Following the response of the case study, some interesting results are given related to three different scenarios named reconfiguration (stages) 1, 2 and 3 (Figure 9a), where each stage is 20 ms long. And for each of the those cases three different responses like: dynamic response, consumption and communication times are given. The response shown, corresponds to three tanks performance during three different dynamic conditions with respect to communication time behaviour. The local response presented in the third scenario is performed due to modification of the response and dynamic conditions with respect to communication time delays. Figure 10) presents communication system response of each element involved on the system. In this case, periodic activation per task and node is presented. Where free time needs to be reorder to be more efficient. Figure 10a) presents periodic time modification when reconfiguration was necessary and second graphic presents time used to communicate elements through the network. For instance, at first 20 seconds every period of activated task are relatively similar. From 20-40 seconds those modified periodic tasks change the relatively. Similar situation is presented from 40-60 seconds. The reader may realize that some tasks reduce their periodic consumption times while others tends to spend more time.

8 Concluding Remarks

As has been shown in this work, fuzzy logic control based upon Takagi-Sugeno structure allows the possibility of control reconfiguration as long as linear models based representation of the plant are available. Despite of local faults and bounded time delays appearance, several conditions should be fulfil in order to be able to follow this proposal, for instance, plant should be observable and controllable during the whole nonlinear behaviour as well as the states should be present during undesirable situations. Several approaches can be pursued such as robustness study during variations from one to another scenarios as well as further work is needed when not every state is available during fault conditions which is a realistic condition. In this case observer design under fault scenario need to be defined looking for smooth transitions between fault and fault free status. Stability conditions are reached as long as system design is pursued by Equations 15-19, where fuzzy logic variables and local control laws are established. Following future work, a non-desirable drawback is related to a peripheral element fault without any propagation to the rest of the system. This can be possible by just considering some specific scenarios during fault propagation from local to general catastrophic faults.

Acknowledgements

The authors would like to thank the financial support of DISCA-IIMAS-UNAM, and UNAM-PAPIIT (IN106100, IN103310 and IN101307), ICyTDF PICCO10-53 México in connection with this work.

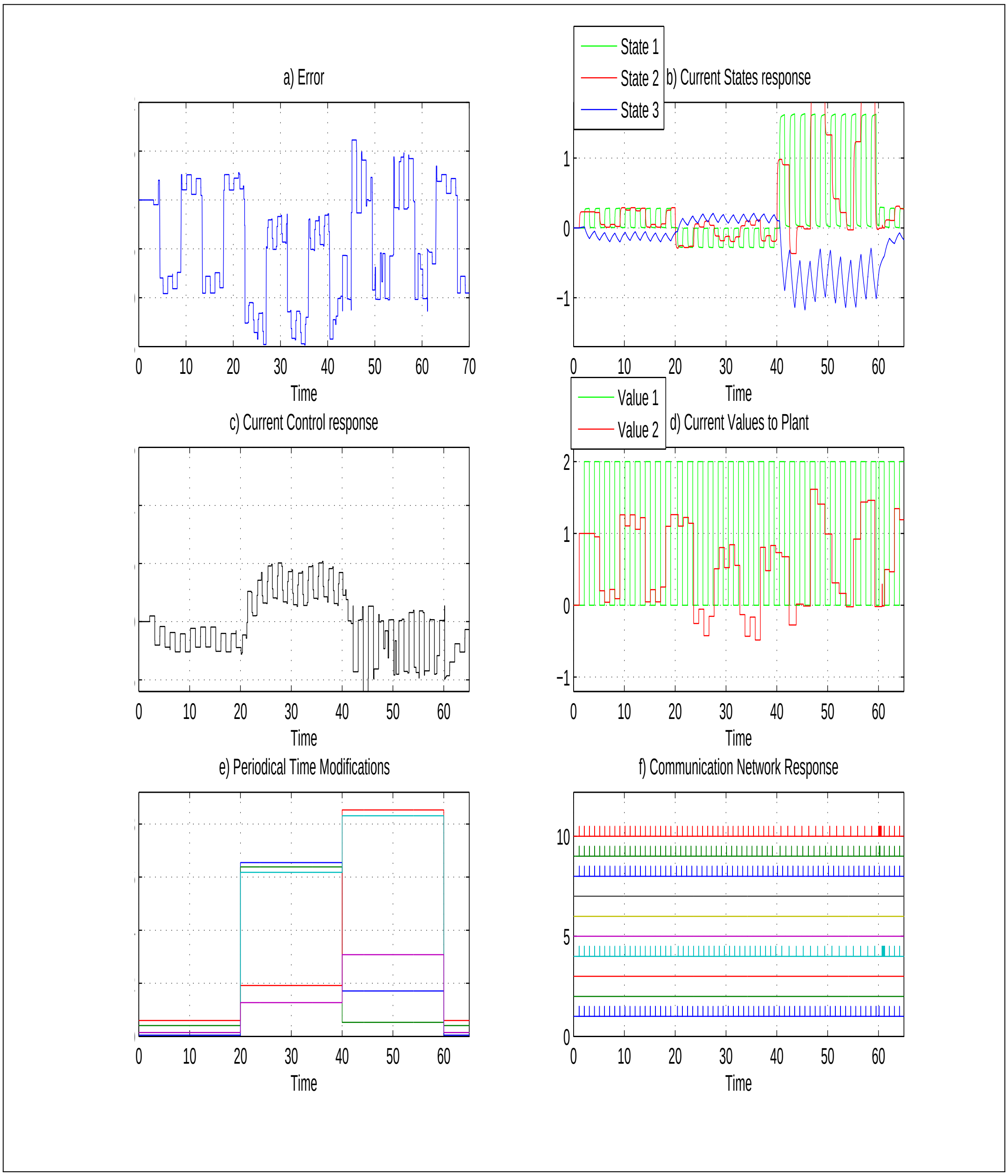


Figure 9: Dynamic Response of Case Study Following Three Scenarios.

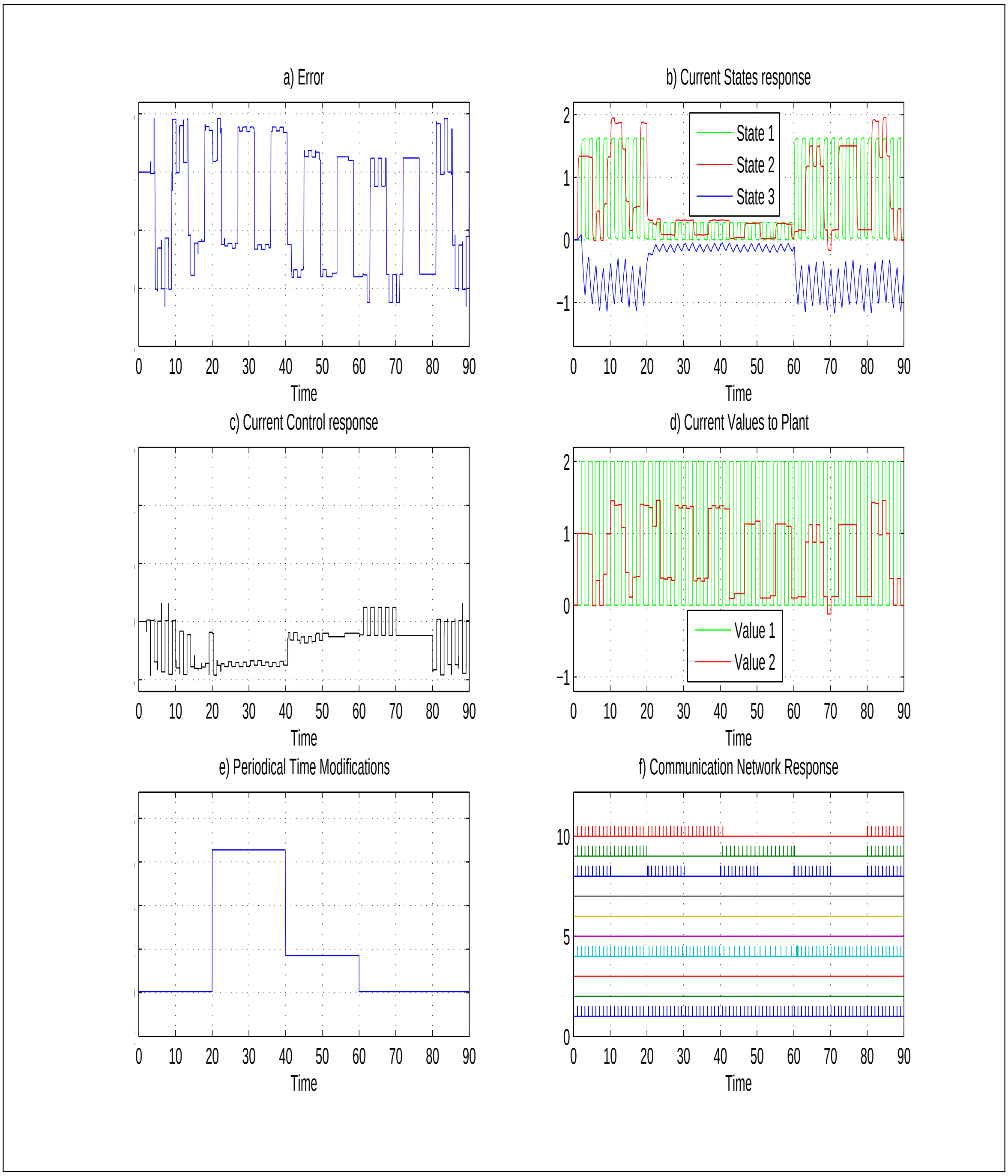


Figure 10: Dynamic Response of Case Study Following Three Scenarios considering periodic communication.

Bibliography

[1] Benítez-Pérez, H., Thompson, H. A. and Fleming, P. J., Simulation of Distributed Fault Tolerant Heterogeneous Architectures for Real-Time Control, *5 th IFAC Workshop on Algorithms and Architectures for Real-Time Control AARTC98, pp. 89-94, Cancún, México*, 1998.

[2] Masten, M. K., Electronics: The intelligence in Intelligent Control *IFAC Symposium on Intelligent Components and Instrument for Control Applications, Annecy, France*, pp. 1-11, 1997.

[3] Nilsson, J., *Real-Time Control with Delays*, PhD. Thesis, Department of Automatic Control, Lund Institute of Technology, Sweden, 1998.

[4] Wu N., Reliability of Reconfigurable Control Systems: A Fuzzy Set Theoretic Perspective,*Proceedings of the 36 th Conference on Decision & Control, IEEE, TP15 5:10, pp. 3352-3356, San- Diego, USA*, 1997.

[5] Jiang J., and Zhao Q.,Reconfigurable Control Based on Imprecise Fault Identification, *Proocedings of the American Control Conference, IEEE, pp. 114-118, San Diego*, June, 1999.

[6] Izadi-Zamanabadi R. and Blanke M.,A Ship Propulsion System as a Benchmark for Fault-Tolerant Control,*Control Engineering Practice, Vol. 7, pp. 227-239*, 1999.

[7] Blanke, M., Kinnaert M., Lunze J., and Staroswiecki M.,*Diagnosis and Fault Tolerant Control*, Springer, 2003.

[8] Benítez-Pérez, H., and García-Nocetti, F., *Reconfigurable Distributed Control*, Springer Verlag, 2005.

[9] Benítez-Pérez, H., and García-Nocetti, F., Reconfigurable Distributed Control based upon Smart Peripheral Elements, *European Control Conference; ECC01, Porto, Portugal*, pp. 3534-3539, 2001

[10] Benítez-Pérez, H., *Smart Distributed Systems*, PhD. Thesis AC & SE Dept. University of Sheffield, UK, 1999.

[11] Driankov, D., Hellendoorn, H., Reinfrank, M., *An Introduction to Fuzzy Logic Control*, Springer- Verlag, 1994.

[12] Benítez-Pérez, H., Cárdenas-Flores F., García-Nocetti, F., *An Implementation of Reconfigurable Network Control based upon Automata Proposal for three Conveyor Belt Case of Study*, International Journal of Computers, Communications & Control, vol. II(2007), No. 4, pp 314-327, 2007.

[13] Benítez-Pérez, H., *Real-time distributed control: A fuzzy an model predictive control approach for a nonlinear problem*, Nonlinear Analysis Hybrid Systems, 2 (2008), pp. 474-490, 2008.

[14] Quiñones-Reyes P., Benítez-Pérez, H., Cárdenas-Flores F., García-Nocetti, F. *An approximation for reconfigurable fuzzy Takagi-Sugeno networked control*, IEEE Potentials, vol. 27, No. 6, pp. 38-44, 2008.

[15] Liu, J.,*Real-Time Systems*, Prentice Hall, 2000.

[16] Krishna C. M. and Shin K. G., *Real-Time Systems*, Mc. Graw Hill, New York, USA, 1997.

[17] Lian F. Moyne J. and Tilbury D., Network Design Consideration for Distributed Control Systems, *IEEE Transactions on Control Systems Technology*, Vol. 10, No. 2, pp. 297-307, March 2002.

[18] Yi, Z., and Heng, P.; ďż˝Stability of Fuzzy Control Systems with Bounded Uncertain Delaysďż˝; IEEE Transactions on Fuzzy Systems, vol. 10, No.1, pp. 92-97, 2002.

[19] Dardinier-Maron V., Hamelin F., and Noure H.; A Fault-Tolerant Control Design Against Major Actuator Failures: Application to a Three-Tank Systemďż˝; IEEE 38th Conference on Decision and Control, pp: 3569-3574, 1999.

[20] Mathworks; System Identification Toolbox Users Guide; MATLAB, 1998.

[21] Cervin, A., Henriksson, D., Lincoln, B., Eker, J., and Arzďż˝n, K.; ďż˝How Does Control Timing Affect Performance?ďż˝; IEEE Control Systems Magazine, Vol. 23, pp. 16-30, 2003

[22] Benítez-Pérez, H., and García-Nocetti, F.,*Reconfigurable Takagi-Sugeno Fuzzy Logic Control for a Class of Nonlinear System considering Communication Time Delays on Peripheral Elements*, Submitted to American Control Conference 2006.

[23] Zadeh L. A., *A theory of commonsense knowledge, Aspects of Vagueness*, H.J. Skala, S. Termini and E. Trillas (eds.), 257-296. Dordrecht: Reidel, 1984.

[24] Zadeh L. A., *Fuzzy probabilities*, Information Processing and Management 19, 148-153, 1984.

[25] Zadeh L. A., *The Role of Fuzzy Logic in Modeling, Identification and Control*, Modeling Identification and Control, 15(3):191-203, 1994.

[26] Zadeh L. A., *Fuzzy Logic and the Calculus of Fuzzy If-Then Rules*, Proceedings of SYNAPSE '91, Tokyo, Japan, 1991.

Int. J. of Computers, Communications & Control, ISSN 1841-9836, E-ISSN 1841-9844
Vol. VI (2011), No. 3 (September), pp. 403-417

Identification of ERD using Fuzzy Inference Systems for Brain-Computer Interface

I. Dzitac, T. Vesselényi, R.C. Ţarcă

Ioan Dzitac
Aurel Vlaicu University of Arad
Department of Mathematics-Informatics
Elena Dragoi St., 2, Complex M
Arad, Romania
E-mail:ioan.dzitac@uav.ro

Tiberiu Vesselényi, Radu Cătălin Ţarcă
University of Oradea
Faculty of Management and Technological Engineering
Universitatii St. 1, 410087
Oradea, Romania
E-mail:tvesselenyi@yahoo.co.uk, rtarca@uoradea.ro

Abstract: A Brain-Computer Interface uses measurements of scalp electric potential (electroencephalography - EEG) reflecting brain activity, to communicate with external devices. Recent developments in electronics and computer sciences have enabled applications that may help users with disabilities and also to develop new types of Human Machine Interfaces. By producing modifications in their brain potential activity, the users can perform control of different devices. In order to perform actions, this EEG signals must be processed with proper algorithms. Our approach is based on a fuzzy inference system used to produce sharp control states from noisy EEG data.
Keywords: Event Related Desynchronization (ERD), Brain-Computer Interface (BCI), electroencephalography (EEG), fuzzy inference system.

1 Introduction

Devices that are using brain (cortical or scalp) electric potentials which allow brain function signals to control machines (or computers) are called brain-computer interfaces (BCI). It had been shown that BCI systems using steady state visual evoked potentials have allowed healthy subjects to communicate with computers [1].

Non-invasive, electroencephalogram (EEG)-based BCI technologies can be used to control a limb prothesis, for the use of a computer program, and for other functions such as environment control or entertainment. BCI technologies can improve the lives of people with neurological disorders and also can restore motor control by guiding activity-dependent brain plasticity [2]. A review of the advantages of BCIs which shows that they can provide even a good candidate for space applications is presented in [3]. Some BCI implementations are using sensory-motor rhythms as it is presented in [4]. The interest shown for BCI technology is motivated by the applicability for helping disabled, for gaming, and as a tool in cognitive neuroscience [5]. In [6] a model is proposed which includes a set of definitions of the typical entities encountered in a BCI, diagrams which explain the structural correlations among them and a detailed description of the timing of a trial.

In this paper we present some of our original experiments with BCI, and EEG data processing methods applied.

Copyright © 2006-2011 by CCC Publications

2 BCI applications

Important application for BCIs can be found in the medical field. A good example for this is presented in [7] where BCI methodology based on self-regulation of slow-cortical potentials of the EEG is used. This kind of two choice task is called "brain-switch" [8], but there are other applications in which multiple choices are employed.

Another important application of BCI is represented by cognitive experiments during which the researchers try to better understand some sequences of brain functions. In [9] for example the researchers use information derived from previous EEG recordings to inform the analysis of functional magnetic resonance imaging (fMRI) data collected for the same behavioral task.

Robotics and environment control applications are also emerging in the studies of this area. For example in [10] a study was made in order to determine the most discriminative features for a BCI system based on statistically significant differences between two energy density maps calculated from EEG signals during two different motor tasks. In [11] it is shown that simple rehabilitation tasks can be performed and household robotic devices can be easily controlled by means of BCI technology. Robotic arms, which have been controlled by means of cortical invasive interfaces in animal studies, could be the next frontier for noninvasive BCI applications. A review of current robotic technologies that are relevant to BCI is presented in [12]. As it can be seen from the references in this field, employment of prothesis and rehabilitation is a largely studied area in the field of BCI applications [13], [14], [15]. Several other applications of BCI are designed for entertainment, gamming and learning to concentrate or to achieve a certain desired mental state.

3 Considerations on EEG signals and signal processing methods used in BCI

Most of the BCI devices developed by now are based on the electric potentials of scalp known as EEG, which is noninvasive opposing to those based on cortical potentials where the electrodes are in direct contact with the cortex (electrocorticography - ECoG). A special method is using EEG signals combined with an appropriate model of volume conduction and of neuro-electrical sources (high-resolution electroencephalography) [16].

In order to intentionally generate the electrical potentials of the brain two kinds of tasks are widely used which are also correlated with different regions of the brain: the sensory-motor (or sometimes motor imagery) tasks and those tasks which imply some kind of processing called mental or cognitive-related tasks (i.e. simple syntactic analysis of words). The sensory-motor tasks are by far mostly used because their direct implication in limb movement [17], [4], [18]. Examples of mental or cognitive-related task studies are given in [19] and [20].

In sensory-motor tasks an often used method is to employ Event-related desynchronization (ERD) which is a reduction in EEG signal amplitude in a specific frequency band (8 - 12 Hz) as a result of a particular event. When a subject is making a hand movement, or even just imagine it, a reduction in sensory-motor signal amplitude can be observed in certain parts of the scalp. In order to use this in a BCI application the proper components of the signal must be found. The signals measured by sensors, particularly those used in noninvasive techniques, are known to be very noisy due to a series of potentials generated by other biological or artificial sources than brain potentials. One of the major challenges of BCI devices is to filter out the unwanted components of the signals. Here both hardware and software solutions are to be applied. The other major task is to uniquely recognize a specific pattern in the useful component of the signals which is meant to trigger a specific command. The number of sensors placed in specific areas of the scalp can vary from 1 to 36 which are read in on different channels of the device. Each of these channels can be filtered separately to monitor a specific, larger or narrower, frequency domain.

Researchers applied a wide range of techniques to filter, preprocess and analyze signals and to produce decisions.

Statistical Approach. In [21] the researchers studied spike sorting by introducing a mathematical model consisting of three Gaussian waveforms, which appropriately represents the general shapes of action potentials. Then they searched for the best-fit waveform for each noise-corrupted spike based on the model, using peak fitting method. The performance of the proposed method was assessed with synthesized neural recordings composed by spike templates and white Gaussian noise in various signal to noise ratio environments. Gaussian process (GP) classification to binary discrimination of motor imagery EEG data is described in [22]. A robust Bayesian linear regression algorithm is presented in [23] that automatically detects relevant features and excludes irrelevant ones, all in a computationally efficient manner.

Spectral Methods. In [24] the researchers investigated spectral power changes both in intracranial ECoG recordings in epilepsy patients and in non-invasive EEG recordings optimized for detecting high gamma band (a specific frequency domain in EEG signals) activity in healthy subjects. In order to characterize the non-Gaussian information contained within the EEG signals, a new feature extraction method based on "bispectrum" is proposed in [25] and applied to the classification of right and left motor imagery.

In [26] a novel method is presented for detecting frequency-frequency coupling between the electrical output of cortical areas as measured by ECoG, electroencephalography EEG and magnetoencephalography MEG, the biphase-locking value. The method is specifically sensitive to non-linear interactions, i.e. quadratic phase coupling across frequencies. This method was employed to study signals measured from pre-motor area, in the motor cortex.

Wavelet Methods. Wavelet transforms are applied in [27] to study motor imagery tasks. In this work the continuous wavelet transform (CWT) was applied together with Student's two-sample t-statistics for 2D time-scale feature extraction, where features are extracted from EEG signals recorded from subjects performing left and right motor imagery. In [28] a novel method was proposed for the feature extraction of electroencephalogram (EEG) based on wavelet packet decomposition (WPD). In [29] the assessment of wavelet transform (WT) as a feature extraction method was used in representing the electrophysiological signals.

Genetic Algorithms. Positive potentials at a latency of about 300 ms in EEG (P300) are largely used to study brain functions. This pattern is used as sign of cognitive function in decision making processes. In [30] P300 detection approach based on some features and a statistical classifier was implemented. The optimal feature set was selected using a genetic algorithm from a primary feature set including some morphological, frequency and wavelet features and was used for the classification of the data.

Artificial Neural Networks and Support Vector Machines. In [31] the authors investigated if an artificial neural network-based model for closed-loop-controlled neural prostheses could use neuromuscular activation recorded from individuals with impaired spinal cord to predict their end-point gait parameters (such as stride length and step width). This ANN-based model allows a seamless incorporation of neuromuscular activity, detected from paralyzed individuals, to adaptively predict their altered gait patterns, which can be employed to provide closed-loop feedback information for neural prostheses. Support vector machine implementations are used in [32], [33], [34]. The results indicate that the proposed algorithms are promising for future use of rehabilitative BCI applications in neurologically impaired patients.

Fuzzy logic. In [35] a new method is introduced, which combines information from different classic time series similarity measures, using a fuzzy fusion framework. This method is accurate and reliable in P300 detection. This framework is used to combine two computationally simple signal detection methods: "peak picking" and "template matching". Fusion takes place in the last step (decision-making step) by means of a fuzzy rule-base. In order to predict the gripping force from the EEG signals in [36] a methodology that uses subsequent signal processing methods is used: filtering, principal component analysis, and the phase-demodulation method. A fuzzy inference system is then used to predict the gripping force from the processed EEG data. Neuro-fuzzy methods are also to be considered as a highly efficient way to deal with uncertain information.

The paper [37] introduces a number of modifications to the learning algorithm of the self-organizing fuzzy neural network (SOFNN) to improve computational efficiency. An analysis of the SOFNN effectiveness when applied in an electroencephalogram (EEG)-based brain-computer interface (BCI) involving the neural-time-series-prediction-preprocessing (NTSPP) framework is also presented, where a sensitivity analysis of the SOFNN hyper-parameters was performed using EEG data recorded from three subjects during left/right-motor-imagery-based BCI experiments. The aim of this analysis was to eliminate the need to choose subject and signal specific parameters for EEG preprocessing. The results indicate that a general set of NTSPP parameters chosen provide the best results when tested in a BCI system.

A new brain-computer interface design using fuzzy ARTMAP (FA) neural network, as well as an application of the design is proposed in [38]. The objective of this BCI-FA design is to classify the best three of the five available mental tasks for each subject using spectral amplitude values of electroencephalogram (EEG) signals. The findings show that the average BCI-FA outputs gave very low error using the best triplets of mental tasks identified from the classification performances of FA. In [39] a subject-based feature extraction method using the fuzzy wavelet packet in brain-computer interfaces (BCIs) is presented. In [40] is presented a new method introduced in [34], which combines information from different classic time series similarity measures, using a simple fuzzy fusion framework. This method is accurate and reliable in P300 (a positive event-related component occurring 300ms after stimulus onset) detection. Fusion takes place in the decision-making step by means of a fuzzy rule-base. Compared to similar works on electroencephalogram-based (EEG-based) BCI datasets, in spite of being computationally simple, this new technique's performance is comparable to very complicated methods. In [41] an analyze of electroencephalogram (EEG) signals of imaginary left and right hand movements, an application of brain-computer interface (BCI) is presented. The researchers propose here to use an adaptive neuron- fuzzy inference system (ANFIS) as the classification algorithm. ANFIS has an advantage over many classification algorithms in that it provides a set of parameters and linguistic rules that can be useful in interpreting the relationship between extracted features. The continuous wavelet transform is used to extract highly representative features from selected scales.

4 Method and results

The 10-14 Hz frequency band of EEG signals is, among others, associated with those cortical areas that are most directly connected to the brain's normal motor output channels. Movement or preparation for movement is typically accompanied by a decrease of amplitude in this band over sensory - motor cortex. This decrease has been labeled "event-related desynchronization" or ERD [42]. Opposite to ERD, amplitude increase, or "event-related synchronization" (ERS) occurs in the post-movement period and with relaxation. Furthermore, and most relevant for BCI applications, ERD and ERS occur also with motor imagery (i.e., imagined movement) and they do not require actual movement.

In this research our goal is to analyze EEG signals by observing ERD and ERS characteristics and find a suitable algorithm to automatically detect these characteristics by using a fuzzy inference system. We consider that given the high noisiness of EEG signals a fuzzy or neuro-fuzzy approach would be desirable.

In our experiments subjects were asked to perform several hand movements at irregular times (with random intervals between two movements) during a session of 40 seconds. The hand movement consists of pushing a switch button. The subjects were also asked to concentrate on the action (hand movement) imagining it before. The EEG signals collected from a single electrode placed in C3 point on the scalp were registered. Electrode placement location is shown in figure 2. The switch on-off states were also registered using a separate acquisition device which is triggered to start data acquisition at the same time as the EEG signal acquisition, so the EEG signals and the switch on - off states data covers the same time period.

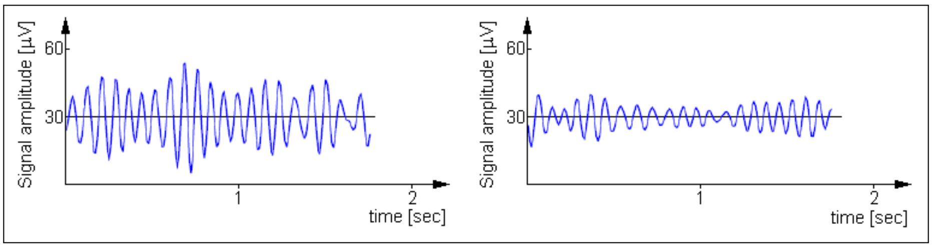


Figure 1: Theoretical EEG signals in 12 Hz frequency band. a) resting hand; b) moving hand (amplitude decrease - ERD).

A total of five healthy male adults (aged 25-48) participated in this study. None of the subjects had participated in the same experiments before and the subjects had a short basic training for the task. The subjects were informed about the nature and purpose of the experiments and consented to participate.

EEG signals were acquired using a BST112 8-channel amplifier from VEB MESSGERATEWERK ZVONIZ and for A/D conversion and transmission of data to a PC, a National Instruments NI-USB 6251 DAQ board had been used. The accuracy of the measuring system is of 1.5259μV. The Ag-AgCl electrode was placed in the C3 point on the scalp, using a conductive gel. As voltage reference, the A1 point was used (figure 2) and the ground was located on the right leg. Data were sampled at a sample rate of 1000 Hz and were transmitted through the USB port to a PC for storage and processing. The data acquisition and processing programs were written in MATLAB language.

Every subject has made 3 training sessions to accommodate with the environment and the equipment and then 4 sessions of 40 seconds each were recorded.

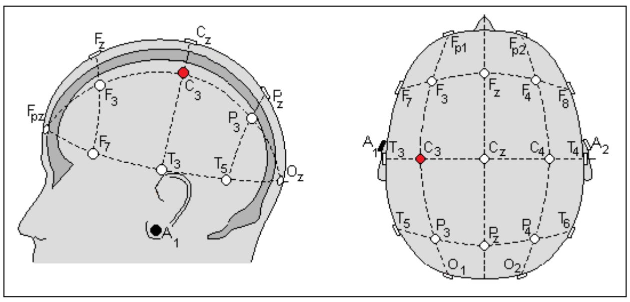


Figure 2: Location of electrode on the scalp in point C3.

The diagram of a sample acquisitioned signal is given in figure 3. As it can be observed the signal is affected by low level noise resulting in a variable offset across it's whole length.

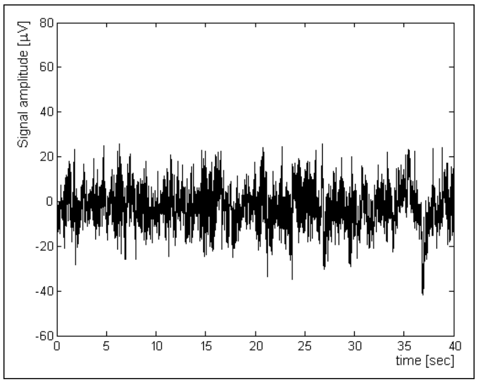


Figure 3: A sample of acquisitioned EEG signal.

This can be eliminated by using a running average filter which gives local average values which then are extracted from the original signal (sEEG1). If n is the length of the signal v (in samples) and m is the running average filter length given by the vector u, then the length of the average will be m+n-1. The running average w is given by convolution of the filter vector with the signal as it is shown in formulae 1:

$$w(k) = \sum_j u(j) \cdot v(k - j + 1) \tag{1}$$

where k is the index of the resulting signal w and j is the index of the filter u. In this case all values of u are equal to 1/m. In this case we used m = 200, which for a sampling rate of 1 kS/sec will not generate errors for the studied frequency band (centered on 12 Hz).

The running average values are shown in figure 4 (red). According to the convolution theorem, that convolution of signal and the digital filter's impulse response is nearly the same as multiplying their Fourier transforms. This can be used to speed up the calculations if it can be shown that the results are accurate up to an acceptable limit.

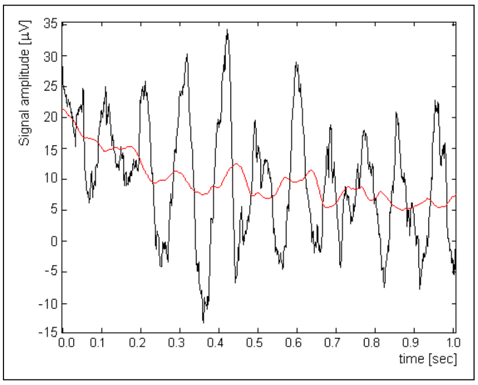


Figure 4: Detail of the acquisitioned signal for a 1 second time span (black). Running average values of the signal (red).

After extraction of the running average the resulting signal (sEEG2), is free of low frequency noises as it can be seen in the diagram given in figure 5.

A further enhancement of the signal can be obtained by reducing the random peaks, employing a running average filter of the same form as before, but this time keeping the resulting average values. This has the effect of smoothing the signal without modifying the studied components.

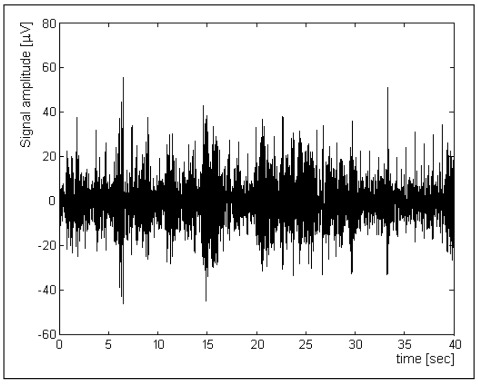


Figure 5: The signal given in figure 3, with running average values extracted.

This time we used m = 20. The resulting signal (sEEG3) is shown in detail for a time span of 1 second in figure 6. Studying this signal we can observe that the dominating frequency is around 12-14 Hz but still having some low frequency components.

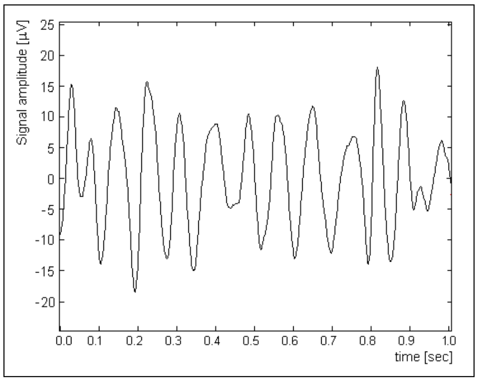


Figure 6: Detail of processed signal given in figure 5, for a time span of 1 second.

In order to further analyze the signal we have to switch to frequency domain analysis because time domain diagrams can no longer be enhanced. In case of frequency domain analysis of acquisitioned digital signals spectral amplitudes are computed using the Fast Fourier Transform (FFT - which is a variant of the Fourier transform used in numeric computing). A FFT amplitude value is given by formulae 2:

$$A(k) = \sum_{j=1}^{N} a(j)\omega_N^{(j-1)(k-1)} \tag{2}$$

where: $\omega_N = e^{\frac{(-2\pi \cdot i)}{N}}$, N - number of samples; a(j) - amplitude of j- th sample in the signal.

In order to obtain the required spectrum we will use only the real part of the obtained FFT amplitudes (real(A(k))).

As we want to analyze changes in the signal the FFT algorithm should be applied not on the entire length of the signal but with a suitably sized window which is moved step by step on the whole length of the signal. In our case we used a window of size NFFT = 256. The number of partitions is NP = integer(m/NFFT). Some samples at the end of the signal will be lost but these can be only a quarter of a second which in our case is reasonable (given a 40 seconds long signal).

Representing the result we will obtain a diagram of FFT amplitudes as a function of time and frequency as it is shown in figure 7. The results had been interpolated using a cubic spline interpolation method in order to smooth the borders between partitions.

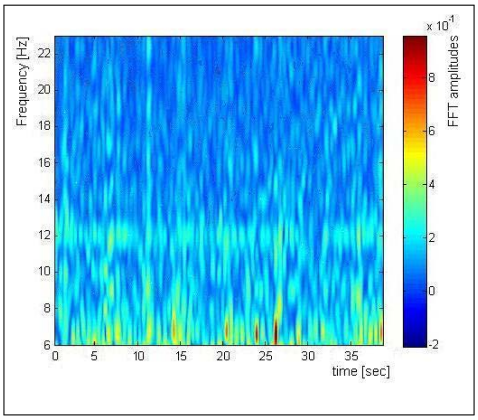


Figure 7: Time frequency diagram of the processed signal. The frequency band centered on 12 Hz is clearly observable.

If we select a slice of this diagram at a specific moment in time we ca obtain the spectra of the signal. Such spectra are given in figures 8 (resting hand) and 9 (moving hand).

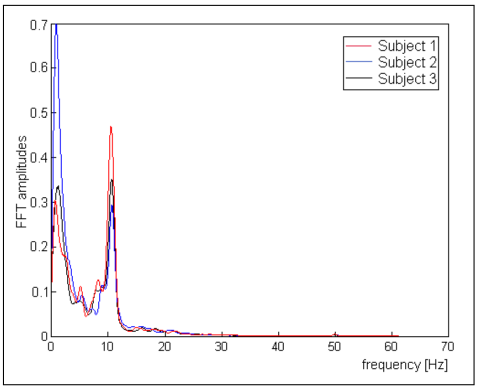


Figure 8: FFT amplitudes of the EEG signal for resting hand.

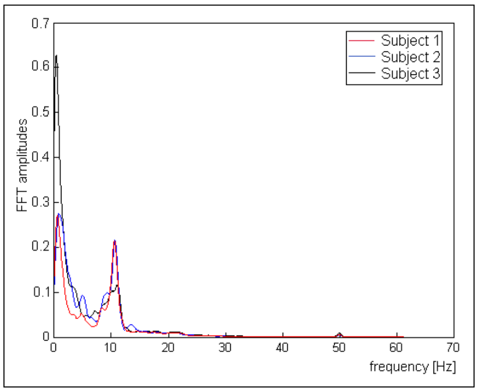


Figure 9: FFT amplitudes of the EEG signal for moving hand.

If we represent a slice of the time-frequency diagram for a specific frequency we will obtain the variation in time of FFT amplitude at that frequency. As ERD and ERS are occurring in a frequency band of 10-14 Hz choosing a center frequency of 12 Hz the resulting diagram is shown in figure 10 (black). On this diagram we had superimposed the switch on-off states (red). We can observe that the "ON" states (which are showing the real movement of the hand) mainly corresponds to the drop of signal amplitude of the chosen frequency. This should show the occurrence of ERD. Also there can be observed a rise of the curve corresponding with time intervals when the switch is not activated (the subject is relaxed). This should show the occurrence of ERS. In case of ERD the anticipation of the real movement, the fact that the amplitude is dropping earlier than the hand movement occurs shows that it happens due to motor imagery.

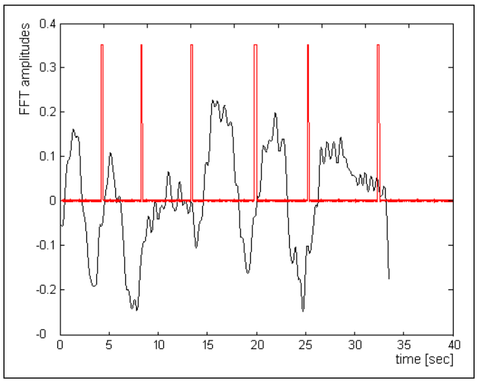


Figure 10: FFT amplitude versus time for the processed signal at the frequency of 12 Hz (black) and switch on-off states diagram (red).

As it can be seen in figure 10, and it had been observed in almost all of the studied cases, these characteristics can not be expressed with simple parameters (i.e. value of amplitude or the slope of the amplitude drop). Hence, in order to automatically detect ERD occurrence in FFT amplitude signals some fuzzy or neuro-fuzzy algorithms can be applied which may be capable of recognizing these kinds of amplitude drops. Fuzzy Logic is widely used in modelling and identification problems and in applications where data sets are obtained with a greater or smaller degree of uncertainty [43], [44], [45].

In the followings we will use an adaptive neuro-fuzzy inference system (ANFIS) to solve this issue. ANFIS is implemented in the Fuzzy Logic Toolbox of the MATLAB environment and uses a hybrid learning algorithm in order to identify parameters of Sugeno-type fuzzy inference systems [46]. It applies a combination of the least-squares method and the backpropagation gradient descent method for training membership function parameters to emulate the given training data set. The data sets obtained and processed, as it has been described, will be used for training and checking of the ANFIS system. After training, this system can be used to identify portions of the EEG signal which meet the ERD characteristics.

From the acquisitioned and processed data we had chosen 14 intervals which represents true ERD characteristics and which we will use to train, check and validate the fuzzy identification algorithm. In order to apply the algorithm the amplitudes must be normalized and the time intervals must be scaled so all the datasets representing ERD characteristics will have the same amplitude domain and the same number of samples. For normalization the datasets were divided with the maximum value of each dataset and for time scaling a linear interpolation algorithm had been used.

Two of the datasets were chosen as input data for training and checking. These datasets are shown in figure 11. Another dataset had been chosen as output dataset shown in figure 12.

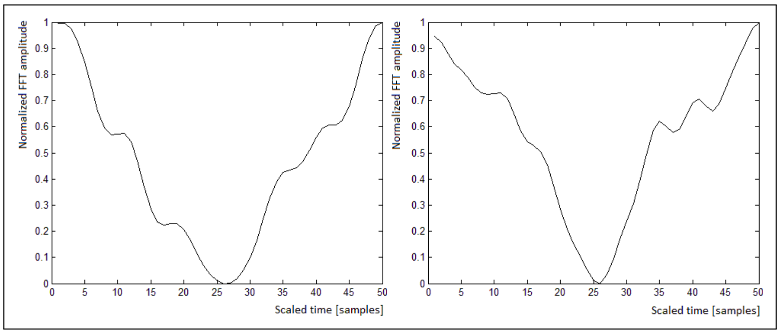


Figure 11: Normalized FFT amplitudes versus scaled time for two ERD characteristics; a. training input, b. checking input.

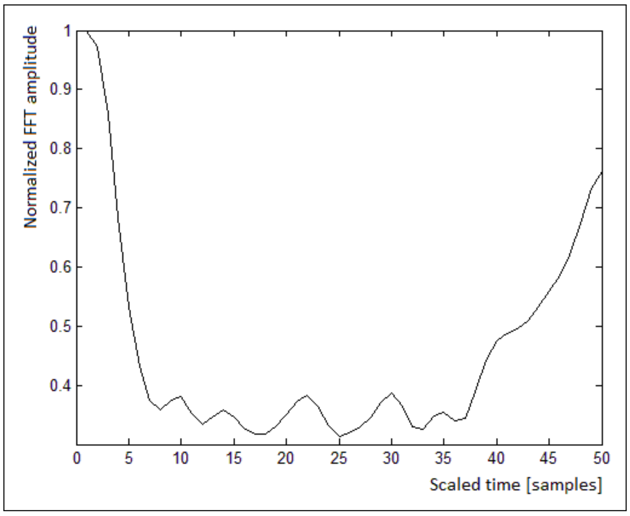


Figure 12: Normalized FFT amplitudes versus scaled time, output dataset.

In the training process the initial membership functions are modified to fit the training input - output dataset and in the checking process these membership functions are tuned in correspondence with the checking input-output datasets. The diagrams of the membership functions obtained after training are given in figure 13.a. and after tuning in figure 13.b.

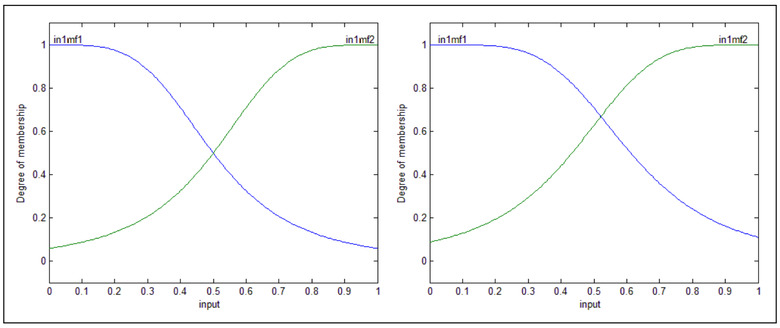


Figure 13: Membership functions obtained after training (a.) and after tuning (b.).

The fuzzy inference system parameters are: number of nodes: 12; number of linear parameters: 4; number of nonlinear parameters: 6; number of fuzzy rules: 2.

Once the fuzzy inference system was trained and tuned we used 14 ERD characteristic datasets to verify their outputs (VO) compared to the results obtained for the dataset used as output in the training process (TO) which is used as reference. System output of the two different validation datasets are given as examples in diagrams of figure 14.

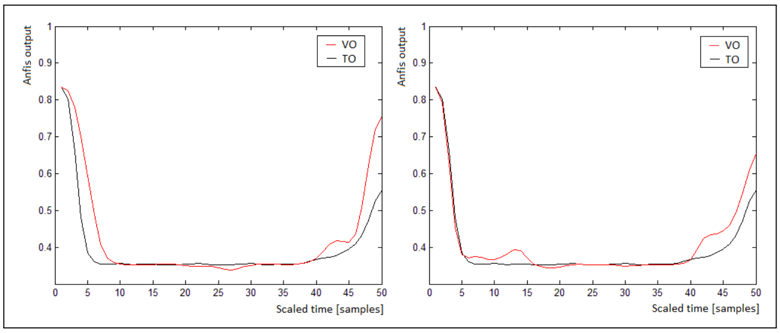


Figure 14: Validation and training system output data diagrams.

For each validation dataset the root mean square (RMS) values of the difference between VO and TO datasets were computed. These values are shown in the diagram in figure 15.

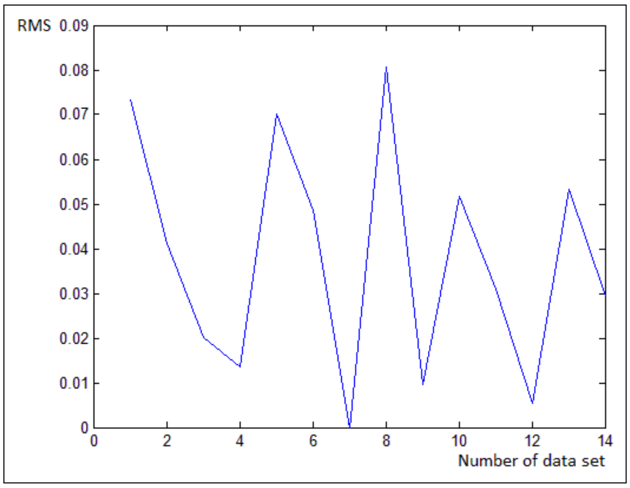


Figure 15: RMS values of difference between VO and TO for 14 validation datasets.

In order to identify if a dataset is a true ERD component or not the RMS values of difference between VO and TO can be used as a classifier. On the basis of RMS values span obtained for known true ERD components we can distinguish ERD components in real time acquired signals using simple or complex threshold techniques.

5 Conclusions and future works

In this work we described our experiments and the development of data processing for EEG signals, focusing on identification of ERD characteristics which can be used in brain - computer interfaces. The experiments showed that EEG signals had to be carefully acquired and processed in order to identify ERD components. Data acquisition equipment quality, careful selection of subjects and subject focus capacity are of major importance in achieving successful results. Data preparation had been done considering simple and time effective algorithms which can be used for real time processing of EEG signals.

Further studies must be made to validate the described fuzzy inference system on a larger number of datasets for both true and false ERD components in order to establish it's robustness. There is still in discussion whether to use output in the training process dataset as reference for ERD identification or to find another more appropriate reference. In a future work we will try to answer these questions.

Acknowledgements

We thank our colleagues Simona Dzitac, Cristina Hora and Ioan Moga for her/his practical contributions in this research.

Also we thank the professors Hariton Costin and Horia-Nicolai Teodorescu for the positive and constructive comments and recommendations on our manuscript, which significantly contributed to improving the quality of this paper.

We are very honored to thank professor Lotfi A. Zadeh for his generous openness in collaborations with Romanian researchers and for his visits in Romanian cities Arad (2003: He was granted as Doctor Honoris Causa of Aurel Vlaicu University of Arad) and Oradea (2008: As Keynote Speaker in our conference ICCCC 2008 and Exploratory Workshop on Natural Language Computation- EWNLC 2008).

Dear professor Zadeh, celebrating the 90th birthday, we wish you a good health, long life, and new interesting achievements!

Bibliography

[1] Brendan Z. Allison, Dennis J. McFarland, Gerwin Schalk, Shi Dong Zheng, Melody Moore Jackson, Jonathan R. Wolpaw, Towards an independent brain-computer interface using steady state visual evoked potentials, *Clinical Neurophysiology*, 119(2): 399-408, 2008

[2] Janis J Daly, Jonathan R., Wolpaw, Brain-computer interfaces in neurological rehabilitation, *The Lancet Neurology*, 7(11): 1032-1043, 2008

[3] Carlo Menon, Cristina de Negueruela, José del R. Millán, Oliver Tonet, Federico Carpi, Michael Broschart, Pierre Ferrez, et. all, Prospects of brain-machine interfaces for space system control, *Acta Astronautica*, 64(4): 448-456, 2009

[4] Elisabeth V.C. Friedrich, Dennis J. McFarlanda, Christa Neuperb, Theresa M. Vaughan, Peter Brunner, Jonathan R. Wolpaw, A scanning protocol for a sensorimotor rhythm-based brain-computer interface, *Biological Psychology*, pp. 80(2): 169-175, 2009

[5] Marcel van Gerven, Ole Jensen, Attention modulations of posterior alpha as a control signal for two-dimensional brain-computer interfaces, *Journal of Neuroscience Methods*, 179(1): 78-84

[6] Lucia Rita Quitadamo, Maria Grazia Marciani, Gian Carlo Cardarilli, Luigi Bianchi, Describing Different Brain Computer Interface Systems Through a Unique Model: A UML Implementation, *Neuroinform*, 6: 81-96, 2008, DOI 10.1007/s12021-008-9015-0

[7] I.H. Iversen, N. Ghanayim, A. Kübler, N. Neumann, N. Birbaumer, J. Kaiser, A brain-computer interface tool to assess cognitive functions in completely paralyzed patients with amyotrophic lateral sclerosis, *Clinical Neurophysiology*, 119(10): 2214-2223, 2008

[8] G. Pfurtscheller, T. Solis-Escalante, Could the beta rebound in the EEG be suitable to realize a "brain switch"? , *Clinical Neurophysiology*, 120(11): 24-29, 2008

[9] Marios G. Philiastides, Paul Sajda, EEG-Informed fMRI Reveals Spatiotemporal Characteristics of Perceptual Decision Making, *The Journal of Neuroscience*, 27(48):13082-13091, 2007

[10] A. Vuckovic, F. Sepulveda, Quantification and visualisation of differences between two motor tasks based on energy density maps for brain-computer interface applications, *Clinical Neurophysiology*, 119(2): 446-458, 2008

[11] Oliver Tonet, Martina Marinelli, Luca Citi, Paolo Maria Rossini, Luca Rossini, Giuseppe Megali, et. all, Defining brain -machine interface applications by matching interface performance with device requirements, *Journal of Neuroscience Methods, Brain-Computer Interfaces (BCIs)*, 167(1): 91-104, 2008

[12] Hyun K. Kim, Shinsuk Park, Mandayam A. Srinivasan, Developments in brain-machine interfaces from the perspective of robotics, *Human Movement Science*, 28(2): 191-203, 2009

[13] Febo Cincotti, Donatella Mattia, Fabio Aloise, Simona Bufalari, Gerwin Schalk, Giuseppe Oriolo, et. all, Non-invasive brain-computer interface system: Towards its application as assistive technology, Brain Research Bulletin, Special Issue: *Robotics and Neuroscience*, 775(6): 796-803, 2008

[14] F. Galán, M. Nuttin, E. Lew, P.W. Ferrez, G. Vanacker, J. Philips, J. del R. Millán, A brain-actuated wheelchair: Asynchronous and non-invasive Brain-computer interfaces for continuous control of robots, *Clinical Neurophysiology*, 119(9): 2159-2169, 2008

[15] Chia-Lin Chang, Zhanpeng Jin, Hou-Cheng Chang, Allen C. Cheng, From neuromuscular activation to end-point locomotion: An artificial neural network-based technique for neural prostheses, *Journal of Biomechanics*, 42(8): 982-988, 2009

[16] Febo Cincotti, Donatella Mattia, Fabio Aloise, Simona Bufalari, Laura Astolfi, Fabrizio De Vico Fallani, et. all, High-resolution EEG techniques for brain-computer interface applications, *Journal of Neuroscience Methods, Brain-Computer Interfaces (BCIs)*, 167(1): 31-42, 2008

[17] Valerie Morash, Ou Bai, Stephen Furlani, Peter Lina, Mark Halletta, Classifying EEG signals preceding right hand, left hand, tongue, and right foot movements and motor imageries, *Clinical Neurophysiology*, 119(11): 2570-2578, 2008

[18] Vadim V. Nikulin, Friederike U. Hohlefeld, Arthur M. Jacobs, Gabriel Curio, Quasi-movements: A novel motor-cognitive phenomenon, *Neuropsychologia*, 46(2): 727-742, 2008

[19] Farhad Faradji, Rabab K. Ward, Gary E. Birch, Plausibility Assessment of a 2-State Self-Paced Mental Task-Based BCI Using the No-Control Performance Analysis, *Journal of Neuroscience Methods*, 180(2): 330-339, 2009

[20] Daniel Pérez-Marcos, Jens-Uwe Knote, Reinhard Both, Galina Ivanova, Quantification of cognitive-induced brain activity: An efficient method for online applications, *Computers in Biology and Medicine*, 38(11-12): 1194-1202, 2008

[21] Seong-eun Ro, Joon Hwan Choi, Taejeong Kim, A new action potential classifier using 3-Gaussian model fitting, Advances in Neural Information Processing (ICONIP 2006) / Brazilian Symposium on Neural Networks (SBRN 2006), *Neurocomputing*, 71(16-18): 3631-3634, 2008

[22] Mingjun Zhong, Fabien Lotte, Mark Girolami, Anatole Lécuyer, Classifying EEG for brain computer interfaces using Gaussian processes, *Pattern Recognition Letters*, 29(3):354-359, 2008

[23] Jo-Anne Ting, Aaron D'Souza, Kenji Yamamoto, Toshinori Yoshioka, Donna Hoffman, Shinji Kakeif, et. all, Variational Bayesian least squares: An application to brain -machine interface data, *Neural Networks*, 21(8): 1112-1131, 2008

[24] Tonio Balla, Evariste Demandt, Isabella Mutschler, Eva Neitzel, Carsten Mehring, Klaus Vogt, et. all, Movement related activity in the high gamma range of the human EEG, *NeuroImage*, 41(2): 302-310, 2008

[25] Shang-Ming Zhou, John Q. Gan, Francisco Sepulveda, Classifying mental tasks based on features of higher-order statistics from EEG signals in brain-computer interface, *Information Sciences*, 178(6): 1629-1640, 2008

[26] F. Darvas, J.G. Ojemann, L.B. Sorensen, Bi-phase locking - a tool for probing non-linear interaction in the human brain, *NeuroImage*, 46(1): 123-132, 2009

[27] Wei-Yen Hsu, Yung-Nien Sun, EEG-based motor imagery analysis using weighted wavelet transform features, *Journal of Neuroscience Methods*, 175(2): 310-318, 2009

[28] Wu Ting, Yan Guo-zheng, Yang Bang-hua, Sun Hong, EEG feature extraction based on wavelet packet decomposition for brain computer interface, *Measurement*, 41(6): 618-62, 2008

[29] Dean Cvetkovic, Elif Derya Übeyli, Irena Cosic, Wavelet transform feature extraction from human PPG, ECG, and EEG signal responses to ELF PEMF exposures: A pilot study, *Digital Signal Processing*, 18(5): 861-874, 2008

[30] Vahid Abootalebi, Mohammad Hassan Moradib, Mohammad Ali Khalilzadeh, A new approach for EEG feature extraction in P300-based lie detection, *Computer Methods and Programs in Biomedicine*, 94(1): 48-57, 2009

[31] R. Murat Demirer, Mehmet Sirac Ozerdem, Coskun Bayrak, Classification of imaginary movements in ECoG with a hybrid approach based on multi-dimensional Hilbert-SVM solution, *Journal of Neuroscience Methods*, 178(1): 214-218, 2009

[32] W. A. Chaovalitwongse, P. M. Pardalos, On the time series support vector machine using dynamic time warping kernel for brain activity classification, *Cybernetics and Systems Analysis*, 44(1): 125-138, 2008

[33] Andreas Trřllund Boy, Ulrik Qvist Kristiansen, Martin Billinger, Omar Feix do Nascimento, Dario Farina , Identification of movement-related cortical potentials with optimized spatial filtering and principal component analysis, *Biomedical Signal Processing and Control*, 3(4): 300-304, 2008

[34] Gholamreza Salimi-Khorshidi, Ali Motie Nasrabadi, Mohammadreza Hashemi Golpayegani, Fusion of classic P300 detection methods' inferences in a framework of fuzzy labels, *Artificial Intelligence in Medicine*, 44(3): 247-259, 2008

[35] Vito Logar, Igor Škrjanc, Aleš Belič, Simon Brežan, Blaž Koritnik, Janez Zidar, Identification of the phase code in an EEG during gripping-force tasks: A possible alternative approach to the development of the brain-computer interfaces, *Artificial Intelligence in Medicine*, 44(1): 41-49, 2008

[36] Dzitac S., Felea I., Dzitac I., Vesselenyi T., (2008), An Application of Neuro-Fuzzy Modelling to Prediction of some Incidence in an Electrical Energy Distribution Center, *International Journal of Computers, Communications and Control*, ISSN 1841-9836, 3(S):287-292, 2008

[37] Coyle D, Prasad G, McGinnity T.M., Faster self-organizing fuzzy neural network training and a hyperparameter analysis for a brain-computer interface, *IEEE Trans Syst Man Cybern B Cybern.*, 39(6):1458-71, 2009

[38] Palaniappan R., Paramesran R., Nishida S., Saiwaki N., A new brain-computer interface design using fuzzy ARTMAP, *IEEE Transactions on neural systems and rehabilitation engineering*, ISSN 1534-4320, 10(3): 140-148, 2002

[39] Bang-hua Y., Guo-zheng Y., Ting W., Rong-guo Y., Subject-based feature extraction using fuzzy wavelet packet in brain-computer interfaces, *Signal Processing*, 87(7): 1569-1574, 2007

[40] Vesselenyi T., Dzitac I., Dzitac S., Hora C., Porumb C., Preliminary Issues On Brain-Machine Contextual Communication Structure Development, *Soft Computing Applications, 2009. SOFA '09. 3rd International Workshop on*, ISBN 978-1-4244-5054-1, pp.37-42, IEEE Xplore DOI 10.1109/SOFA.2009.5254882, 2009

[41] Darvishi, S.; Al-Ani, A. Brain-Computer Interface Analysis using Continuous Wavelet Transform and Adaptive Neuro-Fuzzy Classifier, Engineering in Medicine and Biology Society, 2007. EMBS 2007. *29th Annual International Conference of the IEEE*, 3220 - 3223.

[42] Pfurtscheller, G.: EEG event-related desynchronization (ERD) and event-related synchronization (ERS), In *E.* Niedermeyer, F.H. Lopes da Silva (eds.) Electroencephalography: basic principles, clinical applications and related fields, 4th edition, Williams and Wilkins, Baltimore, MD, 958-967, 1999

[43] Zadeh L. A., The Role of Fuzzy Logic in Modeling, Identification and Control, Modeling, *Identification and Control*, 15(3):191-203, 1994

[44] Zadeh L. A., Fuzzy Logic and the Calculus of Fuzzy If-Then Rules, *P*roceedings of SYNAPSE'91, Tokyo, Japan, 1991

[45] Zadeh L. A., Toward extended fuzzy logic-A first step, *Fuzzy Sets and Systems*, 160: 3175-3181, 2009

[46] ***, *MATLAB, User Guide*, MathWorks, 2008

Int. J. of Computers, Communications & Control, ISSN 1841-9836, E-ISSN 1841-9844
Vol. VI (2011), No. 3 (September), pp. 418-427

Uncertain Fractional Order Chaotic Systems Tracking Design via Adaptive Hybrid Fuzzy Sliding Mode Control

T.C. Lin, C.H. Kuo, V.E. Balas

Tsung-Chih Lin
Feng-Chia University, 40724, Taichung, Taiwan
E-mail: tclin@fcu.edu.tw

Chia-Hao Kuo
Ph.D Program in Electrical and Communications Engineering
Feng-Chia University, Taichung, Taiwan
E-mail: peterqo022@hotmail.com

Valentina E. Balas
Aurel Vlaicu University of Arad, Romania
B-dul Revolutiei 77, 310130 Arad, Romania
E-mail: balas@drbalas.ro

Abstract: In this paper, in order to achieve tracking performance of uncertain fractional order chaotic systems an adaptive hybrid fuzzy controller is proposed. During the design procedure, a hybrid learning algorithm combining sliding mode control and Lyapunov stability criterion is adopted to tune the free parameters on line by output feedback control law and adaptive law. A weighting factor, which can be adjusted by the trade-off between plant knowledge and control knowledge, is adopted to sum together the control efforts from indirect adaptive fuzzy controller and direct adaptive fuzzy controller. To confirm effectiveness of the proposed control scheme, the fractional order chaotic response system is fully illustrated to track the trajectory generated from the fractional order chaotic drive system. The numerical results show that tracking error and control effort can be made smaller and the proposed hybrid intelligent control structure is more flexible during the design process.
Keywords: Fractional order chaotic systems; fuzzy logic control, adaptive hybrid control.

1 Introduction in domain

Due mainly to its demonstrated applications in numerous seemingly diverse and widespread fields of science and engineering, fractional calculus has gained considerable popularity and importance during past three decades [1]- [2]. In control system, due to the fact that the theoretical aspects are well established, fractional order controllers are successfully used to enhance the performance of the feedback control loop. It is observed that the description of some systems is more accurate when the fractional derivative is used. Nowadays, many fractional-order differential systems behave chaotically, such as the fractional-order Chua's system [3], the fractional-order Duffing system [4], the fractional-order system, the fractional-order Chen's system [5], the fractional-order cellular neural network [6], the fractional-order neural network [7]. The tracking problem of fractional order chaotic systems is first investigated by Deng and Li [21] who carried out tracking in case of the two fractional Lü systems. Afterwards, they studied chaos tracking of the Chen system with a fractional order in a different manner [22]-[24].

Based on the universal approximation theorem, [9]- [20] (fuzzy logic controllers are general enough to perform any nonlinear control actions) there is rapidly growing interest in systematic

Copyright © 2006-2011 by CCC Publications

design methodologies for a class of nonlinear systems using fuzzy adaptive control schemes. Like the conventional adaptive control, the adaptive fuzzy control is classified into direct and indirect fuzzy adaptive control categories [9], [17]- [19]. A direct adaptive fuzzy controller uses fuzzy logic systems as controller in which linguistic fuzzy control rules can be directly incorporated into the controller. On the other hand, an indirect adaptive fuzzy controller uses fuzzy descriptions to model the plant in which fuzzy IF-THEN rules describing the plant can be directly incorporated into the indirect fuzzy controller. Moreover, a hybrid adaptive fuzzy controller can be constructed using a weighting factor to sum together the control efforts from indirect adaptive fuzzy controller and direct adaptive fuzzy controller.

Although the concept of sliding mode control (SMC) and the theory of fractional order system are well known, their integration, fractional sliding mode control, is an interesting filed of research dwelt on this paper with some applications [8]. The motivation of this paper stands on two driving forces: One, most systems in the reality display behavior characterized best in time domain of fractional operators, the other, the uncertainties on the process dynamics can appropriately be alleviated by utilizing SMC technique.

In this paper, by combining the approximate mathematical model, linguistic model description and linguistic control rules into a single adaptive fuzzy controller, an adaptive hybrid fuzzy controller is proposed to achieve prescribed tracking performance of fractional order chaotic systems. A new adaptive hybrid fuzzy SMC algorithm incorporated Lyapunov stability criterion is proposed so that not only the stability of adaptive fuzzy control system is guaranteed but also the influence of the approximation error and external disturbance on the tracking error can be attenuated to an arbitrarily prescribed level.

This paper is organized as follows: In section2, an introduction to fractional derivative and its relation to the approximation solution will be addressed. Section 3 generally proposes adaptive hybrid fuzzy SMC of uncertain fractional order systems in presence of uncertainty and its stability analysis. In Section 4, application of the proposed method on fractional order expression chaotic system is investigated. Finally, the simulation results and conclusion will be presented in Section 5.

2 Basic definition and preliminaries for fractional order systems

The concept of fractional calculus is popularly believed to have steamed from a question raised in the year 1695 by Marquis de L'Hoptial to Gottfried Wilhelm Leibniz. It is a generalization of integration and differentiation to non-integer order fundamental operator, denoted by ${}_aD_t^q$, where a and t are the limits of the operator. This operator is a notation for taking both the fractional integral and functional derivative in a single expression defined as [1]

$$ {}_aD_t^q = \begin{cases} \frac{d^q}{dt^q}, & q > 0 \\ 1 & q = 0 \\ \int_t^a (d\tau)^{-q}, & q < 0 \end{cases} \tag{1} $$

There are some basic definitions for the general fractional and the commonly used definitions are Grunwald-Letnikov and Riemann-Liouville [1]. The Grunwald-Letnikov definition is expressed as

$$ {}_aD_t^q f(t) = \lim_{h \to 0} \sum_{j=0}^{\left[\frac{t-a}{h}\right]} (-1)^j \begin{pmatrix} a \\ b \end{pmatrix} f(t - jh) \tag{2} $$

where [.] is the integer part. The simplest and easiest definition is Riemann-Liouville definition given as

$$ {}_aD_t^q f(t) = \frac{1}{\Gamma(n-q)} \frac{d^n}{dt^n} \int_0^t \frac{f(\tau)}{(t-\tau)^{q-n+1}} d\tau \tag{3} $$

where n is the first integer which is not less q, i.e., $n-1 < q < n$, and Γ is the Gamma function.

The numerical simulation of a fractional differential equation is not simple as that of an ordinary differential equation. In this paper, the algorithm which is an improved version of Adams-Bashforth-Moulton algorithm to find an approximation for fractional order systems based on predictor-correctors is given. Consider the following differential equation

$$ {}_aD_t^q y(t) = r(y(t), t), \quad 0 \le t \le T \ and \ y^{(k)}(0) = y_o^{(k)}, k = 0, 1, 2, ..., m-1 \tag{4} $$

where

$$ {}_aD_t^q y(t) = \begin{cases} \frac{1}{\Gamma(m-q)} \int_0^t \frac{f^{(m)}(\tau)}{(t-\tau)^{q-m+1}} d\tau, & m-1 < q < m \\ \\ \frac{d^m}{dt^m} y(t), & q = m \end{cases} \tag{5} $$

and m is the first integer larger the q. The solution of the equation (4) is equivalent to Volterra integral equation [1] described as

$$ y(t) = \sum_{k=0}^{[q]-1} y_0^{(k)} \frac{t^k}{k!} + \frac{1}{\Gamma(q)} \int_0^t (t-\lambda)^{q-1} r(y(\lambda), \lambda) d\lambda \tag{6} $$

Let h=T/N, $t_n = nh$, n=0,1,2,$\cdots$ N . Then (6) can be discretized as follows.

$$ y_h(t_{n+1}) = \sum_{k=0}^{[q]-1} y_0^{(k)} \frac{t_{n+1}^k}{k!} + \frac{h^q}{\Gamma(q+2)} r(y_h^p(t_{n+1}), t_{n+1}) + \frac{h^q}{\Gamma(q+2)} \sum_{j=0}^{n} a_{j,n+1} r(y_h(t_j), t_j) \tag{7} $$

where predict value $y_h^p(t_{n+1})$ is determined by

$$ y_h^p(t_{n+1}) = \sum_{k=0}^{[q]-1} y_0^{(k)} \frac{t_{n+1}^k}{k!} + \frac{h^q}{\Gamma(q)} \sum_{j=0}^{n} b_{j,n+1} r(y_h(t_j), t_j) \tag{8} $$

and

$$ a_{j,n+1} = \begin{cases} n^{q+1} - (n-q)(n+1)^q, & j = 0 \\ (n-j+2)^{q+1} + (n-j)^{q+1} - 2(n-j+1)^{q+1} & 1 \le j \le n \\ 1 & j = n+1 \end{cases} \tag{9} $$

$$ b_{j,n+1} = \frac{h^q}{q}((n+1-j)^q - (n-j)^q) \tag{10} $$

The approximation error is given as

$$ \max_{j=0,1,2,\cdots N} |y(t_j) - y_h(t_j)| = O(h^p) \tag{11} $$

where p=min(2,1+q). Therefore, the numerical solution of a fraction order chaotic system discussed in this paper can be obtained by applying the above mentioned algorithm.

3 Adaptive hybrid fuzzy sliding mode control of uncertain fractional order chaotic systems

In this section, we study adaptive hybrid fuzzy tracking control of uncertain fractional order chaotic systems, i.e., to force output trajectory which is obtained by the algorithm mentioned in section 2 of the response system to track output trajectory of the drive system.

Consider a fractional order chaotic dynamic system

$$x^{(nq)} = f(\underline{x},t) + g(\underline{x},t)u + d(t), \quad y = x_1 \tag{12}$$

where $\underline{x} = [x_1, x_2, ..., x_n]^T = [x, x^{(q)}, x^{(2q)}, ..., x^{((n-1)q)}]^T$ is the state vector, $f(\underline{x},t)$ and $g(\underline{x},t)$ are unknown but bounded nonlinear functions which express system dynamics, d(t) is the external bounded disturbance, $|d(t)| \leq D$, and u(t) is the control input. The control objective is to force the system output y to follow a bounded reference signal y_d which is the output trajectory of a drive system, under the constraint that all signals involved must be bounded. To begin with, the reference signal vector $\underline{y}_d$ and the tracking error vector $\underline{e}$ will be defined as

$$\underline{y}_d = \left[y_d, y_d^{(q)}, ..., y_d^{((n-1)q)}\right]^T \in R^n,$$

$$\underline{e} = \underline{y}_d - \underline{x} = \left[e, e^{(q)}, e^{(2q)}, ..., e^{((n-1)q)}\right]^T \in R^n, e^{(iq)} = y_d^{(iq)} - x^{(iq)} = y_d^{(iq)} - y^{(iq)}$$

In general, in the space of the error state a sliding surface is defined by

$$s(\underline{x},t) = -(\underline{k}\underline{e}) = -\left(k_1 e + k_2 e^{(q)} + ... + k_{n-1} e^{(n-2)q} + e^{(n-1)q}\right) \tag{13}$$

where $\underline{\underline{k}} = [k_1, k_2, ..., k_{n-1}, 1]$ in which the k_i's are all real and are chosen such that $h(r) = \sum_{i=1}^{n} k_i r^{(i-1)q}, k_n = 1$ is a Hurwitz polynomial where r is a Laplace operator. The tracking problem will be considered as the state error vector $\underline{e}$ remaining on the sliding surface $s(\underline{x},t) = 0$ for all $t \geq 0$. The sliding mode control process can be classified into two phases, the approaching phase with $s(\underline{x},t) \neq 0$ and the sliding phase with s(x,t)= 0 for initial error $\underline{e}(0) = 0$. In order to guarantee that the trajectory of the state error vector $\underline{e}$ will translate from the approaching phase to the sliding phase, the sufficient condition

$$s(\underline{x},t)\dot{s}(\underline{x},t) \leq -\eta > 0 \tag{14}$$

must be satisfied. Two type of control law must be derived separately for those two phases described above. In the sliding phase, it implies $s(\underline{x},t) = 0$ and $s^{(q)}(\underline{x},t) = 0$. In order to force the system dynamics to stay on the sliding surface, the equivalent control $\overline{u}$ can be derived as follows:

If $f(\underline{x},t)$ and $g(\underline{x},t)$ are known and free of external disturbance, i.e., d(t)=0, taking the derivative of the sliding surface with respective to time, we get

$$s^{(q)} = -\left(\sum_{i=1}^{n-1} c_i e^{(iq)} + e^{(nq)}\right) = -\left(\sum_{i=1}^{n-1} k_i e^{(iq)} + y_d^{(nq) - y^{(nq)}}\right)$$

$$= -\left(\sum_{i=1}^{n-1} k_i e^{(iq) - f(\underline{x},t) - g(\underline{x},t)} u_{eq}\right) - y_d^{(n)} = -\sum_{i=1}^{n-1} k_i e^{(i)} + f(\underline{x}) + b(\underline{x})\overline{u}(t) - x_d^{(n)} = 0 \tag{15}$$

Therefore, the equivalent control can be obtained as

$$\overline{u} = \frac{1}{g(\underline{x},t)}\left(\sum_{i=1}^{n-1} k_i e^{(iq)} - f(\underline{x},t) + y_d^{(nq)}\right) \tag{16}$$

On the contrary, in the approaching phase, $s(\underline{x},t) \neq 0$, an approaching-type control u_{ap} must be added in order satisfy the sufficient condition (4) and the complete sliding mode control will be expressed as

$$u = \overline{u} - u_{ap}, \quad u_{ap} = \psi_h sgn(s) \tag{17}$$

where $\psi_h \geq \eta > 0$.

