

Using A Hybrid Genetic Algorithm and Fuzzy Logic for Metabolic Modeling

John Yen and Bogju Lee

Center for Fuzzy Logic, Robotics, and Intelligent Systems
Department of Computer Science
Texas A&M University
College Station, TX77843-3112
yen@cs.tamu.edu

James C. Liao

Department of Chemical Engineering
Texas A&M University
College Station, TX77843-3122

Abstract

The identification of metabolic systems is a complex task due to the complexity of the system and limited knowledge about the model. Mathematical equations and ODE's have been used to capture the structure of the model, and the conventional optimization techniques have been used to identify the parameters of the model. In general, however, a pure mathematical formulation of the model is difficult due to parametric uncertainty and incomplete knowledge of mechanisms. In this paper, we propose a modeling approach that (1) uses fuzzy rule-based model to augment algebraic enzyme models that are incomplete, and (2) uses a hybrid genetic algorithm to identify uncertain parameters in the model. The hybrid genetic algorithm (GA) integrates a GA with the simplex method in functional optimization to improve the GA's convergence rate. We have applied this approach to modeling the rate of three enzyme reactions in *E. coli* central metabolism. The proposed modeling strategy allows (1) easy incorporation of qualitative insights into a pure mathematical model and (2) adaptive identification and optimization of key parameters to fit system behaviors observed in biochemical experiments.

Introduction

Very often, chemical reactions happen as a series of steps instead of as a single basic action. Therefore, a chemical research problem has been to capture or describe the series of steps called *pathway* of a chemical reaction. To do this, chemical engineers perform experiments with the reaction: measure the overall stoichiometry, detect reaction intermediates, hypothesize relations among the products, plot concentrations over time, and so on. A classic example of this in biomodeling is the pathway of glucose metabolic model which is shown in Figure 1. Each node describes a metabolite participating in the pathway, while each reaction is shown in the pathway as an arrow, which is labeled by the variable V denoting the rate of the reaction.

Extensive studies have unveiled numerous functions crucial to living cells, such as metabolic pathways, en-

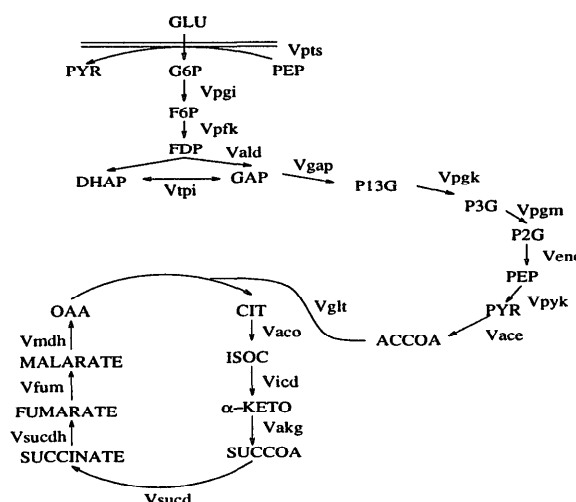


Figure 1: Pathway of glucose metabolic model

zyme actions, gene regulations, and global physiological controls. Several attempts have been reported to simulate or predict system behavior based on individual component models. For example, enzyme kinetic equations have been derived and assembled to model metabolic pathways (Achs & Garfinkel 1977; Heinrich & Rapoport 1974; Liao *et al.* 1988); components of DNA replication and gene expression have been modeled to simulate the replication of plasmids (Lee & Baily 1984; Straus, Walter, & Gross 1988); and key aspects of cellular functions have been represented mathematically to describe the overall cellular behavior (Schuler & Domach 1983). On the other hand, descriptive models either unstructured, structured, or based on optimization principles have been developed (Fredrickson 1976; Kompala, Ramkrishna, & Tsao 1984; Ramkrishna 1983). As a consequence of the reductionist approach and the fast progress of molecular biology, mechanisms at the molecular level are reasonably well established. These molecular mechanisms are combined to explain system behavior, most often in an intuitive manner. For example, reg-

ulation of enzyme activity has been used to explain the regulation of metabolic pathways, and the action of each component in a gene regulation network, or regulon, is used to explain the overall response of the network.

