A qualitative-fuzzy framework for nonlinear black-box system identification

R. Bellazzi
Dipartimento di Informatica e Sistemistica
Università di Pavia
Via Ferrata 1
27100 Pavia, Italy
ric@aim.unipv.it

R. Guglielmann and **L. Ironi**
Istituto di Analisi Numerica
Consiglio Nazionale delle Ricerche
Via Abbiategrasso 209
27100 Pavia, Italy
{raffy, ironi}@dragon.ian.pv.cnr.it

Abstract

Reasoning about a natural or man-made system often requires a model that gives an accurate *quantitative* description of its dynamics. Such a model may be built from observations on the system either using structural models of the underlying physics or learning an input-output relation directly from observed data. The process of modeling based on experimentation is called *System Identification*, and in the end, whichever approach is used, it reduces to an optimization procedure for parameter estimation. A central problem of optimization techniques deals with the choice of a good initialization. This paper presents a novel approach to nonlinear black-box system identification which combines QR methods with fuzzy logic systems: such a method aims at building a good initialization of a fuzzy identifier, so that it will converge to the input-output relation which captures the nonlinear dynamics of the system. Fuzzy inference procedures are initialized with a rule-base predefined by the human expert: when such a base is not available or poorly defined, the inference procedure becomes extremely inefficient. Our method aims at solving the problem of the construction of a meaningful rule-base: fuzzy rules are automatically generated by encoding the knowledge of the system dynamics described by the outcomes of its qualitative simulation. Both efficiency and robustness of the method is demonstrated by its application to different domains: in this paper, we consider the problem of identifying the dynamics of Thiamine (vitamin B_1) and its phosphoesters in the cells of the intestine tissue.

Introduction

System Identification (SI) deals with the development and analysis of methods which infer a *quantitative* model of the system dynamics from observed data (Haber & Unbehauen 1990; Juditsky *et al.* 1995; Ljung 1987; Söderström & Stoica 1989). SI is applied to perform both grey and black-box modeling tasks; we refer to grey models when the available knowledge of the underlying physics is such that the structural model equation may be explicitely formulated but the numeric values of some parameters are unknown, and to black-box models when the structural equation may not be written because either the knowledge about functional relationships is incomplete or unknown or no first principles are available. In the former case (*parametric* approach), the problem reduces to the optimal estimation of the unknown parameter values in the equation by using the measured data. More generally, parametric SI may require to choose, in accordance with suitable criteria, also the "best" structural equation within a set of plausible candidate models.

In black-box SI (*non-parametric* approach), the problem is much more complex as it requires to choose an identifier model (Haber & Unbehauen 1990) as well as to estimate its parameters. Both the identifier and parameters do not necessarily refer to the underlying physical reality: as a matter of fact, the parameters do not have a physical meaning but are a means for adjusting the fit to the data.

A common problem to both approaches deals with the optimization procedure for parameter estimation which may terminate at local extrema if the numerical search for the optimal value is not properly initialized. Recent work within the Qualitative Reasoning (QR) research framework has addressed the problem of the parameter estimation phase in the model building process (Bradley 1994; Bradley, O'Gallagher, & Rogers 1997; Capelo, Ironi, & Tentoni 1996; 1998), namely the crucial issue of automatically providing a "good" initial guess to start the optimization procedure. The implemented QR techniques proposed in the mentioned papers are components of computational environments which integrate a variety of techniques with the goal of automating model building. Such papers consider ODE models which are automatically generated by exploiting physical knowledge either of a specific physical domain (Capelo, Ironi, & Tentoni 1996; 1998) or explicitely supplied by the user's input (Bradley 1994; Bradley, O'Gallagher, & Rogers 1997). The candidate ODE models for parametric SI are selected in accordance with their consistency with the observations. An interesting method, based on semi-quantitative infer-

ences, to reduce the candidate model space is given in (Kay 1996). In the same paper, it is also proposed a non-parametric identifier that fits a monotonic function to noisy data.

The problem we are addressing here is how QR techniques can be used to improve the performance of non-parametric approaches. It exists a rich and well-established theory for black-box modeling of linear physical systems (Ljung 1987; Söderström & Stoica 1989). Nonlinear SI, which is much more complex and problematic, has been studied for a long time, under the label *non-parametric regression*, within the statistics community. Recently, due to the paucity of directly applicable results, nonlinear SI has received more and more attention in the control community with a consequent development of a number of new approaches capable to describe the nonlinear dynamics of a real system from input-output data. Neural networks, multi-variate splines and fuzzy logic systems are the most known approximation schemes used for learning an input-output relation from data (Jang 1993; Khannah 1990; Wang 1994). Although these approaches are successfully applied to a variety of domains, they are affected by two main drawbacks: first, the model identification procedure usually requires a large amount of data and is often extremely inefficient; second, the identification result, a nonlinear function, does not capture any structural knowledge. With the goal to overcome these drawbacks, we propose a novel method which combines the qualitative and non-parametric modeling frameworks. Therefore, such a method is applicable whenever the incompleteness of the available structural knowledge of the system under study is not so strong as to prevent from formulating a qualitative model of its dynamics through Qualitative Differential Equations (QDE).

We believe that the efficiency and robustness of nonlinear black-box SI methods may improve only if they incorporate and exploit *all* available knowledge of the system, namely the structural and human knowledge, and the experimental one. Qualitative models and linguistic rules represent properly the structural and human expert knowledge, respectively. As qualitative modeling formalism we have chosen QSIM (Kuipers 1994) because of both its expressive power to represent QDE and its reasonable predictive capacity. As nonlinear identifiers we have chosen fuzzy logic systems since various classes of Fuzzy Systems (FS) can be proved to have the universal approximation property (Wang 1994), i.e. they can approximate any real continuous function on a compact set at any degree of accuracy. A clear advantage of using FSs deals with their capability to incorporate in the same framework both linguistic descriptions of the unknown system dynamics, in the form of IF-THEN rules, and experimental data. Moreover, the meaning of their parameters is linked to the input-output data. Fuzzy inference procedures are initialized with a rule-base which defines the structure of the input-output relation where parameters occur. When such a base is not available or poorly defined, also the fuzzy inference may become extremely inefficient as the inference structure has to be determined using only numerical evidence. Our method aims at solving the central problem of the construction of a meaningful rule-base: fuzzy rules are automatically generated by encoding the knowledge of the system dynamics captured by its qualitative simulated behaviors. In outline, our method, which we label FS-QM, may be summerized as follows:

- formulation and simulation of a qualitative model of the system at hand;
- automatic generation of both a fuzzy rule base from the outcomes of the qualitative simulations and its associated FS;
- parameter estimation of the FS from the experimental data.