To obtain the sliding mode control (17), the system functions $f(\underline{x},t), g(\underline{x},t)$ and switching parameter ψ_h must be known in advance. However, $f(\underline{x},t)$ and $g(\underline{x},t)$ are unknown and external disturbance, $d(t) \neq 0$, the ideal control effort (16) cannot be implemented. We replace $f(\underline{x},t)$, $g(\underline{x},t)$ and u_{ap} by the fuzzy logic system $f(\underline{x}|\underline{\theta}_f)$, $g(\underline{y}|\underline{\theta}_g)$ and $h(s|\theta_h)$ in specified form as [9], [17]-[19], i.e.,

$$f(\underline{x}|\underline{\theta}_f) = \xi^T(\underline{x})\underline{\theta}_f, g(\underline{x}|\underline{\theta}_g) = \xi^T(\underline{x})\underline{\theta}_g, \quad h(s|\underline{\theta}_h) = \emptyset^T(s)\underline{\theta}_h \tag{18}$$

let $|h(s|\underline{\theta}_h)| = D + \psi_h + \omega_{max}$ when $s(\underline{x},t)$ is outside the boundary layer. Here the fuzzy basis functions $\xi(\underline{x})$ and $\emptyset(s)$ depend on the fuzzy membership functions and is supposed to be fixed, while θ_f, θ_g and θ_h are adjusted by adaptive laws based on Lyapunov stability criterion. Therefore, depending on plant knowledge and control knowledge, a hybrid adaptive fuzzy controller can be constructed by incorporating both fuzzy description and fuzzy control rules using a weighting factor α to combine the indirect adaptive fuzzy controller and the direct adaptive fuzzy controller. Based on the trade-off between plant knowledge and control knowledge, the weighting factor $\alpha \in [1,1]$ can be adjusted. Therefore, the total control effort can be expressed as

$$u_c = \alpha u_i + (1-\alpha)u_d \tag{19}$$

where the direct adaptive fuzzy controller u_d and the indirect adaptive fuzzy controller u_i are given as follows:

$$u_d(\underline{x}) = u_D(\underline{x}|\underline{\theta}_D) - \frac{h(s|\underline{\theta}_h)}{g(\underline{x},t)} \; and \; u_i(\underline{x}) = \frac{1}{g(\underline{x}|\underline{\theta}_g)}\left[\sum_{i=1}^{n-1} k_i e^{(iq)} + y_d^{(nq)} - f(\underline{x},|\underline{\theta}_f) - h(s|\underline{\theta}_h)\right] \tag{20}$$

where $u_D(\underline{x}|\underline{\theta})$ is obtained by fuzzy logic system specified as

$$u_D(\underline{x}|\underline{\theta}_D) = \xi^T(\underline{x})\underline{\theta}_D \tag{21}$$

The optimal parameter estimations $\underline{\theta}_f^*, \underline{\theta}_g^*$, $\underline{\theta}_h^*$ and $\underline{\theta}_D^*$ are defined as

$$\underline{\theta}_f^* = \arg\; min_{\underline{\theta}_f \in \Omega_f}\left[\sup_{x\in\Omega_x} |f(\underline{x}|\underline{\theta}_f) - f(\underline{x},t)|\right], \; \underline{\theta}_g^* = \arg min_{\underline{\theta}_g \in \Omega_g}\left[\sup_{x\in\Omega_x} |g(\underline{x}|\underline{\theta}_g) - g(\underline{x},t)|\right]$$

$$\underline{\theta}_D^* = \arg\; min_{\underline{\theta}_D \in \Omega_D}\left[\sup_{x\in\Omega_x} |u_D(\underline{x}|\underline{\theta}_g) - \overline{u}|\right], \; \underline{\theta}_h^* = \arg min_{\underline{\theta}_h \in \Omega_h}\left[\sup_{\underline{x}\in\Omega_{\underline{x}}} |h(s|\underline{\theta}_h) - u_{ap}|\right]$$

where $\Omega_f, \Omega_g, \Omega_D$ and $\Omega_{\underline{x}}$ are constraint sets of suitable bounds on $\underline{\theta}_f, \underline{\theta}_g, \underline{\theta}_h^*, \underline{\theta}_D$ and x respectively and they are defined as $\Omega_f = \{\underline{\theta}_f | \underline{\theta}_f| \leq M_f\}$, $\Omega_g = \{\underline{\theta}_g | \underline{\theta}_g| \leq M_g\}$, $\Omega_D = \{\underline{\theta}_D | \underline{\theta}_D| \leq M_D\}$, $\Omega_h = \{\underline{\theta}_h | \underline{\theta}_h| \leq M_h\}$ and $\Omega_x = \{x||x| \leq M_x\}$, where M_f, M_g, M_D, M_h and are positive constants.

By using (20), (21), sliding surface equation (15) can be rewritten as

$$\begin{aligned} s^{(q)} = \omega + \alpha \left[f(\underline{x}|\underline{\theta}_f^*) - f(\underline{x}|\underline{\theta}_f)\right] + \alpha \left[g(\underline{x}|\underline{\theta}_g^*) - g(\underline{x}|\underline{\theta}_g)\right] u_i - (1-\alpha)h(\underline{s}|\underline{\theta}_h^*) \\ -\alpha h(\underline{s}|\underline{\theta}_h) - (1-\alpha)g(\underline{x})\left[u_D(\underline{x}|\underline{\theta}_D^*) - u_D(\underline{x}|\underline{\theta}_D)\right] + \alpha h(\underline{s}|\underline{\theta}_h^*) - \alpha h(\underline{s}|\underline{\theta}_h^*) \\ +(1-\alpha)h(\underline{s}|\underline{\theta}_h^*) - (1-\alpha)h(\underline{s}|\underline{\theta}_h^*) + d(t) \end{aligned} \tag{22}$$

where the minimum approximation errors is defined as

$$\omega = \alpha \left[f(\underline{x}) - f(\underline{x}|\underline{\theta}_f^*)\right] + \alpha \left[g(\underline{x}) - g(\underline{x}|\underline{\theta}_g^*)\right] u_i + (1-\alpha)\left[u_D(\underline{x}|\underline{\theta}^*) - u_D\right] \tag{23}$$

If $\underline{\tilde{\theta}}_f = \underline{\theta}_f - \underline{\theta}_f^*$, $\underline{\tilde{\theta}}_g = \underline{\theta}_g - \underline{\theta}_g^*$ and, $\underline{\tilde{\theta}}_D = \underline{\theta}_D - \underline{\theta}_D^*$, we have

$$\begin{aligned} s^{(q)} = -(1-\alpha)h(\underline{s}|\underline{\theta}_h^*) + \omega - \alpha\underline{\tilde{\theta}}_h^T \emptyset(s) - \alpha\underline{\tilde{\theta}}_f^T \xi(\underline{x}) - \alpha\underline{\tilde{\theta}}_g^T \xi(\underline{x})u_i \\ +(1-\alpha)g(\underline{x})\underline{\tilde{\theta}}_D^T \xi(\underline{x}) - \alpha h(\underline{s}|\underline{\theta}_h^*) - (1-\alpha)\underline{\tilde{\theta}}_h^T \emptyset + d(t) \end{aligned} \tag{24}$$

Following the proceeding consideration, the following theorem can be obtained.

Theorem: Consider the fractional order SISO nonlinear chaotic system (12) with control input (19), if the fuzzy-based adaptive laws are chosen as

$$\theta_f^{(q)} = r_1 s\xi(\underline{x}), \quad \theta_g^{(q)} = r_2 s\xi(\underline{x})u_i, \quad \theta_D^{(q)} = r_3 s\emptyset(\underline{s}) \ and \ \theta_h^{(q)} = -r_4 s g(\underline{x})\xi(\underline{x}) \tag{25}$$

where $r_i > 0, i = 1 \sim 4$. Then, the overall adaptive scheme guarantees the global stability of the resulting closed-loop system in the sense that all signals involved are uniformly bounded and the tracking error will converge to zero asymptotically.

Proof: In order to analyze the closed-loop stability, the Lyapunov function candidate is chosen as

$$V = \frac{1}{2}s^2 + \frac{\alpha}{2r_1}\underline{\tilde{\theta}}_f^T\underline{\tilde{\theta}}_f + \frac{\alpha}{2r_2}\underline{\tilde{\theta}}_g^T\underline{\tilde{\theta}}_g + \frac{\alpha}{2r_4}\underline{\tilde{\theta}}_h^T\underline{\tilde{\theta}}_h + \frac{(1-\alpha)}{2r_3}\underline{\tilde{\theta}}_D^T\underline{\tilde{\theta}}_D \frac{(1-\alpha)}{2r_4}\underline{\tilde{\theta}}_h^T\underline{\tilde{\theta}}_h \tag{26}$$

Taking the derivative of the (26) with respect to time, we get

$$\begin{aligned} V^{(q)} &= ss^{(q)} + \frac{\alpha}{r_1}\underline{\tilde{\theta}}_f^T\underline{\tilde{\theta}}_f^{(q)} + \frac{\alpha}{r_2}\underline{\tilde{\theta}}_g^T\underline{\tilde{\theta}}_g^{(q)} + \frac{\alpha}{r_4}\underline{\tilde{\theta}}_h^T\underline{\tilde{\theta}}_h^{(q)} + \frac{(1-\alpha)}{r_3}\underline{\tilde{\theta}}^T\underline{\tilde{\theta}}^q + \frac{(1-\alpha)}{r_4}\underline{\tilde{\theta}}_h^T\underline{\tilde{\theta}}_h^{(q)} \\ &= -(1-\alpha)sh\left(\underline{s}|\underline{\theta}_h^*\right) + s\omega - \alpha s\underline{\tilde{\theta}}_h^T\emptyset(s) - \alpha s\underline{\tilde{\theta}}_h^T\xi(\underline{x}) - \alpha s\underline{\tilde{\theta}}_g^T\xi(\underline{x})u_i + (1-\alpha)sg(\underline{x})\underline{\tilde{\theta}}^T\xi(\underline{x}) \\ &-ash\left(\underline{s}|\underline{\theta}_h^*\right) - (1-\alpha)s\underline{\tilde{\theta}}_h^T\emptyset + sd(t) + \frac{\alpha}{r_1}\underline{\tilde{\theta}}_f^T\underline{\tilde{\theta}}_f^{(q)} + \frac{\alpha}{r_2}\underline{\tilde{\theta}}_g^T\underline{\tilde{\theta}}_g^{(q)} + \frac{\alpha}{r_4}\underline{\tilde{\theta}}_h^T\underline{\tilde{\theta}}_h^{(q)} + \frac{(1-\alpha)}{r_3}\underline{\tilde{\theta}}^T\underline{\tilde{\theta}}^q + \frac{(1-\alpha)}{r_4}\underline{\tilde{\theta}}_h^T\underline{\tilde{\theta}}_h^{(q)} \\ &\leq \frac{\alpha}{r_1}\underline{\tilde{\theta}}_f^T\left(\underline{\tilde{\theta}}_f^{(q)} - r_1 s\xi(\underline{x})\right) + \frac{\alpha}{r_2}\underline{\tilde{\theta}}_g^T\left(\underline{\tilde{\theta}}_g^{(q)} - r_2 s\xi(\underline{x})u_i\right) + \frac{1}{r_4}\underline{\tilde{\theta}}_h^T\left(\underline{\tilde{\theta}}_h^{(q)} - r_4 s\emptyset(\underline{s})\right) - \alpha s(D+\eta)sgn(s) \\ &+\frac{1-\alpha}{r_3}\underline{\tilde{\theta}}^T\left(\underline{\tilde{\theta}}^{(q)} + r_3 sg(\underline{x})\xi(\underline{x})\right) - (1-\alpha)s(D+\psi_h)sgn(s) + sd(t) + s\omega \end{aligned} \tag{27}$$

From the robust compensator u_a and the fuzzy-based adaptive laws are given (25), after simple manipulation, we have

$$V^{(q)} \leq s\omega - s\psi_h sgn(s) = s\omega - |s|\psi_h \tag{28}$$

Using the corollary of Barbalat's Lemma [16]-[19], we have $\lim_{t\to\infty}|s(\underline{x},t)| = 0$. Therefore, $\lim_{t\to\infty}|e(t)| = 0$. The proof is completed.

4 Simulation example

In this section, we will apply our adaptive hybrid fuzzy sliding mode controller to force the fractional order chaotic gyro response system to track the trajectory of the fractional order chaotic gyro drive system.

Example: The fractional order chaotic gyro drive and response systems are given as follows:

Drive System:

$$\begin{cases} y_1^{(q)} = y_2 \\ y_2^{(q)} = -100\left(\frac{y_1}{4}\right) + \frac{y_1^3}{12} - 0.5y_2 - 0.05y_2^3 + \sin(y_1) + 35.5\sin(2t)y_1 - \frac{x_1^3}{6} + d(t) \end{cases}$$

Response System:

$$\begin{cases} x_1^{(q)} = x_2 \\ x_2^{(q)} = -100\left(\frac{x_1}{4}\right) + \frac{x_1^3}{12} - 0.7x_2 - 0.08x_2^3 + \sin(x_1) + 33\sin(2t)x_1 - \frac{x_1^3}{6} + \Delta f(x_1, x_2) + d(t) + u_c(t) \end{cases}$$

where structured uncertainty $\Delta f(x_1, x_2) = -0.1sin(x_1)$ and external disturbance $d(t) = 0.2cos(\pi t)$. The main objective is to control the trajectories of the response system to track the reference trajectories obtained from the drive system. The initial conditions of drive and response systems are chosen as $[y_1(0), y_2(0)]^T = [1, -1]^T$ and $[x_1(0), x_2(0)]^T = [1.6, 0.8]^T$, respectively. For q=0.95, $\alpha = 0.7$ and all design constants are specified as $k_1 = k_2 = 1, r_1 = 150, r_2 = 20, r_3 = 1, r_4 = 1$ and step size $h = 0.01$. The phase portrait of the drive and response systems for free of control input is given in Figure 1. It is obvious that the tracking performance is bad without control effort supplied to response system.

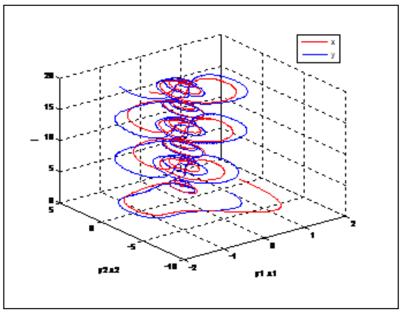


Figure 1: Phase portrait of chaotic drive and response systems

The membership functions for x_i i=1,2 are selected as follows:

$\mu_{F_1^i}(x_i) = exp\left[-0.5\left(\frac{x_i-4}{2}\right)^2\right]$, $\mu_{F_2^i}(x_i) = exp\left[-0.5\left(\frac{x_i-2.7}{2}\right)^2\right]$, $\mu_{F_3^i}(x_i) = exp\left[-0.5\left(\frac{x_i-1.2}{2}\right)^2\right]$,
$\mu_{F_4^i}(x_i) = exp\left[-0.5\left(\frac{x_i}{2}\right)^2\right]$, $\mu_{F_5^i}(x_i) = exp\left[-0.5\left(\frac{x_i+1.2}{2}\right)^2\right]$, $\mu_{F_6^i}(x_i) = exp\left[-0.5\left(\frac{x_i+2.7}{2}\right)^2\right]$,
$\mu_{F_7^i}(x_i) = exp\left[-0.5\left(\frac{x_i+4}{2}\right)^2\right]$,

From the adaptive laws (25)-(28), the control effort of the response system can be obtained as

$$u_c = \alpha u_i + (1-\alpha)u_d \tag{29}$$

Figure 2 shows the trajectories of the states x_i, y_i and x_2, y_2, respectively. Control effort trajectory is given in Figure 3 and phase portrait, tracking performance, of the drive and response

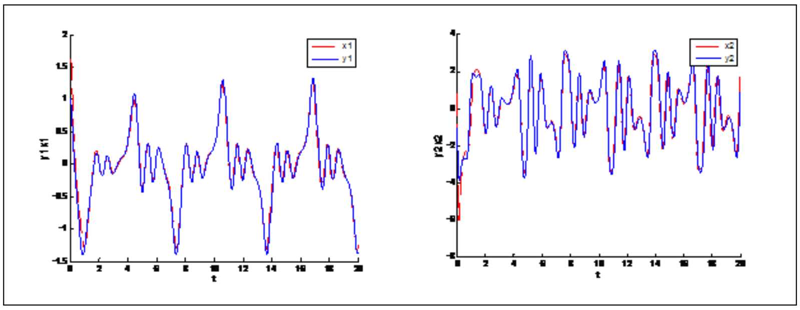


Figure 2: The trajectories of the states x_i, y_i and x_2, y_2

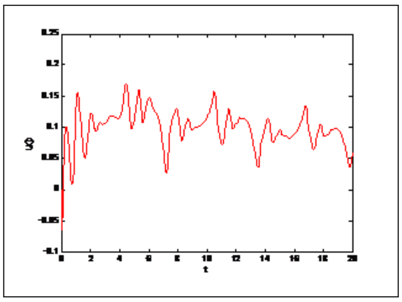


Figure 3: Trajectory of the control effort

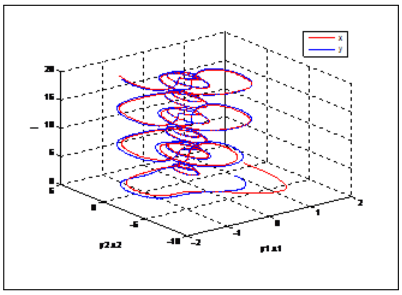


Figure 4: Phase portrait, tracking performance, of the drive and response systems

systems is shown in Figure 4. Trajectory of the sliding surface is given is Figure 5. The maximum value of $V^{(q)}(t)$ is -1.711e-4 which is always negative defined and consequently is stable.

In order to show the robustness of the proposed adaptive hybrid fuzzy sliding mode control, the control effort is activated at 5 second. The phase portrait, tracking performance, of the drive and response systems is given in Figure 6. Figure 7 shows the trajectories of the states x_i, y_i and x_2, y_2 respectively. We can see that a fast tracking of drive and response is achieved as the control effort is activated. Control effort trajectory is given in Figure 8. Trajectory of the sliding surface is given is Figure 9. The maximum value of $V^{(q)}(t)$ is -1.732e-4 which is always negative defined and consequently is stable.

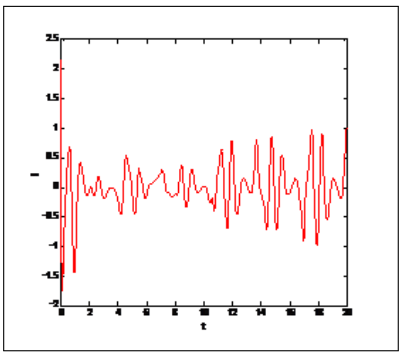


Figure 5: Trajectory of the sliding surface

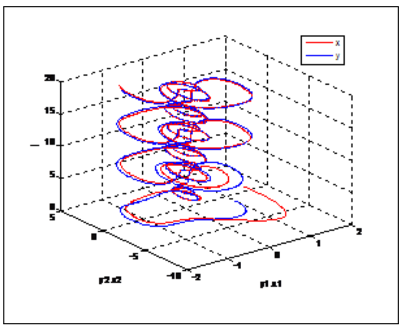


Figure 6: Phase portrait, tracking performance, of the drive and response systems

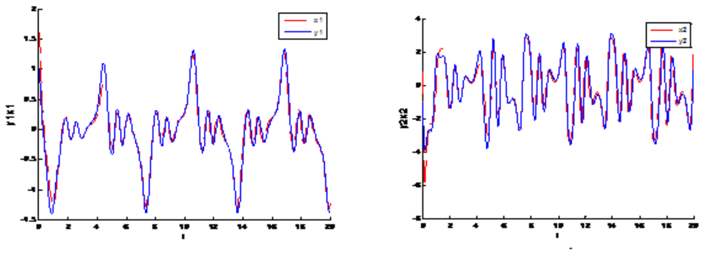


Figure 7: The trajectories of the states x_i, y_i and x_2, y_2

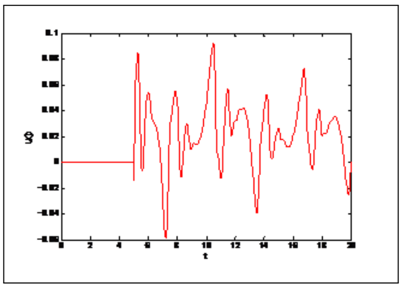


Figure 8: Trajectory of the control effort

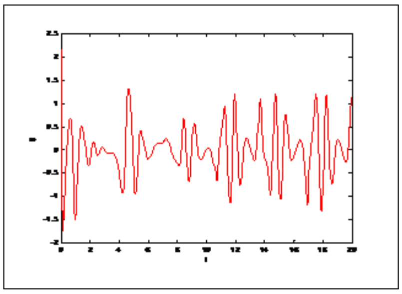


Figure 9: Trajectory of the sliding surface

5 Conclusions

A novel adaptive hybrid fuzzy sliding mode controller is proposed to achieve tracking performance of fractional order chaotic systems in this paper. It is a flexible design methodology by the trade-off between plant knowledge and control knowledge using a weighting factor ? adopted to sum together the control effort from indirect adaptive fuzzy controller and direct adaptive fuzzy controller. Based on the Lyapunov synthesis approach, free parameters of the adaptive fuzzy controller can be tuned on line by output feedback control law and adaptive laws. The simulation example, the output trajectory of the fractional order chaotic response system to tracking the trajectory of the fractional order chaotic drive system, is given to demonstrate the effectiveness of the proposed methodology.

Bibliography

[1] M. S. Tavazoei, M. Haeri, Synchronization of chaotic fractional-order systems via active sliding mode controller, Physica A, Vol. 387, pp. 57-70, 2008.

[2] A.A. Kilbas, H.M. Srivastava and J.J. Trujillo, Theory and applications of fractional differential equations, North-Holland Math. Studies 204, Elsevier, Amsterdam, 2006.

[3] I. Petras, A note on the fractional-order Chua's system, Chaos, Solitons & Fractals, 38 (1), pp. 140-14,7 2008.

[4] X. Gao and J. Yu, Chaos in the fractional order periodically forced complex Duffing's oscillators, Chaos, Solitons & Fractals 26, pp. 1125-1133, 2005.

[5] J.G. Lu and G. Chen, A note on the fractional-order Chen system, Chaos, Solitons & Fractals 27, pp. 685-688, 2006.

[6] P. Arena and R. Caponetto, Bifurcation and chaos in non-integer order cellular neural networks, Int J Bifurcat Chaos 8 (7), pp. 1527-1539, 1998.

[7] Petras I., A note on the fractional-order cellular neural networks. In: Proceedings of the IEEE world congress on computational intelligence, international joint conference on neural networks, Vancouver, Canada; pp.16-21, 2006.

[8] S. H. Hosseninnia, R. Ghaderi, A. Ranjbar N., M. Mahmoudian, S. Momani, Sliding mode synchronization of an uncertain fractional order chaotic system, Journal Computers & Mathematics with Applications, Vol. 59, pp. 1637-1643, 2010.

[9] L.A. Zadeh, Fuzzy logic, neural networks and soft computing. Commun. ACM 37 3, pp. 77-84, 1994.

[10] X. Z. Zhang; Y.N. Wang; X. F. Yuan, H? Robust T-S Fuzzy Design for Uncertain Nonlinear Systems with State Delays Based on Sliding Mode Control, International Journal of Computers Communications & Control, 5(4):592-602, 2010.

[11] Balas, M.M., Balas, V.E., World Knowledge for Control Applications by Fuzzy-Interpolative Systems, International Journal of Computers Communications and Control, Vol. 3, Supplement: Suppl. S, pp. 28-32, 2008.

[12] L. X. Wang, and J. M. Mendel, Fuzzy basis function, universal approximation, and orthogonal least square learning, IEEE Trans. Neural Networks, vol. 3, no. 5, pp. 807-814, 1992.

[13] L. X. Wang, Adaptive Fuzzy Systems and Control: Design and Stability Analysis. Englewood Cliffs, NJ: Prentice-Hall, 1994.

[14] K. Diethelm, An algorithm for the numerical solution of differential equations of fractional order, Elec. Trans. Numer. Anal. 5, pp. 1-6, 1997.

[15] J. L. Castro, "Fuzzy logical controllers are universal approximators," IEEE Trans. Syst., Man, Cybern., 25, pp. 629-635, 1995.

[16] S. S. Ge, T. H. Lee, C. J. Harris, Adaptive Neural Network Control of Robotic Manipulators, World Scientific Publishing Co., Singapore, 1998.

[17] C. H. Wang, T. C. Lin, T. T.. Lee and H. L. Liu, "Adaptive Hybrid Intelligent Control for Unknown Nonlinear Dynamical Systems", IEEE Transaction on Systems, Man, and Cybernetics Part B, 32 (5), October, pp. 583-597, 2002.

[18] C. H. Wang, H. L. Liu and T. C. Lin, "Direct Adaptive Fuzzy-Neural Control with Observer and Supervisory Control for Unknown Nonlinear Systems", IEEE Transaction on Fuzzy Systems, 10(1) , pp. 39-49, 2002.

[19] T. C. Lin, C. H. Wang and H. L. Liu, "Observer-based Indirect Adaptive Fuzzy-Neural Tracking Control for Nonlinear SISO Systems Using VSS and ", Fuzzy Sets and Systems, 143, pp.211-232, 2004.

[20] L. A. Zadeh, Knowledge Representation in Fuzzy Logic, IEEE Trans. Knowledge and Data Engineering, 1 (1), pp. 89-100, 1989.

Int. J. of Computers, Communications & Control, ISSN 1841-9836, E-ISSN 1841-9844
Vol. VI (2011), No. 3 (September), pp. 428-441

Applications of Fuzzy Technology in Business Intelligence

A. Meyer, H.-J. Zimmermann

Andreas Meyer
INFORM Institut f. Operations Research und Management GmbH
Risk & Fraud Division
Pascalstr. 23
52076 Aachen, Germany E-mail: andreas.meyer@inform-software.com

Hans-Jürgen Zimmermann
RWTH, Aachen Institute of Technology
Templer Graben 55
52062 Aachen, Germany
E-mail: zi@or.rwth-aachen.de

Abstract: Fuzzy Set Theory has been developed during the last decades to a demanding mathematical theory. There exist more than 50,000 publications in this area by now. Unluckily the number of reports on applications of fuzzy technology has become very scarce. The reasons for that are manifold: Real applications are normally not single-method-applications but rather complex combinations of different techniques, which are not suited for a publication in a journal. Sometimes considerations of competition my play a role, and sometimes the theoretical core of an application is not suited for publication. In this paper we shall focus on applications of fuzzy technology on real problems in business management. Two versions of fuzzy technology will be used: Fuzzy Knowledge based systems and fuzzy clustering. It is assumed that the reader is familiar with basic fuzzy set theory and the goal of the paper is, to show that the potential of applying fuzzy technology in management is still very large and hardly exploited so far.
Keywords: fuzzy technology in business intelligence, fraud detection, risk assessment, intelligent data mining, fuzzy expert systems

1 Introduction

In his first paper on Fuzzy Sets L.A. Zadeh [20] already mentioned as one motivation of his Theory of Fuzzy Sets: "The fact remains that such imprecisely defined 'classes' play an important role in human thinking, particularly in the domains of pattern recognition, communication of information, and abstraction." This statement has become even more true in the meantime since we have moved from a time of scarce data into the period of data warehouses etc., i.e., into a world of abundance of data, in which people try hard to extract useful information from masses of data ([24]).
From the point of view of applications we still consider as some of the most important goals of fuzzy set theory to extract information from data and model it visually or otherwise in such a way, that people can understand it, communicate it and to model problems adequately. They can use it to solve their problems better than by a purely dichotomous modeling language. In doing this professionally, they can combine the high computing power of EDP with human experience and creativity.
Extremely successful fuzzy models were first used in engineering intelligence in areas such as Fuzzy Control ([13, 24, 26, 30]). Controlling cranes, cement kilns, video cameras, washing machines, ABS, and even subway systems by fuzzy control turned out to be almost sensational. Most

Copyright © 2006-2011 by CCC Publications

of these applications had one feature in common: These were manmade systems, the control of which was often nonlinear (and therefore difficult to model traditionally). However the controls could be decomposed into linear systems by modeling human experience by fuzzy technology ([27]) and one could then determine the adequate parameters, operators ([6]) and membership functions, as well as the defuzzification models, by trial and error. Practitioners loved these controllers because they used predominantly rather basic fuzzy set theory and operators as well as membership functions could be defined rather than determined on the basis of human knowledge. Many of these models have become regular teaching material in control engineering courses.
In the meantime fuzzy set theory was further developed, it became more powerful mathematically and it became more strenuous to learn and understand it. When one started to apply it to business intelligence and to human decision making another problem became visible: Many applications do not permit a trial and error calibration as in fuzzy control because the results of a fuzzy model cannot easily be compared with the results or the behavior of the real system. Think of strategic decisions, of evaluations of long term vulnerability of companies or persons, the determination of the creditworthiness ([29]) of persons or institutions. Here the human knowledge that goes into the fuzzy model has to be modeled properly in advance. That means, that operators ([17]), membership functions, inference methods ([27]) etc. have to properly map the counterparts in the human mind, in which they are very often very context dependent. This is no longer only a mathematical problem but predominantly a problem of psycho-linguistics or similar disciplines ([7, 25]). Unluckily this part of science is much less developed than the mathematics of fuzzy set theory. Hence, in applications one often has to rely again on assumptions rather than on scientific results when modeling operators ([19, 28]), membership functions ([5, 9, 16]) and other parts of fuzzy models. The justification of assumptions, of course, also depends on whether one wants to build descriptive or prescriptive models.
In the following sections we want to show, that, never-the-less, practical applications of fuzzy models in the area of business intelligence are also possible and very useful.

2 RiskShield - A Software Solution Relying on fuzzy technology

This chapter describes the real world software application RiskShield [11] which has predominantly been developed in the years 1999 to 2010. RiskShield is one of the leading tools in the European payments/banking and insurance sector to fight fraud. Well-known air cargo organizations currently consider its possible use to enhance air freight security at acceptable costs. We will report on the practical use of RiskShield in the mentioned industries in Chapter 3.

To briefly describe the solution independently of its purpose and industry in which it is used (payments/banking, insurance, air cargo) we will describe the solution rather in the context of banking, i.e. the term 'transaction' may refer to a financial transaction in the payments sector, to a claim in the insurance sector or to a set of information concerning a parcel in the air cargo sector.

2.1 Software Description

The Basics

RiskShield assesses the risk of incoming transactions individually in real-time within milliseconds. The software works for the human user invisibly in the background. An IT application of the client (e.g. the bank or insurance company) sends every transaction to RiskShield. The software's multi channel approach includes real-time evaluation of additional statistical data and patterns. It builds up internal histories for all relevant entities in the process. Typically

RiskShield installations include 10 to 20 profiles which are constantly updated with any incoming event. Predefined groups can be established to detect any untypical behavior. Fuzzy technology is used to evaluate all different aspects - in many cases the combination of current information with dynamic patterns and histories is needed to identify critical events.
For every incoming transaction RiskShield returns a risk score which triggers the possible decision for (manual) special treatment of the transaction. Depending on configuration the software product can further create a case (alarm) to be further manually investigated and returns the main indicators for the suspicion to bring the human investigators in a more comfortable position to start the special treatment of the transaction (see Figure 1). Every transaction is scored against the same RiskShield decision system to assure consistent results. To evaluate all information provided to the system several rules and patterns of the decision system may contribute to the assessment of an individual transaction.
The type of fuzzy inference used in RiskShield corresponds to the inference described for fuzzyTECH in [15], [30, pp. 255 ff.].

Other Logics and Techniques of RiskShield

RiskShield utilizes several complementary techniques like fuzzy technology [10], matching of patterns against dynamic profiles, recognition of networks (see below) and external lists to ensure a good risk assessment quality. A combination of those techniques forms the decision system within the software solution to assess the risk of a single transaction, i.e. the threat potential which may be caused by the respective transaction.

Matching of patterns against dynamic profiles:
The decision making system can use several profiles which can be defined according to the requirements. Based on these profiles RiskShield recognizes defined patterns in real-time which can comfortably edited by in-house risk and fraud experts without internal IT support or support by the vendor INFORM ([11]).

Recognition of networks and external lists:
Through RiskShield it is possible to detect social and historical relationships based on historical data as well as to use external lists. The meaning of an external list can be defined within the decision system. Therefore these external lists can be used in different ways, e.g. to define positive/white lists or negative/black lists. An example for a white list in Internet banking systems might be a list of account numbers or a list of beneficiary accounts; in credit card issuing systems a black list might consist of BIN-ranges (first x digits of the credit card number); in the air freight area a black list of addressees could be derived from a list of PEP (politically exposed persons) which also receive a special treatment in RiskShield applications for anti money laundering. External lists can not only form white or black lists but also any other list which is relevant for the decision making system. The entries of each list can be associated to a valid-until timestamp, if required.

The combination of these complementary techniques offers the RiskShield user exhaustive possibilities to build a complete safety net dealing with multi-channel fraud attacks and threats.

Response and Alert Management

The risk score RiskShield has returned is used for a decision and an automated workflow control (possibly special treatment of a transaction). This score represents the decision which has been made regarding the current transaction (Approval, Referral, Decline, Hold Movement/ Hold Payment).

Additional scores and information can be defined to be returned by the software solution, e.g.

- Case classes: Groupings of alerts regarding the class of the transaction or the assumed type of fraud.
- Importance: Alerts can be rated regarding their importance so that the most important alerts can be worked with first.
- Hints: A list of rules which had a main impact to the score of the transaction.
- Actions: Criteria can be defined to trigger consecutive actions, e.g. send email/fax/sms, update master data, provide information to CRM, get address scoring.

RiskShield offers the possibility to document the investigation process of the respective alert through a web front end. Additionally, the system offers further analysis functionality to monitor and trace each assessment in detail (drill-down functionality).

RiskShield's use of fuzzy inference combined with matching of patterns against dynamic profiles, recognition of social and historical relations and external lists provide superb fraud prevention qualities for the payments, insurance and probably also for the air cargo security sector. As a result users will obtain almost only relevant hits due to the low false positives output. Research has shown that the automated identification of suspicious transactions for further investigation matches senior analyst evaluations to a very high percentage of cases.

Other Key Features of RiskShield

- Real-time decision making
  - Response and other actions within milliseconds
  - Up to 4,100 single assessments per second
- Processing
  - Automatic threat hit matching
  - Analysis and simulation
  - Analysis and complex simulation with production data (chart, table, and record filtering)
  - All decisions and decision criteria are transparent to the threat analyst
- Implementation
  - Standard interfaces (XML, CSV, JMS)
  - Pre- and Post-Processing plug-ins for quick implementation
  - Software Updates within minutes
  - external support not needed
  - backward compatibility
  - Supports a high secure PCI-DSS compliant implementation (relevant to credit card business)

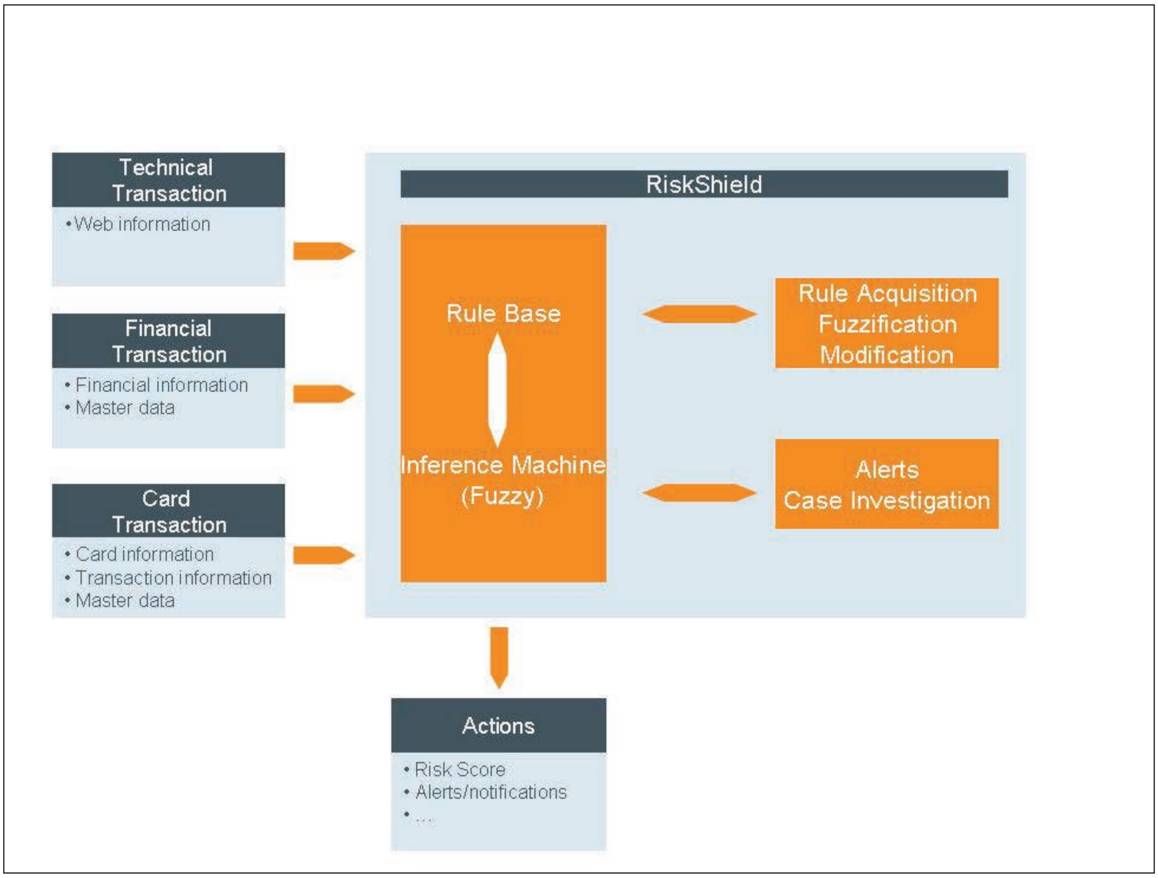


Figure 1: Scheme of RiskShield (input data exemplarily from payments sector)

2.2 Benefits of Fuzzy Logic from Real World Users' Perspective

Real world business users from the industries banking/payments, insurance and air cargo perceive that fuzzy technology as a basis of automated software solutions gives a competitive edge for the following reasons:

- Decision systems using fuzzy logic represent experience of experts adequately:
  Fuzzy logic lets you rely on the experience of people who already understand your system.

- Fuzzy logic is flexible:
  With any given system, it is easy to layer on more functionality without starting again from scratch.

- Fuzzy rules consider fraud/threat and non-fraud/non-threat indicators:
  Like a human expert all aspects of decision making can be integrated into a decision system using fuzzy rules.

- Decision systems using fuzzy logic are easier and more cost-effectively to maintain:
  Before changing an existing rule or adding a new one the impact of the changes to the overall decision system has to be analyzed. As fuzzy logic systems require much less rules than conventional rule based systems the impact of possible changes can be anticipated much easier. Decision systems using fuzzy logic result in shorter reaction times to new threat patterns As explained new rules can be set into productive use faster.

- Fuzzy logic helps minimize false-positives:
  The concept of fuzzy logic represents fuzzy bounds. Thus, the usage of hard bounds which typically result in false-positives is obsolete.

- Decision systems using fuzzy logic are more reliable:
  Conventional rule bases systems have to consider a lot of special cases. One missing case may lead to completely unexpected behavior of the system. Within fuzzy systems the differentiation between special cases is not needed so much.

3 Fraud Prevention and Risk Assessment

In this chapter we report on the practical use of fuzzy logic through the software product RiskShield (see the description in Chapter 2) to prevent fraud in the payments industry (Section 3.1), to detect insurance fraud (Section 3.2). Furthermore we describe its possible use to enhance air freight security at acceptable costs (Section 3.3).

3.1 Payments Industry

The Challenge
In the last few years, fraudsters have become well organized and well equipped to work professionally to attack the payments industry. They prefer to attack payment systems with a low ratio of effort/success. This is in particular the case for online banking platforms where it is possible to use automated scripts like Trojans and viruses that capture the customer online banking session to communicate directly with the attacking server platform (man-in-the-middle and man-in-the-browser attack). Furthermore, phishing attacks are popular as they require not much manual effort.
In the past hardware based IT security measurements have been sufficient for many years to protect Internet banking transactions against fraud, as fraud attacks followed relatively simple patterns. This has changed completely in the last few years. Hardware based IT security products used by end-customers are static regarding its security functionality. That means once they have been issued they won't change their behavior. So when it turns out in practice that the hardware products are not resistant to some online attack banks had to implement an additional security measurement or to replace at high overall costs the hardware products for all customers. Today even banks with the best equipped platforms realize, that relying solely on hardware based protection is too risky (as reaction to new fraud patterns takes too much time in practice) and too expensive.

The Patterns
Different data is required to identify a fraud pattern and protect the account. We need to combine all information we have available in the environment as there are many questions to be answered.
What is the usual type and nature of the customer's transactions? Which beneficiary accounts is he using and what do we know about it? Does the amount of the current transaction exceed the usual payments done by this customer?
Are there other account holders for this account and does this transaction fit to one of their profiles? The transaction might be not typical for the account holder, but may be it is typical for the beneficiary account? What other products does the customer have and how does he use them? Is he a share holder? How many cards does he have?
What do we know concerning the IP address? Is the IP address well known in combination with

this customer? Or is a foreign IP address unexpected for this customer that is always using one of the domestic providers?
What happens during the internet session? Does the IP address change suddenly? Is the usage of the online banking web application as expected?

The Solution
The approach to effectively fight against the current sophisticated attacks should be a combination of expert knowledge ([26]) and a generic software solution. The technical expert must have knowledge of the design, workflow, and architecture of the existing online banking platform and the business expert must have knowledge about the bank's products and customer profiles.
Fortunately, fraud patterns have fingerprints that help to identify them. Once the pattern is identified, a software solution is required that is capable to process all required data sources, to combine the information streams, to identify the sequence of events, to be enhanced on-the-fly, and that is connected to the existing IT environment to stop the payment and to generate alerts.
Fuzzy technology ([21–23]) can easily model expert knowledge about fraud patterns and put it into a software solution. For example, a rule like the following could cross the mind of a human expert and could easily be entered into a fuzzy logic based solution like RiskShield:
In case of a payment from a customer account with an amount of untypical high value, an IP address from a risky region instead of the well known IP address for this account, stop payment for manual intervention.
An effective ant-fraud solution has to address further important issues:

Noise Reduction: An online banking fraud prevention software solution would be useless, if the quality of the generated alerts would be so bad that it led to an inflation of alerts which hide the single valid fraud alarm. So "noise" reduction is required to get the most important alerts only. As RiskShield implementations have shown, this is possible by a most powerful anti fraud pattern definition that filters out undesired alerts down to an acceptable false-positive rate. When evaluating measurable input data RiskShield uses fuzzy bounds instead of hard bounds to minimize false-positives. Particularly in this regard, fuzzy logic is an important underlying concept for the system.

Alerts: Once the payment has been put to hold by RiskShield the case investigation user has time to access the alert data, to verify all the data displayed, to switch the view from the account layer into the customer or beneficiary layer, and to review the profiles and statistical data. This helps to get an overview about the situation in total.

Actions: After investigating the case, the user has the possibility to initiate further actions like releasing or finally stopping the payment by one mouse click. Other actions can be attached to the workflow as well, like an online update of the RiskShield-Server to have this information available for all further decisions. Once all necessary actions have been initiated, the user closes the case which is then sent to the RiskShield reports to check the efficiency of the case queues and users.

Regarding the use of fuzzy technology RiskShield supports its users e.g. through graphical editors to define membership functions, to determine operators, to automatically perform the inference process and to determine a suitable defuzzyfication process.

In the payments industry RiskShield is also in use to detect fraud in other areas ([4, 8]) (so-called channels) including: debit cards and credit cards from the perspective of the issuing bank or from the merchant's or acquirer's perspective, ATM skimming, loan application and others. The software product's multi-channel approach combines findings about a customer across different channels to better assess whether a current transaction from a specific channel has indeed been authorized by the supposed customer.

3.2 Insurance Industry

Due to insurance fraud insurance companies across the globe lose significant amounts of money. The sum of total payments to fraudsters may even by far exceed the profits generated by an insurance company. Insurance experts assess that, depending on the line of business and country, more than 10 percent of all claims notifications have a fraudulent background.

Insurance companies are nowadays battling strongly to keep customer's affections. However, they perceive too high fraud losses to be unacceptable. Regarding their claims handling process insurance companies are typically faced with the challenge to reduce settlement and handling costs of claims on the one hand and to increase customer satisfaction on the other hand. Increase of customer satisfaction includes low insurance premiums and fast, unbureaucratic settlement of legitimate claims. Reduction of settlement and handling costs seems to be a contradictory objective. Because due to the typically high percentage of fraudulent claims, it seems insurance companies have to manually review each single claim very carefully. But this would increase the claims handling costs and would annoy the honest customers awaiting fast settlement of their legitimate claim.

Recently we have recognized the trend among the rather big insurance companies to look for the introduction of some kind of intelligent software which ensures that simultaneously legitimate claims are assigned to a fast settlement track and suspicious claims are assigned to senior fraud experts of the company for further manual review. To do so, the software should automatically assess each single claim and make a decision on how to proceed with the claim. The outputs of the software have to be transparent and traceable. Claims handlers and fraud experts of the insurance company may rather accept a fraud detection software in case they feel confident the software assesses the claims in the same way they would do. That is why expert systems based on fuzzy technology, e.g. RiskShield from INFORM, seem to be well suited to assess insurance claims. In doing so, RiskShield takes both incriminating and exonerating indicators into account. The system balances these indicators as a human expert would do and finally consolidates them to an overall risk score. Expert rules are formulated in a linguistic if-then context, similar to the human language. Furthermore, RiskShield gives additional outputs to indicate the reason of the assessment or to advice what to do next in terms of claims handling. Moreover, RiskShield checks, as a human claims handler would do, for each incoming claim whether the claim related information is already sufficient for a qualified evaluation.

The process of gauging like a human expert when evaluating a claim is illustrated in the following example:

- A new claim, let's see ...
- A new contract ... (incriminating)
- A 2,000 claim, thus a minor loss of property ... (exonerating)
- A brand new Mercedes ... (incriminating)
- No police on scene ... (incriminating)
- The car will be repaired at a partner workshop ... (exonerating)
- A minor damage with no injuries ... (exonerating)
- The policy holder has had only few previous claims ... (exonerating)

- It is a long time customer re other lines of business . . . (exonerating)
- Conclusion: I think we should pay this claim.

INFORM typically delivers its product RiskShield accompanied by consulting services. These services include the provision of customized fraud detection and claims assessment logics which turned out to combine superb fraud detection quality with minimum false positives: Claims marked by RiskShield as suspicious for further investigation match senior analyst evaluations typically in more than 90RiskShield's fraud detection rules consider different types of information like the policy holder, the claims history, the policy itself, the policy history, the loss, and others. Fuzzy technology is useful in the given context to reflect human thinking and to use soft boundaries instead of crisp boundaries which in general will be easily explored and misused by fraudsters.

An insurance claim automatically classified by RiskShield as inconspicuous can be settled faster and more cost-effectively and so strengthening customer loyalty. The resulting saving in time at claims handlers and senior fraud experts for a detailed investigation of claims which have been classified as unusual leads to the reduction of fraud losses which can save a insurance company millions every year.

The fuzzy technology based solution RiskShield is successfully used in more than 35 insurance portfolios across Europe and North America for automatic claim evaluation and fraud risk assessment. Recently the software solution has also been applied to detect medical billing fraud. In a pilot scheme for one of the leading US car insurance companies INFORM analyzed through RiskShield four years of data of all medical billing claims submitted to the insurance company in one US state. INFORM figured out that approximately 18% of all medical bills are to a significant extent incorrect resp. fraudulent.

3.3 Air Freight Security

The latest terror attack attempts using explosive parcels in air freight result in the need to increase security through extended security controls of the parcels at affordable additional costs. A manual reliable security check of every parcel to be delivered cannot be implemented in practice from both organizational and costs perspectives. This subsection describes the idea of applying software-based monitoring solutions which have been proven in the payments/banking sector to fight fraud and to detect irregular patterns, to the air freight sector to identify in real-time suspicious parcels to be delivered. It is reasonable to assume that an air cargo organization already has the required information like sender, addressee, place of posting, freight declaration and other specifications electronically available and possibly stored in one central or a few central applications or data bases. The combination of such pieces of information plus some specific knowledge of the typical behavior of geographical regions as senders (typical previous movements) and the parcels an addressee typically receives should be sufficient from our view to identify suspicious items of freight. To fully automate this process air cargo organizations may consider the use of INFORM's software RiskShield which relies on fuzzy logic as an underlying core technology to keep false alarms acceptably low. Items which have been identified by RiskShield as suspicious will in the air cargo scenario be referred for manual investigation. Air cargo companies currently consider the use of the software solution as it allows economically justifiable close checks to increase airfreight security to a much higher level. In case an air cargo organization already have implemented comprehensive X-ray checks, possibly for every single parcel due to regulatory provisions like from Luftfahrtbundesamt in Germany, a software like RiskShield would nevertheless increase the security level as in practice one possibly cannot trust in the reliability of a single X-ray check, especially if such X-ray checks have been taken abroad in a "risky country".

4 Customer Segmentation in Financial Services

4.1 Crisp vs. Fuzzy Clustering

Banks have different products to offer to their customers ([14, pp. 3-9]), they have (very) different customers for which the products are differently suited, and they have a lot of information (data) of their customers. An efficient marketing of these products requires, as in other markets, a segmentation of the customers such that only suitable products are offered to the appropriate class of customers. Even to day often this segmentation is performed by using single criteria such as age, property, income etc. This obviously does not lead to adequate classes of customers. A better way to classify customers is to use classical clustering, in which a number of attributes can be used to describe a customer. This, however, leaves the problem of objects (e.g. customers) that can not be assigned unequivocally to any particular cluster. With fuzzy cluster analysis ([18]), such as the fuzzy c-means algorithm ([2, 3]), each customer is not assigned to one and only one cluster but rather with different degrees of membership to each class or segment. For outlyers the degrees of membership might not be meaningful, but the use of possibilistic clustering ([12]) cures even this weakness by normalizing the sum of all assigned degrees of membership for each element (customer) to be 1.

4.2 Segmenting Customers of Banks

Let us consider a (smaller) data base of 120,000 records of customers, from which 300 records were samples randomly ([14, p. 3]). The sample records contain all the data fields which are available via the bank's customer information system, such as: age, sex, marital status, number of children, income, assets, information on other banking activities and also the bank's profit margin for each individual customer. The next step would be to select those features from all the characteristic describing a customer, which are relevant for customer segmentation. After discussions with experts the following features may be relevant:

- age (in years)
- net monthly income
- financial assets deposited with the bank
- loans taken out from the bank
- annual marginal gain for the bank. This value denotes the sum which the bank earns or loses from the individual customer per year.

Now the optimal number of clusters has to be determined. A number of criteria have to be satisfied: It should be possible to interpret the different clusters in terms of the expert's vocabulary; the classification should make sense in terms of the different products the bank wants to sell, etc. Technically two measures vary with the number of clusters: the partition coefficient (increasing with the number of clusters) and the partition entropy (falling with the number of clusters). Both tend towards monotone behavior depending on the number of clusters. As the optimal number of clusters can be considered the number where the entropy lies below the rising trend and where the value of the partition coefficient lies above the falling trend. The "optimal" number of clusters is determined by simulation. In our case the best number of clusters is 7 and the cluster centers are as shown in figure 2 (from [14, p. 16]). In this table the clusters have already been "labeled" by experts.

Figure 3 (from [14, p. 17]) shows the membership values of the cluster centers. It is obvious, that a crisp clustering would never have arrived at such a classification. The rest of the customers can now be classified according to the 7 classes. This can be done directly by using DataEngine or normally by the respective c-means algorithm in the software system of the bank. The goal of such a classification run can either be to segment customers in a databank or to assign a new customer to any of these classes. Customers are not exclusively assigned to one cluster, but they may belong to several clusters to different degrees. This corresponds more to reality, even though the labels of the clusters may suggest that the clusters are highly homogeneous with respect to the elements in them. Since the segmentation process does not lead to a once and for all segmentation (customers become older, richer or change their position) a repetition of the process may lead to other clusters and to other customers in existing clusters. This dynamic segmentation ([1]) may also indicate changes in the structure of the customers which otherwise could not be recognized. To avoid misunderstandings it should be mentioned, that the customer segmentation described above is only the core of a larger "Financial Suite" that is used in banks. There is an extensive preprocessing of customer data before segmentation and the results of the segmentation are then used for several purposes, which will be sketched in the next section.

Cluster.fcm (Cluster Centers)

		Age [Years]	Income [$/Month]	Assets [$]	Credit [$]	Contr_Margin [$/Year]
1	Senior_Management	39,47	3.871,825	11.096,456	-4.995,550	44,402
2	Early_Retired	59,92	1.886,855	8.716,017	-2.422,165	51,125
3	Middle_Management	44,57	3.921,299	13.248,014	-92.558,286	546,815
4	Young_Executives	32,07	1.835,261	5.968,807	-3.271,856	16,233
5	Pensioner	76,20	1.497,464	21.746,804	-1.804,408	199,306
6	Apprentices	49,50	3.654,464	30.242,750	-5.923,647	145,061
7	Young_People	10,07	111,358	2.485,742	-238,935	4,772

Figure 2: Centers of the 7 clusters

4.3 Extending Customer Segmentation to other Areas of Banking

One of uses could be the allocation of recourses (sales persons, advisors, etc.) to different classes of customers. After they have been assigned to the different segments, profit potentials can be determined for customer classes or single customers. On the basis of these potentials targets for salespersons can be defined, which can then be used to judge the performance of sales persons in different segments. At the same time sales campaigns can be designed and planned for different products or different segments. So far the analysis (segmentation) and the measures derived from them were static, i.e. they investigated structures at a certain point in time. The management of banks is, however, not only interested in static structures but also in the dynamic behavior of customers. They might want to know, for instance, when a customer will leave the bank, long before that happens. Only then counter measure can be taken. This leads to one major complication: Customers are now no longer defined by numerical or other features (at a certain point in time), but customer behavior is defined by vectors or functions over time. Classical fuzzy cluster methods, however, cluster points and not functions. Hence, either those methods have to be modified or new methods have to be developed. It would exceed the scope of this paper to describe in detail these methods and their application to dynamic problems of

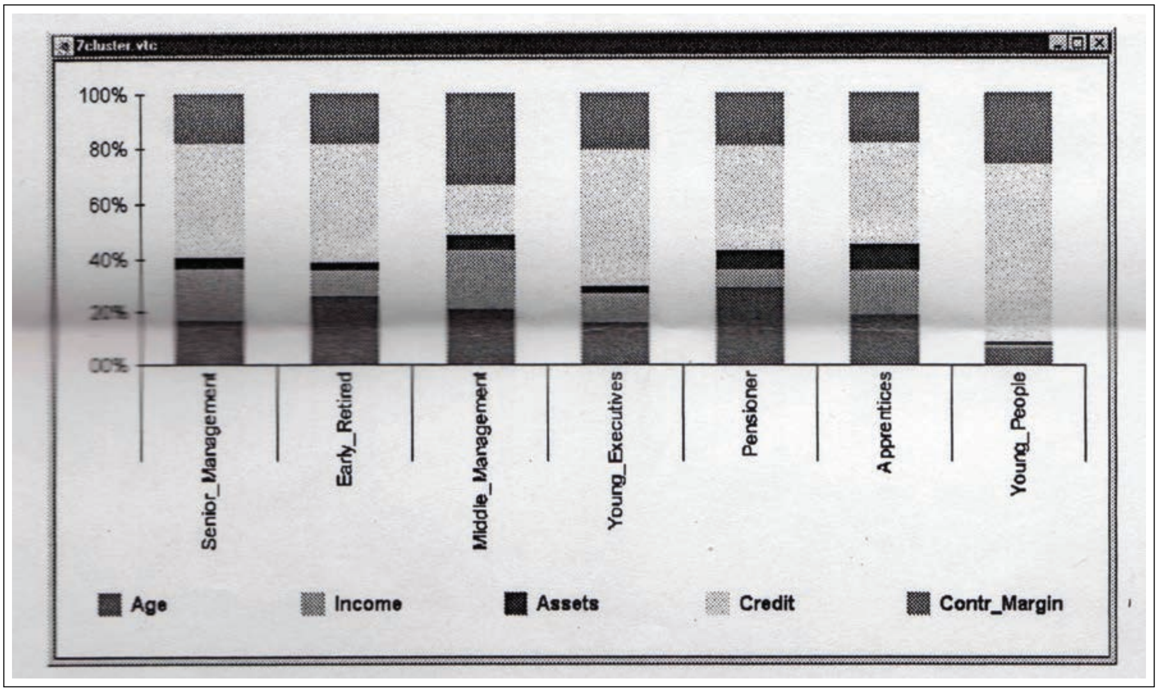


Figure 3: Bar chart of the cluster centers

banks. The interested reader is referred to Angstenberger ([1]) or Zimmermann ([30]) for more detailed descriptions and examples of applications.

5 Conclusions

Mathematically Fuzzy Set Theory has been developed very far. The application of fuzzy control in engineering problems is also quite accepted and wide spread. In business intelligence, however, applications of fuzzy technology are still rather scarce, even though the original motivation and justification of fuzzy set theory was to model and improve human communication. One reason for this might be, that the appropriate modeling of human e is still difficult and context dependent rather than generic. More contributions of psycho-linguistics in this direction are still desirable. Another reason may be, that the successful application of fuzzy technology into management systems requires the acceptance of this technology by managers. This in turn is strengthened by considering management needs, views, and attitudes already when developing applications in this area. Successful pilot applications my also be useful. In this paper such applications are described in the area of financial management. Technologically fuzzy expert systems as well as fuzzy cluster approaches are used. It would have exceeded the scope of this paper to describe the theoretical bases on which the applications rest in detail. We have rather tried to stress those features which are considered to be important from a manager's point of view. All the applications described are real world applications which have shown to be very successful. We hope to have presented attractive problem solutions which can be used as pilot applications and, hence, increase the application of fuzzy technology in the management area.

Bibliography

[1] L. Angstenberger, Dynamic Fuzzy Pattern Recognition with Applications to Finance and Engineering, Kluwer Academic Publishers, Boston, Dodrecht, London, 2001

[2] J.C. Bezdek, Pattern Recognition with fuzzy objective function algorithms, New York, London, 1981

[3] J.C. Bezdek, J.D. Harris, Fuzzy partitions and relations, Fuzzy Sets and Systems 1, 111-127, 1978

[4] B.Bouchon-Meunier, R.R.Yager, L.A.Zadeh (edtrs.) Uncertainty in Intelligent and Information Systems, World Scientific, Singapore, 2000

[5] H. Dishkant, About membership function estimation, Fuzzy Sets and Systems 5, pp.141-147, 1981

[6] D. Dubois, H. Prade, A review of fuzzy set aggregation connectives, Information Science 36, pp.85-121, 1985

[7] Ch. Freksa, Linguistic description of human judgments in expert systems and in the soft sciences, in: M.M. Gupta, E. Sanchez (edtrs), Approximate Reasoning in Decision Analysis, Amsterdam, New York, Oxford , 1982

[8] J.M. Hammerbacher, R.R Yager, The personalization of security:An application of fuzzy set theory, Fuzzy Sets and Systems 5, pp. 1-9, 1981

[9] H.M. Hersh, A. Caramazza, H.H. Brownell, Effects of context on fuzzy membership functions, in:M.M. Gupta et.al.(edtrs) Advances in Fuzzy Set Theory and Applications, Amsterdam, New York, Oxford, 1979

[10] INFORM, fuzzyTECH 5.8, User Manual, available via www.fuzzytech.com, Aachen, 2010

[11] INFORM, RiskShield 4.0, User Manual, available via www.riskshield.com, Aachen, 2011

[12] R.Krishnapuram, J.M.Keller, A possibilistic approach to clustering, IEEE Trans. Fuzzy Systmes 1, 98-110, 1993

[13] E.H. Mamdani, Application of fuzzy logic to approximate reasoning, IEEE Trans.Comput. 26, pp. 1182-1191, 1977

[14] MIT DataEngine Manual 2.1, MIT GmbH, Aachen, 1997

[15] M. Mizumoto, H.-J. Zimmermann, Comparison of fuzzy reasoning methods, Fuzzy Sets and Systems,8, pp.253-283, 1982

[16] A.M. Norwich, L.B. Turksen, A model for the measurement of membership and consequences of its empirical implementation, Fuzzy Sets and Systems 12, pp.1-25, 1984

[17] D. Ruan, A critical study of widely used fuzzy implication operators and the inference rules in fuzzy expert systems, Ph.D. Thesis, Gent 1990

[18] A.Salski, Ecological Modeling and Data Analysis, in: H.-J. Zimmermann (edtr.): Practical Applications of Fuzzy Technologies, Kluwer Academic Publ., Boston 1999, pp.247-266

[19] U. Thole, H.-J. Zimmermann, P. Zysno, On the suitability of minimum and product operators for the intersection of fuzzy sets, Fuzzy Sets and Systems 2, pp.167-180, 1979

[20] L.A. Zadeh, Fuzzy Sets, Information and Control 8, pp.338-353, 1965

[21] L.A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning, Memorandum ERL-M 411, Berkeley 1973

[22] L.A. Zadeh, Outline of a new approach to the analysis of complex systems and decision processes, IEEE Trans.Syst.Man Cybernet. 3, pp. 28-44, 1973

[23] L.A. Zadeh, The role of fuzzy logic in the management of uncertainty in expert systems, Fuzzy Sets and Systems 11, 199-227, 1983

[24] L.A. Zadeh, A New Frontier in Computation - Computation with Information Described in Natural Language, in: From Natural Language to Soft Computing: New Paradigms in Artificial Intelligence .L.A.Zadeh, Dan Tufis, F.G.Filip, I.Dzitac (Edtrs), Editura Academiei Romane, Bäile Felix (Rom.), 2008

[25] H.-J. Zimmermann, Testability and meaning of mathematical models in social sciences, Mathematical Modelling 1, pp.123-139, 1980

[26] H.-J. Zimmermann, Fuzzy Sets, Decision Making, and Expert Systems, Kluwer, Boston, Dodrecht, Lancaster, 1987

[27] H.-J. Zimmermann, Fuzzy set theory - and inference mechanism, in: G.Mitra (edr.) Mathematical models for decision support, Berlin , Heidelberg 1988

[28] H.-J. Zimmermann, P.Zysno, Latent connectives in human decision making, Fuzzy Sets and Systems 4, pp.37-51, 1980

[29] H.-J. Zimmermann, P. Zysno, Decisions and evaluations by hierarchical aggregation of information, Fuzzy Sets and Systems 10, pp.243-266, 1983

[30] H.-J. Zimmermann, Fuzzy Set Theory and its Applications, fourth edition, Boston, Dodrecht, London, 2001

Int. J. of Computers, Communications & Control, ISSN 1841-9836, E-ISSN 1841-9844
Vol. VI (2011), No. 3 (September), pp. 442-457

Decision Support System for Evaluating Existing Apartment Buildings Based on Fuzzy Signatures

G.I. Molnárka, L.T. Kóczy

Gergely I. Molnárka, László T. Kóczy
Széchenyi István University, Györ,
Hungary, H-9026 Györ, Egyetem tér 1.
E-mail: mgergo@sze.hu, koczy@sze.hu

Abstract: In historical district at European cities it is a major problem how to take decision on renovating or replacing existing buildings. This problem is imminent in Budapest (Hungary) in many traditional districts such as the Ferencváros district where we selected a compound area for further examination. By financial aid for the renovation of these buildings which awarded by Municipal Assembly of this district in question there is much uncertainty and confusion concerning how to decide whether or not and how to reconstruct a building where new private owners apply for support.
In this paper we propose a formal evaluation method based on fuzzy signature rule bases (the formal being a special case of L-fuzzy object). Using the available expert knowledge we propose a fuzzy signature model including relevance weights and weighted aggregations for each node and parent node, respectively, so that as a result a single membership value may be calculated for each building in question. Linguistic labels for decision (such as worthless, average, highly valuable, etc.) are generated from the values thus obtained. Such linguistic calculations might be of help for the Municipal Assembly awarding financial support. A complete example wit 26 buildings is presented.
Keywords: building condition evaluation, fuzzy signatures, weighted relevance aggregation operators.

1 Introduction

In the last twenty years due to the transition to democracy changes have happened in the property structure of building stocks, among others, in the Ferencváros ("Francis Town") district of Budapest (capital of Hungary) [1], [2]. The tendency of enhancing the role of private ownership resulted in difficulties of the maintenance process, since the expenses of a general intervention were multiples of the owners' capitals available. In practice, numerous external financial corporations are involved nowadays in supporting renovation and rehabilitation. This solution, however, increases the citizens' level of indebtedness and, consequently, influences the national debt harmfully.

For supporting renovation municipalities offer non-repayable direct financial assistance to the communities of flat owners. This financial aid is achievable annually by application: awarding is decided by the Municipal Assembly. The (non-financial) background of the decision on awarding consists mainly of urban aspects (façade and roof renovation), and is based at present exclusively on the data given by the applicant in the application form.

However these aspects are important and conform to public interest, numerous factors are disregarded and therefore the efficiency of the municipal financial assistance is challengeable. From the point of public view it may also be important to include other attributes of the given building, e.g. physical (static condition, energetic efficiency, etc.) and psychological (technical

Copyright © 2006-2011 by CCC Publications

ageing, architectural, etc.) values. Based on further information an objective ranking of buildings may be created, which can help consider the condition of the applicants' buildings during the process of awarding.

In the present article a proposal for a decision-support tool is outlined that is based on a collection of technical information of the involved buildings. The attributes of buildings can be arranged in a predefined structure, where all information on the buildings can be evaluated systematically. Without any detailed explanation it is obvious that the goal (decision-support tool) and the circumstances (classified data structure, involvement of experts) strongly recommend the application of an expert system. Issuing from the character of data on hand and the observation methods applied, the expert system may well be based on fuzzy set theory [3], especially on the approach of fuzzy signatures which is an efficient tool for describing and characterizing objects with multiple and vague uncertain attributes.

2 Professional Overview of the Examined Building Stock

The Ferencváros district has four main residential areas. The Mid-part is located close to the city centre (see Figure 1).

Figure 1: Overall view of the examined area (Middle Ferencváros)

The urban structure of this area was formed in the most effective period of urbanization in Budapest (1875-1920). The majority of the still existing apartment buildings were built in the first decades of the twentieth century or in the interwar period. Until the mid-nineties this district represented the slum area in the urban texture of the capital.

In the period of the '90s this area came into the attention of international urban designers and politicians by the expressed rehabilitation process. Among others the project manager SEM IX Municipal Company won the FIABCI Prize (a notable award of International Real Estate Federation) in the category of public sector in 1998. Beside the attractive intervention for renovating series of valuable buildings and public areas these changes caused serious relocation of citizens.

The main axis of the rehabilitated urban structure is the Tompa Street that links a busy traffic route (the Middle Ring Street) with a cozy and silent square in the centre of the district. At this semi-intensive area 26 apartment houses can be found on both sides of the 400 m long street. The overview of the area can be seen in Figure 2.

For introducing the apartment houses in the examined area it is worthwhile to describe the buildings by the age of construction.

Figure 2: Tompa Street as the axis of a rehabilitation project

Two buildings (8%) were built in the Age of Historicism (1875-1905, a period using various historical styles), although their constructions had been changed in the next periods. Therefore the global description of these buildings is given together with the next set of buildings.

The larger part of the buildings (14 pieces; 54%) was constructed in the Age of Academic Style (1905-1920). The main attributes of these apartment houses (formerly rental houses) are collected in Table 1.

Three buildings were built in the interwar period (12%) - those reflect to the Age of Modern with their attributes (see Table 2).

The rest (seven buildings; 26%) was constructed during the rehabilitation period (1990-2004).

These apartment houses represent the attributes of this age (see Table 3).

The date of construction can be estimated by an overall look at the external architectural elements applied (cornices, decorations, etc.) and on the shape of the building. With the knowledge of the assortments of each period, a global description of buildings can be done.

However the collected attributes may represent the entire building stock in the examined area, it has to be remarked that every case is different. In Figure 3 a collection of different buildings represents the motley stock.

In spite of the rehabilitation project, some buildings remained in quite bad state with evidences of lack of maintenance (Fig.3a). In case of the larger part of the historic buildings only the façade was renovated (Fig. 3b). The poor energetic performance (see Table 4) and the natural ageing of associated building constructions (roof constructions, chimneys, waterproofing, etc.) may cause serious problems in the near future for the owners. Beside this, the setting of buildings, the application of narrow courtyards produces apartments with bad attributes (lack of sunlight in rear side apartments, badly ventilated courts, lack of vegetation, etc.).

Fig.3.c. represents a good example for good quality renovation and modernization at the same time: the two-storey building was enlarged with two more storeys and a penthouse with retaining the aesthetic values of the original building. In this case the transformation of the building resulted in good life circumstances (it received a common garden with the recently built adjoined block; only the well-orientated apartments were kept).

The buildings from the modern era (Fig 3e, 3f) are supposed to be in fairly good condition. However, although it is invisible outside, the reinforced concrete beams and pillars cause serious geometric thermal bridges (resulting in bad energetic attributes; even because of accumulation of vapor the appearance of mould on the interior surface of external walls is not excludable). It is also remarkable that the reinforced bauxite concrete load-bearing constructions need to be revised periodically (decennially).

The recently built apartment houses (Fig 3g, 3h) correspond to all regulations; the building

Type of Construction (estimated)
footing made of mixed materials (stones, bricks and concrete) traditional brickwork vertical load bearing system (clay brick with lime-mortar) Prussian vault type slab over the cellar (steel beams with clay brick fields) reinforced concrete slabs in intermediate storey (different solutions of early era of reinforced concrete) dowelled beam floor as loft slab (made of timber) traditional timber roof framing with ceramic tiles or zinc plates coupled openings made of timber horizontal waterproof layer under load-bearing walls (paper-based bitumen layer) natural stone skirting and simple rendering external drainage system
Façade Ornaments
geometric ornaments, cornices that articulate the façade simple statues (in low artistic quality)
Type of Installation
inner court ventilated cellar ventilated loft light well differentiated arrangement of apartments by social status
Former Reconstructions
local strengthening of floors, replacement in some cases façade renovation (drainage system, rendering, skirting)

Table 1: Global attributes of rental houses built in period of 1875-1920

Type of Construction (estimated)
pad footing made of concrete and stones post and beam framework made of reinforced concrete with infilling brickwork wall reinforced concrete slabs (in one case: utilization of bauxite as accelerating admixture) flat roof with slag heat insulation and paper-based bitumen layer covering coupled openings made of timber waterproof layer under the entire building (paper-based bitumen layer) natural stone skirting and simple rendering or natural stone cladding and brick façade internal drainage system
Façade Ornaments
-
Type of Installation
integrated arrangement to an enclosed stairway modern spatial structure with bright apartments
Former Reconstructions
flat roof reconstructions façade renovation (broken stone replacements, rendering) periodic revision of reinforced bauxite concrete structures

Table 2: Global attributes of apartment houses built in the interwar era

Type of Construction (estimated)
footing made of concrete post and beam framework made of reinforced concrete with infilling masonry unit reinforced concrete slabs heat insulated pitched roof with concrete tiles plastic openings with insulated glazing waterproof layer under the entire building (reinforced modified bitumen layer) simple rendering external drainage system
Façade Ornaments
simple and valueless decorations
Type of Installation
integrated arrangement to an enclosed stairway modern spatial structure with bright apartments
Former Reconstructions
-

Table 3: Global attributes of apartment houses built in the '90s

	Heat Transmission Coefficient ($U[W/m^2K]$)	
	Mean values (estimated)	**current regulation**
External walls	1.44	0.45
Loft slabs	0.80	0.30
Openings	2.80	1.60

Table 4: Present energetic performance of former rental houses

constructions and the associated elements are supposed to be acceptable in quality. In spite of these facts, it has to be pointed out that the public interest could not be achieved here. From the aspect of cityscape these buildings are negligible.

3 Application of Fuzzy Signatures

In the case of the planned decision support tool the generalization of fuzzy sets first introduced by Zadeh [4] has to be mentioned; Zadeh's early student, Goguen proposed the concept of L-fuzzy sets [5]. L-fuzzy membership grades are elements of an arbitrary lattice L:

$$A : x \to L, \forall x \in X \tag{1}$$

Vector-valued fuzzy sets are introduced in [6]. They are special L-fuzzy sets, where L is the lattice of n-dimensional fuzzy vectors, L=[0,1]n in (1).

Vector valued fuzzy sets assign to each element of X a set of quantitative features rather than a single degree - this way providing additional information about the specific element.

Fuzzy signatures [7] are generalized vector valued fuzzy sets, where each vector component is possibly another nested vector. This generalization can be continued recursively to any finite depth, thus forming a signature with depth m (2).

$$A_s : x \to [a_i]_{i=1}^k, a_i = \begin{cases} [0,1] \\ [a_{ij}]_{j=1}^k \end{cases}, a_{ij} = \begin{cases} [0,1] \\ [a_{ijl}]_{l=1}^k \end{cases}, \forall x \in X \quad (2)$$

Figure 3: Representative buildings of the examined building stock

The structure of fuzzy signatures can be represented both in vector form and also as a tree graph (Figure 4 represents both the vector form and the tree graph of the proposed structure).