This intuitive approach has been successful to the extent of first approximation, but has rapidly become unsatisfactory as one demands a detailed explanation of system behavior. Furthermore, when an explanation based on intuitive synthesis of molecular mechanisms fails, it is difficult to determine whether the observation is a manifestation of novel molecular mechanism or is a complex interaction of known mechanisms. Some observations cannot be explained simply by intuitive synthesis of existing mechanisms. In general, complete mechanistic models are rare because of parametric uncertainty and incomplete knowledge of mechanisms, whereas descriptive models lack the ability to link component properties to system behavior. Moreover, when model prediction does not agree with experimental observations, it is difficult to distinguish between errors in parameters and errors in model structures.

In this paper, we propose a modeling approach that (1) uses fuzzy rule-based model to augment algebraic enzyme models that are incomplete, and (2) uses a hybrid genetic algorithm to identify uncertain parameters in the model. We have applied this approach to modeling the rate of enzyme reactions in *E. coli* central metabolism. The proposed modeling strategy allows (1) easy incorporation of qualitative insights into a pure mathematical model and (2) adaptive identification and optimization of key parameters to fit system behaviors observed in biochemical experiments.

The next section describes the basics of fuzzy logic-based modeling and the hybrid genetic algorithm (GA), which integrates a GA with the simplex method in functional optimization to improve the GA's convergence rate. We then describe the proposed modeling strategy and its application to modeling *E. coli* central metabolism. Finally, we discuss issues to be addressed in our future research and make some concluding remarks in section .

Background

Fuzzy Logic-based Modeling

It has been demonstrated that fuzzy modeling can be used to model complex systems that are not well understood (Takagi & Sugeno 1985; Sugeno & Kang 1988). The main contribution of fuzzy logic to system modeling is to introduce a new paradigm of modeling through three fundamental concepts that are closely related: *fuzzy partition*, *fuzzy rules* and *interpolative reasoning*. A fuzzy partition divides an input space to partially overlapping regions using fuzzy sets. Each subregion is associated with a local model for the region through a fuzzy rule. In areas where subregions partially overlap, the corresponding local models are combined to form a global model through a process (called interpolative

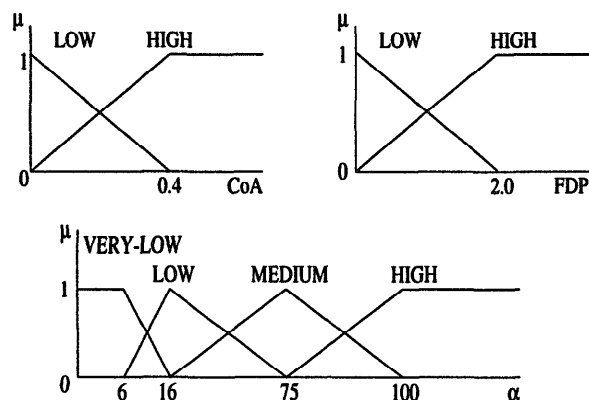


Figure 2: Fuzzy sets for V_{ppc} modeling

reasoning or fuzzy inference) that is analogous to linear interpolation.

A fuzzy partition generalizes classical partitions and divides a space into a collection of disjoint subspaces to allow smooth transitions from one subspace into a neighboring one. This is accomplished using fuzzy sets, which were developed by Lofti A. Zadeh to allow objects to take partial membership in a vague concept (i.e., a concept without sharp boundaries) (Zadeh 1965). The degree to which an object belongs to a fuzzy set, which is a real number between 0 and 1, is called the *membership value* in the set. The meaning of a fuzzy set is thus characterized by a *membership function* that maps elements in a universe of discourse (i.e., the domain of interest) to their corresponding membership values. Figure 2 shows the membership functions of the fuzzy sets for modeling enzyme PPC. CoA represents acetyl-coA, and μ denotes the membership value in the fuzzy sets.

