Efficient Numerical Optimization Algorithm Based on Genetic Algorithm for Inverse Problem

Daisuke Tominaga
Dept. of Biochem. Eng. & Sci.
Kyushu Institute of Technology
Fukuoka 820-8502, Japan

Nobuto Koga
Dept. of Biochem. Eng. & Sci.
Kyushu Institute of Technology
Fukuoka 820-8502, Japan

Masahiro Okamoto
Dept. of Biochem. Eng. & Sci.
Kyushu Institute of Technology
Fukuoka 820-8502, Japan

Abstract

We have developed an efficient algorithm based on the Genetic Algorithm(GA) for optimization of a model of a nonlinear system. Estimation of the interaction mechanisms among system components by using experimentally observed dynamic responses (time-courses) of some of the system components is generally referred to as "inverse problem". The S-system, which belongs to power-law formalism, is one of the best representations to solve such an inverse problem; the S-system is rich enough in structure to capture all relevant dynamics. In this paper, for the purpose of solving the inverse problem, we introduce the GA and propose an efficient procedure for the estimation of large numbers of parameters in the S-system formalism. We applied our method to a simple oscillatory system and a gene expression network.

1 INTRODUCTION

Organizationally complex systems such as gene expression networks and metabolic pathways are comprised of numerous, richly interacting components. In the case where the details of the molecular mechanism that govern interactions among system components (state variables) are not well known, however, how do we mathematically model such complex processes?; most of these processes are nonlinear. Description of these processes requires a representation that is general enough to capture the essence of the experimentally observed response. One of the best approaches that satisfy this requirement is the "S-system"(Savageau, 1976, Voit, 1991,Tominaga and Okamoto, 1998) which is a type of power-law formalism because it is based on a particular type of ordinary differential equation in which the component processes are characterized by power-law functions;

$$\frac{dX_i}{dt} = \alpha_i \prod_{j=1}^{n} X_j^{g_{ij}} - \beta_i \prod_{j=1}^{n} X_j^{h_{ij}} \qquad (1)$$

where n is the number of state variables or reactants (X_i), i,j $(1 \leq i,j \leq n)$ are suffixes of state variables. The terms g_{ij} and h_{ij} are interactive effectivity of X_j to X_i. The first term represents all influences that increase X_i, whereas the second term represents all influences that decrease X_i. In a biochemical engineering context, the non-negative parameters α_i and β_i are called rate constants, and real-valued exponents g_{ij} and h_{ij} are referred to as kinetic orders. The S-system is rich enough in structure to capture all relevant dynamics; an observed response(dynamic response) may be monotone or oscillatory, it may contain limit cycles or exhibit deterministic chaos. As long as it can be formulated as a system of ordinary differential equations, the S-system can be formulated as a canonical model. Furthermore, the simple homogeneous structure of S-system has a great advantage in terms of system analysis and control design, because the structure allows analytical and computational methods to be customized specifically for this structure(Irvine and Savageau, 1990). However, the S-system formalism has a major disadvantage in that this formalism includes a large number of parameters that must be estimated(α_i, β_i, g_{ij} and h_{ij}). The estimation of these parameters often causes a bottleneck, and matching the model to the experimentally observed responses(time courses of relevant state variables or reactants) is almost never straightforward and is almost always difficult. The number of estimated parameters in S-system formalism is $2n(n+1)$, where n is the number of state variables(X_i). In this paper, we should propose an algorithm and procedures for the estimation(optimization) of large numbers of parameters (Okamoto et.al., 1997, Tominaga et al., 1996, 1999). The basic idea is as follows: the Genetic Algorithm (GA)(Baker, 1985, Goldberg, 1989, Davis, 1991) as a nonlinear numerical optimization method which is much less likely to be stranded in local minima. Furthermore, in order to find the skeletal structure (small-size system) of S-system formalism that matches the experimentally observed responses, some of the parameters (g_{ij} and h_{ij}), absolute values of which are less than a given threshold value, are to be removed (reset to 0) during optimization procedures. By introducing this algorithm reffered to as structure skeletalizing

(Tominaga and Okamoto, 1998, Tominaga et al., 1999), that optimized essential S-system model that matches to the experimentally observed responses should be possible.

2 OPTIMIZATION PROCEDURES

Since the S-system is a formalism of ordinary nonlinear differential equation, the system can easily be solved numerically by using a suitable numerical calculation program such as Runge-Kutta-Gill method.

