Continuous Mammalian Cortical Area Development Model: Transforming From Qualitative To Quantitative Model

Mohammad A. Alsharaiah, Mutaz K. Khazaaleh, Laith H. Baniata

Abstract: The understanding of mammalian cortical area development network has long been a core objective in Systems Biology. The cerebral cortex is spliced into many functionally distinct areas. The evolution of these areas during neural growth is relied on a set of genes expression patterns. For instance, through the anterior-posterior axis, gradients of set of genes such as Fgf8, Emx2, Pax6, Coup-tfl, and Sp8 control the role in specifying a real identity. The Knowledge about this network is mainly of qualitative nature and incomplete. Therefore, we utilized a computational method to understand the complex dynamic interactions behavior between the aforementioned genes. There is a need to understand which interactions, and other combinations of interactions would be appeared in networks that mimic the anterior-posterior expression patterns. In addition, a continuous concentration measurements are considered as significant indicators for proteins activities. The concentration level during the interaction can provide more insight to understand system activities and supports both of diagnosis and drugs treatment studies. Thus, to achieve this task, herein, we present a two-step approaches to facilitate the computational model. The first step is employing a Boolean network model since the available knowledge in qualitative form, this model mimics the genes interactions. The Boolean model treats expression levels as Boolean since the nature of the expression data is currently available in the qualitative form. However, even the discrete model is a powerful tool to study and understand the dynamic behavior of genes interaction, but it can never provide detailed time courses of concentration levels. Therefore, in the second step, we transformed the discrete-Boolean model to continuous simulation to reproduce a continuous protein concentration. We engaged a standard method such as multivariate polynomial interpolation to transform the logic operations to ordinary differential equations model (ODE).

Index Terms: mammalian cortical development, modeling, qualitative model, quantitative model, ordinary differential equations model (ODE), multivariate polynomial interpolation, Hill-cube.

1. INTRODUCTION

His research investigate a very significant biological system such as mammalian cerebral cortex. It is a complex system that has a precise structure which is also separated into several functionally distinct areas [2]. These areas are controlled by set of different combinations of gene. Precisely, the available two-dimensional field experiments such as anterior-posterior patterning have recognized many genes and expressed embryonically which are serious to the locating of cortical areas in adult [3]. Therefore, we interested in the patterning that can be found in this axis For instance, in mammalian system such as in the embryonic day 8 (E8) for the mouse, a special gene called morphogen (Fgf8) is expressed exactly at the anterior pole of the developing telencephalon (Figure 1A) [4] , [5]. The figure shows the anterior neural ridge or commissural plate (blue), it demonstrate the patterning centre in the developing forebrain that secretes the morphogen Fgf8. Since the protein is secreted, it is hypothesized that it diffuses to form a gradient [5]. As a consequent, and instantly after Fgf8 expression is initiated in mouse, set of transcription factors (TFs) - Emx2, Pax6, Coup-tfi and Sp8- are expressed in gradients across the surface of the cortex (Figure 1B) [6] [7]. The literature proved that these four TFs could provide a unique coordinate system for realization [8]. In addition, applying any modification of these genes can effect the development process of the system. For instance, altering Fgf8 and the four TFs shifts area positions in late embryonic stages and in adult [9][10].

Fig. 1A. Gene expression in the developing neocortex[4]. Where: the directions which named by A, P, D, V, indicate Anterior, Posterior, Dorsal, and Ventral individually.

Fig. 1B. the four transcription factors, expressed in spatial mRNA and protein gradients through the developing forebrain [5]. M and L indicate Medial and lateral individually.

• Dr. Mohammad Abdillah Alsharaiah, Ph.D. degree in computer science and information systems from Lincoln University, Canterbury, New Zealand, E-mail: MOHdsharaiah@gmail.com.
• Dr. Mutaz Khazal Khazaaleh, Ph.D. degree in computer science and information systems from Lincoln University, Canterbury, New Zealand, lecturer in Computer Science Department, AlBaqa’ Applied University, Jordan, E-mail: Mutaz.khazaaleh@bau.edu.jo.
• Dr. Laith Hasan Baniata, was born in Amman, Ph.D. degree in computer science and engineering from Kyungpook National University, South Korea, E-mail: L.baniata@ju.edu.jo.
However, in this research we explore the mammalian cortical area development network in supplementary details due to scarce models that available in the literature [1]. Furthermore, herein we present new productive method which can provide new protein quantitative data. This data explains the intermediate states for portions activities. Protein intermediate state has not be explained in the previous available models. Therefore, the next section explain the method to generate mammalian cortical area development network, then, the next section exposes the model design and simulation. After that, result and discussion are coming. Then and there, transforming the discreet model to continuous model are coming in specific section within their results and discussions. Last, conclusion section is presented in the final section.

