Evolutionary variance of gene network model via simulated annealing

Kyoungmin Roh
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd
Part of the Bioinformatics Commons

Recommended Citation
https://lib.dr.iastate.edu/rtd/15396

This Thesis is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
Evolutionary variance of gene network model via simulated annealing

by

Kyoungmin Roh

A thesis submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Bioinformatics and Computational Biology

Program of Study Committee:
Stephen Proulx, Major Professor
Zhijun Wu
Michael Smilely
Jeffrey Essner

Iowa State University
Ames, Iowa
2008

Copyright © Kyoungmin Roh, 2008. All rights reserved.
UMI Number: 1454703

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI Microform 1454703
Copyright2008 by ProQuest LLC
All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346
TABLE OF CONTENTS

LIST OF TABLES ................................................. iv

LIST OF FIGURES ................................................. v

ACKNOWLEDGEMENTS .............................................. viii

CHAPTER 1. GENERAL INTRODUCTION ......................... 1
  1.1 Evolutionary gene network ................................ 1
  1.2 Simulated annealing algorithm ............................. 2
  1.3 Reference .................................................. 3

CHAPTER 2. EVOLUTIONARY VARIANCE OF GENE NETWORK MODEL
  via SIMULATED ANNEALING ................................. 4
  2.1 Abstract .................................................. 4
  2.2 Introduction .............................................. 5
  2.3 Methods .................................................. 6
    2.3.1 Dynamics of models and Simulated annealing ......... 6
    2.3.2 Parameter analysis .................................. 11
  2.4 Results .................................................. 12
    2.4.1 High fitnesses of models after simulation ............. 12
    2.4.2 V and µ parameters .................................. 23
  2.5 Discussion ............................................... 35
    2.5.1 A high fitness of gene network model ................. 35
    2.5.2 Parameters ........................................... 36
    2.5.3 Computation .......................................... 36
LIST OF TABLES

Table 2.1 Parameter description. ................................................................. 10
Table 2.2 Initial parameters and a fitness, and optimal parameters and a high
  fitness of Model 1. ................................................................................... 16
Table 2.3 Initial parameters and a fitness, and optimal parameters and a high
  fitness of Model 2. ................................................................................... 18
Table 2.4 Initial parameters and a fitness, and optimal parameters and a high
  fitness of Model 3. ................................................................................... 21
Table 2.5 Relations of $v$ parameter and $\mu$ parameter: $v_{21}$ is directly proportional
  to $\mu_1 \times (1 + K_1)$, and $v_{12}$ is directly proportional to $\mu_2 \times (1 + K_2)$. . 33
LIST OF FIGURES

Figure 2.1 Model 1: Gene 1 activates Gene 2 and Gene 2 activates Gene 1 but each of them inhibits itself. ............................................. 6
Figure 2.2 Model 2: Both genes inhibit each other, but each of them activates itself. ....................................................... 7
Figure 2.3 Model 3: There is no interaction each other, but each of them inhibits itself. ..................................................... 8
Figure 2.4 Flow Chart of Simulation Program .................................................. 12
Figure 2.5 On the left side graphs show the change of expression of genes and the fitness before the simulation, and on the right side graphs show the change of expression of genes and the fitness of Model 1 with evolutionary time. ........................................... 15
Figure 2.6 The track of fitness shows how new fitness is accepted during simulation. This simulation start with the initial temperature, 1. .................. 16
Figure 2.7 It indicates the track of fitness with the initial temperature 0.3487. .. 17
Figure 2.8 On the left side graphs indicate the change of expression of genes and the fitness before the simulation, and on the right side graphs show the change of expression of genes and the fitness of Model 2. .................. 19
Figure 2.9 The track of fitness indicates how new fitness is accepted during simulation. .................................................. 20
Figure 2.10 On the left side graphs indicate the change of expression of genes and fitness before the simulation, and on the right side graphs show the change of expression of genes and fitness of Model 3. .................. 21
Figure 2.11  The track of fitness indicates how new fitness is accepted during simulation.

Figure 2.12  Blue line is model 1 and red line is model 2. Gene 1 of model 1 is expressed faster than gene 1 of model 2 when the environments are changed.

Figure 2.13  200 times simulation: Blue line is a simulation with the initial parameters. Red line is other 60 simulation results. The cutoff value is 0.6. After collecting high fitnesses, we found the pattern of parameters inducing a high fitness.

Figure 2.14  Special cases of relations $v_{21}$ and $\mu_{1}(1 + K_{1})$ parameters. $v_{21}$ and $v_{22}$ is big changed in environment 2.

Figure 2.15  $v_{21}$ and $\mu_{1}(1 + K_{1})$ parameters are in direct proportion.

Figure 2.16  Red line is the change of concentration of gene 1 and blue line is the change of concentration of gene 2. Point 30 shows both genes are expressed well in both environments.

Figure 2.17  Red line is the change of concentration of gene 1 and blue line is the change of concentration of gene 2. Point 60 shows gene 2 is expressed well in environment 1 but gene 1 is not expressed well in environment 2.

Figure 2.18  Special cases of relations $v_{12}$ and $\mu_{2}(1 + K_{2})$ parameters. $v_{11}$ and $v_{12}$ is big changed in environment 1.

Figure 2.19  $v_{12}$ and $\mu_{2}(1 + K_{2})$ parameters are in direct proportion.

Figure 2.20  Red line is the change of concentration of gene 1 and blue line is the change of concentration of gene 2. Point 40 shows both genes are expressed well in both environments.

Figure 2.21  Red line is the change of concentration of gene 1 and blue line is the change of concentration of gene 2. Point 16 shows gene 1 is expressed well in environment 2 but gene 2 is not expressed well in environment 1.

Figure 2.22  X-axis is $v_{21}-\mu_{1}(1+K_{1})$, and y-axis is $v_{12}-\mu_{2}(1+K_{2})$.  
Figure 2.23  Red line is the change of concentration of gene 1 and blue line is the change of concentration of gene 2. ................................. 32
Figure 2.24  Red line is the change of concentration of gene 1 and blue line is the change of concentration of gene 2. ................................. 32
Figure .1  Track of fitness of model 1 with the initial temperature, 1. ............ 41
Figure .2  Track of fitness of model 1 with the initial temperature, 0.4305. ....... 42
Figure .3  Track of Parameters of Model 1 .................................................. 43
Figure .4  Track of Parameters of Model 2 .................................................. 44
Figure .5  Track of Parameters of Model 3 .................................................. 45
Figure .6  200 times simulation: Blue line is a simulation with the initial parameters. Red line is other 166 simulation results. the cutoff value is 0.4 . 46
Figure .7  Special cases of relations $v_{21}$ and $\mu_1(1+K_1)$ parameters. ........... 47
Figure .8  $v_{21}$ and $\mu_1(1+K_1)$ parameters are in direct proportion. ............. 48
Figure .9  Special cases of relations $v_{12}$ and $\mu_2(1+K_2)$ parameters. .......... 49
Figure .10  $v_{12}$ and $\mu_2(1+K_2)$ parameters are in direct proportion. .......... 50
ACKNOWLEDGEMENTS

I would like to take this opportunity to thank my major professor, Stephen Proulx, who helped me with various aspects of conducting research and the writing of this thesis. His lab has been a great fit for me and I appreciate all his support. I also should mention other lab members, including Olaf Bochmann, Lisa Ellen Schwanz, and Hong Hua. They helped to understand this study area.

I would like to thank my committee members for their efforts and contributions to this work: Dr. Michael Smiley, Dr. Zhijun Wu, and Dr. Jeffrey Essner.

I would also like to thank my BCB friends and my best friend, Olga for their loving guidance during the writing of this work.

Finally and most importantly, I would like to dedicate this thesis to my husband, Bongkeun Kim and to my other family members without whose support I would not have been able to complete this work.
CHAPTER 1. GENERAL INTRODUCTION

1.1 Evolutionary gene network

One the most important problems in computational biology is the analysis of gene networks. Gene networks show which genes are expressed, to what level, and where and when for understanding of the function of organisms at the molecular level (Klipp et al., 2005).

Gene network has been described with boolean networks, bayesian networks, stochastic equations, directed graphs, and ordinary and partial differential equation. Even though many insights have already been gained from the modeling of particular processes or of the regulation of individual sets of genes, understanding of the regulation of genes is still a scientific challenge (Klipp et al., 2005; Proulx et al., 2005; Proulx and Phillips., 2005; Force et al., 2005).