However, when an adequate time-course of relevant state variable is given, a set of parameter values α_i, β_i, g_{ij} and h_{ij}, in many cases, will not be uniquely determined, because it is highly possible that the other sets of parameter values will also show a similar time-course. Therefore, even if one set of parameter values that matches the observed time-courses is obtained, this set is still one of the best candidates that explain the observed time-courses. Our strategy is to explore and exploit these candidates within the immense huge searching space of parameter values.

In this optimization problem, each set of parameter values to be estimated is evaluated using the following procedure: Suppose that $X_{i,cal,t}$ is numerically calculated time-course at time t of state variable X_i and $X_{i,exp,t}$ represents the experimentally observed time-course at time t of X_i. Sum the relative error between $X_{i,cal,t}$ and $X_{i,exp,t}$ to get the total error f

$$f = \sum_{i=1}^{N} \sum_{t=1}^{T} \left\{ \left(\frac{X_{i,cal,t} - X_{i,exp,t}}{X_{i,exp,t}} \right)^2 \right\} \tag{2}$$

where N is the number of experimentally observable state variables, T is the number of sampling points of the experimental data. The problem is to find a set of parameters that minimizes f.

The proposed method is based on simple GA, and the structure of the genome(design code) of each individual(each set of parameter values) is shown in Figure 1.

A genome(corresponds to one individual) contains a set of S-system parameters (n α_is and β_is, and $n \times n$ g_{ij}s and h_{ij}s) which forms an $n \times (2n+2)$ matrix. An individual represents one S-system model. Each small square in Figure 1 corresponds to each parameter that has a real value. We introduced the following coding for representing real numbers: a 32bit unsigned integer format within a given searching region, that is, each dimensional region to be searched is divided into 2^{32} discrete points and is numbered using a unsigned

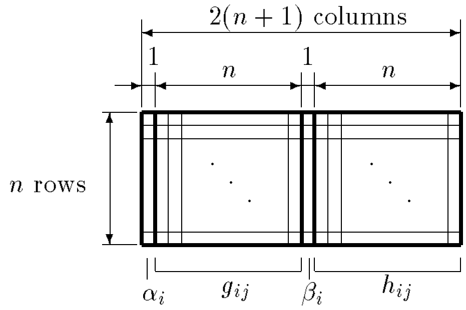

Figure 1: Design code of an individual; two n vectors of α_i and β_i, and two $n \times n$ matrices of $n \times n$ g_{ij} and h_{ij} form $n \times (2n+2)$ matrix. This matrix represents one S-system model.

integer. A real value within a searching region is represented by scaling a unsigned integer with offset.

2.1 GENETIC ALGORITHM

The optimization procedure in GA is as follows:
(0) Prepare a set of experimentally observed time-course data of n state variables. The number of sampling point(T) is common for each state variable. Determine the number of individual P, maximum limit of generation G_{max}, search regions for $\alpha_i(\beta_i)$ and $g_{ij}(h_{ij})$, initial mutation rate m_0, and the threshold of structure skeletalizing S for g_{ij} and h_{ij}.
(1) P initial guesses(P sets of α_i, β_i g_{ij} and h_{ij}) are randomly created. Each matrix element of an individual (Figure 1) has a real number and is randomly set within a given searching region.
(2) Evaluate each individual. Solve a simultaneous differential equation (Equation 1) for each individual, and calculate total error f using Equation 2. The fitness of each individual is calculated as the reciprocal of total error f.
Within the group of P-individuals in every generation, select the individual having the largest fitness(elite individual) and check whether the total error f of this individual is less than a given threshold value of convergence. If the largest fitness converges or the number of generations is beyond the specified count G_{max}, or if the fitness of elite individual does not change during $G_n(< G_{max})$ generations, exit the loop and terminate. The elite individual is reported as the best matched model.
(3) Otherwise continue searching by introducing genetic operations (crossover and mutation), as shown in Figure 2. Elite strategy is introduced over the generations.

