University of California, Los Angeles

Teaching Mathematical Modeling of Protein Biosynthesis Using Animated Morphing Transitions

Doug Dang, BE 199 Course Report
March 20, 2018

Abstract
Using a very simplified cartoon of the DNA to protein production process as an example, this project aims to show how to convert a cartoon model into a directed graph model, and then into an ordinary differential equation model. This is accomplished by morphing the three separate models into one another, each model representing the same protein production process in a unique way. By isolating each step of the biosystems model and reconstructing each piece as a component of the mathematical model, the modeling process concepts are connected in a new manner, including explaining each component and where it comes from. By integrating animation and morphing tools into the development of mathematical models from biosystems cartoons, connections between observable, experimental biology and abstract, mathematical modeling methodologies are pedagogically facilitated. Work has begun on modeling a protein dimerization process in a positive feedback loop.

Introduction
Our premise is that systems modeling education can be facilitated by using a more inductive analysis approach differently than with traditional and deductive textbook education alone. Inductive instruction is defined as presenting specific examples or models and challenging students to notice generalizations and patterns to learn how to solve problems. In comparison, deductive instruction presents generalizations and asks students to apply these rules to specific scenarios.[1] By challenging students to think inductively, students are able to advance their perspective on problems, training as problem-solvers and critical thinkers; students who learn deductively and are less-equipped to tackle real-world problems.[2] Positive aspects to visual learning in education also tie into an enhanced learning experience for engineering students. Visual learning paired with traditional lecture learning coalesce to activate both the creative and the logical side of the brain, promoting a comprehensive view and analysis when collecting intaking information.[3] We develop visual animation and morphing tools here into a novel and inductive learning modality that directly links abstract modeling and biological systems concepts in a way that maximally facilitates learning.

Mathematical modeling of the DNA to protein production process and protein dimerization aims to combine two different academic disciplines by taking two organic, biological systems and making them predictable and quantitative through mathematics. By presenting the concept of mathematical modeling in computational biology through an educational module that physically morphs one graph to the next allows a student to consume the information by logical and visual means simultaneously. This process emulates project-based learning because it connects concepts, constructs relationships, and combines creative representation and traditional lecturing of technical material, promoting inductive analysis from one, specific model to general principles of mathematical modeling of biological systems.

Methods

To build the educational module, the qualitative structure and parameters of a simple model of DNA regulation of the protein production process in a cell was first defined in traditional cartoon form (figure 1). The model has three state variables, *DNA*, *mRNA*, and protein *P*, and five parameters: k_1 represents the rate coefficient for DNA transcribed into the mRNA for the protein P; k_2 represents the fractional degradation rate of mRNA in the cell; k_3 represents the rate coefficient for protein production regulated by mRNA; k_4 represents the fractional degradation rate for protein in the cell. Additionally, there is a substantial time delay in the pathway for protein production regulated by represented by a time delay constant, τ. The modeling process involves converting the cell diagram cartoon model (figure 1-2), then abstract directed graph model (figure 3), and finally into a maximally abstract set of modeling equations (figures 4-6).

The first step of the modeling process was to represent the DNA to protein production process through a cartoon cell diagram that illustrated the human cell, including the nucleus, ribosomes, and cytoplasm, as well as the main variables of the process, DNA, mRNA, and protein, with the corresponding coefficients, via Microsoft PowerPoint on slides two through four of the module.[4] After the cartoon model was defined, each variable is connected by a forward arrow paired with its respective production coefficient, along with the time delay in the mRNA to protein translation process. To illustrate the cartoon more simply, the original cartoon was re-conceptualized into a more abstract cartoon that more closely replicated the colors and animations used in the remainder of the module while taking out parts of the biological process that do not have influence on the mathematical representation (figure 1-2).

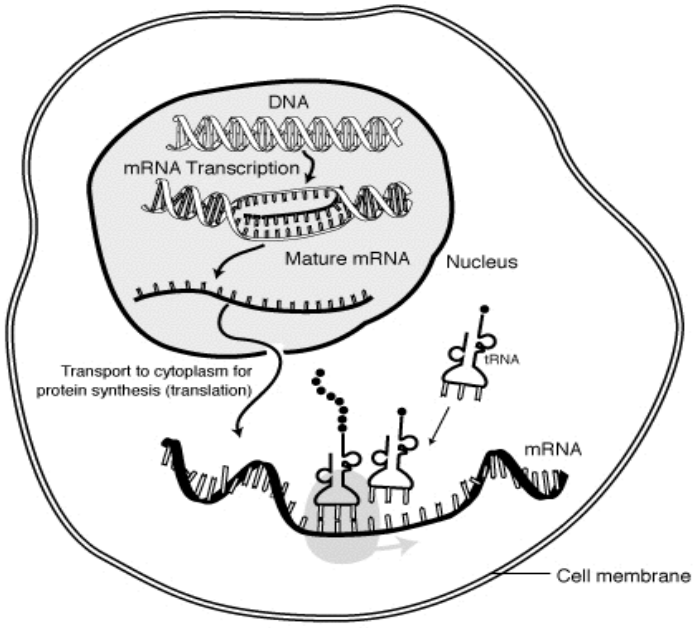

Figure 1: A cell diagram cartoon model representing a very simplified version of the DNA to protein production process.