Fuzzy signatures can be considered as special, multidimensional constructions that are applicable for storing structured fuzzy data. In this structure the dimensions are interrelated in the meaning that a sub-group of variables determines a character on a higher level. Therefore, complex and interdependent data components can be described and evaluated in a compact way.

In many applications, the obtained information of experts can be described in different ways, even the structure of observation can be different; nevertheless decisions have to be taken by these data. With the assistance of signatures these alterations in structures can be handled. The

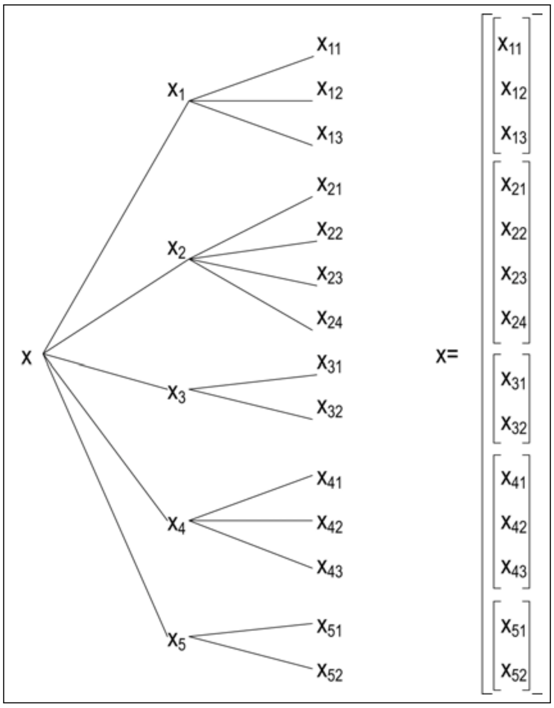


Figure 4: Tree structure and vector form of fuzzy signature

main advantage of the application of fuzzy signatures is that they can handle situations with uneven data structures and information.

Furthermore, the model created for the given task can be arranged hierarchically [8]; this feature is very similar to the way of thinking of human experts. This fact underlines the argument that fuzzy signatures are deployable on the area of decision making.

The advantage of fuzzy signatures is that they organize the available data components into hierarchical structures. This hierarchy determines the basic structure of fuzzy signature-based observations. It may occur that some elements are missing at several observations. Therefore, it is necessary to have a kind of structure modifier operator for comparing signatures with quite different structures. It is advisable to apply aggregation operators for reducing sub-trees to their parent node. In case of a multileveled hierarchy, a recursive process leads to obtaining the aggregated value of the parent node.

In our case, as the most important question the aggregation operators had to be defined. The structure of the fuzzy signature supports the use of different aggregation operators for each node.

4 The Proposed New Approach

As it was described in Section 2, due to several different aspects it is difficult to compare the elements of inhomogeneous building stock. This fact makes the judgements professionally unfounded in the awarding procedure presented in the introduction.

In the following a formal system will be proposed that may support the decisions of the Municipal Assembly in the awarding procedure. Similarly to [9], the proposed tool offers an effective solution for applying the experts' knowledge in responsible decision making. This system is based on fuzzy signatures described in Section 3.

In the recent decades the importance of visual examination on site increased in comparison with destructive analyses among building diagnostic experts. The shapes of the deteriorations, the information collected about checked building constructions and circumstances may offer enough evidence for experienced specialist to esteem the causes and consequences of building failures. Taking advantage of this, some new diagnostic tools were created (cf. Koppány, [10]).

As a comparable study, a comprehensive assessment of rental houses in Lisbon was compiled; furthermore, an objective evaluation of rental houses was done with a method based on visual examination on site [11].

The application of fuzzy set theory was proposed in bridge maintenance systems as well recently [12]. As Agárdy implies [13], soft computing techniques may help differentiating sets of objects of built environment by several aspects. As a decision support tool the fuzzy expert system was also proposed for determining the importance of intervention in case of building failures [14].

The circumstances of the evaluation of building stocks discussed above are quite different from the core problem of the above approaches. The methods of the assessments that were implied in the mentioned papers are applicable for examining the physical condition of the buildings and their elements without considering the narrower and wider context in their place. The aim of the present paper is to obtain a comprehensive and comparable data set for each building; therefore it was not possible to narrow the aspects to the physical and measurable conditions without losing important information.

This intent resulted in a large scale set of examination aspects that is difficult to handle without any classification abilities. This reason also explains why the adaptation of hierarchical system of fuzzy signature makes the support tool effective.

The model.

As the first step, the basic structure of the fuzzy signature developed from the available data and the experts' knowledge; then the adequate fuzzy sets of data elements must be determined. The next step is the identification of fuzzy signature based rules applying the experts' knowledge and the available input-output pairs. When the rule base is ready, the fuzzy signature based observation can be directly evaluated, thus generating a suggested decision.

Let us overview the concrete structure of the fuzzy signature suitable for representing the set of attributes and their respective relations used in the building evaluation approach.

In the context of this study the accessible information on a building may be arranged in five main groups, which groups constitute the first level of the fuzzy signature structure (nodes x1; x2; x3; x4 and x5).

For aggregating sub-trees within the fuzzy signatures the WRAO operator (Weighted Relevance Aggregation Operator, introduced by Mendis et al. [15]) was applied. With the application of weighted aggregations more expert knowledge can be involved in the examination. The initiated relevance weight determines the relevancy of a child node on a higher level. For determining the relevance weights by observation Mendis et al. [16] propose a method.

Let us denote the value and relevant weight of l child of n pieces of aggregating nodes with xl and wl. The WRAO function is denoted with @ in (3).

$$@(x_1, x_2, ..., x_n; w_1, w_2, ..., w_3) = \left(\frac{1}{n} \sum_{l=1}^{n} (w_l \cdot x_l)^p \right)^{\frac{1}{p}} \tag{3}$$

where the p is the aggregator factor p ∈ R, p ≠ 0.

For describing different components of the structure linguistic variables may be applied. The next step is to define these linguistic variables and their membership functions.

The global analysis of apartment houses is a complex task. A large number of factors has to be considered while a building is evaluated. The goal of the discussed study was to obtain a comprehensive character of each building. Therefore, the parent nodes in the fuzzy signature structure are the five important segments of building sciences: urban planning (cityscape attributes); art and architecture (architectural quality); engineering (constructional quality); urban sociology (life quality) and protection of monuments (aesthetic attributes).

In the following the parent and child nodes are described in groups: their relevance weights are denoted with w. For the proper application of fuzzy set signatures it has to be confirmed that at the leaves membership functions shall be applied over the [0,1] interval: for this reason the basic sets of the determined attributes have to be normalized to [0,1]. The closer the value of observation is to 1 the better is the quality in each respect.

At the nodes and at the leaves of the fuzzy signature structure partitions of triangular or trapezoidal sets are applied forming Ruspini partitions [17]($\sum_i A_i(x_0) = 1$ for every $x_0 \in X[0,1]$).

A. Cityscape Attributes ($x_1; w_1$=0.4).

The urban values of buildings are collected in the group of cityscape attributes. The matching of the examined building into the urban texture can be described by the following three child nodes (their membership functions are presented on Figure 5):

Matching to the character of the street ($x_{11}; w_{11}$=0.8) This attribute represents how the examined building harmonizes with the surroundings in shape, color, etc.

Matching to the position ($x_{12}; w_{12}$=0.4) This attribute represents how the relation is between the examined building and its position in the street (corner or other prevailing position).

Matching of the height of the façade ($x_{13}; w_{13}$=0.2) This attribute represents how the examined building differs from the surroundings vertically.

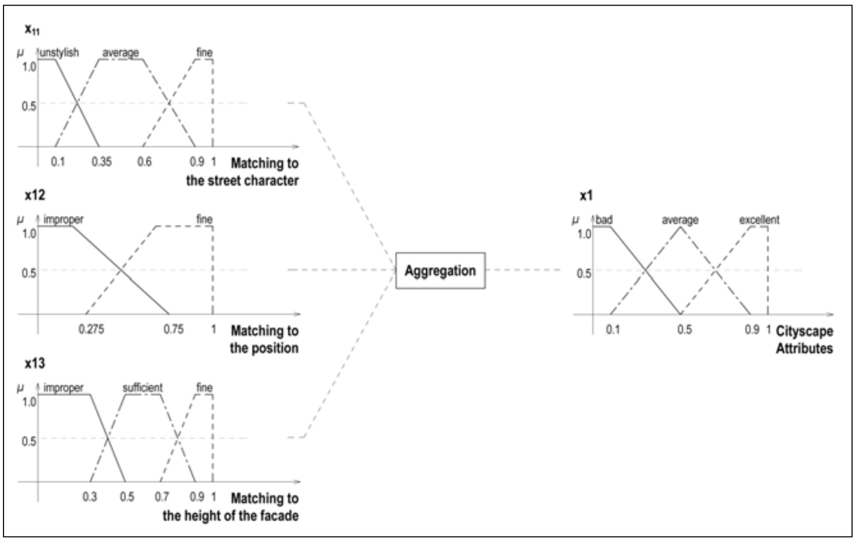


Figure 5: Membership functions of the Cityscape Attribute and of its child nodes

B. Architectural quality ($x_2; w_2$=0.75).

The broadly defined architectural quality represents the value of the building itself. As a product a building may represent extra meanings: it is more than a simple apartment house, it reflects the era of its design and construction. In this respect, four different attributes were established. The membership functions are presented in Figure 6.

Reproducibility ($x_{21}; w_{21}$=0.8) From the point of view of the reproducibility the buildings can be examined to what degree their construction could be realized with up-to-date technologies; considering available techniques and current regulations. In this case, mutatis mutandis, the lower value means that the given building is easier replaceable.

Authenticity ($x_{22}; w_{22}$=0.3) This feature represents that how the examined building corresponds to the atmosphere of the age of its construction.

Value of in-built materials ($x_{23}; w_{22}$=0.4) The proper application of fine materials (e.g. natural stone, wrought iron, hand-manufactured bricks as claddings, etc.) increases the quality of the buildings. This feature represents in what rate the examined building contains additional attributes.

Body shaping ($x_{24}; w_{24}$=0.7) Although the architectural tools applied in design had changed during the ages, the plasticity of buildings was always important, even in the Age of Modern. Beside the plasticity it is also important whether the body shaping is in harmony with the surroundings.

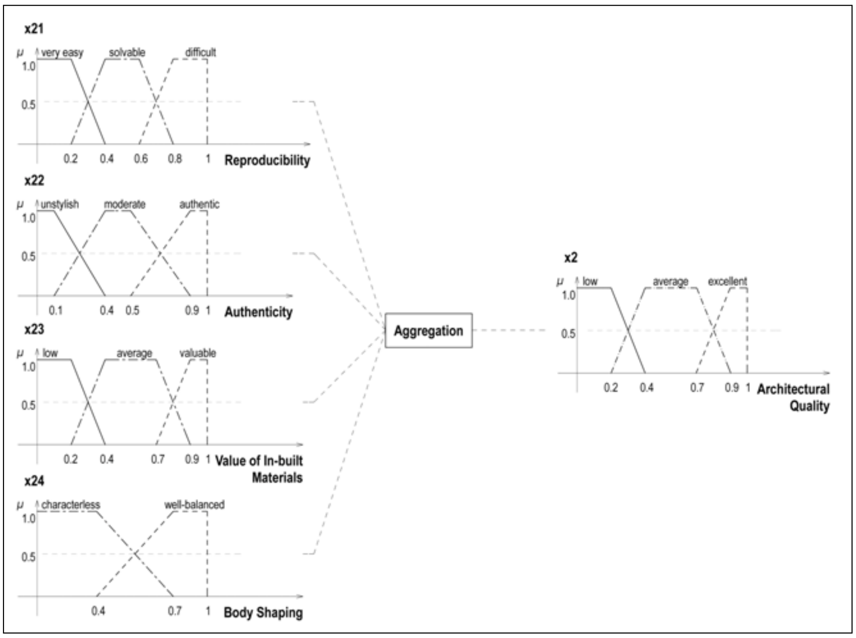


Figure 6: Membership functions of Architectural Quality and of its child nodes

C. Constructional quality ($x_3; w_3$=0.85).

The condition and the energetic performance of building constructions determine heavily the overall quality of the buildings. The exploring of the in-built constructions and their qualifying are done here only by visual examination on site. The discussed study does not contain a detailed and objective survey. The two components of this quality descriptor are as follows:

Consistence of building constructions ($x_{31}; w_{31}$=0.8) Although this fuzzy sub-signature can be divided into more sub-trees, here only single merged values are applied.

Energetic Attributes ($x_{32}; w_{32}$=0.4) The comparative analysis of external building constructions is based on estimated values for homogenous unit portions of the external wall, the openings and the loft slab. As a standard here the national regulations have to be also taken into account.

The values that represent the observations are not equal to the index number applied for determining the energetic quality of buildings.

D. Life quality ($x_4; w_4$=0.45).

This aspect has three components. Figure 8 represents the child nodes and the Life Quality membership functions.

Spatial arrangement of buildings ($x_{41}; w_{41}$=0.6) The spatial arrangement of flats may be concluded from the shape and the layout of the buildings. Although there can be a great variation of spatial arrangements of flats in a house itself, the type of the flats, the accommodation density

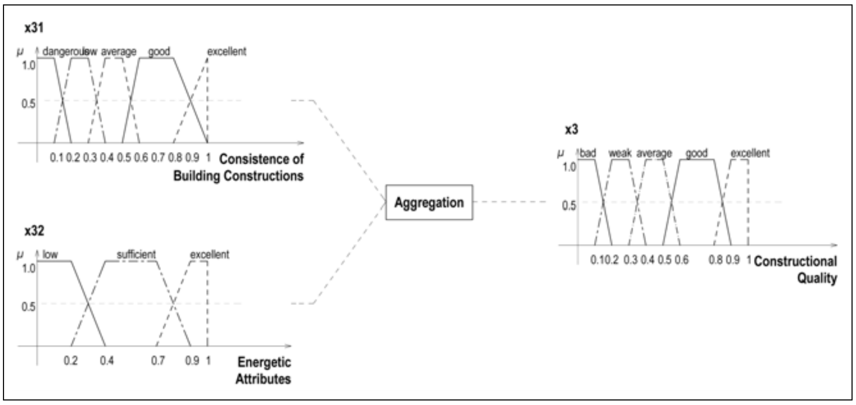


Figure 7: The Membership functions of Constructional Quality and of its child nodes

together determine the life quality of the examined building.

Existence of connection with private parks, gardens ($x_{42}; w_{42}$=0.4).

A part of the buildings have direct connection with private parks or gardens that influence the value of the building positively. The index that represents the existence of connection is binary (yes or no, 0/1).

Recreational quality ($x_{43}; w_{43}$=0.6) The distance of the examined buildings from the traffic route or the busy intersections influence the noise and pollution level of the interior spaces of the buildings. The recreational quality index represents this aspect as an accumulated value.

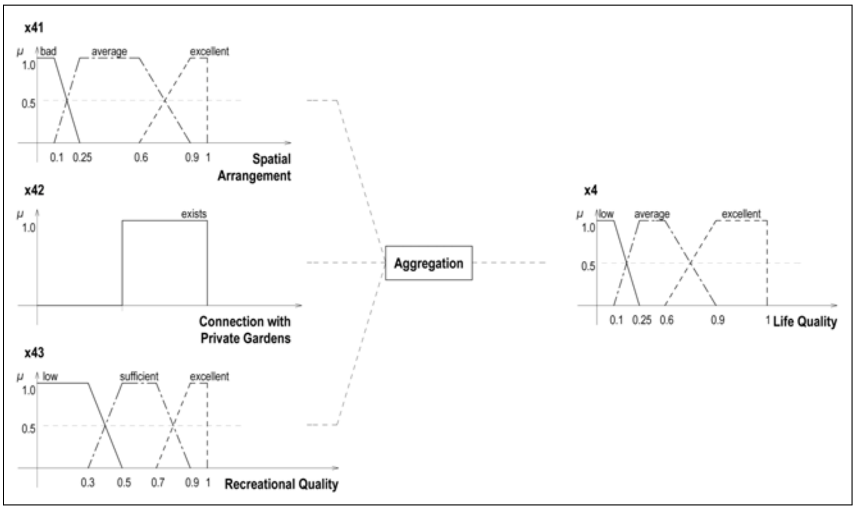


Figure 8: Membership functions of Life Quality and of its child nodes

E. Aesthetic Attributes ($x_5; w_5$=*0.5).*

Although the National Heritage Office has its professional tool for appreciating the historic value of buildings, in the present analysis only the external marks are taken into account. This attribute and its leaves can be seen in Figure 9.

Professional Judgement ($x_{51}; w_{51}$=0.8) The professional attributes of aesthetic evaluation are obtained from the analysis of the proportion of the buildings and the application of architectural elements. This analysis is founded on the qualification of the harmony between the design and the realization.

Public Judgement ($x_{52}; w_{52}$=0.3) For a comprehensive opinion about the examined buildings a public poll was done. 29 non-professionals filled questionnaires for representing the public opinion about the examined buildings. For evaluating the questionnaires a linguistic approximation according to [18] was applied. Some answers differed from the typical radically; the peaks and "needles" originating from this phenomenon were cut off ("noise cutoff").

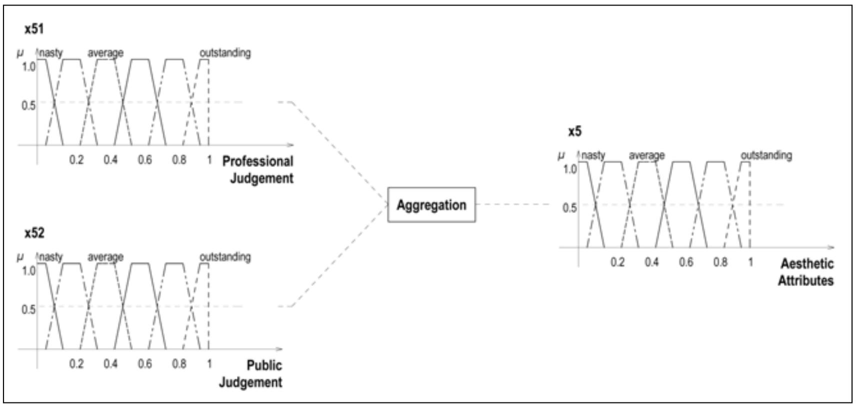


Figure 9: Membership functions of Aesthetic Attributes and of its child nodes

The Rule Base.

At the first level the number of available input variables is five. Their partitions (3;3;5;3;6 partitions) result that the total number of rules in the rule base is 810. At the level of child nodes a different (much higher) number of rules can be identified from the partitions of input variables. The detailed description of these input variables and their partitions can be above. The hierarchical structure of the signatures presents dealing with a rule base of very high complexity.

Inference Method.

Basing on Mamdani-type inference [19], Tamás [20] introduced the generalized method that operates on signature based rule bases. In this method the alteration is only in the first step, where the degree of matching between observation and the antecedents of rules is determined.

In the discussed procedure this method is applied, where the minimum conjunction is taken as the aggregation operator for reducing the signature structure. Figure 10 represents the outline of a general fuzzy inference system.

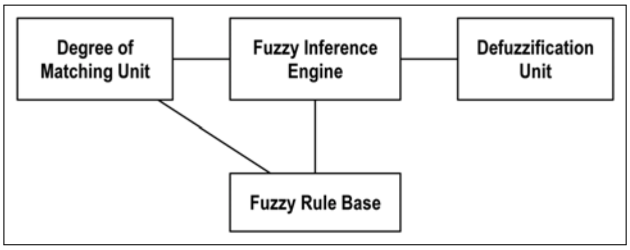


Figure 10: Outline of Fuzzy Inference Systems

5 Results

The examined buildings were evaluated with the method discussed above. The partial results are collected in Table 5.

As an example the values obtained for building 17 in Fig 3c are presented here:

The result of fuzzy signature inference is a fuzzy set that can be defuzzified for determining the consequence value. In the present examination the COG (center of gravity) method was

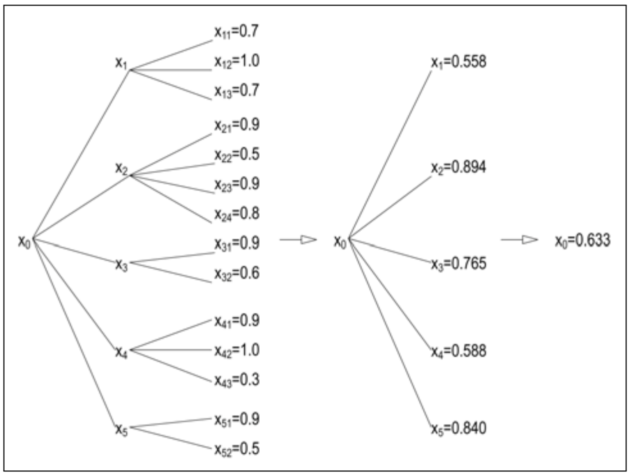


Figure 11: Fuzzy signature and aggregated values for building 17 (Fig. 3c)

	Citcyscape Atts				Architectural Quality					Constr Qty			Life Quality				Aesthetic Atts			Protection
(Nr)	x_{11}	x_{12}	x_{13}	x_1	x_{21}	x_{22}	x_{23}	x_{24}	x_2	x_{31}	x_{32}	x_3	x_{41}	x_{42}	x_{43}	x_4	x_{51}	x_{52}	x_5	Value
5	0,9	0,9	0,9	0,831	0,9	0,9	0,8	0,7	0,870	0,4	0,3	0,226	0,3	0,0	0,1	0,090	0,6	0,6	0,600	0,4540
6	0,9	0,9	0,7	0,831	0,8	0,8	0,4	0,8	0,710	0,3	0,2	0,128	0,3	0,0	0,1	0,090	0,7	0,8	0,700	0,3750
7	0,5	0,8	0,9	0,500	0,1	0,2	0,2	0,6	0,160	0,9	0,7	0,600	0,6	1,0	0,2	0,401	0,6	0,8	0,600	0,5000
8	0,9	0,9	0,9	0,831	0,6	0,6	0,6	0,6	0,550	0,4	0,4	0,250	0,4	1,0	0,2	0,102	0,5	0,6	0,500	0,2570
9	0,4	0,9	0,9	0,500	0,5	0,3	0,3	0,5	0,368	0,4	0,3	0,250	0,3	1,0	0,3	0,090	0,4	0,5	0,400	0,2480
10	1,0	0,9	0,9	0,831	0,2	0,5	0,5	0,6	0,474	0,4	0,3	0,250	0,6	1,0	0,2	0,401	0,5	0,6	0,500	0,4340
11	0,9	0,9	0,5	0,831	0,8	0,8	0,4	0,5	0,592	0,4	0,4	0,250	0,3	1,0	0,3	0,090	1,0	0,6	0,949	0,5000
12	0,8	0,9	0,9	0,656	0,7	0,9	0,4	0,8	0,627	0,3	0,3	0,185	0,6	1,0	0,3	0,401	0,7	0,9	0,700	0,4410
13	0,8	0,9	0,7	0,656	0,9	0,8	0,4	0,8	0,713	0,4	0,4	0,250	0,2	1,0	0,4	0,411	0,9	0,6	0,840	0,5000
14	0,7	0,9	0,9	0,558	0,8	0,4	0,4	0,4	0,550	0,3	0,3	0,185	0,2	0,0	0,3	0,100	0,7	0,6	0,700	0,3750
15	0,6	0,8	0,7	0,500	0,7	0,7	0,4	0,7	0,627	0,3	0,4	0,250	0,3	0,0	0,4	0,105	0,4	0,5	0,400	0,2500
16	1,0	0,9	0,9	0,831	0,5	0,3	0,6	0,6	0,431	0,5	0,3	0,250	0,5	1,0	0,4	0,411	0,8	0,6	0,800	0,5000
17	0,7	1,0	0,7	0,558	0,9	0,5	0,9	0,8	0,894	0,9	0,6	0,765	0,9	1,0	0,3	0,588	0,9	0,5	0,840	0,6330
18	1,0	0,8	0,9	0,831	0,7	0,7	0,5	0,8	0,627	0,5	0,4	0,250	0,5	1,0	0,4	0,411	0,7	0,7	0,700	0,5000
19	0,7	1,0	0,9	0,558	0,9	0,2	0,8	0,7	0,601	0,4	0,3	0,250	0,3	0,0	0,3	0,090	0,8	0,8	0,800	0,5000
20	0,3	0,9	0,9	0,500	0,3	0,2	0,5	0,7	0,423	0,5	0,3	0,250	0,7	1,0	0,4	0,411	0,5	0,4	0,500	0,3750
21	0,9	0,7	0,7	0,762	0,1	0,2	0,3	0,6	0,368	0,7	0,5	0,450	0,6	1,0	0,5	0,477	0,1	0,5	0,160	0,2500
22	0,8	0,9	0,7	0,656	0,9	0,9	0,2	0,5	0,550	0,2	0,3	0,185	0,8	1,0	0,3	0,388	0,9	0,2	0,700	0,4410
23	0,2	0,9	0,8	0,500	0,5	0,3	0,4	0,7	0,550	0,3	0,3	0,185	0,3	1,0	0,4	0,411	0,4	0,8	0,400	0,2500
24	0,2	0,3	0,4	0,386	0,1	0,1	0,5	0,7	0,153	0,8	0,5	0,450	0,8	1,0	0,3	0,388	0,1	0,6	0,160	0,2500
25	0,2	0,5	0,2	0,381	0,1	0,2	0,3	0,8	0,368	0,5	0,6	0,250	0,4	1,0	0,4	0,411	0,2	0,3	0,200	0,2500
26	0,3	0,6	0,4	0,437	0,2	0,1	0,3	0,8	0,368	0,9	0,5	0,600	0,8	1,0	0,6	0,467	0,2	0,3	0,200	0,2500
27	0,3	0,7	0,3	0,470	0,1	0,2	0,3	0,8	0,368	0,6	0,8	0,525	0,5	1,0	0,6	0,477	0,2	0,3	0,200	0,2500
28	0,2	0,5	0,7	0,392	0,1	0,1	0,3	0,6	0,174	0,8	0,5	0,450	0,7	1,0	0,4	0,411	0,1	0,3	0,160	0,2500
29	0,8	0,8	0,8	0,627	0,7	0,3	0,4	0,4	0,459	0,3	0,3	0,185	0,3	1,0	0,8	0,482	0,4	0,6	0,400	0,3360
30	0,1	0,1	0,2	0,169	0,1	0,1	0,2	0,8	0,153	0,8	0,7	0,450	0,7	1,0	0,8	0,571	0,1	0,6	0,160	0,2500

Table 5: Evaluation of the examined buildings

applied.

The applied categories for the assessment of buildings at the side of consequence are visible in Table 6 where the categories and corresponding ranges of values for the ranked buildings are presented.

Categories	**Values**	**Buildings (Nr.)**
worthless	0-19%	
below average	20-39%	6;8;9;14;15;20;21;23;24;25;26;27;28;29;30
average	40-74%	5;7;10,12;13;16;17;18;19;22
valuable	75-89%	
highly valuable	90-100%	

Table 6: Categories of overall value of the examined buildings

The result of the evaluation demonstrates that the building stock has a balanced value. In spite the fact that each building has different conditions (age, aesthetical, constructional quality), the deviation among their category indices is low. Among other conclusions this result may confirm the perception that the buildings constructed recently did not increase the overall value of the examined building stocks.

6 Conclusions and Future Work

As a summary it can be stated that the results of calculations correlate to the former assumptions. The utilization of linguistic variables supported the evaluation of non-measurable (e.g. aesthetic) and uncertain (e.g. energetic performance) values. The application of weighted relevance aggregators in the reducing phase also maintained the professional aspects in examination. The unique hierarchical construction of fuzzy signature structure enhanced the adequacy of the overall evaluation. These statements certify the previously expressed supposition that a complex and well defined decision support tool based on fuzzy signature structure may help in the mentioned awarding procedure.

Therefore, after several refinements of the tool, it is going to be presented to the Municipal Assembly. These refinements may take place in a further research project in the following directions.

The complexity of the building (different structural and sub-structural components, spatial object with complicated relation to the users and the surroundings, etc.) requires more complex analyses. That is why it is necessary to increase the number of attributes and to create more precise linguistic variables. For avoiding the confusion in the system in the process of the enlargement of the structure it is advisable to create further sub-trees (sub-sub-trees).

It is also worth considering that a confidence index can be involved for the process of determination of the matching the rules that may support differentiating the observations.

The developed decision support tool may take the intermittently determined aspects of the Assembly. Therefore the application of controllable relevance weight on the parent nodes has to be examined.

Acknowledgement
The authors wish to record their gratitude to Attila Koppány for his advices and professional assistance during each phase of the project.

Bibliography

[1] G. Locsmándi, Large-Scale Restructuring Processes in the Urban Space of Budapest, in R. Dalla D. Longa (ed.) Urban Models and Public-Private Partnership, Berlin, Springer Berlin Heidelberg, 2011

[2] S. Lowe, A tale of two cities - rental housing in Budapest and Sofia in the 1990s, Journal of Housing and the Built Environment, Kluwer Academic Publishers, pp. 249-266, 2000

[3] L. T. Kóczy and D. Tikk, Fuzzy Systems (in Hungarian), Budapest, Typotex, 2000

[4] L. A. Zadeh. Fuzzy Sets, Information and Control, 8, pp. 338-353, 1965

[5] J. A.Goguen, L-Fuzzy Sets, Journal of Mathematical Analysis and Applications, 18, pp. 154-174, 1967

[6] L. T. Kóczy, Vector Valued Fuzzy Sets, BUSEFAL, Tolouse, pp. 41-57, 1980

[7] L. T. Kóczy, T. Vámos, G. Biró, Fuzzy Signatures, Proceedings of EUROFUSEC-SIC'99, Budapest, pp. 210-217, 1999

[8] K. W. Wong, A. Chong, T. D. Gedeon, L. T. Kóczy, Hierarchial Fuzzy Signature Structure for Complex Structured Data, Proceedings of the International Symposium on Computational Intelligence and Intelligent Informatics, Nabeul, Tunisia, pp. 105-109, 2003

[9] I. Anica-Popa, G. Cucui, A Framework for Enhancing Competitive Intelligence Capabilities using Decision Support System based on Web Mining Techniques, International Journal of Computers Communications & Control, ISSN 1841-9836, 4(4):326-334, 2009

[10] A. Koppány, A Diagnostic System Created for Evaluation the Quality of Building Constructions, Proceedings of 11 DBMC International Conference on Durability of Building Materials and Components, Istambul, 2008

[11] C. O. Pedro et al., Portuguese Method for Building Condition Assessment, Emerald, Lisbon, 26, pp. 322-335, 2008

[12] I. Szaradics, Adaptability of Fuzzy Control Systems in Bridge Management (M.Sc thesis), Györ, Széchenyi István University, 2007

[13] Gy. Agárdy, Fuzzy Logic Techniques in Bridge Management Systems, Acta Technica Jaurinensis, Győr, Széchenyi István University, 1, pp. 49-60, 2008

[14] G. I. Molnàrka, The Construction of the Knowledge Base of Fuzzy Expert System in Building Pathology - From a Functional Access, Proceedings of CIB World Building Congress, Rotterdam, pp. 1406-1417, 2007

[15] B. S. U. Mendis, T. D. Gedeon, J. Botzheim, L. T. Kóczy, Generalised Weighted Relevance Aggregation Operators for Hierarchical Fuzzy Signatures, Proceedings of the International Conference on Computational Intelligence for Modelling, Control and Automation (CIMCA 2006), Sydney, 2006

[16] B. S. U. Mendis, T. D. Gedeon,L. T. Kóczy, On the Issue of Learning Weights from Observations for Fuzzy Signatures, Proceedings of the World Automation Congress 2006, Budapest, pp. 1-6, 2006

[17] E. H. Ruspini, A New Approach to Clustering, Information Control, 15, pp. 22-32, 1969

[18] A. Mehler, K. Reinhard, G. Klir, K. Sentz, On the Issue of Linguistic Approximation. Aspects of Automatic Text Analysis, Studies in Fuzziness and Soft Computing, Berlin, Springer Berlin, Heidelberg, 209, pp. 61-78, 2007

[19] E.H.Mamdani, S.Assilian, An Experiment in linguistic variables with a fuzzy logic controller, International Journal of Man Machine Studies, 7, pp. 1-13, 1975

[20] K. Tamás, L. T. Kóczy, Mamdani-type Inference in Fuzzy Signature Based Rule Bases, 8th International Symposium of Hungarian Researchers on Computational Intelligence and Informatics, Budapest, pp. 513-525, 2007

Int. J. of Computers, Communications & Control, ISSN 1841-9836, E-ISSN 1841-9844
Vol. VI (2011), No. 3 (September), pp. 458-461

Remembering the Beginnings

C.V. Negoita

Constantin Virgil Negoita
Computer Science, Hunter College,
City University of New York
http://www.hunter.cuny.edu/cs/Faculty/Negoita
E-mail: cnegoita@hunter.cuny.edu

Abstract: This paper wants to provide a personal view of the beginnings of fuzzy logic which was first introduced in the 1960's by Lotfi Zadeh, when fuzzy sets were presented as models of any vague concept, making possible the computation with words.
Keywords: fuzzy sets, fuzzy systems, vagueness

1 Introduction

We know Zadeh as we know no other eminent man who have made the last half of the last century memorable in science history. A mere glance at the materials to which we have access still suffice to show that our information regarding him is of such a kind as to leave scarcely anything to be desired. In the first place we have his papers. They are written with talent. He has not only left a minute record of his research during a space of nearly five decades but he found real pleasure in communicating on paper.

As a scientist, he was consistent. His memoir of himself remains unfortunately a fragment, but enough was completed to illustrate that portion of his career during which fuzzy sets were promoted. But if we owe much to the communicativeness of Zadeh himself, we owe much also to the communicativeness of his peers. The admiration of his followers is well known, but the picture which the critics drew of him was not always a pleasing one. They saw him not as he presented himself to the fascinated eye of friendship.

It should be remembered that they knew him only from a vivid immediacy that derives from the speech of witnesses, full of slang of the moment yet extremely serious because the fuzzy set seemed of the utmost cultural significance. What they painted was what they understood, and what they understood was very little.

2 Fuzzy Sets and Generalized Sets

It is common for logicians to give truth conditions for predicates in terms of set theory. "John is tall" is defined to be true just in case the individual John is in the set of tall men. Zadeh used the same path and said that "tall men" is a fuzzy set when the membership is a matter of degree.

In presenting the theory of fuzzy sets, we hoped to break through the bars of the prison of set theory. To understand a fuzzy set, imagine a two-dimensional world called Flatland. Each Flatlander is incarcerated in a flat set. We can peel him off and place him back somewhere else. If we fling a Flatlander into our three-dimensional world, he can see only two-dimensional cross sections of our world, a family of crisp sets. Simply put, by adding another dimension, we can capture more features. This is what a fuzzy set does. It adds a new dimension: our evaluation of the membership. Using classical flat mathematics, a fuzzy set can be represented by a family of crisp sets, projected on the Flatland.

Copyright © 2006-2011 by CCC Publications

The classical rules of logic are represented by operations on the set with only two elements: true and false (0 and 1).

In the universe of all sets - call it the category of sets - this important set is called classifier. When we tried to investigate the category of fuzzy sets, it was impossible to find a similar classifier. In fact in fuzzy set theory there is no fuzzy set of fuzzy subsets of a fuzzy set. The point is that there are two predicates in set theory: membership and equality. In the category of generalized sets, both can be fuzzy, but in fuzzy set theory only membership is allowed to be. This fact puzzled a lot of people.

Some critics bordered on the vituperative, and the tenets of fuzzy logic were dismissed as comical. Its arguments were declared frivolous and idle exercises in irrelevance and blasphemy.

In 1977 Arbib wrote bad reviews for the periodicals, and in 1984 Zeleny published a paper on the (ir)relevancy of fuzzy set theory. That, no doubt, explains why Herbert Toth, in 1987, in his PhD Thesis at the University of Vienna suspected that probably something has gone wrong in the development and interpretation of the theory.

Toth didn't deny Zadeh's original definition to be natural, immediate and elegant. This assertion was sufficiently justified by the vast amount of literature in an epidemically growing number of papers and books.

3 Fuzzy Systems and the Postmodern Times

Perhaps to deter us fuzzy people from further abuse, or perhaps only to improve our connection, Jim Bezdeck established our annual meetings. These are the origins of the international meetings that have today become pilgrimages, where all controversis were clarified when fuzzy sets were represented as families of crisp sets, and the Japanese started to implement fuzzy comptrollers, and, later, fuzzy systems were implemented as neural networks. But, still then, some mathematicians never understood fuzziness, because, for them, the precise specification of a set could be given only by binary logic.

The anatomy of a boom is simple. Over time, most ideas will rise in value. As this happens, people are attracted to them and this causes the ideas to rise more. This further gain attracts more people and gradually, perhaps over a period of a century, the number of people looking for this increase in value comes to determine what ideas are worth. The knowledgeable man, as he unwisely considers himself, is now concerned with the way an idea is attracting interest. That, rightly for the moment, determines its value.

The binary logic, beautiful, useful, and promising, determined the modern era, obsessed with mathematical models. Scientific truths became the pillars of progress. Zadeh, speaking of degrees of truth, shocked the foundation of modernity, and became the postmodern of information sciences. More than that, he defined postmodernity as a return to premodernity.

Why do postmodern scientists, in their advancing years, when love for precision survives, but only barely, cleave to the fluid vagueness as though it were cable for rappelling and not a tightrope any longer?

To embrace the whole from one point of view. The remarks made are less detailed, but more sure. You perceive each object less distinctly, but you describe the facts with more certainty. The details of the immense picture are lost in the shade, but you conceive a clear idea of the entire object.

In the philosophy of science this fact has been known for a long time. In the sixth century, Leontius from Bizantium observed that our impression of the world is vague, not revealing the details. If we attempt to particularize by division into species and individuals, the general view is lost: we are heading not towards truth but towards an infinite regress. In 1906, in France, Pierre Duhem, in a book about physics, its object and structure, distinguished between practical facts expressed in vague, ordinary language and theoretical facts expressed in precise, quantitative language. He argued that confidence in the truth of a vague assertion may be justified just because of its vagueness, which makes the assertion compatible with a whole range of observed facts. There is a balance between precision and certainty: one is increased only to the detriment of the other.

In Fuzzy Systems [1], a book I published in England, in 1981, I noted that one can build an infinite number of functions with values in an unit interval. In other words, we can say that a set never becomes exhausted in its evaluations. The set is massive, complete in itself, totally and wholly given, equivalent to the inert world of objects and things. As a result of an evaluation, a fuzzy set corresponds to the human consciousness. The set is both logically and ontologically prior to the fuzzy set. The latter is dependent upon the former. The fuzzy set is inconceivable without the set and is derived from it trough ordering. The fuzzy set is a borrowed being, an ordered set. It is through man that fuzziness enters the world. What is about the being of man that occasions fuzziness? The need of synthesis, the freedom to move up in the universe of concepts in order to find out stability. The reason that vague descriptions are used is to achieve stability.

Although awareness of the vagueness in our descriptions has existed for a very long time, only Zadeh's inquiries, which come under the heading "fuzzy", have brought matters to a head. It was Zadeh who coined the term when he spoke about fuzzy sets, blurred, indistinct in shape or outline, frayed, fluffy number of things of the same kind, together because they are similar, with different degrees of togetherness.

The description is fuzzy, not the things. People in the business of making exact descriptions concluded that an exact description is virtually impossible. This is a fact we have to accept and adjust to. Inexactness is not a liability. On the contrary, it is a blessing in the sufficient information it can convey with less effort. The vague description is easier to remember. That is inexactness makes for greater efficiency. Goguen in 1969 and Belman in 1970, agreed with Zadeh, when he said that precise quantitative analyses are not likely to have much relevance to the real-world problems which involve humans either as individuals or in gropus. He saw that as the complexity of a system increases, our ability to make precise and yet significant statements about its behavior diminishes. He saw all these things. He said also that precision and certainty are incompatible.

The 1960s were an age of innovation. At every corner critical thought and change were eating away the foundations of the traditional modern era.

In 1969 I defended a PHD thesis on Information Retrieval. To retrieve means to specify descriptors and rules for combining them. Immediately the question of vagueness came up. In my dissertation, I have questioned the meaning of vagueness in linguistic description with obsessive insistence. As my research went on, I began first to suspect, then to dislike, and finally to detest everything connected with boundaries between classes because a continuum exists that makes it impossible to do so satisfactorily.

In 1971 I published Information Storage and Retrieval which sounded another note of the famous war waged by common sense against the abuse of binary logic, a war in which so many were to engage. Out of this book grew Fuzzy Sets and their Applications, written together with Dan Ralescu in 1974, where we introduced the theorem of representation of a fuzzy set as a family of crisp sets.

4 Some Final Remarks

The history of fuzzy systems is marked by different visions. On the one hand a fuzzy set is a function. On the other hand a fuzzy set is a family of sets. The first vision took logic as its paradigm. The second vision made possible the recent embracing of connectionism.

For the idea of a model of vagueness, I was indebted to Zadeh whose paper on "Fuzzy Sets" appeared in 1965. It was he who first proved, it would seem, conclusively, that the meaning of words could be represented by numbers. To him we owe the first full and satisfactory representation of a vague word as a function.

Berkeley was the place where Zadeh was preaching his gospel, but his onslaughts were mercy compared with those terrible philippics in which, at Vanderbilt University in Nashville, Georgescu Roegen gave vent to his rage against arithmomorphism, the worship of numbers. In 1971, in an extraordinary book, The Entropy Law and the Economic Process, among the thesis that he defends is the claim that concepts are not arithmomorphic. They do not overlap. Concepts like "good" or "tall" have no boundaries. Instead they are surrounded by a penumbra within which they overlap with their opposites. At a particular historical moment, he notes, a nation may be both a democracy and a dictatorship just as there is an age when a man is both young and old. To the category of concepts we cannot apply the fundamental law of the binary logic, the principle of excluded middle (X cannot be both A and non-A).

This axiom has so deeply imbued the modern era that it was felt to be natural and self-evident. Its denial seemed to be nonsensical. In opposition to the binary logic is what one might call premodern dogmatic logic, which assumes that A and non-A do not exclude each other as predicates. This logic was predominant in the first centuries. For the premodern Church Fathers, the opposition was a category of man's mind, not in itself an element of realty. The idea was that thought can only perceive in contradictions. The only way in which the world can be grasped ultimately lies, not in thought, but in the experience of oneness. (Three divine persons were grasped as one God).

Basically, fuzzy logic is based on the same feeling, and its applications allow engineers to create machines that approach human responses to stimuli, working with incomplete and unclear data to generate positive actions. Using fuzzy logic, Japanese washing machines are able to decide how dirty clothes are, how much water and soap should be used to wash them, and how long it should take to get them clean, all things an experienced launderer would know how to do instinctively. To me, the most impressive accomplishment was a fuzzy system built by Michio Sugeno, in 1985, at the Tokyo Institute of Technology, when he stabilized a helicopter that lost a rotor blade. No human pilot can manage that, and no mathematical model either.

Fuzzy systems, linguistically inspired, are a direct consequence of the seminal papers of Zadeh, published in the 1970s.

In 1975 classical fuzzy set theory had reached its apogee, since solutions of its basic problems were now at hand. Classical fuzzy set theory then changed from an heroic phase, in which we addressed ourselves to hitherto unfathomable questions, to an academic phase, in which a wealth of detail, albeit most important detail, was worked on by an army of competent scholars and technologists following well-established lines. The period of trail blazing was over, though most of the practical benefits were yet to be reaped.

Bibliography

[1] Negoita, C.V., *Fuzzy Systems*, Abacus Press, Tunbridge Wells, UK, 1981

[2] Negoita, C.V., *Expert Systems and Fuzzy Systems*, Benjamin Cummings, Menlopark, CA, 1985

[3] Negoita, C.V., *Fuzzy Sets*, Newfalcon Publications, Tempe, AZ, 2000

Int. J. of Computers, Communications & Control, ISSN 1841-9836, E-ISSN 1841-9844
Vol. VI (2011), No. 3 (September), pp. 462-472

Distance Based Triggering and Dynamic Sampling Rate Estimation for Fuzzy Systems in Communication Networks

C.N. Nyirenda, F. Dong, K. Hirota

Clement N. Nyirenda, Fangyan Dong, Kaoru Hirota
Department of Computational Intelligence and Systems Science
Tokyo Institute of Technology
G3-49, 4259 Nagatsuta, Midori-ku, Yokohama 226-8502, Japan.
E-mail: {nyirendac, tou, hirota}@hrt.dis.titech.ac.jp

Abstract: To reduce computational cost in fuzzy systems in communication networks, distance based triggering and sampling rate adaptation probabilities are proposed based on the concept of *probability via expectation*. The triggering probability, which is calculated by using the square of distance between subsequent input vectors, governs the rate at which the fuzzy system is triggered. The dynamic sampling rate probability, which governs the adaptation of the sampling rate, is computed by using the exponentially weighted moving average (EWMA) of the triggering probability. A stopping criterion, based on convergence tests, is also proposed to ensure that the mechanism switches off when the sampling period has converged. The triggering mechanism reduces the number of computations in the Fuzzy Logic Congestion Detection (FLCD) in wireless Local Area Networks (WLANs) by more than 45%. Performance, in terms of packet loss rate, delay, jitter, and throughput, however, remains virtually the same. On the other hand, the dynamic sampling rate mechanism leads to more than 150% improvement in sampling rate and more than 70% reduction in fuzzy computations while performance in the other key metrics remains virtually the same. As part of future work, the proposed mechanism will be tested in fuzzy systems in wireless sensor/actuator networks.
Keywords: communication networks, fuzzy systems, sampling rate.

1 Introduction

The number of fuzzy logic based applications in communication networks is increasing rapidly. This development is motivated [1] by the difficulties experienced when modeling communication networks by using conventional analytical methods. Some of the fuzzy applications include power control [2] in cellular systems; congestion control in IP netwoks [3], [4]; routing [5] and data fusion [6] in wireless sensor networks; and Quality of Service management in wireless sensor and actuator networks [7]. Input parameters are, generally, sampled at a fixed rate and the fuzzy system is triggered accordingly. In some cases, an external signal is used in order to trigger the systems. The fuzzy computations are invoked even when there are no significant differences between the subsequent input parameters, at the expense of precious CPU and memory resources. Furthermore, for systems that employ a sampling rate, the rate is chosen by trial and error such that it is difficult to tell if it is optimal.

This work proposes a distance based triggering mechanism for fuzzy systems in communication networks. This generic framework can be applied to any fuzzy system with minor customization. A preliminary version of this work was presented in [8]. In this work, the concept of *probability via expectation* [9] is used to calculate the triggering probability by using the square of the distance between two subsequent input vectors. If the new input parameter vector is very

Copyright © 2006-2011 by CCC Publications

far from the previous one, the triggering probability is 1. If subsequent input vector parameters are deemed to be very close to each other, the triggering probability is 0.

A sampling rate adaptation probability, which is based on the transformed Lorentzian function [10] of the exponentially weighted moving average (EWMA) of the triggering probability, is also proposed. When the EWMA is very high, it implies that the system's input vector is under-sampled. As a result, the sampling rate is increased. On the other hand, when the EWMA is low, it implies that the system's input vector is oversampled. The sampling rate is, therefore, decreased. A stopping criterion, based on convergence tests [11], [12] of subsequences of the sequence of the sampling period, is also proposed in order to ensure that the mechanism switches off when the sampling period has converged.

These mechanisms are tested on the Fuzzy Logic Congestion Detection (FLCD) mechanism [13] in the wireless Local Area Network (WLAN) environment using simulations on the NS2 platform [14] running on the Ubuntu 9.10 OS. The computing hardware is composed of a 4GB RAM and an Intel Core i7 860 2.80GHz CPU. The impact of the reduction in fuzzy system evaluations on packet loss rate, delay, jitter and throughput is also evaluated.

The rest of the paper is organized as follows. The proposed distance based triggering and the dynamic sampling rate mechanisms are presented in 2. An overview of the FLCD approach in WLANs is presented in 3. The evaluation of the proposed mechanisms is presented in 4.

2 The Distance Based Triggering and Dynamic Sampling Rate Mechanisms

The distance based triggering and the dynamic sampling rate mechanisms are incorporated in the input mechanism of the generic fuzzy logic control framework. Every τ seconds, the new crisp inputs $x_1(t), ..., x_N(t)$ are normalized to the range $[0, 1]$. The triggering probability $P(Tr)$, which governs the system's triggering rate, is calculated by using the distance between the new normalized crisp input vector $(x_1^*(t), ..., x_N^*(t))$ and the previous normalized crisp input vector $(x_1^*(t-\tau), ..., x_N^*(t-\tau))$. If the system is not triggered, the previous crisp outputs are used for decision making or control action. When the system has been triggered, fuzzy computations [15] [16] are carried out to generate new output(s).

The dynamic sampling rate mechanism uses $P(Tr)$ to compute the sampling rate adaptation probability $P(Ra)$, based on which the value of τ is adapted to an optimal level. Both $P(Tr)$ and $P(Ra)$ are developed by using the concept of *probability via expectation* [9]. According to this concept, an event A, which, in a given case, either occurs or does not, corresponds to a set of realizations ω. This set, also denoted by A, is a subset of Ω, which denotes the sample space of all possible realizations. The *probability of A*, the expected proportion of cases in which event A actually occurs, is defined as $P(A) = E(I(A, \omega))$, where $I(A, \omega)$ is the *indicator function of A*, defined by

$$I(A, \omega) = \begin{cases} 1 & (\omega \in A) \\ 0 & (\omega \notin A). \end{cases} \quad (1)$$

2.1 Distance Based Triggering Probability

The event of interest is the triggering of the fuzzy system, denoted by Tr. A realization for which the event Tr takes place is denoted by ω_1, while Ω_1 denotes the sample space of all possible realizations for which triggering is considered.

To reduce computational overhead when calculating the distance based triggering probability $P(Tr)$, the square of distance $d^2(t)$ between $(x_1^*(t), ..., x_N^*(t))$ and $(x_1^*(t-\tau), ..., x_N^*(t-\tau))$ is used, where $d^2(t) = (x_1^*(t) - x_1^*(t-\tau))^2 + ... + (x_N^*(t) - x_N^*(t-\tau))^2$. Let $v_1 : [0,1]^N \longrightarrow [0,1]$ denote the distance parameter that defines the variation of ω_1 with respect to Ω_1. This parameter is defined by using

$$v_1 = \begin{cases} \frac{d^2(t)}{\Phi} & \text{if } d^2(t) < \Phi \\ \\ 1.0 & \text{otherwise.} \end{cases} \tag{2}$$

where Φ is a normalizing constant. The triggering probability $P(Tr)$ is defined as $P(Tr) = E(I(Tr, \omega_1))$; the indicator function $I(Tr, \omega_1)$ is defined by using

$$I(Tr, \omega_1) = \begin{cases} 1 & (v_1 > R_1) \\ \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

where $R_1 \in [0,1]$ is a random number.

A high $P(Tr)$ implies that the distance between the new input vector and the previous one is large. Therefore, there is sufficient new information such that the fuzzy system has to be triggered. On the other hand, a low $P(Tr)$ implies that the change in the input vector is not significant. Therefore, the system should use the previous crisp outputs thereby preserving CPU and memory resources.

2.2 Sampling Rate Adaptation Mechanism

Apart from governing the triggering rate, $P(Tr)$ also gives information on whether the sampling period τ is optimal or not. If $P(Tr)$ is very high for long periods, it implies that the system is predominantly under-sampled such that there is a need to increase the sampling rate. Conversely, if $P(Tr)$ is very low for long periods, it implies that the inputs are more or less static. The sampling rate must be reduced because the frequent processing of the inputs is just a waste of computing resources. An inverted bell shaped function must, therefore, be employed in order to ensure that there are little or no changes to τ when $P(Tr)$ is close to 0.5, which signifies an optimal sampling rate. On the other hand, in extreme regions, τ must either be decreased or increased by an adaptation factor α. A transformed Lorentzian function is employed to generate a bell shaped function depicting this behavior. The traditional three-parameter Lorentzian function [10], from which it is derived, is defined by

$$L(z) = A_p \left[\frac{\gamma^2}{(z - z_0)^2 + \gamma^2} \right], \tag{4}$$

where z_0 is the centre; γ is the width parameter; and A_p determines the peak.

Sampling Rate Adaptation Probability

To track the variations of $P(Tr)$ at time t, the *exponentially weighted moving average (EWMA)* of $P(Tr, t)$, denoted by $\overline{P(Tr, t)}$, is determined by using

$$\overline{P(Tr, t)} = w * \overline{P(Tr, (t - \tau))} + (1 - w) * P(Tr, t), \tag{5}$$

where $w \in [0, 1]$ is a weighting factor. To ensure that previous values of $\overline{P(Tr, t)}$ are discounted at a medium rate, $w = 0.5$ is used in this study.

The event depicting the adaptation of the sampling rate of the fuzzy system is denoted by Ra. A realization for which this event takes place is denoted by ω_2, while Ω_2 denotes the sample space of all possible realizations for which the adaptation of the sampling rate is considered. In a similar approach to 2.1, let $v_2 : [0,1] \longrightarrow [0,1]$ denote the rate adaptation parameter that defines the variation of ω_2 with respect to Ω_2 by using the transformed Lorentzian function. The relationship between v_2 and $\overline{P(Tr,t)}$ is, therefore, defined by

$$v_2 = L^*(\overline{P(Tr,t)}) = \left[\frac{(\overline{P(Tr,t)} - 0.5)^2}{(\overline{P(Tr,t)} - 0.5)^2 + \gamma^2}\right]. \tag{6}$$

where γ is defined by

$$\gamma = f(\overline{P(Tr,t)}) = \begin{cases} \overline{P(Tr,t)} & \text{if } \overline{P(Tr,t)} \leq 0.5 \\ \\ 1 - \overline{P(Tr,t)} & \text{if } \overline{P(Tr,t)} > 0.5 \end{cases} \tag{7}$$

The sampling rate adaptation probability is defined as $P(Ra) = E(I(Ra,\omega_2))$; the indicator function $I(Ra,\omega_2)$ is defined by

$$I(Ra,\omega_2) = \begin{cases} 1 & (v_2 > R_2) \\ \\ 0 & \text{otherwise,} \end{cases} \tag{8}$$

where $R_2 \in [0,1]$ is a random number. If $P(Ra) = 1$, τ is adjusted by using

$$\tau = \begin{cases} (1-\alpha)\tau & \text{if } \overline{P(Tr,t)} \geq 0.5 \\ \\ (1+\alpha)\tau & \text{otherwise.} \end{cases} \tag{9}$$

This mechanism will help to ensure that $\overline{P(Tr,t)} \to 0.5$. The sampling period can be initialized randomly within a particular range. This mechanism will optimize it on-line based on the variations in the inputs. This characteristic is essential for new links and in situations where new nodes or traffic patterns have been introduced on the already existing links.

Stopping Criterion for Sampling Rate Adaptation

Once the dynamic sampling rate mechanism starts running, there is a need for a stopping criterion; otherwise, this mechanism will end up consuming precious CPU and memory resources even when the sampling period has converged. The evolution of the sampling rate τ can be presented as a sequence (τ_k) based on the concept of sequences of real numbers discussed in [11], [12]. When convergence tests show that the sequence has converged, it implies that an optimal sampling period has been realized. Therefore, the sampling rate estimation mechanism must be stopped.

Definition 1. A sequence (τ_k) converges to τ if for every $\epsilon > 0$, there exists a $K \in \mathbb{N}$ such that $|\tau_k - \tau| < \epsilon$ for all $k \geq K$. The point τ is called the limit of (τ_k).

If the terms of the sequence get arbitrarily close together, a sequence is said to be *Cauchy*.

Definition 2. A sequence (τ_k) is said to be a Cauchy sequence if for every $\epsilon > 0$, there exists a $K \in \mathbb{N}$ such that $|\tau_k - \tau_m| < \epsilon$ for all $k, m \geq K$.

Every convergent sequence is a Cauchy sequence. For real numbers, the converse is also true; every Cauchy sequence is convergent. In addition, a sequence is convergent if and only if all of its subsequences converge toward the same limit. A *subsequence* of the sequence (τ_k) is a sequence of the form (τ_{k_j}), where for each $j \in \mathbb{N}$, there is $k_j \in \mathbb{N}$, and $k_j < k_{j+1}$ for all j. From these observations, we have the following theorem

Theorem 3. *If the first term of a convergent subsequence, $\tau_{k_1} = \tau_K$, then for the first p terms,*

$$\frac{1}{p-1}\sum_{j=1}^{p-1}\left|\tau_{k_{j+1}} - \tau_{k_j}\right| < \epsilon. \tag{10}$$

Proof: This follows from *Definition 1* and *Definition 2*. Because the subsequence (τ_{k_j}) is convergent, it is also a *Cauchy sequence*. Therefore for every $\epsilon > 0$, there exists a $K \in \mathbb{N}$ such that $\left|\tau_{k_j} - \tau_{k_n}\right| < \epsilon$ for all $k_j, k_n \geq K$. This implies that the average of the absolute values of the differences of subsequent p terms is also less than ϵ. □

By comparing the average of the absolute values of the differences of subsequent terms in a subsequence (τ_{k_j}) with ϵ, it is possible to determine whether the original sequence (τ_k) has converged or not. The terms of the subsequence are extracted from (τ_k) every $L * \tau(t)$ seconds, where $L \in \mathbb{N}$. To reduce computational overhead and memory requirements, p is set to 4; on the other hand, $\epsilon = 1 \times 10^{-5}$. While the focus is on detecting convergence of (τ_k), it must be pointed out that cases where this sequence does not converge should also be anticipated. In such cases, the sampling rate estimation mechanism, once activated, will remain on until the system administrator decides to stop it.

The proposed mechanisms are tested on the zero-order Takagi-Sugeno [16] inference based FLCD approach [13] in WLANs. Next, an overview of the FLCD algorithm is described.

3 An Overview of the FLCD Approach in WLANs

The FLCD mechanism is a 2-input 1-output system. It is composed of the Fuzzy Logic Control Unit (FLCU), the Congestion Notification Unit (CNU), and the CHOKe [17] Activator (CA) as shown in **Figure 1**. The FLCU uses the backlog (queue size) factor x_1 and the packet arrival factor x_2 to generate the packet marking probability p_b. The queue on the outgoing link is sampled at a period $\tau = 2$ msec in order to obtain the two inputs. The CNU either marks (if ECN is enabled) or drops packets with a probability p_b. Responsive flows such as TCP react to these events by reducing their sending rates thereby reducing congestion at the bottleneck link.

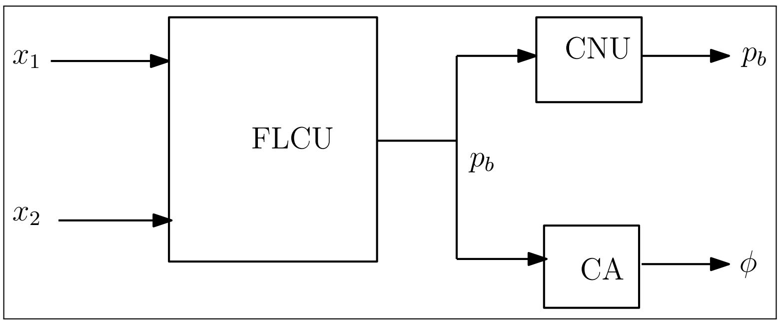


Figure 1: The FLCD Mechanism.

For purposes of fairness, in light of non-responsive flows and network anomalies such as Denial of Service (DoS) attacks and routing loops, which may flood the network as the responsive flows back off, the CA uses p_b to generate a parameter $\phi \in [0, 1]$, where $\phi = p_b^3$. The CA probabilistically picks an arriving packet picked based on the value of ϕ. This packet is compared with a randomly chosen packet from the buffer. If they have the same flow ID, they are both dropped. Otherwise the randomly chosen packet is left unchanged and the arriving packet is queued if the buffer is not full; otherwise it is dropped. As a result, more packets from non-responsive and TCP-unfriendly flows are dropped at the bottleneck link.

4 Evaluation of the Distance Based Triggering and the Dynamic Sampling Mechanisms

The distance based triggering and the dynamic sampling rate mechanisms are implemented at the input of the FLCD mechanism which is used for congestion control in the access point (AP) of the WLAN topology shown in **Figure 2**. The objective is to reduce congestion for traffic flowing from the servers on the high-speed wired/cabled network to the nodes in the bandwidth constrained wireless network. Simulations are implemented on the NS2 simulation platform [14]. In **Figure 2**, servers S1, S2 are connected to the Gateway which is connected to AP. BW1 and

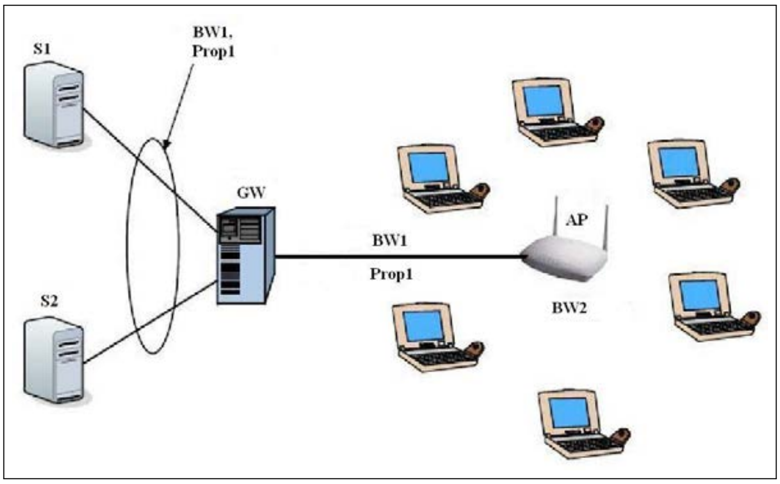


Figure 2: The used WLAN topology.

Prop1 denote the bandwidth and propagation delay between the servers and the Gateway and between the Gateway and the AP while BW2 denotes the wireless channel capacity. The wireless nodes are equidistant from the AP. For brevity, only 5 nodes are shown in **Figure 2** but more than 5 nodes are used in the simulations. Half of the nodes in the wireless network are fixed while the other half consists of mobile nodes. To depict a *high-speed topology*, BW1 and Prop1 are set to 10Gbps and 1ms respectively while the capacity, BW2, is set to 144 Mbps.

Two sets of experiments are conducted. The first one evaluates the impact of the distance based triggering mechanism on the FLCD algorithm. It is aimed at comparing the number of fuzzy system evaluations in a FLCD algorithm with and without the distance based triggering mechanism. The other objective is to find out the impact of the distance based triggering mechanism on system performance. Metrics for system performance include packet loss rate, link utilization, packet delay, and jitter. The second experiment evaluates the efficiency of the dynamic sampling rate mechanism. The objective is to find out if it really manages to guide the system toward the optimal sampling rate. Convergence times are also captured.

Simulations are configured as follows. Each simulation run takes 100 seconds. In all runs, one FTP flow and one web traffic flow are configured to flow from Server 1 to each of the nodes in the WLAN between 5 seconds and 95 seconds leaving enough time for the simulation to start up and also to shut down gracefully. The standard web traffic generator included in the NS2 platform is used with the following parameter settings: an average of 30 web pages per session, an inter-page parameter of 0.8, an average page size of 10 objects, an average object size of 400 packets and a ParetoII shape parameter of 1.002. Each web traffic flow has 4 sessions. UDP traffic is also configured to flow from Server 2 to 10 nodes in the intervals [25s-30s] and [80s-90s]. Parameters of UDP traffic are as follows: Packet Size of 1500bytes, packet interval of 12.5ms and a flow rate of 120kbytes/sec. TCP type is New Reno with a data packet size of 1000 bytes and ACK packet size of 40 bytes. The buffer size is set to 200 bytes.

4.1 Experiment 1 - Testing the impact of the distance based triggering mechanism

Only the distance based triggering mechanism is activated. The number of WLAN nodes is varied by using 10, 20, 30 up to 100 nodes. The value of the normalizing constant Φ in (2) is set to 2×10^{-2} based on several trial runs. After that, simulations are carried out. The overall results are average values over 20 independent runs, which are conducted at each and every testing point. These values are plotted in the graphs along with the error bars representing the 99% confidence intervals of the averages. The FLCD mechanism that employs the proposed triggering mechanism is labeled FLCD+D while the normal one is labeled FLCD. **Figure 3** - **Figure 7** show the results.

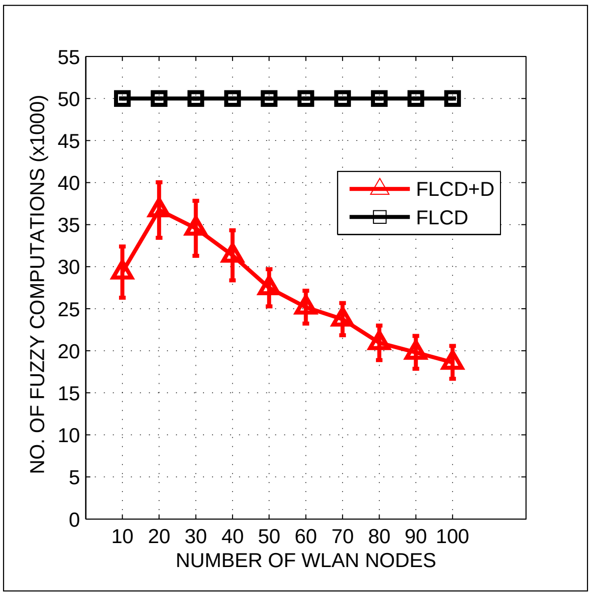


Figure 3: Number of fuzzy computations as congestion level increases

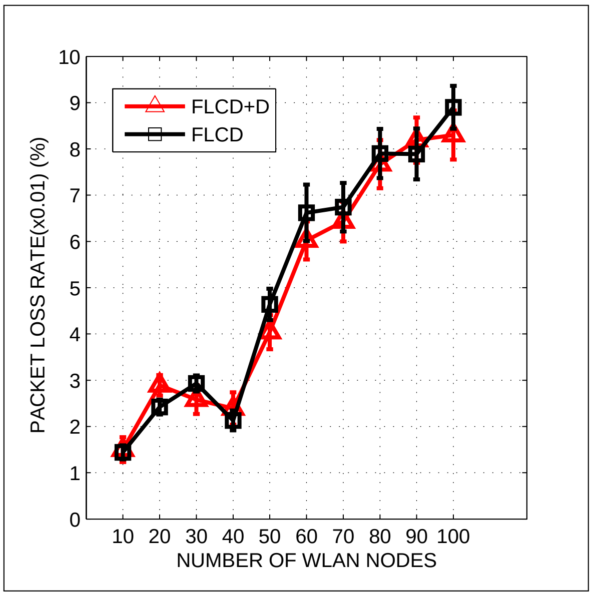


Figure 4: Packet loss rate as congestion level increases

In **Figure 3**, the distance based triggering mechanism reduces the number of fuzzy computations by more than 45% while in **Figure 4** - **Figure 7**, packet loss rate, delay, jitter, and throughput remain virtually the same when the distance based triggering mechanism is employed. These results confirm the fact that a great deal of computing power is lost due to redundant computations in fuzzy systems in communication networks.

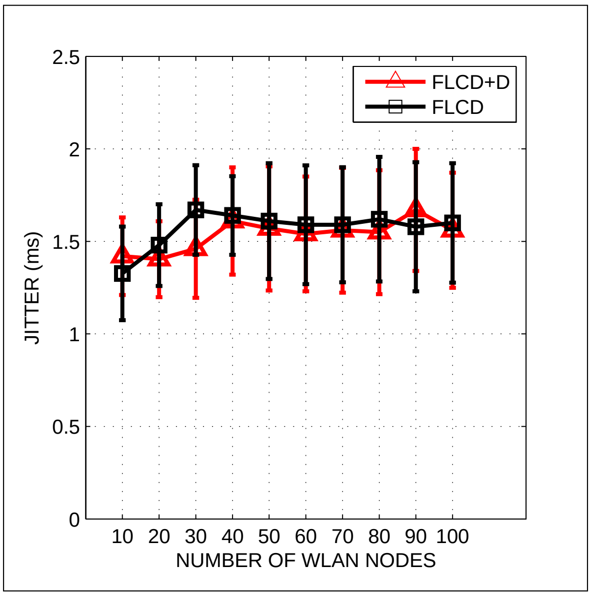


Figure 5: Packet jitter as congestion level increases

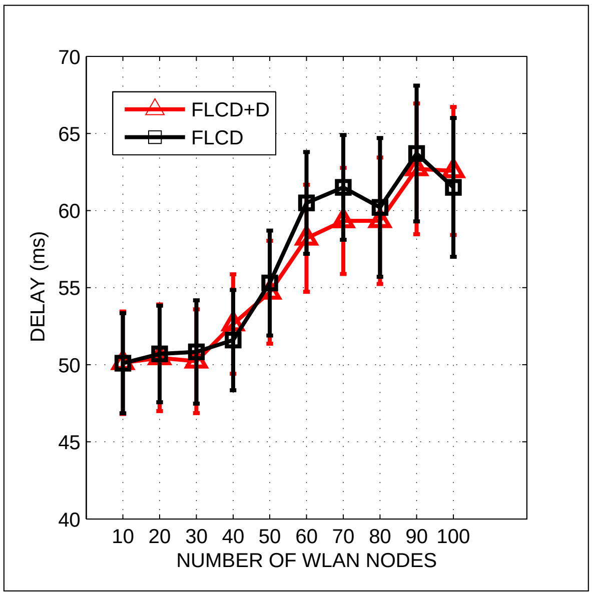


Figure 6: Packet delay as congestion level increases

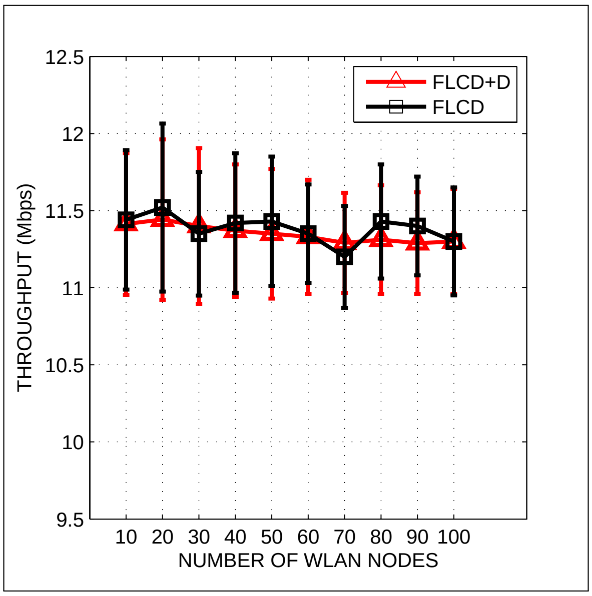


Figure 7: Throughput as congestion level increases

4.2 Experiment 2 - Efficiency of the dynamic sampling rate mechanism

Both the distance based triggering and the dynamic sampling rate mechanisms are activated. The FLCD mechanism that employs the proposed distance based triggering mechanism is labeled FLCD+DS while the normal one is labeled FLCD. The following parameter values are used: $L = 10$, $\alpha = 1 \times 10^{-3}$. These values are determined after several trial runs. Simulations are carried out with 50 nodes on the wireless network on every run. The overall results are average values over 20 independent runs, which are conducted at each and every testing point. In FLCD+DS, the sampling period τ is initialized randomly within (0,1] seconds. Based on the existing congestion control literature [3], $\tau = 2$ msec in FLCD.

Apart from the performance metrics used in Experiment 1, the averages of the converged sampling periods τ and the convergence times are also presented. To statistically establish if two corresponding averages are significantly different from each other, the one-way ANOVA test, at a significance level of 1%, is applied. **Table 1** shows the results. Columns 2 and 3 denote the corresponding average and standard deviation values for FLCD+DS and the FLCD respectively. Column 4 denotes the percentage improvement that FLCD+DS exhibits over FLCD. Statistically significant percentage improvements are shown in bold.

Table 1: Dynamic Sampling Rate Mechanism vs the Conventional approach for 50 nodes.