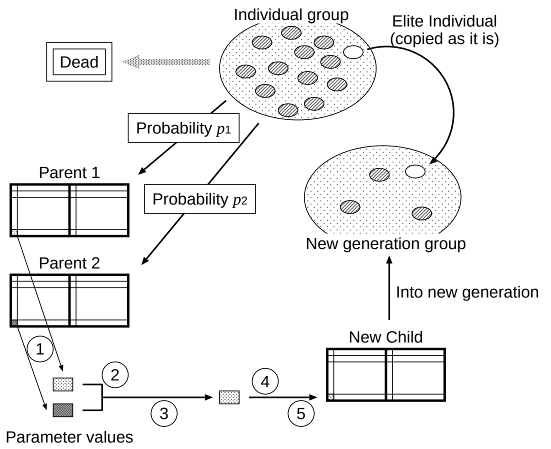

Figure 2: Crossover, mutation and structure skeletalizing. In order to create a child individual for the next generation, select two individuals from the individual group according to the probabilities proportional to their fitness. Genes of a child are created from the parents' genes by crossover. For each gene (S-system parameter), (1)Pick the gene from either parents. (2)(3)Determine which gene will be written into the child, by using uniform random number. (4)Mutation is applied to the gene with a certain probability (mutation rate), using normal distributed random number. (5)If the value of the gene is less than a given threshold, the value is reset to zero(in every specified generation) (structure skeletalizing).

(4) With the exception of the elite individual, the design code of each child is created based on the design codes of two parents. Two parents are selected by the roulette-wheel selection with ranking strategy. For each individual, the probability to be selected(p_i) is calculated by the following equation:

$$p_i = \frac{1}{N}\left(\eta^+ - (\eta^+ - \eta^-)\frac{i-1}{N-1}\right) \qquad (1 \le i \le N) \tag{3}$$

where N is the number of individuals, η^+ and η^- are the maximun and minimum probability of selection. We set the η^+ and η^- to 1.1 and 0.9, respectively.
(5) From the two selected parents, each matrix element of the child is alternatively chosen from the corresponding elements of either parent at the same probability. Since the number of matrix elements is $2n(n+1)$, where n is the number of state variables, this operation corresponds to the $2n(n+1)-1$ point crossover. The value of each matrix element is not changed by crossover.
(6) Furthermore, miss-writing (mutation) is supposed to occurr at mutation rate m when the matrix elements of the parents are copied to a child. We introduced Gaussian mutation in which average and distribution are defined by the original value and d, respectively. Set d to d_0 at the initial generation, and change the value according to the following conditions: (i)When the fitness of the elite individual does not change during G_m generations, change to $d_1(< d_0)$. (ii)After the change, if the fitness of the elite individual does not increase during next G_d generations, change to $qd_0(q > 1)$. (iii)If the fitness is not improved during next G_d generations, the value of distribution d is returned to d_0. (iv)When the fitness of the elite individual increase due to the (i)-(iii) strategies, the distribution d is reset to d_0. The mutation rate m is changed to k-times($k > 1$) when the fitness of the elite individual does not go up during G_m generations.
(7) By repeating the procedures(4)-(6), $(P-1)$ children are created for next generation. One child is the elite individual whose design code is not submitted to the genetic operations.
(8) Structure skeletalizing; the details are described in the following section.
(9) Return to step (2).

2.2 STRUCTURE SKELETALIZING

In S-system formalism, the number of estimated parameters increases with the order of n^2, which leads to a considerably large computational cost and ambiguous capture of interactive mechanism among state variables (X_i). In order to capture optimized essential interactions among X_i, the following structure skeletalizing procedure is introduced: 1)The term g_{ij} (h_{ij}), absolute value of which is less than a given threshold value (skeletalizing value) is reset to zero. 2)This procedure is performed at every specified generation and at the termination of optimization.

3 APPLICATIONS

In order to examine the effectiveness of the proposed procedures, we have applied our method to determine a set of parameters of the S-system.

3.1 OSCILLATORY SYSTEM

Figure 3(A) shows the calculated time-courses of X_1 and X_2 of the S-system($N = 2$ in Equation 1), parameter values of which are shown in Table 1. Given these time-courses of X_1 and X_2(Figure 3(A) is presented as experimental data($X_{i,exp,t}$, $N = 2$, $T = 50$ in Equation 2)), we examined whether our proposed optimization procedures could explore and exploit a set of parameter values which can provide the best fitted time-course shown in Figure 3(A). The major optimizing conditions are as follows: searching ranges are [0.0, 5.0] for α_i and β_i, [-3.0, 3.0] for g_{ij} and h_{ij}, $T = 50$,

