Atlantic Symposium Comp Biol Genomic Info Systems Technol September. 950-953

Using Virtual Reality to Understand Complex Metabolic Networks

J.A. Dickerson*, Y. Yang*, K. Blom*, A. Reinot*, J. Lie[#], C. Cruz-Neira*, and E. S. Wurtele[#]
Electrical and Computer Engineering, [#]Genetics, Development and Cell Biology, Iowa State University, Ames, IA,

Abstract

Metabolic networks combine metabolism and regulation. These complex networks are difficult to understand and visualize due to the diverse types of information that need to be represented. Three-dimensional graph visualization coupled with visualization of the physical structure of a cell help create a novel integrated information workspace for the study of metabolic networks. This paper shows how biologists can interact with a virtual reality representation to get a more complete view of network behavior.

Introduction

Complex interactive metabolic pathways contain many different types of information, which presents a challenge to computationally model and visualize the interactions. Most current methods for visualizing the pathways focus on two-dimensional (2D) graph models to represent pathways [1-4]. These 2D graph-based models of metabolic networks are overloaded since edges and nodes can have multiple meanings [5-8]. The edges and nodes of the graph represent a variety of different concepts such as chemical reactions, rates, cell compartment identification, lab test results, etc. Three-dimensional graph visualization coupled with visualization of the physical structure of a cell help create a novel integrated information workspace for the study of metabolic networks.

Metabolic pathways create considerable amounts of information, which is almost impossible to investigate and understand through traditional biology research and analysis methods. We are combining advanced biological knowledge with automated mathematical logic, complex data structures, fuzzy cognitive maps, interactive graph visualization, and other computational tools to create a novel analytical suite of tools for the biologists.

Virtual reality (VR) and immersive environments are a relatively new research tools. VR strives to present the user with a convincing, interactive three-dimensional (3-D) environment. The user views the 3-D environment stereoscopically, typically with the aid of specialized glasses. The user's position is tracked by a computer so the virtual environment can respond to the user's movements. Projection-based virtual reality systems, such as Iowa State University's enclosed cube structure called the C6, have stereo images projected onto the display surfaces.

One aspect of this research is the creation of a virtual metabolic network environment. This task requires visualizing and navigating complex graphs in a scalable immersive environment, and integrating physical models and graphical representations of cell metabolism. The goal is to create a seamless system to enable biologists to gain insight on cell metabolism and to provide an educational tool to communicate their findings through virtual reality experiences.

3-D Network Visualization

Metabolic networks contain multiple pathways. A pathway is a collection of interconnected biochemical reactions. A biochemical reaction includes substrates, enzymes, and products. Each reaction can be modeled as a directed graph, G (V, E), where V = {substrate nodes, enzyme nodes, product nodes}, E = {edges connecting the nodes}. Using this modeling method recursively upon the entire set of reactions, we get one directed graph to represent the metabolic networks. So the problem of visualizing metabolic networks converts to the problem of visualizing a directed graph.

Although most of the existing methods to draw metabolic pathways are in 2D space, there are a few 3D graph drawing algorithms, such as force-directed drawing and orthogonal drawing [9, 10]. However, these algorithms produce graphs which are difficult to interpret in three dimensions due to edge-crossings and graph complexity. Figure 1 shows a three-dimensional force-directed layout called GEM (www.tulip.org) used for the Arabidopsis metabolic network.

Metabolic networks often consist of tens of thousands of nodes, making any global network layout very complex and difficult to interpret. Instead of visualizing the total network at one time, we visualize a portion of the network at one time. The biologist can choose any node as a focus node for a reaction of interest (ROI).

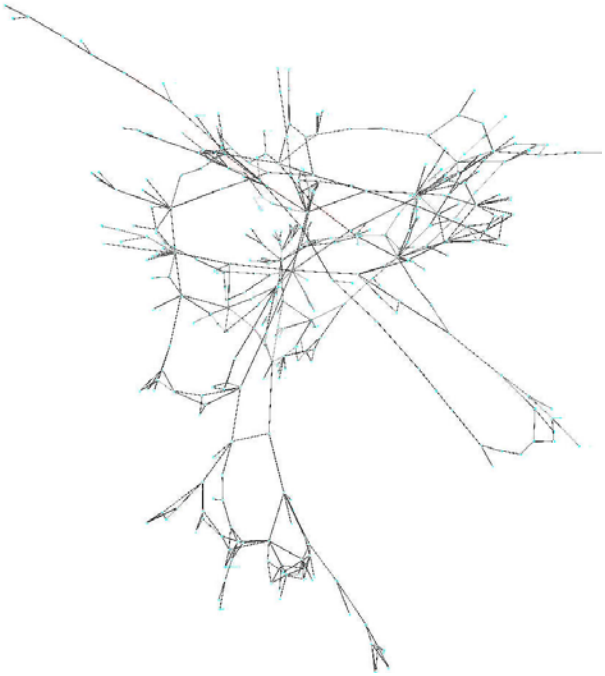

Figure 1: Partial graph of metabolic pathways in Arabidopsis. The graph layout uses a 3-D force-directed layout which puts highly connected nodes in the center of the graph.