Metric	**FLCD+DS**	**FLCD**	**Improv.**
Loss rate (%)	0.055 ± 0.0047	0.057 ± 0.0051	3.5%
Delay (sec)	54.59 ± 4.26	52.91 ± 5.54	−3.01%
Jitter (sec)	1.576 ± 0.136	1.59 ± 0.123	0.88%
Throughput (Mbps)	11.47 ± 0.372	11.49 ± 0.415	0.17%
Evaluations	12687 ± 1149	50000	**74%**
τ (msec)	5.05 ± 1.634	2	**152.5%**
Conv. time (msec)	15.69 ± 5.37	–	–

The proposed mechanisms reduce the number of evaluations by more than 70%. The converged sampling period in FLCD+DS is also increased by more than 150%. Again, there is no significant difference between the two approaches in terms of the other key performance metrics; the sampling period converges within the first 20 msec. The dynamic sampling rate mechanism in FLCD+DS performs on-line optimization of τ from random values in (0,1] sec to 5.05 ± 1.634 msec. This leads to drastic reductions in fuzzy system evaluations while overall system performance remains virtually the same as in FLCD.

5 Conclusions and Future Works

To reduce computational cost in fuzzy systems in communication networks, a distance based triggering probability, which governs the triggering rate of the fuzzy system, is proposed. When the system has not been triggered, the previous output(s) are used. A dynamic sampling rate probability, based on the transformed Lorentzian function of the exponentially weighted moving average (EWMA) of the triggering probability, is also proposed. When the EWMA is tending to the extremes, the sampling period is updated in two ways. If the EWMA is higher than 0.5, it implies that the system is under-sampled such that the sampling period is decreased probabilistically. Conversely, if the EWMA is lower than 0.5, it implies that the system is over-sampled such that the sampling period is increased probabilistically. A stopping criterion is also proposed to ensure that this mechanism is switched off when the sampling period has converged.

The triggering probability reduces the number of fuzzy computations in the Fuzzy Logic Congestion Detection (FLCD) in WLANs by more than 45%. Performance, in terms of packet loss rate, delay, jitter, and throughput, however, remains virtually the same. On the other hand, the dynamic sampling rate mechanism leads to more than 150% improvement in sampling rate while the number of fuzzy computations in the FLCD mechanisn in WLANs is reduced by more than 70%. Again, performance in the other key metrics remains virtually the same.

This work shows that the triggering probability can help to alleviate redundant computations in fuzzy systems in communication networks while preserving the performance levels. It also shows that the system can optimize the sampling period just by using the distance information. As part of future work, the proposed mechanism will be implemented in fuzzy systems in wireless sensor/actuator networks.

Bibliography

[1] M. Spott, K. Leiviska, and T. Martin: Roadmap Contribution IBA C Applications in Telecommunications, Multimedia and Services, *European Network on Intelligent Technologies (EUNITE) for Smart Adaptive Systems (SAS)*, July 2004.

[2] Y.L. Chen, J.W. Wang, Y.S. Lin, and J.H. Wen: Combined Fuzzy-Based Power Control with Window-Based Transmission Rate Management in Multimedia CDMA Cellular Systems, *International Journal of Electronics and Communications*, doi:10.1016/j.aeue.2010.04.009, 2 June 2010.

[3] C. Chrysostomou, A. Pitsillides, A. Sekercioglu: Fuzzy Explicit Marking: A Unified Congestion Controller for Best-effort and Diff-serv Networks, *Computer Networks Journal (COMNET)*, Vol. 53, No. 5, pp.650-667, 9 April 2009.

[4] C.N. Nyirenda and D.S. Dawoud: Multi-objective Particle Swarm Optimization for Fuzzy Logic Based Active Queue Management, in *Proc. of the IEEE International Conference in Fuzzy Systems*, Vancouver, Canada, pp. 2231-2238, July 2006.

[5] M. Balakrishnan and E. E. Johnson: Fuzzy diffusion analysis: Decision significance and applicable scenarios, *Proc. of IEEE Military Communications Conference*, no. 1, pp. 2175-2181, October 2006.

[6] M. Marin-Perianu and P. J. M.Havinga: D-FLER: A distributed fuzzy logic enginefor rule-based wireless sensor networks, *Proc. of International Symposium on Ubiquitous Computing Systems (UCS)*, pp. 86–101, 2007.

[7] F. Xia, W.H. Zhao, Y.X. Sun, Y.C. Tian: Fuzzy Logic Control Based QoS Management in Wireless Sensor/Actuator Networks, *Sensors*, Vol. 7, No.12, pp. 3179-3191, 2007.

[8] C.N. Nyirenda, F. Dong, and K. Hirota: Euclidean Distance Based Triggering of Fuzzy Systems in Communication Networks, In proceedings of *the International Symposium on Intelligent Systems (iFAN 2010)*,Tokyo, Japan, September 2010.

[9] P. Whittle, *Probability via Expectation*, 4th ed., Springer-Verlag, New York, 2000.

[10] D.A. McQuarrie: *Mathematics for physical chemistry*, pp. 124, Univ. Science Books, 2008.

[11] J.K. Hunter and B. Nachtergaele: *Applied analysis*, World Scientific, 2001.

[12] B. Richmond and T. Richmond, *A Discrete Transition to Advanced Mathematics*, AMS Bookstore, 2009.

[13] C.N. Nyirenda, D.S. Dawoud, F. Dong, M. Negnevitsky, and K. Hirota: A Fuzzy Multiobjective Particle Swarm Optimized TS Fuzzy Logic Congestion Controller for Wireless Local Area Networks, *Journal of Advanced Computational Intelligence and Intelligent Informatics (JACIII)*, Vol.15, No.1, pp. 41-54, January 2011.

[14] *NS2 network simulator*, http://www.isi.edu/nsnam/ns/, Accessed on 28 June, 2010.

[15] E.H. Mamdani and S. Assilian, An experiment in linguistic synthesis with a fuzzy logic controller, *International Journal of Man-Machine Studies*, Vol. 7, No. 1, pp. 1-13, 1975.

[16] T. Takagi and M. Sugeno, Fuzzy identification of systems and its applications to modeling and control, *IEEE Trans. on Systems,Man and Cybernetics*, Vol.15, 116-132, 1985.

[17] R. Pan, B. Prabhakar, and K. Psounis: Choke - a stateless active queue management scheme for approximating fair bandwidth, *Proc. of INFOCOM*, pp. 942-951, March 2000.

Int. J. of Computers, Communications & Control, ISSN 1841-9836, E-ISSN 1841-9844
Vol. VI (2011), No. 3 (September), pp. 473-502

Human-inspired Identification of High-level Concepts using OWA and Linguistic Quantifiers

M.Z. Reformat, R.R. Yager, Z. Li, N. Alajlan

Marek Z. Reformat
thinkS^2*: thinking Software and Systems laboratory*
Electrical and Computer Engineering
University of Alberta, Canada
E-mail: Marek.Reformat@ualberta.ca

Ronald R. Yager
1. Machine Intelligence Institute
Iona Collage, New Rochelle, NY, USA
E-mail: yager@panix.com
2. Visiting Distinguished Scientist
King Saud University, Riyadh, Saudi Arabia

Zhan Li
thinkS^2*: thinking Software and Systems laboratory*
Electrical and Computer Engineering
University of Alberta, Canada
E-mail: zhanli@ualberta.ca

Naif Alajlan
Advanced Lab for Intelligent Systems Research
College of Computer and Information Sciences
King Saud University, Riyadh, Saudi Arabia
E-mail: najlan@ksu.edu.sa

Abstract: Intelligent agent based system can be used to identify high-level concepts matching sets of keywords provided by users. A new human-inspired approach to concept identification in documents is introduced here. The proposed method takes keywords and builds concept structures based on them. These concept structures are represented as hierarchies of concepts (HofC). The ontology is used to enrich HofCs with terms and other concepts (sub-concepts) based on concept definitions, as well as with related concepts. Additionally, the approach uses levels of importance of terms defining the concepts. The levels of importance of terms are continuously updated based on a flow of documents using an Adaptive Assignment of Term Importance (AATI) schema. The levels of activation of concepts identified in a document that match these in the HofC are estimated using ordered weighted averaging (OWA) operators with linguistic quantifiers. A simple case study presented in the paper is designed to illustrate the approach.
Keywords: concept identification, text documents, ontology, hierarchy of concepts, ordered weighted averaging operator, importance of concepts.

1 Introduction

Keyword-matching is the most popular approach used in current search engines for finding useful texts [1]. However, the meaning of words used as keywords is context-dependent. In

Copyright © 2006-2011 by CCC Publications

addition, users tend to use very few keywords (three or less) what created difficulties in identifying their proper meaning [2, 3].

In this paper we focus on human-like analysis of text, i.e., processes of identification of concepts based on terms/words from a document. The proposed approach reflects a real life scenario where a person links the information in a document to a specific topic by recognizing multiple terms that describe concepts associated with this topic. The terms that appear in a document "activate" terms known to the person that are related to the topic's concepts. In such a way, a net of relevant words is activated. The more words related to a specific concept are activated, the more confidence the person has that the document contains this specific concept. In other words, in concept-based models, we try to "understand" the meaning of a set of keywords and attempt to determine presence of concepts defined by these keywords in documents.

The proposed approach starts with a simple set of keywords provided by the user and representing an entity/item the user is interested in. This set is transformed into a Hierarchy of Concepts (HofC). The HofC is further enhanced with pieces of relevant knowledge obtained from ontology-based knowledge base. This ontology-based knowledge base is created based on ontology domains. Its elements are equipped with values representing levels of contribution of individual keywords towards definitions of concepts. These values are automatically determined using AATI – a schema proposed in ... for determining importance values of keywords based on continuous flow of documents.

The proposed approach relies on the following idea: a process of identification of a concept in a document is equivalent to determining a level of activation of HofC representing this concept. Once a HofC is built and enhanced using an ontology-based knowledge base, it is "checked against" a document. The words found in document are matched with terms and concepts from HofC. Every time a term/concept is found it is activated. The activation levels of terms and concepts are aggregated using Ordered Weighted Averaging (OWA) operator. The weights of OWA are determined based on activation levels and linguistic quantifiers, such as *MOST*, *AT LEAST HALF*, *ALL* and *ABOUT ONE THIRD*. The activation and aggregation propagate in a bottom-up fashion. A final activation level – level of activation of the HofC – is treated as an indicator of how well the concept is present in the document.

The paper is organized in the following way. Section 2 contains a brief description of work done so far in the area of concept-based analysis of text with emphasis on applications of ontology to support definition and identification of concepts. The background is presented in Section 3. Section 4 contains an overview of the proposed approach. The next three section are dedicated to specific elements of that approach: the construction of ontology-based knowledge base is discussed in Section 5, Section 6 contains description of the process of building and enhancing HofCs, while the process of determining activation level of HofC for a document is described in Section 7. A case study with two different linguistic quantifiers is included in Section 8. The papers finishes with conclusions.

2 Related work

There are two most popular approached for concept-based analysis of documents: classifier-based approach, and concept structure-based approach. Because our work is focused on a concept-identification method that involved ontology we do not describe the classifier-based techniques for identification of concepts. Some example of the work related to this topic can be found in [4], [5], [6], and [7].

2.1 Concept Knowledge Bases

It can be noted that classifier-based techniques requires human-labor to set categories, build training sets, update models, etc. Therefore, constructing a background concept structure as a knowledge base to define concepts is a more suitable approach. Some example of such structures are:

- **Synonym thesauri** are treated as repositories of terms"related" to keywords provided by users. For example, Anick [8] has proposed a system that automatically generates an extended condition: "a boolean expression is composed by ORing each query term with any stored synonyms and then ANDing these clusters together." That is, the whole query is ORed together. Each OR term is composed of synonyms from an online thesaurus.

- **Conceptual taxonomy** is seen as a hierarchical organization of concepts. Each concept in a conceptual taxonomy connects to both its superconcepts and its subconcepts. Therefore, it provides a topological structure for efficient conceptual search and retrieval. In the project by Sun Microsystems, a conceptual indexing technique was proposed to automatically generate conceptual taxonomies [9].

- **Ontology** is a model of a domain or a problem. Ontologies can be used to provide formal semantics (meanings, concept-based information) to any sort of information, such as databases, Web documents, etc. A commonly accepted definition states that "an ontology is an explicit and formal specification of a conceptualization of a domain of interest" [10]. In general, an ontology is a representation of a set of concepts and relations between those concepts in a domain.

Ontologies are gaining popularity because of their applications to text analysis. Some of the popular ontologies are: WordNet, SENSUS, and Gene Ontology.

WordNet is an English lexical ontology built at Princeton University, which includes explanations of the terms and relations between terms (synonyms, antonyms, etc.) [11]. As the most popular linguistic ontology, WordNet is used by many researchers for expanded queries [12] [13]. However, the systems applying WordNet suffered from word sense disambiguation (WSD) because of polysemy[1]. This problem has been addressed in [14], [15], and [16]. Additionally, researchers report difficulties in applying linguistic-based ontologies to non-linguistic applications [18].

SENSUS is a natural language-based ontology developed by the Natural Language Group at Information Sciences Institute (ISI) [17]. It is an extension and reorganization of WordNet. It has been used for concept-based retrieval from online yellow pages and product catalogs [18].

Gene Ontology is composed of three structured controlled vocabularies describing gene products with their associated biological processes, cellular components and molecular functions [19]. This ontology is a part of an information retrieval system, KiPar, to facilitate access to the literature relevant to kinetic modelling of a given metabolic pathway in yeast [20].

2.2 Ontology-based Concept Identification

There are different ways to identify concepts in a text using ontology as a knowledge structure. We have divided those approaches into four major categories.

[1] Polysemy means more than one words with the same meanings

Regular Expressions, Rules, and Ontology

In this type of approaches, regular expressions are used to identify concepts from texts. Several predefined rules are then used to aggregate the identifications of single concepts.

Embley has proposed the use of information extraction ontologies, which are formalized over regular expressions [21]. Muller et al. have constructed an ontology-based information retrieval and extraction system for biological literature, which is called *Textpresso* [22] where biological concepts (e.g., gene, allele, cell or cell group, phenotype, etc.) are presented as regular expressions.

Natural Language Processing (NLP) with Ontology

Informally, NLP aims to solve problems by making computers understand and process human language. There are two types of approach in NLP: *deep* approach and *shallow* approach. As Styltsvig stated, "Deep approaches presume access to a comprehensive body of world knowledge. These approaches are not very successful in practice, mainly because access to such a body of knowledge does not exist, except in very limited domains" [23]. Shallow NLP techniques normally rely on simple rules to perform analysis.

Cimiano et al. have presented the LexOnto model consisting of a domain ontology and a corresponding ontology for associating lexical information to entities of the given domain ontology [24]. Morneau et al. have proposed SeseiOnto using NLP to identify concepts from texts [25]. Compared with the LexOnto model, this approach involves more shallow NLP. First of all, SeseiOnto removes articles and prepositions from NLP-based user queries (sentences) by tagging. The remaining words in the queries are matched to the concepts in the ontology.

Utilizing NLP and an ontology may be the most intuitive application to build concept-based systems. However, the discussions on how much NLP should be involved never stops. In LexiOnto [24], NLP exists in the entire process; while in SeseiOnto [25], NLP handles conversion from texts to concept-based structures. Though the NLP techniques allow computer agents to recognize concepts from texts, their complexity cannot be ignored.

Vector Space and Ontology

This type of solution is the most popular. In general, the ontology is utilized to expand the query, and vector space approach is used to evaluate similarity.

Vallet et al. have proposed an ontology based information retrieval system applying a vector-space model to retrieve relevant documents [26]. After terms and concepts are connected through ontology-driven weighted annotations on the documents, a classic vector-space model is utilized to evaluate the relevance between documents and queries. The weights are based on the frequency of occurrence of the instances in each document. Another ontology-based framework for semantic information retrieval has been proposed in [27]. With annotations based on GATE [28] as an information extraction module, metadata for documents are generated. The metadata are extracted concepts from the document text. Based on the term weighting technique $cf \cdot idf$ (concept frequency - inverted document frequency), which is similar to $tf \cdot idf$, concepts in the documents are recognized and indexed. Castells et al. [29] have presented a complete concept-based information retrieval model. In their model, first a set of root ontology classes is constructed from three main base classes: DomainConcept, Topic, and Document. Documents are annotated with concept instances from the ontology. The annotations are weighted based on an adaption of the $tf \cdot idf$ scheme, and the ranking is then achieved through computing similarity values between queries and documents through the classic vector-space approach. In [30], Tomassen et al. has presented an ontology-driven system WebOdIR, where each concept is extended by associating it with a vector of key phrases describing the concept. The system starts from ranking concepts in the ontology according to ontology relevance. It then generates a query for each concept based on relations with other concepts. After submitting the queries to the underlying search engine, a set of documents for each concept is retrieved and then clustered.

Latent Semantic Indexing and Ontology

Latent semantic indexing (LSI) is similar to the vector space approach that presents documents/queries in the manner of vectors. Like vector space approaches, LSI relies on a term-document matrix in which each column represents a document and each row lists frequencies of a term in different documents. LSI uses the singular value decomposition (SVD) to reduce the dimensions of the term-document matrix and approximate the most important part of the original matrix.

Snasel et al. mapped LSI concepts to Wordnet [31]. Wordnet provides synonyms to expand queries: this improves recall through sacrificing precision, because it increased the number of keywords. LSI helps to retrieve the most relevant k terms/concepts from term matrix T_{mk}. The authors argued that the expansion of these k terms instead of all keywords balanced the conflicts between precision and recall.

3 Background

3.1 Semantic Web and Ontology

The Semantic Web [32], as an extension of the World Wide Web (WWW), makes the web content more machine-processable. Tim Berners-Lee, who invented the WWW in the late 1980s, introduced this special vision of the Web, in which the meaning of the content in Web documents would play a much more crucial role than it does today. To accomplish that the Semantic Web needs a more firmly structured language for machine agents to process. The crucial first step has been the adaptation of the language RDF (resource description framework) [33] as a standard for the Semantic Web.

The RDF is more a model than a language; it is designed to present information about Web resources. RDF presents statement in the form of triples, i.e., *subject-predicate-object*. RDF is applied to build information-sharing models. In the triples, the *subject* denotes the resources, the *object* denotes the properties of the subject, and the *predicate* denotes the relations between the subject and the object. The RDF provides a foundation for defining the main data structure of the Semantic Web – ontology.

The most popular definition of an ontology, in the context of the Semantic Web, is "an explicit and formal specification of a conceptualization of a domain of interest" [10]. Ontologies are more than just a vocabulary, they are sources of knowledge of a specific domain. Currently, most of ontologies are implemented in OWL (Web ontology language) which is based on RDF and designed by W3C.

The most important aspect of ontologies used for Semantic Web applications is related to identifying two ontology layers: *the ontology definition layer*, and *the ontology instance layer*. The ontology definition layer represents a framework used for establishing an ontology structure and for defining classes (concepts) existing in a given domain. The structure of an ontology is primarily based on a relation *is-a* between classes. This relation represents a *subClassOf* connection between a superclass and a subclass. In such a way, a hierarchy of classes is built.

The ontology definition contains descriptions of all classes of the ontology. The classes are defined using datatype properties and object properties. Both property types provide an accurate

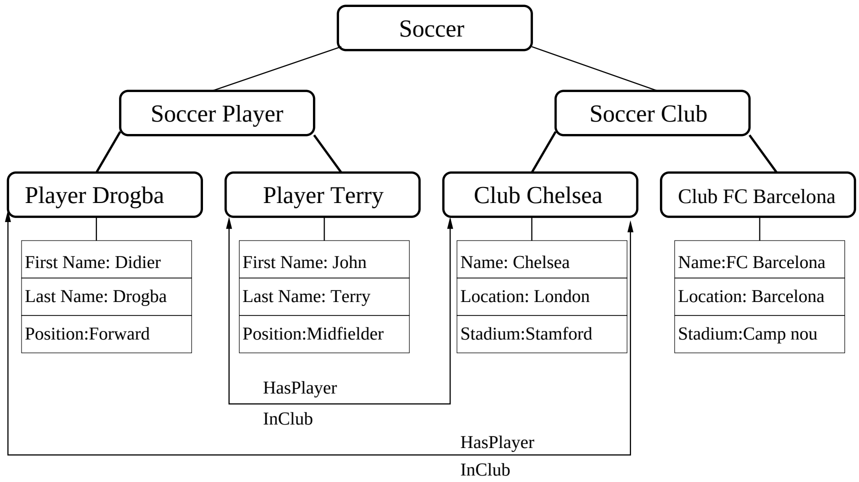


Figure 1: Ontology Snippet

and complete description of a class, as described below:

- The *datatype property* focuses on describing features of a class; datatype properties can be expressed as values of data types such as boolean, float, integer, string, and many more (for example, byte, date, decimal, time);
- The *object property* defines other than *is-a* relations among classes (nodes); these relations follow the notion of the RDF that is based on the triple subject-predicate-object, where: *subject* identifies what object the triple is describing; *predicate* defines the piece of data in the object a value is given to; and *object* is the actual value of the property; for example, in the triple John likes books, John is the subject, likes is the predicate and books is the object.

Both types of property are important for defining ontologies. The possibility of defining class properties and relations between classes creates a versatile framework capable of expressing complex situations with sophisticated classes and the multiple different kinds of relationships existing among them.

Once an ontology definition is constructed, its instances, called individuals, can be built. The properties of classes are filled out: real data values are assigned to datatype properties, and links to instances of other classes (individuals) are assigned to object properties.

A fragment of an ontology is presented in Fig. 1. It contains three classes: *Soccer*, *Soccer_ Club* and *Soccer_ Club*. Both *Soccer_ Club* and *Soccer_ Player* are *subClassOf* of the class *Soccer*. Class *Soccer_ Club* has two individuals: *Club_ Chelsea* and *Club_ FC_ Barcelona*; while Class *Soccer_ Player* has two players as individuals: *Player_ Drogba* and *Player_ Terry*.

Class *Soccer_ Player* is defined by three data properties and one object property, as presented in Fig. 1. The three data properties are *First_ Name*, *Last_ Name* and *Position*. The one object property is *InClub*. Class *Soccer_ Club* is also defined by three data properties and one object property. The three data properties are *Name*, *Stadium* and *Location*. The one object property is *HasPlayer*.

Two individuals of the class *Soccer Player* are *Player Drogba* and *Player Terry*. Individual *Player Drogba* is defined by the term "Didier" as the value of a datatype property *First Name*, by the term "Drogba" as the value of datatype property *Last Name*, by the term "Forward" as the value of datatype property *Position*, and by individual *Club Chelsea* as the value of the object property *InClub*. Similarly, individual *Player Terry* is described by "John" as *First Name*, "Terry" as *Last Name*, "Midfielder" as *Position*, and individual *Club Chelsea* as the value of *InClub*.

Individuals of the class *Soccer Club* (*Club Chelsea* and *Club FC Barcelona*) are also described by three datatype properties and an object property. The three datatype properties listed are *Name*, *Location*, and *Stadium*. The object property is *HasPlayer*, which is the inverse property[2] of the object property *InClub*. The values of the datatype properties and the object property for individuals *Club Chelsea* and *Club FC Barcelona* are shown in Fig. 1.

3.2 Ordered weighted averaging (OWA) operators

Basic principles of OWA

Aggregation of different pieces of information is a common aspect of any system that has to infer a single outcome from multiple facts. An interesting class of aggregation, ordered weighted averaging (OWA) [34] operators, is a weighted sum over ordered pieces of information.

In a formal representation, the OWA operator, defined on the unit interval I and having dimension n (n arguments), is a mapping $F_w : I^n \rightarrow I$ such that

$$F_w(a_1, ..., a_n) = \sum_{j=1}^{n}(w_j \cdot b_j) \tag{1}$$

where b_j is the jth largest of all arguments a_1, a_2, ..., a_n, and w_j is a weight such that w_j is in [0,1] and $\sum_{j=1}^{n} w_j = 1$.

If $id(j)$ is the index of the jth largest of a_i then $a_{id(j)} = b_j$ and $F_w(a_1, ..., a_n) = \sum_{j=1}^{n}(w_j \cdot a_{id(j)})$. If W is an n-dimensional vector whose jth component is w_j, and B is an n-dimensional vector whose jth component is b_j, then $F_w(a_1, a_2, ..., a_n) = W^T B$. In this formulation, W is referred to as the OWA weighing vector and B is called the ordered argument vector.

At the beginning of the 1980s, Zadeh [35] introduced the concept of linguistic quantifiers. Those quantifiers describe a proportion of objects. According to Zadeh, a person knows a vast array of terms that are used to express information about proportions. Some examples are *MOST*, *AT LEAST HALF*, *ALL* and *ABOUT ONE THIRD*. The important issue is to formally represent those quantifiers.

In the mid-1990s, Yager [36] showed how we can use a linguistic quantifier to obtain a weighing vector associated with an OWA aggregation. Yager introduced parameterized families of regular increasing monotone(RIM) quantifiers. These quantifiers are able to guide aggregation procedures by verbally expressed concepts in a description independent dimension. A RIM quantifier is a fuzzy subset Q over $I = [0, 1]$ in which for any proportion $r \in I$, $Q(r)$ indicates the degree to which r satisfies the concept indicated by the quantifier Q [36]. Assuming a RIM quantifier, we can associate with Q an OWA weighing vector W such that for $j = 1$ to n,

$$w_j = Q(\frac{j}{n}) - Q(\frac{j-1}{n}) \tag{2}$$

where n is a number of pieces of information to be aggregated. This expression indicates that the weighing vector W is a manifestation of the quantifier underlying the aggregation process. Using this expression the values of the weighing vector can be obtained directly from the expression representing the quantifier.

For example, let us take a look at the parameterized family $Q(r) = r^p$, where $p \in [0, \infty)$. Here if $p = 0$, $w_1 = 1$ and $w_j = 0$ for $j \neq 1$, and we obtain the *existential* (*max*) quantifier, which makes the OWA operator F closer to an *or* operator; when $p- > \infty$, $w_n = 1$ and $w_j = 0$ for $j \neq n$, and we have the quantifier *for all* (*min*), which makes the OWA operator F closer

[2] If an object property o_1 points object A from object B, and another object o_2 points B from A, o_1 and o_2 are regarded as inverse properties.

to an *and* operator; and when $p = 1$ we have $Q(r) = r$, $w_j = \frac{1}{n}$ and we deal with the quantifier $SOME$.

OWA with Argument Importance

In the previous section we showed how a quantifier Q indicating interaction between pieces of information can be used to calculate an OWA weighing vector W. However, not all pieces of information are of the same importance. A user may desire to assign different weights (importance) to different arguments (pieces of information).

Let $m_i \in [0, 1]$ be a value associated with an argument a_1 indicating its importance. In such a case, let M be a n-dimensional importance vector $[m_1, m_2, ...m_n]$, and the weighing vector W has to be calculated based on both Q and M.

The first step is to calculate the ordered argument vector B (Eq. 1), such that b_j is the jth largest of all arguments a_1, a_2, ..., a_n. Furthermore, we assume μ_j denotes the importance weight associated with the attribute that has the jth largest value. Thus if a_3 is the largest value, then $b_1 = a_3$ and $\mu_1 = m_3$. The next step is to calculate the OWA weighing vector W using a modified version of Eq.(2),

$$w_j = Q(\frac{X_j}{T}) - Q(\frac{X_{j-1}}{T}) \tag{3}$$

where $X_j = \sum_{k=1}^{j} \mu_k$ and $T = X_n = \sum_{k=1}^{n} \mu_k$.

So, X_j is a sum of the importance of the jth most satisfied arguments, and T is the sum of all importance. When all arguments have the same importance, Eq. (3) simplifies to Eq. (2).

3.3 Hierarchy of concepts (HofC)

The idea of representing concepts as a hierarchy was introduced by Yager [37], who represented concepts with atomic attributes, words or other concepts. Using this method, a tree-like structure is established where each vertex is a concept, and terminal vertices (leaves) are attributes. The edges of the hierarchy of concepts (HofC) represent relationships that help to define concepts with other concepts and/or attributes. These edges (connections) are of significant importance to the HofC. If we assume that concept C_1 is defined by two other concepts C_2 and C_3, then the hierarchy will have two edges connecting C_2 with C_1, and C_3 with C_1. The concept C_1 is called a superconcept, and C_2 and C_3 are subconcepts. This also means that *activation* of concepts C_2 and C_3 leads to activation of C_1.

The HofC introduces a very important element, the activation of a superconcept by active sub-concepts is fully controlled by a user. There are two controlling components: importance vector M and linguistic quantifier Q. The vector M indicates the significance of each subconcept in defining a superconcept. In other words, M determines the weight of each participating subconcepts in identifying an activation level of a superconcept. The linguistic quantifier Q guides the aggregation of subconcept activations. Both M and Q determine how activation levels of subconcepts should be combined using the OWA operator.

A simple example of HofC is shown in the Fig. 2. According to the hierarchy, *conceptA* is defined as:

$$conceptA = (conceptB, conceptC, M_A, Q_A)$$

This means that *conceptA* is defined by *conceptB* and *conceptC*. M_A determines the importance of both of *conceptB* and *conceptC* in defining *conceptA*. In this case, it is a two-value vector $M_A = [M_{A-B}, M_{A-C}]$ that implies the importance of activations of both subconcepts during calculation of the activation level of the *conceptA*. The quantifier Q_A can be of any

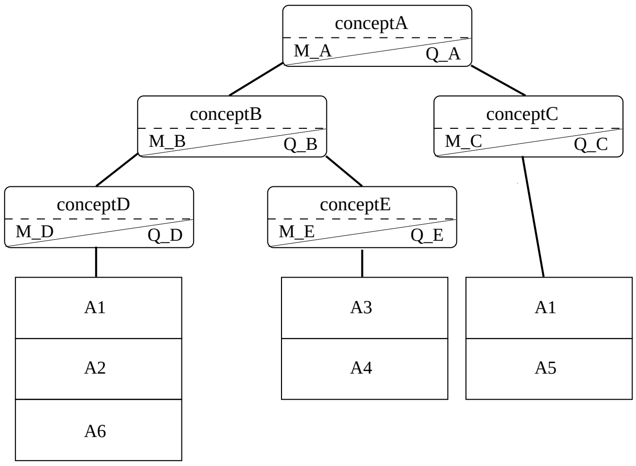


Figure 2: Simple HofC

type, for example, *most* or *some*, and identifies a mechanism of combining activation levels of subconcepts. The rest of the concepts are defined in the following way,

$$conceptB = (conceptD, conceptE, M_B, Q_B)$$
$$conceptD = (A_1, A_2, A_6, M_D, Q_D)$$
$$conceptE = (A_3, A_4, M_E, Q_E)$$
$$conceptC = (A_1, A_5, M_C, Q_C)$$

As we can see, *conceptD*, *conceptE*, and *conceptC* are defined by attributes only. Activation levels of these concepts are calculated by aggregating activations of attributes. Activation of an attribute means that the attribute is present, for example, in a Web page. The aggregation process for each concept is controlled via the M and Q associated with that concept. For *conceptD*, aggregation of activations of attributes A_1, A_2, and A_6 is performed by the OWA operators with a weighing vector determined by M_D and Q_D.

4 Concept Identification Process: Overview

A concept-based information retrieval finds relevant documents based not on simple keyword matching but on concepts that are associated with terms from users' sets of keywords. In order to accomplish that the process of concept identification is composed of the following stages:

- firstly, a user's set of keywords – called hereafter **See**d **K**eywords for **Con**cept Identification: **SeeKCon** – is translated into a structure of concepts; this is done by "mapping" keywords from SeeKCon to definitions of concepts, and constructing an equivalent Hierarchy of Concepts – HofC;
- secondly, the constructed HofC is expanded using knowledge extracted from an Ontology-based Knowledge Base – *ObKB*; SeeKCons provided by users, who usually are not experts, contains simple terms; those terms are "linked" to definitions of concepts included in ontology; the terms, other concepts, and relations associated with those definitions are used to enhance the terms that are included in the equivalent HofC; this process leads to the expended HofC;

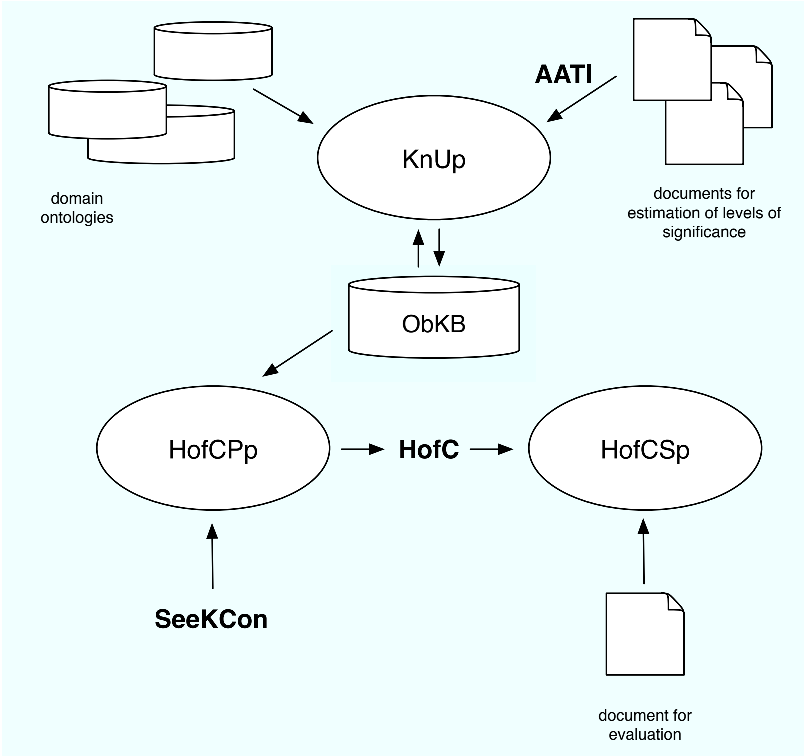


Figure 3: Overview of the Proposed Approach for Concept Identification

- finally, documents are checked how well they satisfy the extended HofC; for an HofC concept defined only by terms (no subconcepts), satisfaction of the concept is decided by the presence of those terms in a document; for an HofC concept defined by a list of terms and subconcepts, satisfaction of this concept is decided by the presence of those terms and the satisfactions of its subconcepts; in the end, the satisfaction of the extended HofC – and indirectly SeekCon – is determined by the satisfactions of concepts defined in this HofC.

An essential part of the proposed approach is a knowledge base containing definitions of concepts. These definitions are composed of (1) terms that describe these concepts together with their importance; (2) related concepts and their importance; (3) relations between terms and concepts. This knowledge base is built around ontology representing domains of interest. Overall, the concept identification process contains a number of activities: a knowledge updating process *KnUp*, a HofC preparation process *HofCPp*, and a HofC satisfaction process *HofCSp*. *KnUP* works with a knowledge base to provide extra definitions of concepts and updates of single keywords importance in defining concepts. *HofCPp* translates users' sets of keywords into HofCs, and expands them with knowledge from a domain knowledge base. *HofCSp* estimated levels of activation of concepts in documents what leads to identification of concepts.

Fig. 3 presents an overview of the proposed concept identification process. Component *KnUp* builds and manages the *ObKB*, the knowledge base. The importance values for terms or concepts in the *ObKB* are obtained by the AATI scheme, as shown in Fig. 3. The *document repository* contains "unknown" to the AATI documents. That is, there is no need for human experts to work on these documents to generate such knowledge as contained concepts, categories they belong to and so on. Therefore, the process of updating the *ObKB* does not stop. This is an outstanding advantage of the approach because no extra human labor is needed to build sample documents and the knowledge in the *ObKB* can be kept up to date.

HofCPp translates users' SeeKCons into equivalent HofCs. It enriches these HofCs with knowledge stored in the *ObKB*. HofCs provide lists of concepts to be identified and their definitions. The definitions of concepts, also known as the knowledge of concepts, include subconcepts, terms, their importance and their relations. As the output of this process, expanded HofCs are utilized for evaluation.

HoCSp determines how HofCs are activated by documents. Documents are semantically annotated first. The annotated terms are pieces of information that satisfy concepts from the HofC. OWA is integrated with the HofC to evaluate the satisfactions of queries by texts in documents. This process is mapped as the aggregation of the activations of the concepts from the HofC.

The proposed process successfully implements concept identification. A term, as a piece of atom information, is not a binary string, but is related to a concept that is a meaningful entity defined in the *ObKB* or in an HofC. That is, it has semantics. This is accomplished based on semantics instead of a simple presence. The main techniques, such as an ontology-based knowledge base (the *ObKB*), an AATI scheme, a HofC format, and OWA operators for aggregation, are applied for this purpose.

- An ontology provides a specification of a conceptualization. That is, it provides lists of concepts, terms and their relations in a domain.
- An AATI scheme assigns importance to terms and in turn to concepts in the *ObKB* by "blindly" reading of documents.
- Users' SeeKCons are organized into concepts in hierarchical structures (HofC).
- OWA operators aggregate the satisfactions of concepts in a HofC to evaluate the satisfaction of the whole HofC. The HofC accepts different linguistic quantifiers that make the approach flexible in defining concepts and representing different user interests.

5 Ontologies and Ontology-based Knowledge Base

In order to identify high-level concepts we need to find out how many terms and subconcepts from concept definitions are present in a document. Concepts which have their definitions "highly activated" are concepts identified in the document. The process of concept identification requires two essential parts:

1) *definitions of concepts* that contain lists of related concepts, subconcepts and terms, as well as relations between them: here, we use ontology as source of such information;
2) *levels of importance of concepts and terms*: they represents levels of contribution of terms/-concepts toward concept activation levels; to address that issue we have developed Adaptive Assignment of Term Importance (AATI) schema.

There are two very important parts of *KnUp*: one that builds ontology-based knowledge base – *ObKB* – from existing domain ontologies, and the other one that updates *ObKB*. The process of updating the *ObKB* is accomplished by the new self-adaptive scheme, AATI (Adaptive Assignment of Term Importance), which assigns importance values to terms/concepts in the *ObKB*, and continuously updates them.

5.1 Ontology-based Knowledge Base (ObKB)

In an ontology, a *class* is described with *properties*. There are two types of properties: datatype property, and object property (Section 3.1). In such a case, a concrete piece of information – called an *individual* – is an instance of a class and is created by assigning values to the *properties*. The ontology as defined above does not contain information required for concept identification in a suitable format. There are few main reasons for that:

1. a HofC contains concepts, while an ontology contains classes and individuals; it is difficult to expand HofCs directly from ontology without a transformation process;
2. HofC connections and ontology connections have different meanings;
3. an ontology does not contain importance (represented by vector M), and linguistic quantifiers (represented by Q) which are necessary for OWA operators to determine levels of activation of concepts contained in HofCs (section 3.2).

Ontology vs. HofC: Component Aspect

In the Semantic Web definition of ontology, classes are abstract and individuals represent concrete information. That is, class defines properties that have no concrete values; while individuals are defined by values assigned to the properties. Unlike an ontology that is designed to store a large amount of organized knowledge, a HofC represents a list of specific information relevant to a high level concept. If a class in an ontology has no individuals, i.e., no concrete information associated with it, it can be provided later by others. But it is not reasonable that a HofC should contain a vertex without a concrete value associated with it. Such a concept is useless because there is no way to identify it in a document. Therefore, both classes as well as individuals from domain ontologies are translated into concepts in the *ObKB*. The following formats and operations pertain here:

- *concepts* are used in the HofC and the *ObKB*; *classes* and *individuals* are used in the ontology;
- *attributes* are used in the HofC and the *ObKB*; *properties* are used in ontology;
- the attributes in the HofC and the *ObKB* are called *terms*, which can be found in documents.
- OWA operators are applied to aggregate activation levels of terms and concepts included in HofCs; OWA operators require M (importance) and Q (linguistic quantifiers)[3].

Ontology vs. HofC: Conncection Aspect

Another important difference between ontologies and HofCs is related to connections between components existing in both structures.

The *ObKB* has the same connections as ontology. These connections have different meanings. For example, let us assume there is a *ObKB* built based on the ontology shown in Fig. 1. The connection between the concept *Player Drogba* and the concept *Soccer Player* is different from the connection between the concept *Player Drogba* and the concept *Club Chelsea*. The former connection means "is-a" or "has-a." That is, *Player Drogba* is-a *Soccer Player* or *Soccer Player* has-a *Player Drogba*. The latter connection means "InClub" or "HasPlayer". That is, *Player Drogba* InClub *Club Chelsea* or *Club Chelsea* HasPlayer *Player Drogba*.

The connections in HofCs represent relative levels in the hierarchy structure, that is, one concept is a subconcept (superconcept) of another concept. HofCs are built to represent SeeKCons where subconcepts contribute to defining their related superconcepts. The connection may be meaningless, because the user can connect any concepts to express his/her interests. For example, let us assume the knowledge in the ontology shown in Fig. 1 is correct. A user can create

[3]Details about M and Q are in section 3.2 and section 3.3

his/her own HofC which has player *Player Drogba* as the superconcept, and *Club FC Barcelona* as the subcocnept if he/she assumes that Player Drogba will transfer to Club FC Barcelona and he/she wants to check if there is any rumour about it. In this case, the satisfaction of *Club FC Barcelona* contributes to the satisfaction of *Player Drogba*, and in turn to the satisfaction of the whole HofC, though it seems the connection should not exist.

In general, subconcepts (superconcepts) are used to describe two connected concepts in a HofC; while the term *related-concepts* is used to describe two connected concepts in the *ObKB*. Overall, we define three types of connections.

1. *TypeI* connection connects concepts that are parts of users' SeeKCons; in other words the concepts provided by users are related to each other through *TypeI* connections; they are the strongest type of connections;
2. *TypeII* connection connects concepts defined in the domain ontology; these connections are "is-a" and "has-a" relations; the first stage of expanding an equivalent HofC is based on this type of connection (Section 6.2);
3. *TypeIII* connection also connects concepts defined in an ontology; however these connections represent all types of relations except "is-a" or "has-a" relation; this is the weakest connection; the second stage of expanding an equivalent HofC is based on this type of connection (Section 6.2).

Mapping from Ontology to ObKB

In a nutshell, the process of building the *ObKB* from domain ontologies consists of a few steps which transforms ontology specific elements into components suitable for construction of HofC. Those steps are:

1. a class from ontology is translated into a concept, however the definitions of properties are not translated;
2. an individual from ontology is translated into a concept, and
   - the values of object properties of this individual are translated into concepts attached to the concept created based on this individual;
   - the values of datatype properties of this individual are translated into attributes attached to the concept created based on this individual;
3. a default value of Q is added to the created concept; the default Q is $MOST$;
4. the default value of M is set as zero for each concept and term.

The *ObKB* is used to expand equivalent HofCs created based on SeeKCons provided by users.

5.2 Estimating Importance Levels

The *ObKB* is a model of a given domain. It contains concepts, terms, and relations that are specific for this domain. The AATI scheme integrated with the *ObKB*, determines importance values that represent levels of contributions of these concepts and terms towards concept activation.

In [39], the AATI was studied and validated through experiments. The AATI scheme is implemented based on the power iteration [38], which is used to find eigenvectors of a matrix in linear algebra. The main steps of AATI implementation are as follows:

1. translate an ontology into the *ObKB*;

2. take a new document;
3. parse the document and annotate it with concept names an terms from the *ObKB*;
4. for each term do one of the following:
   (a) *if Term Weight* is not zero (it means the term has already appeared in documents), take this as its new *Term Weight*;
   (b) *if Term Weight* is zero (the term has not been found in any documents), randomly generate a number between 0 and 1, and assign it as its *Term Weight*;
5. calculate the *Page Value* of this document;
6. update the *Term Weights* of those terms found in the document;
7. normalize *Term Weights* across all terms (make the sum of *Term Weights* equal to 1);
8. if there are no more documents, *STOP*; otherwise go back to Step 2.

The AATI continuously updates *Term Weights* based on available documents. Intuitively, we can say that the *Page Value* is high when terms occurring in a document have high *Term Weights*. At the same time, we can state that the *Term Weight* of a term is high when the *Page Values* of documents that contain the term are high.

5.3 ObKB-based Annotation of Documents

We have developed a Java annotation module based on the UIMA (Unstructured Information Management Architecture) library created by IBM [40], which adds extra information to the texts of documents.

Annotation is the basis for concept identification processes. Terms, which are literal values of the datatype properties – attributes in the *ObKB* – are used to identify concepts in texts. All occurrences of *ObKB* terms in a document are identified, and annotation information is added to them. This information includes its position (begin, end) in a document, the concept it belongs to, and the attribute it belongs to. Annotation connects the *ObKB* with documents: the frequency of a term presented in a document is stored and used later for determination how well the document satisfies the concept, to which the term is attached in the *ObKB*.

Fig.4 shows how a document is annotated. For example, the term “Chelsea” from *ObKB* is the attribute *Name* attached to the concept *Club Chelsea*. The annotation module finds that this term is present twice in the document. So, for the *ObKB* and the annotation module, a string “Chelsea,” which to most computer agents is meaningless (nothing but a binary string), has its own semantics/meaning: it is the name of a soccer club.

6 HofC Preparation Process

The *HofCPp* preprocesses SeeKCons, builds HofCs from them, and expands these HofCs with knowledge from the *ObKB*.

The preprocessing of SeeKCons removes stop words and stems words. Stop words include articles such as “the” and “a” and prepositions such as “in” and “from.” The stemming process removes suffixes and prefixes reducing words to their stems. For example, the term “extended” can be stemmed to “extend.”

The process of building equivalent HofCs translates SeeKCons into concept structures as explained in Section 6.1. Next, knowledge is retrieved from the *ObKB* to enhance the definitions of concepts. The details are presented in Section 6.2.

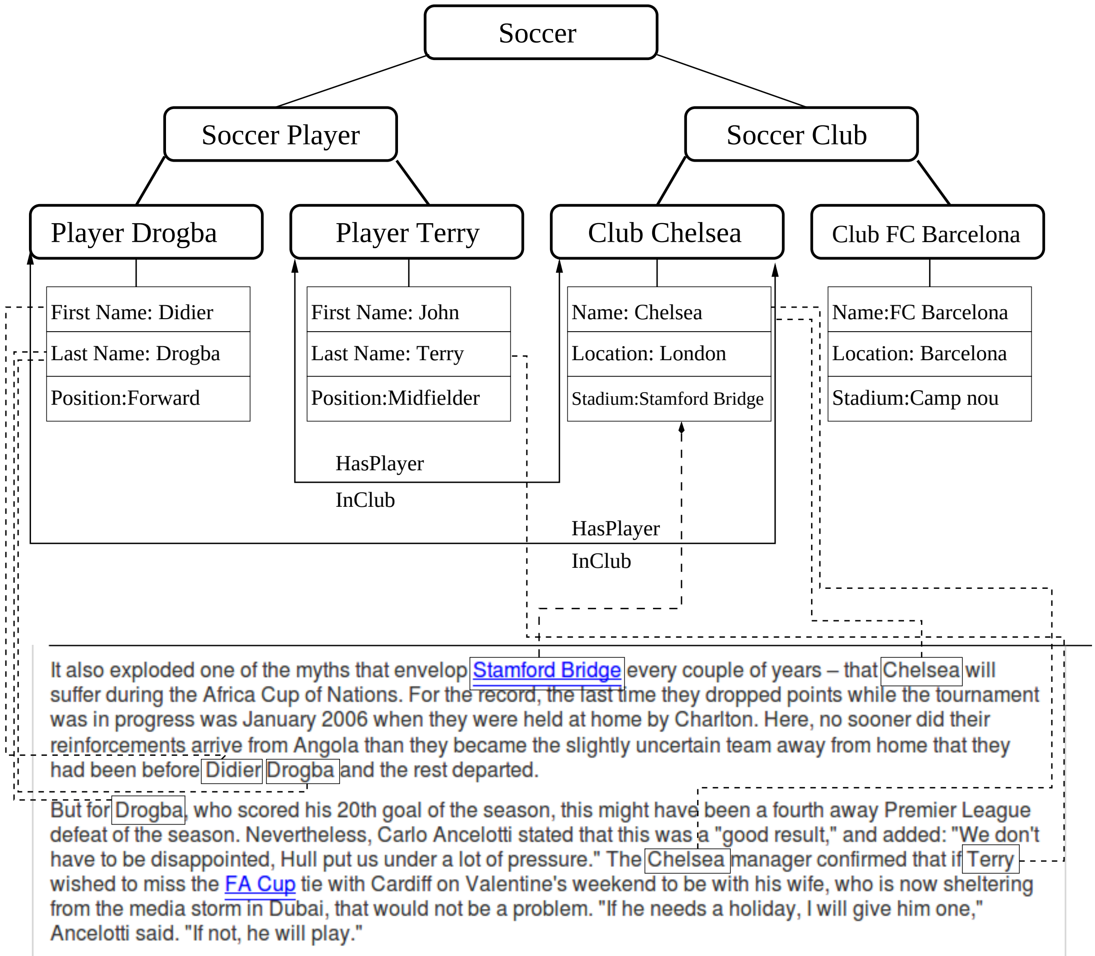


Figure 4: Document Annotation: Example

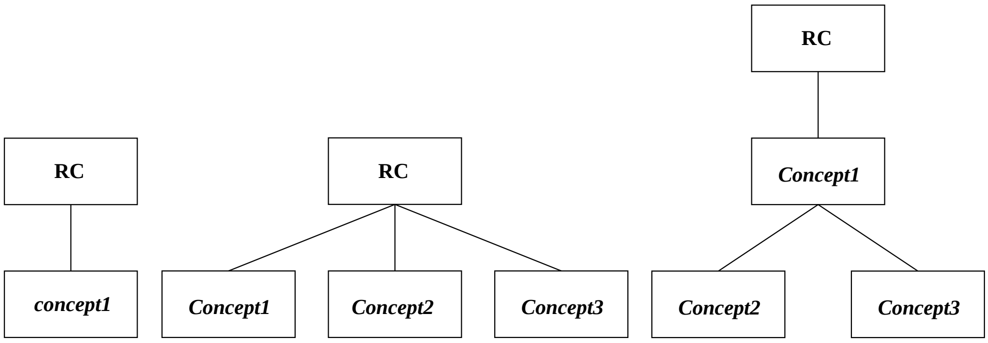
6.1 Representation of Users' Sets of Keywords as HofCs

A SeeKCon is a statement representing the interest of the user. The user provides keywords that describe objects that the user desires to identify in a document. The keywords are represented as a flat structure, that is, there is no indication that the keywords are related to each other, or that some keywords depend on others. This is a simplification suitable for automatic processing but it is not a realistic representation of a human-like way of finding concepts. For a human, all keywords are interconnected. They constitute a network of words representing concepts. The activation of a single word initiates activation of a concept and related concepts.

A SeeKCon is represented by a HofC – a tree-like structure built of concepts (vertices) and terms (terminal vertices/leaves) (Section 3.3). This structured form of a SeeKCon, that resembles a network of words and concepts, provides a more intuitive way of expressing things the user is looking for. A document that satisfies an equivalent HofC contains a high-level concept related to this HofC. The satisfaction of the HofC is a result of an aggregation of satisfactions performed at each vertex of HofC using OWA, and associated with it the linguistic quantifier Q, and the vector M of importance values.

As presented in Section 3.2, a linguistic quantifier Q associated with a vertex is used at the time of the aggregation of satisfactions of terms and other concepts (subconcepts) attached to this vertex. The possible quantifiers are *SOME*, *MORE*, etc. The notion of linguistic quantifiers plays an important role in representing different ways of combining satisfactions.

A vector M (Section 3.2) represents the importance of each term and concept that is taken into consideration during an aggregation process. Not all terms and concepts contribute uniformly to the activation of a concept associated with a vertex of HofC.



(a) Keyword (b) Multiple Keywords (c) Structured Keywords

Figure 5: Sample Root Concepts (RCs)

A SeeKCon represents a high-level concept to be identified in documents. The goal is to find out if a given document contained terms and subconcepts that are part of the equivalent HofC. This translates into a process of estimating a level of activation (or satisfaction) of the HofC. In order to simplify this process, a special concept, called *Root Concept (RC)*, is automatically added to each equivalent HofC. In this case, the activation of the HofC is the same as the activation of the top level concept of the equivalent HofC.

Some possible structures of HofCs are presented in Fig 5. In Fig. 5a, a SeeKCon contains one single concept; in Fig. 5b, a SeeKOnt contains multiple concepts in a flat structure; and in Fig. 5c, a SeeKCon contains a few structured concepts. Activation of these HofCs is simplified to the activation of a single root concept (RC), regardless the complexity of HofCs.

6.2 Building and Expanding HofC

The process of identifying high-level concepts in a document depends on the ability to determine if HofC concepts can be inferred based on the texts of the document. This depends on the contents of the HofC itself – the more terms and concepts are used to build the HofC, the higher the chance of proper evaluation of presence of a high-level concept in a document. We do not expect users to provide comprehensive SeeKCons containing a large number of terms and concepts. The model accepts conventional keyword-based SeeKCons that are automatically translated into HofCs (Section 6.1). That is, the keywords are translated into concepts with a specific hierarchy representing how those concepts are related to each other. The keywords from SeeKCon are regarded as connected with a *typeI* connection.

The names of concepts from an equivalent HofC (the same as keywords provided by the user) are used to retrieve knowledge of equivalent concepts from the *ObKB*. The knowledge includes subconcepts, attributes, linguistic quantifiers (Q), and importance vectors (M). The expansion process is as follows.

for each *concept* from the HofC **do**
  search *ObKB* for a concept equivalent to the *concept*
  if an equivalent concept is found **then**
    1. copy all concepts from the *ObKB* connected via connection $TypeII$ to the found concept as subconcepts of the *concept*

2. copy all attributes of all just copied concepts from *ObKB* as attributes of the subconcepts
3. copy all concepts, and their attributes, connected to the concept found in *ObKB* via $TypeIII$ connection

else

add a string-type attribute "NAME" to the *concept*; the value of this attribute is the name of the *concept*

end if

end for

A HofC includes concepts and attributes arranged in a hierarchy. Identification of a concept in the *ObKB* through a keyword means finding the matching keyword among the concept names or concept attributes in the knowledge base. Once an equivalent concept is found, all knowledge regarding it present in the *ObKB* is used to enhance the HofC.

It is also possible that the user provides other linguistic quantifiers Q different from the default quantifier contained in the *ObKB*. Different linguistic quantifiers mean different ways of aggregation of activation levels (Section 3.2).

7 Identification of Concepts: HofC Satisfaction Process

HofC Satisfaction Process (*HofCSp*) accepts expanded HofC (output from *HofCPp*), and identifies concepts in documents.

An extended HofC contains not only concepts (user's keywords) themselves but also concepts and attributes from the definitions of keyword-based concepts. Those "new" concepts and attributes can be identified in documents.

The evaluation of relevance is mapped as the "activation" of the HofC based on OWA (Section 3.2). The more of the HofC is activated, the higher the chance that the document contains a high-level concept represented by HofC.

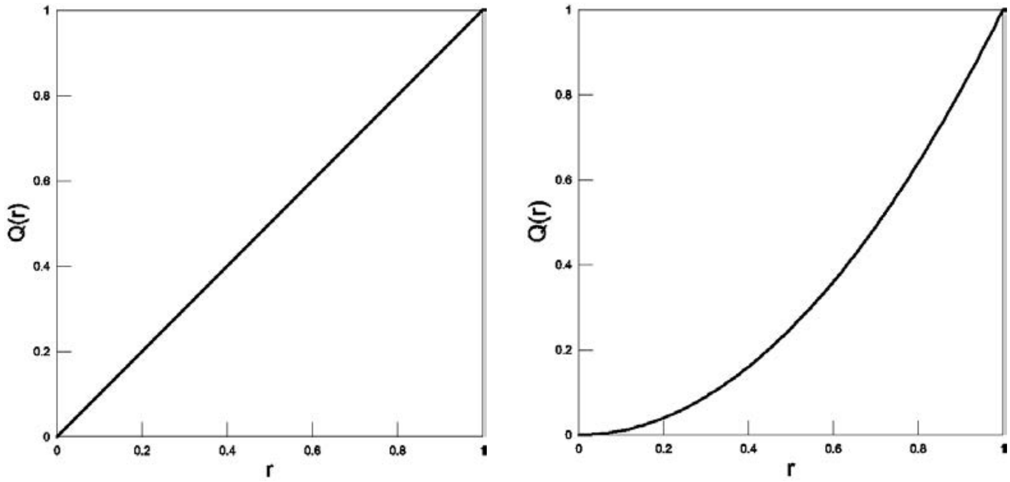
7.1 Satisfaction, Importance and Aggregation weights

Before explaining how *HofCSp* works, three terms used here: satisfaction, importance, and aggregation weights, need to be defined

- *Satisfaction* represents how well a *criterion* in the HofC is satisfied by the text from a document. A *criterion* can be an attribute, a concept or even a complete HofC. The satisfaction of the root concept (RC) (Section 6.1) is equivalent to the satisfaction of the whole HofC.
- *Importance* represents how important a criterion is. With the AATI scheme, each entity (term or concept) in the *ObKB* includes a measure of importance.
- *Aggregation weight* is the weight of a criterion calculated by the OWA operator (Section 3.2). Aggregation weight depends not only on the importance of the criterion but also on the linguistic quantifier. In Section 7.2, the use of linguistic quantifiers and their importance in calculating aggregation weights is explained.

7.2 Linguistic Quantifier

The concept of linguistic quantifiers was introduced by Zadeh [35] in the early 1980s. Some examples are *ALL*, *MOST*, *AT LEAST HALF* and *ABOUT ONE THIRD*. To formally represent those quantifiers, Zadeh suggested using a fuzzy subset $Q(r)$ as a linguistic expression corresponding to a quantifier that indicates the degree to satisfy the concept for any proportion $r \in [0, 1]$.



(a) *some* (b) *most*

Figure 6: Linguistic quantifiers

Yager [36] used the linguistic expression $Q(r)$ to obtain a weighting vector W associated with an OWA aggregation. These quantifiers were then able to guide aggregation procedures by verbally expressed concepts in a description independent dimension.

In the presented work, two linguistic quantifiers are considered:

- $Q(r) = r$ represents the linguistic quantifier *some*, Fig. 6a. The equation means that the aggregation weight of a criterion depends only on its importance value. That is, the relations between concepts do not change their aggregation weights.
- $Q(r) = r^2$ represents linguistic quantifier *most*, Fig. 6b. Using this equation, the aggregation weight of a criterion tends to be large when it is less satisfied (more details are in section 3.2).

7.3 Satisfaction of Concept with Attributes Only

A criterion for a single concept can be an attribute or a subconcept. The simplest situation is when a concept has only attributes attached to it.

The literal values of attributes are called *terms*. The presence (more precisely, frequencies) of a term or terms in a document decides how well the criterion (attribute) is satisfied. Mathematically, the satisfaction of a term s_i is expressed with the following formula:

$$s_i = e^{-\frac{1}{freq_i}} \tag{4}$$

where $freq_i$ is the frequency of the term in the document. The reasons for using Eq. 4 are:

1. It is a monotonically increasing function, so when its frequency increases, the satisfaction increases.
2. The function "saturates" with increasing values of the argument, so if a given term occurs many times in a document, its satisfaction does not overshadow the satisfactions values of other terms.

Each criterion has an assigned value of importance. Assuming there are n different datatype attributes attached to a single concept C, the importance vector M contains n elements, where each element represents the importance of a criterion. The satisfaction vector S also contains n elements, where each element represent the satisfaction level calculated for a criterion based on a given document. In such case, the ordered satisfaction vector S^{prime} (Eq. 1) is formed based on S, in which s_j^{prime} is the jth largest of all satisfactions s_1, s_2, ..., s_n. Furthermore, we assume μ_j denotes the importance weight associated with the attribute that has the jth largest value. Thus if s_3 is the largest value, then $s_1^{prime} = s_3$ and $\mu_1 = m_3$. Based on this, the OWA weighing vector W is obtained by $Q(r)$ in Eq. 5,

$$w_j = Q(\frac{X_j}{T}) - Q(\frac{X_{j-1}}{T}) \tag{5}$$

where $X_j = \sum_{k=1}^{j} \mu_k$ and $T = X_n = \sum_{k=1}^{n} \mu_k$. That is, X_j is the sum of the importance weights of the jth most satisfied arguments, and T is the sum of all importance weights.

So, the satisfaction of concept S_c is:

$$S_c = \sum_{k=1}^{n} w_k \cdot s_k^{prime} \tag{6}$$

7.4 Satisfactions of Concept with Attributes and Subconcepts

As mentioned earlier, a criterion can be an attribute or a subconcept. The satisfaction of a subconcept is the satisfaction of a concept. The calculations of the satisfaction of a criterion based on terms (attributes) and the calculation of the satisfaction of a criterion based on concepts (subconcepts) are different. The calculation of the satisfaction of a criterion based on terms is quite simple, as shown in Eq. 4, while the calculation of the satisfaction of a criterion based on concepts includes more complicated aggregation. When both attributes and subconcepts are attached to the concept, we aggregate separately satisfactions of attributes and satisfactions of subconcepts, and then aggregate the results of both aggregations.

For example, for the *conceptD* in Fig. 7 its satisfaction is calculated using Eq. 6 using frequencies of terms. In the case of the *conceptA* in Fig. 7, its satisfaction is calculated in two steps: firstly, the satisfaction of the $TermA1$ is determined as well as satisfactions of the *conceptB* and the *conceptC*; secondly, all satisfactions are aggregated using OWA with M_A and Q_A.

7.5 Satisfaction of HofC

Activation of subconcepts the HofC is propagated upward as the process to calculate satisfaction of the HofC continues. That is, calculation of the satisfaction of the HofC starts from the attributes/concepts at the lowest level and ends with the concepts at the highest level. The presence of terms/attributes (for example, $TermD1, TermD2, TermC1$, and $TermC2$, see Fig. 7) in a document contributes to the satisfaction of the corresponding concepts ($ConceptD$ and $ConceptC$, respectively, Fig. 7). The aggregation of the satisfactions of those concepts and other terms contributes to the satisfaction of the higher-level concepts (for example, $TermB1$ and $ConceptD$ lead to activation of $ConceptB$, Fig. 7). The process is repeated until the satisfaction of the concept on the top level is calculated.

The structure of HofC determines which terms/attributes and concepts contribute to the satisfactions of which concepts – satisfaction of a single concept is calculated based on the satisfactions of all terms and concepts that are attached (from the bottom) to it. The aggregation of satisfactions is performed at each concept of HofC using OWA (Eq. 6).

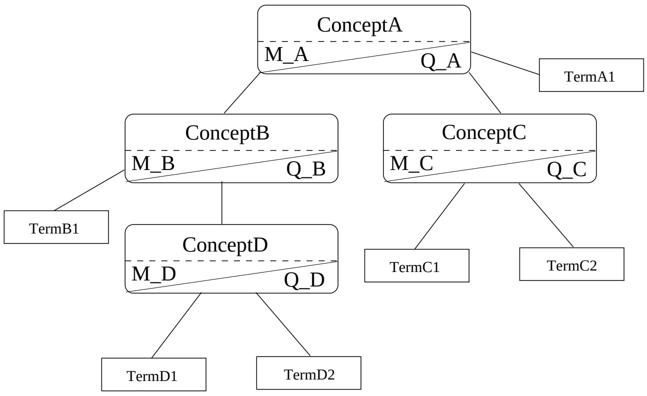


Figure 7: Enhanced HofC

8 Case Study

In order to illustrate the proposed process for concept identification we present a simple case study where two SeeKCons: A:Chelsea and B:Drogba;Terry;Chelsea, are "checked" against two documents Doc.1 and Doc.2. The process is performed with two linguistic quantifiers *MOST* ($Q(r) = r^2$) and *SOME* ($Q(r) = r$). The frequencies of terms in Doc.1 and Doc.2 are listed in Table 1. Roughly, Doc.1 contains more information about *Concept Chelsea* and Doc.2 contains more information about *Player Drogba* and *Player Terry*.

Table 1: Term frequencies in documents

	Doc.1	**Doc.2**
Concept Chelsea		
Name:Chelsea	6	0
Location:London	2	1
Stadium:Stamford	2	1
Concept Player Drogba		
First Name:Didier	0	2
Last Name:Drogba	1	2
Position:Forward	0	1
Concept Player Terry		
First Name:John	0	2
Last Name:Terry	1	3
Position:Midfielder	0	0

8.1 Building and Expanding HofC

Both SeeKCons A and B are used to create HofCs. HofC_A is built from the keyword "Chelsea", Fig. 8. The keyword "Chelsea" itself is a piece of meaningless text for the machine. However, *KnUp* tries to figure out the concept embedded in this keyword. It finds the concept *Club Chelsea* in the *ObKB* because the term "Chelsea" is the value of its attribute *Name*. Then definitions of the concept in the knowledge base are added into the SeeKCon including the

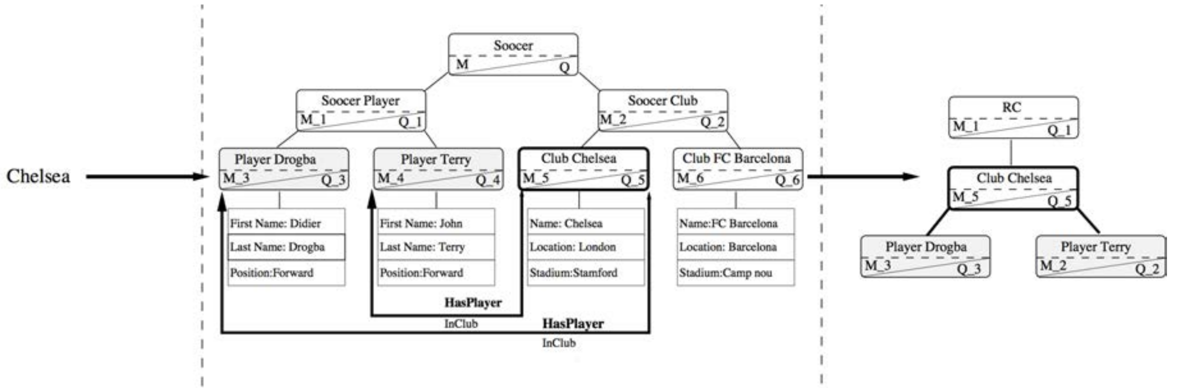


Figure 8: Construction of HofC_A

values of its datatype attributes such as “London” (attribute *Location*) and “Stamford”(attribute *Stadium*), and the values of the object attribute *HasPlayer* associated with the two concepts: concept *Player Drogba* and *Player Terry.* The connection between concept *Club Chelsea* and concept *Player Drogba* and the connection between concept *Club Chelsea* and concept *Player Terry* are all $TypeIII$ connections (Section 6.1 for more details). For simplicity, in this chapter we treat $TypeI$, $TypeII$ and $TypeIII$ connections equally in calculation.

Similarly, HofC_B is built from keywords “Drogba,” “Terry,” and “ Chelsea”, Fig. 9. Since there is no indication about their relations, the process *HofCPp* places the three concepts in a flat (equal level) structure. The connections in this HofC are $TypeI$.

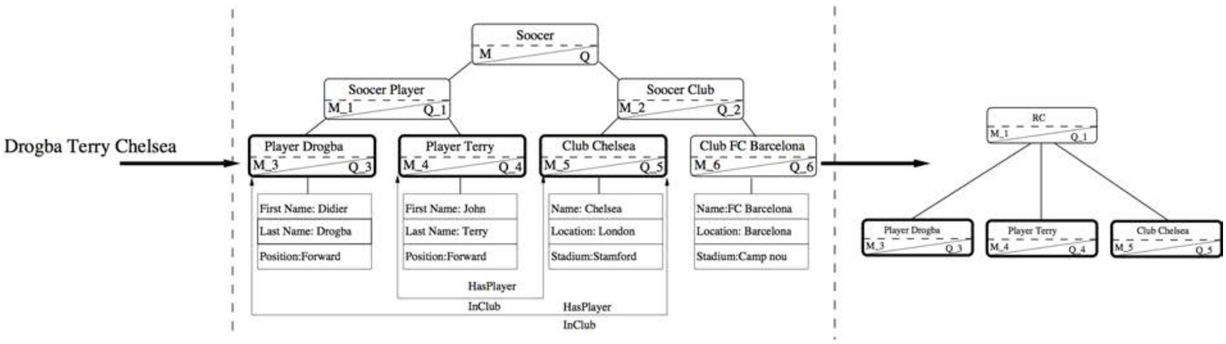


Figure 9: Construction of HofC_B

8.2 Importance of Terms

The importance (TW) of terms is assigned by the AATI. The values are displayed in the table below:

Concept Player Drogba			
Term	Didier	Drogba	Forward
Importance	0.06	0.11	0.02
Concept Player Terry			
Term	John	Terry	Midfielder
Importance	0.03	0.07	0.06
Concept Club Chelsea			
Term	Chelsea	London	Stamford
Importance	0.15	0.02	0.08

8.3 Evaluation

In total, there are four cases using the linguistic quantifiers *MOST* and *SOME*: HofC_A(most), HofC_A(some), and HofC_B(most), HofC_B(some). Cases HofC_(most) and HofC_A(some) utilize HofC_A; cases HofC_B(most) and HofC_B(some) utilize HofC_B. Table 2 shows the design.

Table 2: Case study design

Case	*Linguistic Quantifier*
HofC_A(most)	most
HofC_A(some)	some
HofC_B(most)	most
HofC_B(some)	some

Due to space limitations we show all computational details for two of those cases: HofC_A(some) and HofC_B(most). The results for the rest of the case are given in the table in Section 8.4.

Case HofC_A(most) for Doc. 1

HofC_A, shown in Fig. 8, is generated from the keyword "Chelsea." It is composed of three concepts, which are *Club Chelsea*, *Player Drogba*, and *Player Terry*. The linguistic quantifier in this case is *MOST*.

Fig. 10 displays the aggregation orders in both cases: HofC_A(most) and HofC_A(some). The ellipses in the figure denote the aggregation processes; the numbers in the ellipses are the orders of the processes. As shown in Fig. 10, satisfactions of the concepts *Player Drogba* and *Player Terry* are calculated firstly. As subconcepts of the concept *Club Chelsea*, their satisfactions are aggregated. Then satisfactions of the terms in concept *Club Chelsea* are calculated and aggregate with the satisfactions of the aggregated subconcepts, which is the satisfaction of the whole HofC.

Player Drogba has no subconcept, so its satisfaction is decided by attached terms. The term satisfaction is calculated based on Eq. (4). We calculate the ordered satisfactions of the terms in *Player Drogba*:

Concept Player Drogba	satisfaction	importance
term: Drogba	0.3679	0.11
term: Didier	0	0.06
term: Forward	0	0.02

Because the linguistic quantifier is *MOST*, which is defined by $Q(r) = r^2$. For the concept *Player Drogba*, the sum of the importance of all attached terms is $0.11 + 0.06 + 0.02 = 0.19$. Based on OWA, the weights of its criteria/terms are:

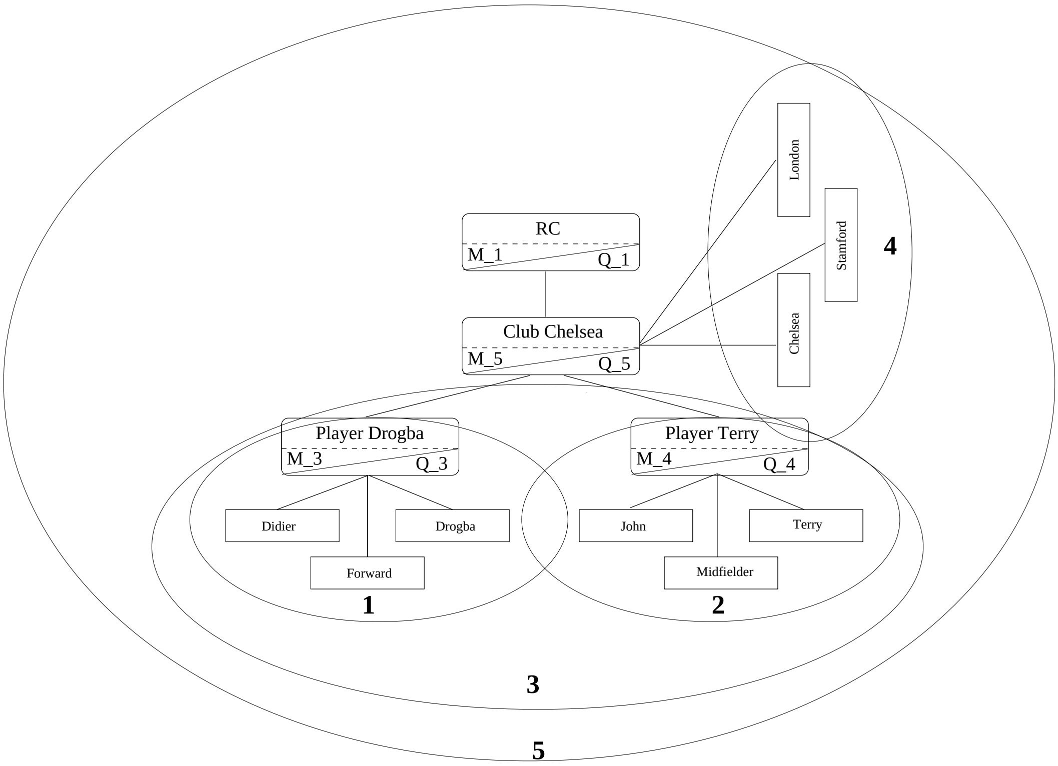


Figure 10: HofC_A: Satisfaction Calculation Process

term Drogba:

$$w_{tDrogba} = Q(\frac{0.11}{0.19}) - Q(\frac{0}{0.19}) = 0.3352$$

term Didier:

$$w_{tDidier} = Q(\frac{0.17}{0.19}) - Q(\frac{0.11}{0.19}) = 0.4654$$

term Forward:

$$w_{tForward} = Q(\frac{0.19}{0.19}) - Q(\frac{0.17}{0.19}) = 0.1994$$

For Doc.1, concept *Player Drogba* has its satisfaction as:

$$\begin{aligned} s_{cDrogba} &= 0.3352 \times 0.3679 + 0.4654 \times 0 + 0.1994 \times 0 \\ &= 0.1233 \end{aligned} \quad (7)$$

Then we calculate the ordered satisfactions of the terms in concept *Player Terry*:

Concept Player Terry	satisfaction	importance
term: Terry	0.3679	0.07
term: John	0	0.03
term: Midfielder	0	0.06

The sum of the importance of all attached terms in concept *Player Terry* is $0.07+0.03+0.06 = 0.16$. Based on OWA, the weights of its criteria/terms are calculated:

term Terry:

$$w_{tTerry} = Q(\frac{0.07}{0.16}) - Q(\frac{0}{0.16}) = 0.1914$$

term John:

$$w_{tJohn} = Q(\frac{0.10}{0.16}) - Q(\frac{0.07}{0.16}) = 0.1992$$

term Forward:

$$w_{tForward} = Q(\frac{0.16}{0.16}) - Q(\frac{0.10}{0.16}) = 0.6094$$

For Doc.1, the concept *Player Terry* has its satisfaction as:

$$\begin{aligned} s_{cTerry} &= 0.1914 \times 0.3679 + 0.1992 \times 0 + 0.6094 \times 0 \\ &= 0.0704 \end{aligned} \tag{8}$$

Defined by the structure of the HofC_A, satisfaction of the concept *Club Chelsea* is obtained by aggregating the satisfactions of its subconcepts (concept *Player Drogba* and concept *Player Terry*) and terms("Chelsea," "London," and "Stamford"). Because the methods for calculateing the satisfaction of subconcepts and terms are different, we calculate the satisfactions separately and then combine them. The details are as follows.