We studied the evolutionary gene networks. Our mathematical approach aimed to study gene networks within the concepts of evolutionary time, and environment changes. Ordinary differential equations (ODE) were used to model the reaction kinetics between genes. A gene network was represented as directed graph, where nodes represent genes, and edges - reactions between genes (Klipp et al., 2005).

Also, Fitness is a central concept in evolutionary theory. Fitness means individual’s ability to propagate its genes. If one individual has a genetic trait that result in low fitness, It would be hard to pass genetic trait to next generation. It is populations that adapt as genetic traits that result in a high fitness in individuals become more common. Thus, we figured out a high fitness of the gene network model. The fitness is influenced by both genes and environment.

When we discuss evolutionary gene networks, it is important to realize epistatic effect (Proulx and Phillips., 2005). The term "epistasis" was introduced as the masking of the expression of one-locus by the alleles at another locus by William Bateson, and Phillips has
recommended the term "epistasis" as a generic term to describe gene interaction (Wolf et al., 2000). Thus, we described two genes’ interactions to analyze evolutionary gene networks and applied simulated annealing algorithm.

1.2 Simulated annealing algorithm

Simulated annealing is a stochastic computational technique to program global optimization. In general, simulated annealing algorithm is used to find the global minimum value of an objective function with many degrees of freedom subject to conflicting constraints and is an NP-complete problem, because the objective function will have many local minima (Davis, 1987). However, we applied the simulated annealing algorithm to find global maximum value of an objective function. Kirkpatrick said the closest analogy with the shifting balance are simulated annealing algorithm and the stochastic process of the simulated annealing leads to alternative solutions, which is analogous to Wright’s equation (Kirkpatrick et al., 1983). There are two part in the simulated annealing. First one is annealing process. As temperature is decreased slowly, good results are chosen and at the low temperature, the optimal results are selected. Second one is stochastic mechanics. Suppose that at time t, the fitness of gene network model is $W$. At time $t + 1$, a candidate fitness $C$ is generated randomly. The criterion for selecting, or rejecting fitness $C$ depends on the difference between the fitness $C$ and $W$. Specifically, the ratio $p$ between the probability of being in $C$ and the probability of being in $W$:

$$p = \exp\left(\frac{C - W}{Temperature}\right)$$  \hspace{1cm} (1.1)

If $p > 0$, that is, the fitness $C$ is greater than the fitness $W$, then fitness $C$ is automatically accepted as the new fitness for time $t + 1$. While $p \leq 0$, that is the fitness $C$ is less than or equal to the fitness $W$, then fitness $C$ is accepted as the new fitness with probability $p$. Finally, when the temperature is low enough and $p \leq 0$, Always $C - W$ is negative and $\frac{C - W}{Temperature}$ is close to negative 0. Then, the probability $p$ is very small and always we can get a high fitness of the gene network (Davis, 1987).
1.3 Reference


CHAPTER 2. EVOLUTIONARY VARIANCE OF GENE NETWORK MODEL via SIMULATED ANNEALING

2.1 Abstract

The traditional approach of molecular biology research was on examining and collecting data on a single gene or a single reaction. However, recently, there has been much interest on the dynamics of gene regulatory networks (Klipp et al., 2005). We applied mathematical approach for modeling of gene network. The models depict the reaction kinetics of the constituent parts and the functions are ultimately made from basic principle of simple expressions derived from Michaelis-Menten enzymatic kinetics, and the functional forms are usually chosen as Hill functions that serve as an approximation for the real molecular dynamics (Klipp et al., 2005). These dynamics depend on many parameters and the parameters strongly influence the behavior of the resulting gene network. Thus, we used simulated annealing algorithm to calculate a high fitness and optimal parameters of the gene network. The simulated annealing algorithm is suitable for calculating many degree of freedom (Tomshine and Kaznessis, 2006), and is the closest analogy with the shifting balance theory of populations (Kirkpatrick et al., 1983). We developed 3 different models that have two genes and experience two different environments, and simulated to describe the behavior of evolutionary gene networks. From simulation, we could obtain a high fitness of each gene network model, and we could indicate how gene network is evolved in evolutionary time from tracks of parameters and a fitness. Also, we analyzed the relations of a high fitness and parameters. We think we can apply to design and optimize other gene network, and these findings are useful to analysis of the evolutionary gene network.
2.2 Introduction

Recently, there has been much interest on the dynamics of gene regulatory networks, being a collection of DNA segments in a cell that interacts with each other and with other substances in the cell (Ideker et al., 2001; Proulx et al., 2005; Proulx and Phillips., 2005). The nodes are genes, the input of the node is the transcription factor, and the output of the node is gene expression in gene networks (Vohradsky, 2001). There are several modeling techniques that have been used. For instance, there are boolean networks, bayesian networks, graphical gaussian models, and stochastic networks (Klipp et al., 2005). Boolean networks describe qualitative gene regulatory interactions, and gene expression to two states: on and off. Bayesian networks depict probabilistic gene regulatory networks and consist of a directed acyclic graph, and a set of probability distributions (Klipp et al., 2005). One of the techniques for studying gene regulatory networks is to use mathematical models describe a gene regulatory network using ordinary differential equations. Ordinary differential equations describe the reactions kinetics of the constituent parts, and the functions are ultimately induced from basic principle of simple expressions derived from Michaelis-Menten enzymatic kinetics. Also, ordinary differential equations are used to describe temporal evolutionary gene networks and the functional forms are usually chosen as Hill functions that serve as an analysis for the real molecular dynamics (Klipp et al., 2005). We used ordinary differential equation to apply temporal changes of gene network models in evolutionary time and these mathematic formulas help to understand the dynamics of gene networks well.

After we described mathematical gene network models, we calculated a fitness of each models. We want to show two things in this paper. First, how we obtained a high fitness of dynamic gene network models. These dynamics depends on parameters, and the parameters are features of each gene in the network. This parameter prediction is major challenge in dynamic gene networks (Piazza et al., 2008). We used the simulated annealing algorithm for calculating a high fitness of gene network models and we obtained the optimal parameters inducing the high fitness. Second, we want to show how parameters influence a fitness of gene networks. Thus, we simulated 200 times, and analyzed the parameters of gene network models (de Visser
et al., 2003; Kitano, 2004). From the simulation using the simulated annealing algorithm, we could obtain a high fitness and the optimal parameters inducing the high fitness, and from these parameters, we could expect the behavior of the resulting gene network. In the future, we will expand our project to deal with interactions of several genes and we believe that this research will help to analyze the evolutionary gene networks.

2.3 Methods

2.3.1 Dynamics of models and Simulated annealing

We have 3 different gene network models that have two genes and depict two different environments: (1) Model 1, (2) Model 2, and (3) Model 3. We developed ordinary differential equations of these models, and simulated these gene networks. We applied the simulated annealing algorithm to get a high fitness from simulation. Each model consists of 4 differential equations describing gene interactions. These equations indicate the expression of gene 1 and gene 2 in each environment.

Here, $\mu_1$, and $\mu_2$ are the first-order rate constants of degradation of gene 1 and gene 2 respectively. $v_{ij}$ denotes the constant rate of expression of gene $j$ in environment $i$, and the Hill term $\frac{v_{ij}}{K_i + G_{ni}}$ describes the formation of gene $j$ is activated by gene $i$ with maximal rate $v_{ij}$, dissociation constant $K_i$, and Hill coefficient $n_i$ (Klipp et al., 2005).

2.3.1.1 Models

![Figure 2.1 Model 1: Gene 1 activates Gene 2 and Gene 2 activates Gene 1 but each of them inhibits itself.](image)
Model 1

The first model has two genes and experiences two different environments. In the first model, Gene 1 activates Gene 2, and Gene 2 activates Gene 1, but each of them inhibits itself [Figure 2.1].