Based on fuzzy set theory, fuzzy logic generalizes modus ponens in classical logic to allow a conclusion to be drawn from a fuzzy if-then rule even when the rule's condition is partially satisfied (Zadeh 1973). The strength of the conclusion is calculated based on the degree to which the antecedent is satisfied by the input data. Conclusions from multiple fuzzy rules are then combined to form a global conclusion. This is the essence of the *interpolative reasoning*.

There are two kinds of fuzzy rule. The first kind of fuzzy model, referred to as the Sugeno-Takagi-Kang model in the literature, uses a linear equation to describe a rule's local model. An example of this type of rule is shown below for a system with two input variables (x , y) and one output variable (z):

$$\text{If } x \text{ is } A \text{ and } y \text{ is } B \text{ then } z = a_0 + a_1x + a_2y$$

where A and B denote fuzzy sets and a_0 , a_1 and a_2 denote constants. Let w_i denotes the degree the input to the model matches the condition of the i -th rule, and y_i denotes the conclusion of the i -th rule. The formula below combines the conclusion of all rules in a Sugeno-

Takagi-Kang model through interpolative reasoning:

$$y = \sum_i w_i y_i / \sum_i w_i$$

The second type of fuzzy rule maps a fuzzy subregion to a fuzzy conclusion as shown below:

If x is A and y is B then z is C

The interpolative reasoning process for this kind of rule is analogous to that of Sugeno-Takagi-Kang fuzzy model. Degree of matching in the premise of a rule is propagated to the consequent to form an inferred fuzzy subsets. These fuzzy subsets are combined and defuzzified if necessary. Both types of fuzzy rule is used in the proposed modeling approach.

Compared to other approximation technique (e.g., piecewise linear approximation, spline, etc.), a fuzzy model is simpler to develop, easier to understand, and more flexible in providing a smooth approximation to a complex nonlinear relationship.

Genetic Algorithms

Genetic algorithms are global search and optimization techniques modeled from natural genetics, exploring search space by incorporating a set of candidate solutions in parallel (Holland 1975). A genetic algorithm (GA) maintains a *population* of candidate solutions where each solution is usually coded as a binary string called a *chromosome*. A chromosome – also referred to as a *genotype* – encodes a parameter set (i.e., a candidate solution) for a set of variables being optimized. Each encoded parameter in a chromosome is called a *gene*. A decoded parameter set is called a *phenotype*. A set of chromosomes forms a population, which is evaluated and ranked by a *fitness evaluation* function. The initial population is usually generated at random.

The evolution from one generation to the next one involves mainly three steps. First, the current population is evaluated using the fitness evaluation function, then ranked based on their fitness values. Second, GA stochastically select “parents” from the current population with a bias that better chromosomes are more likely to be selected. This is accomplished using a selection probability that is determined by the fitness value or the ranking of a chromosome. Third, the GA reproduces “children” from selected “parents” using two *genetic operations*: crossover and mutation. This cycle of evaluation, selection, and reproduction terminates when an acceptable solution is found, when a convergence criterion is met, or when a predetermined limit on the number of iterations is reached. The GA has been shown to be an effective search techniques on a wide range of difficult optimization problems (Dejong 1975; Holland 1975). The randomness and parallelism of GA often enable it to find a global optimum without being trapped in a local optimum. The GA has been proved to outperform conventional gradient

search technique on difficult problems involving discontinuous, noisy, high dimensional, and multimodal objective functions (Goldberg 1989).

However, the computational cost of a GA to find a global optimum is typically very high. That is, it usually requires a large number of generations before it converges to an acceptable solution. This issue is especially important for applying a GA to the parameter identification of metabolic and physiological systems due to the high computational cost of the fitness evaluation function. To evaluate a particular guess for a set of parameters in a model for such systems, one needs to (1) simulate the model based on the guessed parameters, and (2) calculate the error between simulation result and the experimental data. Even though efficient simulation packages are available, the computational cost of simulating many (e.g., a hundred) complex models for thousands of generations is extremely high.