The ordered satisfactions of sub-concepts in concept *Club Chelsea* are:

Concept Player Chelsea	satisfaction	importance	weights
concept: Drogba	0.1233	0.19	0.2947
concept: Terry	0.0704	0.16	0.7053

Satisfaction of the combination of sub-concepts in concept *Club Chelsea* is:

$$\begin{aligned} s^{prime}_{cChelsea} &= 0.1233 \times 0.2947 + 0.0704 \times 0.7053 \\ &= 0.0860 \end{aligned} \tag{9}$$

The ordered satisfactions of the terms in *Club Chelsea* are:

Concept Player Chelsea	satisfaction	importance	weights
term: Chelsea	0.8465	0.15	0.3600
term: London	0.6065	0.02	0.1024
term: Stamford	0.6065	0.08	0.5376

Satisfaction of the combination of the terms in *Club Chelsea* is:

$$\begin{aligned} s"_{cChelsea} &= 0.8465 \times 0.3600 + 0.6065 \times 0.1024 + 0.6065 \times 0.5376 \\ &= 0.6929 \end{aligned} \tag{10}$$

Satisfaction of *Club Chelsea* is calculated by the combination of subconcepts $s^{prime}_{cChelsea}$ and terms $s"_{cChelsea}$ as:

Concept Club Chelsea	satisfaction	importance	weights
comb. of terms	0.6929	0.25	0.1736
comb. of sub-concepts	0.0860	0.35	0.8264

Finally, when applied to Doc.1, satisfaction of the HofC_A is the same as satisfaction of concept *Club Chelsea*:

$$\begin{aligned} s_{cChelsea} &= 0.6929 \times 0.1736 + 0.0860 \times 0.8264 \\ &= 0.1914 \end{aligned} \tag{11}$$

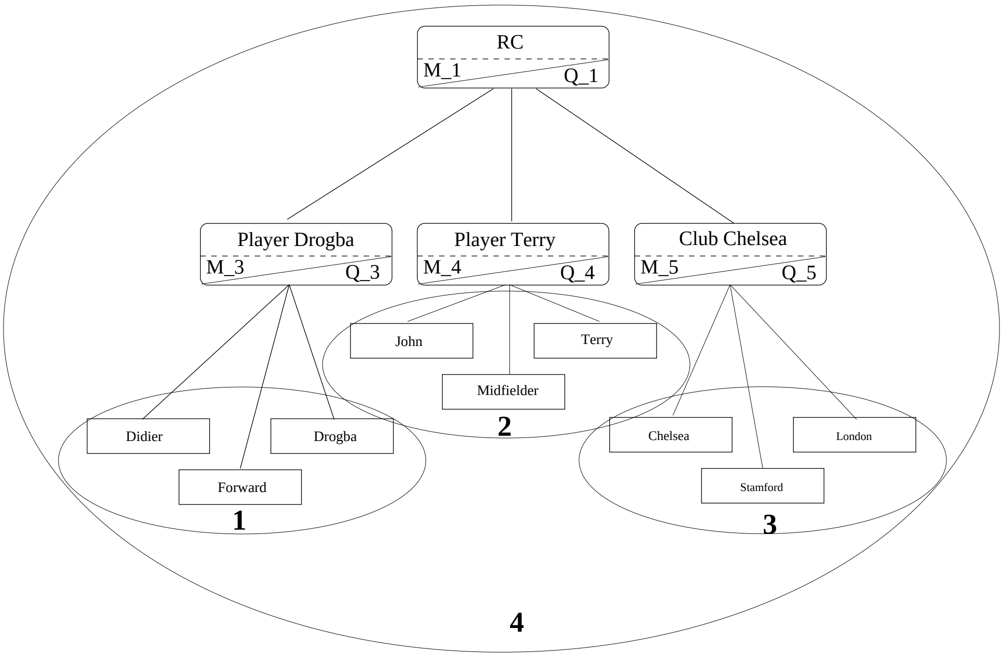


Figure 11: HofC_B: Satisfaction Calculation Process

Case HofC_B(some) for Doc. 1

If a SeeKCon "Chelsea Drogba Terry" is provided, a new HofC is created as shown in Fig. 9.

Fig. 11 displays the aggregation order for HofC_B in cases HofC_B(most) and HofC_B(some). The ellipses in the figure denote the aggregation processes; the numbers in the ellipses are the orders of the processes. As shown in Fig 11, the three single concepts (concepts *Player Drogba*, *Player Terry*, and *Club Chelsea*) are processed first. The satisfaction of the three separated concepts are aggregated as the satisfaction of the whole HofC.

First, we calculate the ordered satisfaction of the terms in *Player Drogba* as:

Concept Player Drogba	satisfaction	importance
term: Drogba	0.3679	0.11
term: Didier	0	0.06
term: Forward	0	0.02

Because the linguistic quantifier is *some*, which is defined by $Q(r) = r$. For *Player Drogba*, the sum of the importance of all attached terms is $0.11 + 0.06 + 0.02 = 0.19$. Based on OWA, the weights of its criteria/terms are:

term Drogba:

$$w_{tDrogba} = Q(\frac{0.11}{0.19}) - Q(\frac{0}{0.19}) = 0.5789$$

term Didier:

$$w_{tDidier} = Q(\frac{0.17}{0.19}) - Q(\frac{0.11}{0.19}) = 0.3158$$

term Forward:

$$w_{tForward} = Q(\frac{0.19}{0.19}) - Q(\frac{0.17}{0.19}) = 0.1053$$

For Doc.1, *concept Drogba* has its satisfaction as:

$$\begin{aligned} s_{cDrogba} &= 0.5789 \times 0.3679 + 0.3158 \times 0 + 0.1994 \times 0 \\ &= 0.2130 \end{aligned} \tag{12}$$

We calculate the ordered satisfactions of the terms in *Player Terry*:

Concept Player Terry	satisfaction	importance
term: Terry	0.3679	0.07
term: John	0	0.03
term: Midfielder	0	0.06

The sum of the importance of all attached terms in *Player Terry* is $0.07+0.03+0.06 = 0.16$. Based on OWA, the weights of its criteria/terms are:

term Terry:

$$w_{tTerry} = Q(\frac{0.07}{0.16}) - Q(\frac{0}{0.16}) = 0.4375$$

term John:

$$w_{tJohn} = Q(\frac{0.10}{0.16}) - Q(\frac{0.07}{0.16}) = 0.1875$$

term Forward:

$$w_{tForward} = Q(\frac{0.16}{0.16}) - Q(\frac{0.10}{0.16}) = 0.3750$$

For Doc.1, *concept Terry* has its satisfaction as:

$$\begin{aligned} s_{cTerry} &= 0.4375 \times 0.3679 + 0.1875 \times 0 + 0.3750 \times 0 \\ &= 0.1610 \end{aligned} \tag{13}$$

The ordered satisfactions of the terms in *Club Chelsea* are:

Concept Player Chelsea	satisfaction	importance	weights
term: Chelsea	0.8465	0.15	0.6000
term: London	0.6065	0.02	0.0800
term: Stamford	0.6065	0.08	0.3200

The satisfaction of the combination of the terms in *Club Chelsea* is:

$$\begin{aligned} s"_{cChelsea} &= 0.8465 \times 0.6000 + 0.6065 \times 0.0800 + 0.6065 \times 0.3200 \\ &= 0.7505 \end{aligned} \tag{14}$$

The ordered satisfactions of the three concepts are:

RC	satisfaction	importance	weights
concept:Chelsea	0.7505	0.35	0.4167
concept:Drogba	0.2130	0.19	0.3167
concept:Terry	0.1610	0.16	0.2667

Therefore, the satisfaction of the RC, which is the satisfaction of HofC_B(some), is:

$$\begin{aligned} s_{cRC} &= 0.4167 \times 0.7505 + 0.3167 \times 0.2130 + 0.2667 \times 0.1610 \\ &= 0.4231 \end{aligned} \tag{15}$$

8.4 Discussion

The results for all cases are summarized in Table 3. There are 12 pieces of information (sum of the frequencies of terms) to satisfy three aimed concepts (Table 1) in both documents. For Doc.1, 10 out of 12 are directly related to *Club Chelsea*, while for Doc.2, 10 out of 12 are directly related to *Player Drogba* and *Player Terry*. In all of the results, the scores for Doc.2 are greater than the scores for Doc.1. This is mainly because different pieces of information have different importance (section 8.2). Moreover, the scores differ when calculated with different hierarchical structures or with different linguistic quantifiers.

Table 3: The results of cases

Case	*Linguistic Quantifier*	*Doc.1*	*Doc.2*
HofC_A(most)	most	0.1914	0.3234
HofC_A(some)	some	0.4231	0.4391
HofC_B(most)	most	0.1977	0.3115
HofC_B(some)	some	0.4231	0.4391

For HoC_A(some) and HofC_B(some) the satisfaction of HofC_A is equal to the satisfaction of HofC_B when the same document is evaluated. That is, with the linguistic quantifier *some*, the hierarchical structures do not affect the calculation of satisfactions. This is because *some* is defined by $Q(r) = r$ (Fig. 6a), which means the aggregation weight of a criterion is decided by its importance value directly. That is, with the linguistic quantifier *some*, the order of aggregation does not change the calculation.

For information aggregated by the linguistic quantifier *most* (defined by $Q(r) = r^2$,Fig. 6b) order is crucial. For example, if there are two pieces of information (i_1 and i_2) with importance $imp_1 = 0.2$ and $imp_2 = 0.3$, respectively, and the order of the aggregation is i_1 then i_2, with *most* their weights are: $wgt_1 = (0.2/0.5)^2 = 0.16$ and $wgt_2 = 1 - (0.2/0.5)^2 = 0.84$, respectively. That is, i_2 obtains larger aggregation weight than i_1. If the order of the aggregation is i_2 then i_1, their weights are: $wgt_2^{prime} = (0.3/0.5)^2 = 0.36$ and $wgt_1^{prime} = 1 - (0.3/0.5)^2 = 0.64$, respectively. That is, i_1 obtains larger aggregation weight than i_2. Thus with *most*, the order of aggregation is important for calculations. Moreover, because the aggregation of pieces of information is in an order from most satisfied to less satisfied, the less satisfied piece of information tends to have larger weights.

The difference in satisfactions between Doc.1 and Doc.2 with the linguistic quantifier *some* is much smaller than that with the linguistic quantifier *most*. For HofC_A, when the linguistic quantifier changes from *some* to *most*, the difference in satisfactions between Doc.1 and Doc.2 changes from $0.4391 - 0.4231 = 0.0160$ to $0.3234 - 0.1914 = 0.1315$. For HofC_B, when the linguistic quantifier changes from *some* to *most*, the difference in satisfaction between Doc.1 and Doc.2 changes from $0.4391 - 0.4231 = 0.0160$ to $0.3115 - 0.1977 = 0.1138$. This is because satisfaction calculated with *most* is affected by hierarchies while satisfaction with *some* is not.

The case studies are good examples of how HOTIR, especially component Eval, ranks documents. Moreover, the detailed calculations illustrate the effects of the linguistic quantifiers *some* and *most* in HOTIR. In the next chapter, the designed experiments examine the performance of HOTIR with *some* and *most*.

9 Conclusions

The constant need for finding relevant documents leads to increased interests in concept-based search techniques. A necessary part of it is concept identification. In the paper we propose a human-inspired approach to detect concepts in text documents. The proposed approach is a fusion of multiple techniques: ontology, term importance evaluation schema, hierarchies of concepts, linguistic quantifiers and the aggregation operator OWA. The following elements are proposed for the implementation of the approach:

- *ObKB* – the ontology-based knowledge base provides supplementary knowledge to define concepts in sets of keywords representing users' interests;
- AATI – the scheme for assigning importance values to terms and concepts;
- Hierarchies of Concepts (HofC) – structures built from concepts, terms/attributes; they are used as representation of users' keywords; they are enhanced with knowledge from *ObKB*;
- linguistic quantifiers and OWA – used to aggregate activation levels of concepts and terms from HofC and estimating overall activation of HofC which is equivalent to identification of concepts in documents.

The case study has been designed to show how the proposed approach performs with different settings. The obtained results and inspection of the calculation processes confirm the ability of the approach to identify concepts in documents in a way similar to humans. A new set of experiments is being planed to verify this in real-world environment.

Bibliography

[1] H.M.Haav and T.-L. Lubi, A survery of concept-based information retrieval tools on the web, *Proceedings of 5th East-European conference ADBIS*, Vilnius, Lithuania, 29-41, 2001.

[2] J. A. Gulla and P. G. Auran and K. M. Risvik, Linguistics in Large-Scale Web Search, *in: Natural Language Processing and Information Systems*, 218-222, 2002.

[3] A. Spink and D. Wolfram and M. B. J. Jansen and T. Saracevic, Searching the Web: the public and their queries, *Journal of the American Society for Information Science and Technology*, 52: 226-234, 2001.

[4] V. Vidulin, M. Lustrek and M. Gams, Training a genre classifier for automatic classification of Web pages, *Journal of Computing and Information Technology*, 15(4): 305-311, 2007.

[5] Y. Aphinyanaphongs, I Tsamardinos, A. Statnikov, H. Douglas and C. F. Aliferis, Text categorization models for high-quality article retrieval in internal medicine, *Journal of the American Medical Information Association*, 12(2): 207-216, 2005.

[6] A. Anagnostopoulos, A. Broder and K. Punera, Effective and efficient classification on a search-engine modeling, *Knowledge and Information Systems*, 16(2): 129-154, 2008.

[7] B. Choi and X. Peng, Dynamic and hierarchical classification of Web pages, *Online Information Review*, 28(2): 139-147, 2004.

[8] P. J. Anick, Adapting a full-text information retrieval system to the computer troubleshooting domain, *Proceedings of the 17th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 349-358, 1994.

[9] W. A. Woods, 1997, Conceptual indexing: a better way to organize knowledge, *Technical Report: TR-97-61, Sun Microsystems, Inc. Mountain View, CA,USA.* http://research.sun.com/techrep/1997/abstract-61.html,

[10] T. E. Gruber, A translation approach to portable ontology specifications, *Knowledge Acquisition*, 5: 199-220, 1993.

[11] G. A. Miller, WordNet: a lexical database for English, *Communications of the ACM*, 38(11): 39-41, 1995.

[12] E. M. Voorhees, Query expansion using lexical-semantic relations, *Proceedings of the 17th Annual ACM SIGIR conference on research and development in information retrieval*, New York, NY, USA, 61-69, 1994.

[13] Z. Gong, C.-W. Cheang and L. H. U, Web query expansion by WordNet, *Lecture notes in computer science*, 3588, 166-175, 2005.

[14] M. Baziz, M. Boughanem, N. Aussenac-Gilles, and C. Chrisment, Semantic Cores for Representing Documents in IR, *Proceedings of 2005 ACM Symposium on Applied Computing*, Santa Fe, New Mexico, 1011-1017, 2005.

[15] S.G. Kolte and S.G. Bhirud, Word Sense Disambiguation Using WordNet Domains, *First International Conference on Emerging Trends in Engineering and Technology, 2008. ICETET '08*, Nagpur, Maharashtra, 1187-1191, July, 2008.

[16] Y.-B. Kim and Y.-S. Kim, Latent Semantic Kernels for WordNet: Transforming a Tree-Like Structure into a Matrix, *ALPIT '08 International Conference on Advanced Language Processing and Web Information Technology*, 76-80, July, 2008.

[17] K. Knight and R. Whitney, Ontology Creation and Use: SENSUS, Information sciences institute, University of Southern California, http://www.isi.edu/natural-language/resources/sensus.html, 1997.

[18] N. Guarino, C. Masolo and G. Vetere, OntoSeek: Content-based Access to the Web, *IEEE Intelligent Systems*, 14(3): 70-80, 1999.

[19] S. E Lewis, Gene Ontology: looking backwards and forwards, *Genome Biology*, 6(1): 103, 2005.

[20] I. Spasic, E. Simeonidis, H. L. Messiha, N. W. Paton, and D. B. Kell, KiPar, a tool for systematic information retrieval regarding parameters for kinetic modelling of yeast metabolic pathways, *Bioinformatics*, 25(11): 1404-1411, 2009.

[21] D.W. Embley, Towards Semantic Understanding - An Approach Based on Information Extraction Ontologies, *Proceedings of the Fifteenth Australasian Database Conference (ADC2004)*, 3-12, 2004.

[22] H.-M.Muller, E. E. Kenny, and P. W. Sternberg, Textpresso: An Ontology-Based Information Retrieval and Extraction System for Biological Literature, *PLoS Biology*, 2(11): 1984-1998, 2004.

[23] H. B. Styltsvig, Ontology-based information retrieval (PHD Thesis), Denmark, 2006.

[24] Ph. Cimiano, P. Haase, M. Herold, M. Mantel, and P.Buitelaar, LexOnto: A Model for Ontology Lexicons for Ontology-based NLP, *Proceedings of the OntoLex (From Text to Knowledge: The Lexicon/Ontology Interface) workshop at ISWC07 (International Semantic Web Conference)*, Busan, South-Korea, 2007.

[25] M. Morneau, G. W. Mineau, and D. Corbett, LexOnto: A Model for Ontology Lexicons for Ontology-based NLP, *Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence*, Hongkong, China, 449-455, 2006.

[26] D. Vallet, M. Fernandez and P. Castells, An ontology-based information retrieval model, *Proceedings of 2nd European Semantic Web Conference, ESWC 2005*, Grete, Greece, 455-470, June, 2005

[27] O. Dridi and M.B. Ahmed, Building an ontology-based framework for semantic information retrieval: application to breast cancer, 3rd International Conference on Information and Communication Technologies: From Theory to Applications, ICTTA 2008. , April, 1-6, 2008.

[28] H. Cunningham, D. Maynard, K. Bontcheva , V. Tablan, C. Ursu, M. Dimitrov, M.Dowman, N. Aswani, I.Roberts, Y. Li, A. Shafirin and A. Funk, Developing Language Processing Components with GATE, The University of Sheffield, April, 2009, http://gate.ac.uk/sale/tao/index.html.

[29] P. Castells, M. Fernandez and D. Vallet, An Adaptation of the Vector-Space Model for Ontology-Based Information Retrieval, *IEEE Transactions on Knowledge and Data Engineering*, 19(2): 261-272, 2007.

[30] S. L. Tomassen, Searching with Document Space Adapted Ontologies, in *Emerging Technologies and Information Systems for the Knowledge Society*, 5288, 513-522, 2008.

[31] V. Snasel, P. Moravec and J. Pokomy, WordNet Ontology Based Model for Web Retrieval, *Proceedings of the International Workshop on Challenges in Web Information Retrieval and Integration*, Tokyo,Japan, 220-225, April, 8-9, 2005.

[32] G. Antoniou and F. van Harmelen, A Semantic Web Primer(2nd Edition), *The MIT Press, Cambridge, Massachusetts, London, England*, 2008.

[33] B. McBride, RDF Primer, Aug., 2008, http://www.w3.org/TR/REC-rdf-syntax.

[34] R.R. Yager, On ordered weighted averaging aggregation operators in multi-criteria decision making, *IEEE Transactions on Systems, Man and Cybernetics*, 18: 183-190, 1988.

[35] L.A. Zadeh, A computational approach to fuzzy quantifiers in natural language, *Computers and Mathematics with Applications* 9: 149-184, 1983.

[36] R.R. Yager, Families of OWA operators, *Fuzzy Sets and Systems*, 59: 125-148, 1993.

[37] R.R.Yager, A Hierarchical Document Retrieval Language, *Information Retrieval*, 3: 357-377, 2000.

[38] Wikipedia, Power Iteration, May, 2008, http://en.wikipedia.org/wiki/Power_method.

[39] Z.Li and M.Reformat, A Schema for Ontology-based Concept Definition and Identification, *International Journal of Computer Applications in Technology*, 38(4): 333-345, 2010.

[40] D. Ferrucci and A. Lally, UIMA: an architectural approach to unstructured information processing in the corporate research environment, *Natural Language Engineering*, 10(3-4): 327-348, 2004.

Int. J. of Computers, Communications & Control, ISSN 1841-9836, E-ISSN 1841-9844
Vol. VI (2011), No. 3 (September), pp. 503-529

Adaptive Non-singleton Type-2 Fuzzy Logic Systems: A Way Forward for Handling Numerical Uncertainties in Real World Applications

N. Sahab, H. Hagras

Nazanin Sahab, Hani Hagras
University of Essex, United Kingdom
The Computational Intelligence Centre
E-mail: nsahab@essex.ac.uk, hani@essex.ac.uk

Abstract: Real world environments are characterized by high levels of linguistic and numerical uncertainties. A Fuzzy Logic System (FLS) is recognized as an adequate methodology to handle the uncertainties and imprecision available in real world environments and applications. Since the invention of fuzzy logic, it has been applied with great success to numerous real world applications such as washing machines, food processors, battery chargers, electrical vehicles, and several other domestic and industrial appliances. The first generation of FLSs were type-1 FLSs in which type-1 fuzzy sets were employed. Later, it was found that using type-2 FLSs can enable the handling of higher levels of uncertainties. Recent works have shown that interval type-2 FLSs can outperform type-1 FLSs in the applications which encompass high uncertainty levels. However, the majority of interval type-2 FLSs handle the linguistic and input numerical uncertainties using singleton interval type-2 FLSs that mix the numerical and linguistic uncertainties to be handled only by the linguistic labels type-2 fuzzy sets. This ignores the fact that if input numerical uncertainties were present, they should affect the incoming inputs to the FLS. Even in the papers that employed non-singleton type-2 FLSs, the input signals were assumed to have a predefined shape (mostly Gaussian or triangular) which might not reflect the real uncertainty distribution which can vary with the associated measurement. In this paper, we will present a new approach which is based on an adaptive non-singleton interval type-2 FLS where the numerical uncertainties will be modeled and handled by non-singleton type-2 fuzzy inputs and the linguistic uncertainties will be handled by interval type-2 fuzzy sets to represent the antecedents' linguistic labels. The non-singleton type-2 fuzzy inputs are dynamic and they are automatically generated from data and they do not assume a specific shape about the distribution associated with the given sensor. We will present several real world experiments using a real world robot which will show how the proposed type-2 non-singleton type-2 FLS will produce a superior performance to its singleton type-1 and type-2 counterparts when encountering high levels of uncertainties.
Keywords: Type-2 Fuzzy logic, non-singleton Fuzzy Logic Systems, interval type-2 fuzzy logic systems, Type-2 non-singleton type-2 FLS.

1 Introduction

Professor Lotfi Zadeh introduced fuzzy sets in 1965 [1]. He has written several papers since then, but his comprehensive paper on the concept of linguistic variables [2] became very famous as a reference on fuzzy logic theory. He introduced fuzzy logic for the application of approximate

Copyright © 2006-2011 by CCC Publications

[a] [b]

Figure 1: Two examples of domestic products that used fuzzy logic. (a) AEG washing machine [8]. (b) Hitachi air conditioner [9].

reasoning and mentioned the need of a humanistic system whose behavior is strongly influenced by human judgment, perception or emotion for the purpose of computing with words to solve the problems of human-oriented fields such as philosophy, psychology, politics, law, sociology and literature. Professor Zadeh also described the use of his fuzzy logic approach in [3] in describing the behaviors of too complex or too ill-defined systems. This has inspired control engineers from over the world to investigate the power of fuzzy logic to control applications which are difficult to mathematically model.

In 1975, the first Fuzzy Logic Controller (FLC) was developed by Mamdani and Assilian for controlling a steam engine [4]. Since then Fuzzy Logic Systems (FLSs) have been applied with great success to many real world applications. It was reported in 1995 that over the last two decades, the number of applications of FLSs have dramatically increased. Several industrial applications of fuzzy logic were reported from which we can mention cement kilns [5], steel furnaces [5], aircraft engine control, steam turbines, power supply controllers, etc [6]. FLSs have also been used domestically in elevators, washing machines [6], fridges [7], air conditioners, automobiles [5], automatic transmission, camera autofocus control, etc [6].

Figure 1 shows two examples of fuzzy logic usage in domestic appliances which are the AEG washing machine [8] (shown in Figure 1a) and the Hitachi air conditioner [9] (shown in Figure 1b). According to the John Lewis website about a top brand washing machine from AEG and reporting about the fuzzy logic washing machine, the website mentions "fuzzy logic circuit detects when the laundry is out of balance and re-jig it accordingly, ensuring minimum wear and tear to the drum bearings. The fuzzy logic also detects half loads, if too much detergent has been added and adds extra rinses if required" [8]. Hence, for the past thirty years fuzzy logic and its applications became embedded in our everyday environments.

Fuzzy logic laid the basis for a successful method to model uncertainty, vagueness, and imprecision [10]. The FLSs are general knowledge base systems [11] with linguistic rules that can be constructed using the knowledge of experts in a given field. During more than four decades from the invention of fuzzy logic, great progress in using FLS has been achieved. While traditionally type-1 FLSs have been widely employed in real world applications, recent years have shown a rapidly growing interest in the research and application of type-2 FLSs. This is because, it has become apparent that type-1 FLSs cannot fully handle the large amounts of uncertainties encountered in real world environments and applications. Interval type-2 FLSs have shown to provide better performance when compared to type-1 FLSs in applications with high levels of uncertainty. The difference in performance has been attributed to the nature of interval type-2 fuzzy sets which can better account for the uncertainties as they incorporate a Footprint of Uncertainty (FOU) and they can be assumed to embed a large number of type-

1 fuzzy sets. However, the majority of the type-2 FLSs [12] handle the linguistic and input numerical uncertainties using singleton interval type-2 FLSs that mix the numerical and linguistic uncertainties to be handled only by the linguistic labels type-2 fuzzy sets. Input numerical uncertainties are associated with the noise, imprecision and uncertainty associated with sensors and input devices. However, the linguistic uncertainties are associated with human words and perceptions. Hence, it seems paradoxical to use singleton type-2 FLSs to handle the input numerical uncertainties, as this ignores the fact that if input numerical uncertainties were present, they should affect the incoming inputs to the FLS. Thus we cannot treat the incoming FLS inputs as perfect signals as in the case of singleton FLSs. Hence, there is a need to employ non-singleton FLSs to handle the input numerical uncertainties by modeling the inputs as fuzzy inputs rather than crisp values.

Although, there are many papers on type-1 FLSs, quite limited number of papers considered the usage of type-1 non-singleton FLSs such as [13]- [25] and even fewer number considered type-2 non-singleton FLS [25]- [31]. The work done so far in (type-1 and type-2) non-singleton FLSs use predefined shapes for the uncertainty distribution affecting the FLSs inputs (mainly Gaussian and triangular) which might not reflect the real uncertainty distribution associated with the given sensor [32]- [35]. For example in the case of a sonar sensor, the numerical input uncertainties depend on many factors such as the shape and type of objects reflecting the sonar signal as well as wind speed, humidity and temperature. Moreover, the uncertainty distribution also varies with the measured values where for the case of a sonar sensor, the amount of uncertainty affecting measurement readings at 20cm distance could be much less than the uncertainties affecting the measurement at large distances such as 3m. Hence, it is not fair to consider that there is specific shape for the uncertainty distribution (triangular, Gaussian, etc.) with fixed parameters (average, standard deviation, etc.) for all the measured values. [12]

In this paper, we will present some of the theoretical basis for generating an adaptive type-2 fuzzy input which is better able to represent the encountered numerical uncertainty at a given measurement. The non-singleton type-2 fuzzy inputs are dynamic and they are automatically generated from data and they are piece-wise linear type-2 fuzzy sets which are changing for different measurements to model the encountered uncertainty. Piece-wise linear fuzzy sets were used in [36] for the representation of membership functions. However, we use piece-wise linear fuzzy sets to represent the type-2 fuzzy input variables. In this paper, we will present an overview on how the adaptive type-2 input based non-singleton interval type-2 FLS can operate in real time. We will present real world experiments using a mobile robot which will show how under high input uncertainty levels, the non-singleton type-2 FLS can give a good performance and outperform its singleton type-2 and type-1 FLSs counterparts.

Section 2 will review the singleton type-1 FLS. Section 3 will review the singleton interval type-2 FLS while Section 4 will review non-singleton FLSs. Section 5 will provide an overview on modeling a data based type-2 fuzzy input. Section 6 will provide an overview on the proposed approach for type-2 non-singleton type-2 FLS with the type-2 fuzzy input. Section 7 will provide an overview on the system operating in real-time in real world environments. Section 8 provides the experiments and results and Section 9 provides the conclusions and future work.

2 Singleton Type-1 FLS

The first generations of the FLSs were singleton type-1 FLSs. In the singleton type-1 FLS, both the antecedents and the consequents of the FLS are type-1 fuzzy sets and the input values are crisp and precise values. A type-1 fuzzy set, A, for a variable x $\in$ X is characterized by a type-1 membership function $\mu_A(x)$ which is limited between [0,1] [32]:

$$A = \{(x, \mu_A(x)) | \forall x \in X\} \tag{1}$$

In FLSs, membership functions are associated with terms that appear in antecedents or consequents of the FLS rules [32]. Figure 2 shows an example of the linguistic labels, Close, Moderate and Far modeled by type-1 fuzzy sets considered for the antecedent of a type-1 FLS.

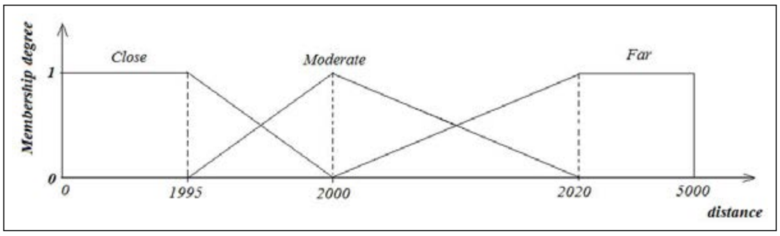


Figure 2: An example of type-1 fuzzy sets which is also used in our experiments as antecedent membership functions.

Figure 3 shows the structure of singleton type-1 FLS (the usual Mamdani type) where the crisp inputs are first fuzzified to input type-1 fuzzy sets which trigger the inference engine and rule base to generate output type-1 fuzzy sets which are then defuzzified to obtain a final crisp output of the FLS. In singleton FLSs, the incoming inputs to the FLS are considered as crisp numbers which assume that the incoming inputs to FLSs are not associated with noise or uncertainty.

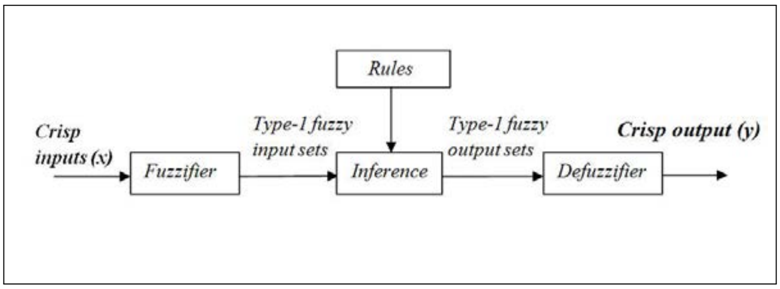


Figure 3: Structure of a singleton type-1 FLS [32].

Considering having p inputs $x_1 \in X_1, ..., x_p \in X_p$ and one output y $\in$ Y, where F_k^l is the antecedent membership function for input k and rule l, and G^l is consequent membership function for rule l, the lth rule is shown as:

$$R^l : If\ x_1 is\ F_1^l\ and\ ...\ x_p\ is\ F_p^l,\ then\ y\ is\ G^l. \qquad l = l, ..., M \tag{2}$$

Singleton fuzzifier maps a crisp input $\overrightarrow{x} = (x_1, ..., x_p) \in X_1 \times X_2 \times ... \times X_p \equiv \overrightarrow{X}$ into a fuzzy set A_x in $\overrightarrow{X}$. Singleton fuzzifier is nothing more than a fuzzy singleton where the output of the singleton fuzzifier related to input k and rule l is $\mu_{F_k^l}(x_k\prime)$ when the incoming input is $x_k = x_k\prime$. The firing strength of rule l can be found as follows [34]:

$$\mu^l = \mu_{F_1^l}(x_1\prime) * \mu_{F_2^l}(x_2\prime) * ... * \mu_{F_p^l}(x_p\prime) \tag{3}$$

Where * represented the t-norm operation. The defuzzification produces a crisp output for the FLS from the fuzzy sets that appear at the output of inference engine. There are many different types of defuzzifiers such as centroid, height, and center of sets defuzzifiers [32]. For example for the center sets defuzzifier (which we have used in our experiments) when c^l is the centroid of the related consequent fuzzy set to rule l and T is t-norm operation, the defuzzified output $y_{cos}(\overrightarrow{x})$ could be written as follows [32]:

$$y_{cos}(\overrightarrow{x}) = \frac{\sum_{l=1}^{M} c^l T_i^p = 1\mu_{F_i^l}(x_i)}{\sum_{l=1}^{M} T_i^p = 1\mu_{F_i^l}(x_i)} \tag{4}$$

There are many sources of uncertainty facing the FLS in dynamic unstructured real world environment and many real world applications; some of them are: [37]

• Uncertainties in inputs to FLS due to noise, change of the conditions of observation, etc.;

• Linguistic uncertainties as the meaning of words that are used in antecedent and consequent linguistic labels can have different meaning to different experts [32];

• Uncertainties associated with the use of noisy training data;

• Uncertainties associated with the change in the operation conditions of controllers;

• Uncertainties in control outputs which can be due to the wear and tear associated with the FLS actuators.

All of these uncertainties translate into uncertainties about membership functions. Type-1 FLS cannot handle linguistic and numerical uncertainties associated with dynamic situations, as they use precise membership functions that the user believes capture uncertainties. When the environment change, the chosen type-1 fuzzy sets might not be appropriate any more. This can cause degradation in the performance of the FLS and the designer might end up wasting time to frequently tune type-1 FLS to deal with various uncertainties [37], [32]. Hence, it was claimed that type-1 fuzzy sets might have limited capabilities to directly handle uncertainties related to dynamic uncertain situations. That sounds paradoxical because the word fuzzy has the connotation of uncertainty. The paradox about type-1 fuzzy set has been known for a long time, that a representation of fuzziness is made using membership values which are themselves crisp numbers [38].

A type-2 fuzzy set is characterized by a fuzzy membership function, i.e. the membership value for each element of this set is a fuzzy set in [0,1], unlike a type-1 fuzzy set where the membership value is a crisp number in [0,1] [32]. The membership functions of type-2 fuzzy sets are three dimensional and include a Footprint Of Uncertainty (FOU), it is the new third-dimension of type-2 fuzzy sets and the FOU that provide additional degrees of freedom that can make it possible to directly model and handle the uncertainties [32]. Therefore FLSs that use type-2 fuzzy sets have the potential to handle higher uncertainty levels than their type-1 counterparts [37], [38].

3 Singleton Interval Type-2 FLS

Type-2 fuzzy logic was also introduced by Zadeh in 1975 in [2], but little works [39] were published about them until the middle of previous century. Until then few people were studying on type-2 FLS, as people were learning what to do with type-1 fuzzy sets [39]. Type-2 FLSs are the extension of the type-1 FLS. In type-2 FLSs, the antecedent and/or consequent membership functions of type-2 FLS are type-2 fuzzy sets [40].

A type-2 fuzzy set $\widetilde{A}$ is characterized by a type-2 membership function $0 \leq \mu_{\widetilde{A}}(x,u) \leq 1$, where x $\in$ X and u $\in J_x \subseteq$ [0,1], as [32]:

$$\widetilde{A} = \{((x,u), \mu_{\widetilde{A}}(x,u)) | \forall x \in X, \forall u \in J_x \subseteq [0,1]\} \tag{5}$$

J_x is called primary membership of x, where $J_x \subseteq [0,1]$ for $\forall x \in X$. The uncertainty in the primary membership grades of a type-2 fuzzy membership function consist of a bounded region, that we call the Footprint of Uncertainty (FOU) of a type-2 membership function. It is the union of all primary membership grades [41]. Figure 4a shows the type-1 membership function which is totally crisp and precise. Figure 4b shows the primary membership function of a type-2 fuzzy set and related FOU. The FOU of a type-2 membership function handles the rich variety of choices that can be made for a type-1 membership function [37], [43]. A type-2 FLS can be thought of as a collection of a large number of type-1 FLS's. When the antecedent and consequent membership grades in the type-2 FLS have a continuous domain, the number of embedded type-1 FLS's is uncountable [40].

Interval type-2 FLS [33] is a special case of general type-2 FLS when the secondary membership functions are interval sets. The FOU can be described in terms of upper and lower membership functions. Figure 4b illustrates an upper membership function and a lower membership function that are two type-1 membership functions that are bounds for the FOU of a type-2 fuzzy set. The upper membership function is denoted by $\overline{\mu}_{\widetilde{A}}(x)$, and the lower membership function is denoted by $\underline{\mu}_{\widetilde{A}}(x)$ [32]. Figure 4c also shows the 3-D view of an interval type-2 fuzzy set.

In a type-2 FLS the rules are defined as Equation (2), however the antecedent and/or consequents are type-2 fuzzy sets. We can mark type-2 fuzzy sets with sign "$\sim$" to be different from type-1 fuzzy sets as [32]. The structure of singleton type-2 FLS is shown in Figure 5 where crisp inputs are first fuzzified into input type-2 fuzzy sets (in singleton fuzzification) which then activate the inference engine and the rule base to produce output type-2 fuzzy sets. These output type-2 fuzzy sets are then processed by the type-reducer which combines the output sets and then performs a centroid calculation, which leads to type-1 fuzzy sets called the type-reduced sets [32]. The defuzzifier can then defuzzify the type-reduced type-1 fuzzy outputs to produce crisp outputs to be fed to the actuators. More information about the singleton interval type-2 FLS can be found in [32], [37], [40].

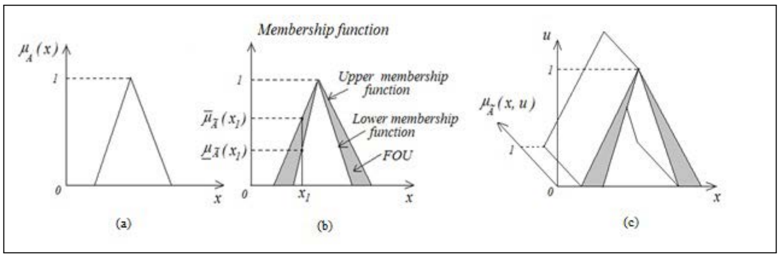


Figure 4: (a) A type-1 fuzzy set. (b) The primary membership function of a type-2 fuzzy set. (c) A 3-D view of interval type-2 fuzzy set.

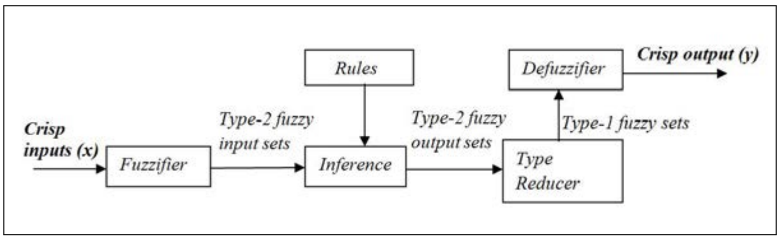


Figure 5: Structure of singleton interval type-2 FLS [37], [32].

Using type-2 fuzzy sets to represent the inputs and outputs of a FLS has many advantages when compared to the type-1 fuzzy sets, we summarize some of these advantages as follows [37]:

- The type-2 fuzzy sets of membership functions are themselves fuzzy and include a FOU, so they can model and handle the linguistic and numerical uncertainty associated with the inputs and outputs of the FLS in changing and dynamic unstructured environments and handle the difficulty associated with determining the exact membership functions for the fuzzy sets [41]. Hence, FLSs based on type-2 fuzzy sets have the potential to produce better performance that type-1 FLS.

• Using type-2 fuzzy sets to represent the FLC inputs and outputs will result in the reduction of the FLC rule base when compared to using type-1 fuzzy sets as the uncertainty represented in the footprint of uncertainty in type-2 fuzzy sets lets us cover the same range as type-1 fuzzy sets with smaller number of labels and the rule reduction will be greater when the number of the FLC inputs increases [32].

• Each input and output will be represented by a large number of type-1 fuzzy sets which are embedded in the type-2 fuzzy sets [42], [41]. The use of such a large number of type-1 fuzzy sets to describe the input and output variables allows for a detailed description of the analytical control surface and gives much smoother control surface and response [37].

During the past decade, type-2 fuzzy sets and systems have become very popular [44]. There are many applications done using type-2 FLSs such as video streaming control [45], induction motor control [46], two-axis motion control [47], hot strip mill temperature control [19], marine diesel engines, DC-DC converter, mobile robots, ambient intelligent environment, etc. [37]. In these applications, the type-2 FLSs have outperformed their type-1 counterparts.

Although, there are many papers on interval type-2 FLSs, the vast majority of these papers used singleton FLSs which mix the numerical and linguistic uncertainties to be handled only by linguistic labels type-2 fuzzy sets. This looks paradoxical, as if numerical uncertainties were present, they should affect the incoming inputs to the FLS, and therefore we cannot treat the incoming FLS inputs as perfect signals in the case of singleton FLSs [48]. Hence, there was a need to consider non-singleton type-2 FLSs to handle the incoming numerical uncertainties.

4 Non-Singleton FLS

The major difference between a non-singleton FLS and a singleton FLS is in the fuzzification part. Zadeh introduced fuzzification in 1973 as "the operation of fuzzification (or, more specifically, support fuzzification) has the effect of transforming a non-fuzzy set into a fuzzy set or increasing the fuzziness of a fuzzy set" [3]. Since 1975 when Mamdani used FLS in a steam engine [4], the majority of the FLSs were using singleton FLS employing singleton fuzzification. The reason for widespread employment of singleton fuzzification is because of its simplicity and speed of computation which allows for real time operation. In singleton fuzzification, inputs are considered to be crisp data and the effect of noise, imprecision, and uncertainty is ignored, while the non-singleton fuzzification models the FLS inputs as fuzzy sets.

FLSs in real world environments [48] can face high levels of uncertainties, which can also be categorized into linguistic and numerical uncertainties. The linguistic uncertainties is related to human words and perceptions, and the input numerical uncertainties are related to noise, imprecision and uncertainty related to input devices and sensors. In non-singleton FLSs, the numerical uncertainties can be modeled and handled by non-singleton fuzzy inputs and the linguistic uncertainties will be handled by antecedent fuzzy sets to represent the linguistic labels. Dealing with numerical uncertainty is especially important in engineering applications in which measurement instruments, sensors, or input devices are employed to produce inputs to FLSs.

There has been some research from instrumentation and measurements researchers in modeling uncertainty measurements using type-1 fuzzy sets and even type-2 fuzzy sets. There are a number of non-ideal situations affecting measurement results such as measurement method, measurement condition, and operator [49]. This has caused the issue of measurement uncertainty in which the true value is unknown. Ferrero [49] also mentioned that up to that date, the most widely known and assessed theory to deal with incomplete information was probability. But probability just covers random uncertainty which is a part of incomplete knowledge and it is needed to consider both random and unknown systematic effect [49].

Quite limited number of papers (compared with large number of paper in FLSs) used non-singleton type-1 fuzzification. Since 1997 Mouzouris and Mendel presented some closed forms of non-singleton fuzzifier [32], [34]- [35]. Liang and Mendel [32]- [33] in 2000 presented some closed form of non-singleton type-2 FLS for the case of Gaussian type-2 fuzzy inputs with uncertain standard deviation, and some limited number of papers have used type-1 non-singleton type-2 FLS in their applications such as [19]- [25] and more limited number of authors have used type-2 non-singleton type-2 [25]- [31]. However, all the papers assume a predefined shape (such as triangular or Gaussian) as the fuzzy input variable. However, these predefined shapes might not accurately represent the real shape of data distribution related to the input device or sensor. Three categories of non-singleton FLSs have been seen so far in papers which are: Non-singleton type-1 FLS, type-1 non-singleton type-2 FLS, and type-2 non-singleton type-2 FLS. In the first two categories, the input is modeled as a type-1 fuzzy set and in the third category type-2 fuzzy sets are used to model incoming inputs to the FLS.

In this paper, we will present a new approach to obtaining an interval type-2 fuzzy input variable [50] directly from data with no assumption of any predefined shapes for inputs. We will then present our novel adaptive type-2 non-singleton type-2 FLS.

5 Overview of Modeling a Data-Based Type-2 Fuzzy Input

We will start by presenting our data based method to model non-symmetric non-specified shapes of convex type-2 fuzzy inputs, which are directly obtained from data and we do not assume a specific shape about the uncertainty distribution with the given sensor [50].

A summary of the process is as follows:

1. Record 10,000 continuous measurements (at a specific measurement condition) and obtain a histogram for that condition to show how many times a measurement value is recorded.

2. Obtain a normal convex type-1 fuzzy set from the data distribution as follows [50]:

- Obtain the first point from the left side in which the distribution value is in its maximum (shown as max_l in Figure 6).

- Consider the graph from the starting point to the first left maximum (max_l) and connect the points at each distribution value if the value is increasing, otherwise keep the line horizontally at previous value.

- Continue the previous step until getting to the point max_l.

- Then Obtain the first point from right side in which the distribution value is in its maximum (max_r).

- Consider the graph from the right ending point to the first right maximum (max_r) and connect the points at each distribution value if the value is increasing or fixed, otherwise keep the line horizontally at the previous value.

- Continue the previous step until getting to the point max_r.

- To having a convex set, the value from max_l to max_r is kept constant equal to maximum distribution value ($distribution\_no(max_l) = distribution\_no(max_r)$).

- Divide the whole set by the maximum value to change the scale of the function to the acceptable interval [0,1] as a normal fuzzy set.

3. Repeat the first and second steps whilst changing different possible conditions which could affect the input measurements. As for the case of a sonar sensor [50], we considered changing temperature, sound noise, and wind while recording measurements at a fixed distance.

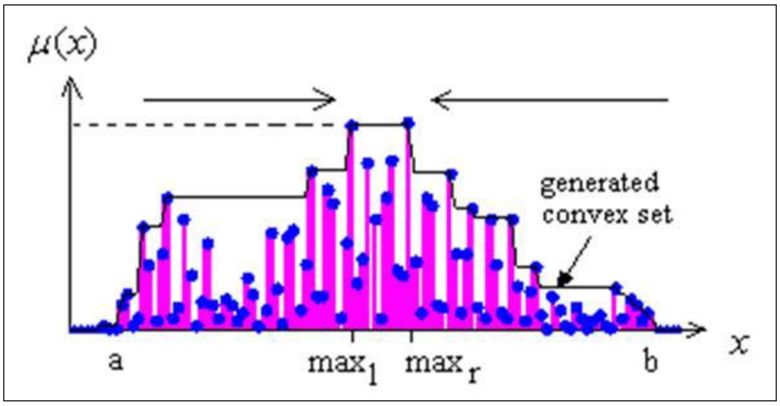


Figure 6: Obtaining a convex function from data distribution [50].

4. Create a convex type-2 fuzzy set from the various type-1 fuzzy input variables, as shown in Figure 7. The method to generate type-2 fuzzy sets from type-1 fuzzy sets is similar to the approach in [51], but we considered non-specified shapes instead of triangular type-1 fuzzy sets. The details of the approach were described in [50]. Another method (which can be done in a discrete form) is to obtain the upper membership function by taking the maximum t-conorm of all the type-1 fuzzy sets, whilst keeping the area from the left point with maximum membership value $\overline{(max_l)}$ to the right point with maximum membership value $\overline{(max_r)}$ constant at one. Lower membership function could be also considered as the minimum of all the type-1 fuzzy sets. [50]

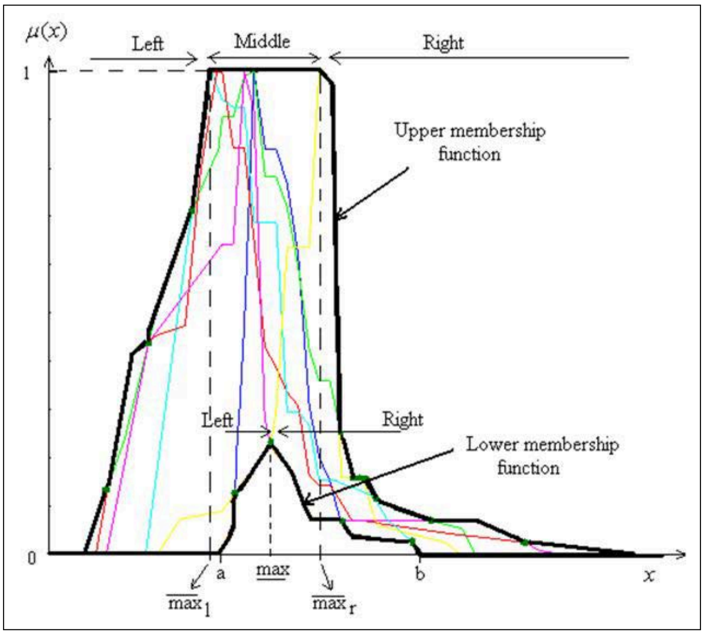


Figure 7: Obtaining type-2 fuzzy input from type-1 fuzzy inputs, the colorful lines show each type-1 fuzzy set and upper membership function and lower membership functions of the resulted interval type-2 fuzzy input is shown by black thick lines [50].

Figure 8 illustrates a type-2 fuzzy input variable modeled from the case of one of the sonar sensors at around 3000mm from a cylindrical object. The black lines define type-2 fuzzy input variable, and the colorful dash lines show the type-1 fuzzy sets from which the type-2 fuzzy input is created.

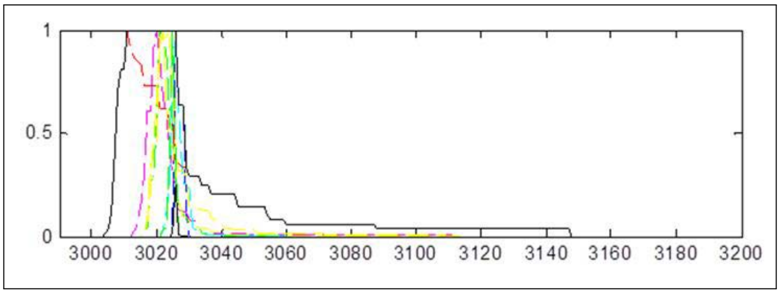


Figure 8: A sample of the modeled type-2 fuzzy inputs at the distance of around 3000mm [12].

In the following section, we present a summarized view on how to employ the type-2 fuzzy input in a type-2 non-singleton type-2 FLS.

6 Type-2 Non-Singleton Type-2 FLS

The type-2 non-singleton type-2 FLS employs the created type-2 fuzzy input variables to handle the encountered numerical uncertainty, and uses type-2 fuzzy sets to model the linguistic labels like "close" , "moderate", and "far" which represent membership functions of antecedents to handle the encountered linguistic uncertainty. Figure 9 illustrates the structure of the type-2 non-singleton type-2 FLS [48]. The following subsections presents the detailed operation of the type-2 non-singleton type-2 FLS.

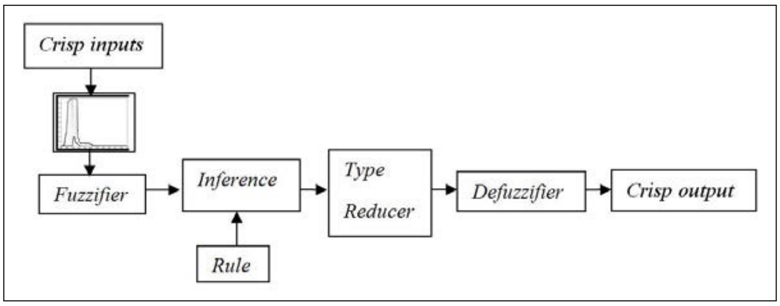


Figure 9: The structure of the type-2 non-singleton type-2 FLS [50].

6.1 Fuzzification

The antecedent type-2 fuzzy sets to represent the linguistic labels were considered as triangular, left shoulder, and right shoulder membership functions (as shown in Figure 10). The input type-2 fuzzy input variables were considered to be a non-specified piece-wise linear shape as explained in Section 5 (as shown in Figure 11). The fuzzifier converts the inputs into linguistic labels with given membership value to activate the rule base and inference engine. The operation between the linguistic labels antecedent fuzzy sets (defined by the interval type-2 fuzzy set) and the fuzzy input variable to provide the lower and upper membership values of input k within rule l is given by Equation (6) and Equation (7) respectively as follows: The antecedent type-2 fuzzy sets to represent the linguistic labels were considered as triangular, left shoulder, and right shoulder membership functions (as shown in Figure 10). The input type-2 fuzzy input variables were considered to be a non-specified piece-wise linear shape as explained in Section 5 (as shown in Figure 11).

The fuzzifier converts the inputs into linguistic labels with given membership value to activate the rule base and inference engine. The operation between the linguistic labels antecedent fuzzy sets (defined by the interval type-2 fuzzy set $\widetilde{F}_k^l$) and the fuzzy input variable to provide the lower and upper membership values of input k within rule l is given by Equation (6) and Equation (7) respectively as follows:

$$\overline{\mu}_k^l(x_k) = \sup_{x_k \in X_k} (\overline{\mu}_{x_k}(x_k) * \overline{\mu}_{F_k^l}(x_k)) \tag{6}$$

$$\underline{\mu}_k^l(x_k) = \sup_{x_k \in X_k} (\underline{\mu}_{x_k}(x_k) * \underline{\mu}_{F_k^l}(x_k)) \tag{7}$$

The $\overline{\mu}_{x_k}(x_k)$ and $\underline{\mu}_{x_k}(x_k)$ are the upper and lower membership functions of the input fuzzy variable, and $\overline{\mu}_{F_k^l}(x_k)$ and $\underline{\mu}_{F_k^l}(x_k)$ are the upper and lower membership functions of the antecedent fuzzy sets representing the linguistic labels respectively. In this work, for the sup operation, we have used the maximum t-conorm, and for the $*$ operation, we have used minimum t-norm. As we have considered three forms of antecedent type-2 fuzzy sets, therefore six algorithms are required to obtain the upper membership value and lower membership values of the output of the fuzzier for the different antecedents.

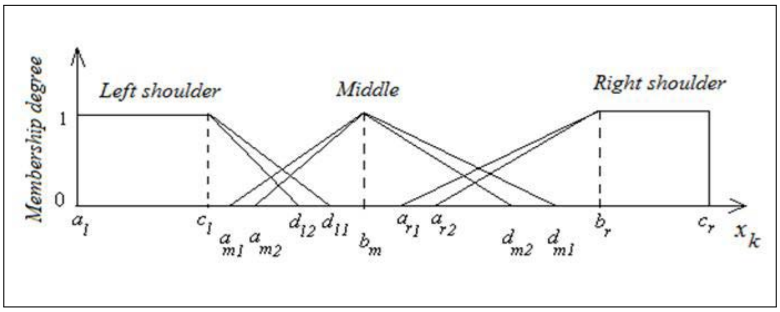


Figure 10: The considered antecedent type-2 fuzzy sets [48].

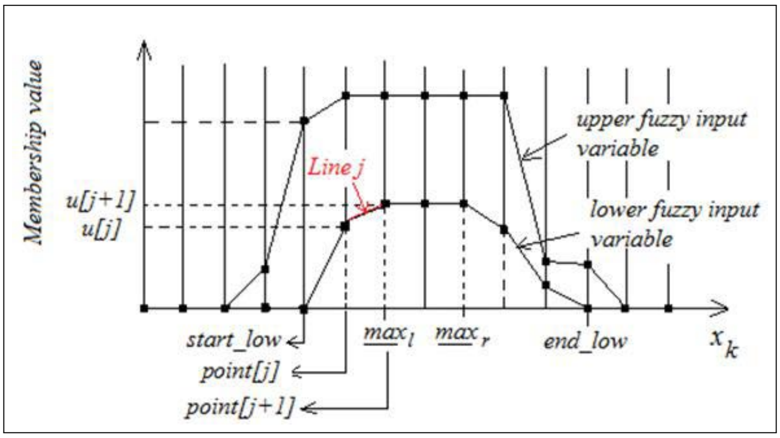


Figure 11: A considered shape of type-2 fuzzy input variable for the non-singleton fuzzification [12].

Due to the limited space, we will present the algorithms to obtain the upper and lower membership values of the type-2 non-singleton fuzzification when only right shoulder antecedent is considered. The rest of the algorithms follow the same logic.

The first example of fuzzification is on finding upper membership value of the fuzzifier, when antecedent membership function is a right shoulder (shown in Figure 12 with green and black colors). Following Equation (6), we need to perform the max-min operation between the upper membership functions of the type-2 fuzzy input variable and the right shoulder antecedent type-2 fuzzy set. The minimum operation between the type-1 membership functions and the type-1 fuzzy input variable will resulted in a type-1 fuzzy set which is shown in Figure 12 by red line, where we need to find its sup (maximum) to find $\overline{\mu}_k^l(x_k)$.

When the ending point of the upper membership function of the fuzzy input variable is larger than a_{r1}, it is required to consider the operations involving the right shoulder antecedent type-2 fuzzy set, as it is not zero. Two cases can happen, where case 1 is shown in Figure 12a and case 2 is shown in Figure 12b and Figure 12c. Case1 happens when $\overline{max}_r$ (shown in Figure 12) is smaller than b_r. In that case, the line of antecedent from a_{r1} to b_r would be considered.

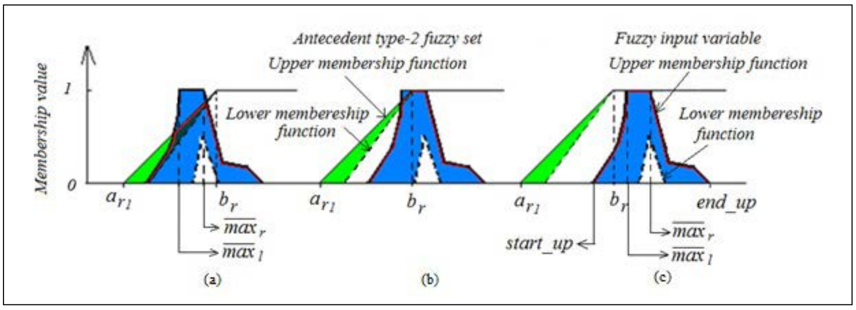


Figure 12: Different cases for obtaining the upper membership value of the fuzzifier output when the right shoulder antecedent membership function is considered in rule (a) case1. (b) and (c) case 2.

The algorithm in Figure 13 shows how to obtain the $\overline{\mu}_k^l(x_k)$ where we find the intersection of the sets of lines j (which form the piece-wise linear type-2 fuzzy input variable) and the antecedent type-2 fuzzy sets.

The given line j (of fuzzy input) ranges from *point[j]* to *point[j+1]* in the x-axis line where j+1 is the line connecting to the line j and located at its right in the piecewise linear function. The *u[j]* and *u[j+1]* are also membership value at *point[j]* and *point[j+1]* respectively. The membership value for any point x_k located on line j is obtained from (8) as follows:

$$\frac{\mu_{x_k}(x_k) - u[j]}{x_k - point[j]} = \frac{u[j+1] - u[j]}{point[j+1] - point[j]}$$

$$\mu_{x_k}(x_k) = \frac{u[j+1] - u[j]}{point[j+1] - point[j]} x_k + \frac{point[j+1]u[j] - point[j]u[j+1]}{point[j+1] - point[j]} \tag{8}$$

The membership value $\mu_{F_k^l}(x_k)$ of any point x_k on a line of the antecedent fuzzy set (starting from an extreme left point "left" (a_{r1} in our case) and ending at an extreme right point "right" (b_r in our case)) is obtained as follows:

$$\frac{\mu_{F_k^l}(x_k) - \mu_{F_k^l}(left)}{x_k - left} = \frac{\mu_{F_k^l}(right) - \mu_{F_k^l}(left)}{right - left}$$

$$\mu_{F_k^l}(x_k) = \frac{\mu_{F_k^l}(right) - \mu_{F_k^l}(left)}{right - left} x_k + \frac{right\mu_{F_k^l}(left) - left\mu_{F_k^l}(right)}{right - left} \tag{9}$$

At the intersection point $x_{k,c}$ between an antecedent line and a line j of the fuzzy input variable, $\mu_{x_k}(x_k)$ in Equation (8) is equal to $\mu_{F_k^l}(x_k)$ in Equation (9). Hence $x_{k,c}$ could be found as follows:

$$x_{k,c} = \frac{\left[\begin{array}{c} (right\mu_{F_k^l}(left) - left\mu_{F_k^l}(right))(point[j+1] - point[j]) + \\ (right - left)(point[j]u[j+1] - point[j+1]u[j]) \end{array}\right]}{\left[\begin{array}{c} (right - left)(u[j+1] - u[j]) + \\ (\mu_{F_k^l}(left) - \mu_{F_k^l}(right))(point[j+1] - point[j]) \end{array}\right]} \tag{10}$$

Subsistuting with the value of $x_k = x_{k,c}$ in Equation (8) or (9), we can find the intersection points when it is needed as $\overline{\mu}_k^l(x_k) = \mu_k^l(x_k)$.

Case 2 happens when $\overline{max}_r$ is larger or equal b_r, which is the case shown in Figure 12b and Figure 12c. Hence the output of fuzzifier for the upper membership value is one. The algorithm given in Figure 13 summarizes the method to obtain $\overline{\mu}_k^l(x_k)$ in the case that the antecedent type-2 fuzzy set is a right shoulder.

Step1: Initially consider $\overline{\mu}_k^l(x_k) = 0$ *(which means no input is fired when right shoulder antecedent and input k is considered.)*

Step2: Check if $end\_up > a_{r1}$*. If that was true, continue; otherwise exit the algorithm.*

*Step3: **(case1)** Check if* $\overline{\max}_r < b_r$*. If that was true continue; otherwise go to case2.*

Step4: Consider the line connecting a_{r1} *and* b_r *of the upper membership function of the antecedent type-2 fuzzy set and set the parameters of antecedent line as (*$left = a_{r1}$ *and* $right = b_r$*, and hence* $\mu_{F_k^l}(left) = 0$ *and* $\mu_{F_k^l}(right) = 1$ *).*

Step5: Consider the first line j (which ends at point[j+1]=end_up) of the upper membership function of the type-2 fuzzy input variable.

Step6: Check if there were no more lines j exists ($point[j] < \overline{\max}_r$ *). If that was true, exit the algorithm; otherwise continue.*

Step7: Obtain the point of intersection from Equation (10).

Step8: Check if the obtained intersection point is acceptable ($point[j] \le x_{k,c} \le point[j+1]$ *). If that was accepted, continue; otherwise consider the next line (decrease j by one) and come back to step6.*

Step9: Obtain $\overline{\mu}_k^l(x_k) = \mu_{xk}(x_{k,c})$ *for the point* $x_{k,c}$ *from equation (8). Then exit the algorithm.*

*Step10: **(case2)** Check if* $\overline{\max}_r \ge b_r$*, if that was true continue; otherwise exit the algorithm.*

Step11: Consider $\overline{\mu}_k^l(x_k) = 1$ *. Then exit from the algorithm.*

Figure 13: The algorithm to obtain the upper membership value of the fuzzifier for input k and rule l (when right shoulder antecedent is considered in the rule).

The second example of fuzzification is on finding lower membership value of the fuzzier, when antecedent membership function is a right shoulder (shown in Figure 14 with green and black colors). Following the Equation (7), we need to obtain the t-norm (minimum) between the type-1 membership function $\underline{\mu}_{x_k}(x_k)$ and $\underline{\mu}_{F_k^l}(x_k)$. This will result in a type-1 fuzzy set (shown by red line in Figure 14), where we need to find its sup (maximum) to find related $\underline{\mu}_k^l(x_k)$.

If the ending point of lower membership function of the type-2 fuzzy input is smaller or equal a_{r2}(the start of the lower membership function of right shoulder antecedent), the lower membership value of the fuzzifier for the right shoulder is zero. If it is not zero, then two cases can happen, where case 1 is shown in Figure 14a and case 2 is shown in Figure 14b. Case 1 happens when the lower membership value of fuzzy input at point $\underline{max}_r$(shown in Figure 14a) is larger than the membership value of the antecedent point $\underline{max}_r$, which can be considered as:

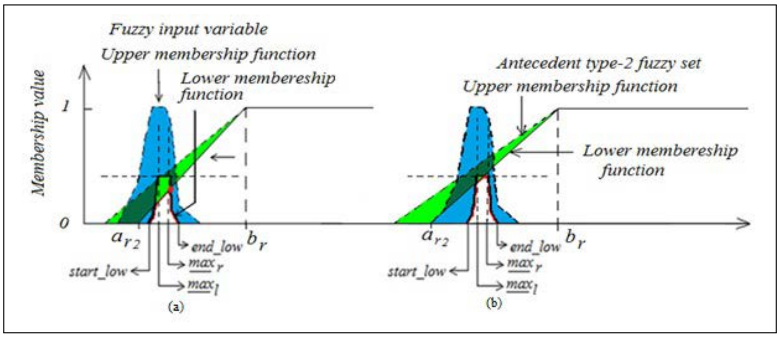


Figure 14: Different cases for obtaining the lower membership value of the fuzzifier output when the right shoulder antecedent membership function is considered in rule (a) case1 (b) case 2.

$$\underline{\max}_r < \frac{right(\mu_{x_k}(\underline{\max}_r) - \mu_{F_k^l}(left)) + left(\mu_{F_k^l}(right) - \mu_{x_k}(\underline{\max}_r))}{\mu_{F_k^l}(right) - \mu_{F_k^l}(left)} \tag{11}$$

In this case, the max-min will be located on the antecedent line from a_{r2} to b_r. Because of the positive slope of the antecedent line, the most right intersection between the lower membership function of the type-2 fuzzy input and the antecedent line has the highest membership value, and its membership value is $\underline{\mu}_k^l(x_k)$, when the right shoulder is considered in rule.

Case 2 happens when the lower membership value of fuzzy input at point $\underline{max}_r$ is not larger that the membership value of the antecedent at point $\underline{max}_r$ (when case1 does not happen). In this case, $\underline{\mu}_k^l(x_k)$ is equal to the maximum membership value of the lower membership function of the type-2 fuzzy input. Figure 15 illustrates the algorithm to obtain the lower membership value of the fuzzifier for input k and rule l when right shoulder antecedent is considered in rule.

Step1: *Initially consider $\underline{\mu}_k^l(x_k) = 0$ (which means no input is fired when right shoulder antecedent and input k is considered.)*

Step2: *Check if $end\_low > a_{r2}$. If that was true, continue; otherwise exit the algorithm.*

Step3: *Consider the line connecting a_{r2} and b_r of the lower membership function of the antecedent type-2 fuzzy set and set the parameters of antecedent line as (left $= a_{r2}$ and right$= b_r$, and hence $\mu_{F_k^l}(left) = 0$ and $\mu_{F_k^l}(right) = 1$).*

Step4: (case1) *Check if the statement (11) is true. If it was true continue; otherwise go to case2.*

Step5: *Consider the first line j (which ends at point[j+1]=end_low) of the upper membership function of the type-2 fuzzy input variable.*

Step6: *Check the considered line j is within the possible area of happening intersection between the lower membership functions of antecedent and fuzzy set with highest membership value ($point[j] \geq \underline{\max}_r$). If that was true, continue; otherwise exit the algorithm.*

Step7: *Obtain the point of intersection from Equation (10).*

Step8: *Check if the obtained intersection point is acceptable ($point[j] \leq x_{k,c} \leq point[j+1]$). If that was accepted, continue; otherwise consider the next line (decrease j by one) and come back to step6.*

Step9: *Obtain $\underline{\mu}_k^l(x_k) = \mu_{xk}(x_{k,c})$ for the point $x_{k,c}$ from equation (8). Then exit the algorithm.*

Step10: (case2) *Check if statement (13) is not true. In this case, $\underline{\mu}_k^l(x_k) = \mu_{xk}(\underline{\max}_l) = \mu_{xk}(\underline{\max}_r)$. (which is actually the maximum membership value of the lower membership function of the fuzzy input variable. Then exit the algorithm.*

Figure 15: The algorithm to obtain the lower membership value of the fuzzifier for input k and rule l (when right shoulder antecedent is considered in the rule).

6.2 Rule Base and Inference Engine

The outputs of fuzzifier trigger the inference engine and the rule base to generate type-2 fuzzy outputs which are then type-reduced and defuzzified to generate the final crisp output of the FLS. The same approach as singleton FLS is applicable to the type-2 non-singleton type-2 FLS after obtaining the output of fuzzifier. The rule base, inference engine, type-reduction and defuzzification follow as:

• Rule Base: The rule base of singleton and non-singleton FLS are similar. Considering a FLS having inputs $x_1 \in X_1, ..., x_p \in X_p$ and one output y ∈ Y, the rule l is defined as [32]:

$$If\ x_1\ is\ \widetilde{F}_1^l\ ...\ and\ x_p\ is \widetilde{F}_p^l,\ then\ y\ is\ \widetilde{G}^l \tag{12}$$

Where l=1,..., M and M is the total number of rules in the rule base. The $\widetilde{F}_k^l$ is the antecedent membership function for input k and rule l, and $\widetilde{G}^l$ is consequent membership function for rule l.

• Inference Engine: We have shown in the fuzzifier subsection, how to find the upper and lower membership values of a given type-2 fuzzy input variable to a given linguistic label. The process needs to be done for all inputs (k=1,..., p). At the end, the firing strength of each rule $F^l = [\underline{f}^l, \overline{f}^l]$ could be found where $\underline{f}^l$ and $\overline{f}^l$ will be obtained as follows:

$$\underline{f}^l = T_{k=1}^p \{\underline{\mu}_k^l(x_k)\} \tag{13}$$

$$\overline{f}^l = T_{k=1}^p \{\overline{\mu}_k^l(x_k)\} \tag{14}$$

6.3 Type-Reduction

After obtaining the lower and upper firing strengths for each rule from the inference engine and obtaining the centroid of consequent membership function, the center of sets type-reducer [32] is used to produce the type-reduced sets.

6.4 Defuzzification

The final defuzzified crisp output is obtained by taking the average of the type-reduced set.