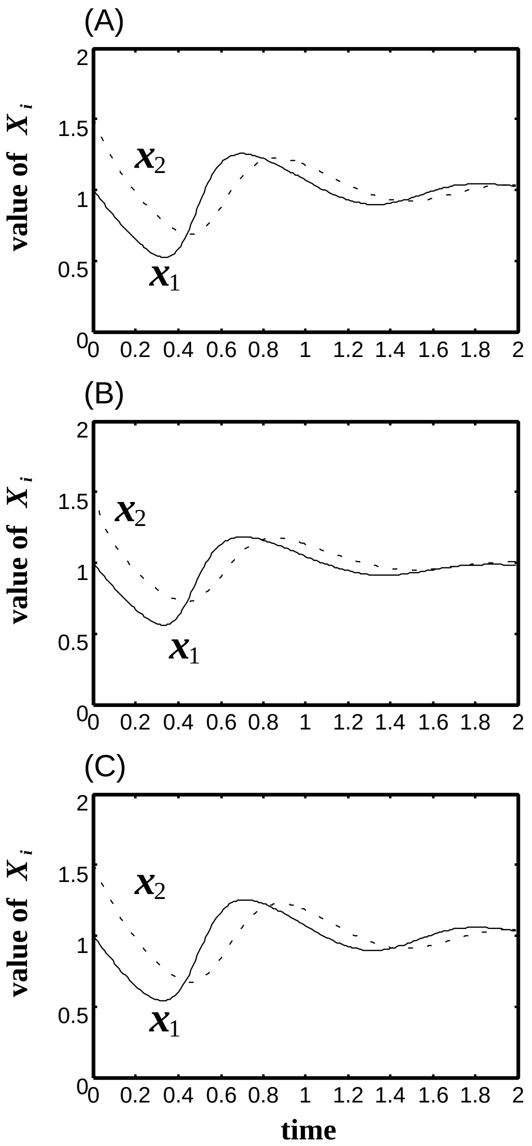

Figure 3: Given time-courses and obtained time-courses of X_1 and X_2. (A), Given time-courses(parameter value are listed in Table 1); (B), obtained time-courses (a set of parameters is listed in Table 2.) (C), obtained time-courses (a set of parameters is listed in Table 4.) Initial values of X_1 and X_2 is 1.0 and 1.5 respectively. In (A), the number of sampling points of experimental data is 50 ($T = 50$ in Equation (2).)

$P = 10000$, $G_{max} = 1000$, $G_n = G_{max}/2$, $m_0 = 0.05$, $G_m = 20$, $d_0 = 4.0$(for α_i and β_i), 1.6(for g_{ij} and h_{ij}), $d_1 = 0.5$(for α_i and β_i), 0.2(for g_{ij} and h_{ij}), $G_d = 10$, $q = 5$, $k = 1.01$, and the structure skeletalizing are performed at every generation and the threshold value of which is 0.5.

Figure 3(B) shows the obtained time-courses under the above optimization condition. The obtained parameter values are listed in Table 2. Compared the time-courses in Figure 3(B) with those in Figure 3(A), the overshoot peak values of X_1 and X_2 in 3(B) is less than those in 3(A), and the time-courses after $t = 1.0$ (arbitrary unit) of X_1 and X_2 in 3(B) are more monotonous than those in 3(A). The total dynamic patterns in 3(B) are, however, very similar to those in 3(A); the average relative error between calculated (3(B)) and given time-courses (3(A)) per sampling point is 3.65% (see the column of algorithm A in Table 3).

Since this result shows that the optimization method might be stranded in local minima, we have improved our algorithm in order to explore and exploit the better parameter space. In the above case, since the number of individuals in one generation is very large ($P = 10000$), most of the individuals the fitness value of which is small might be disappeared by ranking strategy, which leads to the survival of the large numbers of individuals which have very similar S-system parameters. In order to avoid this situation, we divided the individuals into many small groups. We made 9 small groups which consists of 100 inidividuals. The alternative algorithm is as follows: First, each small group independently performs optimization according to the proposed algorithm except for $P = 100$, $G_{max} = 10000$. After the independent searching, 9 best individuals are obtained from each small group. Prepared 91 individuals except for 9 best individuals, we make the final group (the number of individuals (P)$= 91 + 9 = 100$) which is allowed to the optimization according to the proposed algorithm ($P = 100$, $G_{max} = 10000$). The total numbers of created individuals is, therefore, $9 \times 100 \times 10000 + 1 \times 100 \times 10000 = 10^7$, which is the same as in the case of one large group; $P \times G_{max} = 10000 \times 1000 = 10^7$.