[Formula 2.1] describes interaction of gene 1 and gene 2, and degradation of gene 1 in environment 1.

\[
\frac{dG_1}{dt} = \frac{v_{11} + (\alpha_{21} \times G_2)}{(G_{n1}^1 + K_1)} - \mu_1 \times G_1
\] (2.1)

[Formula 2.2] describes interaction of gene 1 and gene 2, and degradation of gene 2 in environment 1.

\[
\frac{dG_2}{dt} = \frac{v_{12} + (\alpha_{12} \times G_1)}{(G_{n2}^2 + K_2)} - \mu_2 \times G_2
\] (2.2)

[Formula 2.3] describes interaction of gene 1 and gene 2, and degradation of gene 1 in environment 2.

\[
\frac{dG_1}{dt} = \frac{v_{21} + (\alpha_{21} \times G_2)}{(G_{n1}^1 + K_1)} - \mu_1 \times G_1
\] (2.3)

[Formula 2.4] describes interaction of gene 1 and gene 2, and degradation of gene 2 in environment 2.

\[
\frac{dG_2}{dt} = \frac{v_{22} + (\alpha_{12} \times G_1)}{(G_{n2}^2 + K_2)} - \mu_2 \times G_2
\] (2.4)

Figure 2.2 Model 2: Both genes inhibit each other, but each of them activates itself.
Model 2

The second model is that gene 1 inhibits gene 2, and gene 2 inhibits gene 1, but each of them activates itself [Figure 2.2].

[Formula 2.5] describes interaction of gene 1 and gene 2, and degradation of gene 1 in environment 1.

\[
\frac{dG_1}{dt} = \frac{v_{11} + (\alpha_{21} \times G_1)}{G_2^{m_1} + K_1} - \mu_1 \times G_1 \tag{2.5}
\]

[Formula 2.6] describes interaction of gene 1 and gene 2, and degradation of gene 2 in environment 1.

\[
\frac{dG_2}{dt} = \frac{v_{12} + (\alpha_{12} \times G_2)}{G_1^{m_2} + K_2} - \mu_2 \times G_2 \tag{2.6}
\]

[Formula 2.7] describes interaction of gene 1 and gene 2, and degradation of gene 1 in environment 2.

\[
\frac{dG_1}{dt} = \frac{v_{21} + (\alpha_{21} \times G_1)}{G_2^{m_1} + K_1} - \mu_1 \times G_1 \tag{2.7}
\]

[Formula 2.8] describes interaction of gene 1 and gene 2, and degradation of gene 2 in environment 2.

\[
\frac{dG_2}{dt} = \frac{v_{22} + (\alpha_{12} \times G_2)}{G_1^{m_2} + K_2} - \mu_2 \times G_2 \tag{2.8}
\]

Figure 2.3 Model 3: There is no interaction each other, but each of them inhibits itself.

Model 3

The third model is that there is no interaction each other but each of them inhibits itself [Figure 2.3].
[Formula 2.9] describes degradation of gene 1 in environment 1.
\[
\frac{dG_1}{dt} = v_{11} - \mu_1 \times G_1
\] (2.9)

\[
\frac{dG_2}{dt} = v_{12} - \mu_2 \times G_2
\] (2.10)

\[
\frac{dG_1}{dt} = v_{21} - \mu_1 \times G_1
\] (2.11)

\[
\frac{dG_2}{dt} = v_{22} - \mu_2 \times G_2
\] (2.12)

2.3.1.2 Simulated annealing algorithm

The simulated annealing algorithm is used to get certain results from unmanageable systems using combinatorial methods (Vicente et al., 2003). Models have two different genes and depict two different environments. The fitness is influenced by changing the concentrations of genes. The concentrations of genes are changed by parameters and two different environments. The simulated annealing algorithm is used to find a high fitness and optimal parameters of gene network models as the temperature is decreased. Here is the ordinary differential equations of fitness of gene networks.

Fitness in environment 1:
\[
\frac{dW}{dt} = \exp(-1 \times \frac{(Z_{11} - G_1)^2}{\sigma_{11}^2}) \times \exp(-1 \times \frac{(Z_{12} - G_2)^2}{\sigma_{12}^2})
\] (2.13)

Fitness in environment 1:
\[
\frac{dW}{dt} = \exp(-1 \times \frac{(Z_{21} - G_1)^2}{\sigma_{21}^2}) \times \exp(-1 \times \frac{(Z_{22} - G_2)^2}{\sigma_{22}^2})
\] (2.14)

We set \(\sigma_{11}, \sigma_{12}, \sigma_{21}, \sigma_{22}\) are same values, 0.5, and \(Z_{11}\) is 0, \(Z_{12}\) is 1, \(Z_{21}\) is 1, and \(Z_{22}\) is 0. Also, there are 12 parameters: \(v_{11}, v_{12}, v_{21}, v_{22}, n_1, n_2, K_1, K_2, \mu_1, \mu_2, \alpha_{12},\) and \(\alpha_{21}\). These parameters are derived from Hill equations. Then we used the simulated annealing algorithm to get a high fitness and optimal parameters of gene regulatory network models. [Table 2.1] describes the minimum and maximum values of these parameters.
Table 2.1 Parameter description.

<table>
<thead>
<tr>
<th>Description</th>
<th>Min</th>
<th>Max</th>
<th>Initial values</th>
</tr>
</thead>
<tbody>
<tr>
<td>$v_{ij}$ Maximal rate of activation of $G_j$ in environment $i$</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>$n_i$ Hill coefficient of $G_i$</td>
<td>0</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>$K_i$ Dissociation constant of $G_i$</td>
<td>0</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>$\mu_i$ Degradation rate of $G_i$</td>
<td>0</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>$\alpha_{ij}$ Maximal rate of activation from $G_i$ to $G_j$</td>
<td>0</td>
<td>1</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Initial conditions

We need to set initial conditions: initial parameters, initial concentration of genes, initial temperature, decreasing temperature rate, loops, and step size. We set the initial parameters; $v_{11} = 1, v_{12} = 1, v_{21} = 1, v_{22} = 1, n_1 = 0.5, n_2 = 0.5, K_1 = 0.5, K_2 = 0.5, \mu_1 = 0.5, \mu_2 = 0.5, \alpha_{12} = 0.5,$ and $\alpha_{21} = 0.5$. They are mean values of the size of each parameter and we set the initial concentration of $G_1$ is 0.5, $G_2$ is 0.5, and $W$ is 0. [Table 2.1] shows the minimum and maximum sizes of each parameter.

Finding the optimal initial conditions is a hard task due to the large number of possible combinations of parameter values. Thus, we set initial temperature was 1. We can set much higher temperature but if we set much higher temperature, we need to simulate more to make the temperature is close to 0 and it means it takes much more time. Also, we can set much lower temperature but we would obtain bad results from simulation because at the low temperature, solutions are not chosen nearly at random over the whole range of possibilities.

Then, the decreasing temperature rate is done as described in Tomshine and Kaznessis (Tomshine and Kaznessis, 2006). It is a simple proportional annealing method. Whereby $T_{i+1} = \alpha \times T_i$, where $\alpha$ is an empirically chosen number with $0 < \alpha < 1$. We found a proper $\alpha$ is 0.9.

We considered how many times we need to simulate to obtain a high fitness of gene network models. Our criterion is the temperature at the end of simulation. During the simulation, we accept or reject a new system depending on the temperature. When the temperature is high, the system is unstable, but when the temperature is low enough, the system is stable. Therefore, we assume the temperature would be very low at the end of annealing. Thus, we
could decide how many simulations are need when the initial temperature was 1, and the
decreasing rate was 0.9. We set 500 decreasing temperature loops and 10 loops to change the
step size. Thus we set total 5000 loops to simulate the program just for once.

Lastly, we developed a step size. The size of the attempted steps in parameter space is
one of critical issue of the optimization process (Tomshine and Kaznessis, 2006). If the step
size is too large, it can cause the system to oscillate too much, and if step size is too small, it
can cause excessive iteration for convergence. Thus, we simulate 10 times at same conditions,
and count how many time new system is accepted. Then if new system is accepted only a few
times, we change the step size to bigger, while if new system is accepted almost, we change
the step size to smaller.

Simulations

A flowchart of the simulation process is shown in [ Figure 2.4 ] based on the simulated
annealing algorithm. We set the initial condition. Then, there are three options largely: (1)
the choice of an evolutionary time function, (2) the choice of a time length, and (3) the choice
of a model,. We have 3 different models and we have 3 different types of time functions: (1)
fixed 10 size time, (2) uniform random time, and (3) exponential random time. We fixed two
different environments are switched 10 times in real time. Thus, the case of (1) fixed 10 size
time will have E(T)= (10 × 10) time length, the case of (2) random time will have E(T)= (10
× uniform random number) time length, and the case of (3) exponential random time period
will have E(T)= (10 × exponential random number) time length.