7 Adaptive Type-2 Input Based Non-Singleton Type-2 FLS

We considered the non-singleton type-2 fuzzy inputs to be dynamic and automatically generate from data and the shapes are changing for different measurement values. As we cannot generate the type-2 fuzzy sets for all the measurement values, we need to interpolate the type-2 fuzzy set for an incoming measurement value by interpolating the type-2 fuzzy sets representing the measurements to the left and right of incoming measurement value. Figure 16 shows the process of the on-line singleton type-2 FLS. [12]. We obtained some sample of original type-2 fuzzy input variables for different measurement values within considered range according to the instructions given in Section 5, and then obtained the rest of the type-2 fuzzy input variables by a linear interpolation defined for type-2 fuzzy sets. The interpolated membership value for the desired crisp incoming input (d) can be obtained as follows:

$$u_d[i] = \frac{(d - d_2)}{(d_1 - d_2)} u_1[i] + \frac{(d - d_1)}{(d_2 - d_1)} u_2[i] \tag{15}$$

The $u_1[i]$ is the membership value of the left original fuzzy input variable, and $u_2[i]$ is the right original fuzzy input variable. The d_1 is the given measurement input for the left fuzzy input variable and d_2 is the given measurement input for the right neighboring fuzzy input variables, and d is the crisp value (new measurement data) in question which is to be modeled by an interpolated type-2 fuzzy input variable.

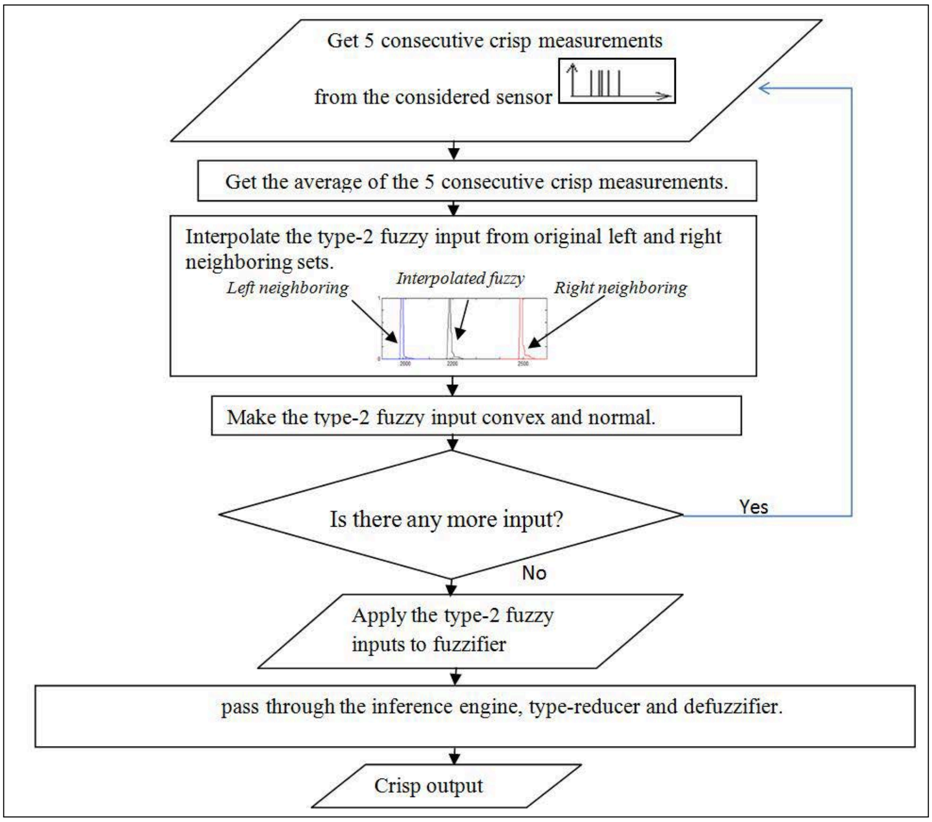


Figure 16: The process of on-line non-singleton type-2 FLS [12].

The interpolation is considered to be done separately for the lower membership functions and the upper membership functions for the fuzzy input. The resulting interval type-2 fuzzy sets need to be convex and normal and if this is not the case, we need to convert the type-2 fuzzy set to a normal convex fuzzy set [48].

8 Experiments and Results

In this section, we will present real world experiments in a real world environment and we will describe the characteristics of the considered FLSs and test environment. We will then present control surface analysis and the real world experiments results.

8.1 Experimental Setup

In this Section, we will present real world experiments which were carried using a real world robot navigating in a real world environment. We will compare the performance of the adaptive type-2 non-singleton type-2 FLS against its singleton type-1 and type-2 FLSs counterparts. In our experiments, we have employed a four-wheel mobile robot (shown in Figure 17) to do a left edge following behavior using two sonar sensors. A hair dryer was located on robot to blow hot air through the path of front sonar sensor to create larger measurement uncertainty while robot is moving.

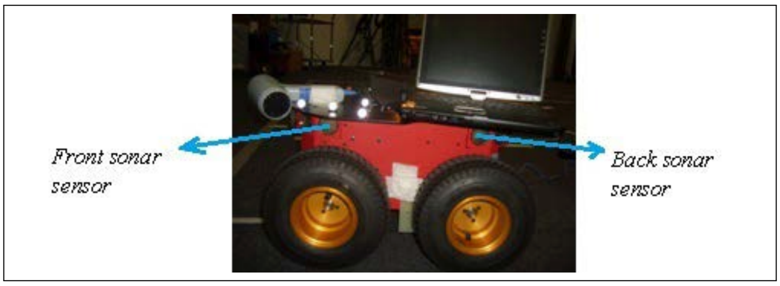


Figure 17: The four wheel mobile robot used in the experiments.

In order to test the robot under unseen input uncertainties [12], we have applied various sources of uncertainties to the input sonar sensors (as shown in Figure 18a, Figure 18b) where we have used a hair dryer stuck on robot to blow hot air while a fan fixed at a certain position on ground to create cold wind. These employed uncertainties are introduced during the online operation of the robot and they have not been seen by the robot during the FLS design. We chose changing the wind speed (through the fan and dryer) and temperature (through the hair dryer) to introduce input uncertainties to the sonar sensors. This is because the wind, temperature variations and atmospheric turbulences are all meteorological conditions that have an effect on sound propagation (as we used sonar sensors) [52]. The speed of sound in air is proportional to the square root of absolute temperature as follows:

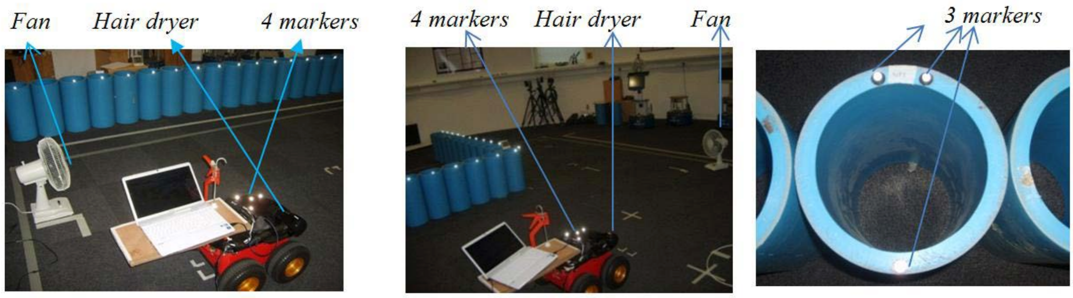
$$c = 20.05\sqrt{T + e/p} \tag{16}$$

Where T is absolute temperature in Kelvin, and e is partial pressure of water vapor, and p is barometric pressure. Wind is also another factor affecting the sound travel. The wind in the same direction of sound will cause the wave propagation to be bent downward, and in the upwind direction sound waves move upward which cause no direct sound penetration [52].

To obtain the position of the robot in a precise way, the VICON system is used in which six cameras are employed. To be able to recognize the robot, four markers are used (as shown in Figure18). The top of the cylinders are also marked to be able to check if the edges of the cylinders are in a same line where the starting and ending cylinders are modeled in VICON to be able to distinguish the certain position of wall (one of them is shown in Figure18c) [12].

In our experiments, we considered a type-1 FLS which was able to give a good performance under the normal lab conditions. The performance of type-1 FLS was considered to be acceptable in a usual room temperature without applying external uncertainty. Two main factors were considered to evaluate the performance of the FLSs, the first one is the Root Mean Square Error (RMSE) which is considered as a measurement factor for the average of error for the distance between the robot's path from the desired path. The second factor is standard deviation (STD) of error in following edge to check the variance of the error around the average error value.

The shapes of the type-2 fuzzy sets used to represent the linguistic labels are shown in Figure 19 where we blurred the type-1 membership functions (shown in thin lines in Figure 19) to obtain the FOU of the type-2 fuzzy sets (shown in Figure 19 by thick line) for the interval (singleton and non-singleton) type-2 FLS. The consequent membership functions (for all FLSs) are type-1 fuzzy sets (shown in Figure 20) as there are no major uncertainties associated with robot outputs. The rule-base (employed for all the FLSs) is shown in Table 1.



(a) (b) (c)

Figure 18: Mobile robot experiments when (a) Left edge following set in a straight line. (b) Left edge following set in straight line with a right angle in the way. (c) One of the cylinders modeled in Vicon system using three markers to know the exact position of wall.

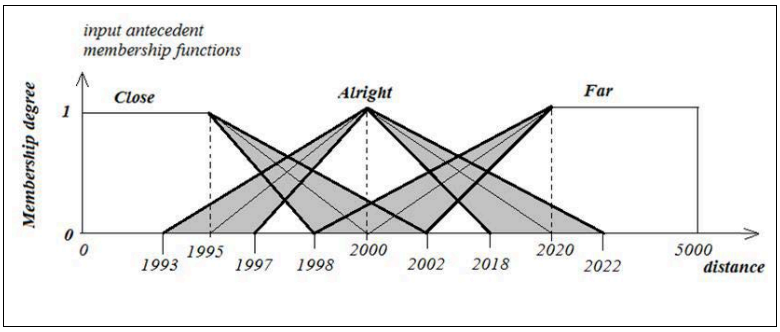


Figure 19: Antecedent membership functions for both the front and back senors. The type-1 fuzzy sets are shown in thin line while the interval type-2 fuzzy sets are shown in thick lines [12].

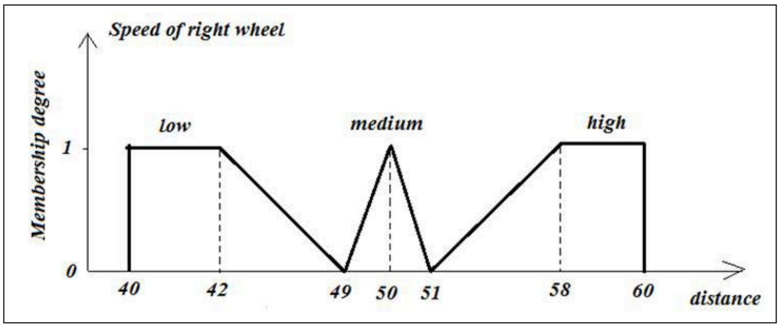


Figure 20: The consequent type-1 fuzzy sets for all the used FLSs [12].

Front sonar	Back sonar		
	Close	Alright	Far
Close	low	low	low
Alright	low	medium	low
Far	high	high	high

Table 1: The rule base for all the FLSs; the output in the rule base is the speed of right wheel [12].

8.2 Control Surface Analysis

The control surface graphically represents the unknown function articulated by the FLS. Figure 21 shows the control surfaces for the singleton type-1 and type-2 FLSs and the type-2 non-singleton type-2 FLS. The control surfaces show the inputs of front and back sensors plotted against the speed of right wheel as output of the FLSs. Note that a smooth shape of a control surface translates to a smooth control response that can deal with uncertainty and imprecision.

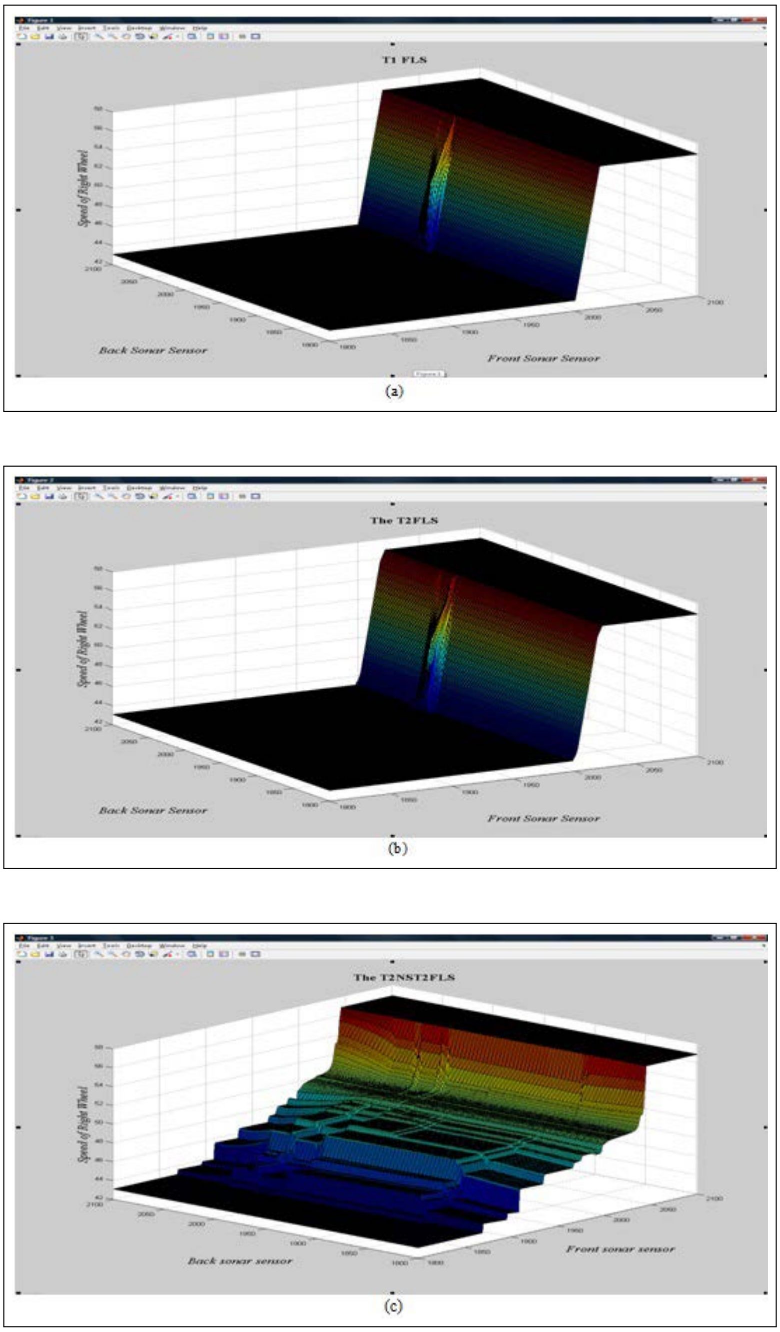


Figure 21: The control surface of (a) The singleton type-1 FLS. (b) The singleton interval type-2 FLS. (c) The type-2 non-singleton intreval type-2 FLS.

As seen from Figure 21a, the control surface of the type-1 FLS is rather steep which means it can be affected by high levels of uncertainties in the inputs where small change in the inputs can result in large output changes which might cause high overshoots/undershoots and steady state errors in addition to moving towards an unstable situation. On the other hand the control surface of the singleton type-2 FLS (shown in Figure 21b) is smoother than the type-1 FLS, which means it can handle higher level of uncertainties. However, the control surface of the type-2 non-singleton type-2 FLS (shown in Figure 21c) is much smoother where there are smooth transitions at all parts of the control surface where any small change in inputs will result in a small change in the outputs. Thus using the type-2 input based non-singleton type-2 FLS result in a control surface which is smoother than the singleton type-1 and type-2 FLSs. This smooth response will consequently give a very good control performance that can handle the high levels of uncertainties and disturbances where small variations in inputs will not cause significant changes to outputs. Thus, it is expected that the type-2 non-singleton type-2 FLS will be able to handle higher levels of uncertainties than its singleton type-1 and type-2 FLSs counterparts. This will be evidenced in the real world experiments reported in the following subsection.

8.3 Real World Experiments

Figure 22a and Figure 22b show the results of the robotic experiments illustrated in Figure 18a, which compare the performance of singleton type-1, singleton type-2 and type-2 input based non-singleton type-2 FLS to realize an edge following behavior at a desired distance of 2m. The robot starts from the distance of 1.8m from the nearest edge. Figure 22a shows the robot's paths of singleton type-1, type-2 and the non-singleton type-2 FLS under no induced uncertainty. As shown in Figure 22a and Table 2 as expected under no uncertainty, the performance of all the FLSs will be quite similar where the type-1 FLS will be slightly better. However, as high levels of uncertainties are introduced (by blowing hot wind through the way of one sonar) as shown in Figure 22b and Table 2, the type-2 non-singleton type-2 FLS gives the best performance in terms of the RMSE and the STD (which could represent the size of the overshoots/undershoots from the average). In terms of control performance, the type-2 non-singleton type-2 FLS provides the best performance followed (by a large margin) by the singleton type-2 FLS and then the singleton type-1 FLS which performs poorly.

Figure 22b evidences the control surface analysis in Figure 21 where the type-2 nonsingleton type-2 FLS recovers quickly from the uncertainties with relatively smaller overshoot/undershoot.

In the other experiments we are presenting in the paper, we consider the same type-1 fuzzy antecedent while we increase the FOU as illustrated in Figure 23 to have larger difference between type-1 FLS and type-2 FLS. The rule base and consequent membership functions were considered the same as Table 1 and Figure 20 respectively. The other differences are introducing a right angle in the middle of the edges set in straight line and changing the position of fan (shown in Figure 18b).

Different Conditions	Different Types of FLS		
	Type-1 FLS	Type-2 FLS	Type-2 nonsingleton Type-2
No uncertainty	STD = 45.5617 RMSE =45.5695	STD =46.5435 RMSE=46.5823	STD= 47.9173 RMSE= 51.2167
large uncertainty	(Stopped) STD = 624.8140 RMSE=1268.2	STD =477.2452 RMSE=1037.8	STD =244.2411 RMSE=634.4

Table 2: Results of following a straight wall for the singleton type-1 FLS, singleton type-2 FLS, and adaptive type-2 input based non-singleton type-2 FLS under different conditions (related to Figure 22).

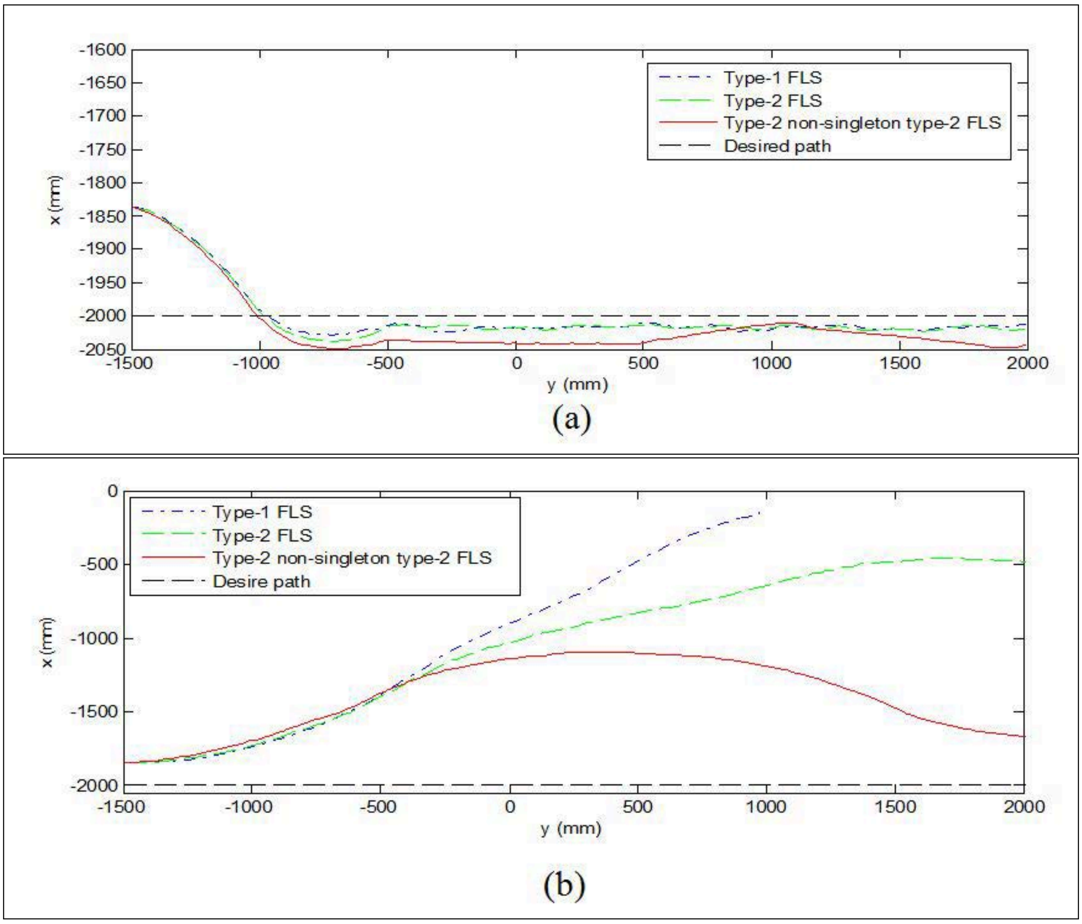


Figure 22: The robot's path when employing singleton type-1 FLS, singleton type-2 FLS, and the adaptive type-2 input based type-2 FLS, while following a straight wall under different conditions (a) No external uncertainty. (b) Large external uncertainty.

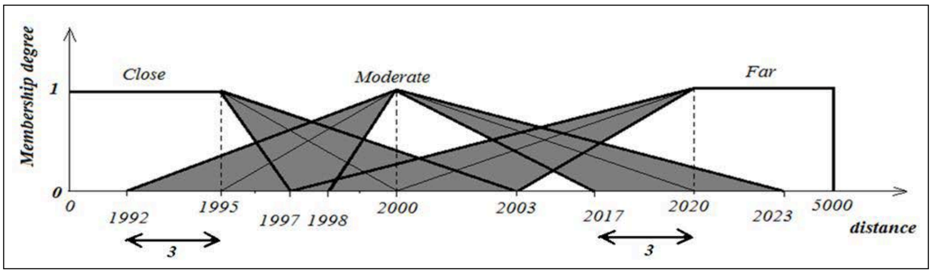


Figure 23: The type-2 fuzzy sets for the antecedent membership functions considered in the experiments whose results are shown in Figure 24 and Figure 25, which is actually considering uncertainty bound of 3 around type-1 fuzzy sets.

Figure 24 shows the robot's paths when employing the type-2 non-singleton type-2 FLS with its singleton type-1 FLS and singleton type-2 FLS counterparts to do left edge following at a desired distance of 2m. The robot was started from the distance of 1.8m from the nearest edge.

Figure 24a shows that the robot's paths under no induced uncertainty for the three FLS are quite similar. Looking at Table 3, it can be seen that the performance of the type-2 input based non-singleton type-2 FLS is slightly better than the other FLSs because it has the least

STD and RMSE compared with its singleton type-1 FLS and singleton type-2 FLS counterparts. When considering moderate uncertainty (the related robot's paths are shown in Figure 24b), the difference between the performance of the adaptive type-2 non-singleton type-2 FLS and the singleton type-1 FLS and singleton type-2 FLS are becoming more distinguishable. The robot's path for the non-singleton type-2 FLS is closer to the desired path specified by black dashed lines, so it has the best performance while the type-2 FLS has also outperformed the type-1 FLS. Table 3 also shows the least STD and RMSE for the non-singleton type-2 FLS, then the singleton type-2 FLS which outperform the type-1 FLS by having less STD and RMSE. When considering large uncertainty (the related robot's paths are shown in Figure 24c), the non-singleton type-2 FLS outperforms its singleton type-2 and singleton type-1 FLS counterparts, while the type-1 FLS is performing very poorly. Table 3 also shows the least STD and RMSE for the non-singleton type-2 FLS.

Figure 25 shows the robot's paths when employing the type-2 non-singleton type-2 FLS and its singleton type-1 FLS and singleton type-2 FLS counterparts to do left edge following at a desired distance of 2m. The robot was started from the distance of 1.9m from the nearest edge. Figure 25a shows the robot's path when employing the FLSs under no induced uncertainty. The performances of the FLSs are quite similar as it can be seen from the Figure 25a while the adaptive type-2 non-singleton type-2 FLS was slightly better as it has the least STD and RMSE (illustrated in Table 4). Figure 25b and Figure 25c also shows that the type-2 input based non-singleton type-2 FLS outperforms its singleton type-2 and type-1 FLS counterparts when moderate and large uncertainty are applied respectively. Table 4 also confirms that the type-2 input based non-singleton type-2 FLS outperforms its counterparts FLSs in the experiments as the type-2 input based non-singleton type-2 FLS has the least STD and RMSE, while the type-1 FLS performs poorly.

Different Conditions	Different Types of FLS		
	Type-1 FLS	Type-2 FLS	Type-2 nonsingleton Type-2
No uncertainty	STD =202.5274 RMSE =202.7028	STD =203.8221 RMSE=204.0164	STD =191.3634 RMSE=191.5175
Moderate uncertainty	STD =907.7865 RMSE =910.7622	STD =700.8891 RMSE=703.5157	STD =533.0486 RMSE=546.5797
Large uncertainty	(Stopped)	STD =694.8356 RMSE=699.5279	STD =669.3059 RMSE=676.0831

Table 3: Summary results of the edge following behviour for the singleton type-1 FLS, singleton type-2 FLS, and adaptive type-2 input based non-singleton type-2 FLS under the different three conditions related to Figure 24.

Different Conditions	Different Types of FLS		
	Type-1 FLS	Type-2 FLS	Type-2 nonsingleton Type-2
No uncertainty	STD = 201.7354 RMSE = 201.7395	STD = 201.7226 RMSE=201.7245	STD = 193.5605 RMSE=193.5647
Moderate uncertainty	STD =562.0358 RMSE=588.3824	STD=548.4835 RMSE=584.5484	STD= 345.9873 RMSE=352.3825
Large uncertainty	(Stopped)	(Stopped)	STD =747.6389 RMSE=748.5229

Table 4: Summary results of the edge following behavior for the singleton type-1 FLS, singleton type-2 FLS, and adaptive type-2 input based non-singleton type-2 FLS under the different three conditions related to Figure 25.

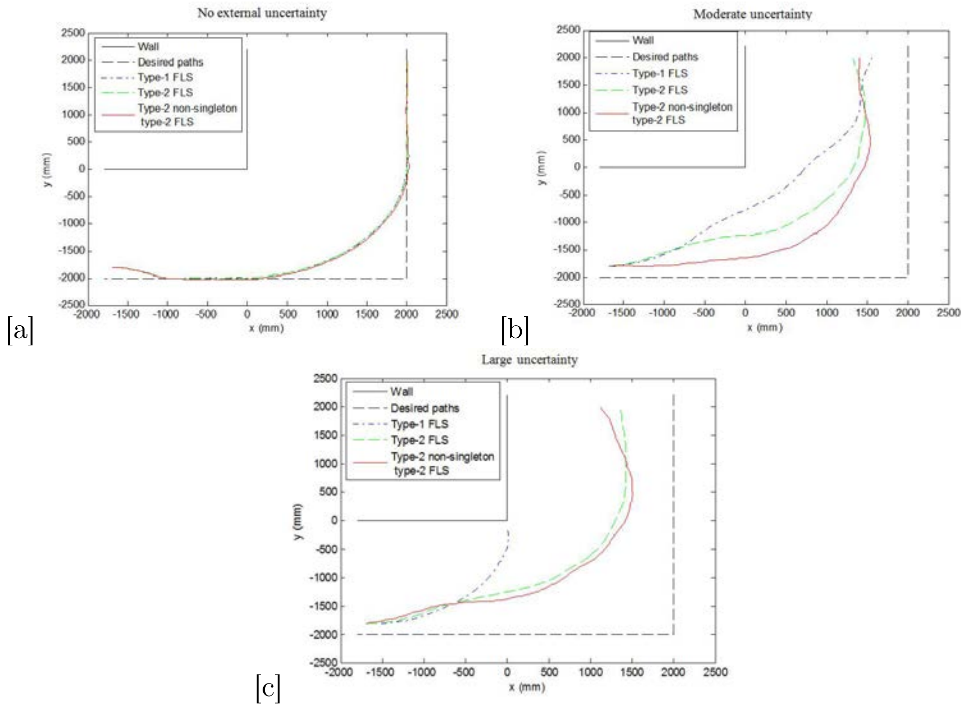
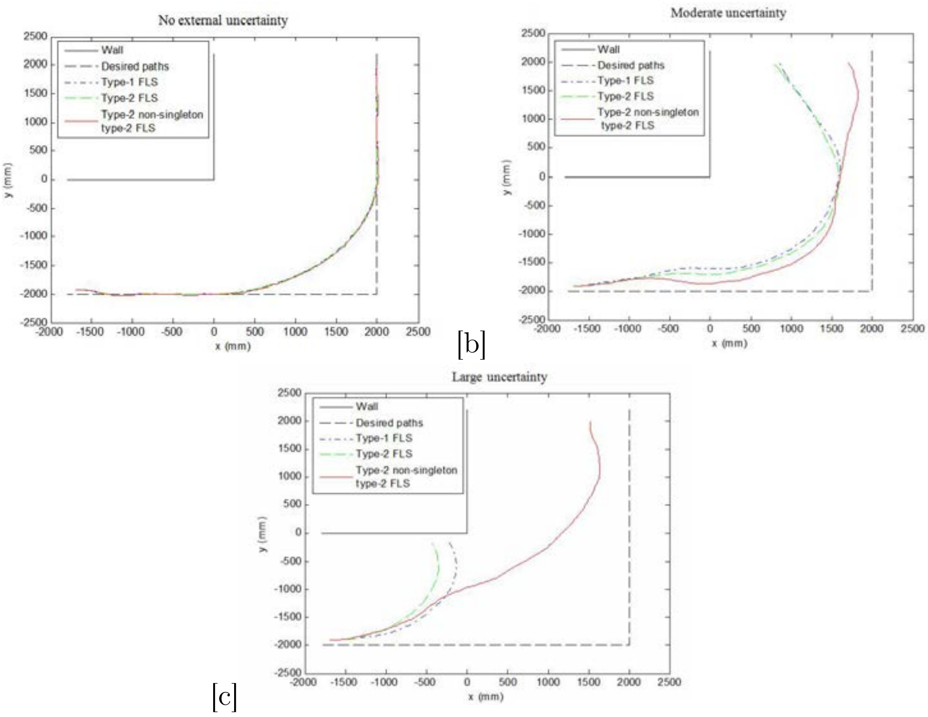


Figure 24: Comparison between singleton type-1 FLS, singleton type-2 FLS, and the adaptive type-2 input based non-singleton type-2 FLS when the robot started from 1.8 meter distance from the edge of wall while applying (a) No external uncertainty (b) Moderate uncertainty (c) Large uncertainty.

9 Conclusions and Future Directions

In this paper, we have presented an adaptive type-2 fuzzy input based non-singleton type-2 FLS which is better able to handle the encountered input uncertainties. The non-singleton type-2 fuzzy inputs are dynamic and they are automatically generated from data and they do not assume a specific shape about the uncertainty distribution associated with the given sensor. We have presented an overview on how the adaptive type-2 input based non-singleton interval type-2 FLS can operate in real time. We have shown through control surface analysis, how the non-singleton type-2 FLS can produce much smoother control surface than the singleton type-1 and type-2 FLSs. This should result in a better control response that handles high uncertainty levels. This was evidenced through the real world experiments using mobile robots where under high uncertainty levels, the non-singleton type-2 FLS can outperform its singleton type-2 and type-1 FLSs counterparts. This work has shown the benefits achieved when using type-2 non-singleton type-2 FLSs in applications characterized by high input uncertainty levels.

In our future work, we are going to tune the parameters of the type-2 FLS by neural networks. We aim also to work towards modeling more generalized type-2 fuzzy input variable or even doing non-singleton fuzzifier in general type-2 FLSs. Last but not least, we are also considering developing a self-tuning adaptive type-2 input based non-singleton interval type-2 FLS, which is updating the shape of the fuzzy input variable whenever the performance of the FLS was deteriorated.



[a] [b] [c]

Figure 25: Comparison between singleton type-1 FLS, singleton type-2 FLS, and the adaptive type-2 input based non-singleton type-2 FLS when the robot started from 1.9 meter distance from the edge of wall while applying (a) No external Uncertainty. (b) Moderate uncertainty. (c) Large uncertainty.

Bibliography

[1] L. A. Zadeh, Fuzzy Sets, Information and Control, 8, 338-353, 1965.

[2] L. A. Zadeh, The Concept of a Linguistic Variable and its Application to Approximate Reasoning-I, Information Science 8, 199-249, 1975.

[3] L. A. Zadeh, Outline of a New Approach to the Analysis of Complex Systems and Decision Processes, IEEE Transactions on Systems, Man, and Cybernetics, pp.28-44, 1973.

[4] E. H. Mamdani, S. Assilian, An Experiment in Linguistic Synthesis with a Fuzzy Logic Controller, International Journal of Man-Machine Studies (1975), Vol. 7, No. 1, pp.1-13, 1975.

[5] G. Feng, A Survey on Analysis and Design of Model-Based Fuzzy Control Systems, IEEE Transactions on Fuzzy Systems, Vol. 14, No. 5, pp.676-697, 2006.

[6] P. P. Bonissone, V. Badami, K. H. Chiang, P. S. Khedkar, K. W. Marcelle, M. J. Schutten, Industrial Applications of Fuzzy Logic at General Electric, Proceedings of IEEE, Vol. 83, No. 3, pp.450-465, 1995.

[7] B. Choi, S. Han, S. Hong, Refrigerator Temperature Control Using Fuzzy Logic and Neural Network, IEEE International Symposium on Industrial Electronics, Vol. 1, pp.186-191, 1998.

[8] http://www.johnlewis.com/

[9] http://www.gandhiappliances.com/index.php?main_page=product_info&cPath=10_224&products_id=5847

[10] R. John, S. Coupland, Type-2 Fuzzy Logic: A Historical View, IEEE Computational Intelligence Magazine, Vol. 2, No. 1, pp.57-62, 2007.

[11] D. Wu, W. W. Tan, Genetic Learning and Performance Evaluation of Interval Type-2 Fuzzy Logic Controllers, Engineering Applications of Artificial Intelligence, Vol. 19 No. 8, pp. 829-841, 2006.

[12] N. Sahab, H. Hagras, An Adaptive Type-2 Input Based Non-Singleton Type-2 Fuzzy Logic System for Real World Applications, IEEE International conference on Fuzzy Systems 2011.

[13] A. Perez-Neira, J. Sueiro, and J. Roca and M. Lagunas, A Dynamic Non-singleton Fuzzy Logic System for DS/CDMA Communications, IEEE International Conference on Fuzzy Systems, Vol. 2, pp. 1494-1499, 1998.

[14] M. Benedetti, A. Casagranda, and M. Donelli, and A. Massa, An Adaptive Multiscaling Imaging Technique Based on a Fuzzy-logic Strategy for Dealing with the Uncertainty of Noisy Scattering Data, IEEE Transactions on Antennas and Propagation, Vol. 55, No. 11, pp. 3265-3278, 2007.

[15] C. K. Makropoulos and D. Butler, Spatial decisions under uncertainty: fuzzy inference in urban water management, Journal of Hydroinformatics, IWA publishing, pp. 3-18, 2004.

[16] C. Quek and R. Zhou, Structure and learning algorithms of a nonsingleton input fuzzy neural network based on the approximate analogical reasoning schema, International Journal on Information Science and Engineering Science (Fuzzy Sets and Systems), Elsevier, vol.157, no. 13, pp. 1814-1831, 2006.

[17] E. Cavallaro, S. Micera, and P. Dario, and W. Jensen, and T. Sinkjaer,On the Intersubject Generalization Ability in Extracting Kinematic Information from Afferent Nervous Signals, IEEE Transactions on Biomedical Engineering, Vol.50, No. 9, pp. 1063-1073, 2003.

[18] T. Chua and W. Tan, Non-singleton genetic fuzzy logic system for arrhythmias classification, Engineering Applications of Artificial Intelligence, Elsevier, Vol. 24, No. 2, Pages 251-259, 2011.

[19] G. M. Mendez, L. L. Lezma, R. Colas, G. M. Perez, J. R. Cuellar, J. J. Lopez, Application of Interval Type-2 Fuzzy Logic Systems for Control of Coiling Entry Temperature in Hot Strip Mill, International conference on Hybrid Artificial intelligent Systems, pp.352-359, Springer-Verlag Berlin 2009.

[20] G. M. Mendez, M. A. Hernandez, Interval Type-1 Non-Singleton Type-2 TSK Fuzzy Logic Systems Using the Hybrid Training Method RLS-BP, IEEE Symposium on Foundations of computational Intelligence, pp.370-374, 2007.

[21] G. M. Mendez, A. Hernandez, Hybrid IT2 NSFLS-1 Used to Predict the Uncertain MXNUSD Exchange Rate, Springer-Verlag Berlin Heidelberg 2008, pp. 575-582, 2008.

[22] G. M .Mendez, A. Hernandez, A. Cavazos, M. T. M. Jimenez, Type-1 Non-singleton Type-2 Takagi-Sugeno-Kang Fuzzy Logic Systems Using the Hybrid Mechanism Composed by a Kalman Type Filter and Back Propagation Methods, Springer-Verlag Berlin Heidelberg 2010, pp.429-437, 2010.

[23] G. M. Mendez, L. A. Leduc, First-Order Interval Type-1 Non-singleton Type-2 TSK Fuzzy Logic Systems, Springer-Verlag Berlin Heidelberg ,pp.81-89, 2006.

[24] G. M. Mendez, Interval Type-1 Non-singleton Type-2 Fuzzy Logic Systems are Type-2 Adaptive Neuro-Fuzzy Inference Systems, International Journal of Reasoning-based Intelligent Systems, Vol. 2, No. 2, pp.95-99, 2010.

[25] G. M. Mendez, I. L. Juarez, L. A. Leduc, R. Soto, A. Cavazos, Temperature Prediction in Hot Strip Mill Bars Using a Hybrid Type-2 Fuzzy algorithm, International Journal of Simulation Systems, Science & Technology, pp.33-43, 2005.

[26] G. M. Mendez, M. A. Hernandez, Hybrid Learning for Interval Type-2 Fuzzy Logic Systems Based on Orthogonal Least-squares and Back-propagation Methods, International Journal on Information Science, Vol. 179, No. 13, pp.2146-2157, 2009.

[27] G. Mendez, A. Hernandez, Hybrid Learning of Interval Type-2 Fuzzy Systems Based on Orthogonal Least Squared and Back Propagation for Manufacturing Applications, Journal of Automation, Mobile Robotics & Intelligent Systems, pp.23-32 , 2008.

[28] H. Wu, J. M. Mendel, "Classification of battlefield ground vehicles using acoustic features and fuzzy logic rule-based classifiers", IEEE transactions on fuzzy systems, Vol. 15, No. 1, pp.56-72, 2007.

[29] G. Mendez, M. Hernandez, IT2 TSK NSFLS2 ANFIS, Ninth Mexican International Conference on Artificial Intelligence, IEEE Computer Society, pp.89-93, 2010.

[30] G. Mendez, M. Hernandez, Interval Type-2 Non-singleton Type-2 Takagi-Sugeno-Kang Fuzzy Logic Systems Using the Hybrid Learning Mechanism Recursive-Least-Square and Back-Propagation Methods, IEEE International Conference on Control, Automation, Robotics and Vision, Singapore, pp.710-714, 2010.

[31] P. Torres, D. Saez, Type-2 Fuzzy Logic Identification Applied to the Modeling of a Robot Hand, IEEE International conference on Fuzzy Systems, pp.854-861, 2008.

[32] J. M. Mendel, Uncertain Rule-based Fuzzy Logic Systems, introduction and new direction, Prentice Hall PTR Upper Saddle river, NJ, 2001.

[33] Q. Liang, J. M. Mendel, Interval Type-2 Fuzzy Logic Systems: Theory and Design, IEEE Transactions on Fuzzy Systems, Vol.8, No.5, pp.535-550, 2000.

[34] G. C. Mouzouris, J. M. Mendel, Nonsingleton Fuzzy Logic Systems: Theory and Application, IEEE Transactions on Fuzzy Systems, Vol.5 , No.1, pp.56-71, 1997.

[35] J. L. S. Neto, L. H. Hoang, G. C. Mouzouris, J. M. Mendel, Comments on "Nonsingleton fuzzy logic systems: theory and applications" and Reply, IEEE Transactions on Fuzzy Systems, Vol.6, No.2, pp.325-326 , 1998.

[36] S. Coupland, R. John, Geometric Interval Type-2 Fuzzy Systems, Proceedings of the joint 4th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT 2005) and the 11th LFA (joint EUSFLAT-LFA 2005), pp.449-454, 2005.

[37] H. Hagras, Type-2 FLCs: A New Generation of Fuzzy Controllers, IEEE Computational Intelligence Magazine, Vol. 2, No. 1, pp.30-43, 2007.

[38] J. M. Mendel, Type-2 Fuzzy Sets: Some Questions and Answers, Feature article in IEEE Neural Network Society, pp.10-13, 2003.

[39] J. M. Mendel, Advances in Type-2 Fuzzy Sets and Systems, Information Sciences 177 (2007) 84-110, 2006.

[40] N. N. Karnik, J. M. Mendel, Q. Liang, Type-2 Fuzzy Logic Systems, IEEE Transactions on Fuzzy Systems, Vol. 7, No. 6, pp. 643-658, 1999.

[41] Q. Liang, N. N. Karnik, J. M. Mendel, Connection Admission Control in ATM Networks Using Survey-Based Type-2 Fuzzy Logic Systems, IEEE Transactions on Systems, Man, and Cybernetics- Part-C: Applications and Reviews, Vol. 30, No. 3, pp.329-339, 2000.

[42] J. M. Mendel, R. B. John, Type-2 Fuzzy Sets Made Simple, IEEE Transactions on Fuzzy Systems, Vol. 10, No. 2, pp.117-127, 2002.

[43] J. Mendel, H. Wu, Uncertainty Versus Choice in Rule-based Fuzzy Logic Systems , IEEE International Conference on Fuzzy Systems, pp.1336-1342, 2002.

[44] J. M. Mendel, Type-2 Fuzzy Sets and Systems: How to Learn About Them, IEEE SMC eNewsletter, 2009.

[45] E. A. Jammeh, M. Fleury, C. Wagner, H. Hagras, M. Ghanbari, Type-2 Fuzzy Logic Congestion Control for Video Streaming Across IP Networks, IEEE Transactions on Fuzzy Systems, Vol. 17, No. 5, pp.1123-1142, 2009.

[46] S. Barkati, E. M. Berkouk , M. S. Boucherit, Application of Type-2 Fuzzy Logic Controller to An Induction Motor Drive with Seven-level Diode-clamped Inverter and Controlled Infeed, Electrical Engineering Journal- Spring Berlin, Vol.90, No. 5, pp.347-359, 2008.

[47] F. J. Lin, P. H. Chou, Adaptive Control of Two-axis Motion Control System Using Interval Type-2 Fuzzy Neural Network, IEEE Transactions on Industrial Electronics, Vol. 56, No. 1, pp.178-193, 2009.

[48] N. Sahab, H. Hagras, A Type-2 Nonsingleton Type-2 Fuzzy Logic System to Handle Linguistic and Numerical Uncertainties in Real World Environments, IEEE International Symposium on Advances in Type-2 Fuzzy Logic Systems 2011.

[49] A. Ferreoro, S. Salicone, Fully comprehensive mathematical approach to the expression of uncertainty in measurement, IEEE Transactions on Instrumentation and Measurement, Vol. 55, No. 3, pp.706-712, 2006.

[50] N. Sahab, H. Hagras, A Hybrid Approach to Modeling Input Variables in Non-Singleton Type-2 Fuzzy Logic Systems, The 10th Annual UK Workshop in Computational Intelligence (UKCI 2010), pp. 1-6, 2010.

[51] F. Liu, J. M. Mendel, An Interval Approach to Fuzzistics for Interval Type-2 Fuzzy Sets, IEEE International Conference on Fuzzy Systems, pp.1-6, 2007.

[52] J. Lamancusa (2000), Outdoor Sound Propagation, Noise Control, ME 458: Engineering Noise Control, Penn State University, pp.10-19, 2009.

Int. J. of Computers, Communications & Control, ISSN 1841-9836, E-ISSN 1841-9844
Vol. VI (2011), No. 3 (September), pp. 530-561

From Electrical Engineering and Computer Science to Fuzzy Languages and the Linguistic Approach of Meaning: The non-technical Episode: 1950-1975

R. Seising

Rudolf Seising
European Centre for Soft Computing
Edificio Científico-Tecnológico 3ª Planta,
C Gonzalo Gutiérrez Quirós S/N
Mieres, Asturias, Spain
E-mail: rudolf.seising@softcomputing.es

Abstract:
In this paper we illuminate the first decade of Fuzzy Sets and Systems (FSS) where nobody thought that this theory would be successful in the field of applied sciences and technology. We show that especially Lotfi A. Zadeh, the founder of the theory of FSS, expected that his theory would have a role in the future of computer systems as well as humanities and social sciences. When Mamdani and Assilian picked up the idea of FSS and particularly Fuzzy Algorithms to establish a first Fuzzy Control system for a small steam engine, this was the Kick-off for the "Fuzzy-Boom" in Japan and later in the whole world and Zadeh's primary intention trailed away for decades. Just in the new millennium a new movement for Fuzzy Sets in Social Sciences and Humanities was launched and, hopefully, will persist!
Keywords: fuzzy T-S model, fuzzy logic systems, nonlinear systems, uncertainties, tracking control.

1 Introduction

About half a decade after his seminal papers "Fuzzy Sets" and "Fuzzy Sets and Systems" have appeared in print [1, 2], Lotfi A. Zadeh (born 1921), the founder of this mathematical theory, notified that he did not expect the incorporation of fuzzy sets and systems (FSS) into the fields of sciences and engineering. He was then professor and chair of Electrical Engineering (EE) at US Berkeley the name of the department changed in the year 1967 to Electrical Engineering and Computer Science (EECS) – and he wrote: "What we still lack, and lack rather acutely, are methods for dealing with systems which are too complex or too ill-defined to admit of precise analysis. Such systems pervade life sciences, social sciences, philosophy, economics, psychology and many other "soft" fields." [3, 4] In the first years after his foundation, the theory of FSS Zadeh was intended to open the field of its applications to humanities and social sciences. Also reading an interview that was printed in the *Azerbaijan International*, in 1994, we can improve this view: when Zadeh was asked, "How did you think Fuzzy Logic would be used at first?" his retrospective answer was: "In many, many fields. I expected people in the social sciences-economics, psychology, philosophy, linguistics, politics, sociology, religion and numerous other areas to pick up on it. It's been somewhat of a mystery to me, why even to this day, so few social scientists have discovered how useful it could be."

In section III we refer to some of the papers that Zadeh has recited, written, or published between 1965 and 1975 that consolidate the perspective that the inventor of FSS wished to establish his new mathematical theory to the humanities, arts and social sciences. Then he was very surprised when Fuzzy Logic (FL) was first in the 1970s "embraced by engineers" and later

Copyright © 2006-2011 by CCC Publications

in that decennium FSS has been successful "used in industrial process controls and in 'smart' consumer products such as hand-held camcorders that cancel out jittering and microwaves that cook your food perfectly at the touch of a single button." He said: "I didn't expect it to play out this way back in 1965." [5]

We prefix in the following section II the pre-history of FSS and FL to show that this great change in science in the 20th century originates from two developments in its first half: computers and system theory, and Zadeh was involved in both of them. Therefore, in section II.1 we focus his work concerning "thinking machines" and in section II. 2 we consider his progress in generalizing system theory[1].

2 The Age of Computers and System Theory

2.1 From "Thinking Machines" and System Theory to "Making Computers think like people"

After the Second World War, computers next to the atomic bomb, the most famous technical product of war research became popular as "electronic brains" or "thinking machines". This "era of computers" started already with the analogue MIT Differential Analyzer of Vannevar Bush (1890-1974) but the digital machines that have been built during the war, ENIAC (Electronic Numerical Integrator and Computer) and EDVAC (Electronic Discrete Variable Computer), both designed by John Presper Eckert (1919-1995) and John William Mauchly (1907-1980), gave this technological development an eminent push.

Figure 1: Title page of the Columbia Engineering Quarterly, January 1950.

In the spring of 1945 the mathematician John von Neumann (1903-1957) was asked to prepare a report on the logical principles of the EDVAC, since the ENIAC had not had any such description and it had been sorely missed. In this report [7] he adopted the neuron model from a paper of Warren Sturgis McCulloch (1898-1968) and Walter H. Pitts (1924-1959) [8] that explained the brain and nervous system to a logical computer and drew the inverse conclusion. The similarity between neurons and electric switching elements was apparently so clear to him that he did not thoroughly question it. When the British mathematician Alan Mathison Turing (1912-1954) published in 1950 his famous article "Computing Machinery and Intelligence" [9] in

[1]For details to the genesis of FSS and FL see: [6]

the journal *Mind* the answer of the following question was very popular and also of philosophical interest: “Can machines think?” Turing proposed the well-known imitation game, now called the “Turing test”, to decide whether a computer or a program could think like a human being or not. In those days the young electrical engineer Lotfi Aliasker Zadeh (Fig. 2, left side) was interested in these new computing machines. In 1949 he had obtained a position at Columbia University in New York as an instructor responsible for teaching the theories of circuits and electromagnetism but after this year, when he had received his Ph. D., he turned his attention to the problems of computers. Inspired by a lecture of Claude E. Shannon (1916-2001) in New York in 1946, two years before his “Mathematical Theory of Communication” would be published [10], and also by Norbert Wiener’s (1894-1964) famous book *Cybernetics* [11], Zadeh served as a moderator at a debate on digital computers at Columbia University between Shannon, Edmund C. Berkeley (1909-1988), the author of the book *Giant Brains or Machines That Think* published in 1949 [12], and Francis J. Murray (1911-1996), a mathematician and consultant to IBM.

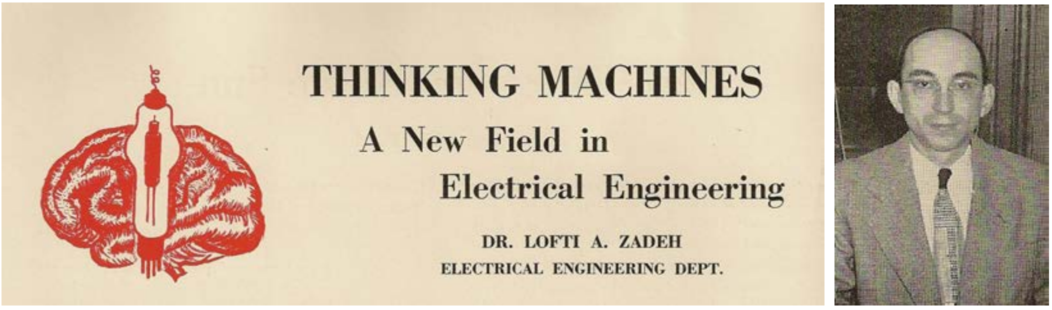


Figure 2: Left: Illustration accompanying Zadeh’s article [13]; right: Lotfi A. Zadeh in the 1940s, picture in [13], p. 13.

Then, in 1950, still unaware of Turing’s philosophical article, Zadeh wrote the paper “Thinking Machines A New Field in Electrical Engineering” (Fig. 2, right side), which appeared in the student journal *The Columbia Engineering Quarterly* (Fig. 1) in New York City in 1950 [13]. Here, Zadeh put up for discussion the questions “How will ’electronic brains’ or ’thinking machines’ affect our way of living?” and “What is the role played by electrical engineers in the design of these devices?” ([13], p. 12.] He was looking for “the principles and organization of machines which behave like a human brain. Such machines were then variously referred to as “thinking machines”, “electronic brains”, “thinking robots”, and similar names. He mentioned that the “same names are frequently ascribed to devices which are not “thinking machines” in the sense used in this article”, therefore he separated as follows: “The distinguishing characteristic of thinking machines is the ability to make logical decisions and to follow these, if necessary, by executive action.” ([13], p. 12.) He stated: “More generally, it can be said, that a thinking machine is a device which arrives at a certain decision or answer through the process of evaluation and selection.” With this definition he decided that the MIT Differential Analyzer was not a thinking machine, but both then built large-scale digital computers, UNIVAC (Universal Automatic Computer) and BINAC (Binary Automatic Computer), were thinking machines because they both were able to make non-trivial decisions. ([13], p. 13.) Zadeh explained in this article “how a thinking machine works” (Fig. 3) and he claimed, “the box labelled Decision Maker is the most important part of the thinking machine”.

Zadeh illustrated his argumentation by peering forward into the year 1965, which was then 15 years in the future. Three years earlier, in this version of the future, the administration at Columbia University had decided, for reasons of economy and efficiency, to close the admissions office and install in its place a thinking machine called the “Electronic Admissions Director”. The

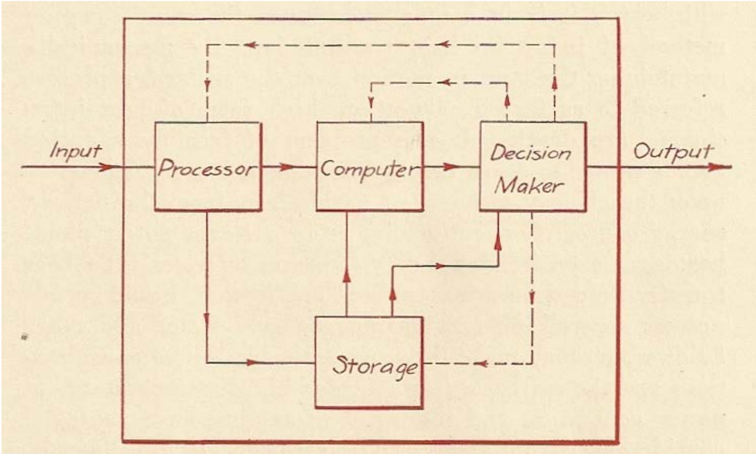


Figure 3: Zadeh's chart for the basic elements of a "Thinking Machine", [13], p. 13.

construction and design of this machine had been entrusted to the electrical engineering department, which completed the installation in 1964. Since then, the "director" has been functioning perfectly and enjoying the unqualified support of the administration, departments and students. This "thinking machine" functions as follows:

1. Human secretaries convert the information from the list of applicants into series of numbers $a_1, a_2, a_3, ..., a_n$; each number represents a characteristic, e. g. a_1 could stand for the applicant's IQ, a_2 for personal character, and so on.

2. The lists coded thusly are provided to the processor, which processes them and then relays some of the data to the computer and another part of the data to storage. On the basis of applicant data as well as university data, the computer calculates the probabilities of various events, such as the probability that a student will fail after the first five years. This information and the saved data are sent to the decision maker to come to final decision on whether to accept the applicant. The decision is then made based on directives, such as these two:

   - Accept if the probability of earning the Bachelor's degree is greater than 60%;
   - Reject if the probability that the applicant will not pass the first year of college is greater than 20%.

Zadeh didn't consider the machine sketched out here to be as fanciful as student readers (and surely others, as well) may have thought: Machines such as this could be commonplace in 10 or 20 years and it is already absolutely certain that thinking machines will play an important role in armed conflicts that may arise in the future. ([13], p. 30) Now, in the year 1950, though, there was still much to be done so that these or similar scenarios of the future could become reality.

"Thinking Machines are essentially electrical devices. But unlike most other electrical devices, they are the brain children of mathematicians and not of electrical engineers. Even at the present time most of the advanced work on Thinking Machines is being done by mathematicians. This situation will last until electrical engineers become more proficient in those fields of mathematics which form the theoretical basis for the design of Thinking Machines. The most important of these fields is that of symbolic logic." ([13], p. 31).

The fundamental principles of “thinking machines”, Zadeh stressed, were developed by mathematicians, but today, after more than 50 years of *Artificial Intelligence* (AI) - a research program that was launched in 1959, that spread to many scientific and technological communities throughout the world and that includes a number of successes - we know that AI has lagged behind expectations. AI became a field of research aimed at developing computers and computer programs that act “intelligently” even though no human being controls these systems. AI methods became methods of computing with numbers and finding exact solutions. As well, humans are able to resolve such tasks very well, as Zadeh mentioned very often over the last decades. In conclusion, Zadeh stated that “thinking machines” do not think as humans do.

In the 1960s Zadeh’s research topic was System Theory, chapter II.2 is concerned with this development in detail, but in the 1970s, Zadeh connected the two research subjects with each other when he distinguished between mechanic (or inanimate or man-made) systems at one hand and humanistic systems at the other hand. He saw the following state of the art in computer technology: “Unquestionably, computers have proved to be highly effective in dealing with mechanistic systems, that is, with inanimate systems whose behaviour is governed by the laws of mechanics, physics, chemistry and electromagnetism. Unfortunately, the same cannot be said about humanistic systems, which so far at least have proved to be rather impervious to mathematical analysis and computer simulation.” He explained that a “humanistic system” is “a system whose behaviour is strongly influenced by human judgement, perception or emotions. Examples of humanistic systems are: economic systems, political systems, legal systems, educational systems, etc. A single individual and his thought processes may also be viewed as a humanistic system.” ([14], p. 200) To summarize, he argued, “that the use of computers has not shed much light on the basic issues arising in philosophy, literature, law, politics, sociology and other human-oriented fields. Nor have computers added significantly to our understanding of human thought processes excepting, perhaps, some examples to the contrary that can be drawn from artificial intelligence and related fields.” ([14], p. 200).

Computers have been very successful in mechanic systems but they could not be that successful humanistic systems in the field of non-exact sciences. Zadeh argued that this is the case because of his so-called *Principle of Incompatibility* that he established in 1973 for the concepts of exactness and complexity: “The closer one looks at a ’real world’ problem, the fuzzier becomes its solution.” [15][3] With this principle there is a difference between system analysis and simulations that are based on precise number computing at one hand and analysis and simulations of humanistic systems at the other hand. Zadeh conjectured that precise quantitative analysis of the behaviour of humanistic systems are not meaningful for “real-world societal, political, economic, and other types of problems which involve humans either as individuals or in groups.” ([15], p. 28).

From the mid-1980s he focused on “Making Computers Think like People”. [16] For this purpose, the machine’s ability to “compute with numbers” was supplemented by an additional ability that was similar to human thinking. The “remarkable human capability [of humans] to perform a wide variety of physical and mental tasks without any measurements and any computations” inspired him and he has given everyday examples of such tasks in many papers: parking a car, playing golf, deciphering sloppy handwriting, and summarizing a story. Underlying this, is the human ability to reason with perceptions “perceptions of time, distance, speed, force, direction, shape, intent, likelihood, truth and other attributes of physical and mental objects.” ([17], p. 903).

[3] More explicitly: “Stated informally, the essence of this principle is, that as the complexity of a system increases, our ability to make precise and yet significant statements about its behaviour diminishes until a threshold is reached beyond which precision and significance (or relevance) become almost mutually exclusive characteristics.” [15]


Making computers think like people

The term 'fuzzy thinking' is pejorative when applied to humans, but fuzzy logic is an asset to machines in applications from expert systems to process control

Figure 4: Headline of L. A. Zadehs 1984-paper [16].

2.2 From Circuit Theory to Fuzzy Systems Theory

Let's go back to the 1950's! Also in the *Columbia Engineering Quarterly* Zadeh published in 1954 the article "System Theory" [18] where, he characterized systems as a "black boxes" with inputs $u_1, ..., u_m$ and outputs $\nu_1, ..., \nu_n$, $(m, n \in N$ (Fig. 5), and in the case that these inputs and outputs are describable as time dependent functions then the dynamic behaviour of the system can be studied mathematically, and the input-output-relationship of the system is

$$(\nu_1, ..., \nu_n) = f(u_1, ..., u_m) \tag{1}$$

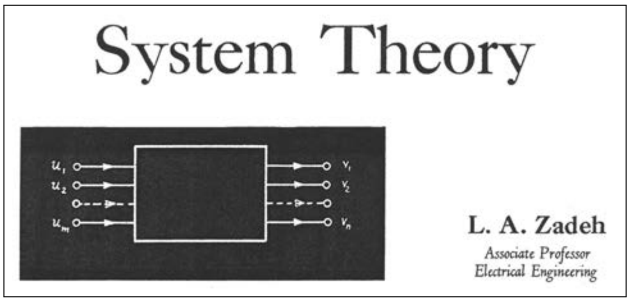


Figure 5: Headline of L. A. Zadehs 1954-paper showing a system with inputs and outputs, [18], p. 16.

Of course, in Zadeh's later system theoretic papers a more sophisticated treatment of these interrelationships is visible: "Suppose that a system A is specified as a given combination of N component systems $A_1, ..., A_N$, each of which is defined as a specified set of input-output pairs. How can one deduce from the knowledge of these sets of input-output pairs and the way in which the components A (that is $A_1, ..., A_N$) are combined the set of input-output pairs which constitutes A? This question presents one of the central problems of system theory." ([18], p. 18).

In the early 1950s, system theory was a rising scientific discipline "to the study of systems per se, regardless of their physical structure". Engineers in that time were, in general, inadequately trained to think in abstract terms, but nevertheless, Zadeh believed that it was only a matter of time before system theory would attain acceptance. It turns out that he was right: Eight years later, when he wrote the article "From Circuit Theory to System Theory" [19] for the anniversary edition of the *Proceedings of the IRE* appeared in May 1962 to mark the 50th year of the Institute of Radio Engineers (IRE), he could describe problems and applications of system theory and its relations to network theory, control theory, and information theory. Furthermore, he pointed out "that the same abstract "systems" notions are operating in various guises in many unrelated fields of science is a relatively recent development. It has been brought about, largely within the past two decades, by the great progress in our understanding of the behaviour of both inanimate and animate systems-progress which resulted on the one hand from a vast expansion in the scientific and technological activities directed toward the development of highly complex systems for such purposes as automatic control, pattern recognition, data-processing, communication, and machine computation, and, on the other hand, by attempts at quantitative analyses of the extremely complex animate and man-machine systems which are encountered in biology, neurophysiology, econometrics, operations research and other fields" ([19], p. 856f.).

After 1958, when Zadeh became a professor of electrical engineering at the University of California, Berkeley he published papers on system theory and two well known books with colleagues at his new department: He authored *Linear System Theory* together with Charles A. Desoer (1926-2010) in 1963 [20] and he edited the volume *System Theory* with Elijah Polak (born 1931) [21]. His own contribution in the latter book has the title "The Concept of State in System Theory" [22]. This concept of state was Zadeh's "new view" in System Theory that he presented also in April 1963, when he participated in the *Second Systems Symposium* at the *Case Institute of Technology* in Cleveland, Ohio. 17 speakers and more than 200 participants, systems scientists in terms of the general systems theory and cybernetics but also technical system scientists tried to discuss relations between technical and nontechnical system science.

The proceedings, published by Mihaljo D. Mesarović (born 1928), were entitled *Views on General Systems Theory* [23] and here Zadeh placed a general notion of state in system theory [24]. His starting points were the fields of dynamical systems and of automata. To present a simple example, Zadeh referred to an important subject in the history of computer science that is named after Alan Turing: the Turing machine. "In From Circuit Theory to System Theory" [19] he came from this idea: "Roughly speaking, a Turing machine is a discrete time ($t = 0, 1, 2, \ldots$) system with a finite number of states or internal configurations, which is subjected to an input having the form of a sequence of symbols (drawn from a finite alphabet) printed on a tape which can move in both directions along its length. The output of the machine at time t is an instruction to print a particular symbol in the square scanned by the machine at time t and to move in one or the other direction by one square. A key feature of the machine is that the output at time $t + 1$ and the state at time $t + 1$ are determined by the state and the input at time t ([19], p. 858). If s_t, u_t, and y_t denote *state, input,* and *output* of the Turing machine at time t, respectively, and if f and g are functions on pairs of s_t and u_t, then the machine-operation is characterized by the following set of state equations:

$$s_{t+1} = f(s_t, u_t),\ t = 0, 1, 2, ... \tag{2}$$

$$y_t = g(s_t, u_t) \tag{3}$$

If the system is a differential system instead of a discrete-state system, state, input, and output of the system are represented by vectors $s(t)$, $y(t)$, and $u(t)$, respectively. With $\dot{s}(t) = d/dt\ s(t)$, state equations assume the forms

$$\dot{s}(t) = f(s(t), u(t)) \tag{4}$$

$$y(t) = g(s(t), u(t)) \tag{5}$$

Some mathematicians and control theorists in the Soviet Union in the 1940s and 1950s, as the Russian mathematician Lew Semjonowitsch Pontrjagin (1908-1988), used these state equations more early than western scientists, and Lotfi Zadeh has been familiar with the scientific progress in the Soviet Union. He referred to the fact that "in the United States, the introduction of the notion of state and related techniques into the theory of optimization of linear as well as nonlinear

Figure 6: Left to right: Charles A. Desoer, Lotfi A. Zadeh, Elijah Polak, Richard E. Bellman

Figure 7: Left to right: Lew S. Pontrjagin, Richard E. Bellman, Lotfi A. Zadeh, Robert Kalaba.

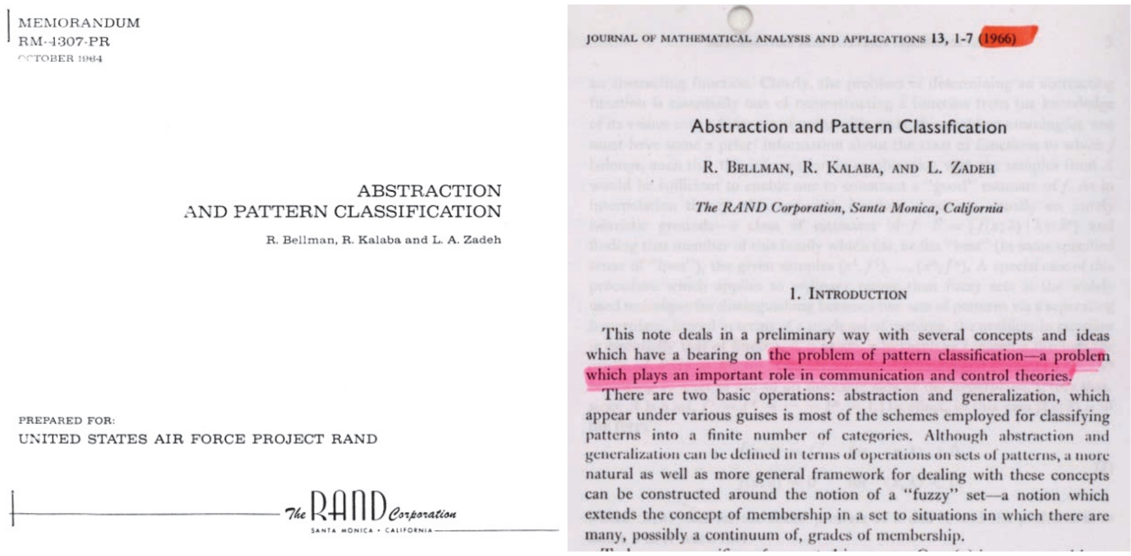
systems is due primarily to Richard Ernest Bellman (1920-1984), whose invention of dynamic programming has contributed by far the most powerful tool since the inception of the variational calculus to the solution of a whole gamut of maximization and minimization problems." ([19], p. 858.).

Bellman and Zadeh have been very close friends and in the summer of 1964 they planned on doing some research together at RAND in Santa Monica where Bellman was employed, but before that time, Zadeh was supposed to give a talk on pattern recognition in the *Wright-Patterson Air Force Base*, Dayton, Ohio.

During this travel he started thinking about pattern recognition problems and grades of membership of an object to be an element of a class as he returned to mind almost 50 years later: "While I was serving as Chair, I continued to do a lot of thinking about basic issues in systems analysis, especially the issue of unsharpness of class boundaries. In July 1964, I was attending a conference in New York and was staying at the home of my parents. They were away. I had a dinner engagement but it had to be canceled. I was alone in the apartment. My thoughts turned to the unsharpness of class boundaries. It was at that point that the simple concept of a fuzzy set occurred to me. It did not take me long to put my thoughts together and write a paper on the subject. This was the genesis of fuzzy set theory." ([26], p.7).[4]

When Zadeh met Bellman in Santa Monica, they both discussed a new theory of membership grades and sets with fuzzy borders. Then, Zadeh wrote a paper on these "fuzzy sets" and some

[4]A more detailled presentation of the history of the theory of FSS and FL is give in the author's book [6]


MEMORANDUM
RM-4307-PR
OCTOBER 1964

ABSTRACTION
AND PATTERN CLASSIFICATION

R. Bellman, R. Kalaba and L. A. Zadeh

PREPARED FOR:
UNITED STATES AIR FORCE PROJECT RAND

The RAND Corporation
SANTA MONICA • CALIFORNIA

JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 13, 1-7 (1966)

Abstraction and Pattern Classification

R. BELLMAN, R. KALABA, AND L. ZADEH

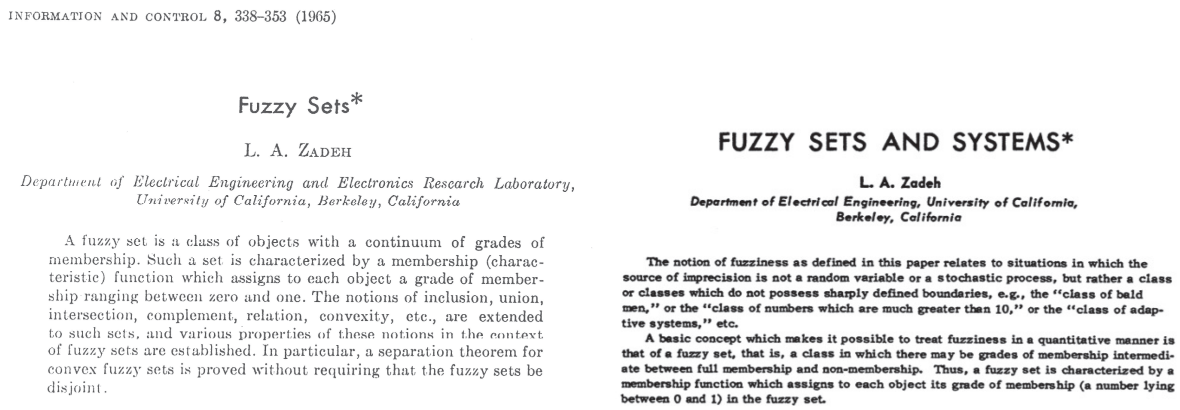
The RAND Corporation, Santa Monica, California

1. INTRODUCTION

This note deals in a preliminary way with several concepts and ideas which have a bearing on the problem of pattern classification—a problem which plays an important role in communication and control theories.

There are two basic operations: abstraction and generalization, which appear under various guises is most of the schemes employed for classifying patterns into a finite number of categories. Although abstraction and generalization can be defined in terms of operations on sets of patterns, a more natural as well as more general framework for dealing with these concepts can be constructed around the notion of a "fuzzy" set—a notion which extends the concept of membership in a set to situations in which there are many, possibly a continuum of, grades of membership.

Figure 8: Left: title page of the Rand-memorandum [27]; right: title page of the article [28].


INFORMATION AND CONTROL 8, 338–353 (1965)

Fuzzy Sets*

L. A. ZADEH

Department of Electrical Engineering and Electronics Research Laboratory, University of California, Berkeley, California

A fuzzy set is a class of objects with a continuum of grades of membership. Such a set is characterized by a membership (characteristic) function which assigns to each object a grade of membership ranging between zero and one. The notions of inclusion, union, intersection, complement, relation, convexity, etc., are extended to such sets, and various properties of these notions in the context of fuzzy sets are established. In particular, a separation theorem for convex fuzzy sets is proved without requiring that the fuzzy sets be disjoint.

FUZZY SETS AND SYSTEMS*

L. A. Zadeh

Department of Electrical Engineering, University of California, Berkeley, California

The notion of fuzziness as defined in this paper relates to situations in which the source of imprecision is not a random variable or a stochastic process, but rather a class or classes which do not possess sharply defined boundaries, e.g., the "class of bald men," or the "class of numbers which are much greater than 10," or the "class of adaptive systems," etc.

A basic concept which makes it possible to treat fuzziness in a quantitative manner is that of a fuzzy set, that is, a class in which there may be grades of membership intermediate between full membership and non-membership. Thus, a fuzzy set is characterized by a membership function which assigns to each object its grade of membership (a number lying between 0 and 1) in the fuzzy set.

Figure 9: Headline of Zadeh's article "Fuzzy Sets" [29] of his proceedings contribution, emanated from his talk "A New View on System Theory" [30].

pattern recognition problems. This paper appeared as a RAND-Corporation memorandum in October 1964 by the authors Bellman, his collaborator Robert Kalaba (1926-2004), and Zadeh [27]. The paper, written by Lotfi Zadeh, contains the first definitions of the theory of fuzzy sets in a scientific text. Two years later the same paper appeared under the same title and authorship in the *International Journal for Applied Mathematics and Applications* [28].