An ROI is defined as all the reactions that the focus node participates in. We also have incorporated a crossing-free 3D layout algorithm to visualize the ROI. We put the reactions evenly in the 3D region around the focus node. Under this schema, an edge crossing will happen only when a node (other than the focus node) participates in more than one reaction. We call these nodes intersection nodes. The crossing-free feature is achieved by splitting the intersection node so that each reaction involving the intersection node has its own copy of the node. Figure 2 shows an example of a selected ROI. The database we are currently using for these graphs is a combination of MetNetDB [11] and ARACYC [12].

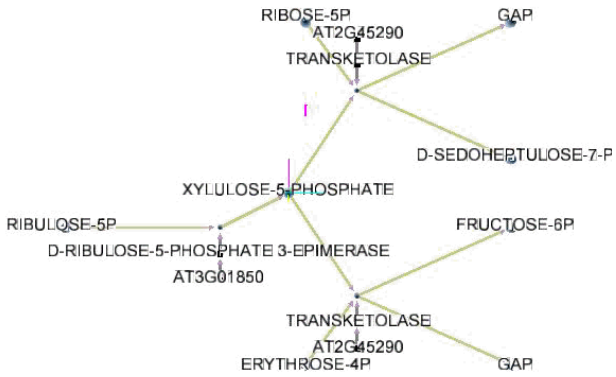

Figure 2: A reaction of interest selected from the Calvin Cycle. The focus node is xylulose-5-phosphate.

Navigation in the Network

Three-dimensional network visualization techniques are one method of representing the quantity and complexity of the metabolic states and pathways. In order to take advantage of this representation, the user must be able to navigate and interact with the network in a meaningful manner. This project introduces a tablet PC into the virtual environment.

Textual information concerning the source, synonyms and other data is important in the study of metabolic pathways. The tablet PC provides a method to present a traditional 2-D desktop interface to the user. Methods for effectively displaying textual information are well known in the desktop environment using a Graphical User Interface (GUI). We have built our tool using Java for the GUI and a tool included in the VR software, Tweek (www.vrjuggler.org). This allows us to run a java GUI on a wireless tablet PC; the PC communicates with the main application to retrieve and display the textual information and steers the VR Environment, as shown in Figure 3.

Figure 3: Interaction with the three-dimensional metabolic network using a tablet PC to select nodes.

Navigation throughout the network becomes challenging as the number of nodes and edges of the network increases. Each of the nodes in our 3-D representation has a text tag giving its name. This is useful for nodes close to the user. The tablet PC's GUI can provide an easily accessible method for the user to find and travel to a node of interest. In this GUI, the user can see tables of the nodes and edges present in the displayed graph. The user can interactively select nodes and edges and see the complete information on

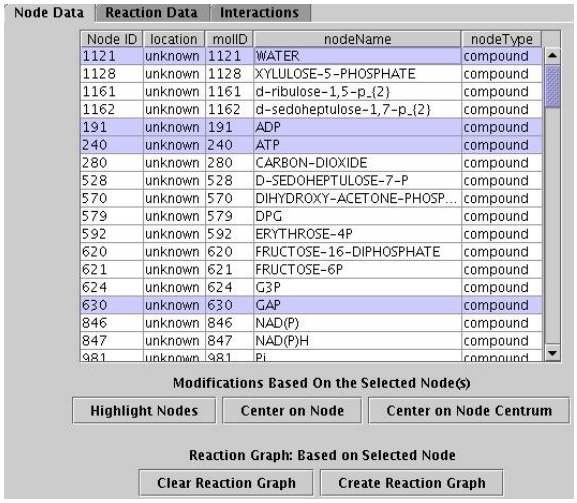

Figure 4: Java GUI for displaying node information for navigation in the graph.

the node or edge on the GUI. The user can also select to have the nodes and/or edges colored to highlight them in the scene or can select a node or edge or have the VR environment moved to show the selected object. An example interaction is "pulling" a reaction-of-interest out from a larger graph as shown in Figure 4. In the future we hope to expand the methods of viewing the network and add different navigational metaphors.

Virtual Cell Representations

The plant cell contains over 20 distinct types of subcellular compartments, called organelles. Each organelle is highly specialized for specific metabolic reactions, keeping groups of metabolites and enzymes in distinct compartments. Metabolic efficiency is increased because metabolites and enzymes of a given pathway can be more highly concentrated. Moreover, unwanted side reactions due to related enzymes of different pathways are minimized.

As a teaching environment for high school and college students, we are developing a virtual cell, and integrating this cell with cellular metabolism and regulation. The student will be able to visualize the cell from the outside, as well as cross-sections of the entire cell. As the student zooms inward, she/he enters the cell and the organelle systems within. From this organelle world, the student can track metabolic pathways, following anabolism and catabolism within the cell. This would encompass visualizing reactions within the organelle, and as a given metabolite leaves the cell, the student would virtually move from organelle to organelle.