2 CONSTRUCTING THE SIX GENES NETWORK INTO BOOLEAN LOGIC FUNCTIONS

Boolean network model has been utilized to demonstrate the dynamic behaviors of biological systems and to modernize biological networks underlying specific functions [12]. The nature of regulatory interaction between these genes provides a complex behavior, and this behavior is hard to understand and therefore more insights are needed to be discovered [13]. Computational models can be helpful in the analysis of this kind of system and can offer greater understanding and insight. However, few models can be found in the literature which explore the regulatory interactions between these five genes such as [1]. Furthermore, there are few analysis that has been performed at the systems level [14]. The nature of the available data for cortex system can be found in the qualitative form. This inspired the biological research to design a Boolean model to construct and study the dynamic behavior for the five genes. Boolean models can be seen as the mathematically demanding illustration of qualitative biological knowledge. Their components, called species, can have only discrete states, usually two; as 1 or 0 ,ON or OFF, present or absent, ‘deactivated’ and ‘activated’. In addition, this kind of model has the ability to capture the essential behavior of a network. [1]. We applied the biological assumption for this model in our study to capture the regulatory relationships between the five genes. Also, we included a new node called “Start “node, to work as an indicator for the days 8. This node is considered as a starting point for our discrete dynamical network over the simulation and it can help to reflect the desired steady state expression levels in the anterior and posterior compartments inside the Boolean model. Table 1 shows all the biological interaction in the system over the development process, where the network was turned into a set of Boolean logic functions. The regulation rule for each component is defined by the Boolean logical operators; AND, OR and NOT.

<table>
<thead>
<tr>
<th>Rule</th>
<th>START=START</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rule 2</td>
<td>Fgf8=START</td>
</tr>
<tr>
<td>Rule 3</td>
<td>Emx2= ~Fgf8 &amp; &amp; ~Sp8 &amp; &amp; Coup_tfi &amp; &amp; ~Sp8</td>
</tr>
<tr>
<td>Rule 4</td>
<td>Pax6= ~Emx2 &amp; &amp; ~Coup_tfi &amp; &amp; Sp8</td>
</tr>
<tr>
<td>Rule 5</td>
<td>Coup_tfi= ~Fgf8 &amp; &amp; ~Sp8</td>
</tr>
<tr>
<td>Rule 6</td>
<td>Sp8=Fgf8 &amp; &amp; ~Emx2</td>
</tr>
</tbody>
</table>

3 BOOLEAN MODEL DESIGN AND SIMULATION

The main objective of this study was to unravel the qualitative characteristics of the dynamic behaviour of mammalian cortical area development network. This investigation can be done using synchronous Boolean approach. The synchronous Boolean approach updates all elements (nodes) in the network simultaneously at each time point. The updating of state for each node is controlled by Boolean rules that are defined earlier in the previous section. This study involves synchronous modelling to produce global dynamics of the system. The results are compared with the latest models such as [1] which provided the base for model advancements in this study. The constructed model must be able to provide results that are in agreement with literature and prior experimental observations; in particular it should reproduce the observed enhancement as well as the essential components of the network that are needed to accomplish system development process. In this study, a simulation model was made using open source software packages such as Odefy-Matlab software Package [1]. In the Boolean network, if we have N species X₁, X₂, ..., Xₙ, where these species represent genes, proteins, etc., each is represented by a variable Xi and taking values in {0, 1}. For each species Xi a set of species RI = {X₁, X₂, ..., Xₙ} ⊂ {X₁, X₂, ..., Xₙ} that influence xi are included and for each species Xi an update function is used to give a new value of Xi at the next step for every possible combination of (Xi₁, Xi₂, ...., Xiₙ) ∈ {0,1}ᵣ.