Zadeh submitted his first article "Fuzzy Sets" to the editors of *Information and Control* in November 1964 and it appeared in this journal in June 1965 [27]. Many years later he wrote: "I knew that the word "fuzzy" would make the theory controversial. Knowing how the real world functions, I submitted my paper to *Information and Control* because I was a member of the Editorial Board. There was just one review – which was very lukewarm. I believe that my paper would have been rejected if I were not on the Editorial Board. Today, with over 26,000 Google Scholar citations, "Fuzzy sets" is by far the highest cited paper in *Information and Control.*" ([27], p.7).

Zadeh's third paper on fuzzy sets came out in the proceedings of the *Symposium on System*

Theory (April 20-22, 1965) at the Polytechnic Institute in Brooklyn, where Zadeh gave a talk with the title "A New View on System Theory"[7]. This view dealt with the concepts of fuzzy sets and in this talk Zadeh defined "fuzzy systems": A system S is a *fuzzy system* if (input) $u(t)$, output $y(t)$, or state $x(t)$ of S or any combination of them ranges over fuzzy sets. ([30], p. 33) He explained that "these concepts relate to situations in which the source of imprecision is not a random variable or a stochastic process but rather a class or classes which do not possess sharply defined boundaries." ([30], p. 29) He noticed, "Such classes are not classes or sets in the usual sense of these terms, since they do not dichotomize all objects into those that belong to the class and those that do not"[8]. He introduced "the concept of a "fuzzy set", that is a class in which there may be a continuous infinity of grades of membership, with the grade of membership of an object x in a fuzzy set A represented by a number $\mu A(x)$ in the interval $[0, 1]$". Zadeh maintained that these new concepts provide a "convenient way of defining *abstraction* - a process which plays a basic role in human thinking and communication." ([30], p. 29).

Also 17 years later, in his contribution to the 2nd volume of Kluwer's series "Frontiers in System Research", titled *Systems Methodology in Social Science Research: Recent Developments*, Zadeh wrote: "The systems theory of the future – the systems theory that will be applicable to the analysis of humanistic systems – is certain to be quite different in spirit as well as in substance from systems theory as we know it today. I will take the liberty of referring to it as fuzzy systems theory because I believe that its distinguishing characteristics will be a conceptual framework for dealing with a key aspect of humanistic systems – namely the pervasive fuzziness of almost all phenomena that are associated with their external as well as internal behavior" [25]. He concluded this paper as follows: "Fuzzy systems theory is not yet an existing theory. What we have at present are merely parts of its foundations. Nevertheless, even at this very early stage of its development, fuzzy systems theory casts some light on the process of approximate reasoning in human decision making, planning, and control. Furthermore, in the years ahead, it is likely to develop into an effective body of concepts and techniques for the analysis of large-scale humanistic as well as mechanistic systems." ([25], p. 39) More than a decade later, in 1994 he presented *perception-based system modeling*: "A system, S, is assumed to be associated with temporal sequences of input $X_1, X_2, \ldots$; output $Y_1, Y_2, \ldots$; and states $S_1, S_2, \ldots$. S_2 is defined by state-transition function f with $S_{t+1} = f(S_t, X_t)$, and output function g *with* $Y_t = g(S_t, X_t)$, for $t = 0, 1, 2, \ldots$ In perception-based system modelling, inputs, outputs and states are assumed to be perceptions, as state-transition function, *f*, and output function, *g*." ([32], p. 77.).

This view on future artificial perception-based systems (Fig. 9) – CW-systems and therefore systems to reasoning with perceptions – is the goal of CTP. This view is closely linked by regarding the human brain as a fundamentally fuzzy system. Only in very few situations do people reason in binary terms, as machines classically do. This human characteristic is reflected in all natural languages, in which very few terms are absolute. The use of language is dependent on specific situations and is very seldom 100% certain. For example, the word "thin" cannot be defined in terms of numbers and there is no measurement at which this term suddenly stops being applicable. Human thinking, language and reasoning can thus indeed be called fuzzy. The theory of fuzzy sets has created a logical system far closer to the functionality of the human mind than any previous logical system.[9]

Fuzzy Sets and Systems and Fuzzy Logic enable them to express uncertainty regarding mea-

[7] In the Proceedings of this symposium there is a shortened manuscript version of the talk with the heading Fuzzy Sets and Systems [30]

[8] [29], p. 29. Zadeh used quotation marks to indicate the difference between usual classes or sets and his new (fuzzy) sets.

[9] For more details concerning the history of the theory of Fuzzy Sets and Systems and also its technical applications see [6].

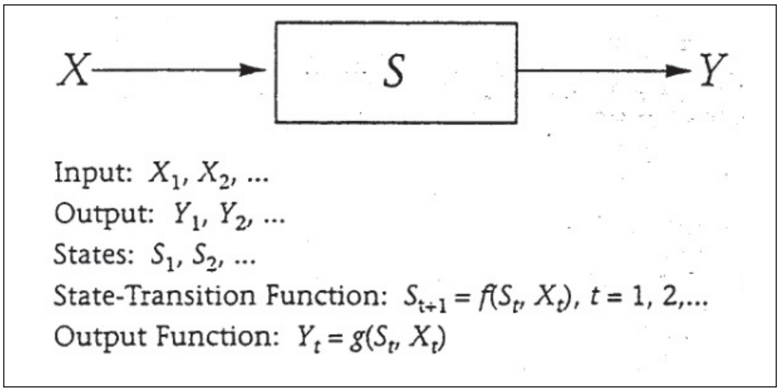


Figure 10: Perception-based system modelling, [32].

surements, diagnostics, evaluations, etc. In theory, this should put the methods of communication used by machines and human beings on levels that are much closer to each other.

3 From Fuzzy Logic to Fuzzy Languages

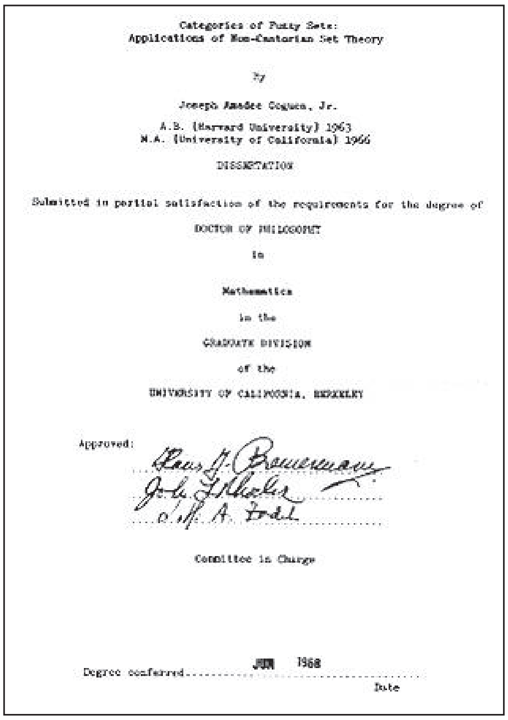
3.1 Fuzzy Logic

To understand what happened from coming from Fuzzy Sets and Systems to the idea of the CTP we have go once again back to the roots. In the 1960s Zadeh looked for applying fuzzy sets in linguistics. This idea led to interdisciplinary scientific exchange on the campus of the University of California at Berkeley between him and the mathematician Joseph Goguen (1941-2006) - who was a Ph. D. student of his, his Berkeley-colleague Hans-Joachim Bremermann (1926-1996), who was then in the mathematics department on the one hand and between the psychologist Eleanor Rosch (Heider) (born 1938) and the Berkeley-linguist George Lakoff (born 1941) on the other. Zadeh had served as first reviewer for Goguens's Ph.D. thesis "Categories of Fuzzy Sets" [33] and Bremermann served as the second. In this work, Goguen generalized the fuzzy sets to so-called "L-sets" [34]. An L-set is a function that maps the fuzzy set carrier X into a partially ordered set $L : A : X \to L$. The partially ordered set L Goguen called the "truth set" of A. The elements of L can thus be interpreted as "truth values"; in this respect, Goguen then also referred to a "Logic of Inexact Concepts" [35].

Since Zadeh's earlier definition had established this truth set as the unit interval, Fuzzy Set Theory was very soon associated with multi-valued logics, and also Lotfi Zadeh mentioned this in later papers, e.g.: "It should be noted that a membership function may be regarded as a predicate in a multivalued logic in which the truth values range over {0, 1]." ([30], p. 131, fn. 2). Goguen's generalization of the set of values to a set L for which the only condition was to be partially ordered cleared up these misunderstandings. Goguen's work was laid out in terms of logical algebra and category theory, and his proof of a representation theorem for L-sets within category theory justified Fuzzy Set Theory as an expansion of set theory.

3.2 Fuzziness for Biology and Computer Science

Also in 1969 Zadeh gave a talk on "Biological Applications of the Theory of Fuzzy Sets and Systems" [36] where he proposed his new mathematical theory to the life scientists, when he wrote: "The great complexity of biological systems may well prove to be an insuperable block to the achievement of a significant measure of success in the application of conventional mathematical techniques to the analysis of systems." [36] "By 'conventional mathematical techniques'


Categories of Fuzzy Sets:
Applications of Non-Cantorian Set Theory

By

Joseph Amadee Goguen, Jr.

A.B. (Harvard University) 1963
M.A. (University of California) 1966

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Mathematics

in the

GRADUATE DIVISION

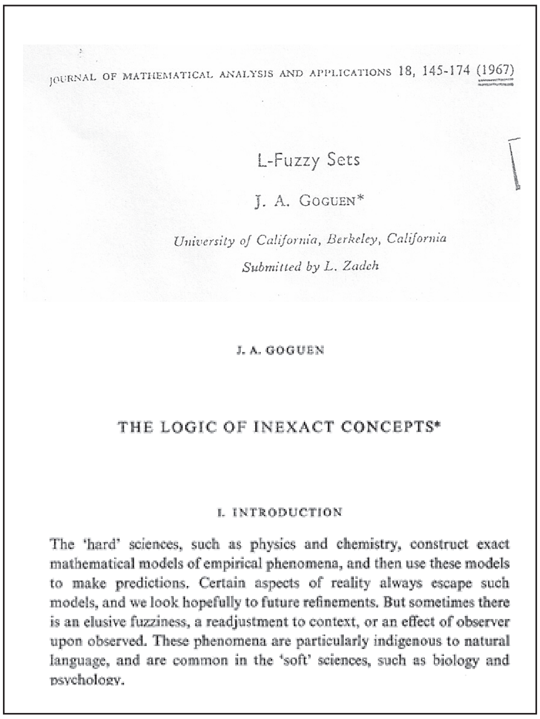
of the

UNIVERSITY OF CALIFORNIA, BERKELEY

Approved:

Committee in Charge

Degree conferred JUN 1968
Date


JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 18, 145-174 (1967)

L-Fuzzy Sets

J. A. GOGUEN*

University of California, Berkeley, California

Submitted by L. Zadeh

J. A. GOGUEN

THE LOGIC OF INEXACT CONCEPTS*

I. INTRODUCTION

The 'hard' sciences, such as physics and chemistry, construct exact mathematical models of empirical phenomena, and then use these models to make predictions. Certain aspects of reality always escape such models, and we look hopefully to future refinements. But sometimes there is an elusive fuzziness, a readjustment to context, or an effect of observer upon observed. These phenomena are particularly indigenous to natural language, and are common in the 'soft' sciences, such as biology and psychology.

Figure 11: Left: Title page of Goguen's PhD thesis [33], right: Title pages of Goguen's articles [34] and [35].

Figure 12: Left to right: Hans J. Bremermann, Joseph Goguen, Eleanor Rosch, George Lakoff.

in this statement, we mean mathematical approaches for which we expect that precise answers to well-chosen precise questions concerning a biological system should have a high degree of relevance to its observed behavior. Indeed, the complexity of biological systems may force us to alter in radical ways our traditional approaches to the analysis of such systems. Thus, we may have to accept as unavoidable a substantial degree of fuzziness in the description of the behavior of biological systems as well as in their characterization." [36] In the same year he wrote more generally: "What we still lack, and lack rather acutely, are methods for dealing with systems which are too complex or too ill-defined to admit of precise analysis. Such systems pervade life sciences, social sciences, philosophy, economics, psychology and many other "soft" fields." [37,38].

Since that time, Zadeh is inspired by the "remarkable human capability to perform a wide variety of physical and mental tasks without any measurements and any computations", e. g. parking a car, playing golf, deciphering sloppy handwriting, and summarizing a story.

However, in 1970, 20 years after later then his first paper on "Thinking machines" [13], Zadeh

Biological Application of the Theory of Fuzzy Sets and Systems

L. A. Zadeh
University of California, Berkeley, California

Toward a Theory of Fuzzy Systems

L. A. Zadeh

Figure 13: Headlines of Zadeh's 1969 papers [36] and [38].

Theoretical Research and Informatic

Man and Computer. Proc. int. Conf., Bordeaux 1970, pp. 130–165
(Karger, Basel 1972)

Fuzzy Languages and their Relation to Human and Machine Intelligence[1]

L. A. Zadeh
Department of Electrical Engineering and Computer Sciences
and the Electronics Research Laboratory, University of California, Berkeley, Calif.

Figure 14: Headlines of Zadeh's 1970 paper "Fuzzy Languages and their Relation to Human and Machine Intelligence" [31].

was aware of Turing's philosophical article [9] when he presented his paper "Fuzzy Languages and their Relations to Human and Machine Intelligence" at the conference "Man and Computer" in Bordeaux, France: "The question of whether or not machines can think has been the subject of many discussions and debates during the past two decades."[11] He continued: "As computers become more powerful and thus more influential in human affairs, the philosophical aspects of this question become increasingly overshadowed by the practical need to develop an operational understanding of the limitations of the machine judgment and decision making ability. Can computers be relied upon to match people, decide on promotions and dismissals, make medical diagnoses, prescribe treatments, act as teachers, formulate business, political and military strategies, and, more generally, perform intellectual tasks of high complexity which in the past required expert human judgment? Clearly, this is already a pressing issue which is certain to grow in importance in the years ahead." ([31], p. 130).

He called it a paradox that the human brain is always solving problems by manipulating "fuzzy concepts" and "multidimensional fuzzy sensory inputs" whereas "the computing power of the most powerful, the most sophisticated digital computer in existence is not able to do this". Therefore, he stated that "in many instances, the solution to a problem need not be exact, so that a considerable measure of fuzziness in its formulation and results may be tolerable. The human brain is designed to take advantage of this tolerance for imprecision whereas a digital computer, with its need for precise data and instructions, is not." ([31], p. 132) One year later these arguments should culminate in Zadeh's *Principle of Incompatibility* that we mentioned already in section II.1, whereas here he intended to push his theory of fuzzy sets to model the imprecise concepts and directives: "Indeed, it may be argued that much, perhaps most, of human thinking and interaction with the outside world involves classes without sharp boundaries in which the

[11]Here, Zadeh gave 10 citations, including that of Turing: [40, 9, 7, 41- 47].

transition from membership to non-membership is gradual rather than abrupt." ([31], p. 131) He stated: "Although present-day computers are not designed to accept fuzzy data or execute fuzzy instructions, they can be programmed to do so indirectly by treating a fuzzy set as a data-type which can be encoded as an array [...]."[12]

Granted that this is not a fully satisfactory approach to the endowment of a computer with an ability to manipulate fuzzy concepts, it is at least a step in the direction of enhancing the ability of machines to emulate human thought processes. It is quite possible, however, that truly significant advances in artificial intelligence will have to await the development of machines that can reason in fuzzy and non-quantitative terms in much the same manner as a human being." ([31], p. 132).

In August 1967, the Filipino electrical engineer William Go Wee (born 1937) at Purdue University in Indiana had submitted his dissertation "On Generalizations of Adaptive Algorithms and Application of the Fuzzy Sets Concept to Pattern Classification" that he had written under King Sun Fu[13], one of the pioneers in the field of pattern recognition (see Fig. 15). Wee had applied the fuzzy sets to iterative learning procedures for pattern classification and had defined a finite automaton based on Zadeh's concept of the fuzzy relation as a model for nonsupervised learning systems: "The decision maker operates deterministically. The learning section is a fuzzy automaton. The performance evaluator serves as an unreliable "teacher" who tries to teach the "student" to make right decisions." ([50], p. 101).

The fuzzy automaton representing the learning section implemented a "nonsupervised" learning fuzzy algorithm and converged monotonously. Wee showed that this fuzzy algorithm could not only be used in the area of pattern classification but could also be translated to control and regulation problems. He also demonstrated that the fuzzy automaton he had defined contained the concepts of deterministic and non-deterministic automata as special cases: "Based on the concept of fuzzy relation defined by Zadeh, a class of fuzzy automata is formulated similar to that of Mealy's definition. A fuzzy automaton behaves in a deterministic fashion. However, it has many properties similar to that of a probabilistic automaton." ([50], p. 88) Working with his doctoral advisor, Wee presented his findings in the article "A Formulation of Fuzzy Automata and its Applications as a Model of Learning Systems" [51].

At the end of the same year the Chinese student Chin-Liang Chang completed his dissertation "Fuzzy Sets and Pattern Recognition" that was an advancement of Zadeh's thoughts on the separation problem in pattern recognition. This was the first Ph D dissertation on Fuzzy Sets that was supervised by Lotfi Zadeh[14] (see Fig. 1). Chang had also had contact with Professor King Sun Fu to whom he expresses gratitude for the conversations they shared. [52].

Two years later, in "Towards a Theory of Fuzzy Systems" that was first printed as a report in 1969 [37, 38], Zadeh's goal was a theory for all systems - including those that were too complex or poorly defined to be accessible to a precise analysis. Alongside the systems of the "soft" fields, the "non-soft" fields were replete with systems that were only "unsharply" defined, namely "when the complexity of a system rules out the possibility of analyzing it by conventional mathematical means, whether with or without the computers". ([38], p. 469f) As he would also do in the same year in Bordeaux [38], Zadeh was already pointing out here the usefulness of fuzzy sets in computer science: In describing their fields of application, he enumerated the problems that would be solved by future computers. Alongside pattern recognition, these included traffic control systems, machine translation, information processing, neuronal networks and games like chess and checkers. We had lost sight of the fact that the class of non-trivial problems for which one

[12] Here, Zadeh referenced to the early article (1970) on "Fuzzy programs" by his student Shi Kuo Chang [48].

[13] King Sun Fu (1930-1985) was a professor at the Purdue School of Electrical Engineering, West Lafayette, Indiana (1960-1985) and was the founding president of the International Association for Pattern Recognition.

[14] A listing of dissertations supervised by Zadeh: http://www.eecs.berkeley.edu/Pubs/Dissertations/Faculty/zadeh.html.

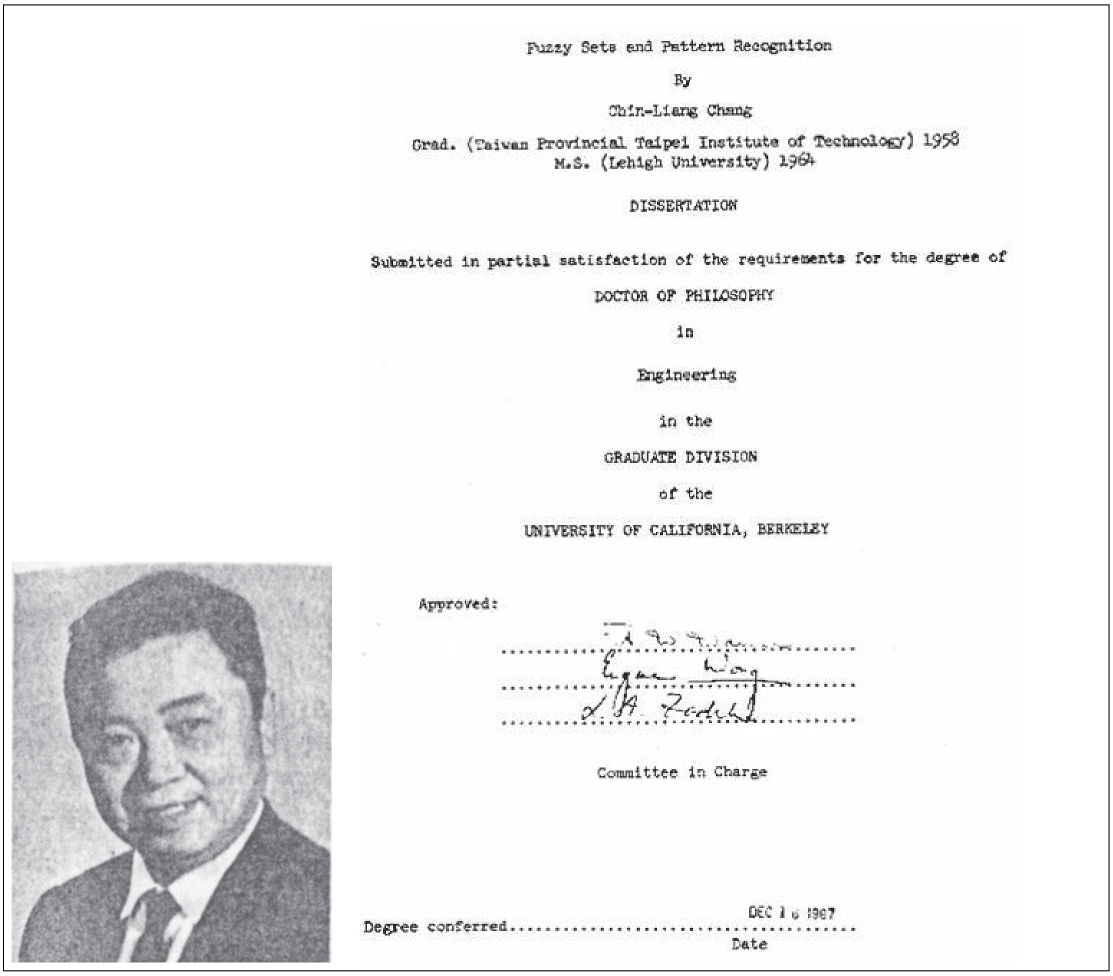
Figure 15: Attendees at the 1st NAFIPS Meeting in Logan, Utah, 1982. King Sun Fu is in the middle of the first row. (Left to right) 1. row: Paul P. Wang, Ebrahim Mamdani, King Sun Fu, James T. P. Yao, L. Saitta. 2. row: Marc Roubens, Philippe Smets, Janet Efstathiou, Richard Tong, Ron R. Yager. Top row: Piero Bonissone, Jim Bezdek, Enrique H. Ruspini, Elie Sanchez.

could find a precise solution algorithm was very limited, he wrote. Most real problems were much too complex and thus either completely unsolvable algorithmically or – if they could be solved in principle – not arithmetically feasible. In chess, for instance, there was in principle an optimal playing strategy for each stage of the game; in reality, however, no computer was capable of sifting through the entire tree of decisions for all of the possible moves with forward and backward repetitions in order to then decide what move would be the best in each phase of the game. The set of good strategies for playing chess had fuzzy limits similar to the set of tall men - these were fuzzy sets. By far the most systems that remained to be solved were fuzzy systems, and in a footnote Zadeh remarks that the automata proposed by Wee and his supervisor Eugene Santos[15] were also considered examples of fuzzy systems. ([38], p. 471, fn. 1).

To make fuzziness a part of system theory, Zadeh presented in 1968 "fuzzy algorithms". With that, he had fuzzified the central concept of computer sciences. "The concept in question will be called a fuzzy algorithm because it may be viewed as a generalization, through the process of fuzzification, of the conventional (nonfuzzy) conception of an algorithm." ([53], p. 94).

Algorithms depend upon precision. An algorithm must be completely unambiguous and error-free in order to result in a solution. The path to a solution amounts to a series of commands which must be executed in succession. Algorithms formulated mathematically or in a programming language are based on set theory. Each constant and variable is precisely defined, every function and procedure has a definition set and a value set. Each command builds upon them. Successfully running a series of commands requires that each result (output) of the execution of a command lies in the definition range of the following command, that it is, in other words, an element of the input set for the series. Not even the smallest inaccuracies may occur when defining these

[15]Eugene S. Santos studied at the Mapua Institute of Technology in Manila until 1963 and afterward at Ohio State University in Columbus, where he earned the Ph.D. in 1965. In 1974, he became a professor in the department of computer sciences and information systems at Youngstown State University in Youngstown, Ohio.



Fuzzy Sets and Pattern Recognition

By

Chin-Liang Chang

Grad. (Taiwan Provincial Taipei Institute of Technology) 1958
M.S. (Lehigh University) 1964

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Engineering

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, BERKELEY

Approved:

Committee in Charge

Degree conferred DEC 1 6 1967
Date

Figure 16: Left: Chin-Liang Chang; right: Title page of Chang's PhD thesis [52].

coordinated definition and value ranges.

After 1965, when Zadeh had fuzzified input, output and state in system theory and had thus founded a theory of fuzzy systems [30], it was obvious to him how to go about fuzzifying algorithms. The commands needed to be fuzzified and so, of course, did their relations! "I began to see that in real life situations people think certain things. They thought like algorithms but not precisely defined algorithms." [53].

Inspired by this idea, he wrote the article "Fuzzy Algorithms" for *Information and Control* in 1968 which uncharacteristically contained neither theorems nor proofs. Many years later he said that "it is not really a mathematical paper. And the reason why it appeared there is because, again, I was on the editorial board. So it could be published quickly. And I do say it's not a mathematical paper but the idea. But then other people who were mathematicians have developed that and added more mathematical and so forth. So, my function was not that of coming up with very precise. It's just an idea. That's little bit like a composer who just hums something, a sort of orchestrazing ..." [54]. In this article he wrote: "Essentially, its purpose is to introduce a basic concept which, though fuzzy rather than precise in nature, may eventually prove to be of use in a wide variety of problems relating to information processing, control, pattern recognition, system identification, artificial intelligence and, more generally, decision processes involving incomplete or uncertain data. The concept in question will be called fuzzy algorithm because it may be viewed as a generalization, through the process of fuzzification, of the conventional (nonfuzzy) conception of an algorithm". ([53], p. 94).

To illustrate, fuzzy algorithms may contain fuzzy instructions such as: (a) "Set y approximately equal to 10 if x is approximately equal to 5," or (b) "If x is large, increase y by several units," or (c) "If x is large, increase y by several units; if x is small, decrease y by several units; otherwise keep y unchanged." The sources of fuzziness in these instructions are fuzzy sets which are identified by their underlined names. ([53], p. 94f).

All people function according to fuzzy algorithms in their daily life, Zadeh wrote – they use recipes for cooking, consult the instruction manual to fix a TV, follow prescriptions to treat illnesses or heed the appropriate guidance to park a car. Even though activities like this are not normally called algorithms: "For our point of view, however, they may be regarded as very crude forms of fuzzy algorithms". ([53], p. 95).

In that time Zadeh wrote also a paper with the title "Toward Fuzziness in Computer Systems. Fuzzy Algorithms and Languages" [55]. I found this typeset-script in Zadeh's office without any reference of publication and date and perhaps it did never appear in print. In the survey-like section on Fuzzy Algorithms he wrote in a footnote "More detailed discussions may be found in [38] and [53]".[16] Therefore this paper was not written before 1969. The next section in this article is titled "Fuzzy Languages". It has also just two pages and a footnote says: "A more detailed discussion of fuzzy languages appears in a forthcoming paper by E. T. Lee and the writer." Because this paper was published in the journal *Information Sciences* in 1969 the article [55] is not younger. As a consequence the date of text [55] is 1969, too.

However, the association of fuzziness and computers in the title of this paper must have sounded surprisingly in the late 1960s and referring to that Zadeh set in the introduction to this paper: "At first glance, it may appear highly incongruous to mention computers and fuzziness in the same breath, since fuzziness connotes imprecision whereas precision is a major desideratum in computer design." ([55], p. 9) In the following paragraphs Zadeh justified this with arguing that the future computer systems will have to perform many more complex information processing tasks than that kind of computers that he and his contemporaries in the 1960s knew. He expected that the future computers have to process more and more imprecise information! "Fuzziness, then, is a concomitant of complexity. This implies that as the complexity of a task, or a system for performing that task, exceeds a certain threshold, the system must necessarily become fuzzy in nature. Thus, with the rapid increase in the complexity of the information processing tasks which the computers are called upon to perform, a point is likely to be reached – perhaps within the next decade - when the computers will have to be designed for processing of information in fuzzy form. In fact, it is this capability - a capability which present-day computers do not possess - that distinguishes human intelligence from machine intelligence. Without such capability we cannot build computers that can summarize written text, translate well from one natural language to another, or perform many other tasks that humans can do with ease because of their ability to manipulate fuzzy concepts." ([55], p. 10) For that purpose, Zadeh pointed out, "intriguing possibilities for computer systems" are offered by fuzzy algorithms and fuzzy languages!

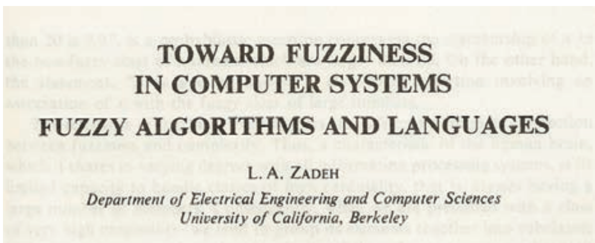
3.3 Fuzzy Languages

Zadeh introduced fuzzy algorithms in the following way: "Roughly speaking, a fuzzy algorithm is an ordered set of instruction containing names of fuzzy sets. An example of such an instruction is "If x is large, set y equal to 2. Otherwise, set y equal to 1"." ([55], p. 13) To execute such fuzzy instructions by computers they have to get an expression in fuzzy programming languages. Consequently the next step for Zadeh was to define fuzzy languages. Beginning the section with this title in [55], he wrote "All languages, whether natural of artificial, tend to evolve and rise

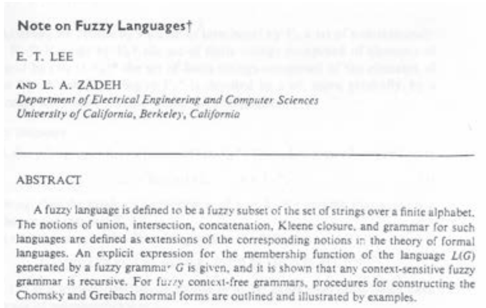
[16] However, in Zadeh's paper these articles have different numbers in the References.

in level through the addition of new words to their vocabulary. These new words are, in effect, names for ordered subsets of names in the vocabulary to which they are added." ([5], p. 16).

Real world phenomena are very complex and rich of members. To characterize or picture these phenomena in terms of our natural languages we use our vocabulary and because this set of words is restricted, Zadeh argued that this process leads to fuzziness: "Consequently, when we are presented with a class of very high cardinality, we tend to group its elements together into subclasses in such a way as to reduce the complexity of the information processing task involved. When a point is reached where the cardinality of the class of subclasses exceeds the information handling capacity of the human brain, the boundaries of the subclasses are forced to become imprecise and fuzziness becomes a manifestation of this imprecision. This is the reason why the limited vocabulary we have for the description of colors makes it necessary that the names of colors such as red, green, bleu [sic.], purple, etc. be, in effect, names of fuzzy rather than non-fuzzy sets. This is why natural languages, which are much higher in level than programming languages, are fuzzy whereas programming languages are not." ([55], p. 10) Here, Zadeh argued explicitly for programming languages that are – because of missing rigidness and preciseness and because of their fuzziness – more like natural languages. He mentioned the concept of *stochastic languages* that was published by the Finnish mathematician Paavo Turakainen in *Information and Control* in the foregoing year [56], being such an approximation to our human languages using randomizations in the productions, but however, he preferred fuzzy productions to achieve a formal fuzzy language. Then, he presented a short sketch of his program to extend non-fuzzy formal languages to fuzzy languages which he published in elaborated form with the co-author Edward T.-Z. Lee in "Note on Fuzzy Languages" [57].[17] His definition in these early papers was given in the terminology of the American computer scientists John Edward Hopcroft (born 1939) and Jeffrey David Ullman (born 1942) that was published in the same year [58].


TOWARD FUZZINESS
IN COMPUTER SYSTEMS
FUZZY ALGORITHMS AND LANGUAGES

L. A. ZADEH
Department of Electrical Engineering and Computer Sciences
University of California, Berkeley


Note on Fuzzy Languages†

E. T. LEE

AND L. A. ZADEH
Department of Electrical Engineering and Computer Sciences
University of California, Berkeley, California

ABSTRACT

A fuzzy language is defined to be a fuzzy subset of the set of strings over a finite alphabet. The notions of union, intersection, concatenation, Kleene closure, and grammar for such languages are defined as extensions of the corresponding notions in the theory of formal languages. An explicit expression for the membership function of the language *L(G)* generated by a fuzzy grammar *G* is given, and it is shown that any context-sensitive fuzzy grammar is recursive. For fuzzy context-free grammars, procedures for constructing the Chomsky and Greibach normal forms are outlined and illustrated by examples.

Figure 17: Left: Headline of Zadeh's script from 1969 [55]; right: Headline of the article of Lee and Zadeh [57].

L is a *fuzzy language* if it is a fuzzy set in the set V_T^*[18] of all finite strings composed of elements of the finite set of terminals V_T, e.g. $V_T = \{a, b, c, ..., z\}$. The membership function $\mu_L(x) : V_T^* \to [0, 1]$ associates with each finite string x, composed of elements in V_T, its grade of membership in L. Here is one of the simple examples that he gave in the early article ([55], p. 16): "Assume that $V_T = \{0,\ 1\}$, and take L to be the fuzzy set $L = \{(0, 0.9), (1, 0.2), (00, 0.8), (01, 0.6), (10, 0.7), (11, 0.3)\}$ with the understanding that all the other strings in V_T^* do not belong to L (i.e., have grade of membership equal to zero)." ([55], p. 16).

In general the language L has high cardinality and therefore it is not usual to define it by a listing of its elements but by a finite set of rules for generating them. Thus, in analogy to the

[17]Later, E. T. Lee finished his Ph D thesis "Fuzzy Languages and Their Relation to Automata" under Zadeh's supervision: [1972]

[18]V_T^* is called the Kleene closure of V_T, named after the American mathematician Stephen Kleene (1909-1994).

case of non-fuzzy languages Zadeh defined a *fuzzy grammar* as "a quadruple $G = (V_N, V_T, P, S)$, where V_N is a set of variables (non-terminals) disjoint from V_T, P is a set of [fuzzy] productions and S is an element of V_N. The elements of V_N (called *[fuzzy] syntactic categories*) and S is an abbreviation for the syntactic category "sentence". The elements of P define conditioned fuzzy sets in $(V_T \cup V_N)^*$." ([55], p. 16).

Turning to the 1970's Zadeh worked out the basic framework of his theory of fuzzy sets and fuzzy relations that gave him the opportunity to characterize fuzzy languages broader than before.

3.4 Fuzzy Relations

In 1971, Zadeh defined similarity relations and fuzzy orderings [59]. In doing so, he was proceeding from the concept of fuzzy relations as a fuzzification of the relation concept known in conventional set theory that he had already defined in his first text on fuzzy sets [1]: If X and Y are conventional sets and if $X \times Y$ is their Cartesian product, then:

- $L(X)$ is the set of all fuzzy sets in X,
- $L(Y)$ is the set of all fuzzy sets in Y and,
- *L(X × Y)* is the set of all fuzzy sets in $X \times Y$.

Relations between X and Y are subsets of their Cartesian product $X \times Y$, and the composition $t = q*$ r of the relation $q \subseteq X \times Y$ with the relation $r \subseteq Y \times Z$ into the new relation $t \subseteq X \times Z$ is given by the following definition: $t = q * r = \{(x, y) \exists y : (x, y) \in q \wedge (y, z) \in r\}$.

Fuzzy relations between sets X and Y are subsets in $L(X \times Y)$. For three conventional sets X, Y and Z, the fuzzy relation Q between X and Y, and the fuzzy relation R between Y and Z are defined: $Q \in L(X \times Y)$ and $R \in L(Y \times Z)$. The combination of these two fuzzy relations into a new fuzzy relation $T \in L(X \times Z)$ between X and Z can then be combined from the fuzzy relations Q and R into the new fuzzy relation $T \in L(X \times Z)$ when the logical conjunctions are replaced by the corresponding ones of the membership functions.

- The above definition of the composition of conventional relations includes a logical AND $(\wedge)$, which, for the "fuzzification", is replaced by the minimum operator that is applied to the corresponding membership functions.[19]
- The above definition of the composition of conventional relations includes the expression "$\exists$ y" ("there exists a y"). The existing $y \in Y$ is the first or the second or the third ... (and so on); written logically: $(\vee)$ sup $y \in Y$. In the "fuzzifications", the logical OR conjunction is replaced by the maximum operator that is applied to the corresponding membership functions.[20]

The fuzzy relation $T = Q * R$ is therefore defined via Zadeh's "combination rule of max-min combination"[21] for the membership functions: $\mu_T(x, y) = max_{y \in Y} min\{\mu_Q(x, y); \mu_r(y, z)\}, y \in Y$.

[19] Of course, the other proposed fuzzy operators can also be used; in those cases, correspondingly different fuzzy relations are obtained.

[20] In addition to max operator, there are also other conjunction operations for the "fuzzy or" which then lead to other fuzzy relations.

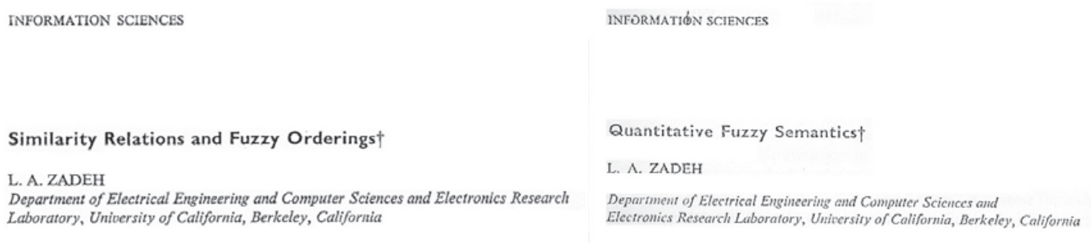
[21] The max-min composition rule is replaced in infinite sets with the sup-min composition rule. However, it is adequate to assume here that all of the sets are finite.

As a generalization of the concept of the equivalence relation Zadeh defined the concept of "similarity", since the similarity relations he defined is reflective, symmetrical and transitive, i.e. for $x, y \in X$ the membership function of S has the following properties:

- Reflexivity: μ_S(x,x) $= 1$;
- Symmetry: μ_S(x,y) $= \mu_S$(y,x);
- Transitivity: $\mu_S(x, y) \geq max_{y \in Y} min\{\mu_S(x, y); \mu_S(y, z)\}$.

4 From Fuzzy Logics and Languages to Fuzzy Semantics and back to Fuzzy Logic

Zadeh's occupation with natural and artificial languages gave rise to his studies in humanities and social sciences, especially on semantics. This intensive work let him to the question "Can the fuzziness of meaning be treated quantitatively, at least in principle?" ([60], p. 160). His 1971 article "Quantitative Fuzzy Semantics" [60] starts with a hint to this studies: "Few concepts are as basic to human thinking and yet as elusive of precise definition as the concept of "meaning". Innumerable papers and books in the fields of philosophy, psychology, and linguistics have dealt at length with the question of what is the meaning of "meaning" without coming up with any definitive answers." ([60], p. 159).[22]


INFORMATION SCIENCES

Similarity Relations and Fuzzy Orderings†

L. A. ZADEH
Department of Electrical Engineering and Computer Sciences and Electronics Research Laboratory, University of California, Berkeley, California

INFORMATION SCIENCES

Quantitative Fuzzy Semantics†

L. A. ZADEH
Department of Electrical Engineering and Computer Sciences and Electronics Research Laboratory, University of California, Berkeley, California

Figure 18: Headlines of Zadeh's articles [59] and [60].

In this paper Zadeh started a new field of research "to point to the possibility of treating the fuzziness of meaning in a quantitative way and suggest a basis for what might be called *quantitative fuzzy semantics*" combining his results on Fuzzy languages and Fuzzy relations. In the section "Meaning" of this paper he set up the basics: "Consider two spaces: (a) a universe of discourse, U, and (b) a set of terms, T, which play the roles of names of subsets of U. Let the generic elements of T and U be denoted by x and y, respectively." Then he started to define the meaning $M(x)$ of a term x as a fuzzy subset of U characterized by a membership function $\mu(y|x)$ which is conditioned on x. One of his examples was: "Let U be the universe of objects which we can see. Let T be the set of terms *white, grey, green, blue, yellow, red, black*. Then each of these terms, e.g., *red*, may be regarded as a name for a fuzzy subset of elements of U which are red in color. Thus, the meaning of red, $M(red)$, is a specified fuzzy subset of U."

In the following section of this paper, that is named "Language", Zadeh regarded a language L as a "fuzzy correspondence", more explicit, a *fuzzy binary relation*, from the term set $T = \{x\}$

[22]In a footnote he named the books of the philosophers, linguists or cognitive scientists Samuel Abraham (born 1923) and Ferenc Kiefer (born 1931) [61], Yehoshua Bar Hillel (1915-1975) [62], Max Black (1909-1988) [63], Rudolf Carnap (1891-1970) [64], Noam Chomsky (born 1928) [65], Alan Fodor (born 1935) and Jerrold J. Katz (1932-2002) [66], Leonard Linsky (born 1922) [67], Sir John Lyons (born 1932) [68], Shimon Ullman (born 1948) [69], Willard Van Orman Quine (1908-2000) [70]), Sebastian K. Shaumyan (1916-2007) [71], Zellig Harris (1909-1992) [72].

to the universe of discourse $U = \{y\}$ that is characterized by the membership function $\mu_L : T \times U \to [0,1]$.

If a term x of T is given, then the membership function $\mu_L(x,y)$ defines a set $M(x) in U$ with the following membership function: $\mu_{M(x)}(y) = \mu_L(x,y)$. Zadeh called the fuzzy set $M(x)$ the *meaning* of the term x; x is thus the name of $M(x)$.

With this framework Zadeh continued in his 1972 article [31] that we mentioned already in chapter III. 2 to establish the basic aspects of a theory of fuzzy languages that is "much broader and more general than that of a formal language in its conventional sense." ([31], p. 134) In the following we quote his definitions of the concepts *fuzzy language, structured fuzzy language* and *meaning*:

Definition 1: A fuzzy language L is a quadruple $L = (U, T, E, N)$, in which U is a non-fuzzy *universe of discourse*; T (called the *term set*) is a fuzzy set of terms which serve as names of fuzzy subsets of U; E (called an *embedding set* for T) is a collection of symbols and their combinations from which the terms are drawn, i.e., T is a fuzzy subset of E; and N is a fuzzy relation from E (or more specifically, the support of T) to U which will be referred to as a *naming relation.*[23]

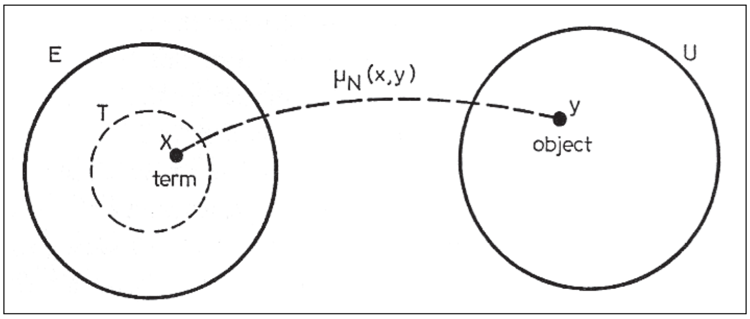


Figure 19: The components of a fuzzy language: U=universe of discourse; T=term set; E=embedding set for T; N=naming relation from E to U; x = *term*; y = *object* in U; $\mu_N(x,y)$=strenght of the relation between x and y; $\mu_T(x)$=grade of membership of x in T.([31], p.136)

In the case that U and T are infinite large sets, there is no table of membership values for $\mu_T(x)$ and $\mu_N(x,y)$ and therefore the values of these membership functions have to be computed. To this end, universe of discourse U and term set T have to be endowed with a structure and therefore Zadeh defined the concept of a structured fuzzy language.

Definition 2: A structured fuzzy language L is a quadruple $L = (U, ST, E, SN)$, in which U is a universe of discourse; E is an embedding set for term set T) S_T is a set of rules, called syntactic rules of L, which collectively provide an algorithm for computing the membership function, μ_T, of the term set T; and S_N is a set of rules, called the semantic rules of L, which collectively provide an algorithm for computing the membership function, μ_N, of the fuzzy naming relation N. The collection of syntactic and semantic rules of L constitute, respectively, the syntax and semantics of L.

To define the concept of meaning, Zadeh characterized the membership function $\mu_N : supp(T) \times U \to [0,1]$ representing the strength of the relation between a term x in T and an object y in U.

However, analogously as in the case of systems and fuzzy systems (see chapter II.2) he clarified: "A language, whether structured or unstructured, will be said to be *fuzzy* if [term set] T or [naming relation] N or both are *fuzzy.* Consequently, an non-fuzzy language is one in which both T and N are non-fuzzy. In particular, a non-fuzzy structured language is a language with both non-fuzzy syntax and non-fuzzy semantics." [31], p. 138).

[23]The support of a fuzzy subset A of X is a non-fuzzy subset $supp(A)$ defined by $supp(A) = \{x | \mu_A(x) > 0\}$.

With these definitions it is clear that natural languages have fuzzy syntax and fuzzy semantics whereas programming languages, as they were usual in the early 1970s, were non-fuzzy structured languages. The membership functions μ_T and μ_N for term set and naming relation, respectively, were two-valued and the compiler used the rules to compute these values 0 or 1. This means that the compiler decides deterministically by using the syntactic rules whether a string x is a term in T or not and it also determines by using the semantic rules whether a term x hits an object y or not. On the other hand we have natural languages, e.g. English, and it is possible that we use sentences that are not completely correct but also not completely incorrect. These sentences have a degree of grammaticality between 0 and 1. Of course, at least native speakers use with high frequency correct sentences. "In most cases, however, the degree of grammatically of a sentence is either zero or one, so that the set of terms in a natural language has a fairly sharply defined boundary between grammatical and ungrammatical sentences", Zadeh wrote ([31], p. 138).[24]

Much more fuzziness we find in semantics of natural languages: Zadeh gave the example "if the universe of discourse is identified with the set of ages from 1 to 100, then the atomic terms *young* and *old* do not correspond to sharply defined subsets of U. The same applies to composite terms such as *not very young*, *not very young and not very old*, etc. In effect, most of the terms in a natural language correspond to fuzzy rather than non-fuzzy subsets of the universe of discourse." ([31], p. 139).

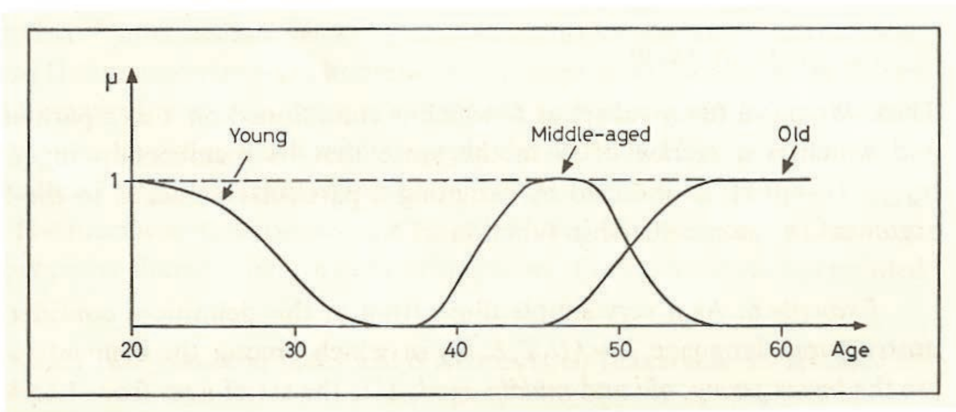


Figure 20: Membership functions of the fuzzy sets $M(young)$, $M(middle - aged)$ and $M(old)$. ([31], p. 140)

Zadeh now identified these fuzzy subsets of the universe of discourse that correspond to terms in natural languages with its "meaning":

Definition 3: The *meaning of a* term x in T is a fuzzy subset $M(x)$ of U in which the grade of membership of an element y of U is given by $\mu_{M(x)}(y) = \mu_N(x, y)$.

Thus, $M(x)$ is a fuzzy subset of U which is conditioned on x as a parameter and which is a section of N in the sense that its membership function, $\mu_{M(x)} : U \to [0, 1]$, is obtained by assigning a particular value, x, to the first argument in the membership function of N.

Zadeh concluded this paper mentioning that "the theory of fuzzy languages is in an embryonic stage" but he expressed his hope that basing on this framework better models for natural

[24]However, Zadeh mentioned that this observation "that natural languages are generally characterized by slightly fuzzy syntax and rather fuzzy semantics does not necessarily hold true when T is associated with an infinite rather than finite alphabet. Thus, when the terms of a language have the form of sounds, pictures, handwritten characters, etc., the fuzziness of its syntax may be quite pronounced. For example, the class of handwritten characters (or sounds) which correspond to a single letter, say R, is rather fuzzy, and this is even more true of concatenation of handwritten characters (or sounds)." [31], p. 139)

languages will be developed than the models of the "restricted framework of the classical theory of formal languages." ([31], p. 163).

Later in the 1970s he published important papers summarizing and developing the concepts we presented above: in 1973 "Outline of a new approach to the analysis of complex systems and decision processes" [73] appeared in the *IEEE Transaction on Systems, Man, and Cybernetics*, in 1975 the three-part article "The concept of a Lingustic Variable and its Application to Approximate Reasoning" [14, 74, 75][25] appeared in the journal *Information Sciences*, in the same year Zadeh published "Fuzzy Logic and Approximate Reasoning" in the philosophical journal Synthese [76] and in 1978 Zadeh published "PRUF - a meaning representation language for natural languages" in the *International Journal of Man-Machine Studies* [77]. [26]


28 IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS, VOL. SMC-3, NO. 1, JANUARY 1973

Outline of a New Approach to the Analysis of Complex Systems and Decision Processes

LOTFI A. ZADEH

L. A. ZADEH*

FUZZY LOGIC AND APPROXIMATE REASONING

(In Memory of Grigore Moisil)

INFORMATION SCIENCES 8, 199-249 (1975)

The Concept of a Linguistic Variable and its Application to Approximate Reasoning–I

L. A. ZADEH

Computer Sciences Division, Department of Electrical Engineering and Computer Sciences, and the Electronics Research Laboratory, University of California, Berkely, California 94720

Int. J. Man-Machine Studies (1978) 10, 395–460

PRUF—a meaning representation language for natural languages†

L. A. ZADEH

Computer Science Division, Department of Electrical Engineering and Computer Sciences and the Electronics Research Laboratory, University of California, Berkeley, CA 94720, U.S.A.

Figure 21: From left above to right below: Headlines of Zadeh's articles [14, 73, 76, 77].

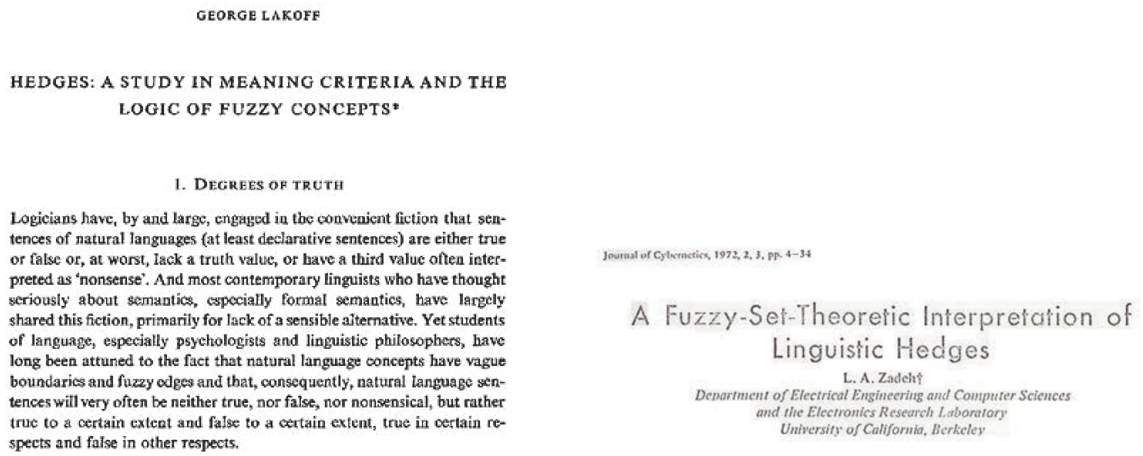
It was in these 1970's when the Berkeley-psychologist Eleonor Rosch developed her prototype theory on the basis of empirical studies. This theory assumes that people perceive objects in the real world by comparing them to prototypes and then ordering them accordingly. In this way, according to Rosch, word meanings are formed from prototypical details and scenes and then incorporated into lexical contexts depending on the context or situation. It could therefore be assumed that different societies process perceptions differently depending on how they go about solving problems [78]. When the linguist George Lakoff heard about Rosch's experiments, he was working at the *Center for Advanced Study in Behavioral Sciences* at Stanford. During a discussion about prototype theory, someone there mentioned Zadeh's name and his idea of linking English words to membership functions and establishing fuzzy categories in this way. Lakoff and Zadeh met in 1971/72 at Stanford to discuss this idea and also the idea of idea of fuzzy logic, after which Lakoff wrote his paper "Hedges: A Study in Meaning Criteria and the Logic of Fuzzy Concepts" [79]. In this work, Lakoff employed "hedges" (meaning barriers) to categorize linguistic expressions and he invented the term "fuzzy logic" whereas Goguen had used "logic of inexact concepts" (see chapter II.1).

Based on his later research, however, Lakoff came to find that fuzzy logic was not an appropriate logic for linguistics: In my interview he said: "It doesn't work for real natural languages, in traditional computer systems it works that way." [80] But: "Inspired and influenced by many discussions with Professor G. Lakoff concerning the meaning of hedges and their interpretation in terms of fuzzy sets," Zadeh had also written an article in 1972 in which he contemplated "linguistic operators", which he called "hedges": "A Fuzzy Set-Theoretic Interpretation of Hedges".

[25] We quoted some paragraphs of the first part of this article [14] already in chapter 2.1

[26] PRUF is an acronym for "Possibilistic Relational Universal Fuzzy".

Here he wrote: "A basic idea suggested in this paper in that a linguistic hedge such as *very, more, more or less, much, essentially, slightly* etc. may be viewed as an operator which acts on the fuzzy set representing the meaning of its operand" [81].


GEORGE LAKOFF

HEDGES: A STUDY IN MEANING CRITERIA AND THE LOGIC OF FUZZY CONCEPTS*

I. DEGREES OF TRUTH

Logicians have, by and large, engaged in the convenient fiction that sentences of natural languages (at least declarative sentences) are either true or false or, at worst, lack a truth value, or have a third value often interpreted as 'nonsense'. And most contemporary linguists who have thought seriously about semantics, especially formal semantics, have largely shared this fiction, primarily for lack of a sensible alternative. Yet students of language, especially psychologists and linguistic philosophers, have long been attuned to the fact that natural language concepts have vague boundaries and fuzzy edges and that, consequently, natural language sentences will very often be neither true, nor false, nor nonsensical, but rather true to a certain extent and false to a certain extent, true in certain respects and false in other respects.

Journal of Cybernetics, 1972, 2, 3, pp. 4–34

A Fuzzy-Set-Theoretic Interpretation of Linguistic Hedges

L. A. Zadeh†

Department of Electrical Engineering and Computer Sciences and the Electronics Research Laboratory University of California, Berkeley

Figure 22: Headlines of Lakoff's article [79] and Zadeh's articles [81].

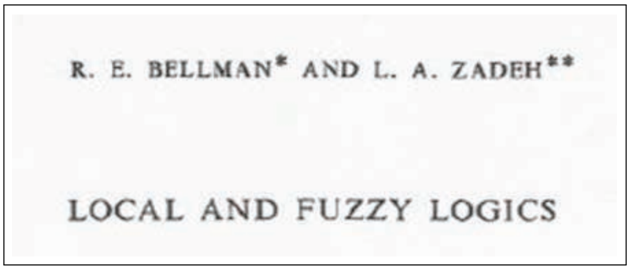
5 The "Fuzzy Boom" with Fuzzy Control – and thereafter?

Richard Bellman left the RAND Corporation in 1965 and was appointed Professor of Mathematics, Electrical Engineering and Medicine at the University of California, Los Angeles (UCLA).[27] This unusual title alone illustrates Bellman's multifaceted gifts and interests. His close friendship with Zadeh resulted in a number of co-authored publications. Bellman had been invited to deliver a lecture at the *International Symposium on Multiple-Valued Logic* at Indiana University in Bloomington, Indiana in May of 1975. He spoke on the subject of "Local Logics". The symposium proceedings includes only a 27-page abstract and a note indicating that the print version would be published in the book *Modern Uses in Multiple-Valued Logic* [82]. This 1977 tome does include the 62-page-long paper "Local and Fuzzy Logics" by Bellman and Zadeh ([83], see Fig. 25), in which the concept of fuzzy sets is carried over to fuzzy logic. Here the authors postulate the following properties of fuzzy logic that we can now trace back to our topics in the last section:

- Truth values here are fuzzy sets of the unit interval that has denominations like "true", "very true", "not very true", "false", "more true" or "less true", etc.
- These truth values are generated by a grammar and they can be interpreted by means of semantic rules.
- Fuzzy logic is local, i.e. both the truth values and their conjunctions such as "AND", "OR" and "IF-THEN" have variable rather than fixed meanings.
- The interference rules of fuzzy logic are not exact but rather approximative.

In the 1970s Lotfi Zadeh expected that Fuzzy Sets, Fuzzy Systems, Fuzzy Logic, Approximate Reasoning, Fuzzy Algorithms, for short his Linguistic Approach "provides an approximate and

[27] For more biographical information on Bellman, see the IEEE History Center website: http://www.ieee.org/organizations/history/center/



R. E. BELLMAN* AND L. A. ZADEH**

LOCAL AND FUZZY LOGICS

Figure 23: Headline of Bellman's and Zadeh's article "Local and Fuzzy Logics", [83].

yet effective means of describing the behavior of systems which are too complex or too ill-defined to admit of precise mathematical analysis." ([73], p. 28) He expected that "even at its present stage of development" his new fuzzy method "can be applied rather effectively to the formulation and approximate solution of a wide variety of practical problems, particularly in such fields as economics, management science, psychology, linguistics, taxonomy, artificial intelligence, information retrieval, medicine and biology. This is particularly true of those problem areas in these fields in which fuzzy algorithms can be drawn upon to provide a means of description of ill-defined concepts, relations, and decision rules." ([73], p. 44).

However, it was in the field of artificial intelligence and first of all it was the concept of Fuzzy Algorithms that fall on fertile ground: Ebrahim Mamdani (1942-2010, see Fig. 15), an electrical engineer in London[28] had read Zadeh's "Outline of a New Approach to the Analysis of Complex Systems and Decision Processes" [73] shortly after it was published and he suggested to his doctoral student Sedrak Assilian that he devise a fuzzy algorithm to control a small model steam engine, as he mentioned in an interview that he gave me in 1998 [84] and he also pointed to Zadeh's 1973-paper in the article that he published together with Assilian after he had finished his PhD thesis: "The true antecedent of the work described here is an outstanding paper by Zadeh (1973) which lays the foundations of what we have termed linguistic synthesis ... and which had also been described by Zadeh as Approximate Reasoning (AR). In the 1973 paper Zadeh shows how vague logical statements can be used to derive inferences (also vague) from vague data. The paper suggests that this method is useful in the treatment of complex humanistic systems. However, it was realized that this method could equally be applied to "hard" systems such as industrial plant controllers. [85], p. 325."

In these times "human control experts" had to provide and understand the control commands and freshmen in this field had to learn these commands. Usually, an expert observed the sequence of processes and knew based on experience how he should intervene if necessary. If any rules governed how he should proceed, they would include linguistically vague expressions, since he would use words like "much", "little", "some", "very" and so forth. Words such as this have been identified in Zadeh's 1973-paper [73] as "linguistic terms" or "modifiers" of "linguistic variables".

Under Mamdani's supervision Assilian realized a fuzzy system under laboratory conditions, and the two designed a fuzzy algorithm to control a small steam engine by a fuzzy rule base system in a few days [86, 87]. They controlled the system with input variables *heat* and *throttle* and output variables *pressure* and *speed* (Fig. 24) by a fuzzy rule base system.

[28] Ebrahim H. Mamdani was born in Tanzania in 1942. He studied electrical engineering at the College of Engineering in Poona, in British India for his first degree. He went to England in 1966 and he joined an MSc course at Queen Mary College, University of London for PhD studies.In his Ph D dissertation (1971) he studied feedback in Neutral Networks. He became professor of electrical engineering at the Queen Mary College and at Westfield College of London University and since 1995, he has been a professor of electrical engeneering at the Imperial College of Science, Technology and Medicine at the University of London. Ebrahim Mamdani passed away at January 22, 2010

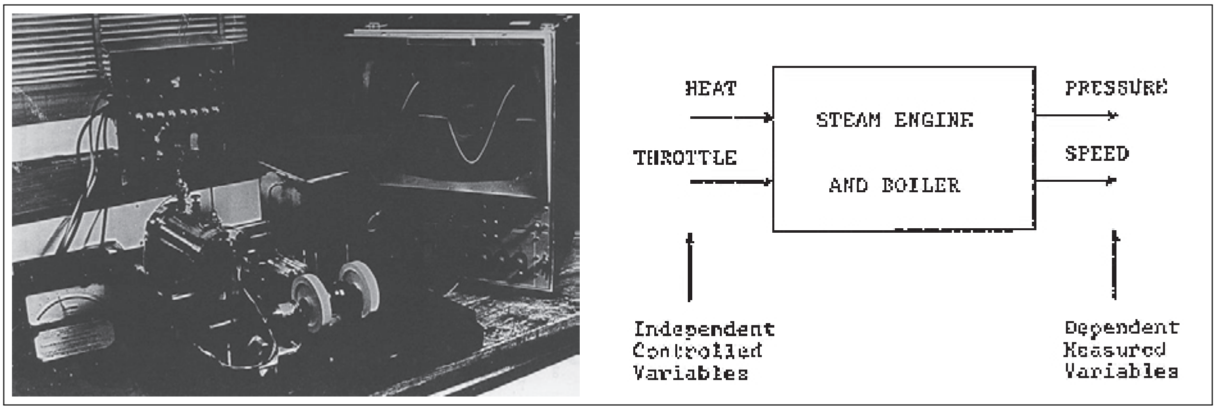


Figure 24: Photograph and schema of the real system of a steam engine [86].

Assilian published his Ph. D. thesis on this first fuzzy control system in 1974 [86]. This steam engine heralded the Fuzzy boom that started in the 1980s in Japan and later pervaded the Western hemisphere. Many fuzzy applications, such as domestic appliances, cameras and other devices appeared in the last two decades of the 20th century.[29]

6 Outlook

In the early years of the research area of Artificial Intelligence (AI) the methods of AI were methods to compute with numbers and to find exact solutions. However, not all problems can be resolved with these methods. On the other hand, humans are able to resolve such tasks very well, as Zadeh mentioned in many speeches and articles over these decades. In conclusion, he stated that "thinking machines" do not think as humans do and from the mid-1980s he focused on "Making Computers Think like People" [16]. For this purpose, the machine's ability "to compute with numbers" was supplemented by an additional ability that was similar to human thinking.

In 1990 he began to formulate a new scientific concept when he wrote that "what might be referred to as soft computing - and, in particular, fuzzy logic - to mimic the ability of the human mind to effectively employ modes of reasoning that are approximate rather than exact. In traditional - hard - computing, the prime desiderata are precision, certainty, and rigor. By contrast, the point of departure in soft computing is the thesis that precision and certainty carry a cost and that computation, reasoning, and decision making should exploit - wherever possible - the tolerance for imprecision and uncertainty. [...] Somewhat later, neural network techniques combined with fuzzy logic began to be employed in a wide variety of consumer products, endowing such products with the capability to adapt and learn from experience. Such neurofuzzy products are likely to become ubiquitous in the years ahead. The same is likely to happen in the realms of robotics, industrial systems, and process control. It is from this perspective that the year 1990 may be viewed as a turning point in the evolution of high MIQ-products and systems. Underlying this evolution was an acceleration in the employment of soft computing - and especially fuzzy logic - in the conception and design of intelligent systems that can exploit the tolerance for imprecision and uncertainty, learn from experience, and adapt to changes in the operation conditions." [16].

At the end of the 20th century Zadeh came back to his early intention to use Fuzzy Sets and Systems and Fuzzy Logic in non-technical areas when he established the method of "Computing with Words" (CW). In 1996 he had published the article "Fuzzy Logic = Computing with Words"

[29]For more details on the history of Fuzzy Control and particularly the Fuzzy Steam Engine see [88] and chapter VI in the author's book [6].

[69] where he proposed CW based on the theories of FSS and these methodologies instead of exact Computing with numbers (CN). Here he wrote that "the main contribution of fuzzy logic is a methodology for computing with words. No other methodology serves this purpose" ([89], p. 103.) and for the new millennium he published his proposal "A New Direction in AI. Toward a Computational Theory of Perceptions" [90]. Once again he clarified that this "Computational theory of perceptions" (CTP) was inspired by the remarkable human capability to operate on, and reason with, perception-based information and he assumed "that progress has been, and continues to be, slow in those areas where a methodology is needed in which the objects of computation are perceptions perceptions of time, distance, form, and other attributes of physical and mental objects." ([90], p. 73).

Since that time, many scientists work hard to contribute with mathematical and logical thinking to establish theories in the areas of CW and CTP. This research lacks the contribution from humanities and social sciences. CW and CTP cannot arise without the fundamentals in these fields and on the other hand: they will lead to new developments in the humanities, such as in linguistics, philosophy or economics.

Acknowledgments
I would like to thank Lotfi A. Zadeh for his generous support to my work on the history of fuzzy sets and systems and its applications for more than 10 years. Especially, I would like to thank him for his offer to let me browse in his private archives in his office in Soda Hall, UC Berkeley and at his home to search and find many important and interesting materials for my historical research. Work leading to this paper was partially supported by the *Foundation for the Advancement of Soft Computing* Mieres, Asturias (Spain).

I would like to thank to this foundation as well as to the Scientific Committee of the *European Centre for Soft Computing* in Mieres, Asturias, Spain. I also thank Enric Trillas and Claudio Moraga very much for important help to write this article.

Bibliography

[1] L. A. Zadeh: *Fuzzy Sets, Information and Control*, 8, 1965, pp. 338-353.

[2] L. A. Zadeh, Fuzzy Sets and Systems. In: Fox, J. (ed.): *System Theory*, New York: Polytechnic Press 1965, pp. 29-37.

[3] L. A. Zadeh, *Toward a Theory of Fuzzy Systems*, Electronic Research Laboratory, University of California, Berkeley 94720, Report ERL-69-2, 6, 1969.

[4] L. A. Zadeh, Towards a theory of fuzzy systems. In: R.E. Kalman, N. DeClaris, (eds.): *Aspects of Network and System Theory*, New York: Holt, Rinehart and Winston, 1971, pp. 469-490.

[5] L. A. Zadeh, *Interview with Lotfi Zadeh, Creator of Fuzzy Logic* by Betty Blair, Azerbaijada International, winter 1994 (2.4): http://www.azer.com /aiweb/categories/-magazine/24_folder/24_articles/24_fuzzylogic.html.

[6] R. Seising, The Fuzzification of Systems. The Genesis of Fuzzy Set Theory and Its Initial Applications – Developments up to the 1970s, *Studies in Fuzziness and Soft Computing* Vol. 216, Berlin, New York, [et al.]: Springer 2007.

[7] J. von Neumann, First Draft of a Report on the EDVAC: http://www.alt.ldv.ei.tum.de/lehre/pent /skript/VonNeumann.pdf.

[8] W. S. McCulloch and W. H. Pitts, A Logical Calculus of the Ideas Immanent in Nervous Activity, *Bulletin of Mathematical Biophysics*, vol. 5, pp. 115-133, 1943.

[9] A. M. Turing, Computing Machinery and Intelligence. In: *Mind*, vol. LIX, 1950, pp. 433-460.

[10] C. E. Shannon, A Mathematical Theory of Communication, *The Bell System Technical Journal*, 27, pp. 379-423 & 623-656, 1948.

[11] N. Wiener, *Cybernetics or Control and Communications in the Animal and the Machine.* Cambridge, Massachusetts, MIT Press, 1948.

[12] E. C. Berkeley, *Giant Brains or Machines that Think*, New York, John Wiley & Sons, 1949.

[13] L. A. Zadeh, Thinking machines - a new field in electrical engineering, *Columbia Engineering Quarterly*, January 1950, pp. 12-13, 30-31.

[14] L. A. Zadeh, The Concept of a Linguistic Variable and its Application to Approximate Reasoning I, *Information Science*, 8, 1975, S. 199-249: 200.

[15] L. A. Zadeh, Outline of a new approach to the analysis of complex systems and decision processes, *IEEE Trans. On Systems, Man, and Cybernetics*, SMC-3 (1), January, 1973, pp. 28-44.

[16] L. A. Zadeh, Making Computers Think like People, *IEEE Spectrum*, 8, 1984, pp. 26-32.

[17] L. A. Zadeh, The Birth and Evolution of Fuzzy Logic - A Personal Perspective, *Journal of Japan Society for Fuzzy Theory and Systems*, vol. 11, No. 6, 1999, pp. 891-905.

[18] L. A. Zadeh, System Theory, *Columbia Engineering Quarterly*, Nov. 1954, pp. 16-19, and 34.

[19] L. A. Zadeh, From Circuit Theory to System Theory, *Proceedings of the IRE*, Vol. 50, 1962, pp. 856-865.

[20] L. A. Zadeh, Desoer, Ch. A., *Linear System Theory: The State Space Approach.* New York; San Francisco, Toronto, London: McGraw-Hill Book Company 1963.

[21] L. A. Zadeh, Polak, E., *System Theory*, Bombay and New Delhi: McGraw-Hill 1969.

[22] L. A. Zadeh, *The Concept of State in System Theory*, [20], pp. 9-42.

[23] Mesarovic, M. D., Views on General Systems Theory, *Proceedings of the Second Systems Symposium at Case Institute of Technology*, Huntington, New York, Robert E. Krieger Publ. Comp. 1964.

[24] L. A. Zadeh, *The Concept of State in System Theory.* In: [23], pp. 39-50.

[25] L. A. Zadeh, Fuzzy Systems Theory: A Framework for the Analysis of Humanistic Systems. In: Roger E. Cavallo (ed.): *Systems Methodology in Social Science Research: Recent Developments*, Kluwer Nijhoff Publishing, Boston and The Hague. 1982 pp. 25-41.

[26] Bellman, R. E., Kalaba, R., Zadeh, L. A., *Abstraction and Pattern Classification.* Memorandum RM-4307-PR, Santa Monica, California: The RAND Corporation, October 1964.

[27] Zadeh, L. A., My Life and Work - a Retrospective View. *Applied and Computatonal Mathematics*, Vol. 10, Nr. 1, Special Issue on Fuzzy Set Theory and Applications, dedicated to the 90th birthday of Prof. Lotfi A. Zadeh, 201, pp. 4-9:

[28] R. E. Bellman, R. Kalaba, L. A. Zadeh, Abstraction and Pattern Classification, *Journal of Mathematical Analysis and Applications*, 13, 1966, S. 1-7.

[29] Zadeh, L. A., Fuzzy Sets, *Information and Control*, 8, 1965, pp. 338-353.

[30] Zadeh, L. A., Fuzzy Sets and Systems. In: Fox, Jerome (ed.): *System Theory.* Microwave Research Institute Symposia Series XV, Brooklyn, New York: Polytechnic Press 1965, p. 29-37.

[31] Zadeh, L. A., Fuzzy Languages and their Relation to Human and Machine Intelligence. In: Man and Computer. *Proc. Int. Conf. Bordeaux* 1970, Karger: Basel, München, New York, 1970, pp. 130-165.

[32] L. A. Zadeh, Fuzzy Logic, Neural Networks, and Soft Computing, *Communications of the ACM*, vol. 37, no. 3, 1994, pp. 77-84.

[33] J. Goguen, *Categories of Fuzzy Sets: Applications of a Non-Cantorian Set Theory.* Ph.D. Thesis. University of California at Berkeley, June 1968.

[34] J. Goguen, L-Fuzzy Sets, *Journal of Mathematical Analysis and Applications*, 18, pp. 145-174, 1967.

[35] J. Goguen, The Logic of Inexact Concepts, *Synthese*, 19, pp. 325-373, 1969.

[36] L. A. Zadeh, "Biological Application of the Theory of Fuzzy Sets and Systems". In: Proctor, LD. (ed.): *The Proceedings of an International Symposium on Biocybernetics of the Central Nervous System.* London: Little, Brown and Comp., 1969, pp. 199-206.

[37] L. A. Zadeh, *Toward a Theory of Fuzzy Systems*, Electronic Research Laboratory, University of California, Berkeley 94720, Report No. ERL-69-2, June 1969.