2.3.2 Parameter analysis

Parameter estimation is complicated and difficult. Initial parameters are important be-
cause, initial parameters can cause different results. If we set the initial parameters inducing a
high fitness, we obtain many parameters inducing a high fitness after 200 simulations. For in-
stance, we set the initial parameters inducing 0.8656 fitness, then after 200 times simulations,
we obtained 152 fitnesses which are higher than 0.7 (152/ 200). On the other hand, when
we set the initial parameters whose size is the mean of parameters’ size, the average fitness
Figure 2.4  Flow Chart of Simulation Program

is 0.4, and the highest fitness is 0.6419 after simulations. Therefore, we conclude the initial parameters influence fitness so much.

However, we simulated 200 times with the initial parameters are the mean values of parameters’ size. Then, we found the pattern of parameters of model 1. After 200 simulations, we obtained same parameters pattern with the previous simulations by optimal parameters inducing a high fitness. Thus, we conclude final fitness value is determined by initial parameter values but the pattern of parameters inducing a high fitness is not changed. Also, we found relations of $v$ parameter and $\mu$ parameter from 200 simulations.

2.4 Results

2.4.1 High fitnesses of models after simulation

Model 1

We simulated total 5000 loops and it took 885.021400 seconds. We obtained much higher
fitness than initial fitness after simulation. Initial temperature was 1 and it was decreased by 10%. Figure 2.5 indicates a result of simulated annealing program with model 1. Table 2.2 shows the number of result of simulated annealing program. This simulation indicates that the initial fitness was 0.0527 and the initial input parameters were $v_{11} = 0.1$, $v_{12} = 1.5$, $v_{21} = 1.5$, $v_{22} = 0.1$, $n_{1} = 0.5$, $n_{2} = 0.5$, $K_{1} = 0.5$, $K_{2} = 0.5$, $\mu_{1} = 0.5$, $\mu_{2} = 0.5$, $\alpha_{12} = 0.1$, and $\alpha_{21} = 0.1$. However, we obtained a high fitness of model 1 after simulation. The ending temperature was $1.3221e^{-23}$, we thought it is low enough. The high fitness was 0.8685 and the optimal parameters were $v_{11} = 0.0512$, $v_{12} = 1.3649$, $v_{21} = 1.9370$, $v_{22} = 0.0746$, $n_{1} = 0.3941$, $n_{2} = 0.2399$, $K_{1} = 0.9527$, $K_{2} = 0.8064$, $\mu_{1} = 0.9883$, $\mu_{2} = 0.7792$, $\alpha_{12} = 0.0092$, and $\alpha_{21} = 0.8963$. In this case, we set the exponential random time function, and the time length was close to 110 made by multiplying between exponential random number and 10. We found that two environments are switched 10 times in 110 time length Figure 2.5. For instance, environment 1 comes out then environment 2 comes out repetitively 10 times.

Figure 2.5 shows the change of the expression of gene 1($G_{1}$), gene 2($G_{2}$) and the fitness($W$) before the simulation on the left side and the right side graphs show the change of the expression of $G_{1}$, $G_{2}$ and the $W$ after the simulation.

From Figure 2.5, we found that always, fitness was increasing. Also, we found that when we gave initial parameters, we set $v_{12}$ and $v_{21}$ is higher than $v_{11}$ and $v_{22}$($v_{11} = 0.1$, $v_{12} = 1.5$, $v_{21} = 1.5$, $v_{22} = 0.1$). It means gene 1 is expressed more in environment 2, and gene 2 is expressed more in environment 1 and the same result is derived from fitness formulas [Formula 2.13] and [Formular 2.14]. Thus, both before simulation and after simulation indicate same pattern show gene 1 and gene 2 are expressed alternatively. On the left side graphs, it shows the expression of gene 1 and gene 2 and fitness in real time. In the environment1, gene 2 was expressed more than gene 1, then gene 1 was expressed more than gene 2 in environment 2. On the right side graphs, it also shows the expression of gene 1 and gene 2 and fitness in real time. Because we set the exponential random time, the time length is flexible. The exponential random time shows similar nature system.

Figure 2.6 indicates the track of fitness. We could find that finally we obtained the
highest fitness during the simulation. The fitness figure of [Figure 2.5] indicates the last point of the fitness of [Figure 2.6]. From the track of fitness, we were wondering why fitness was too much oscillatory. The reason is when the temperature is still high, the system is unstable and smaller new fitness can be accepted with the high probability $p$. Thus, even though new fitness was smaller than current fitness, new fitness could be accepted. However, we expect when the system is stable, always fitness is going up at the end of simulation. [Figure 2.6] is shown always new fitness is higher than current fitness after 65 of the fitness track. We tracked other parameters also, and there are more track figures of parameters in Appendix.

We had a question about initial temperature. Usually we set the initial temperature was 10, 1, or 0.1. If we set the initial temperature was 10, we needed more loops to make the ending temperature is close to 0, and if we set the initial temperature was 0.1, we needed less loops but we could not obtain good results because, already temperature is low and the system is already stable. Thus, We set the initial temperature was 1. However, we found interesting thing. The [Figure 2.6] is too much oscillatory. Thus, we pick 49 times fitness from the fitness track. 49 times fitness was 0.4822 and other parameters were $v_{11} = 0.0997, v_{12} = 1.604, v_{21} = 1.011, v_{22} = 0.02526, n_1 = 0.7716, n_2 = 0.08392, K_1 = 0.9787, K_2 = 0.6782, \mu_1 = 0.6733, \mu_2 = 0.8982, \alpha_{12} = 0.213, \text{ and } \alpha_{21} = 0.1226$. At that point, the temperature was 0.4305, and the concentration of G1 and G2 were 0.8288 and 0.148. Thus, we simulated with an initial temperature is 0.3487 with above parameters and fitness and the concentrations of G1 and G2.

Finally, [Figure 2.7] indicates even though the graph of simulation with the initial temperature 0.3487 is not same with the graph of simulation with the initial temperature 1, finally both fitness tracks go up to 0.86. The fitness of simulation with the initial temperature 0.3487 was 0.8650, and the fitness of simulation with the initial temperature 1 was 0.8685. Thus, we conclude that we can save running time by finding a properly initial temperature. If we set the initial temperature 0.3487, we need less loops than before to simulation and we can save running time. We have more examples in appendix.
Figure 2.5 On the left side graphs show the change of expression of genes and the fitness before the simulation, and on the right side graphs show the change of expression of genes and the fitness of Model 1 with evolutionary time.
Table 2.2  Initial parameters and a fitness, and optimal parameters and a high fitness of Model 1.

<table>
<thead>
<tr>
<th>Element</th>
<th>Initial Parameters</th>
<th>Optimal Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fitness</td>
<td>0.0527</td>
<td>0.8685</td>
</tr>
<tr>
<td>v11</td>
<td>0.1</td>
<td>0.0512</td>
</tr>
<tr>
<td>v12</td>
<td>1.5</td>
<td>1.3649</td>
</tr>
<tr>
<td>v21</td>
<td>1.5</td>
<td>1.9370</td>
</tr>
<tr>
<td>v22</td>
<td>0.1</td>
<td>0.0746</td>
</tr>
<tr>
<td>n1</td>
<td>0.5</td>
<td>0.3941</td>
</tr>
<tr>
<td>n2</td>
<td>0.5</td>
<td>0.2399</td>
</tr>
<tr>
<td>K1</td>
<td>0.5</td>
<td>0.9527</td>
</tr>
<tr>
<td>K2</td>
<td>0.5</td>
<td>0.8064</td>
</tr>
<tr>
<td>mu1</td>
<td>0.5</td>
<td>0.9883</td>
</tr>
<tr>
<td>mu2</td>
<td>0.5</td>
<td>0.7792</td>
</tr>
<tr>
<td>alpha12</td>
<td>0.1</td>
<td>0.0092</td>
</tr>
<tr>
<td>alpha22</td>
<td>0.1</td>
<td>0.8963</td>
</tr>
</tbody>
</table>

Figure 2.6  The track of fitness shows how new fitness is accepted during simulation. This simulation start with the initial temperature, 1.
Figure 2.7 It indicates the track of fitness with the initial temperature 0.3487.

