



Visualization:

Learning to See Mathematically via Image Analysis, Networks, Generative Models
HHMI Workshop on Quantitative Biology 21-24 July 2008 HHMI headquarters, MD

Workshop Leader:

John R. Jungck

BioQUEST Curriculum Consortium

Department of Biology

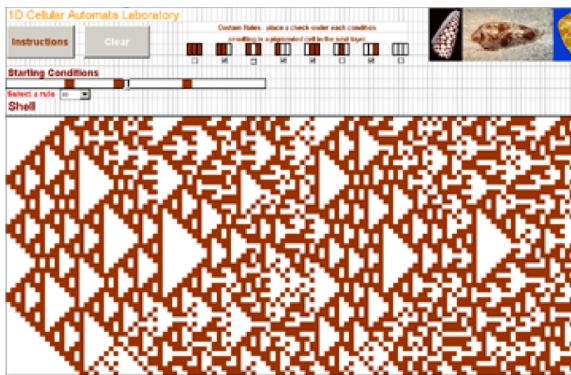
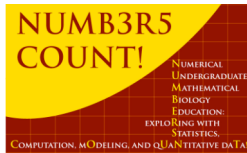
Beloit College

Beloit, WI 53511

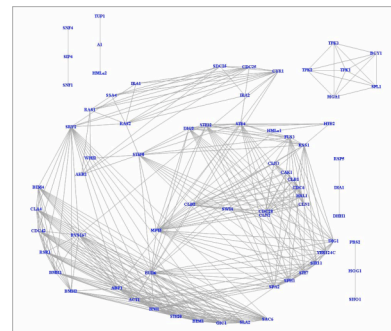
(HHMI Subcontract from Claudia Neuhauser,

HHMI Professor @ University of Minnesota for:

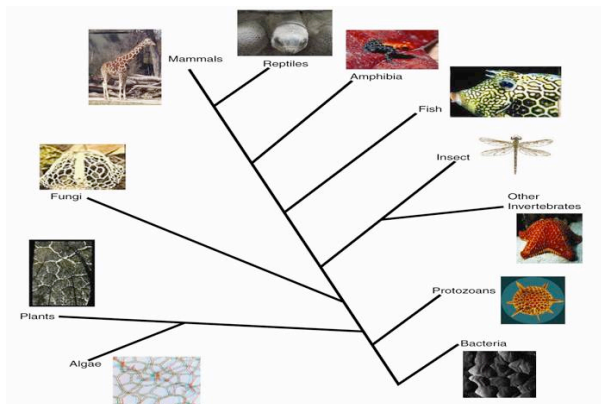
NUMB3R5 COUNT: Numerical Undergraduate Mathematical Biology Education ...)



“the intersection of art,
science, philosophy, and
technology”



“Visualization offers a
method for seeing the
unseen.”



“visual literacy ...
means participation
and makes those
who have achieved
it less passive
observers.”

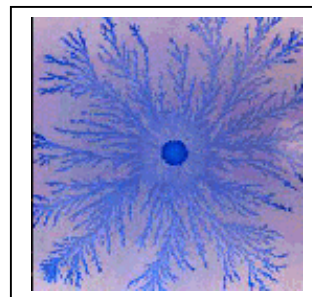
“[V]isual displays are distinctly involved ... in the very ‘construction’ of scientific facts.”

Michael Lynch. (1980).

WORKSHOP OUTLINE

In this workshop, I will focus on nine different kinds of mathematical analysis of visuals in biology:

- (1) Phenomenological
- (2) Topological
- (3) Geometric
- (4) Spatial Statistics
- (5) Networks
- (6) Iterative Generation
- (7) Fractal Measurement
- (8) Bioorthogonal Transforms
- (9) Models and Measurement

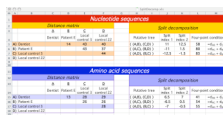


In each case, I argue that visual representations are testable hypotheses, help us reason about biological causation, and help us communicate our inferences. For students to be empowered as scientific investigators, I argue that they need more visual tools than linear regression of an X-Y scatterplot of points or a histogram display of frequencies or compositions.

We use many kinds of visuals that have a mathematical base. I have chosen examples that illustrate some principles that often get neglected and yet represent major themes in biological thinking.



