

Designing Emergent Systems Microworlds to learn computational thinking in the context of synthetic biology

Sugat Dabholkar

Northwestern University

There exists a huge disparity between high school biology instruction and the research practices of modern biologists (Wilensky & Reisman, 2006). On one hand, the nature of biology research has changed significantly with the incorporation of newer technological tools and research methods. For example, the use of computational methods for modeling and data analysis, and an increasing focus on complex systems thinking have significantly changed the nature of research in biology, ranging from molecular genetic networks to ecological networks (Kitano, 2002; 2017). On the other hand, learning scientists have stressed the importance of having authentic scientific inquiry tasks in science curricula for students to learn disciplinary inquiry practices (Chinn & Malhotra, 2002). In my work, I seek to address this by combining two powerful design approaches in learning sciences, namely, agent-based modeling of emergent systems and constructionism (Wilensky, 2001; Kafai & Resnick, 1996). We call this design approach Emergent Systems Microworlds (ESM) (Dabholkar, Anton & Wilensky, 2018).

Emergent complex systems perspective involves understanding how simple interactions between autonomous elements can result in complex emergent patterns at the system level (Jacobson & Wilensky, 2006). This perspective of understanding several natural phenomena has become a focus of real-world scientific investigations as well as recent science education reforms (Yoon et al., 2018). Researchers of science education have argued for and demonstrated effectiveness of emergent systems perspective for understanding natural phenomena (Wilensky & Jacobson, 2015; Wilensky & Resisman, 2006; Hmelo-Silver & Azevedo, 2006).

In an ESM-based curriculum, students explore and learn about emergent phenomena, using agent-based computational models that are designed in NetLogo (Wilensky, 1999) in the form of a microworld. In such models, an agent is a computational object with particular properties and actions. An 'emergent' phenomenon is modelled in terms of agents and their interactions (Wilensky & Rand, 2015). Microworlds are encapsulated open-ended computational exploratory environments in which a set of concepts can be explored, through interactions that lead to knowledge construction (Papert, 1980, Edwards, 1995). ESMs are specifically designed to support students in exploring, and developing and sharing virtual models of systems that exhibit emergent phenomena. ESM-based curricula engage students in actively constructing knowledge in a computational microworld using scientific inquiry practices (SIPs) in the similar fashion as scientists construct knowledge about the real world (Figure 1).

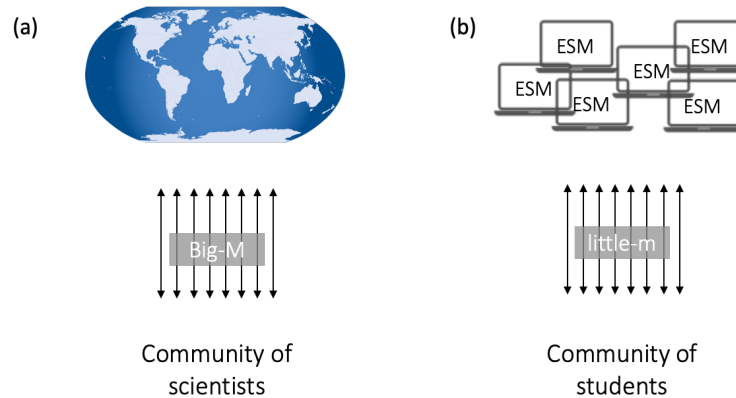


Figure 1: Knowledge co-construction using ESMs (Dabholkar & Wilensky, 2019)

(A) A community of scientists engaging in specific practices to construct knowledge about the world in the form of explanatory models (Big-M) (B) A community of students potentially engaging to construct knowledge in the form of their contextual and case-specific understandings (little-m) by interacting with an ESM

An Emergent Systems Microworld to Learn Synthetic Biology

iTune Computational Lab is an Emergent Systems Microworld designed using NetLogo (Wilensky, 1999) to be used in conjunction with the iTune Device Lab of the BioBuilder program (BioBuilder Educational Foundation, 2019). It uses a computational model of the same genetic circuit which is based on the Lac-Operon of bacterium *E. coli*. The computational model can be accessed at the following website: <http://tinyurl.com/itunecomplab>

By completion of this computational activity the students should be able to:

1. Explain how computational models can be used to investigate real-world phenomena.
2. Explain molecular genetics terms such as promoter, rbs, terminator in the context of synthetic biology.
3. Conduct a systematic investigation with a series of computational experiments by collecting and analyzing data.
4. Explain how stochastic/random variation in cellular processes result in differences in the final results.
5. Explain how scientists can identify general trends in patterns despite randomness in underlying biological processes.