Model 2

We simulated model 2 as model 1. it was 5000 loops and it took 1409.117486 seconds. We obtained higher fitness than initial fitness. [ Figure 2.8 ] shows a result of simulated annealing algorithm with model 2. [ Table 2.3 ] indicates the number of result of simulated annealing program. In [ Figure 2.8 ], it indicates the change of expression of $G_1$, $G_2$ and $W$ before the simulation on the left side and the right side graphs indicate the change of expression of $G_1$, $G_2$ and $W$ after simulation. This simulation indicates that the initial fitness was 0.0065 and the initial input parameters were $v_{11} = 0.1, v_{12} = 1.5, v_{21} = 1.5, v_{22} = 0.1, n_1 = 0.5, n_2 = 0.5, K_1 = 0.5, K_2 = 0.5, \mu_1 = 0.5, \mu_2 = 0.5, \alpha_{12} = 0.1$, and $\alpha_{21} = 0.1$. However, we obtained a high fitness of model 2 after simulation. The high fitness was 0.8528 and the optimal parameters were $v_{11} = 0.0284, v_{12} = 0.5202, v_{21} = 0.3092, v_{22} = 0.0580, n_1 = 0.8640, n_2 = 0.9324, K_1 = 0.5582, K_2 = 0.7614, \mu_1 = 0.5448, \mu_2 = 0.7661, \alpha_{12} = 0.8816$, and $\alpha_{21} = 0.0492$ [ Table 2.3 ]. We set the exponential random time function, and the time length was close to 200 made by multiplying between exponential random number and 10. Therefore, two different environments are switched 10 times in 200 time length.

From [ Figure 2.8 ], we found that the fitness was increasing always but, on the left side graphs, fitness was increasing quickly during the environment 1 and it was increasing very
Table 2.3 Initial parameters and a fitness, and optimal parameters and a high fitness of Model 2.

<table>
<thead>
<tr>
<th>Element</th>
<th>Initial Parameters</th>
<th>Optimal Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fitness</td>
<td>0.0065</td>
<td>0.8528</td>
</tr>
<tr>
<td>v11</td>
<td>0.1</td>
<td>0.0284</td>
</tr>
<tr>
<td>v12</td>
<td>1.5</td>
<td>0.5202</td>
</tr>
<tr>
<td>v21</td>
<td>1.5</td>
<td>0.3092</td>
</tr>
<tr>
<td>v22</td>
<td>0.1</td>
<td>0.580</td>
</tr>
<tr>
<td>n1</td>
<td>0.5</td>
<td>0.8640</td>
</tr>
<tr>
<td>n2</td>
<td>0.5</td>
<td>0.9324</td>
</tr>
<tr>
<td>K1</td>
<td>0.5</td>
<td>0.5582</td>
</tr>
<tr>
<td>K2</td>
<td>0.5</td>
<td>0.7614</td>
</tr>
<tr>
<td>mu1</td>
<td>0.5</td>
<td>0.5448</td>
</tr>
<tr>
<td>mu2</td>
<td>0.5</td>
<td>0.7661</td>
</tr>
<tr>
<td>alpha12</td>
<td>0.1</td>
<td>0.8816</td>
</tr>
<tr>
<td>alpha22</td>
<td>0.1</td>
<td>0.0492</td>
</tr>
</tbody>
</table>

slowly. We thought the reason is model 2 is that genes activate by themselves and inhibit each other. Thus, model 2 indicates different figures with the simulation of model 1.

[Figure 2.9] indicates the track of fitness. This model also shows that the highest fitness is obtained at the end of simulation. Also, after 38 times fitness, we could find always the fitness climb up. There are more track figures of parameters in Appendix.
Figure 2.8 On the left side graphs indicate the change of expression of genes and the fitness before the simulation, and on the right side graphs show the change of expression of genes and the fitness of Model 2.
Figure 2.9  The track of fitness indicates how new fitness is accepted during simulation.

**Model 3**

We simulated total 5000 loops and it took 944.029041 seconds. We obtained much higher fitness than the initial fitness. [Figure 2.10] indicates a result of simulated annealing algorithm with model 3. [Table 2.4] indicates the number of result of simulated annealing program. In [Figure 2.10], it shows the change of the expression of $G_1$, $G_2$ and $W$ before simulation on the left side and the right side graphs indicate the change of the expression of $G_1$, $G_2$ and $W$ after simulation. This simulation indicates that the initial fitness was 0.0048 and the initial input parameters were $v_{11} = 0.1, v_{12} = 1.5, v_{21} = 1.5, v_{22} = 0.1, \mu_1 = 0.5$, and $\mu_2 = 0.5$. However, we obtained a high fitness of model 3 after simulation. The high fitness was 0.8505 and the optimal parameters were $v_{11} = 0.0302, v_{12} = 0.7453, v_{21} = 1.0581, v_{22} = 0.0112, \mu_1 = 0.9734$, and $\mu_2 = 0.6681$ [Table 2.4]. Also, we set the exponential random time period function, and the time length is close to 120 made by multiplying between 10 and the exponential random number. Therefore, two different environments are switched 10 times in 120 time length.

Model 3 is that there is no interaction between genes, and there is only degradation of each gene. Thus, we thought the fitness is not changed like model 1. [Figure 2.11] indicates the track of fitness. It also shows that finally we obtained the highest fitness during the simulation.
Table 2.4  Initial parameters and a fitness, and optimal parameters and a high fitness of Model 3.

<table>
<thead>
<tr>
<th>Element</th>
<th>Initial Parameters</th>
<th>Optimal Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fitness</td>
<td>0.0048</td>
<td>0.8505</td>
</tr>
<tr>
<td>v11</td>
<td>0.1</td>
<td>0.0302</td>
</tr>
<tr>
<td>v12</td>
<td>1.5</td>
<td>0.7453</td>
</tr>
<tr>
<td>v21</td>
<td>1.5</td>
<td>1.0581</td>
</tr>
<tr>
<td>v22</td>
<td>0.1</td>
<td>0.0112</td>
</tr>
<tr>
<td>mu1</td>
<td>0.5</td>
<td>0.9734</td>
</tr>
<tr>
<td>mu2</td>
<td>0.5</td>
<td>0.6681</td>
</tr>
</tbody>
</table>

Figure 2.10  On the left side graphs indicate the change of expression of genes and fitness before the simulation, and on the right side graphs show the change of expression of genes and fitness of Model 3.
2.4.1.1 Simulated annealing

When we set the initial temperature was 1, and the decreasing temperature rate was 10 \%, the ending temperature was $1.3221 \times 10^{-23}$. It means the initial temperature, 1 was cooling slowly and finally the temperature became $1.3221 \times 10^{-23}$. We think it is low temperature enough because the fitness was not going down at the end of simulation. If the temperature is low enough, the fitness only climbs up on the landscape of fitness because the system become a stable and the smaller fitness is not accepted.

Also, we set the empirical step size, $S$ is 0.8 and when new fitness is accepted a few times, $S$ is changed by $S(1/1.2)$, then the new step size is bigger than before. On the other hand, when new fitness is almost accepted, $S$ is changed by $S(1.2)$, then the new step size is smaller than before. Thus, we can prevent that the fitness jumps around far away.

After simulating of each model, we compared between model 1 and model 2. From the optimal parameters, we found that the concentration of gene 1 of model 1 is higher than the concentration of gene 1 of model 2. Thus, we could show gene 1 of model 1 is expressed faster than gene 1 of model 2 when the environments are changed[Figure 2.12].

We obtained interesting results from the simulation. First, we could obtain a high fitness and optimal parameters inducing the high fitness in each model. Second, we could show how
Figure 2.12 Blue line is model 1 and red line is model 2. Gene 1 of model 1 is expressed faster than gene 1 of model 2 when the environments are changed.

the gene network is evolved by parameters, and environments.

2.4.2 V and μ parameters

We simulated 200 times to find the pattern of parameters. Then, [Figure 2.13] shows the results: blue line is a simulation with the initial parameters and red line are other 60 simulation results of 200 simulations. We had 0.6, cutoff fitness value. We could find a parameters’ pattern of model 1 from [Figure 2.13]: \(v_{11}\) is small, \(v_{21}\) is large, \(v_{21}\) is large, and \(v_{22}\) is small and \(\mu_1\), and \(\mu_2\) are always large. This results prove [Formula 2.13] and [Formula 2.14]. When we assumed other parameters didn’t affect the fitness of gene network model except \(v\) and \(\mu\), we could expect gene 1 activates more than gene 2 in environment 2 and gene 2 activates more than gene 1 in environment 1 from [Formula 2.13] and [Formula 2.14]. There is more information in appendix.