Each node in the Boolean network represents a biological species. Si(t∈{0,1}) signifies the state of node i at time t. Furthermore, if the state of a node is 1 (On) at time t and one of its inhibitor is activated, then Si(t + 1) = 0 regardless of all of the activation terms. The simulation for the Boolean model can be started when we defined the initial condition of the model. And for this, we assumed the biological start condition is to select all nodes with OFF (inactive) state, while the proteins stat condition is ON (active). The model simulation can be made more efficient if initial conditions are known. It is always preferable to start the simulation with relevant initial conditions if possible. There is agreement in literature about some of the elements that are present and active.
We have done two different simulations with different initial state. Each initial state has six digits of 1’s and 0’s representing the state of the network species (nodes) [START, Fgf8, Emx2, Pax6, Coup_tfi, Sp8]. For illustration, the first initial states [0,0,0,0,0,0] represents the silent states for network activity, while the second initial states [1,0,0,0,0,0] denotes START protein is active, while all other proteins are inactive and this states can fire the network simulation.

4 RESULTS AND DISCUSSION

The simulation result shows that the first initial states [0,0,0,0,0,0] - (the silent states for network activity when all proteins are silent; the 0 for each protein indicates that the state is not active) - changed the posterior starting state [0,0,0,0,0,0] to the posterior steady state [0,0,1,0,1,0]; this is our first finding that it recognizes the factors that show the stable state. The Boolean descriptions of the desired posterior steady states are shown in Figure 2 and Table 2. These result represent the posterior state where only Emx2 and Coup_tfi were active and the other elements were inactive. The result also shows that both of Emx2 and Coup_tfi are active, and this agree with the biological fact where the two aforementioned elements are expressed in high posterior–low anterior gradients [1]. Also, any change to/in these elements will not affect the dynamic behavior for the network. For example, if Emx2 and Coup_tfi are swapped, the dynamics of the network don’t change.

<p>| TABLE 2 |
| PROTEINS ACTIVITIES INSIDE THE POSTERIOR FOR THE MAMMALIAN CEREBRAL CORTEX |</p>
<table>
<thead>
<tr>
<th>START</th>
<th>Fgf8</th>
<th>Emx2</th>
<th>Pax6</th>
<th>Coup_tfi</th>
<th>Sp8</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

The second simulation provides another form of system activity. The initial states [1,0,0,0,0,0] denotes the START node is active and system activity are initiated. This state represents the case when the Boolean regulatory network is initiated through an external mechanism which make other species active such as Fgf8 [15]. Indeed, it resembles the early state in the anterior Compartent, specifically at around E8, where Fgf8 becomes active. This state provides the anterior steady state [1,1,0,1,0,1]. Moreover, this simulation mimics all proteins activity inside the anterior over development stage as illustrated in Table 3 and Figure 3.

<p>| TABLE 3 |
| PROTEINS ACTIVITIES INSIDE THE ANTERIOR FOR THE MAMMALIAN CEREBRAL CORTEX |</p>
<table>
<thead>
<tr>
<th>START</th>
<th>Fgf8</th>
<th>Emx2</th>
<th>Pax6</th>
<th>Coup_tfi</th>
<th>Sp8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

The results of this experiment demonstrate that Fgf8 activity at the anterior pole can control other system activity. For instance, Fgf8 is considered as a sufficient factor to represent the simple exist interactions between system transcription factors. It can drive the correct spatial patterning of Emx2, Pax6, Coup_tfi and Sp8. The result also shows how Fgf8 can play a durable change in the developing process of the brain [16]. Furthermore, the result explains the relations and the direct interaction between Fgf8, Emx2 and Coup_tfi. All of which lined up with literature results [17][18][19] and how it control the anterior and posterior.

In the simulation part we utilized a heuristic search based simulations. This type of simulation shows identical results which reveal realistic dynamics of controlling the proteins of interest in the anterior-posterior patterning of cortical areas. The synchronous simulation captured the main dynamics of the system species, such as Fgf8 domination on the network; also these results agree with experimental result that is available in the literature [1]. Information in our model is compatible with the biological knowledge from the models [1]. However the previous models have limitation that they are in qualitative form. For instance [1] model reported that there is a need for quantitative result to support the qualitative ones. They argue that the quantitative result can provide promising result to explain protein concentration measurements. Revising proteins concentration is essential to investigate the health and the growth of the brain cells and tissues. Accordingly, the next section presents the transformation of the discrete model to continuous model by utilizing a productive method.