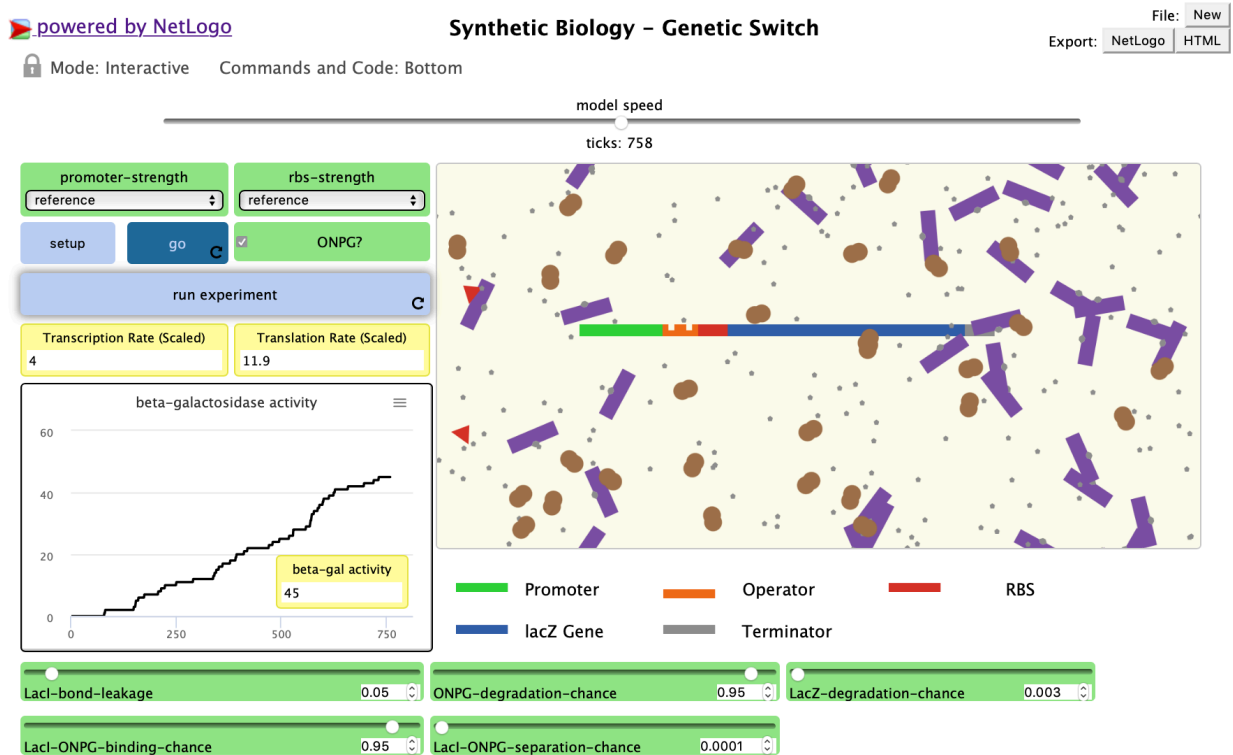


Figure 2: A screenshot of Synthetic Biology - Genetic Switch Model

Please refer to the appendix for model description. A user can select the promoter and RBS strengths, add ONPG (by making ‘CONST-ONPG’ ON) and run the model. The model simulates interactions between the components of the genetic circuit that results in an emergent cellular behavior. The cellular behavior of interest in this model is LacZ (beta-galactosidase) activity which can be observed in a graph and is also represented in the change in the color of the cell to yellow. Beta-galactosidase cleaves ONPG to produce an intensely yellow colored compound. Learners can use the computational model to understand the input-output relationship in a synthetically constructed genetic circuit. They can perform computational assays of enzyme activity of beta-galactosidase enzyme.