FS-QM may be applied to a number of different domains. As test-benchmark we have considered problems from the medical domain. In particular, we applied it for the identification of the dynamics of the Blood Glucose metabolism in insulin dependent diabetic patients in response to different perturbations, such as meals and conventional insulin therapies (Bellazzi *et al.*), as well as for the identification of the dynamics of Thiamine (vitamin B_1) and its phosphoesters in the cells of the intestine tissue. In both cases, FS-QM has shown good performance as for efficiency and robustness.

In this paper we discuss the latter application problem. Thiamine is transformed within the cells through an enzyme-mediated chemical reaction in higher energy forms (Piro-Posphate Thiamine, ThPP, and Mono-Posphate Thiamine, ThMP) that are used in the Krebbs cycle. The chemical reaction is nonlinear, and conventional modeling is hampered by experimental identification problems. By using FS-QM we build an approximator of the system dynamics, that can be used to describe the different metabolic responses of subjects with different pathological conditions.

Brief introduction to fuzzy identifiers

In this section we recall the basic concepts and definitions of fuzzy identifiers which are relevant to a clear description of the method.

A *Fuzzy set* F in U is a generalization of the concept of ordinary set. F is characterized by a membership function $\mu_F : U \rightarrow [0, 1]$, where U, *Universe of discourse*, is the collection of objects that we would like to reason about, and $\mu_F(u)$ represents the degree of membership of $u \in U$ to the set F.

A Fuzzy System (FS) exploits fuzzy concepts for reasoning about a set of objects. A *Linguistic Variable* is a variable whose values are words in natural language, and that can be represented through a collection of fuzzy sets.

A *Fuzzy Rule Base* (FRB) is the knowledge base that we use to reason about the objects in U. In particular, a FRB is a collection of rules of the kind:

IF x_1 is F_1 and ... and x_n is F_n THEN y is G

where $F_1, \ldots, F_n$ and G are fuzzy sets, and $x_1, \ldots, x_n$ and y are linguistic variables.

The most popular FSs proposed in the literature basically differ in the way they map input (antecedents) and output (consequents) information into a rule (Wang 1994). We restrict to the class of FSs that are useful for function approximation, i.e. to those FSs which are proved to be universal approximators. In particular, we would like to approximate continuous functions from $\Re^n \rightarrow \Re$: the value of an output variable y defined on $V \subset \Re$ is inferred by using a FRB with n input variables defined on $U_i \subset \Re$, such that $U = U_1 \times \ldots \times U_n$. This means that each continuous variable should be properly fuzzified. The *fuzzification* operation performs a mapping from $x_i \in U_i$ into a fuzzy set. We exploit the *singleton fuzzifier*, that transforms a real number $x_i \in U_i$ into a fuzzy set with membership function defined over U_i, such that it is equal to 1 in x_i and to 0 elsewhere. Once a quantitative measure of the inputs is fuzzified, it is evaluated as an input of all the rules in the FRB.

In order to perform inferences, it is necessary to resort to a machinery that gives a precise interpretation of the terms is, and and then that appear in a fuzzy rule. For our purposes, it is sufficient to say that it is possible to select a Fuzzy Inference Engine (FIE), in which the fuzzy logic principles are used to map fuzzy sets in U into fuzzy sets in V. In particular, we have chosen as FIE the product-inference rule (Wang 1994): given an input in U, we obtain a collection of M fuzzy sets as outputs, where M is the number of rules of the FRB.

For the purpose of function approximation, we must transform into a real number the output of the application of FIE. This may be done through a *defuzzification* operation (Mamdani 1974): we exploit the *Center Average Defuzzifier*, in which the collection of M fuzzy sets is averaged as follows:

$$y(\underline{x}) = \frac{\sum_{j=1}^{M} \hat{y}_j \prod_{i=1}^{n} \mu_{F_i^j}(x_i)}{\sum_{j=1}^{M} \prod_{i=1}^{n} \mu_{F_i^j}(x_i)} \qquad (1)$$

where $\mu_{F_i^j}$ is the membership function associated with the linguistic variable of x_i that appears in the j-th rule, and $\hat{y}_j$ is the point in V where G_j reaches its maximum value. This functional form can be easily exploited in function approximation without requiring strong assumptions on the input-output relation. The system described by the equation (1) can be rewritten as:

$$y(\underline{x}) = \sum_{j=1}^{M} \hat{y}_j \ \Phi_j(\underline{x}) \qquad (2)$$

where

$$\Phi_j(\underline{x}) = \frac{\prod_{i=1}^{n} \mu_{F_i^j}(x_i)}{\sum_{j=1}^{M} [\prod_{i=1}^{n} \mu_{F_i^j}(x_i)]}$$

and $\mu_{F_i^j}$ has the same meaning as in the equation (1). The functions $\Phi_j()$s are called Fuzzy Basis Functions (FBF); the FS is hence viewed as a linear combination of such FBFs.

In the following, we will exploit a useful class of FS for fuzzy approximation with Gaussian membership functions. Then, FS assumes the mathematical form:

$$y(\underline{x}) = \frac{\sum_{j=1}^{M} \hat{y}_j [\prod_{i=1}^{n} exp(-(\frac{x_i - \hat{x}_i^j}{\sigma_i^j})^2)]}{\sum_{j=1}^{M} [\prod_{i=1}^{n} exp(-(\frac{x_i - \hat{x}_i^j}{\sigma_i^j})^2)]} \qquad (3)$$

Such an expression allows us to interpret the nonlinear function approximation problem with FS as the process of identifying the vector of parameters $(\hat{y}_j, \hat{x}_i, \sigma_i)$ of a known nonlinear function from a set of data. Moreover, the FS defined by the equation (3) holds the universal approximation property (Wang 1994). Other properties can be found in (Kim & Mendel 1995; Wang 1994).