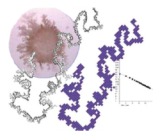
(1) Complexity and Cellular Automata: The one dimensional case (Wolfram’s rules) is used to represent patterns in textile cone seashells and the two dimensional case (Game of Life in three different topological spatial domains) is used to illustrate three philosophical points: (a) determinism is unequal to prediction (i.e., Laplace is wrong) because these systems are completely determined yet unpredictable; the future has to be computed; (b) complex patterns can emerge from simple rules; (c) disproof of structures that could not exist (Garden of Eden theorem, etc.) - uncoupling of complete knowledge of microlevel rules (John Horton Conway’s “the Game of Life”) from emergent macrobehavior (oscillators, gliders, etc.) – analogy of distinction of population genetics of microevolution and paleontological analysis of speciation and macroevolution.



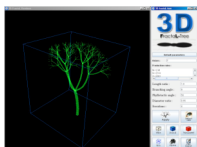
(2) Phylogenetic Trees as Hypotheses: Most publications on phylogeny publish individual trees from the Avogadro’s number of more possible topologies; however, is it informative? How do we test whether two sequences or taxa are evolutionarily distinguishable from two others? We will draw upon two sources: the New Zealand molecular evolution community’s notion of split decomposition (*EvolSeq* and *Split Decomposition* in the Biological ESTEEM project) and the work of participant Sam Donovan (Baum, Smith, & Donovan, 2005; his “Tree Thinking Group website: (<<http://www.lrdc.pitt.edu/donovan/index.html>>)).



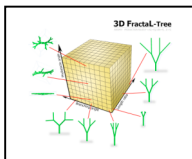
(3) Image Analysis: Every image is full of quantitative data. Student's digital camera's (both stand alones and in their cell phones) are powerful scientific data acquisition instruments with tremendous technological power; yet little of that power is usually employed in classrooms, labs, and field work. Add digital video microscopy and freeware such as *ImageJ* (<http://rsbweb.nih.gov/ij/>) or commercial software for kinematic analysis such as *Videopoint* (<http://www.lsw.com/videopoint/vp/index.html>) and one is empowered for an enormous variety of scientific investigation. We will use two modules from our *Microbes Count!* (Jungck, Fass, and Stanley, 2003; copies are provided to you for this workshop) called "Modeling Growth" and "Valuing Variegated Variation" to illustrate natural growth environments of microbes and viruses offer an alternative to the laboratory culture of domesticated variation.



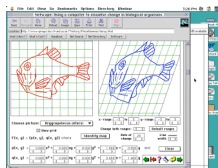
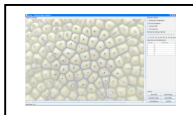
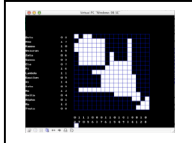
(4) Fractal Measurement: To pick up on a variant of "the laboratory culture of domesticated variation," we will use another module in our *Microbes Count!* (Jungck, Fass, and Stanley, 2003) called "Shaped to Survive" to illustrate the power of using *Fractal Dimension* (Stanley *et al.*, 1989; <http://polymer.bu.edu/oqaf/>; "Patterns in Nature") to measure the "irregularity" or "roughness" of diffusion limited growth colonies of bacteria on hard substrates with minimal food.



(5) Fractal Construction: Following in this fractal genre' we will use a variant called graph grammar rewriting systems or Lindenmayer Systems (L-systems) (*Algorithmic Beauty of Plants*, Prusinkiewicz and Lindenmayer, 1996; available for free download: (<http://algorithmicbotany.org/papers/#abop>)) to engage students in understanding the power of a few genes and a few rules to generate self-similar realistic structures. By the simple use of a ruler, protractor, calipers, a spreadsheet in the Biological ESTEEM project, and our software "*3D Fractal Tree*" (Khripet *et al.* 2005), students can make measurements from a branch of a tree and construct a rotating 3D tree image to compare whether these few measurements capture much of the total appearance of this phytoarchitecture. A similar approach will be quickly demonstrated for the construction of seashells. Similar approaches have also been employed for corals and sponges.



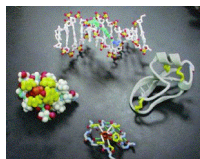
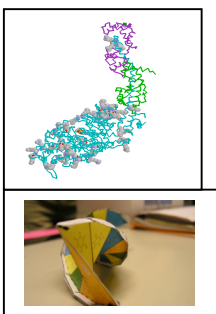
(6) Network Analysis: Most of the important relationships in biology: parent-child, substrate-product, predator-prey, repressor-operator, ancestor-descendent, and fertilized egg-differentiated cell, axon-dendrite synapses, etc. are illustrated as graphs: pedigrees, metabolic pathways, food webs, genetic control circuits and interactomes, phylogenetic trees, fate maps, neuronal networks, etc. Yet few biologists are familiar that graph theory