[38] L. A. Zadeh, Towards a theory of fuzzy systems. In: R.E. Kalman, N. DeClaris, Eds., *Aspects of Network and System Theory*, New York: Holt, Rinehart and Winston, 1971, pp. 469-490.

[39] L. A. Zadeh, The Concept of a Linguistic Variable and its Application to Approximate Reasoning – I, *Information Science* 8, 1975, S. 199-249.

[40] N. Rashevsky: The logical mechanisms of logical thinking. Bulletin of Mathematical Biophysics, vol. 8, 1946.

[41] W. R. Ashby, *Design for Brain.* New York: Wiley, 1960.

[42] J. Pfeiffer, *Thinking Machine*, Philadelphia: Lippincott, 1962.

[43] M. Minsky, Matter, mind and models. *Proceedings of the IFIP Congress*, 1965, pp. 45-49.

[44] H. L. Dreyfus: Alchemy and artificial intelligence. Rand paper P-3244, Santa Monica: Rand-Corporation, 1965.

[45] Papert, S., The artificial intelligence of H. L. Dreyfus. A budget of fallacies, Memorandum 154, Project MAC, (Cambridge: MIT Press, 1968.

[46] H. L. Dreyfus, *A critique of artificial reasoning.* London: Harper & Row, 1971.

[47] R. E. Bellman: Man, society and the computer, Conference Man and Computer, Bordeaux: Institute de la vie, 1970.

[48] S. K. Chang, On the Execution of Fuzzy Programs Using Finite-State Machines, *IEEE Transactions on Computers*, March 1972 (vol. 21 no. 3), pp. 241-253. Or: "Fuzzy Programs - Theory and Applications", Proceedings of the Symposium on Computers and Automata, Polytechnic Institute of Brooklyn, Vol. 21, April 1971, 147-164.

[49] R. S. Interview with L. A. Zadeh on July 5, 2002, University of California, Berkeley, Soda Hall.

[50] William Go Wee and King Sun Fu, A Formulation of Fuzzy Automata and its Application as a Model of Learning Systems. In: *IEEE Transactions on Systems Science and Cybernetics.* Vol. SSC-5, No. 3, pp. 215-223, 1969.

[51] William Go Wee, *On a Generalization of Adaptive Algorithms and Applications of the Fuzzy Set Concept to Pattern Classification.* Ph.D. Thesis, Purdue University, Technical Report, 67, 7, 1967.

[52] Chin-Liang Chang, *Fuzzy Sets and Pattern Recognition.* Ph.D. Thesis. Department of Electrical Engineering, University of California, Berkeley, December 1967.

[53] L. A. Zadeh, Fuzzy Algorithms. In: *Information and Control*, vol. 12, pp. 99-102, February 1968.

[54] R. Seising: Interview with L. A. Zadeh (19.6. 2001), University of California, Berkeley, Soda Hall (unpublished).

[55] L. A. Zadeh, Toward Fuzziness in Computer Systems. *Fuzzy Algorithms and Languages.* (Script, without date and publisher)

[56] P. Turakainen, On Stochastic Languages, *Information and Control*, vol. 12, April 1968, pp. 304-313.

[57] E T. Lee and L. A. Zadeh, Note on Fuzzy Languages, *Information Sciences*, vol. 1, 1969, pp. 421-434.

[58] J. E. Hopcroft and J. D. Ullman, Formal Languages and their Relation to Automata. Reading, Mass.: Addison Wesley, 1969.

[59] L. A. Zadeh, Similarity Relations and Fuzzy Orderings. In: *Information Sciences*, vol. 3, pp. 177-200, 1971.

[60] L. A. Zadeh: Quantitative Fuzzy Semantics. In: Information Sciences, vol. 3, pp. 159-176, 1971.

[61] S. Abraham and S. Kiefer, A Theory of Structural Semantics, The Hague: Mouton, 1965.

[62] Y. Bar-Hillel, Language and Information, Reading, Mass.: Addison-Wesley, 1964.

[63] M. Black, *The Labyrinth of Language*, New York: Mentor Books, 1968.

[64] R. Carnap, *Meaning and Necessity*, Chicago: University of Chicago Press, 1956.

[65] N. Chomsky, *Cartesian Linguistics*, New ok: Harper & Row, 1966.

[66] J. A. Fodor and J. J. Katz, *The Structure of Language*, Englewood Cliffs, N. J.: Prentice-Hall, 1964.

[67] L. Linsky, *Semantics and the Philosophy of Language*, Urbana, Ill.: University of Illinois Press, 1952.

[68] J. Lyons, *Introduction to Theoretical Linguistics*, Cambridge: Cambridge University Press, 1968.

[69] S. Ullmann, *Semantics: An Introduction to the Science of Meaning*. Oxford: Blackwell, 1962.

[70] W. Van Orman Quine, *Word and Object*, Cambridge: MIT Press, 1960.

[71] S.K. Shaumyan, *Structural Linguistics*, Moscow: Nauka, 1967.

[72] Z. Harris, *Mathematical Structures of Language*, New York: Interscience, 1968.

[73] L. A. Zadeh, Outline of a new approach to the analysis of complex systems and decision processes, *IEEE Trans. On Systems, Man, and Cybernetics*, SMC-3 (1), January, 1973, pp. 28-44.

[74] L. A. Zadeh, The Concept of a Linguistic Variable and its Application to Approximate Reasoning II, *Information Science*, 8, 1975, S. 301-357.

[75] L. A. Zadeh, The Concept of a Linguistic Variable and its Application to Approximate Reasoning III, *Information Science*, 9, 1975, S. 43-80.

[76] L. A. Zadeh, Fuzzy Logic and Approximate Reasoning, *Synthese*, vol. 30, 1975, pp. 407-428.

[77] L. A. Zadeh, PRUF - a meaning representation language for natural languages, *International Journal of Man-Machine Studies*, vol. 10, 1978, pp. 395-460.

[78] E. Rosch, Natural Categories, *Cognitive Psychology*, 4, pp. 328-350, 1973.

[79] G. Lakoff Hedges, A Study in Meaning Criteria and the Logic of Fuzzy Concepts, *Journal of Philosophical Logic*, 2, pp. 458-508, 1973.

[80] R. Seising, Interview with George Lakoff (6.8. 2002), University of California, Berkeley, (unpublished).

[81] L.A. Zadeh, A Fuzzy-Set-Theoretic Interpretation of Linguistic Hedges, *Journal of Cybernetics*, vol. 2, pp. 4-34, 1972.

[82] R. E. Bellman, Local Logics. In: *IEEE Logic*, p. 175, 1975.

[83] R. E. Bellman, and L. A. Zadeh, Local and Fuzzy Logics. In: M. J. Dunn and G. Epstein (eds.): *Modern Uses of Multiple-Valued Logic*. In: Invited Papers from the Fifth International Symposium on Multiple-Valued Logic. Held at Indiana University, Bloomington, Indiana, May 13-16, 1975. Dordrecht, Boston: D. Reidel, 1977, pp. 105-165, 1977.

[84] R. S. Interview with E. Mamdani (9.9, 1998), Aachen, RWTH Aachen, at the margin of the European Congress on Intelligent Techniques and Soft Computing EUFIT 1998.

[85] E. H. Mamdani, Advances in the Linguistic Synthesis of Fuzzy Controllers, *International Journal of Man-Machine Studies*, vol. 8, pp. 669-678, 1976.

[86] Assilian, A., *Artificial Intelligence in the Control of Real Dynamic Systems*. Ph. D. Thesis DX193553, University London, August 1974.

[87] E. H. Mamdani and S. Assilian, An experiment in linguistic synthesis with a fuzzy logic controller, *International Journal of Man-Machine Studies*, 7 (1), 1975, pp. 1-13.

[88] R. Seising, The Experimenter and the Theoretician – Linguistic Synthesis to tell Machines what to do. In: Enric Trillas, Piero Bonissone, Luis Magdalena, Janusz Kacprycz (eds.): *Experimentation and Theory – Homage to Abe Mamdani*, Physica-Verlag Series 'Studies in Fuzziness and Soft Computing', 2011 (to appear).

[89] L.A. Zadeh, Fuzzy Logic = Computing with Words, *IEEE Transactions on Fuzzy Systems*, vol. 4, no. 2, 1996, pp. 103-111.

[90] L.A. Zadeh, A New Direction in AI. Toward a Computational Theory of Perceptions, *AI-Magazine*, vol. 22, no. 1, 2001, pp. 73-84.

Int. J. of Computers, Communications & Control, ISSN 1841-9836, E-ISSN 1841-9844
Vol. VI (2011), No. 3 (September), pp. 562-576

Axiomatic Theory of Complex Fuzzy Logic and Complex Fuzzy Classes

D.E. Tamir, A. Kandel

Dan E. Tamir
Texas State University, CS Department
601 University Drive San Marcos, Texas 78666
E-mail: dt19@txstate.edu

Abraham Kandel
University of South Florida, CSE Department
4202 E. Fowler Ave Tampa, Florida 33620
E-mail: kandel@cse.usf.edu

Abstract: Complex fuzzy sets, classes, and logic have an important role in applications, such as prediction of periodic events and advanced control systems, where several fuzzy variables interact with each other in a multifaceted way that cannot be represented effectively via simple fuzzy operations such as union, intersection, complement, negation, conjunction and disjunction. The initial formulation of these terms stems from the definition of complex fuzzy grade of membership. The problem, however, with these definitions are twofold: 1) the complex fuzzy membership is limited to polar representation with only one fuzzy component. 2) The definition is based on grade of membership and is lacking the rigor of axiomatic formulation.
A new interpretation of complex fuzzy membership enables polar and Cartesian representation of the membership function where the two function components carry uncertain information. Moreover, the new interpretation is used to define complex fuzzy classes and develop an axiomatic based theory of complex propositional fuzzy logic. Additionally, the generalization of the theory to multidimensional fuzzy grades of membership has been demonstrated.
In this paper we propose an axiomatic framework for first order predicate complex fuzzy logic and use this framework for axiomatic definition of complex fuzzy classes. We use these rigorous definitions to exemplify inference in complex economic systems. The new framework overcomes the main limitations of current theory and provides several advantages. First, the derivation of the new theory is based on axiomatic approach and does not assume the existence of complex fuzzy sets or complex fuzzy classes. Second, the new form significantly improves the expressive power and inference capability of complex fuzzy logic and class theory. The paper surveys the current state of complex fuzzy sets, complex fuzzy classes, and complex fuzzy logic; and provides an axiomatic basis for first order predicate complex fuzzy logic and complex class theory.
Keywords: Fuzzy Logic, Fuzzy Class Theory, Complex Fuzzy Logic, Complex fuzzy Class theory.

1 Introduction

In 1965, L. A. Zadeh has established the theory of fuzzy sets where the degree of membership of an item in a set can get any value in the interval [0, 1] rather than the two values $\{\notin, \in\}$ [1,2].

Copyright © 2006-2011 by CCC Publications

In addition he introduced the notion of fuzzy logic [1]- [4]. Which is a multilevel extension of classical logic where propositions can get truth-values in the interval [0, 1], and are not limited to one of the two values {True, False} (or {0, 1}) [3]. The four decades that followed his pioneering work has shown a multitude of research work and applications related to signal processing [5], knowledge representation [6], control theory [4], reasoning [7]- [9], and data mining [10]. In 1975 Zadeh introduced the concept of linguistic variable and the induced concept of type-2 (type-n) fuzzy sets [11, 12]. In recent years, type-1 and type-2 along with interval type1/type-2 fuzzy logic and fuzzy systems have been applied in many areas including signal processing [12], fuzzy clustering [13], data mining [10], and software testing [14].

Complex fuzzy sets, classes, and logic have an important role in applications, such as prediction of periodic events and advanced control systems, where several fuzzy variables interact with each other in a complex way that cannot be represented effectively via simple fuzzy operations such as union, intersection, complement, negation, conjunction and disjunction. The initial formulation of these terms stems from the definition of complex fuzzy grade of membership.

Ramot et al. propose an extension of fuzzy set theory and fuzzy logic where the range of degrees of membership and the range of truth-values is the complex unit circle [15, 16]. These definitions, however, have several constraints. First, the derivation relies on a complex fuzzy membership which is limited to polar representation and carries fuzzy information only in the magnitude component of the function. Second, they do not provide an axiomatic approach for their theory. Finally, the derivation of complex fuzzy logic assumes complex fuzzy relations thereby presumes the existence of complex fuzzy sets. Dick has extended the work by Ramot et al., [17], yet his approach is also limited to polar representation with single fuzzy component.

Tamir et al., provide further generalization of the concept of complex fuzzy membership function and use a Cartesian complex fuzzy membership function where both the real part and the imaginary part can be fuzzy functions. Alternatively, polar representation where both the magnitude and the phase values of the complex membership function convey fuzzy information, can be utilized [18]. Furthermore, they provide a new interpretation of complex fuzzy grades of membership as a representation of a complex fuzzy class along with complex fuzzy class operations. While this formulation has several advantages over previous formalisms, it is still based on the definition of grade of membership which is a limiting factor in the ability to provide a rigorous, axiomatic based, theory. In ref. [19], Tamir et al., develop an axiomatic based propositional complex fuzzy logic theory which is independent of complex fuzzy sets, classes and relations. In addition, they demonstrate the potential use of this formalism for inference in complex systems.

In this paper we expand the definition of complex fuzzy logic to include axiomatic based definition of first order predicate complex fuzzy logic and use the new formalism to establish an axiomatic framework for complex fuzzy classes. The new theory is compatible with classical logic, as well as with first order predicate fuzzy logic [20]- [24]. Furthermore, the new theory supports Cartesian as well as polar representation of complex logical fuzzy propositions with two components of ambiguous information. Hence, this form significantly improves the expressive power and inference capability of complex fuzzy logic.

The paper reviews the current state of complex fuzzy sets theory, provides a brief overview of complex fuzzy classes, and complex fuzzy logic; and introduces a new and generalized complex fuzzy propositional logic theory. The new formalism can be used in advanced complex fuzzy logic systems and provides ways for extension into multidimensional fuzzy propositional and first order logic. Furthermore, it can be used for inference with type 2 (or higher) fuzzy sets [11, 12].

The rest of the paper is organized in the following way: section 2 introduces current research based on the concepts of complex fuzzy sets [15]- [18]. Section 3 provides the axiomatic definition of first order complex fuzzy predicate logic [19]. Based on the formulation in section 3, section

4 contains an axiomatic theory of complex fuzzy classes while section 5 includes conclusions and directions for further research.

2 Current and Related Research on Complex Fuzzy Sets

Ramot et al., observe that the expressive power of fuzzy set theory and fuzzy logic and the utility of derived applications can be significantly improved via the introduction of complex fuzzy sets [15,16]. Their observation is mainly motivated by fuzzy processes that contain periodical behavior such as the cycles in economic markets. In order to capture these phenomena in reasoning, they introduce a complex grade of membership and derive the definition of complex fuzzy sets. Later, they introduce complex fuzzy logic via relations on complex fuzzy sets. Their formalism, however, is limited due to the fact that they restrict the membership function to representation using polar coordinate system where only the magnitude carries fuzzy information. Motivated by similar considerations, Tamir et al., extends the formalism proposed by Ramot et al., and introduce complex class theory where both component of a complex fuzzy grade of membership carries fuzzy information [18]. This enables reasoning about processes with multi-dimensional components where each component is carrying fuzzy information and the interaction between the components cannot be decomposed and represented via primitive, one dimensional, fuzzy set theory and fuzzy logic operations such as union, intersection, conjunction, and disjunction. The current formulation of complex fuzzy sets, complex fuzzy classes, and propositional fuzzy logic is based on the definition of grade of membership. This section briefly reviews the first two formalisms [15]- [18]. The next section introduces axiomatic fuzzy logic and complex fuzzy logic theory and includes a discussion on propositional complex fuzzy logic [19].

2.1 Membership-Grade based Definition of Complex Fuzzy Sets

A complex fuzzy set S on a universe of discourse U is defined by a complex-valued grade of membership function $\mu_S(x)$; (Ramot et al. [15,16]):

$$\mu_S(x) = r_S(x)e^{j\omega_S(x)} \tag{1}$$

Where $j = \sqrt{-1}$. This definition utilizes polar representation of complex numbers along with conventional fuzzy set definition; where $r_S(x)$, the amplitude part of the grade of membership, is a fuzzy function defined in the interval [0, 1]. And $\omega_S(x)$ is a real number standing for the phase part of the grade of membership. Ramot et al. propose the following operations on complex fuzzy sets:

1) A directional fuzzy complement which induces the following membership function

$$c(\mu_S(x)) = c(r_S(x))e^{j(\omega_s(x)+\theta)} \tag{2}$$

Where C(f(x))=1-f(x).

2) The union function, ∪, of two complex fuzzy sets A and B induces a membership function $\mu_{A\cup B}(x)$ defined to be:

$$\mu_{A\cup B}(x) = [r_A(x) \oplus r_B(x)] \cdot e^{j\omega_{A\cup B}(x)} \tag{3}$$

where ⊕ represents a *t-conorm* function [24]; and $\omega_{A\cup B}(x)$ is a real function.

3) The intersection function ∩, of two complex fuzzy sets A and B induces a membership function $\mu_{A\cap B}(x)$ defined to be:

$$\mu_{A\cap B}(x) = [r_A(x) \oplus r_B(x)] \cdot e^{j\omega_{A\cap B}(x)} \tag{4}$$

where $\oplus$ represents a *t-conorm* function [24]; and $\omega_{A\cap B}(x)$ is a real function.

Dick expands the research on complex fuzzy sets [17], [25]- [30]. In addition, he observes the relations between complex fuzzy sets and complex neural networks where the excitation, the outputs, and weights can obtain complex values [31]- [35]. Buckley introduces the definition of complex fuzzy numbers [36, 37]. Complex fuzzy numbers have been utilized in several numerical applications [38, 39]. Nevertheless, the concept of complex fuzzy numbers is different from the concept of complex fuzzy sets, classes, and logic [15]- [19], it is also different from the concepts presented in the current paper. Buckley is concerned with generalizing number theory, while references [15]- [19] as well as the current paper are concerned with the generalization of fuzzy set theory and fuzzy logic.

2.2 Membership-Grade based Definition of Complex Fuzzy Classes

In the following sections, a class is denoted by an upper case Greek letter, a crisp set is denoted by an upper case Latin letter, and a member of a set is denoted by a lower case Latin letter.

The Cartesian and polar representation of a complex grade of membership, respectively are given by:

$$\mu(V,z) = \mu_r(V) + j\mu_i(z), \mu(V,x) = r(V)e^{j\sigma\phi(z)} \tag{5}$$

where $\mu_r(V), \mu_i(z), r(V)$ and $\phi(z)$, the real and imaginary components as well as the amplitude and phase components of the complex fuzzy grade of membership, are real value functions over the interval [0,1]. The scaling factor σ is in the interval $(0, 2\pi]$. It is used to control the behavior of the phase within the unit circle according to the specific application.

We adopt the definition of fuzzy classes given by Běhounek et al., and define a Complex Fuzzy class Γ to be a fuzzy class of order 2 [44,45]. Formally, let U be a universe of discourse and let 2^U be the power set of U. Let f_1 be a function from 2^U to [0, 1] and let f_2 be a function that maps elements of U to the interval [0, 1]. For $V \in 2^U$ and z $\in$ U define $\mu_\Gamma(V,z)$ to be:

$$\mu_\Gamma(V,z) = \mu_r(V) + j\mu_i(z) = f_1(V) + jf_2(z) \tag{6}$$

Then, $\mu_\Gamma(V,z)$ defines a fuzzy class of order two, where for every $V \in 2^U$, and for every z $\in$ U, $\mu_\Gamma(V,z)$; is the degree of membership of z in V and the degree of membership of V in Γ. Hence, a complex fuzzy class Γ can be represented as the set of ordered triples:

$$\Gamma = \{V, z, \mu_\Gamma(V,z) | V \in 2^U, z \in U\} \tag{7}$$

Depending on the form of $\mu_\Gamma()$ (Cartesian or polar), $\mu_r()$, $\mu_r()$, r(), and $\phi()$ denote the degree of membership of z in V and / or the degree of membership of V in Γ. The role of the real and imaginary parts of the membership function can be interchanged, the same applies to the components of the polar representation. Furthermore, the usual definition of coordinate transformation:

$$\mu_\Gamma(V,z) \Leftrightarrow r(V)e^{j\phi(x)}; \mu_\Gamma(V,z) \Leftrightarrow r(V) + j\phi(x) \tag{8}$$

maintains the semantics of complex fuzzy classes. Thus, without loss of generality we assume that: $\mu_r(V)$ and r(V) denote the degree of membership of V in Γ and that $\mu_i(z)$ and $\phi(z)$ denote the degree of membership of membership of z in V for the Cartesian and the polar representations respectively. One can consider the grade of membership based definition of a fuzzy class of order 1 as a mapping into a one dimensional space and the grade based definition of fuzzy class of order 2 (see [40]) as a mapping into a two dimensional space. Hence, it is possible to consider a degree of membership of order n as a mapping into an n-dimensional space.

Operations on Complex Fuzzy Classes.

Consider the two complex fuzzy classes: 1. $\Gamma = \{V, z, \mu_\Gamma(V, z) | V \in 2^U, z \in U\}$. 2. $\Psi = \{T, z, \mu_\Psi(T, z) | T \in 2^U, z \in U\}$. Where V and T are fuzzy classes. Assume that Γ and Ψ are defined over a universe of discourse U and let 2^U denote the power set of U. Further assume that the degree of membership of an object $z \in V$, and an object $y \in T$ is given by: $\mu_\Gamma(V, z) = \mu_{\Gamma r}(V) + j\mu_{\Gamma i}(z); \mu_\Psi(T, y) = \mu_{\Psi r}(T) + j\mu_{\Psi i}(y)$ respectively, where $\mu_{\Gamma r}(\alpha), \mu_{\Psi r}(\alpha), \mu_{\Gamma i}(\alpha)$ and $\mu_{\Psi i}(\alpha)$ stand for the real and imaginary parts of $\mu_\Gamma(V, x)$ and $\mu_\Psi(T, y)$. Finally, let $W \in 2^U$, let $\oplus$ denote a t-conorm operation, and let $\oplus$ denote a t-norm operation.

1. The complex fuzzy class complement is defined to be:

$$c(\mu_\Gamma(V, z)) = c(\mu_r(V)) + jc(\mu_i(z)) \tag{9}$$

2. The complex fuzzy class union is defined to be:

$$\mu_{\Gamma \cup \Psi}(W, z) = (\mu_{\Gamma r}(V) \oplus \mu_{\Psi r}(T) + j(\mu_{\Gamma i}(z) \oplus \mu_{\Psi i}(z))) \tag{10}$$

3. The complex fuzzy class intersection operation is given by:

$$\mu_{\Gamma \cap \Psi}(W, z) = (\mu_{\Gamma r}(V) \oplus \mu_{\Psi r}(T) + j(\mu_{\Gamma i}(z) \oplus \mu_{\Psi i}(z))) \tag{11}$$

Reference [18] includes several examples of complex fuzzy classes and their applications.

3 Axiomatic Complex Fuzzy Logic

Two main considerations are motivating this section. The first relates to the observations stated in section 2. Namely, there are multidimensional fuzzy processes which cannot be decomposed into simple single dimensional logical operations. Second, many researchers has observed that the grade based membership approach is limited in its capability to deliver a concise and precise formalism for fuzzy logic [20]- [24], [40]. Instead, current research in fuzzy logic, fuzzy class theory, fuzzy mathematics, and its applications is based on axiomatic theory. In ref. [19] we have presented an axiomatic framework for propositional complex fuzzy logic (CFL). In this section we extend the definition to include first order predicates CFL. The presented framework is based on the basic fuzzy propositional and predicate logic (BL) along with the fuzzy Łukasiewicz (Ł) and fuzzy product (Π) logical systems [21, 22, 40]. We refer to the propositional logic system as ŁΠ and to the first order predicate fuzzy logic system as ŁΠ∀. This section introduces the ŁΠ and ŁΠ∀ fuzzy logic system as well as the ŁΠ and ŁΠ∀ CFL. We use the following notation: a variable is denoted by a lower case Latin letter and a predicate/proposition is denoted by an upper case Latin letter.

3.1 Propositional Fuzzy Logic

Several axiom based logical systems have been investigated [20]- [24], [40]. Běhounek, et al., ([40]) use the ŁΠ/ ŁΠ∀ as the basis for the definition of fuzzy class theory (FCT). In this section we closely follow ŁΠ the system used by Běhounek, et al. For clarity, we reintroduce some of the important notions, notations, and concepts from that paper.

A fuzzy proposition P can get any truth-value in the real interval [0,1], where '0' denotes "False," and '1' denotes "True." Furthermore, the relation $\leq$, over the interval [0,1] implies a monotonically increasing ordering on the truth-values associated with the proposition. A fuzzy interpretation of a proposition P is an assignment of a fuzzy truth-value to P. Let P and Q

denote fuzzy propositions and let i(P) denotes the fuzzy interpretation of P. Table 1, includes the basic connectives of ŁΠ. Table 2 includes connectives that can be derived from the basic connectives. The constant 0 is assumed; and the constant 1 can be derived from 0 and the basic connectives.

Operation	Interpretation
Ł-Implication	$i(P \rightarrow_L Q) = \min(1, 1 - i(P) + i(Q))$
Π-Implication	$i(P \rightarrow_\Pi Q) = \min(1, i(P)/i(Q))$
Π-Conjunction	$i(P \otimes Q) = i(P) \cdot i(Q)$

Table 1: Basic ŁΠ Connectives

Operation	Interpretation
Ł-Negation	$i(\neg P) = 1 - i(P)$
Π-Delta	$\Delta(i(P)) = 1$ if $i(P)=1$; else $\Delta(i(P))=0$
Equivalence	$i(P \leftrightarrow Q) = i(P \rightarrow_L Q) \otimes i(Q \rightarrow_L P)$
$P \ominus Q$	$i(P \ominus Q) = \max(0, i(P) - i(Q))$

Table 2: Derived ŁΠ Connectives

Běhounek, et al., use the basic and derived connectives along the truth-constants and the following set of axioms [40]:

1) The Łukasiewicz set of axioms [21].
2) The product set of axioms [21].
3) The Łukasiewicz Delta axiom:

$$\Delta(P \rightarrow_L Q) \rightarrow_L (P \rightarrow_\Pi Q) \tag{12}$$

4) The Product Delta axiom:

$$\Delta(P \rightarrow_\Pi Q) \rightarrow_L (P \rightarrow_L Q) \tag{13}$$

5) The axiom:

$$R \otimes (P \ominus Q) \leftrightarrow_L (R \otimes P) \ominus (R \otimes Q) \tag{14}$$

The rules of inference are:

1) Modus ponens.
2) Product necessitation (infer ΔQ from P).

Reference [40] includes several theorems that follow from the definition of ŁΠ propositional fuzzy logic. In the next section, we define ŁΠ first order predicate fuzzy logic (ŁΠ∀).

3.2 First Order Predicate Fuzzy Logic

Following classical logic, ŁΠ first order predicate fuzzy Logic, referred to as ŁΠ∀, extends ŁΠ propositional fuzzy logic. The primitives include constants, variables, arbitrary arity functions and arbitrary arity predicates. Formulae are constructed using the basic connectives defined in table 1, derived connectives such as the connectives presented in table 2, the truth-constants, the quantifier ∀ and the identity sign =.The quantifier ∃ can be used to abbreviate formulae derived from the basic primitives and connectives. A fuzzy interpretation, of a proposition $P(x_1, ..., x_n)$

over a domain M is a mapping that assigns a fuzzy truth-value to each n-tuple of elements of M. As in the case of ŁΠ, we closely follow the system used in ref. [40].

Assuming that y can be substituted for x in P and x is not free in Q the following axioms are used.

1) Instances of the axioms of ŁΠ obtained through substitution.

2) Universal axiom I:

$$(\forall x)P(x) \rightarrow P(y) \tag{15}$$

3) Universal axiom II:

$$(\forall x)(P \rightarrow_L Q) \rightarrow (P \rightarrow_L (\forall x)Q) \tag{16}$$

4) Identity axiom I:

$$x = x \tag{17}$$

5) Identity axiom II:

$$(x = y) \rightarrow \Delta(P(x) \leftrightarrow P(Y)) \tag{18}$$

Modus ponens, product necessitation, and generalization are used for inference. In the next section, we define propositional and first order predicate CFL.

3.3 Propositional and First Order Predicate Complex Fuzzy Logic

As noted, the main motivation for developing CFL is twofold: First, we have observed that many multidimensional processes cannot be adequately represented as a combination of single dimension processes expressed via single dimension fuzzy logic. Second, we are interested in an axiomatic based formal system that can be used for a rigorous definition of complex fuzzy classes.

A complex fuzzy proposition P is a composition of two propositions each of which can accept a truth-value in the interval [0,1]. In other words; the interpretation of a complex fuzzy proposition is a pair of truth-values from the Cartesian interval [0,1] × [0,1]. Alternatively, the interpretation can be formulated as a mapping to the unit circle. Formally a fuzzy interpretation of a complex fuzzy proposition P is an assignment of fuzzy truth-value of the form $i(p_r) + j \cdot (p_i)$, or of the form $i(r(p))e^{j\sigma i(\theta(p))}$, where σ is a scaling factor in the interval (0,2Π], to P.

For example, consider a proposition of the form "x ... A ... B ...," along the definition of linguistic variables and constants. Namely, a linguistic variable is a variable whose domain of values is formal or natural language words [1]. Generally, a linguistic variable is related to a fuzzy set such as {*very young male, young male, old male, very old male*} and can get any value from the set. A linguistic constant has a fixed and unmodified linguistic value i.e., a single word or phrase from formal or natural language.

Thus, in a proposition of the form "x ... A ... B ...," where A and B are linguistic variables, i(p_r)(i(r(p))) can be assigned to the term A and i(p_i)(i(θ(p))) can be assigned to term B.

Propositional CFL extends the definition of propositional fuzzy logic and first order predicate CFL extends the notion of first order predicate fuzzy logic. Nevertheless, since propositional CFL is a special case of first order predicate CFL we only present the formalism for first order predicates CFL here. Reference [19] presents propositional CFL. Hence, this section augments ref. [19] with first order predicate CFL. It is the first work that provides this formalism.

Tables 3 and 4 present the basic and derived connectives of ŁΠ∀ CFL. In essence, the connectives are symmetric with respect to the real and imaginary parts of the predicates.

Operation	Interpretation
Ł-Implication	i(P$\rightarrow_L Q$) $= \min(1, 1-i(p_r)+i(q_r)) + j\cdot \min(1, 1-i(p_i)+i(q_i))$
Π-Implication	i(P$\rightarrow_\Pi Q$) $= \min(0, i(p_r)/i(q_r)) + j\cdot \min(0, 1(p_i)/i(q_i))$
Π-Conjunction	i$(P\otimes Q) = i(p_r)\cdot i(q_r) + j\cdot (i(p_i)\cdot i(q_i))$

Table 3: Basic ŁΠ∀ CFL Connectives

Operation	Interpretation
Ł-Negation	i(¬P) = 1 + j1 - i(P)
Π-Delta	Δ(i(P)) = 1 if i(P)=1 + j1; else Δ(i(P))=0 + j0
Equivalence	$i(P\leftrightarrow Q) = i(P_r\rightarrow_L Q_r)\otimes i(Q_r\rightarrow_L P_r) + j\cdot (i(P_i\rightarrow_L Q_i)\otimes i(Q_i\rightarrow_L P_i))$
P⊖ Q	i(P ⊖ Q) $= \max(0, 1(p_r)-i(q_r)) + j\cdot \max(0, i(p_i)-i(q_i))$

Table 4: Derived ŁΠ∀ CFL Connectives

Following classical logic, ŁΠ∀ CFL extends, ŁΠ CFL. The primitives include constants, variables, arbitrary arity functions and arbitrary arity predicates. Formulae are constructed using the basic connectives defined in table 3, derived connectives such as the connectives presented in table 4, the truth-constants, the quantifier ∀ and the identity sign =. The quantifier ∃ can be used to abbreviate formulae derived from the basic primitives and connectives. A fuzzy interpretation of a proposition $P(x_1,...x_n) = P_r(x_1,...x_n) + j\cdot P_i(x_1,...x_m)$ over a domain M is a mapping that assigns a fuzzy truth-value to each (n-tuple) × (m-tuple) of elements of M. As in the case of ŁΠ fuzzy logic, we closely follow the system used in ref. [40].

The same axioms used for first order predicate fuzzy logic are used for first order predicate complex fuzzy logic, and Modus ponens as well as product necessitation, and generalization are the rules of inference.

Complex Fuzzy Propositions and Connectives Examples.

One form of a fuzzy proposition is: "x ... A ...;" where A is a linguistic variable such as "young male" or "tall person" and '...' denote natural language constants such as "Moses." "is," "portfolio," "mutual fund," etc. For example, under one interpretation, the fuzzy truth-value associated with the fuzzy proposition: P = "Moses is *young male*" can be 0.3, and under another interpretation, the fuzzy truth-value associated with the proposition P can be 0.9. In this case, the linguistic variable is "young male," and it is distinguished from the Fuzzy constants "Moses" and "is" by its Italics font.

One form of a complex fuzzy is: "x ... A ... B ..." where A and B are values assigned to linguistic variables and '...' denotes natural language constants. For example, under one interpretation, the complex fuzzy truth-value associated with the complex fuzzy proposition: "x is a *volatile stock in a strong-portfolio*," can be 0.1 + j0.5. Alternatively, in another context, the same proposition can be interpreted as having the complex truth-value $0.3e^{j0.2}$. Consider the following propositions(P and Q respectively):

1) "x is a *volatile stock in a strong-portfolio.*"

2) "x is a stock in a *decline-trend in a strong-portfolio*"

Hence, P is of the form: "x is a A in a B" and Q is of the form "x is a stock C in B" In this case, "*volatile stock*," "*a strong-portfolio*," and "*a decline-trend*" are values assigned to the linguistic variables {A,B,C}. Assume that the complex fuzzy interpretation (i.e., degree of confidence or complex fuzzy truth-value) of P is $i(p_r) + j\cdot(p_i)$, while the complex fuzzy interpretation of Q is $i(q_r) + j\cdot(p_i)$. Thus, the truth-value of "*x is a volatile stock*" is $i(p_r)$, the truth-value of "*x is in a strong-portfolio*" is $i(p_i)$, and the truth-value of "*x is in a decline-trend*" is $i(q_r)$ Suppose that the term "*non*-volatile" stands for "*not volatile*" the term "*weak*" stands for "*not strong*," and the term "*rising*" stands for "*not declining*," Then, the complex fuzzy interpretation of the following composite propositions is:

1) i(¬P)= $(1 - i(p_r)) + j \cdot (1 - i(p_i))$.

That is, $\neg P$ denotes the proposition "x is a non-volatile stock in a weak portfolio." The confidence level in $\neg P$ is $(1 - i(p_r)) + j(1 - i(p_i))$; where the fuzzy truth-value of the term "x is a non-volatile stock" is $(1 - i(p_r))$ and the fuzzy truth-value of the term "weak portfolio" is $(1 - i(p_i))$.

2) i(¬P $\rightarrow_L$ ¬Q) $= min(1, 1 + i(p_r) - i(q_r)) + j \cdot min(1, 1 + i(p_i) - i(q_i))$.

Thus, (¬P$\rightarrow_L$ ¬Q) denotes the proposition "If x is a non-volatile stock in a weak portfolio, THEN x is a stock in a rising-trend in a strong-portfolio." The truth-values of individual terms, as well as the truth-value of ¬P$\rightarrow_L$ ¬Q are calculated according to table 3.

3) i(P ∨ Q) $= max(i(p_r) \cdot i(q_r)) + j \cdot max(i(p_i) \cdot i(q_i))$.

That is, (P∨¬Q) denotes the proposition "x is a volatile stock in a strong-portfolio" OR "x is a stock in a rising-trend in a week portfolio." The truth-values of individual terms, as well as the truth-value of P∨¬Q are calculated according to tables 3 and 4.

4) i(¬P ⊗ Q) $= min(1 - i(p_r), i(q_r)) + j \cdot min(1 - i(p_i), i(q_i))$.

That is, (¬P ⊗ Q) denotes the proposition "x is a volatile stock in a strong-portfolio" AND "x is a stock in a rising-trend in a strong-portfolio." The truth-values of individual terms, as well as the truth-value of ¬P ⊗ Q are calculated according to table 3.

Complex Fuzzy Inference Example.

Assume that the degree of confidence in the proposition R = ¬P (P is defined above) is $i(r_r) + j \cdot (r_i)$ let S = ¬Q and assume that the degree of confidence in the fuzzy implication T = R $\rightarrow_L$S is i(t_r) + j · (t_i). Then, using Modus ponens:

R
$\underline{R \rightarrow S}$
S

one can infer S with a degree of confidence min(1,i(r_r) · i(t_r)) + j · min(1,i(r_i) · i(t_i)).
In other words one using:
"x is a *non-volatile stock* in a *weak portfolio*."
IF
"x is a *non-volatile* stock in a weak portfolio," THEN
"x is a stock in a *rising-trend* in a *weak-portfolio*."
"x is a stock in a *rising-trend* in a *weak-portfolio*"
then, one can infer "x is a stock in a *rising-trend* in a *weak-portfolio*" with a degree of confidence min(1,i(r_r) · i(t_r)) + j · min(1,i(r_i) · i(t_i)).

4 Complex Fuzzy Class Theory

Many natural phenomena are complex and cannot be modeled using one dimensional classes and / or one dimensional variables. For example, in pattern recognition, objects can be represented by a set of measurements and are considered as vectors in a multidimensional space. These patterns might be clustered into several clusters where different clusters relate to different subsets of features. Often, it is not practical to assume that this multidimensional information can be represented via a one-dimensional combination of variables and operators. As a relatively simple example consider the stock market; at any given time each stock can be evaluated by its current value and trend. Each of these parameters can be represented by a fuzzy set and inference can be carried by fuzzy logic. Nevertheless, ignoring the current state of the entire market and many other parameters provides a very limited capability to make sound decisions. Moreover, often the relations between the current value of the stock and the current market trends cannot be adequately represented via simple fuzzy logic operations over individual parameters. Other examples for the need in a mechanism that can represent complex objects and complex sets occur in signal processing where many signals are represented using complex variables. In a noisy environment these signals might form a "complex fuzzy class." Finally, consider a set of values where each value is a member of a fuzzy set. This set referred to as fuzzy set of type-2 cannot be represented by basic operations on fuzzy sets of type-1 [11, 12]. This type of sets however, can be represented via complex classes. Motivated by these considerations Ramon et al., and Dick propose complex fuzzy sets. The limitation of their approach is described in section 2. Tamir et al. introduce a new interpretation of complex fuzzy grade of membership and derive the concept of complex fuzzy classes using the framework of a complex fuzzy grade of membership [18]. This form is limited since it cannot be easily axiomatized and rigorously formalized. In this paper we provide an alternative formulation for complex fuzzy classes which is based on first order predicate complex fuzzy logic (ŁΠ∀ CFL). Based on the logic system presented in section 3.1, the next section introduces the axiomatic based formulation of fuzzy class theory (FCT) as developed in ref. [40]; section 4.2 provides the extension of FCT to complex fuzzy class theory (CFCT).

4.1 Axiomatic Based Fuzzy Class Theory

The axiomatic fuzzy logic can serve as a basis for establishing an axiomatic FCT. Similarly, axiomatic based complex fuzzy logic can serve as the basis for formal definition of complex fuzzy classes. Several variants of FCT exists, most of them use a similar approach and mainly differ in the selection of the logic base. Another difference between various approaches is the selection of class theory axioms. Běhounek, et al., present and analyze a few variants of FCT. In ref. [40] they present a ŁΠ∀ based FCT. Their formalism serves as the starting point for the complex FCT presented in this paper. Next, we provide an overview of the ŁΠ∀ based FCT.

The main components of FCT are:

1) Variables.

a. Variables denoting objects;

b. Variables denoting crisp sets, i.e., a universe of discourse and its subsets;

c. Variables denoting fuzzy classes of order 1;

d. Variables denoting fuzzy classes of order n, that is, fuzzy classes of fuzzy classes of order n-1.

We use the following notation for variables: objects are denoted by lower case Latin letters, crisp sets are denoted by upper case Latin letters, fuzzy classes of order n are denoted by upper case Greek letter of the form Γ^n (n is omitted when n=1).

2) The ŁΠ∀ fuzzy logic system along with its variables, connectives, predicates, and axioms defined in section 3.1 and ref. [40].

3) Additional predicates.

a. A binary predicate '$x \in \Gamma$' denoting membership of objects in fuzzy classes.

4) Additional Axioms.

a. Instances of the comprehension schema (further explained below)

$$(\exists\Gamma)\Delta(\forall x)(x \in \Gamma \leftrightarrow p(X)) \tag{19}$$

b. The axiom of extensionality

$$(\forall x)\Delta(x \in \Gamma \leftrightarrow x \in \Psi) \to \Gamma = \Psi \tag{20}$$

Note that a grade of membership is not a part of the above specified terms; yet it can be derived or defined using these terms.

The comprehension schema is used to "construct" classes. It has the basic form of: $(\forall X)(x \in \Gamma \leftrightarrow P(x))$. Intuitively, this schema refers to the class Γ of all the objects x that satisfy the predicate P(). Instances of this scheme have the generic form: $(\exists\Gamma)(\forall\ \mathrm{x})(\mathrm{x} \in \Gamma \leftrightarrow \mathrm{P(x)})$. Associated with this schema are comprehension terms of the form: $\mathrm{y} \in \{x|P(x) \leftrightarrow \mathrm{P(y)}\}$. The Δ operation introduced in equation 20, is used to produce precise instances of the extensionality schema and ensure the conservatism of comprehension terms.

Fixing a standard model over the FCT, enables the definition of commonly used terms, set operations, and definitions, as well as proving FCT theorems [45]. Some of these elements are listed here:

1) The characteristic function $\chi_{x\in\Gamma} \equiv \chi_\Gamma$ and the grade of membership function $\mu_{x\in\Gamma} \equiv \mu_\Gamma$.

2) Class constants, α-cuts, iterated complements, and primitive binary operations such as union and intersection etc. These operations are constructed using the schema $O_p(\Gamma) \equiv \{x|P(x \in \Gamma)\}$. Table 5 lists some of these elements.

3) Uniform and supreme relations defined in ref. [40] enable the definition of fuzzy class relations such as inclusion.

4) Theorems on axioms, primitive fuzzy class operations, and fuzzy class relations [40].

Term	Symbol	P	Comments
Empty Class	Θ	0	
Universal Class	Φ	1	
Strict Complement	$\backslash\Gamma$	$\sim$	$\sim$ stands for Gödel(G) negation
Class Intersection	$\cap$	$\oplus$	$\oplus$ stands for a G, Ł, or Π conjunction T-norm
Class Union	$\cup$	$\vee$	$\vee$ stands for a G, Ł, or Π disjunction

Table 5: Primitive Fuzzy Class Operations

4.2 Axiomatic based Complex Fuzzy Class Theory

The axiomatic fuzzy logic can serve as a basis for establishing an axiomatic FCT. Similarly, axiomatic based complex fuzzy logic can serve as the basis for formal definition of complex fuzzy classes. In this section we provide a formulation of complex fuzzy class theory (CFCT) that is based on the logic theory presented in section 3.2. This is the first work that provides this formalism.

A new concept that is introduced in this section is the concept of complex objects. Intuitively this may sound odd since we are used to think about set objects as one-dimensional or dimensionless. Nevertheless, in reality, most of the objects of interest in many applications and research projects are represented by a set of measurements and are considered as multidimensional. In the context of this paper, a multi-dimensional object is an object that is associated with (or perceived by) more than one measurement or stimuli. A simple example is human color perception which associates the "qualities" of hue and saturation with many colors. A complex object is represented by two measurements or stimuli also referred to as features.

The main components of FCT are:

1) Variables.

a. Variables denoting objects (potentially complex objects);

b. Variables denoting crisp sets, i.e., a universe of discourse and its subsets;

c. Variables denoting complex fuzzy classes of order 1;

d. Variables denoting complex fuzzy classes of order n, that is, complex fuzzy classes of complex fuzzy classes of order n-1.

3) Additional predicates.

a. A binary predicate '$x \in \Gamma$' denoting membership of objects in complex fuzzy classes.

To elaborate, assume that Γ is a complex class, and let x be a complex object of the form $x = x_r + j \cdot x_i$, then the predicate $x \in \Gamma$ yields a complex truth-level of the form $x \in \Gamma = [x_r \in \Gamma + j \cdot x_i \in \Gamma]$. Intuitively this truth-value represents fuzzy validity of the compounded assertion "$[x_r, x_i]$ belongs to Γ," or "$[x_r, x_i]$ is a member of Γ".

4) Additional Axioms.

a. Instances of the comprehension schema (further explained below)

$$(\exists \Gamma)\Delta(\forall x)(x \in \Gamma \leftrightarrow P(x)) \tag{21}$$

Where x is a complex fuzzy object, Γ is a complex fuzzy class, and P() is a complex fuzzy predicate.

b. The axiom of extensionality

$$(\forall x)\Delta(x \in \Gamma \leftrightarrow x \in \Psi) \rightarrow \Gamma = \Psi \tag{22}$$

Again, x is a complex fuzzy object, Γ is a complex fuzzy class, and P() is a complex fuzzy predicate.

Note that a grade of membership is not a part of the above specified terms; yet it can be derived or defined using these terms.

The comprehension schema is used to "construct" classes. It has the basic form of: $(\forall x)(x \in \Gamma \leftrightarrow P(x))$. Intuitively, this schema refers to the class Γ of all the objects x that satisfy the predicate P(). Instances of this scheme have the generic form: $(\exists \Gamma)(\forall x)(x \in \Gamma \leftrightarrow P(x))$. Associated with this schema are comprehension terms of the form: $y \in \{x|P(x) \leftrightarrow P(y)\}$. The Δ operation introduced in equation 22, is used to produce precise instances of the extensionality schema and ensure the conservatism of comprehension terms.

Fixing a standard model over the CFCT enables the definition of commonly used terms, set operations, and definitions, as well as proving CFCT theorems. Some of these elements are listed here:

1) The complex characteristic function $\chi_{x \in \Gamma} \equiv \chi_\Gamma$ and the complex grade of membership function $\mu_{x \in \Gamma} \equiv \mu_\Gamma$.

2) Complex class constants, α-cuts, iterated complements, and primitive binary operations such as union and intersection etc. These operations are constructed using the schema $O_p(\Gamma) \equiv \{x|P(x \in \Gamma)\}$. Table 6 lists some of these elements.

3) Uniform and supreme relations defined in ref. [40] enable the definition of fuzzy class relations such as inclusion.

4) Theorems on axioms, primitive fuzzy class operations, and fuzzy class relations [40].

Following the axiomatic based definition of grade of membership equations (9-11) can be used as a basis for the definition of "membership-grade based" complement, union, and intersection.

We are currently performing research aimed to provide a comprehensive list of theorems, complex class operations, and relations.

Complex Fuzzy Propositions and Connectives Examples.

In order to provide a concrete example we define the following complex fuzzy classes using the comprehension schema. Let the universe of discourse be the set of all the stocks that were

Term	Symbol	P	Comments
Empty Complex Class	Θ	0	
Universal Complex Class	Φ	1	
Strict Complement	$\backslash\Gamma$	$\sim$	$\sim$ stands for Gödel(G) negation
Complex Class Intersection	$\cap$	$\oplus$	$\oplus$ stands for a G, Ł, or Π conjunction T-norm
Complex Class Union	$\cup$	$\vee$	$\vee$ stands for a G, Ł, or Π disjunction

Table 6: Derived Primitive Class Operations

available for trade on the opening of the New York stock exchange (NYSE) market on January 3, 2011 along with a set of attributes related to historical price performance of each of these stocks. Consider the following complex propositions:

1) P(x)≡ "x is a volatile stock in a strong-portfolio,"

2) Q(X)≡ "x is a stock in a decline-trend in a strong-portfolio,"

Then, the proposition: $(\exists\Gamma)\Delta(\forall x)(x \in \Gamma \leftrightarrow (P(x) \otimes Q(x))$ where x is any member of the universe of discourse defines a complex fuzzy class Γ that can be "described" as the class of "*volatile stocks in a decline-trend in strong-portfolios*," On the other hand, the proposition: $(\exists\Gamma)\Delta(\forall x)(x \in \Gamma \leftrightarrow (\neg P(x) \vee Q(x))$ where x is any member of the universe of discourse, defines a complex fuzzy class Γ that can be "described" as the class of "*non-volatile stocks in a decline-trend in strong-portfolios.*"

5 Conclusions and Directions for Further Research

A new and innovative formal definition of complex fuzzy logic (CFL), referred to as ŁΠ∀ CFL as well as a formalism of complex fuzzy class theory (CFCT) that is based on CFL, is presented in this paper. The new form significantly improves the expressive power and inference capability of CFL and CFCT based systems. In addition, it enables axiomatic and rigorous development of the relevant theory. In the future, we plan to extend the theory. In addition, we plan to apply the new formalism to multidimensional fuzzy propositional and predicate logic; and further explore its potential for usage in advanced complex fuzzy logic systems as well as inference with type-2 (or higher) fuzzy sets.

Bibliography

[1] Zadeh, L.A., *The concept of a linguistic variable and its application to approximate reasoning - Part I.* Information Sciences, 1975, 7: p. 199-249.

[2] Klir, G.J., Tina, A., *Fuzzy sets, uncertainty, and information.* 1988, Upper Saddle River, NJ: Prentice Hall.

[3] Kandel, A., *Fuzzy Mathematical Techniques with Applications.* 1987, New York: Addison Wesley.

[4] Hossu D., Fagarasan I., Hossu A., Iliescu S.S., Evolved Fuzzy Control System for a Steam Generator, *International Journal of Computers Communications & Control*, 2010, 5(2):P. 179-192.

[5] Mhalla A., Jerbi N., Dutilleul S.C., Craye E., Benrejeb M., Fuzzy Filtering of Sensors Signals in Manufacturing Systems with Time Constraints , *International Journal of Computers Communications & Control*, 2010, 5(3):p. 362-374.

[6] Cordova F.M., Leyton G., A Fuzzy Control Heuristic Applied to Non-linear Dynamic System Using a Fuzzy Knowledge Representation, *International Journal of Computers Communications & Control*, 2010, 5(5): P. 664-674

[7] Tamir, D.E. and A. Kandel, *An axiomatic approach to fuzzy set theory.* Information Sciences, 1990, 52: p. 75-83.

[8] Tamir, D.E., Kandel, A., *Fuzzy semantic analysis and formal specification of conceptual knowledge.* Information Sciences, Intelligent systems, 1995, 82(3-4): p. 181-196.

[9] Ronen, M., Shabtai, R., Guterman, H., *Hybrid model building methodology using unsupervised fuzzy clustering and supervised neural networks.* Biotechnology and Bioengineering, 2002, 77(4): p. 420-429.

[10] Bhattacharya M., Das A., Genetic Algorithm Based Feature Selection In a Recognition Scheme Using Adaptive Neuro Fuzzy Techniques, *International Journal of Computers Communications & Control*, 2010, 5(4): p. 458-468.

[11] Zadeh, L.A., *The concept of a linguistic variable and its application to approximate reasoning - Part I.* Information Sciences, 1975, 7: p. 199-249.

[12] Mendel, J.M., *Uncertainty, fuzzy logic, and signal processing.* Signal Processing, 2000, 80: 913-933

[13] Tamir, D.E., Kandel, A., "The Pyramid Fuzzy C-means Algorithm," *The International Journal of Computational Intelligence in Control*, December 2010.

[14] Agarwal, D., Tamir, D. E., Last, M., Kandel, A., *A comparative study of software testing using artificial neural networks and Info-Fuzzy networks.* 2011, (Submitted).

[15] Ramot, D., Milo, R., Friedman, M., Kandel, A., *Complex fuzzy sets.* IEEE Transactions on Fuzzy Systems 2002, 10(2): p. 171-186.

[16] Ramot, D., Friedman, M., Langholz, G., Kandel, A., *Complex fuzzy logic.* IEEE Transactions on Fuzzy Systems, 2003, 11(4): p. 450-461.

[17] Dick, S., *Towards complex fuzzy logic.* IEEE Transaction on Fuzzy Systems, 2005, 13: p. 405-414.

[18] Tamir, D.E., Lin, J., Kandel, A., *A New Interpretation of Complex Membership Grade*, Accepted for publication in the International Journal of Intelligent Systems, 2011.

[19] Tamir, D.E., Last, M., Kandel, A., *Generalized Complex Fuzzy Propositional Logic.* Accepted for publication in the World Conference on Soft computing, San Francisco, 2011.

[20] Baaz, M., Hajek, P., Montagna, F., Veith, H., *Complexity of t-tautologies.* Annals of Pure and Applied Logic, 2002, 113(1-3): p. 3-11.

[21] Cintula, P., *Weakly implicative fuzzy logics.* Archive for Mathematical Logic, 2006, 45(6): p. 673-704.

[22] Cintula, P., *Advances in Ł and Ł1/2 logics.* Archives of Mathematical Logic, 2003, 42: p. 449-468.

[23] Montagna, F., *On the predicate logics of continuous t-norm BL-algebras.* Archives of Mathematical Logic, 2005, 44: p. 97-114.

[24] Hájek, P., *Fuzzy logic and arithmetical hierarchy.* Fuzzy Sets and Systems, 1995, 3(8): p. 359–363.

[25] Moses, D., Degani, O., Teodorescu, H., Friedman, M., Kandel, A. *Linguistic coordinate transformations for complex fuzzy sets. in 1999 IEEE International conference on Fuzzy Systems.* 1999, Seoul, Korea.

[26] Zhang, G., Dillon, T. S., Cai, K., Ma, J., Lu, J., *Operation properties and delta equalities of complex fuzzy sets.* International Journal on Approximate Reasoning, 2009, 50(8): p. 1227-1249.

[27] Nguyen, H.T., Kandel, A., Kreinovich, V. *Complex fuzzy sets: towards new foundations. in Proceedings, 2000 IEEE International Conference on Fuzzy Systems.* 2000, San Antonio, Texas.

[28] Deshmukh, A.Y., Bavaskar, A. B., Bajaj, P. R., Keskar A. G., *Implementation of complex fuzzy logic modules with VLSI approach.* International Journal on Computer Science and Network Security, 2008, 8: p. 172-178.

[29] Chen, Z., Aghakhani, S., Man, J., Dick, S., *ANCFIS: A Neuro-Fuzzy Architecture Employing Complex Fuzzy Sets.* 2009, (submitted).

[30] Man, J., Chen, Z., Dick, S. *Towards inductive learning of complex fuzzy inference systems. in Proceedings of the International Conference of the North American Fuzzy Information Processing.* 2007, San Diego, CA.

[31] Hirose, A., *Complex-valued neural networks.* 2006: Springer-Verlag.

[32] Michel, H.E., Awwal, A. A. S., Rancour, D. *Artificial neural networks using complex numbers and phase encoded weights-electronic and optical implementations. in Proceedings of the International Joint Conference on Neural Networks.* 2006, Vancouver, BC.

[33] Noest, A.J., *Discrete-state phasor neural nets.* Physics Review A, 1988, 38: p. 2196-2199.

[34] Leung, S.H., Hyakin, S., *The complex back propagation algorithm.* IEEE Transactions on Signal Processing, 1991, 39(9): p. 2101-2104.

[35] Iritani, N.T., Sakakibara, K., *Improvements of the traffic signal control by complex-valued Hopfield networks. in Proceedings of the International Joint Conference on Neural Networks.* 2006, Vancouver, BC.

[36] Buckley, J.J., *Fuzzy complex numbers.* Fuzzy Sets and Systems, 1989, 33: p. 333-345.

[37] Buckley, J.J., Qu, Y., *Fuzzy complex analysis I: differentiation.* Fuzzy Sets and Systems, 1991, 41: p. 269-284.

[38] Zhang, G., *Fuzzy limit theory of fuzzy complex numbers.* Fuzzy Sets and Systems, 1992, 46(2): p. 227-2352.

[39] Wu, C., Qiu, J., *Some remarks for fuzzy complex analysis.* Fuzzy Sets and Systems, 1999, 106: p. 231-238.

[40] Běhounek, L., Cintula, P., *Fuzzy class theory.* Fuzzy Sets and Systems, 2005, 154(1): p. 34-55.

Int. J. of Computers, Communications & Control, ISSN 1841-9836, E-ISSN 1841-9844
Vol. VI (2011), No. 3 (September), pp. 577-580

On the meaning of approximate reasoning – An unassuming subsidiary to Lotfi Zadeh's paper dedicated to the memory of Grigore Moisil –

H.-N. L. Teodorescu

Horia-Nicolai L. Teodorescu
Gheorghe Asachi Technical University,
Iasi, Romania
Institute for Computer Science,
Romanian Academy, Iasi Branch

Abstract: The concept of "approximate reasoning" is central to Zadeh's contributions in logic. Standard fuzzy logic as we use today is only one potential interpretation of Zadeh's concept. I discuss various meanings for the syntagme "approximate reasoning" as intuitively presented in the paper Zadeh dedicated to the memory of Grigore C. Moisil in 1975.
Keywords: logic, truth value, natural language, inference.

1 Introduction

I perceive three central ideas in Zadeh's wide conceptual construction in his work until now. The first one is that words and propositions in natural languages, consequently human thoughts are not representable by standard sets and standard binary predicates. The second central idea in many of Zadeh's papers is that humans perform computations in an approximate manner that real numbers can not represent well. The third idea, which Zadeh presented in his more recent works, is that of granularity of human mental representations and reasoning. These three strong points were represented by Zadeh in various forms and synthesized in the title of his paper on "computing with words" [1].

Most frequently - and Zadeh himself is doing no exception - authors cite as the initial paper clearly stating the approximate reasoning model the one published in the journal Information Sciences, July 1975 [2]. However, another paper published in Synthese [3], during the same year, deals with the main ideas in approximate reasoning, not to mention a conference paper published in 1974.

The paper published in Synthese is important in several ways. First, it offers clear explanations of the meanings associated by Zadeh to the syntagme "approximate reasoning". Second, the paper was published in a well-established journal that "spans the topics of Epistemology, Methodology and Philosophy of Science," thus boldly reaching to a large audience in a set of fundamental disciplines and daringly affirming the importance of the new concepts. Third, the paper is important for the Romanian scientists because it re-enforces the understanding of Grigore Moisil's contributions and the deep connections between Zadeh and Moisil. For an in-depth coverage of the last topic, see [4], [5].

Zadeh starts the paper [3] from the common-sense remark that "It is a truism that much of human reasoning is approximate rather than precise in nature." From that remark, he builds a broad program for research.

The program that Zadeh establishes remains, in my opinion, unfinished. In the abstract of the paper discussed here, Zadeh states "Since T [the truth-value set] is not closed under the operations of negation, conjunction, disjunction and implication, the results of an operation on truth-values in T requires, in general, a linguistic approximation by a truth value in T." With

Copyright © 2006-2011 by CCC Publications

this clarification, classic fuzzy logic as we know today is not a proper representation of Zadeh's original ideas.

There are two ways to interpret the above quotation. The first interpretation is in the frame of the classic thinking and runs as follows. Because language is a set of propositions (we restrict the discussion to truth-valuable propositions only), under whatever logic, all simple and composed propositions have a truth value. Denoting the set of truth values by T, the above remark by Zadeh has no effect. Moreover, when defining the truth-valuation function we already need to know the set of truth values, which again makes ineffective Zadeh's remark. Standard fuzzy logic pursue this direction of thinking, as it starts with . Under this approach, as T is given, what is needed is to define the logic operators for the respective logic.

The second interpretation is constructive. It may start with the assumption that language is not predetermined and must be constructed as a recursive, dynamic process, as poets, writers, other language professionals, and laymen do every day. Whenever a new proposition is invented, it is assigned a meaning. The meaning includes what we conveniently name "truth", a coverage degree of the reality that we need for making inferences. The *truth* may or may not be numerically representable. Moreover, the truth of a composed proposition may not be representable by the truth of each of the initial propositions. Therefore, the "set of truth values" evolves continuously. That, in turn, creates a stumbling block. Because we are supposed to know the truth of the original sentences, but not of the result, how are we supposed to infer? The answer proposed by Zadeh is that we still have an approximate truth in the initial set of truth values, approximating the truth of the new proposition. He hypothesizes that there is always an inverse application, which I will name *projection*, from the new truth set to the original one, indicating a truth degree in the original set that approximates the truth degree of the composed proposition. That makes our reasoning possible, if approximate.

This way of thinking, more or less directly suggested by Zadeh, opens the door to several formal descriptions. I sketch below a loose formalism for the recursive approach.

Consider a language $L = (\{p\}, c_1, c_2, \cdots c_r)$, where p denotes extant propositions and c_k logic operators (either unary or binary). Notice that L is an initial language, meaning that the initial set $\{p\}$ is evolvable, that is, it is recursively increased by adding propositions correctly formed from the initial ones through logic operations. Assume that any proposition has a characteristic named truth value. Thus, there is some set T of truth values (I use for this set the same notation as Zadeh's), as well as an application such that for any proposition p there is a truth value in T. Also assume that any propositions can be concatenated to produce a new or extant proposition. Whenever such proposition is new in the language, it has a characteristic truth value which is not necessarily in T. We can regard the creation of new propositions (including those used for reasoning) as producing an application $T \to T\prime$, as for the negation operator, or $T \times T \times \cdots \to T\prime$, as for the connectives. The new $T-$set which includes the original one. Moreover, for satisfying Zadeh's hypothesis, there is a projection operation ϑ. Notice that whenever a new proposition is produced, L is modified.

Because of T non-closure to logic operations, we can not talk about a T (proper) set for a language. Instead, we can conceive an object named T^*-set that can be seen as a string of sets, each generated recursively from the previous T-set or from a product of T-sets. The string must be complemented by a string of applications from a T-set (or from a product of such sets) to the next one in the sequence. Next, we need to allow for a construction where the term *approximate* makes sense. Several mathematical constructions may be used in this respect, namely the concept of topological space, where approximate may mean "in a vicinity of", the concept of metric space or one of its variants, as pseudo-metric space, where the term *approximate* may be interpreted as "at a distance less than epsilon", or a measurable space, where the truth may be conceived as elements of the space and the approximation is interpreted as "the difference between the measures of the respective two truth values (here, a value can be represented for example by a measurable set) is less than epsilon." The version based on the measurable spaces corresponds to the probabilities. The same version is adopted below, but with no direct connection to probabilities.

The main requirements for the T^*-set are: i) The truth-attribute of a proposition is represented by an application $\theta : L \to T$, where T is measurable. ii) For any two extant propositions p_1 and p_2 and for any connective c, the new proposition p_1cp_2 has a truth degree such that $T \cup \{\theta(p_1cp_2)\}$ is measurable and includes the measurable space T. iii) For any valid linguistic construction that uses logic operators, there is an application $\vartheta : T\prime \to T\prime$, named back-projection. iv) For any connective c, there is a formula f_c such that $\vartheta(\theta(p_1cp_2))$ and $f_c(\theta(p_1), \theta(p_2))$ are two points close enough in T.

Standard fuzzy logic obviously satisfies the above conditions, with $T = T\prime$. Replacing the measurable space with a metric one, we obtain another construction, which may be closer to Zadeh's intuition in [3]. Using distances $d : T \to T$, the conditions to satisfy are:

$$d(\vartheta(\neg p), 1 - \theta(p)) < \varepsilon$$

$$d(\vartheta(p \bigwedge q), min(\theta(p), \theta(q)) < \varepsilon$$

$$d(\vartheta(p \bigvee q), max(\theta(p), \theta(q)) < \varepsilon,$$

where ε is the allowed approximation error.

As Grigore Moisil has done, Zadeh emphasizes in all his work how much language is fundamentally creative. By stating that " $T \cdots$ is not closed under the logic operations," he highlights that there is always generation of new meaning, or at least an evolution of the meaning, whenever a sentence is uttered. I believe that this is one of the fundamental contributions made by Zadeh's work until now.

Understanding the creative process in languages and in reasoning was significantly elucidated by Zadeh's work, yet much remains to be done for deriving conclusions and theoretical developments in the directions pointed to by his works. The developments informally suggested above may indicate several such directions, yet others may be put forward in a more formal manner during the years to come.

Acknowledgment. I thank several colleagues and the referees for their comments and suggestions. The note was not supported by, but is related to research performed in the frame of the Romanian Academy.

Bibliography

[1] L.A. Zadeh, *Fuzzy logic = computing with words*. Fuzzy Systems, IEEE Transactions on, Volume: 4 Issue 2, pp. 103–111 (May 1996)

[2] L.A. Zadeh, *The concept of a linguistic variable and its application to approximate reasoning - I*, Information Sciences, vol. 8, no. 3, pp. 199-249, July 1975.

[3] L.A. Zadeh, *Fuzzy Logic and Approximate Reasoning (In Memory of Grigore Moisil)*. Synthese, vol. 30, (1975), pp. 407-428, D. Reidel Publishing Co, Dordrecht, Holland

[4] G. Georgescu, A. Iorgulescu, S. Rudeanu, *Grigore C. Moisil (1906 - 1973) and his School in Algebraic Logic*. Int. J. Computers, Communications & Control, Vol. I (2006), No. 1, pp. 81-99

[5] Z. Lemnaouar and A. Abdelaziz, *On the representation of L-M algebra by intuitionistic fuzzy subsets*. ARIMA J., Volume 4 (2006), pp. 72-85

Int. J. of Computers, Communications & Control, ISSN 1841-9836, E-ISSN 1841-9844
Vol. VI (2011), No. 3 (September), pp. 581-582

"The Disassembly Line: Balancing and Modeling" -Book Review-

F.G. Filip

Title of book: The Disassembly Line: Balancing and Modeling
Authors: Seamus M. McGovern and Surendra M. Gupta
Editing House: McGraw Hill Companies Inc., New York
ISBN: 978-0-07-162287-5
Pages: XVII+373

This book is about the *disassembly* of end-of-life products with particular emphasis on methods and techniques for solving the *Disassembly Line Balancing Problem.*

Disassembly is viewed as "the systematic separation and extracting valuable entities for possible future re-usage". In fact, disassembly is a distinct phase of the product lifecycle. It follows the "before life" phases (such as design and economical evaluation), "useful period" phases (such as manufacturing, distribution, usage and maintenance) and "end of life" phases (such as collecting, sorting). Disassembly might represent the essential first phase of the future activities, such as re-use and re-manufacturing and recycle. Due to the ever higher public awareness, the more and more strict regulations concerning environment quality preservation and increasing economic effectiveness and attractiveness for industry, the activities of recovering valuable parts and subassemblies have become a desirable alternative to the old fashioned disposal processes of end-of-life products.

The authors state, in the preface of the book, that the "disassembly line seems to be the most efficient way to disassemble a product". Consequently, the primary concern of the book is "the complete disassembly of [end-of-life] products on a paced disassembly line for component/material recovery purposes". The authors aim at investigating "the qualitative and quantitative aspects of multi-criteria solution sequences using the various combinatorial optimization techniques" (page 16) to solve the Disassembly Line Balancing Problem (DLPB). The DLBP consists in finding a disassembly feasible solution sequence which preserves precedence constraints and aims at attaining several objectives, such as minimizing the number of work stations and total idle time, ensuring similar idle time at each work station, while attempting to remove hazardous parts and materials and extracting highly demanded product components at the earliest moments possible and minimizing the number of direction changes required for disassembly (removing parts with similar part removal directions together), (page 102).

The book is composed of 29 chapters grouped into three parts entitled "Disassembly Background", "Disassembly-Line Balancing" and "Further Disassembly-Line Considerations" which address general aspects concerning disassembly processes, variations of methods and techniques to solve the DLBP, and other problems related to the disassembly line, respectively.

Part I comprises six chapters which are meant to set the stage for the subsequent chapters. Various information concerning disassembly processes, assembly lines, disassembly lines, other related researches, graphical representations and computational complexity of combinatorial problems are provided.

Part II is made up of 20 chapters and addresses the statement and analysis of the DLBP and several specific variations of methods and techniques which were adapted for solving the problem, tested on four application cases exprimental instances and compared. The objectives of this part of the book are: stating the mathematical model of DLBP, establishing the difficulty of the problem by using the complexity theory and determining the data sets and evaluation criteria to be used in analyzing the problem and solving techniques which are selected (page 99).

Copyright © 2006-2011 by CCC Publications

It is demonstrated (in chapter 9) that the DLBP is a complex NP complete problem in the strong sense and necessitates specialized solution techniques. Accordingly, authors plea for combinatorial optimization approaches and select several algorithms to solve the problem. The techniques to be utilized to solve the DLBP are introduced in chapter 10 and their usage and performances in solving the problem are presented in chapters 12 through 19. There are seven techniques which are adapted, tested and compared. The *exhaustive search* is used to provide the optimal solution. Two *metaheuristic* approaches (genetic algorithms and ant colony optimization) are next studied. Two *purely deterministic searches* (the greedy algorithm and the "hunter-killer" search) and two *2-phase hybrid* methods are adapted and tested.

The four *experimental instances* (the eight-part personal computer, the enhanced 10-part DLBP case, the 25-part cellular instance, and the size independent "a priori" benchmark with a known optimal solution) are described in chapter 11. Chapter 20 contains a detailed comparison of the six heuristic and metaheuristic techniques as applied to the DLBP with respect to several performance measures. Several complementary research results are reviewed in chapter 21 together with future research directions.

Disassembly processes interact with other "before life", "useful", and "after life" periods of product usage and recovery. As a result, to make the picture complete, **Part III** addresses other areas of disassembly research such as product planning, line and facility design, operations scheduling and sequencing, inventory, "just-in-time", revenue and unbalanced lines (chapter 22 through 29).

The authors of the book form a team who may be viewed as a fine and synergic combination of two complementary experiences and backgrounds from academia and industry. Seamus McGovern, an Electronics Engineer at the Volpe National Transportation System Center, holds a commission in the US Navy as an aerospace duty engineer as well as a part-time industrial engineering faculty appointment at the Northeastern University. Surendra M. Gupta is a professor of Mechanical and Industrial Engineering and a director of the Laboratory for Responsible Manufacturing at the Northeastern University. He has authored/co-authored over 400 technical papers and is a pioneer in the domain of the book.

This book represents a very valuable work in a rather young research domain, which may be viewed as opened by the pioneering paper of Güngör and Gupta entitled "Disassembly Line in product Recovery (*International Journal of Production Recovery*, 40 (11), 2002). The volume mainly reflects the original studies of authors and their colleagues. It also makes an exhaustive and systematic review of the results which are reported in the domain scientific literature and are due to other scientists. The organization of the document is well thought and the presentation style is rigorous and clear. Subsequently, though information content is very dense and diverse, the book is accessible and its study is scientifically rewarding. Special remarks can be made to the uniform and coherent notation which is used throughout the book and to graphical illustrations. A final remark of appreciation is to be made to the excellent quality of editing and printing of the book due to the staff of McGraw-Hill Companies.

In conclusion, the book is a timely work which contains relevant, inspiring and challenging information. Therefore, this reviewer warmly recommends it to the readers of academia and industry as well who are interested in modern manufacturing issues and combinatorial optimization methods and software.

Florin Gheorghe Filip
Romanian Academy, Bucharest, Romania
E-mail: ffilip@acad.ro
http://www.biblacad.ro/FlorinFilip.html

Author index

L.A. Zadeh, D. Tufis, F.G. Filip, I. Dzitac (eds.), *From Natural Language to Soft Computing: New Paradigms in Artificial Intelligence*, Editing House of Romanian Academy, ISBN 978-973-27-1678-6, 2008, 268 pages

Authors of chapters: Ahmed S., Balas M., Balas V., Buciu I., Castaneda C., Chira C., Cordova F.M., Dumitrescu D., Dzitac I., Dzitac S., Felea I., Fodor J., Lefranc G., Nafornita I., Nikolova N., Nunez R., Ortega A., Pintea C-M., Quezada L.E., Radojevic D.R., del Rosal E., Rudas I.J., Stefanescu G., Tenekedjiev K., Toneva D., Tufis D., Vesselenyi T., Zadeh L.A.

Some questions from this volume:

"What is meant by Computation with Information Described in Natural Language, or NL-Computation, for short? Does NL-Computation constitute a new frontier in computation? Do existing bivalent-logic-based approaches to natural language processing provide a basis for NL-Computation? What are the basic concepts and ideas which underlie NL-Computation? These are some of the issues which are addressed in the following. What is computation with information described in natural language?" [...]
(*L.A. Zadeh, A New Frontier in Computation - Computation with Information Described in Natural Language*)

Corresponding editor:
Prof. Ioan Dzitac
Aurel Vlaicu University of Arad, Romania
Email: ioan.dzitac@uav.ro
http://www.uav.ro/en/people/exact-sciences/dzitac-ioan