Fuzzy System identification

Once that the FS has been defined, a function approximator $y(\underline{x})$ may be obtained. Such a function approximator is dependent on a parameter vector $\underline{\theta}$, such that $y(\underline{x}) = y(\underline{x}, \underline{\theta})$, where $\underline{\theta} = \{\underline{\hat{y}}, \underline{\hat{x}}, \underline{\sigma}\}$ with $\underline{\hat{y}} = \{\hat{y}_j\}$, $\underline{\hat{x}} = \{\hat{x}_i^j\}$ and $\underline{\sigma} = \{\sigma_i^j\}$, where $i = 1, \ldots, n$, $j = 1, \ldots, M$.

Numeric values, which express the prior knowledge included in the system, and which provide for an initial guess of the system dynamics, are assigned to the parameter vector $\underline{\theta}$ in the construction phase of the FS. Given a set of experimental data, the estimate of $\underline{\theta}$ is then refined through optimization procedures. The identification process can be carried out considering the meaning of the antecedents and consequents of each derived fuzzy rule.

In FS-QM, the description of the system dynamics is performed through rules that give the next value y_{k+1} of the system output as a function of the values of the current inputs $(\underline{u}_k)$ and output (y_k). Then, the output behavior can be described with a Nonlinear AutoRegressive eXogenous input model (NARX) of the kind:

$$y_{k+1} = y(\underline{x}_k, \underline{\theta}) + v_k \qquad (4)$$

where k is a discrete time index, $\underline{x}_k = \{\underline{u}_k, y_k\}$, v_k is the measurement error and $y(\cdot)$ has the same meaning of equation (3).

Since the model (4) is nonlinear in the model parameters, it is necessary to resort to nonlinear identification techniques to estimate the vector $\underline{\theta}$. Within FS-QM a number of approaches may be considered, but two of them seem to take particularly advantage of it.

In the first approach, the FS is represented as feed-forward networks, as described in (Wang 1994), and the identification of $\underline{\theta}$ is performed by using the back-propagation (BP) technique. Such a technique allows us to estimate all the FS parameters through an iterative search in the solution space by employing a gradient descent search. If not properly initialized, such a search procedure may be either trapped at a local minimum or converge very slowly.

Another very interesting strategy to be investigated consists in fixing the parameters $\hat{x}_i^j$ and σ_i^j of the vector $\underline{\theta}$, so that only the parameters $\hat{y}_j$ are estimated from the data. Since the equation (3) is linear in the parameters $\hat{y}_j$, the identification problem expressed by equation (4) is also linear, and therefore it is possible to resort to efficient methods, such as classical linear least squares. The major advantage offered by this choice is that it allows us to preserve the structure of the FS initialized on the basis of the a–priori knowledge represented by the qualitative model.

Herein, we demonstrate that FS-QM provides for a significant initialization of BP algorithm, with a consequent fast convergence to the global minimum.

The FS-QM method

Fuzzy set theory and QR techniques share the motivation of facing too complex or ill-defined physical systems to be analyzed through conventional techniques. Then, the idea of unifying both frameworks with the goal of producing effective solutions to an extended range of application problems is a matter of course. With reference to this, we would like to mention the works proposed in (Shen & Leitch 1993; Vescovi & Travé-Massuyès 1992) which combine fuzzy sets and QR to produce, respectively, a fuzzy qualitative simulation algorithm (FuSim) that allows a semi-quantitative extension to QSIM and an extension to the numeric Euler's method so that it can handle qualitative coefficients.

Our method, FS-QM, is grounded on the intuition that analogies between elements of the QSIM and fuzzy formalisms can be highlighted (Fig. 1).

At a first level, namely the modeling one, given a physical system S characterized by $\{x_i,\ i = 1, \ldots, m\}$ state variables, the quantity space L_i associated with the variable x_i may find its semantic correspondence with the universe of discourse U_i of x_i.

The set $L_i \cup \{(l_j, l_{j+1}),\ l_j \in L_i\}_{j=1,j'-1}$, where j' is the cardinality of L_i, contains all significant distinct qualitative magnitude values ($qmag$) of x_i, whereas the elements of U_i are the supports of the membership functions μ which represent the linguistic values associated with x_i.

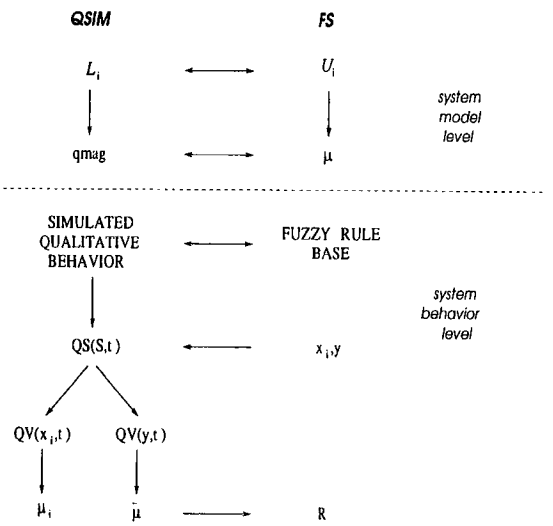


Figure 1: Analogies between the QSIM and FS formalisms. Bi-directional arrows indicate a semantic correspondence, whereas the other ones have the usual meaning.

At a second level, more related to the simulated system behavior, the qualitative system dynamics, which is described by the simulated Qualitative Behaviors (QB), may find its correspondence in fuzzy rule-bases. Each QB is defined by a sequence $QS(S, t_k)$ of qualitative states of the system, where $QS(S, t_k)$ is a m-tuple of the qualitative values of each individual system variable. Given the input and output variables, all their possible dynamics are extracted from the behavior tree, suitably manipulated, and automatically mapped into rule sets where the input and output variables are, respectively, the antecedents and consequent in the IF–THEN rules.

Remark 1. Let us observe that we map a quantity space, which is a set made up of landmarks and intervals, into fuzzy sets: then, apparently, we map a non-uniform set into a uniform one. Landmarks are symbolic names denoting particular real numbers that separate qualitatively distinct regions. However, in many applications, and in particular the medical domain we are interested in, a landmark value may not have a crisp representation but, in its turn, be defined by an interval with "soft" boundaries: as a matter of fact, we consider this case, and then we map a set made up of intervals into fuzzy sets. A semi-quantitative model representation (Kuipers 1994; Shen & Leitch 1993) would be formally more appropriate but not always feasible as we are considering systems whose knowledge is highly incomplete with regards to non-observable variables.