5 TRANSFORMATION OF THE DISCRETE MODEL TO CONTINUOUS MODEL

The discrete models has the drawback of forcing the modelers
to find a new method which can cross over the limitations in the models. Consequently, continuous models among several general approaches have been introduced to make discrete models more productive. Almost all transformations approaches are based on multivariate polynomial interpolation [20], in order to obtain a continuous model from a Boolean one [21] (see Methods). For instance, the function Bi in equation 1 can be defined as a polynomial form after applying this conversion. In the first step we change the discrete variables xi by continuous variables \( \tilde{X}_i \) with values that have range between [0, 1], in this case we normalize proteins concentrations to the unit interval. Consequently, we assume to 'extend' the functions Bi to functions \( \tilde{B}_i : [0, 1]^N \rightarrow [0, 1] \). Then we call the functions \( \tilde{B}_i \) continuous homologues of the Boolean functions Bi. Mainly, the realistic continuous models should compute proteins concentration levels; this includes the production and degradation rates (Kinetic parameters) which is limited in the biological knowledge. Therefore, the adopt the assumption that the production of Xi is to be given by \( \tilde{B}_i \), while the degradation to be proportional to \( \tilde{X}_i \). Then the development of \( \tilde{X}_i \) over time is directed by the ordinary differential equation (ODE):

\[
\dot{X}_i = \left(\frac{y_i}{a_i}\right) \tilde{X}_i . \text{ it now follows that}
\]

\[
\tilde{X}_i = \frac{Y_i}{a_i} \tilde{X}_i = \frac{y_i}{a_i} \left( a_i \tilde{B}_i(\tilde{X}_1, \tilde{X}_2, ..., \tilde{X}_N) - Y_i \tilde{x}_i \right) = \frac{y_i}{a_i} \left( a_i \tilde{B}_i(\tilde{X}_1, \tilde{X}_2, ..., \tilde{X}_N) - \tilde{x}_i \right) \quad (2)
\]

And by setting \( r_i = \frac{y_i}{a_i} \) and \( \tilde{B}_i(\tilde{X}_1, \tilde{X}_2, ..., \tilde{X}_N) = \tilde{B}_i(\frac{a_i}{y_i} \tilde{X}_1, \frac{a_i}{y_i} \tilde{X}_2, ..., \frac{a_i}{y_i} \tilde{X}_N) \),

\[
\tilde{B}_i(\tilde{X}_1, \tilde{X}_2, ..., \tilde{X}_N) = \tilde{B}_i(z_1, z_2, ..., z_N) \quad (3)
\]

To obtain the ODEs, a new system of ODEs can be defined by replacing the Bi in equation (2) by the Hill-Cubes [20], Bi, the definition for the Hill function \( f_{ij} \) with parameters \( n_{ij} \) and \( k_{ij} \) for every interaction and state new functions:

\[
\tilde{B}_i^H(\tilde{X}_1, \tilde{X}_2, ..., \tilde{X}_N) = \tilde{B}_i( f_{i1}(\tilde{X}_1), f_{i2}(\tilde{X}_2), ..., f_{iN}(\tilde{X}_N) ) \quad (4)
\]

6 CONTINUOUS MODEL SIMULATION RESULTS AND DISCUSSION

Transforming the discreet model to continuous model by utilizing Hill-cube function [20] provides promising results. For instance, the simulation for the continuous model related to the posterior part demonstrates the intermediate states for the active proteins. Figure 4 elucidates how each element moved from the silent states to active state gradually, on the contrary of the discrete result that founded in Figure 2. Furthermore, Figure 5 illuminates the continuous changes in proteins concentration levels system species in posterior phase over the simulation. In addition, a continuous simulation has been done to mimics the anterior compartments. This simulation mimics all proteins activity inside the anterior over development stage as shown in Figure 6.
7 CONCLUSION

In this paper, we presented a two-step method that combines the discrete model and the continuous model to generate a new and useful continuous results rather than discrete results. In our approach, first, a Boolean network model was applied to generate dynamic networks that are capable of executing the target protein functions. Then, continuous transformation simulation was conducted to quantitatively explore protein concentration level in mammalian cortical area development model. This shows that from the continuous model one may obtain biological insights not evident from the discrete one. Furthermore, the presented methodology will assist the interaction between modeling and experiments. Besides, it compromises a straightforward technique to apply quantitative analysis methods to qualitatively defined systems. The utilized transformation methods provide a promise result, the result describes the continuous changes in proteins measurements levels over the development process, and it includes the intermediate state for proteins activities such as proteins intermediates states form inactive to active and vice versa inside the mammalian cerebral cortex. Herein, we crossed over the limitation form the previous (Clare E. Giacomantonio model [1] which provide only discrete result. Indeed, we offered a new continuous results for proteins concentration which exposes the development process for the mammalian cerebral cortex.

REFERENCES