provides formal mathematical tools for the analysis of such measures as connectedness and diameter nor the presence of subgraphs with informative properties (maximal cliques, Eulerian circuits, Hamiltonian paths, minimal spanning trees, etc.) nor overall characteristics such as small world, random, or scale free. We have developed a spreadsheet front end to import or enter biological data for a graph visualization package and a set of measurement tools (*BioGrapher*, Viswanathan *et al.* 2005) to provide a more general environment for students to realize that problems from molecular biology to genetics to developmental biology to neurobiology to ecology to evolution all employ a common mathematical perspective. *BioGrapher* (Viswanathan 2005) is an Excel front-end for the *AT&T GraphViz* graphical visualization package (<http://www.graphviz.org/>). *BioGrapher* is unique in that it (1) allows users to enter data in one of three ways: adjacency matrices, nodal lists, and Newick format for phylogenetic trees; (2) displays graphs in four different ways: radially, circularly, hierarchically, and in tree format; (3) properties of graphs such as diameter, connectedness, and average clustering can be computed and displayed histographically and as scatter plots; (4) data can be read from standard csv formatted databases.

(7) Spatial Statistics and Graph Theoretical Analysis of Images: Anytime that irregular polygonal tessellations exist in biology or questions arise about data represented by a set of points distributed in space, hypotheses about aggregation, avoidance, relationships, and clustering can be addressed. We have developed *Ka-me'*: *Voronoi Image Analyzer* (Khiripet, Kwanthan, and Jungck, 2007) to combine biology, graph theory, data structures, and spatial statistics into one highly visual, easy to use computer interface. This software could be used at the cellular level (cancer pathology), organism level (parastichies), and ecological level (tree canopies, fish territories, forest fire outbreaks, epidemiological spread) as well as in data mining and classification.

(8) Bioorthogonal Transformations: Warping, Morphing, Morphometrics, Landmarks: D'Arcy Thompson (1917) illustrated that ontogeny and phylogeny (despite his own anti-Darwinism) could be understood by continuous transformations between one structure and another by stretching one gridded graph paper image (domain) onto another gridded graph paper image (range) in a one-to-one mapping of X-Y points in one image to the other. We'll illustrate a preliminary package (Koch, Panks, Arganbright, Chaavey, and Jungck) for this approach using matrix algebra and online freeware for doing the transforms without the mathematics.



(9) Molecular Visualization: I could not bring along a full three dimensional projection system; however, I refer you to the GeoWall Consortium (<<http://geowall.geo.lsa.umich.edu/home.html>>) and our GeoWall Stereovisualization Software Guide site (<<http://epic.bu.edu/geowall>>) to learn how you can build a portable 3D projection system for classroom use for under \$5,000. Unfortunately, while Molecular Visualization has been an extremely popular use of computer graphics, it has not been often used by students as a tool for doing science. Too often students are entertained by zooming, rotating, and translating, but they often have no sense of why 3D visualization of proteins is so important to testing hypotheses in drug design or immunological receptor interaction. We have found that a small peptide, TrpCage, is a very helpful starter for engaging students in structural bioinformatics (BioQUEST's BEDROCK Problem Spaces (Bioinformatics Education Dissemination: Reaching Out, Connecting, and Knitting-together) – see Tim Herman, Rama Viswanathan, Tia Johnson. Graham Walker, HHMI Professor at MIT along with BioQUESTer Ivica Ceraj and colleagues have developed a new viewer: StarBiochem which allows beginning students to explore three dimensional visualizations of proteins with a user-friendly interface and to make quantitative measurements. Also, Paul R. McCreary, a mathematician from Evergreen State College (<<http://academic.evergreen.edu/m/mccreary/>>), has developed a plug-in for viewing proteins in hyperbolic space (flying through molecules) on a microcomputer and understanding holes in space better as well as a different viewing perspective.

All of the software that will be demonstrated is freely available through the BioQUEST Curriculum Consortium (<http://www.bioquest.org>) or a number of other sites with whom we have wonderful success in using their tools and that have been maintained for a long time in the public domain. The notions of Creative Commons licenses, copyleft, open access, open source, and open science will be discussed in this context.

"What are the advantages for nonartists in developing their visual acuity and expressive potential? The first and crucial value lies in development of criteria that extend beyond natural response and personal or conditioned tastes and preferences. Only those who are visually sophisticated can rise above fashion and fad to make their own choices and judgment of what is appropriate and aesthetically pleasing. At a slightly higher level of involvement, in effect, visual literacy precludes the "Emperor's clothes" syndrome and makes of judgment a higher action than acceptance (or rejection) of a visual statement based on intuition alone. Visual literacy means increased visual intelligence."

Donis A. Donis, (1973).