The \(v_{11}\) parameter means how much gene 1 affect gene 2 in environment 1, \(v_{12}\) parameter means how much gene 2 affect gene 1 in environment 1, \(v_{21}\) parameter means how much gene 1 affect gene 2 in environment 2, and \(v_{22}\) parameter means how much gene 2 affect gene 1 in environment 2. Also, we set the \(Z_{11} = 0\), \(Z_{12} = 1\), \(Z_{21} = 1\), and \(Z_{22} = 0\) in the [Formula 2.13]
] and [ Formula 2.14 ]. Thus, to obtain a high fitness, in the environment 1, v11 should be small, v12 should be large, and in the environment 2, v21 should be large, v22 should be small in the environment 2. Finally, [ Figure 2.13 ] shows same results with what would be expected from these formulas. Thus, we could prove our simulation is working well from same results.

We analyzed the relation of v and $\mu$ parameters and we assumed other parameters didn’t influence the fitness. From the [ Formula 2.3 ], we could analogize that

$$\frac{dG_1}{dt} = \frac{v_{21} + (\alpha_{21} \times G_2)}{(G_1^{m1} + K_1)} - \mu_1 \times G_1 = 0,$$ 

(2.15)

Because we set gene 1 is activated more than gene 2 in environment 1, we could assume $G_1 = 1, G_2 = 0$. Then [ Formula 2.15 ] can be changed like this:

$$\frac{v_{21} + 0}{(1 + K_1)} - \mu_1 = 0,$$ 

(2.16)

$$\mu_1 \times (1 + K_1) = v_{21}$$ 

(2.17)

Therefore, we could expect $v_{21}$ has direct proportion with $\mu_1 \times (1 + K_1)$ because gene 1 is activated more in environment 2. Also, from the [ Formula 2.4 ], we could expect $v_{12}$ has direct proportion with $\mu_2 \times (1 + K_2)$ because gene 2 is activated more in environment 1.

[ Table 2.5 ] shows the relation of $v_{21}$ and $\mu_1 \times (1 + K_1)$, and $v_{12}$ and $\mu_2 \times (1 + K_2)$. It’s not perfectly proportional but [ Figure 2.15 ], and [ Figure 2.19 ] show the direct proportional relation of v and $\mu$ parameters.

We made [ Table 2.5 ] to show the relations of $v_{21}$ and $\mu_1 \times (1 + K_1)$, and $v_{12}$ and $\mu_2 \times (1 + K_2)$, then we draw [ Figure 2.15 ], and [ Figure 2.19 ] to prove the relations of $v_{21}$ and $\mu_1 \times (1 + K_1)$, and $v_{12}$ and $\mu_2 \times (1 + K_2)$. There are some points are out of the direct proportion, and we draw [ Figure 2.14 ], and [ Figure 2.18 ] to find the pattern of these points.

There are several lines are out of pattern in [ Figure 2.13 ]. We think it is possible because it is multiple parameters combination. Thus, sometime, it is happens that $v_{11}$ is large, $v_{12}$ is small, or $v_{21}$ is small, and $v_{22}$ is large when we obtain a high fitness. We choose five cases from [ Figure 2.15 ]. [ Figure 2.14 ] shows the parameters of out of pattern points. In the [ Figure 2.14 ], blue lines are points have directly proportion and red lines indicate points are out of the pattern. We could find the concentration of gene 1 and gene 2 are opposite, and $v_{21}$
and $v_{22}$ is big changed in environment 2. To find the difference between these points, we picked up 30, 60 points. Then [Figure 2.16, Figure 2.17] indicate the change of genes expression in real time. We found that both genes are expressed well in both environments in [Figure 2.16], on the other hand, we found that gene 2 is expressed well in environment 1 but gene 1 is not expressed well in environment 2 from [Figure 2.17].

Also, [Figure 2.18] shows the parameters of out of pattern points, and red lines are points are out of pattern of [Figure 2.19] and blue lines are points have directly proportion. From [Figure 2.18], we could find the concentration of gene 1 and gene 2 are opposite, and $v_{11}$ and $v_{12}$ is big changed in environment 1. We could find the difference between points have directly proportion and points are out of the pattern from [Figure 2.20, Figure 2.21]. [Figure 2.20] indicates the change of genes expression and both genes are expressed well in both environments, on the other hand, [Figure 2.21] shows that gene 1 is expressed well in environment 2 but gene 2 is not expressed well in environment 1 from [Figure 2.21].

Thus, we could conclude points are out of the pattern indicate one of genes is not expressed well in one of environments. These figures [Figure 2.14, Figure 2.15, Figure 2.18, Figure 2.19, Figure 2.16, Figure 2.17, Figure 2.20, Figure 2.21] explain why there is several lines are out of the pattern of parameters [Figure 2.13].
26

Figure 2.13  200 times simulation: Blue line is a simulation with the initial parameters. Red line is other 60 simulation results. The cut-off value is 0.6. After collecting high fitnesses, we found the pattern of parameters inducing a high fitness.
Figure 2.14  Special cases of relations $v_{21}$ and $\mu_1(1 + K1)$ parameters. $v_{21}$ and $v_{22}$ is big changed in environment 2.

Figure 2.15  $v_{21}$ and $\mu_1(1 + K1)$ parameters are in direct proportion.
Figure 2.16 Red line is the change of concentration of gene 1 and blue line is the change of concentration of gene 2. Point 30 shows both genes are expressed well in both environments.

Figure 2.17 Red line is the change of concentration of gene 1 and blue line is the change of concentration of gene 2. Point 60 shows gene 2 is expressed well in environment 1 but gene 1 is not expressed well in environment 2.
Figure 2.18  Special cases of relations $v_{12}$ and $\mu_2(1 + K_2)$ parameters. $v_{11}$ and $v_{12}$ is big changed in environment 1.

Figure 2.19  $v_{12}$ and $\mu_2(1 + K_2)$ parameters are in direct proportion.
Figure 2.20  Red line is the change of concentration of gene 1 and blue line is the change of concentration of gene 2. Point 40 shows both genes are expressed well in both environments.

Figure 2.21  Red line is the change of concentration of gene 1 and blue line is the change of concentration of gene 2. Point 16 shows gene 1 is expressed well in environment 2 but gene 2 is not expressed well in environment 1.
Figure 2.22  X-axis is $v_{21} - mu_1(1+K_1)$, and y-axis is $v_{12} - mu_2(1+K_2)$.

We drew figures[ Figure 2.22, Figure 2.23, Figure 2.24 ], then we could check that gene 1 is expressed more than gene 2 in the environment 2 from [ Figure 2.24] and gene 2 is expressed more than gene 1 in the environment 1 from [ Figure 2.23]. Also we already show both genes are expressed well in both environments in 30 point from [ Figure 2.16].
Figure 2.23  Red line is the change of concentration of gene 1 and blue line is the change of concentration of gene 2.

Figure 2.24  Red line is the change of concentration of gene 1 and blue line is the change of concentration of gene 2.
Table 2.5  Relations of $v$ parameter and $\mu$ parameter: $v_{21}$ is directly proportional to $\mu_1 \times (1 + K_1)$, and $v_{12}$ is directly proportional to $\mu_2 \times (1 + K_2)$.