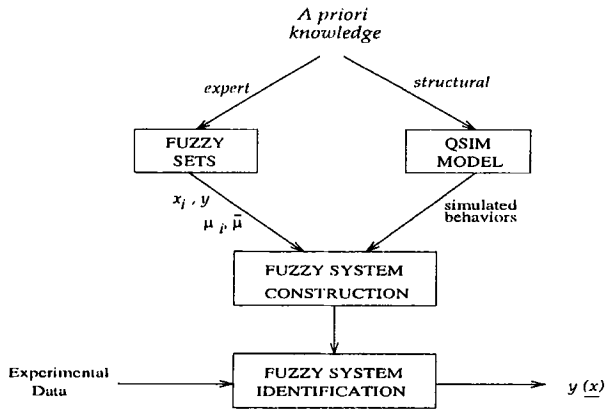


Figure 2: The basic step of FS-QM.

The overall system identification procedure proceeds in three main phases (Fig. 2):

1. QSIM *model formulation and simulation, and definition of the fuzzy elements.* The prior structural knowledge of the system at study must be organized so that its behavioral model can be defined. More precisely, the variables of interest and the network of interactions between them, along with their mathematical descriptions, must be specified: variables are described by their respective quantity spaces, whereas their interactions by a set of qualitative constraints which include both functional dependencies and equations governing the system dynamics. Then, an initial state of the system, which may describe a perturbation on it, has to be provided to simulate its behavior. An attainable envisionment, which does not generate any new landmarks, is performed to produce in one run all possible behaviors that could follow from the given possibly incomplete specification, and shows us the entire range of possible system dynamics at once. The idea underlying the mapping of a QB into a set of FRs exploits the semantic correspondence between the elements of quantity spaces and linguistic values defined through fuzzy membership functions μ. Then, the fuzzy quantity spaces of the input-output variables, i.e the linguistic variables associated with the quantity spaces of the input-output variables, and their corresponding μ, have to be defined. Such definitions are suggested by the expert knowledge.
2. *Construction of the* FS. The prior knowledge of the dynamics of the system captured by the QBs derived at phase 1 together with the fuzzy quantity spaces are exploited to generate automatically a base of FRs whose antecedents and consequent are the input and output variables used to quantitatively identify the dynamics of the system. The choice of the FIE, the fuzzification and defuzzification methods, complete the definition of the FS.
3. *Identification of the* FS *from the experimental data.* The generated rules, interpreted in accordance with the FIE selected at the previous phase, initialize an optimization procedure for the identification of the parameters in the FS, which learns from the available experimental data an accurate input-output relation.

The next subsection details the algorithms concerning the phase 2. The phases 1 and 3 will be exemplified in the next section through the application of the overall procedure to the identification of the dynamics of Thiamine and its phosphoesteres in the cells of the intestine tissue.

Construction of the FS

Given n input variables x_i ($n \leq m$), and the output variable y, a QB is automatically mapped into a FRB in the following steps of the method:

1. from the time set T, draw out its subset $\overline{T} = \cup_{i=1}^{n} T_{x_i} \cup T_y$, whose elements are the significant time-instants of both x_i and y (T_{x_i} and T_y are the sets of distinguished time-points of x_i and y, respectively);
2. from the elements of $\overline{T}$, built $\mathcal{T} = \{t_0\} \cup \{(t_{j-1}, t_j), t_j\}_{j=1\ldots k}$, where k is the cardinality of $\overline{T}$:
3. $\forall t_k \in \mathcal{T}$, where t_k may be either a time-point or a time-interval, repeat:
   (*a*) consider $QS(S, t_k)$;
   (*b*) from $QS(S, t_k)$, draw out the qualitative values of input and output variables, namely $QV(x_i, t_k)$ and $QV(y, t_k)$ (necessarily, $QV(y, t_{k+1})$ if $\{x_i\} \cap \{y\} \neq \emptyset$);
   (*c*) from $QV(x_i, t_k)$ and $QV(y, t_k)$, draw out its related qualitative magnitude *qmag*, i.e. $qmag(x_i, t_k)$ and $qmag(y, t_k)$;
   (*d*) consider the linguistic values, and therefore the membership functions μ_i and $\overline{\mu}$ which are associated with $qmag(x_i, t_k)$ and $qmag(y, t_k)$, respectively;
   (*e*) generate a linguistic rule R where x_i are the antecedents, y the consequent, and then the correspondent fuzzy rule where μ_i and $\overline{\mu}$ are the fuzzy sets.

Remark 2. The number of rules generated in correspondence with each QB is not greater than the cardinality of $\mathcal{T}$. Identical rules may happen to be generated as variables may have the same *qmag* either at different time-points or time-intervals: in such a case we group the equivalent rules and keep only one of them as representative of an input-output relation. Qualitative values of the variables of interest which differ only in the value of the output variable are likely generated. Then, it is probable that some conflict rules, i.e. rules

which have the same antecedents but a different consequent are produced. We store such rules and leave the conflicts be resolved in accordance with the degree of rule calculated on the data pairs (Wang 1994).

Remark 3. The entire range of possible system dynamics is captured by the whole tree of behaviors. Such a tree, as it includes all logically possible outcomes of the given qualitative information, may contain ambiguous results and may be quite large to be efficiently explored. Such ambiguities, as well as the dimension of the behavior tree, may be significantly reduced by additional knowledge. Moreover, let us remark that many behaviors in the tree may not present any difference as far as the input-output variables are concerned. Then, behavior aggregation procedures, which aim at taking the significant distinctions out of the tree, have to be performed to cope with the problem of an efficient generation of a complete and meaningful fuzzy rule base. Whenever reasonable physical assumptions allow us to define equivalent behavior classes, the portion of tree considered is further reduced by keeping one representative for each class.

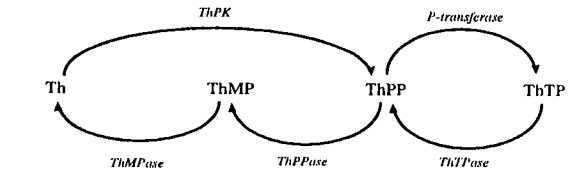
Given the behavior tree, the algorithm for the generation of the FBR which describes the overall system dynamics is performed as follows:

1. selection of an individual behavior;
2. generation of the FRB which corresponds to each individual behavior;
3. union of all the FRBs generated at step 2, and filtering of the equivalent rules.