The obtained parameter set by the revised algorithm is listed in Table 4, and the calculated time-courses are shown in Figure 3(C). The oscillatory profile in Figure 3(C) is quite similar to that in 3(A); the average of relative error between calculated (3(C)) and given time-courses (3(A)) per sampling point is 1.39% (see the column of algorithm B). Compared the obtained S-system parameters (Table 4) with those in Table 1, the structure (sign and magnitude of parameter) is very

Table 1: Given S-system parameters which provide the dynamics shown in Figure 3(A).

i	α_i	g_{i1}	g_{i2}	β_i	h_{i1}	h_{i2}
1	3.0	0.0	-2.5	3.0	-1.0	0.0
2	3.0	2.5	0.0	3.0	0.0	2.0

Table 2: Obtained S-system parameters which provide the dynamics shown in Figure 3(B).

i	α_i	g_{i1}	g_{i2}	β_i	h_{i1}	h_{i2}
1	3.17	-0.51	-2.46	3.10	1.58	0.00
2	3.81	2.17	0.83	3.60	0.00	2.88

similar.

Table 5 shows the obtained results for optimization with the original algorithm(algorithm A) and with the revised algorithm(algorithm B). The values in the table show the average value of 30 trials with the corresponding algorithm. The values in the parentheses represent standard deviations. These results show the superiority of the revised algorithm(algorithm B) to the original one(algorithm A).

3.2 GENE NETWORK

Next, the proposed method was applied to determine a set of parameters of the S-system which represents a typical model of a gene network.

As a case study, we created several sets of time series data, shown in Figure 4, which were numerically calculated using the scheme shown in Figure 5(Savageau, 1998). The S-system parameters in Figure 5 are shown

Table 3: Obtained results of optimization. A: One large group, B: Nine small groups. Time-course data which are calculated from the obtained models are shown in Figure 3. Time for optimization represents a cpu-time until the optimization was performed, which was measured on Tempest 2 (Concurrent Systems Inc., Japan (processor:Alpha 21164A, 600MHz, SPECfp95: 21.3, SPECint95: 18.6))

Algorithm	A	B
Average relative error(%)	3.65	1.39
Maximun fitness value	7.50	51.6
Total explored generations	1000	55186
Total created individuals	10^7	5.5×10^6
Time for optimization(sec)	58803	31677

Table 4: Obtained S-system parameters which provide the dynamics shown in Figure 3(C).

i	α_i	g_{i1}	g_{i2}	β_i	h_{i1}	h_{i2}
1	3.71	0.00	-1.91	3.74	-0.77	0.00
2	2.93	2.58	0.00	3.00	0.00	1.88

Table 5: Obtained results of optimization. A: One large group, B: Nine small groups Shown values are the average of 30 trial runs and numeral with parentheses represents standard deviations. Time for optimization shows a cpu-time(second) until the optimization was performed, which was measured on Tempest2(Concurrent Systems Inc., Japan(processor: Alpha 21164A, 600MHz, SPECfp95: 21.3, SPECint95: 18.6))

Algorithm	A	B
Fitness value	2.86(1.24)	17.55(12.5)
Time for optimization	58392(1061)	36459(4303)

in Table 6. The time series data in Figure 4(A) were calculated with the condition that $\alpha_1 = 15.0$, $\alpha_4 = 8.0$ in Table 6, which corresponds to the situation in which both gene 1 producing X_1(mRNA) and gene 4 producing X_4(mRNA) are active (wild-type).

In contrast, the time series data in Figures 4(B) and 4(C) were obtained under the condition that $\alpha_1 = 0$, $\alpha_4 = 8.0$ and that $\alpha_1 = 15.0$, $\alpha_4 = 0$, respectively; Figure 4(B) shows the case of the disruption of gene 1(corresponds to $\alpha_1 = 0$), and 4(C) is the case of the disruption of gene 4(corresponds to $\alpha_4 = 0$). The optimization task is as follows: Can the proposed algorithm explore and exploit the best S-system parameters matching the observed time courses shown in Figure 4(A)-(C)?