To reduce the computational cost of GA-based approaches to the identification of parameters for metabolic systems, we have developed a hybrid approach that integrates the GA and the simplex method to speed up the rate of convergence while avoiding being easily entrapped at a local optimum (Yen *et al.* 1995). This is described in the next two sections in detail.

Simplex Method

Simplex method is a local search technique that uses the evaluation of the current data set to determine the promising direction of search. The simplex method was first introduced by Spendley *et al.* (Spendley, Hext, & Himsworth 1962) and later modified by Nelder and Mead (Nelder & Mead 1965). A simplex is defined by $N + 1$ points where N is the dimension of the search space. The method continuously forms new simplices by replacing the worst point in the simplex with a new point generated by reflecting the worst point over the centroid of the remaining points. This cycle of evaluation and reflection iterates until the simplex converges to an optimum.

We chose the simplex method rather than a gradient-based method (e.g., steepest descent, Newton strategies) as the local search technique for our hybrid GA because the relationships between the modeling parameters and the modeling objectives (i.e., close fitness between the model prediction and the experimental data) are too complex to be formulated. Consequently, it is difficult to compute the derivatives needed by the gradient-based methods.

A Hybrid Genetic Algorithm Using Simplex Method

We developed a hybrid GA method by introducing the simplex method as an additional local search operator in the genetic algorithm (Yen *et al.* 1995). The hybrid of the simplex method and the genetic algorithm

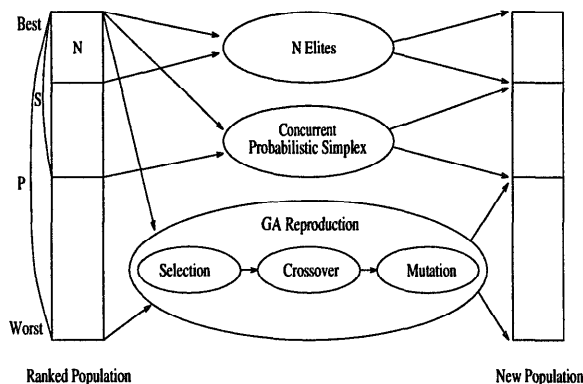


Figure 3: Reproduction in simplex-GA hybrid

applies the simplex method to the top S chromosomes in the GA population to produce $S - N$ children. The top N chromosomes are copied to the next generation. The remaining $P - S$ chromosomes are generated using the GA's reproduction scheme (i.e., selection, crossover, and mutation) where P is the population size in the GA. Figure 3 depicts the reproduction stage of the hybrid approach.

Empirical results obtained by applying the hybrid method to a subset of the biomodeling problem showed that the hybrid method outperformed the GA in terms of the speed of convergence and the quality of solution (Yen *et al.* 1995).

Metabolic Modeling

The Proposed Modeling Strategy

The proposed modeling strategy treats the system at two levels: a *component (molecular) level* such as enzyme reactions, protein-protein interactions and protein-DNA binding, and a *system level* such as metabolic networks, signal transduction pathways, genetic regulation systems, and global responses. In the component level, behavior of the components is described by algebraic equations derived from known molecular mechanisms and/or fuzzy logic models based on descriptive and/or incomplete information. Model parameters at this level are typically estimated from component data using the hybrid GA.

A component level model describes the rate of a reaction as an algebraic equation, whose structure reflects a known molecular mechanism. Several basic types of component level models are described in the next section. Parameters in these nonlinear models can be identified using a nonlinear optimization technique in system identification (e.g., extended Kalman filter, GA). We used the hybrid Simplex-GA to identify these parameters.