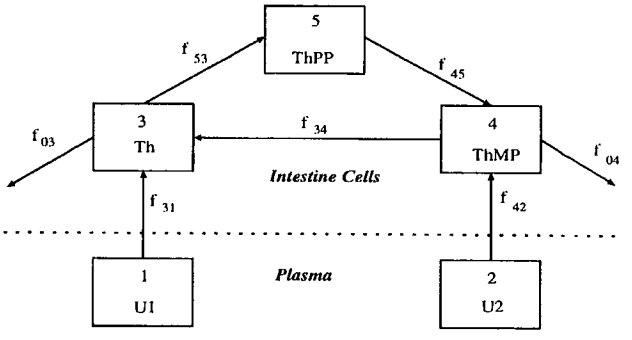
The FS-QM method in Biomedical applications

FS-QM can be applied to a variety of biomedical applications. As other black-box methods, it can be successfully used in problems of recognition (of signals or patterns) and prediction (as in adaptive control). In the latter area, we have studied the problem of predicting the Blood Glucose (BG) concentration in Diabetic patients (Bellazzi *et al.*). That work aimed at deriving a black-box predictor of the BG dynamics in response to exogenous insulin delivery and meal ingestion. The performance of the method was tested in a two-step experiment: a *training phase*, in which the approximator $y(\underline{x})$ was identified on a set of experimental data and a *validation phase*, in which the prediction accuracy of $y(\underline{x})$ was evaluated on a new data set. We obtained a one-step ahead predictor that can be used for therapeutic purposes and in on-line (adaptive) control applications.

FS-QM can be applied also in other contexts, like the approximation of the kinetic of molecules and drugs, whenever the relationships between variables are so complex that parameter identification in conventional modeling is very troublesome.



(A)



(B)

Figure 3: (A) - The Thiamine intracellular cycle. The arcs are labelled by the enzyme which is responsible for the chemical reaction. (B) - Compartmental model of the Thiamine kinetics in the intestine tissue.

In this paper we consider the problem of describing the dynamics of Thiamine (vitamin B_1) and its phosphoesters in the cells of the intestine tissue. Thiamine is transformed within the cells through an enzyme-mediated chemical reaction in a higher energy form that is used in the carbohydrates metabolism. The chemical reaction is nonlinear, and modeling through ordinary differential equations is hampered by identification problems. By using FS-QM we build an approximator of the system dynamics, that can be used to describe the different metabolic response of subjects with different pathological conditions.

Modeling intracellular Thiamine kinetics

Thiamine is transported in the extracellular fluids (blood and vessels) in two different chemical forms, simple Thiamine (Th) and Thiamine Mono-Phosphate (ThMP). As mentioned above, after passing through the cell membrane, the Th molecules are modified in order to be used in the carbohydrate metabolism. Fig. 3A shows the chemical reaction cycle: Th is directly transformed into the Thiamine Pyro-Phosphate (ThPP); ThPP is in equilibrium with Thiamine Tri-

Phosphate (ThTP). ThMP is firstly transformed into Th, and then from Th to ThPP. Finally, ThPP is dephosphorilated to ThMP. Only the lower energetic forms of Thiamine, Th and ThMP can be transported out of the cellular membrane and hence catabolized.

A simulator of the intracellular distribution of Thiamine can be of crucial help in describing syndromes with Thiamine deficiency, like severe liver diseases. A complete model of the human metabolism is quite difficult to derive, also because the data on intracellular Thiamine distribution cannot be measured 'in vivo'. So, the researchers in the field are trying to identify a model on animals that are known to have a human-like Thiamine metabolism.

As described above, chemical processes are mediated by enzymes; since the quantity of enzyme available in the time unit is limited, these reactions are saturable: the resulting models should hence be naturally nonlinear.

A compartmental model of the Thiamine kinetics in the intestine tissue is shown in Fig. 3B. In that model (Rindi *et al.* 1980), the ThTP form is not considered, since it is known that its quantity is quite small, and very difficult to be measured. The variables $U1$ and $U2$ represent the amount of Th and ThMP in the plasma, respectively. The flows (f_{ij}, $i = 0,..,5$, $j = 1,..,5$) (chemical reactions) express nonlinear relationships between quantities entering and leaving a compartment. In the literature, only linear compartmental models has been exploited to describe the Thiamine transformation process, since the nonlinear model is hardly identifiable. On the other hand, the models identified with the linear approximation are quite unsatisfactory in terms of curve fitting, and consequently as simulator of the system kinetics.

We aim at providing for an approximator of the intracellular Thiamine distribution that takes into account the overall complexity of the chemical reactions involved.

An approximator of the intracellular Thiamine kinetics

The data set available for the identification of the approximator is quite rich as we have data on each state variable of the model in Fig. 3B. We can completely express the Thiamine intracellular kinetics by subdividing the overall identification phase into the identification of three approximators as follows:

$$Th_{t+1} = \mathcal{Y}_1(Th_t, ThMP_t, U1_t) \quad (5)$$

$$ThMP_{t+1} = \mathcal{Y}_2(ThMP_t, ThPP_t, U2_t) \quad (6)$$

$$ThPP_{t+1} = \mathcal{Y}_3(ThPP_t, Th_t) \quad (7)$$

The approximators (5), (6), (7) can be easily exploited as a simulator when only the inputs ($U1$ and $U2$) are given. In the following we will describe in detail how the approximator of ThPP has been derived.

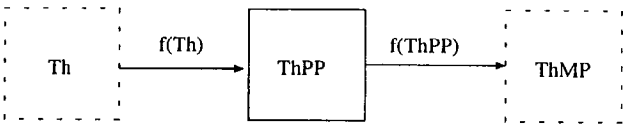


Figure 4: Compartmental model of the Th-ThPP pathway. $f(Th)$ and $f(thPP)$ are labeled in Fig. 3B as f_{53} and f_{45}, respectively.