<table>
<thead>
<tr>
<th>Simulation</th>
<th>$v_{21}$</th>
<th>$\mu_1 \times (1 + K_1)$</th>
<th>$v_{12}$</th>
<th>$\mu_2 \times (1 + K_2)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.0000</td>
<td>0.7500</td>
<td>1.0000</td>
<td>0.7500</td>
</tr>
<tr>
<td>2</td>
<td>1.3786</td>
<td>1.3856</td>
<td>1.2014</td>
<td>1.0688</td>
</tr>
<tr>
<td>3</td>
<td>0.4297</td>
<td>0.4748</td>
<td>1.2413</td>
<td>1.4935</td>
</tr>
<tr>
<td>4</td>
<td>1.1212</td>
<td>1.1729</td>
<td>1.9755</td>
<td>1.3752</td>
</tr>
<tr>
<td>5</td>
<td>1.5027</td>
<td>1.1694</td>
<td>0.6267</td>
<td>0.6865</td>
</tr>
<tr>
<td>6</td>
<td>1.9480</td>
<td>1.8038</td>
<td>1.3984</td>
<td>1.2318</td>
</tr>
<tr>
<td>7</td>
<td>1.4325</td>
<td>1.4612</td>
<td>0.9230</td>
<td>1.0519</td>
</tr>
<tr>
<td>8</td>
<td>1.5201</td>
<td>1.7061</td>
<td>1.1105</td>
<td>1.1870</td>
</tr>
<tr>
<td>9</td>
<td>1.1777</td>
<td>1.1162</td>
<td>0.7490</td>
<td>1.3378</td>
</tr>
<tr>
<td>10</td>
<td>1.8521</td>
<td>1.7650</td>
<td>0.5182</td>
<td>1.7864</td>
</tr>
<tr>
<td>11</td>
<td>0.4020</td>
<td>0.8440</td>
<td>1.0994</td>
<td>1.3158</td>
</tr>
<tr>
<td>12</td>
<td>1.7802</td>
<td>1.6434</td>
<td>0.5071</td>
<td>0.8802</td>
</tr>
<tr>
<td>13</td>
<td>0.9958</td>
<td>1.0596</td>
<td>1.2962</td>
<td>1.2367</td>
</tr>
<tr>
<td>14</td>
<td>1.1980</td>
<td>1.4508</td>
<td>0.2268</td>
<td>1.2148</td>
</tr>
<tr>
<td>15</td>
<td>1.6851</td>
<td>1.3855</td>
<td>0.7961</td>
<td>0.7616</td>
</tr>
<tr>
<td>16</td>
<td>0.6580</td>
<td>1.0273</td>
<td>0.2609</td>
<td>1.7202</td>
</tr>
<tr>
<td>17</td>
<td>0.6557</td>
<td>0.7504</td>
<td>0.3472</td>
<td>0.9917</td>
</tr>
<tr>
<td>18</td>
<td>1.5867</td>
<td>1.8843</td>
<td>1.9245</td>
<td>1.5849</td>
</tr>
<tr>
<td>19</td>
<td>0.7827</td>
<td>0.8452</td>
<td>1.2929</td>
<td>1.7374</td>
</tr>
<tr>
<td>20</td>
<td>1.9706</td>
<td>1.7550</td>
<td>1.3304</td>
<td>1.6223</td>
</tr>
<tr>
<td>21</td>
<td>0.0177</td>
<td>1.0552</td>
<td>0.8552</td>
<td>1.0345</td>
</tr>
<tr>
<td>22</td>
<td>1.9016</td>
<td>1.8243</td>
<td>1.1386</td>
<td>1.2775</td>
</tr>
<tr>
<td>23</td>
<td>0.4141</td>
<td>0.7095</td>
<td>1.1554</td>
<td>1.2722</td>
</tr>
<tr>
<td>24</td>
<td>0.7053</td>
<td>0.9656</td>
<td>1.2872</td>
<td>1.8486</td>
</tr>
<tr>
<td>25</td>
<td>0.5684</td>
<td>0.6131</td>
<td>0.3587</td>
<td>1.6680</td>
</tr>
<tr>
<td>26</td>
<td>1.6662</td>
<td>1.9410</td>
<td>0.8586</td>
<td>0.8428</td>
</tr>
<tr>
<td>27</td>
<td>1.3950</td>
<td>1.7909</td>
<td>1.2988</td>
<td>1.2247</td>
</tr>
<tr>
<td>28</td>
<td>1.5137</td>
<td>1.3419</td>
<td>0.4911</td>
<td>1.0410</td>
</tr>
<tr>
<td>29</td>
<td>1.6305</td>
<td>1.7275</td>
<td>0.2653</td>
<td>1.0222</td>
</tr>
<tr>
<td>30</td>
<td>1.0707</td>
<td>1.0548</td>
<td>1.4723</td>
<td>1.3013</td>
</tr>
<tr>
<td>Simulation</td>
<td>$v_{12}$</td>
<td>$\mu_1 \times (1 + K_1)$</td>
<td>$v_{12}$</td>
<td>$\mu_2 \times (1 + K_2)$</td>
</tr>
<tr>
<td>------------</td>
<td>---------</td>
<td>------------------------</td>
<td>---------</td>
<td>------------------------</td>
</tr>
<tr>
<td>31</td>
<td>1.3482</td>
<td>1.3199</td>
<td>0.8656</td>
<td>1.0216</td>
</tr>
<tr>
<td>32</td>
<td>0.7840</td>
<td>1.0235</td>
<td>0.2862</td>
<td>1.6779</td>
</tr>
<tr>
<td>33</td>
<td>1.6374</td>
<td>1.5268</td>
<td>0.8555</td>
<td>0.8588</td>
</tr>
<tr>
<td>34</td>
<td>1.2773</td>
<td>1.6578</td>
<td>1.1085</td>
<td>0.9441</td>
</tr>
<tr>
<td>35</td>
<td>1.2588</td>
<td>1.3898</td>
<td>1.2990</td>
<td>1.3598</td>
</tr>
<tr>
<td>36</td>
<td>1.6759</td>
<td>1.5682</td>
<td>1.3358</td>
<td>1.0352</td>
</tr>
<tr>
<td>37</td>
<td>0.7564</td>
<td>0.8600</td>
<td>1.6547</td>
<td>1.5547</td>
</tr>
<tr>
<td>38</td>
<td>1.4645</td>
<td>1.2947</td>
<td>0.8362</td>
<td>1.7974</td>
</tr>
<tr>
<td>39</td>
<td>1.6780</td>
<td>1.8485</td>
<td>0.9191</td>
<td>1.2472</td>
</tr>
<tr>
<td>40</td>
<td>1.0376</td>
<td>1.3758</td>
<td>0.7550</td>
<td>0.7220</td>
</tr>
<tr>
<td>41</td>
<td>0.7559</td>
<td>0.7025</td>
<td>0.0300</td>
<td>1.5501</td>
</tr>
<tr>
<td>42</td>
<td>1.0740</td>
<td>1.2857</td>
<td>0.9955</td>
<td>1.0747</td>
</tr>
<tr>
<td>43</td>
<td>0.7121</td>
<td>0.8836</td>
<td>0.9773</td>
<td>1.7675</td>
</tr>
<tr>
<td>44</td>
<td>0.0273</td>
<td>0.6329</td>
<td>1.2480</td>
<td>1.0930</td>
</tr>
<tr>
<td>45</td>
<td>1.1812</td>
<td>1.3164</td>
<td>0.6622</td>
<td>1.5462</td>
</tr>
<tr>
<td>46</td>
<td>1.0127</td>
<td>1.0454</td>
<td>1.5690</td>
<td>1.4778</td>
</tr>
<tr>
<td>47</td>
<td>0.8284</td>
<td>1.2932</td>
<td>0.1372</td>
<td>1.5615</td>
</tr>
<tr>
<td>48</td>
<td>1.5828</td>
<td>1.6675</td>
<td>0.7339</td>
<td>0.8263</td>
</tr>
<tr>
<td>49</td>
<td>0.4247</td>
<td>0.5463</td>
<td>0.4799</td>
<td>1.1163</td>
</tr>
<tr>
<td>50</td>
<td>0.8656</td>
<td>0.8622</td>
<td>0.0975</td>
<td>1.2544</td>
</tr>
<tr>
<td>51</td>
<td>0.7295</td>
<td>1.4961</td>
<td>1.5238</td>
<td>1.5548</td>
</tr>
<tr>
<td>52</td>
<td>0.0137</td>
<td>0.6762</td>
<td>1.8351</td>
<td>1.9287</td>
</tr>
<tr>
<td>53</td>
<td>1.5415</td>
<td>1.2135</td>
<td>1.9261</td>
<td>1.7206</td>
</tr>
<tr>
<td>54</td>
<td>0.0387</td>
<td>1.6347</td>
<td>1.0435</td>
<td>0.9590</td>
</tr>
<tr>
<td>55</td>
<td>1.1190</td>
<td>0.8942</td>
<td>0.5574</td>
<td>0.9124</td>
</tr>
<tr>
<td>56</td>
<td>1.3778</td>
<td>1.0634</td>
<td>0.4936</td>
<td>0.5299</td>
</tr>
<tr>
<td>57</td>
<td>0.4303</td>
<td>1.5068</td>
<td>0.8272</td>
<td>1.0204</td>
</tr>
<tr>
<td>58</td>
<td>1.1092</td>
<td>1.4861</td>
<td>1.0146</td>
<td>1.1579</td>
</tr>
<tr>
<td>59</td>
<td>1.2589</td>
<td>1.0825</td>
<td>1.1916</td>
<td>1.1831</td>
</tr>
<tr>
<td>60</td>
<td>0.7848</td>
<td>1.6060</td>
<td>0.8708</td>
<td>0.9558</td>
</tr>
<tr>
<td>61</td>
<td>1.2739</td>
<td>1.5565</td>
<td>0.4591</td>
<td>1.1904</td>
</tr>
</tbody>
</table>
2.5 Discussion

2.5.1 A high fitness of gene network model

We described 3 different gene network models by ordinary differential equations. The first model is gene 1 activates gene 2, and gene 2 activates gene 1, but each of them inhibits itself. The second model is gene 1 inhibits gene 2, and gene 2 inhibits gene 1, but each of them activates itself. Lastly, the third model is there is no interaction each other but each of them inhibits itself.