"Barbara Stafford is at the forefront of a growing movement that calls for the humanities to confront the brain's material realities. In *Echo Objects* she argues that humanists should seize upon the exciting neuroscientific discoveries that are illuminating the underpinnings of cultural objects. In turn, she contends, brain scientists could enrich their investigations of mental activity by incorporating phenomenological considerations — particularly the intricate ways that images focus intentional behavior and allow us to feel thought.

This, then, is a book for both sides of the aisle, a stunningly broad exploration of how complex images — or patterns that compress space and time — make visible the invisible ordering of human consciousness. Stafford demonstrates, for example, how the compound formats of emblems, symbols, collage, and electronic media reveal the brain's grappling to construct mental objects that are redoubled by prior associations. On the other hand, she compellingly shows that findings in evolutionary biology and the neurosciences are providing profound opportunities for understanding aesthetic conundrums as old and deep-seated as the human urge to imitate, the mapping of inner space, and the role of narrative and nonnarrative representation.

As precise in her discussions of firing neurons as she is about the coordinating dynamics of image making, Stafford locates these major transdisciplinary issues at the intersection of art, science, philosophy, and technology. Ultimately, she makes an impassioned plea for a common purpose — for the acknowledgement that, at the most basic level, these separate projects belong to a single investigation."

Dust jacket, *Echo Objects* (2007)

Visualization is a method of computing. It transforms the symbolic into the geometric, enabling researchers to observe their simulations and computations. Visualization offers a method for seeing the unseen. It enriches the process of scientific discovery and fosters profound and unexpected insights. In many fields it is already revolutionizing the way scientists do science."

B. McCormick, T. DeFanti, and M. Brown. (1987).

"Images of natural phenomena hold a particular fascination for us, if they show objects and events outside our normal range of perceptions or experience. ... But we all share a special fascination for the mysterious and extraordinary phenomena that our senses are not normally capable of showing us. ... However, the simple notion that an image of something will tell us all, or nearly all that we need to know about it is far from true, since the fundamental nature of an image is that it is dependent on the object being observed, the wavelength of the radiation being used to make the observation, and the technology being used to collect and display the observed data.

Robert S. Wolff and Larry Yaeger. (1993).

References:

- Bajaj, Chandrajit (1998). "Visualization Paradigms," In C. Bajaj, editor, *Data Visualization Techniques*. John Wiley and Sons: New York.
- Baum, D. A., Smith, S. D., & Donovan, S. S. (2005). The tree-thinking challenge. *Science*, 310(5750), pp. 979-980 plus see 29 page supplement.
- Ben-Jacob, E., & Levine, H. (1998, October). The artistry of microorganisms: Colonies of bacteria or amoebas form complex patterns that blur the boundary between life and nonlife. *Scientific American* 279: 82-87.
- Canning, David R., and James R. Cox. (2001). Teaching the Structural Nature of Biological Molecules: Molecular Visualization in the Classroom and in the Hands of Students. *Chemistry Education: Research and Practice in Europe* 2 (2): 109-122.
- Cox, James R. (2000). Teaching Noncovalent Interactions in the Biochemistry Curriculum through Molecular Visualization: The Search for Pi Interactions. *Journal of Chemical Education* 77: 1424-1428 (November).
- Domik, Gitta O. (1993). An Agenda for Education in Scientific Visualization. Visualization '92 Workshop Report. *Computer Graphics* 27 (1): 6 (January). Report of the subcommittee "Education for Visualization", or EVC, of the ACM SIGGRAPH Education Committee.
- Donis, Donis A. (1973). *A Primer of Visual Literacy*. MIT Press: Cambridge, Massachusetts, p.185.
- Fujikawa, H. and Matsushita, M. (1991). Bacterial fractal growth in the concentration field of nutrient. *J Phys Soc Japan* 60:88-94.
- Goldberger AL, West BJ. (1987). Fractals in physiology and medicine. *Yale J. Biol Med* 60 (5): 421-35.
- Greenberg, Richard. (1996). "The Image Processing for Teaching Project." In Charles Fisher, David C. Dwyer, and Keith Yocam, Editors, *Education and Technology: Reflections on Computing in Classrooms*, Jossey-Bass Publishers (Apple Press): San Francisco, CA, pp. 221-234.
- Guyon, E., & Stanley, H. E. (Eds.). (1991). *Fractal forms*. Elsevier: Haarlem, Holland. (Text that goes along with the "Patterns in Nature" materials from Boston University.)
- Herman, Tim, Jennifer Morris, Shannon Colton, Ann Batiza, Michael Patrick, Margaret Franzen, and David S. Goodsell. (2006). Tactile teaching: Exploring protein structure/function using physical models. *Biochemistry and Molecular Biology Education* 34 (4): 247-254 (July).
- Jones, C. L.; Lonergan, G. T.; and Mainwaring, D. E. (1993). A rapid method for the fractal analysis of fungal colony growth using image processing. *Binary* 5: 171-180.
- Jungck, John R., Marion Field Fass, and Ethel Stanley. (2003). *Microbes Count!: Problem Posing, Problem Solving, and Persuading Peers in Microbiology*. ASM Press: Washington, D.C.
- Kaandorp, J. A. (1994). *Fractal modeling: Growth and form in biology*. Springer-Verlag: New York.