Modeling the Th–ThPP pathway

The chemical reactions from Th to ThPP and from ThPP to ThMP can be modeled through a single compartment, as shown in the Fig. 4. Such reactions are denoted by $f(Th)$ and $f(ThPP)$. Th, simple Thiamine in the cells, acts as input to the subsystem considered. The qualitative model of the system dynamics is described by:

$$\frac{dThPP}{dt} = f(Th) - f(ThPP) \quad (8)$$

where

1. Th is a triangular shaped function, which represents the input signal. Its quantity space is defined as $(0\ Th^*\ \infty)$, where Th^* is the saturation threshold of $f(Th)$.
2. $ThPP$ has quantity space $(0\ ThPP^*\ \infty)$, where $ThPP^*$ is the saturation threshold of $f(ThPP)$.
3. Both $f(Th)$ and $f(ThPP)$ are represented in QSIM by the functional constraint S^+. From the physiology, we know that $f(Th^*) > f(ThPP^*)$.

Th			
QV	LV	$\hat{x}$ (nCi/g)	σ(nCi/g)
0	Zero	0	2
(0,Th*)	Low	15	7
Th*	Medium	35	8
(Th*,∞)	High	65	12

ThPP			
QV	LV	$\hat{x}$ (nCi/g)	σ(nCi/g)
0	Zero	0	8
(0,ThPP*)	Low	30	11
ThPP*	Medium	90	20
(ThPP*,∞)	High	170	30

Table 1: Correspondence between qualitative values (QV), linguistic values (LV) and fuzzy sets. $\hat{x}$ and σ denote Mean Values and Standard Deviations, respectively, which define Gaussian membership functions.

1. "If Th_t is Zero and $ThPP_t$ is Zero then $ThPP_{t+1}$ is Low"
2. "If Th_t is Low and $ThPP_t$ is Low then $ThPP_{t+1}$ is Medium"
3. "If Th_t is Low and $ThPP_t$ is Medium then $ThPP_{t+1}$ is High"
4. "If Th_t is Low and $ThPP_t$ is High then $ThPP_{t+1}$ is High"
5. "If Th_t is Medium and $ThPP_t$ is High then $ThPP_{t+1}$ is High"
6. "If Th_t is High and $ThPP_t$ is High then $ThPP_{t+1}$ is High"
7. "If Th_t is Low and $ThPP_t$ is High then $ThPP_{t+1}$ is Medium"
8. "If Th_t is Low and $ThPP_t$ is Medium then $ThPP_{t+1}$ is Low"
9. "If Th_t is Low and $ThPP_t$ is Low then $ThPP_{t+1}$ is Low"
10. "If Th_t is Zero and $ThPP_t$ is Low then $ThPP_{t+1}$ is Low"
11. "If Th_t is Zero and $ThPP_t$ is Low then $ThPP_{t+1}$ is Zero"
12. "If Th_t is Low and $ThPP_t$ is Low then $ThPP_{t+1}$ is Zero"
13. "If Th_t is Medium and $ThPP_t$ is Low then $ThPP_{t+1}$ is Low"
14. "If Th_t is High and $ThPP_t$ is Low then $ThPP_{t+1}$ is Medium"
15. "If Th_t is High and $ThPP_t$ is Medium then $ThPP_{t+1}$ is High"

Table 2: FRB derived from the simulated behaviors in Fig. 6

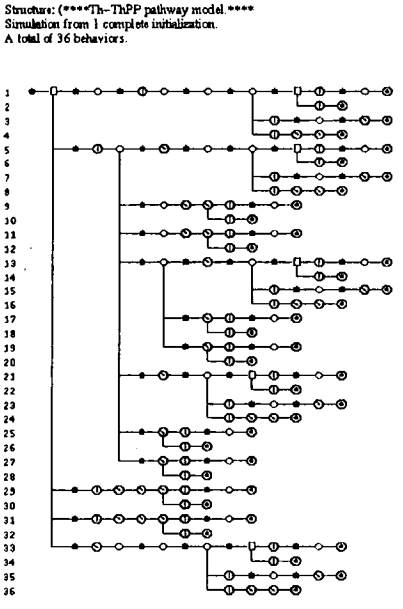


Figure 5: Qualitative simulation results of the $Th - ThPP$ pathway model: behavior tree.

The FRB automatically derived from the qualitative simulation of the model exploits Th_t and $ThPP_t$ as antecedents and $ThPP_{t+1}$ as consequent. Then, the quantity spaces of both Th and $ThPP$ are mapped into their fuzzy representation (see Table 1): the centers and the standard deviations of the membership functions have been derived on the basis of the available physiological knowledge.

Since the data used for the system identification come from experiments with tracers, the initial value of the variables involved in the simulation is set to 0.

The simulation of the QSIM model produces a tree of

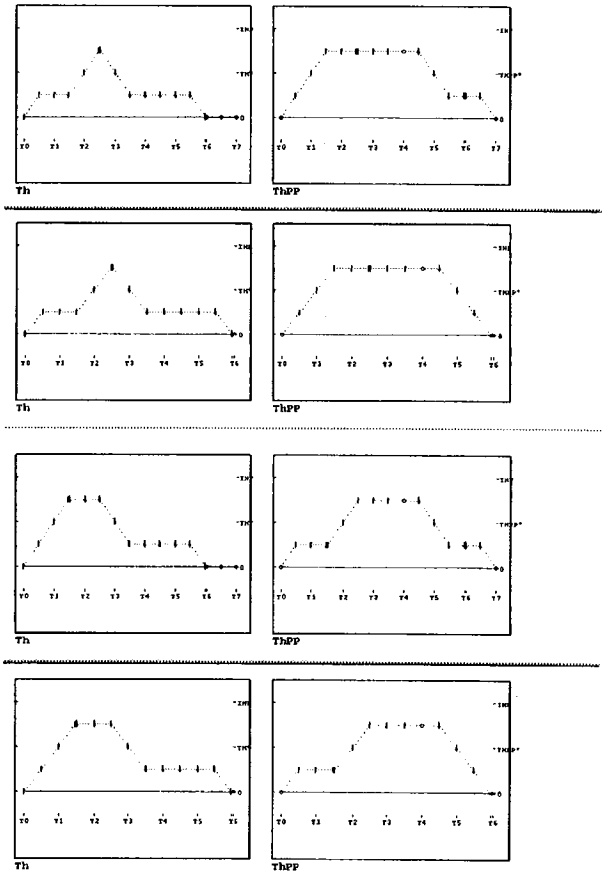


Figure 6: Qualitative simulation results of the $Th - ThPP$ pathway model: each row reports the plot of a representative behaviors of the dynamics of Th and $ThPP$.