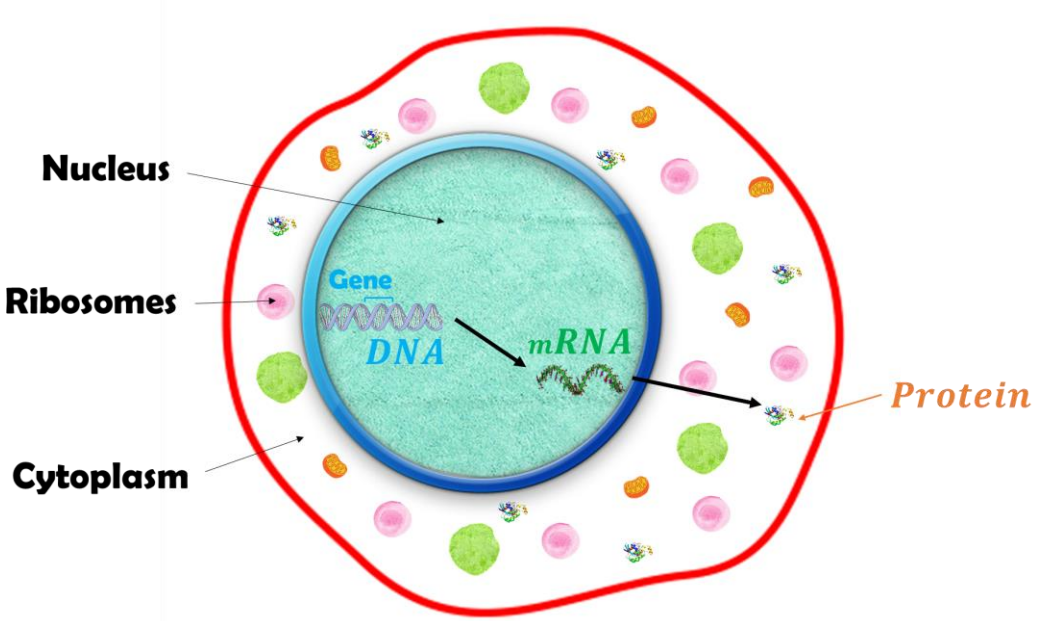

Figure 2: A more abstract cell diagram cartoon model representing a very simplified version of the DNA to protein production process.

In the following slides, five through seven, the cartoon model morphed into the directed graph model, illustrated by using the morph transition function on PowerPoint, which allowed the variables and arrows to physically migrate to respective positions in the directed graph model. Additionally, arrows that represented the degradation of mRNA and proteins were added to the directed graph model, along with their respective degradation rate coefficients (figure 2).

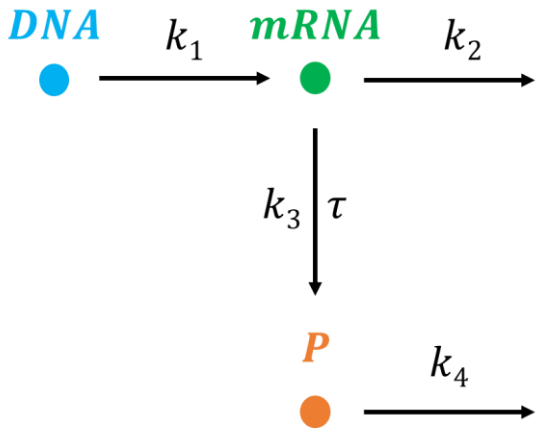

Figure 2: A directed graph model representing an abstracted, visual intermediary between the cartoon model and the mathematical model.

In slides eight to thirteen, the mathematical model is constructed by morphing the directed graph model to the modeling equations. Each variable is moved, along with its respective rate constant to its respective equations. The first equation of the model represents the change in mRNA concentration in the cell with respect to time, on the directed graph model, two processes represented by arrows, are considered. The arrow from the *DNA* variable to the *mRNA* variable

corresponds to a positive addition to the change in mRNA present in the cell with respect to time. The value of the positive addition is dictated by rate constant k_1 and *DNA* variable as a constant not depended on time. Next, the arrow beginning at the *mRNA* node and going to the right represents degradation of mRNA, which is a negative addition to the change in mRNA present in the cell with respect to time. The rate of the degradation is dictated by constant k_2 and *mRNA* variable as a function of time (figure 3).

$$\frac{dmRNA(t)}{dt} = k_1 DNA - k_2 mRNA(t)$$

Figure 3: The change in mRNA present in the cell with respect to time is dependent on rate constants k_1, k_2, and k_3 and variables DNA, M and M again as functions of time t, t, and t-τ, respectively.