We attemped to estimate part of the system parameters in Table 6. The targets of optimization are twelve parameters; α_1, $g_{11}, \ldots, g_{15}$, α_4, $g_{41}, \ldots, g_{45}$. These represent the interaction coefficients that give increasing effects of X_1 and X_4(expression and regulation of mRNA), respectively. The major optimizing conditions are as follows: searching ranges are [0.0, 20.0] for α_i and β_i, [-3.0, 3.0] for g_{ij} and h_{ij}, $T = 50$, $P = 5000$, $G_{max} = 400$, $G_n = G_{max}/2$, $m_0 = 0.004$, $G_m = 10$, $d_0 = 4.0$(for α_i and β_i), 1.2(for g_{ij} and h_{ij}), $d_1 = 0.5$(for α_i and β_i), 0.15(for g_{ij} and h_{ij}), $G_d = 10$, $q = 5$, $k = 1.01$, and the structure skeletalizing were performed at every generation and threshold value of which is 0.05.

Table 7: Obtained results of optimization in gene network. The time for optimization was measured on Tempest 2 (Concurrent Systems Inc., Japan (processor:Alpha 21164A, 533MHz, SPECfp95: 20.1, SPECint95: 16.6))

Average relative error(%)	1.0×10^{-3}
Total explored generations	266
Time for optimization	39368sec

At the 266th generation(CPU-time is about 11 hours), we found the parameter set (α_1, g_{11}, g_{12}, g_{13}, g_{14}, g_{15}, α_4, g_{41}, g_{42}, g_{43}, g_{44}, g_{45}) to be completely identical to that shown in Table 6. In an early step of total optimization(at the 78th generation), a model which reflects the same structure as the given model was obtained, as shown in Table 8. The differences in parameter values between Table 6 and Table 8 are not remarkable, indicating that the structure of the system was obtained by global searching in early stages of the optimization, and local searching (fine tuning of parameters) was performed after the 78th generation.

4 DISCUSSION

Equation 2 corresponds to the cost function or objective function to be minimized during optimization. Since this function is not explicit but functional of the parameters to be estimited (α_i, β_i, g_{ij}, h_{ij}), simultaneous ordinary differential equations shown in Equation 1 must be solved numerically in order to get $X_{i,cal,t}$ in Equation 2. This suggests that because of the lack of numerical stability, the situation where differential equations with the estimated parameter set can not be numerically solved is highly possible. Since the fitness value can not be calculated in this case, the fitness value was supposed to be zero in this study. The GA escapes the parameter set which provides zero fitness, therefore, we have to explore and exploit the parameter set within the searching space having many local minima. In the application to the oscillatory system, we have proposed two algorithms; A)one large group ($P = 10000$, $G_{max} = 1000$) B)nine small groups ($P = 100$, $G_{max} = 10000$). As shown in Tables 3 and 5, it is revealed that the algorithm B) is more effective than the algorithm A); both average relative error and time for optimization in B) are superior to those in A). This indicates that plural small groups can explore more local minima in the searching space and have higher possibility to find better parameter set.