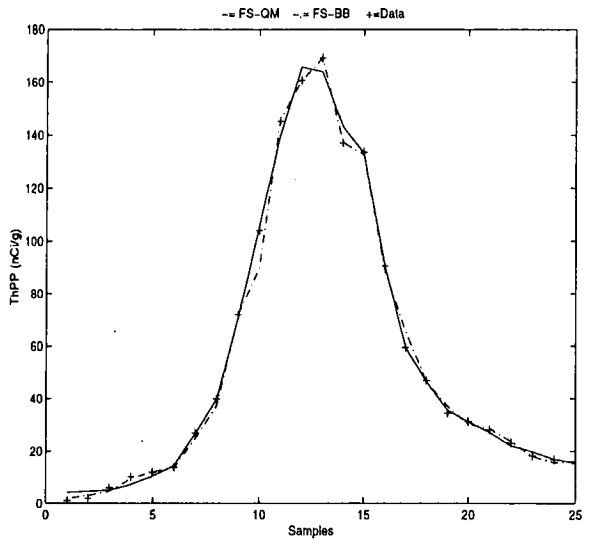


Figure 7: Identification phase - The plot compares the results obtained with 30 loops of BP by applying both FS-QM and FS-BB on the experimental data set.

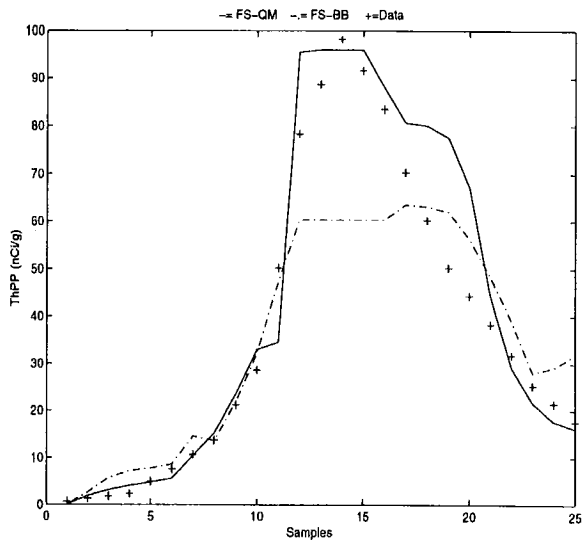


Figure 8: Forecasting phase - The plot compares the performance of the two approximators in forecasting the dynamics of $ThPP$ in diabetic rats. Both approximators have been obtained with 100 loops of BP.

36 quiescent behaviors (Fig. 5): 16 of them are filtered out as physiologically inconsistent with the hypothesis $f(Th^*) > f(ThPP^*)$. Among the remaining behaviors, four of them (Fig. 6) are representative of all the possible dynamics of Th and $ThPP$.

The complete FRB automatically derived from the qualitative behaviors is shown in Table 2.

Results

In this application problem, the main goal is to build a good simulator of the intracellular Thiamine dynamics. Such a goal has been achieved through two phases: an *identification phase*, in which the parameters of the derived FS are refined by using BP algorithm on a first set of data, and a *forecasting phase*, in which the identified FS has been used as simulator, and the results are matched for validation against a new data set.

The data used for identification come from an experiment on a group of rats whose intestine tissue was analyzed after an intravenous bolus of $30\mu g$ of thiazole-$[2^{14}C]$Thiamine, with a radioactivity of 1.25 μCi, for a period of 240 h, sampled with irregular time intervals.

The performance of the approximator $y(\underline{x})$ obtained by using FS-QM has been compared with that one identified by using only the data (denoted by FS-BB), as proposed by Wang (Wang 1994). The comparison has been performed through the calculation, in both L_2 and L_∞ norms, of the absolute errors between the data and the calculated values.

If the number of BP loops (nl) is kept low (≤ 25), we get absolutely better results of the FS-QM over FS-BB both in the identification and forecasting phase. For a greater number of loops, both methods show similar performance (Fig. 7): for $nl \rightarrow \infty$, the errors calculated go to zero in both cases with a slight difference in convergence velocity.

In spite of the comparable identification performance, in the forecasting phase $y(\underline{x})$ performs quite well, whereas the approximator obtained by FS-BB is not able to reproduce the data of $ThPP$ measured in different experimental settings. This is not surprising and may be explained by two occurrences: the data are noisy, and the number of samples is smaller than the number of parameters to be identified. This means that FS-BB, which learns only from data, is more likely driven to identify also the noise.

Fig. 8 and Fig. 9 compare the results obtained by both methods when applied to simulate the dynamics of $ThPP$ in response to two different input values for Th, which correspond to the values of Th injected on two groups of diabetic rats, one of them under therapy.

As it is highlighted by the plot comparison, the performance of the FS-QM approximator is good, while the FS-BB approximator does not reproduce at all the dynamics of $ThPP$.

Open problems and conclusion

The application of the presented framework to both the problem of predicting and simulating the dynamics of nonlinear systems has given good results which confirm its validity in terms of efficiency and robustness. However, several methodological problems are still open, and they will need further work. First of all, the method requires for a better mathematical formalization which defines its range of validity and ap-

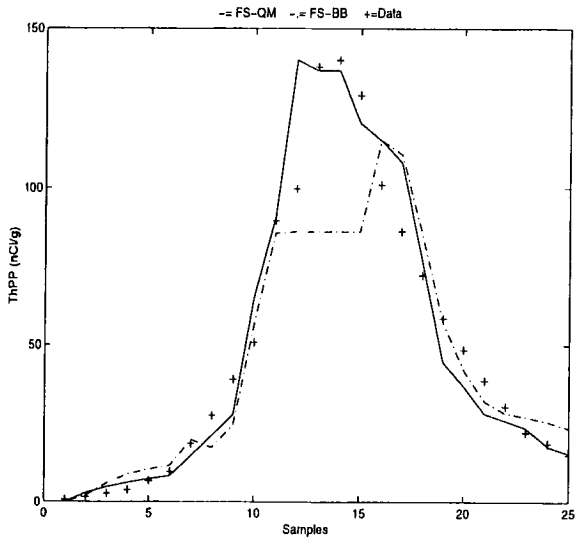


Figure 9: Forecasting phase - The plot compares the performance of the two approximators in forecasting the dynamics of $ThPP$ in diabetic rats under therapy. Both approximators have been obtained with 100 loops of BP.