I have conducted two teacher professional development sessions about using this model as an Emergent Systems Microworld and use an ESM-based curricular unit to conduct a computational lab in conjunction with the biobuilders iTunes-device lab. I conducted interview of an educator who used this model in a classroom setting. However, I have not formally studied its use in an educational setting.

In this *learn.design.compute with bio* workshop, I will briefly discuss the idea of ESMs to engage students in computational thinking, my work regarding other ESMs that have been used and studied in biology classrooms (Dabholkar et.al, 2018; Dabholkar & Wilensky, 2019), and the Synthetic Biology - Genetic Switch model that I described in this abstract.

References

- Berland, L. K., Schwarz, C. V., Krist, C., Kenyon, L., Lo, A. S., & Reiser, B. J. (2016). Epistemologies in practice: Making scientific practices meaningful for students. *Journal of Research in Science Teaching*, 53(7), 1082-1112.
- Biobuilder Educational Foundation (2019) iTunes Device: Evaluate promoter and RBS combinations to optimize beta-galactosidase output (Retrieved from: <https://biobuilder.org/lab/itunes-device/>)
- Chinn, C. A., & Malhotra, B. A. (2002). Epistemologically authentic inquiry in schools: A theoretical framework for evaluating inquiry tasks. *Science Education*, 86(2), 175-218.
- Dabholkar, S., Anton, G., & Wilensky, U. (2018) GenEvo - An emergent systems microworld for model-based scientific inquiry in the context of genetics and evolution. Proceedings of the International Conference for the Learning Sciences, 2018, London, UK.
- Dabholkar, S. & Wilensky, U. (2019). Designing ESM-mediated collaborative activity systems for science learning. International Conference of Computer Supported Collaborative Learning 2019, Lyon, France.
- Edwards, L. (1995). Microworlds as representations. In A. A. diSessa, C. Hoyles, & R. Noss (Eds.), *Computers and exploratory learning* (pp. 127–154). New York: Springer.
- Jacobson, M. J., & Wilensky, U. (2006). Complex systems in education: Scientific and educational importance and implications for the learning sciences. *The Journal of the learning sciences*, 15(1), 11-34.
- Kafai, Y. B., & Resnick, M. (2012). Introduction. In *Constructionism in practice* (pp. 13-20). Routledge.
- Kitano, H. (2002). Computational systems biology. *Nature*, 420(6912), 206-210.
- Kitano, H. (2017). Biological Complexity and the Need for Computational Approaches. In *Philosophy of Systems Biology* (pp. 169-180). Springer International Publishing.
- Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas*. Basic Books, Inc..
- Russ, R. S., Scherr, R. E., Hammer, D., & Mikeska, J. (2008). Recognizing mechanistic reasoning in student scientific inquiry: A framework for discourse analysis developed from philosophy of science. *Science Education*, 92(3), 499-525.
- Wilensky, U., & Reisman, K. (2006). Thinking like a wolf, a sheep, or a firefly: Learning biology through constructing and testing computational theories—an embodied modeling approach. *Cognition and instruction*, 24(2), 171-209.
- Wilensky, U., & Resnick, M. (1999). Thinking in levels: A dynamic systems approach to making sense of the world. *Journal of Science Education and technology*, 8(1), 3-19.
- Wilensky, U. (1999). NetLogo. <http://ccl.northwestern.edu/netlogo/>. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.
- Yoon, S. A., Goh, S. E., & Park, M. (2018). Teaching and Learning About Complex Systems in K–12 Science Education: A Review of Empirical Studies 1995–2015. *Review of Educational Research*, 88(2), 285-325.

Appendix

Synthetic Biology – Genetic Switch Model Description:

WHAT IS IT?

This is a multi-agent model of a genetic circuit in a bacterial cell and is an extension of the GenEvo 1 model. This model shows how biologists can use laboratory techniques to tweak certain aspects of a genetic circuit in order to affect the cell's behavior.