From these models, we tried to calculate a fitness. Even though there are many options and conditions for simulation, we found proper conditions and obtained a high fitness of each model from simulated annealing algorithm. Also, we found which parameters induce the high fitness, and how strongly these parameters influence the gene network models. Therefore, we found how the gene network model is evolved by other condition such as environment, and the concentration of genes.

We have two genes and describe two environments of each gene network model. We makes different gene interaction as we can do. All three cases show similar fitness values at the end of simulation. However, the process of evolution of each gene network model are different. Model 1 indicates the similar concentrations of gene 1 and gene 2 through the real time. Because the model 1 is designed both genes activate each other. On the other hand, the model 2 is designed both genes inhibit each other. Thus, one of gene’s concentration is much higher than other. Also, after simulation, we found the concentration of gene 1 of model 1 is higher than the concentration of gene 1 of model 2 from the optimal parameters.

We developed fitness functions, and set gene 1 activates more in environment 2, and gene 2 activates more in environment 1. Then, we applied simulated annealing algorithm to obtain a high fitness. We could prove the simulated annealing algorithm is working well in this simulation by results. Because we mostly obtain same results with what would be expected from formula. Tomshine and Kaznessis also explored the gene network model using simulated annealing algorithm (Tomshine and Kaznessis, 2006). They studied three genes and described
the complete network of reactions well, but they didn’t mention about evolutionary concept at all. On the other hand, we show how gene network models are evolve in evolutionary time from the track of fitness. Also, we can reduce a running time from finding properly initial temperature for simulation.

2.5.2 Parameters

We simulated 200 times to find the pattern of parameters. We could find the pattern: \( v_{11} \) is small, \( v_{21} \) is large, \( v_{21} \) is large, and \( v_{22} \) is small and \( \mu_1 \), and \( \mu_2 \) are always large. However, there are special cases out of the patterns. Thus, we picked five cases of them and show what the difference is between special cases and the pattern of parameters [Figure 2.14, and Figure 2.18]. The pattern of parameters indicates the maximal rate of activation of gene, \( v_{12} \) is higher than \( v_{11} \) in environment 1, and \( v_{21} \) is higher than \( v_{22} \) in environment 2. However, the special cases of [Figure 2.15] show \( v_{22} \) is higher than \( v_{21} \) in environment 2, and the special cases of [Figure 2.19] show \( v_{11} \) is higher than \( v_{12} \) in environment 1. Even though it is not easy to explain why these special cases can make high fitnesses, it’s possible because the fitness is made by many parameters combinations.

2.5.3 Computation

We simulated a program based on the simulated annealing algorithm, and it took nearly around 18 minutes for each model. The simulation iterated 5000 times to find the optimal parameters and a high fitness of each gene network model. It didn’t take a long time, but when we simulated 200 times to find the pattern of parameters [Figure 2.13], it took 24.08 hours. Thus, the running time is still a problem. However, we could save the running time by finding of the properly initial temperature.

2.5.4 Conclusion

There are many modeling techniques for gene regulatory network. However, we choose mathematical method because, we thought it makes easy to understand these dynamic pro-
cesses. We described several gene network models, showed the dynamic process of gene network, and explored the gene network from evolutionary view point. From this research, we could find how gene network evolve with parameters and two different environments, and we could apply to design and optimize other gene network. Furthermore, these findings of proper parameters and a high fitness of gene network model, and the track of fitness are useful to analysis of the evolutionary gene network.

However, we think the gene network consisted of two genes is still small to apply a real complicated gene network. Thus, we will expand our research to deal with several genes. We thought about duplication concept. Gene duplication is one way to create new genes in genomes, and regulatory interactions are inherited from the ancestral genes after duplication. Thus, a duplicated gene has the same components of interaction. We will apply the duplication concept to expand the gene network model. Finally, we believe that this research will help to analyze the evolutionary gene networks.
2.6 Reference


Ideker, T., Thorsson, V., Ranish, J. A., Christmas, R., Buhler, J., Eng, J. K., Bumgarner,


CHAPTER 3. GENERAL CONCLUSIONS

In this thesis, we explored the evolutionary gene network. We could develop 3 different simple gene network models, and could describe these networks by mathematical approaches. From the evolutionary viewpoint, we figured out high fitnesses of gene network models, and tracked the fitness to show how gene networks are evolved. We used ordinary differential equations because it is useful to apply evolutionary time. Also, when we simulated to obtain a high fitness of gene networks, we applied simulated annealing algorithm. The simulated annealing algorithm is fit for finding good combinations of parameters that produce the high fitness of the gene network and it is the closest analogy with the shifting balance theory in populations (Kirkpatrick, et al., 1983). It’s not easy to find proper conditions for simulation, but we found empirical values from many simulations.

After we developed the fitness formula, we could expect the results, then we obtain same results with what would be expected from simulation. Thus, we could prove our formula is working well and the simulation is working well from the results.

In summary, 3 different gene network models depicted to ordinary differential equations and simulation applied simulated annealing algorithm. We could show how each model evolve with two environments in evolutionary time, and how the parameters of each model affect to the fitness of the gene network model. Furthermore, this study can be applied to design and optimize other gene networks. Therefore, we think this research can encourage the evolutionary gene network study. Also, we think this research is a good start to expand analysis of the evolutionary gene network.
APPENDIX. ADDITIONAL MATERIAL

Track of parameters

We simulated with an initial temperature, 0.4305. First, we simulated with the initial temperature, 1 and we choose 38 times fitness point from [Figure 1]. [Figure 2] indicates a result of simulation with the initial temperature 0.4305 and finally it shows the fitness clime up until 0.8414. The previous simulation with the initial temperature, 1 indicates the fitness is 0.8668 at the end of simulation. Both fitnesses are close to each other. Thus, this example also proves we find the properly initial temperature.

There are 3 figures indicate the track of other parameters of 3 different gene network models [Figure 3, Figure 4, and Figure 5].

![Figure 1](image.png)  
Figure 1  Track of fitness of model 1 with the initial temperature,1.
Figure 2 Track of fitness of model 1 with the initial temperature, 0.4305.
Figure 3 Track of Parameters of Model 1
Figure 4  Track of Parameters of Model 2
Figure 5  Track of Parameters of Model 3
Figure 6  200 times simulation: Blue line is a simulation with the initial parameters. Red line is other 166 simulation results. the cutoff value is 0.4

200 simulations to show the pattern of parameters

We simulated 200 times and set the cutoff value is 0.4. 0.4 is not big enough but it makes the pattern of parameters is much clearly. [ Figure 6] shows same pattern with [ Figure 2.13 ].

Also, there are 4 other figures to explain why there are several lines are out of the pattern of parameters in [ Figure 6 ].
Figure 7 Special cases of relations v21 and $\mu_1(1+K1)$ parameters.
Figure 8  $v_{21}$ and $\mu_1(1+K1)$ parameters are in direct proportion.
Figure 9: Special cases of relations v12 and \( \mu_2(1+K2) \) parameters.
Figure 10  $v_{12}$ and $\mu_2(1+K2)$ parameters are in direct proportion.
BIBLIOGRAPHY