plicability. Important issues which need a thorough study are listed below:

Membership functions
The choice of the membership functions is an important step of the design of a FS, and becomes crucial when the FS has to be tuned on a set of experimental data. We have chosen Gaussian Membership Functions (G-MFs) as they possess the universal approximation properties (Wang 1994), and they are functions of only two parameters (mean and variance), that can be estimated using the BP-technique. Moreover, G-MFs provide for good generalization properties as they ensure completeness ($\forall \underline{x} \in U, \mu(\underline{x}) > 0$), even when the fuzzy rule base spans only a subset of the Cartesian products of the input space. On the other hand, G-MFs are symmetric and with a maximum value located in a single point: this means that some desirable properties, such as different shapes of the μ_i associated with either a landmark or an interval, cannot be represented. In order to preserve the advantages coming from G-MFs but to improve the capability of expressing prior knowledge, in the future we will investigate the use of pseudo-trapezoidal functions (Zeng & Singh 1996).

Time
The major problem deals with the definition of the mapping of the sampling time set into the qualitative time set, i.e. of the mapping of the measuring grid into the "event" one. This is feasible only if the experiment has been designed so that the data set is informative enough to produce the system dynamics. Such an issue would be facilitated if semi-quantitative information on qualitative times would be available.

Hybrid models
The more complete is the a-priori knowledge exploited in modeling, the more close to the solution is the initial guess generated, with a consequent improved efficiency of FS-QM. A semi-quantitative formulation and simulation of the model (Kuipers 1994; Shen & Leitch 1993) is hence preferable. Unfortunately, the quantitative information may be insufficient for a semi-quantitative formulation with the mentioned approaches. Therefore, methods for dealing with hybrid models, where different knowledge sources can coexist would be the ideal way to get as much information as possible from the prior knowledge.

Other identification procedures
If the parameters, $\hat{x}_i^j$ and σ_i^j, of the membership functions in the equation (3) are fixed to their prior values, the identification of the FS turns from a nonlinear problem into a linear one. This choice provides for a clear advantage from a computational point of view, and also allows us to preserve the structure of the FS initialized on the basis of the a–priori knowledge captured by the qualitative model. But, since the FS is fixed, this choice could prevent from identifying a really "good" approximation of the unknown function. Therefore, possible solutions could be given by two-step identification procedures.

References

Bellazzi, R.; Ironi, L.; Guglielmann, R.; and Stefanelli, M. Qualitative models and fuzzy systems: an integrated approach for learning from data. *Artificial Intelligence in Medicine* Special Issue on AIME '97 (to appear).

Bradley, E.; O'Gallagher, A.; and Rogers, J. 1997. Global solutions for nonlinear systems using qualitative reasoning. In Ironi, L., ed., *Proc. 11th International Workshop on Qualitative Reasoning*, 31–40. Cortona: Istituto di Analisi Numerica - C.N.R., Pavia.

Bradley, E. 1994. Automatic construction of accurate models of physical systems. In *Proc. 8th International Workshop on Qualitative Reasoning*, 13–23.

Capelo, A. C.; Ironi, L.; and Tentoni, S. 1996. The need for qualitative reasoning in automated modeling: a case study. In *Proc. 10th International Workshop on Qualitative Reasoning*, 32–39.

Capelo, A.; Ironi, L.; and Tentoni, S. 1998. Automated mathematical modeling from experimental data: an application to material science. *IEEE Trans. SMC* 28(3).

Haber, R., and Unbehauen, H. 1990. Structure identification of nonlinear dynamic systems. *Automatica* 26:651–677.

Jang, J. 1993. Anfis: Adaptive network based fuzzy inference system. *IEEE Trans. on Systems, Man and Cybernetics* 23:665–685.

Juditsky, A.; Hjalmarsson, H.; Benveniste, A.; Deylon, B.; Ljung, L.; Sjöberg, J.; and Zhang, Q. 1995. Nonlinear black-box models in system identification: Mathematical foundations. *Automatica* 31:1725–1750.

Kay, H. 1996. *Refining imprecise models and their behaviors.* Ph.D. Dissertation, The University of Texas at Austin.

Khannah, T. 1990. *Foundations of neural networks.* Reading, MA: Addison-Wesley.

Kim, H., and Mendel, J. 1995. Fuzzy basis functions: Comparison with other basis functions. *IEEE Trans. Fuzzy Systems* 3:158–168.

Kuipers, B. J. 1994. *Qualitative Reasoning: modeling and simulation with incomplete knowledge.* Cambridge MA: MIT Press.

Ljung, L. 1987. *System Identification - Theory for the User.* Englewood Cliffs: Prentice-Hall.

Mamdani, E. 1974. Applications of fuzzy algorithms for simple dynamic plant. *Proc. IEE* 121(12):1585–1588.

Rindi, G.; Patrini, C.; Comincioli, V.; and Reggiani, C. 1980. Thiamine content and turnover rates of some rat nervous region, using labeled thiamine as a tracer. *Brain Res.* 181:369–380.

Shen, Q., and Leitch, R. 1993. Fuzzy qualitative simulation. *IEEE Tran. on Systems, Man and Cybernetics* 23:1038–1061.

Söderström, T., and Stoica, P. 1989. *System Identification.* Cambridge: Prentice–Hall.

Vescovi, M., and Travé-Massuyès, L. 1992. A constructive approach to qualitative fuzzy simulation. In *Proc. 6th International Workshop on Qualitative Reasoning,* 268–280. Edinburgh: Heriot-Watt Univ.

Wang, L. 1994. *Adaptive Fuzzy Systems and Control: design and stability analysis.* University of California at Berkeley: Englewood Cliff, NJ:Prentice–Hall.

Zeng, X., and Singh, M. 1996. Approximation accuracy analysis of fuzzy systems as functions approximators. *IEEE Transactions on Fuzzy Systems* 4(1):44–63.