- Lynch, Michael. (1980). "The externalized retina: Selection and mathematization in visual documentation of objects in the life sciences." In Michael Lynch and Steve Woolgar, editors, *Representation in Scientific Practice*, MIT Press: Cambridge, Mass., pp. 153-186.
- Mandelbrot, B. B., 1982. *The Fractal Geometry of Nature*. Freeman: New York.
- McCormick, B., T. DeFanti, and M. Brown. (1987). "Visualization in Scientific Computing." *Computer Graphics (AC-SIGGRAPH)* 21 (6): 1-15.
- McNeal, Ann P. and Charlene D'Avanzo, eds. (1997). *Student-Active Science: Models of Innovation in College Science Teaching*. Saunders College: Fort Worth, Texas.
- Niklas, Karl J. 1997. Adaptive Walks Through Fitness Landscapes For Early Vascular Land Plants. *American Journal of Botany* 84 (1): 16-25.
- Niklas, Karl J. 2000. Modeling fossil plant form-function relationships: a critique. *Paleobiology* 26: 289-304.
- O'Hara, Robert J. (1992). Telling the tree: narrative representation and the study of evolutionary history. *Biology and Philosophy* 7: 135-160.
- O'Hara, Robert J. (1997). Population thinking and tree thinking in systematics. *Zoologica Scripta* 26: 323-329.
- Poundstone, William. (1984). *The Recursive Universe: Cosmic Complexity and the Limits of Scientific Knowledge*. William Morrow: New York.
- Prusinkiewicz, Przesmyslaw, and Aristid Lindenmayer. (1996). *Algorithmic Beauty of Plants*. Springer-Verlag: New York.
- Rucker, Rudy. (2005). *The Lifebox, the Seashell, and the Soul: What Gnarly Computation Taught Me About Ultimate Reality, the Meaning of Life, and How to Be Happy*. Basic Books (Thunder's Mouth Press): New York.
- Stafford, Barbara Maria. (1994). *Artful Science: Enlightenment Entertainment and the Eclipse of Visual Education*. MIT Press: Cambridge, Massachusetts.
- Stafford, Barbara Maria. (2007). *Echo Objects: The Cognitive Work of Images*, and (1994). *Artful Science: Enlightenment Entertainment and the Eclipse of Visual Education*. MIT Press: Cambridge, Mass.
- Thompson, D'Arcy Wentworth. (1917). *On Growth and Form*. Cambridge University Press: Cambridge, U.K.
- Viswanathan, Rama. (20056). *BioGrapher and WDisplay: Excel Modules for Data Visualization*. *BioQUEST Notes* 15 (1): 6-7.
- Walker, Graham, and Melissa Kosinski-Collins. (visited 16 July 2008). The Use of 3D in the Biology Undergraduate Laboratory Course. Website: (<<http://www.cfkeep.org/html/snapshot.php?id=68340453>>).
- Wolff, Robert S., and Larry Yaeger. (1993). *Visualization of Natural Phenomena*. Telos: The Electronic Library of Science (Springer-Verlag): New York, p. 1.
- Wolfram, Stephen. (2002). *A New Kind of Science*. Wolfram Media: Champaign, Illinois.

BioQUEST Curriculum Consortium Home Page



BioQUEST
Curriculum Consortium

[Sign In](#)

New!! Search the entire BioQUEST Site:

[About BioQUEST](#)
[Calendars](#)
[Projects and Collaborations](#)
[Summer Faculty Workshops](#)
[Publications](#)
[Contact Us](#)

FEATURED PROJECTS





The BioQUEST Curriculum Consortium actively supports educators interested in the reform of undergraduate biology and engages in the collaborative development of curricula.

We encourage the use of simulations, databases, and tools to construct learning environments where students are able to engage in activities like those of practicing scientists.

Workshop in Progress: myPlant IT

BioQUEST is a founding member of





The revised edition of **Microbes Count!** features 40 lab activities and includes online access to data and software tools.