In the application to the gene network, the final op-

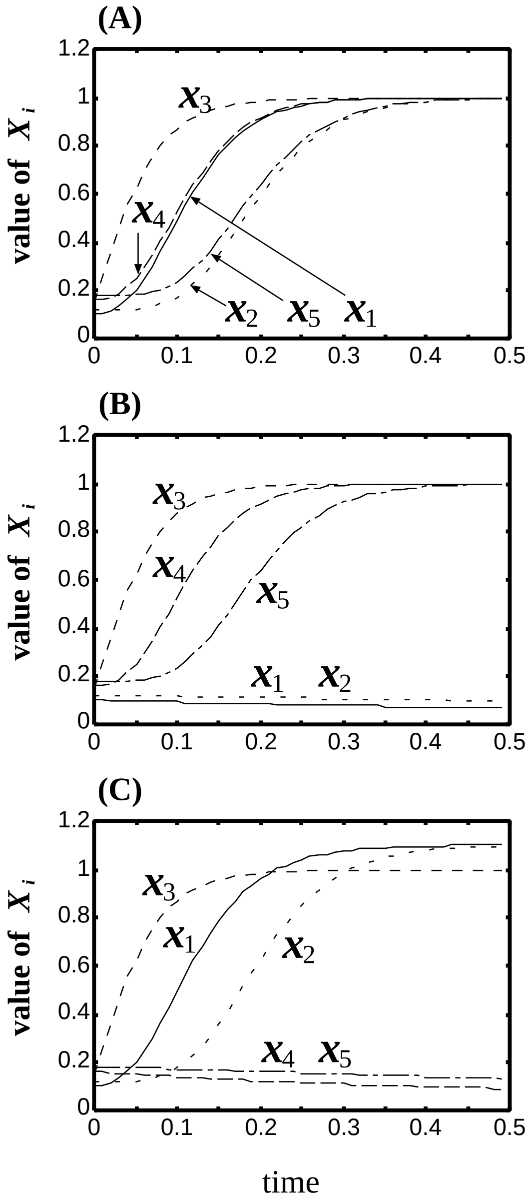

Figure 4: Time series data calculated from the S-system shown in Table 6; (A), wild type($\alpha_1 = 15.0$, $\alpha_4 = 8.0$); (B)disruption of gene 1($\alpha_1 = 0.0$, $\alpha_4 = 8.0$); (C)disruption of gene 4($\alpha_1 = 15.0$, $\alpha_4 = 0.0$); There are 51(sampling points) times 5(components) points in these time-courses data.

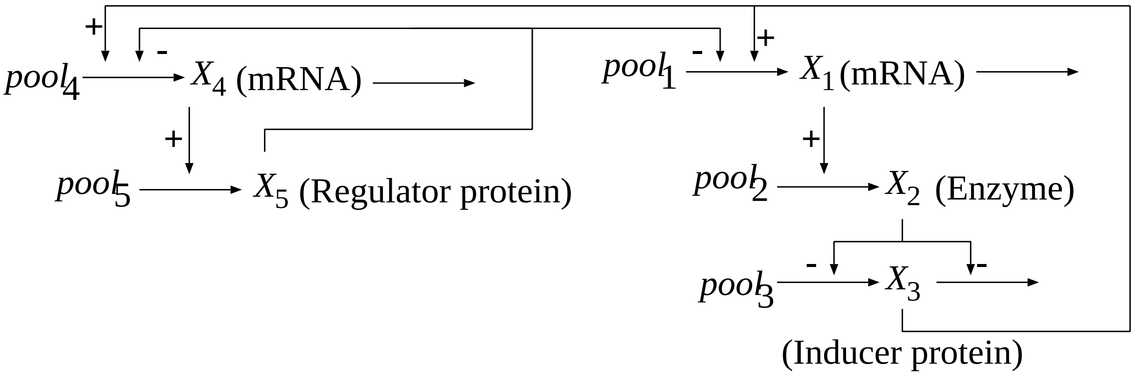

Figure 5: Typical type of gene network.

Table 6: Given S-system parameters in Figure 5. The pool size(constant value, $pool_1$ to $pool_5$ in the Figure 5) correspond to the values of α_1 to α_5.

i	α_i	g_{i1}	g_{i2}	g_{i3}	g_{i4}	g_{i5}	β_i	h_{i1}	h_{i2}	h_{i3}	h_{i4}	h_{i5}
1	15.0	0.0	0.0	1.0	0.0	-0.1	10.0	2.0	0.0	0.0	0.0	0.0
2	10.0	2.0	0.0	0.0	0.0	0.0	10.0	0.0	2.0	0.0	0.0	0.0
3	10.0	0.0	-0.1	0.0	0.0	0.0	10.0	0.0	-0.1	2.0	0.0	0.0
4	8.0	0.0	0.0	2.0	0.0	-1.0	10.0	0.0	0.0	0.0	2.0	0.0
5	10.0	0.0	0.0	0.0	2.0	0.0	10.0	0.0	0.0	0.0	0.0	2.0

Table 8: Obtained S-system parameters at the 78th generation in the gene network. The α_1, g_{11}, g_{12}, g_{13}, g_{14}, g_{15}, α_4, g_{41}, g_{42}, g_{43}, g_{44}, g_{45}, were optimized.