The second equation composing the mathematical model represents the change in protein concentration in the cell with respect to time. On the directed graph model, there are two processes that need to be considered, represented by two arrows. The arrow from the *mRNA* variable to the *P* variable corresponds to a positive addition to the change in protein present in the cell with respect to time and time delay τ. The value of the positive addition is dictated by rate constant k_3 and *mRNA* variable as a function of time minus τ. For the last part of the model equation, the arrow beginning at the *P* node and going to the right represents degradation of protein, which is a negative addition to the change in protein present with respect to time. The value of the degradation is dictated by rate constant k_4 and *P* variable as a function of time (figure 4).

$$\frac{dP(t)}{dt} = k_3 mRNA(t-\tau) - k_4 P(t)$$

Figure 4: The change in protein present in the cell with respect to time is dependent on rate constants k_3 and k_4 and variables M and P as functions of time t and t-τ respectively.

On slide 13, the ordinary differential equations mathematical model is complete and illustrated adjacent to the directed graph model, showing each variable and constant in two different forms (figure 5). The abstraction is complete, but to further highlight the ODE mathematical model as the product of the morph, slide 14 shows the ODE alone (figure 6). Finally, in summary, slide 15 shows all three models: the simplified cartoon model, directed graph model, and ODE model, side-by-side to allow a complete flow from concept to concept (figure 7).

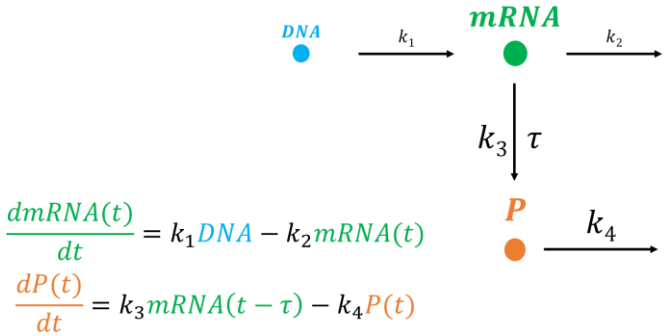

Figure 5: The mathematical model and directed graph model represented adjacently.

$$\frac{dmRNA(t)}{dt} = k_1 DNA - k_2 mRNA(t)$$

$$\frac{dP(t)}{dt} = k_3 mRNA(t-\tau) - k_4 P(t)$$

Figure 6: ODE mathematical model depicted to highlight the equations as the products of the morph.

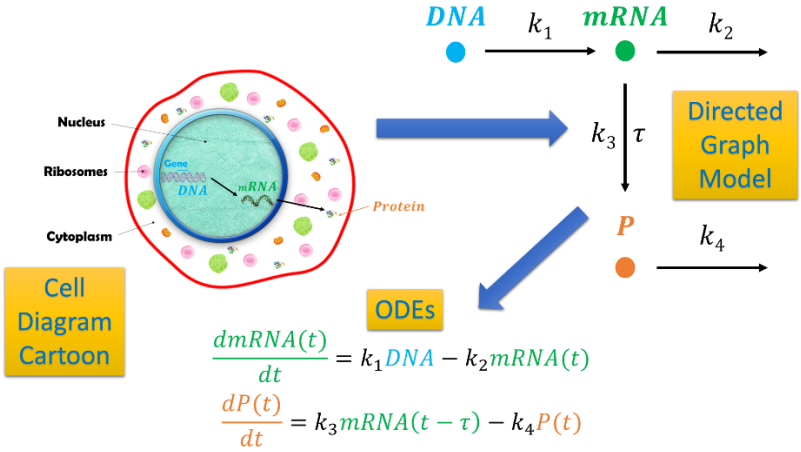

Figure 7: All three model stages (cartoon model, directed graph model, mathematical model) represented in one summary slide showing each graph and how it morphs into the next.

Discussion

The cell diagram cartoon model represents the biological process of DNA to protein production in a familiar way that clearly defined each process leading to the next and each variable having a direct influence on the rate of production of the next variable. Re-conceptualizing the cartoon model facilitates a more natural flow from biological process to cartoon to graph to mathematical model. The directed graph model serves as an intermediate step that facilitates transitioning of the qualitative cartoon model to the quantitative mathematical model. The directed graph model also serves to incorporate the degradation rate coefficients in a logical manner. The process of taking production and degradation process and taking the time to morph the graph to the equations showed each variable and coefficient and where they come from and how they function in each equation of the set and the set of equations altogether. Finally, the summary slides at the end of the module serves to reinforce the information displayed and show the relationship of each model to the other ones. The educational module project embodied the visualization of mathematical modeling and combining the traditional, spoken communication of computational biology and the computerized, visual representation of the same biological models. The module conveyed the information in a digestible way that is going to be used for and implemented in UCLA computational biology and computer science classes.

In the upcoming quarter, we will be investigating ways to make a template to automate biological mathematical modeling in a cell as well as protein dimerization in a positive feedback loop as another biological process represented by a Hill function in a set of more dynamic system of equations.

References

1. " Inductive and Deductive Instruction." Best of Bilash, Saskatchewan Education, sites.educ.ualberta.ca/staff/olenka.bilash/Best%20of%20Bilash/inductivedeductive.html.
2. Felder, Richard M. Learning and Teaching Styles in Engineering Education. Engr. Education, 1988.
3. Heywood, John. Engineering Education: Research and Development in Curriculum and Instruction. IEEE Press, 2005.
4. "Microsoft PowerPoint." Microsoft Corp., 2016.