Although extensively investigated, mechanisms of many enzymes of interest are still only partially understood. Complete enzyme regulation mechanisms, such

as inhibition or activation, are often undetermined, except for a few enzymes with known crystal structure. Therefore, mechanisms describing substrate binding (e.g., random/ordered BiBi or Ping Pong BiBi) may be available, but mechanisms for inhibitor or activator actions are often incompletely known. We use fuzzy logic-based modeling to model the aspects of enzyme reactions that are not characterized mechanistically. More specifically, when experimental data suggests inhibition or activation factors not accounted for by a component model, we use fuzzy logic rules to augment the model by describing a mapping from the inhibiting or activation factors to their effects on the model. These rules are designed by first analyzing the inhibition and/or activation effects for identifying parameters in the original model whose values seem to change when inhibition or activation occurs. A set of fuzzy rules is then designed, each of which maps a specific inhibition or activation situation to a linear equation or a fuzzy set that characterizes the desired value of the parameter for the situation. Parameters in the algebraic equation as well as parameters in the fuzzy logic rules are identified using the hybrid GA.

The component models are then synthesized into system models based on known or hypothesized pathways. Models at this level consist of ordinary differential equations (ODE's). The system level introduces additional parameters into the model which are estimated from system behavior. The hypothesized mechanism usually defines the structure of model while leaving model parameters unspecified. We use the hybrid GA to identify the unspecified system parameters. The hybrid GA's fitness evaluation involves simulating the system-level model for each candidate system parameter set in the population. To simulate the behavior of the system, we used an existing simulation software, DDASAC (Caracotsios & Stewart 1984), for solving non-linear ODE's numerically.

In the two remaining sections, we focus our discussion on the component level modeling. A more detailed discussion on our system level modeling approach can be found in (Yen *et al.* 1995).

Mechanistic Modeling of Enzyme Kinetics

In this section, we introduce four types of mechanisms and the structure of their corresponding mathematical models derived from the mechanisms.

Since enzymes form complexes with their substrates, the rate of reaction is limited by the concentration of enzyme-substrate complex. When the level of a substrate is varied, the initial velocity with which the reaction begins is generally given by an equation. If the reversibility of the reaction is considered, the equation has the following form:

$$V = \frac{(V_s/K_s)S - (V_p/K_p)P}{1 + S/K_s + P/K_p}$$

where S and P are the substrate and product, respec-

tively, and K_s , K_p , and V_s , and V_p are kinetic parameters.

If the reaction involves two substrates and two products (*BiBi reaction*), as in many metabolic systems, the kinetic mechanism may involve ternary-complexes or binary complexes. In the former case, the binding may be either random or ordered, and the reaction rate can be expressed as:

$$V = V_{max} \frac{(A)(B)}{K_1 + K_A(B) + K_B(A) + (A)(B)}$$

where A and B are the concentrations of the two substrates, and V_{max} , K_1 , K_A , and K_B are parameters determined from the initial reaction rate experiments.

The binary-complex mechanism involves a covalent intermediate as the enzyme goes to a modified form. In this case, substrate A first reacts and modifies the enzyme, producing the first product P . Then the second substrate B reacts with the modified enzyme, producing the second product Q . This mechanism is termed *Ping-Pong BiBi* reaction and the reaction rate can be expressed as:

$$V = V_{max} \frac{(A)(B)}{K_A(B) + K_B(A) + (A)(B)}$$

Very often, the reaction rates are inhibited or activated by products, substrates, or other metabolites not participating in the reaction. When such allosteric effects exist, the *Monod-Wyman-Changeux* (MWC) model or its variations can be used. The reaction rates involving inhibition and activation are described by:

$$V = \frac{A(1+A)^{n-1} + LcA(1+cA)^{n-1}}{(1+A)^n + L(1+cA)^n}, L = \frac{L_1(1+B)^n}{(1+C)^n}$$

We will refer to some of these models in the next section.

Integrating Fuzzy Logic with Mechanistic Modeling

For enzymes with incomplete mechanisms, fuzzy models are incorporated to mend the deficiency of the incomplete mechanistic model. An example is the PPC reaction. The dots in Figure 4 summarized the experimental data in the literature about the reaction rate of PPC in different PEP concentration with different activators (Izui *et al.* 1981). The following observations can be made from the figure. (1) Without any activator, the reaction proceeds at a very low rate. (2) Acetyl-CoA is a very powerful activator. (3) FDP exhibits no activation alone. (4) FDP produced a strong synergistic activation with acetyl-CoA. Because these activations change the saturation reaction rate (V_{max}), we modify the original mechanistic equation with a fuzzy logic factor (α) into the following component model:

$$V_{ppc} = \alpha V_{max} \frac{PEP}{K_m + PEP}$$

The fuzzy factor α is modeled by the following four fuzzy rules:

If CoA is LOW and FDP is LOW then α is VERY-LOW
If CoA is LOW and FDP is HIGH then α is LOW
If CoA is HIGH and FDP is LOW then α is MEDIUM
If CoA is HIGH and FDP is HIGH then α is HIGH

where VERY-LOW, LOW, MEDIUM, HIGH are fuzzy sets. The membership functions of these fuzzy sets are shown in Figure 2.

The second example is PYK reaction. It is activated by FDP and inhibited by CoA and ATP (dot data in Figure 6, 7, and 8). This reaction is again modeled with the following mechanistic equation with fuzzy numbers F , L , and c which are determined by fuzzy if-then rules.

$$V_{pyk} = F \frac{PEP(1+PEP)^3 + L \times c \times PEP(1+c \times PEP)^3}{PEP(1+PEP)^4 + L(1+c \times PEP)^4}$$

Modeling of ATP, CoA inhibition is described by the fuzzy factor F which is determined by the following fuzzy rules.

If FDP is LOW and ATP is LOW and CoA is LOW then $F = F_1$
If FDP is LOW and ATP is LOW and CoA is HIGH then $F = F_2$
If FDP is LOW and ATP is HIGH and CoA is LOW then $F = F_3$
If FDP is LOW and ATP is HIGH and CoA is HIGH then $F = F_4$
If FDP is HIGH and ATP is LOW and CoA is LOW then $F = F_5$
If FDP is HIGH and ATP is LOW and CoA is HIGH then $F = F_6$
If FDP is HIGH and ATP is HIGH and CoA is LOW then $F = F_7$
If FDP is HIGH and ATP is HIGH and CoA is HIGH then $F = F_8$

where $F_1, F_2, F_3, F_4, F_5, F_6, F_7$, and F_8 are parameters in the fuzzy model. Modeling of FDP activation is achieved by introducing two fuzzy numbers c and L in the MWC equation. L and c are used for changing the shape of the curve in a MWC model (Cantor & Schimmel 1980).

If FDP is LOW then $c = c_1, L = L_1$

If FDP is HIGH then $c = c_2, L = L_2$

where c_1, c_2, L_1 , and L_2 constants.

Results

We applied the hybrid genetic algorithm to identify the parameters in the proposed model. The fitness of a candidate parameter set is the root means square error between the real experimental data reported in the literature and the candidate model by the GA. Figure 5 plots the fitness versus trials for modeling the reaction rate V_{ppc} . The behavior of the identified model is shown in Figure 4. The figure shows a good fit between dots representing real experimental data and the lines representing the prediction of the model identified.

Similarly, the behavior of the identified model for V_{pyk} and corresponding experimental data is shown in the figures 6, 7, and 8.

Summary

In this paper, we have proposed a novel methodology to integrate fuzzy logic techniques with mechanistic modeling method to model the component level and system level structures of metabolic systems. We also use a hybrid genetic algorithm to identify the key parameters of the model. The strategy here allows one