i	α_i	g_{i1}	g_{i2}	g_{i3}	g_{i4}	g_{i5}	β_i	h_{i1}	h_{i2}	h_{i3}	h_{i4}	h_{i5}
1	16.28	0.0	0.0	1.07	0.0	-0.08	10.0	2.0	0.0	0.0	0.0	0.0
2	10.0	2.0	0.0	0.0	0.0	0.0	10.0	0.0	2.0	0.0	0.0	0.0
3	10.0	0.0	-0.1	0.0	0.0	0.0	10.0	0.0	-0.1	2.0	0.0	0.0
4	8.55	0.0	0.0	2.09	0.0	-1.01	10.0	0.0	0.0	0.0	2.0	0.0
5	10.0	0.0	0.0	0.0	2.0	0.0	10.0	0.0	0.0	0.0	0.0	2.0

timized structure was found at the 266th generation and its essential structure was found at the 78th generation. The fitness value of the best model (obtained at the 266th generation) is 133.6. The relative error of calculated time series data to the given data (Figure 4(A)-(C)) per sampling point is about 1.0×10^{-3}(%). The fitness value and relative error at the 78th generation is 3.86 and 3.45×10^{-2}(%), respectively. The fitness was remarkably improved during the 78th and the 266th generations, however, the relative error per sampling point decreased slightly (from 1.0×10^{-3}(%) to 3.45×10^{-2}(%)).

For practical use, the experimentally observed data generally include a $\pm 10\%$ measurement error, which indicates that such large error range means the parameters obtained at the 78th generation are considered to be the best fitted parameters. In Figure 4(A)to Figure 4(C), we supposed several experimental conditions of gene network, such as wild-type, and disruption of gene. Estimation of parameters in the S-system using experimentally observed time-courses is generally referred to as inverse problem and these time-courses correspond to the restricted conditions for an inverse problem. Since the proposed algorithm proposes candidates (parameter sets) matching the restricted conditions, the best candidate can most likely be found by preparing more time-course data under the disruption and overexpressions in the gene network.

References

Baker,J.E.(1985). Adaptive selection methods for genetic algorithms, *Proceedings of the International Conference on Genetic Algorithms*, 101-111, Lawrene Erlbaum Associates.

Davis, L(1991). *Handbook of genetic algorithms*, Van Nostrand Reinhold.

Goldberg, D.E.(1989). Genetic Algorithms in Search, Optimization and Machine Learning, Addison Wesley.

Irvine, D.H, Savageau, M.A(1990). Efficient solution of nonlinear ordinary differential equations expressed in S-system canonical form, *SIAM JOURNAL OF NUMERICAL ANALYSIS* **27**, 704-735.

Okamoto, M., Morita, Y., Tominaga, D., Tanaka, K., Kinoshita, N., Ueno, J-I., Miura, Y.(1997). Toword a Virtual-Labo-System for Metabolic Engineering: Development of Biochemical Engineering System Analyzing Tool-Kit(BEST-KIT), *Proceedings of Pacific Symposium on Biocomputing '97*, 304-315.

Savageau, M.A.(1998) Rules for the evolution of gene circuitry, *Proceeding of Pacific Symposium on Biocomputing '98*, 54-65, World Scientific.

Savageau, M.A.(1976). *Biochemical system analysis: a study of function and design in molecular biology*, Addison-Wesley.

Tominaga, D., Ueno, J., Miura, Y., Okamoto, M.(1996). Discovery of a Skeletal Network Describing Complex Nonlinear Dynamics: Optimized Essential Model for Temporal Input-Output Matching, *Tutorials of the 4th International Conference on Soft Computing*, 65-79.

Tominaga, D., Okamoto, M.(1998). Design of Canonical Model Describing Complex Nonlinear Dynamics, *Proceeding of Computer Applications in Biotechnology 1998*, 85-90, Elsevier Science.

Tominaga, D., Okamoto, M., Maki, Y., Watanabe, S., Eguchi, Y.(1999). Nonlinear Numerical Optimization Technique Based on Genetic Algorithm for Inverse Problem: Towards the Inference of Genetic Networks, *Computer Science and Biology(Proceedings of the German Conference on Bioinfomatics)*, 127-140.

Tominaga, D., Ueno, J., Miura, Y., Okamoto, M. (1996). Discovery of a Skeletal Network Describing Complex Nonlinear Dynamics: Optimized Essential Model for Temporal Input-Output Matching, In: *Tutorials of the 4th International Conference on Soft Computing (IIZUKA 96)*, **2**, 65-80, Fuzzy Logic Systems Institute.

Voit, E.O.(1991). Canonical nonlinear modeling: S-system approach to understanding complexity, Van Nostrand Reinhold.