Synthetic biology allows biologists to design and test their own genetic circuits. For example, a biologist could design a genetic circuit that caused a bacterium to glow when it was placed in water with a high lead content. This kind of biological engineering is a new frontier being actively explored by scientists around the globe.

HOW IT WORKS

The genetic circuit modelled here has the following components:

1. *promoter with a lac operator* – Transcription starts at the promoter if the repressor protein (LacI) is not bound to the lac operator region. The probability of an RNA polymerase binding to the promoter and starting transcription depends on the promoter strength.
2. *RBS* – A ribosome binding site is downstream to the promoter. The number of proteins produced per transcription depends on the strength of the RBS.
3. *lacZ gene* – A gene that codes for the LacZ protein
4. a terminator – RNA polymerase separates from the DNA when it reaches this region.
5. RNA polymerases - These are represented by brown blobs in the model. This model does not include mRNAs.
6. *LacI repressor proteins* - The purple-colored shapes in the model represent a repressor (LacI proteins). They bind to the operator region (see below) of the DNA and do not let RNAP to pass along the gene, thus stopping protein synthesis. When lactose binds to LacI, they form LacI-lactose complexes (shown by a purple shape with a grey dot attached to it). These complexes cannot bind to the operator region of the DNA.
7. *ONPG molecules* – These are grey pentagons in the model. [ONPG](#) is a chemical that mimics lactose. It is normally colorless. ONPG is hydrolyzed by LacZ enzyme to produce yellow color which is used to check for enzyme activity. Typically, [IPTG](#), another chemical that mimics lactose, is used with ONPG. For simplicity, we have not incorporated IPTG in this model. In this model, ONPG molecules bind to LacI repressor proteins that changes the shape of LacIs preventing them binding to the operator region of DNA.

The model explicitly incorporates transcription by showing the movement of RNA polymerases across DNA. It implicitly incorporates translation and does not incorporate mRNAs or ribosomes.

A user can select the promoter and RBS strengths, add ONPG (by making 'CONST-ONPG' ON) and run the model. The model simulates interactions between the components of the genetic circuit that results in an emergent cellular behavior. The cellular behavior of interest in this model is LacZ (beta-galactosidase) activity which can be observed in a graph and is also represented in the change in the color of the cell to yellow. Beta-galactosidase cleaves ONPG to produce an intensely yellow colored compound.

HOW TO USE IT?

Select the promoter strength and RBS strength using the two choosers.

Press SETUP to initialize the components in the model.

Press GO to run the model.

You can use the RUN EXPERIMENT button to run experiments for a specified time duration (2500 ticks). This is useful for comparing the behavior of the cell in different simulations of the same conditions. You could also use this button to run a timed experiment for different initial conditions (e.g. different promoter and RBS strengths).

'CONST-ONPG?' is a switch which keeps ONPG concentration constant throughout the simulation. This switch can be used to emulate situations where ONPG concentration in the medium is excess and not a limiting factor.

THINGS TO NOTICE

Run the model with 'ONPG?' switch OFF. Notice the molecular interactions inside the cell

- interaction of the LacI protein with the operator
- RNAPs binding to promoter
- RNAPs moving along the DNA
- proteins being generated after an RNAP transcribes the DNA

Observe the same interactions when 'ONPG?' is ON.

Run the model with a set PROMOTER-STRENGTH and RBS-STRENGTH and observe changes in the scaled transcription and translation rates. Also, observe changes in the LacZ activity in the graph as well as in the simulation.

Run it multiple times and observe the differences.

THINGS TO TRY

Change the PROMOTER-STRENGTH and RBS-STRENGTH combination and observe the behavior again.

See which combination has the most robust and optimum behavior.

Change the parameter values of LACI-BOND-LEAKAGE, ONPG-DEGRADATION-CHANCE, COMPLEX-SEPARATION-CHANCE, COMPLEX-FORMATION-CHANCE, and LACZ-DEGRADATION-CHANCE. Notice how these changes affects the behavior of the model.