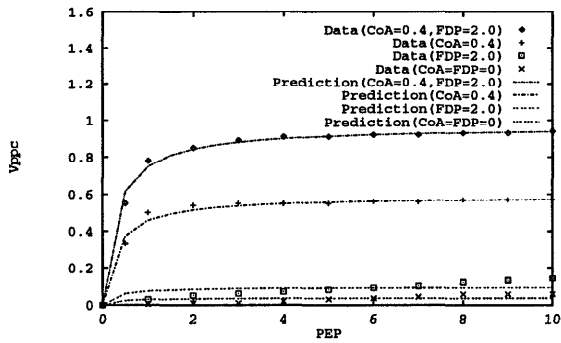


Figure 4: Data and model prediction for the reaction rate of PPC with activators

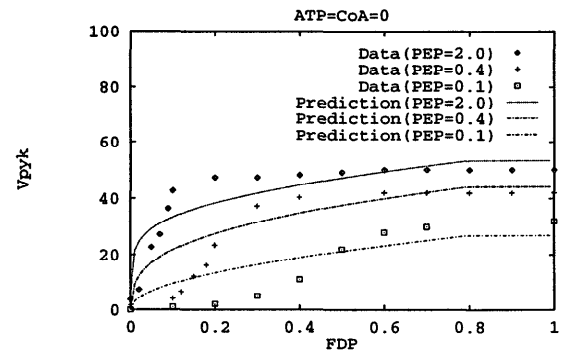


Figure 7: Data and model prediction for PEP activation in PYK reaction

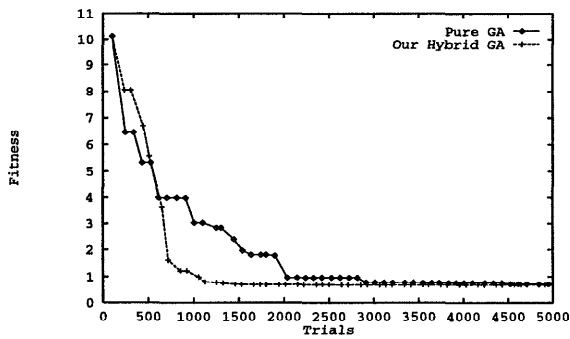


Figure 5: Performance of the hybrid GA on modeling V_{ppc}

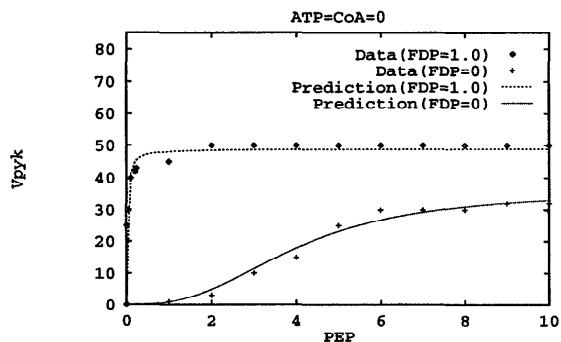


Figure 6: Data and model prediction for FDP activation in PYK reaction

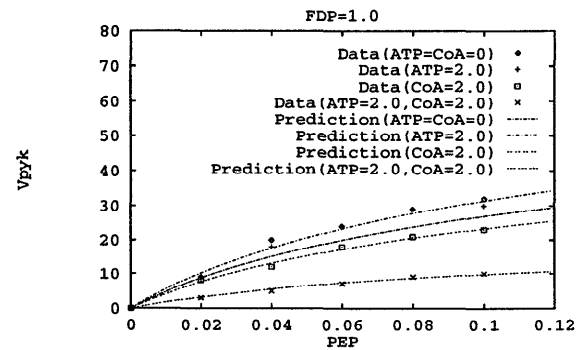


Figure 8: Data and model prediction for ATP and CoA inhibition in PYK reaction

to easily incorporate incomplete information and qualitative description into a mathematical formulation of the model. The modeling approach is promising for the elucidation of the unknown interactions between central metabolism and global regulation, which is essential for understanding biological signal transduction and rational design of metabolic systems for a desired purpose.

One of the most important issues remained to be addressed in our future research is to develop a scalable approach for dealing with the large search space at the system level, for the number of system parameters that may need to be adjusted to fit experimental data are typically very large. We are currently developing a supervisory architecture for dynamically selecting parameters to be optimized based on heuristics, insights about the model, and sensitivity analysis.

Acknowledgements

This research is currently supported by NSF Award BES-9511737 and was partially supported by NSF Young Investigator Awards IRI-9257293 and BCS-9257351. The software package for model simulation DDASAC was originated from M. Caracotsios and W. E. Stewart. The GENESIS implementation of a GA was developed by John J. Grefenstette.

References

Achs, M. J., and Garfinkel, D. 1977. Computer simulation of rat heart metabolism after adding glucose to the perfusate. *Am. J. Physiol.* 232:175-184.

Cantor, C. R., and Schimmel, P. R. 1980. *Biophysical Chemistry - Part III: The Behavior of Biological Macromolecules*. W. H. Freeman and Company.

Caracotsios, M., and Stewart, W. E. 1984. *DDASAC - Double precision Differential Algebraic Sensitivity Analysis Code*.

Dejong, K. A. 1975. *Analysis of the behavior of a class of genetic adaptive systems*. Ph.D. Dissertation, Department of Computer and Communication Sciences, University of Michigan.

Fredrickson, A. G. 1976. Formulation of structured growth models. *Biotechnol. Bioeng.* 18:1481-.

Goldberg, D. E. 1989. *Genetic Algorithms in Search, Optimization and Machine Learning*. MA: Addison-Wesley.

Heinrich, R., and Rapoport, T. A. 1974. A linear steady state treatment of enzymatic chains, general properties, control and effect strength. *Eur. J. Biochem.* 42:89-95.

Holland, J. H. 1975. *Adaptation in Natural and Artificial Systems*. Ann Arbor, MI: University of Michigan Press.

Izui, K.; Taguchi, M.; Morikawa, M.; and Katsuki, H. 1981. Regulation of escherichia coli phosphoenolpyruvate carboxylase by multiple effectors in vivo. *Journal of Biochemistry* 90:1321-1331.

Kompala, D. S.; Ramkrishna, D.; and Tsao, G. T. 1984. Cybernetic modeling of microbial growth on multiple substrates. *Biotechnol. Bioeng.* 26:1272-.

Lee, S. B., and Bailly, J. E. 1984. *Plasmid* 11:166-.

Lee, M., and Takagi, H. 1993. Integrating design stages of fuzzy systems using genetic algorithm. In *Proceedings of 2nd International Conference on Fuzzy Systems*.

Liao, J. C.; Lightfoot, E. N.; Jolly, S. O.; and Jacobson, G. K. 1988. Application of characteristic reaction paths: Rate-limiting capacity of phosphofructokinase in yeast fermentation. *Biotech. Bioeng.* 31:855-868.

Nelder, J. A., and Mead, R. 1965. A simplex method for function minimization. *Computer Journal* 7:308-313.

Ramkrishna, D. 1983. A cybernetic perspective of microbial growth. In Blanch, H. W.; Papoutsakis, E. T.; and Stephanopoulos, G., eds., *Foundations of Biochemical Engineering, American Chemical Society*. Washington, DC: American Chemical Society. 161.

Schuler, M. L., and Domach, M. M. 1983. Mathematical models of the growth of the individual cells. In Blanch, H. W.; Papoutsakis, E. T.; and Stephanopoulos, G., eds., *Foundations of Biochemical Engineering, American Chemical Society*. Washington, DC: American Chemical Society. 101.

Spendley, W.; Hext, G. R.; and Himsworth, F. R. 1962. Sequential application of simplex designs in optimization and evolutionary operation. *Technometrics* 4:441-461.

Straus, D. B.; Walter, W. A.; and Gross, C. A. 1988. Escherichia coli heat shock gene mutants are deficient in proteolysis. *Genes Dev.* 2:1851-1858.

Sugeno, M., and Kang, G. T. 1988. Structure identification of fuzzy model. *Fuzzy Sets and Systems* 28:315-334.

Takagi, T., and Sugeno, M. 1985. Fuzzy identification of systems and its applications to modeling and control. *IEEE Transactions on Systems, Man, and Cybernetics* 15:116-132.

Yen, J.; Liao, J. C.; Randolph, D.; and Lee, B. 1995. A hybrid approach to modeling metabolic systems using genetic algorithm and simplex method. In *Proceedings of the 11th IEEE Conference on Artificial Intelligence for Applications (CAIA95)*, 277-283.

Zadeh, L. A. 1965. Fuzzy sets. *Information Control* 8:338-353.

Zadeh, L. A. 1973. Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Transactions on Systems, Man, and Cybernetics* 3:28-44.