We are initially focusing on three metabolic processes distributed across five subcellular compartments: 1) the Calvin cycle, in which atmospheric $C0_2$ is fixed into sugar using chemical energy derived from the sun; 2) the TCA cycle, in which $C0_2$ from metabolites is released back into the atmosphere, metabolically useful chemical energy is generated, and many chemical building blocks are formed 3) the acetyl-CoA network, leading to the synthesis of membrane lipids, oils, waxes, pigments, and many specialized bioactive plant metabolites.

Conclusions

The 3D virtual environment offers exciting new opportunities for visualizing complex networks. The ability to link detailed physical models with representations of regulatory and metabolic flow will lead to new teaching methods in biology.

Acknowledgements

Funding for this project is provided by grants from the National Science Foundation in the Arabidopsis 2010 (DBI-0209809) and Information Technology Research (IBN-0219366) Programs. Seed funding was also provided by the Iowa State University Plant Sciences Institute and the Roy J. Carver Foundation.

References

[1] M. Becker, and I. Rojas, "A graph layout algorithm for drawing metabolic pathways," *Bioinformatics*, vol. 17, pp. 461-467, 2001.

[2] P. D. Karp, M. Krummenacker, S. Paley, and J. Wagg, "Integrated pathway/genome databases and their role in drug discovery," *Trends in Biotechnology*, vol. 17, pp. 275-281, 1999.

[3] P. D. Karp, "Pathway databases: a case study in computational symbolic theories," *Science*, vol. 293, pp. 2040-4, 2001.

[4] A. J. Hartemink, D. K. Gifford, T. S. Jaakkola, and R. A. Young, "Using Graphical Models and Genomic Expression Data to Statistically Validate Models of Genetic Regulatory Networks," Proceedings of the Pacific Symposium on Biocomputing, Hawaii, 2001.

[5] J. A. Dickerson, D. Berleant, Z. Cox, D. Ashlock, A. W. Fulmer, and E. S. Wurtele, "Creating and Modeling Metabolic and Regulatory Networks Using Text Mining and Fuzzy Expert Systems," in *Computational Biology and Genome Informatics*, C. H. Wu, P. Wang, and J. T. L. Wang, Eds. Hong Kong: World Scientific, 2002.

[6] J. A. Dickerson, D. Berleant, Z. Cox, W. Qi, and E. Wurtele, "Creating Metabolic Network Models using Text Mining and Expert Knowledge," Proceedings of the Atlantic Symposium on Molecular Biology and Genome Information Systems and Technology (CBGIST 2001), Durham, North Carolina, 2001.

[7] J. A. Dickerson, Z. Cox, E. S. Wurtele, and A. W. Fulmer, "Creating Metabolic and Regulatory

Network Models using Fuzzy Cognitive Maps," Proceedings of the North American Fuzzy Information Processing Conference (NAFIPS), Vancouver, B.C., 2001.

[8] J. A. Dickerson, D. Berleant, Z. Cox, D. Ashlock, A. W. Fulmer, and E. S. Wurtele, "Creating and Modeling Metabolic and Regulatory Networks Using Text Mining and Fuzzy Expert Systems," in *Computational Biology and Genome Informatics*, C. H. Wu, P. Wang, and J. T. L. Wang, Eds. Hong Kong: World Scientific, 2002.

[9] M. Closson, S. Gartshore, J. Johansen, and S. K. Wismath, "Fully Dynamic 3-Dimensional Orthogonal Graph Drawing," in *7th International Symposium, GD'99, Stirín Castle, Czech Republic*, vol. 1731, *Lecture Notes in Computer Science*, J. Kratochvíl, Ed. Berlin: Springer Verlag, 1999, pp. 49-58.

[10] B. Landgraf, "3D Graph Drawing," in *Drawing Graphs: Methods and Models*, vol. 2025, *Lecture Notes in Computer Science*, D. W. M. Kaufmann, Ed. Berlin: Springer Verlag, 2001, pp. 172-192.

[11] E. S. Wurtele, J. Lie, L. Diao, H. Zhang, C. M. Foster, B. Fatland, J. A. Dickerson, A. Brown, Z. Cox, D. Cook, E.-K. Lee, and H. Hofmann, "MetNet: software to build and model the biogenetic lattice of Arabidopsis," *Comparative and Functional Genomics*, vol. 4, pp. 239-245, 2003.

[12] S. Y. Rhee, W. Beavis, T. Z. Berardini, G. Chen, D. Dixon, A. Doyle, M. Garcia-Hernandez, E. Huala, G. Lander, M. Montoya, N. Miller, L. A. Mueller, S. Mundodi, L. Reiser, J. Tacklind, D. C. Weems, Y. Wu, I. Xu, D. Yoo, J. Yoon, and P. Zhang, "The Arabidopsis Information Resource (TAIR): a model organism database providing a centralized, curated gateway to Arabidopsis biology, research materials and community," *Nucleic Acids Res.*, vol. 31, pp. 224-228, 2003.

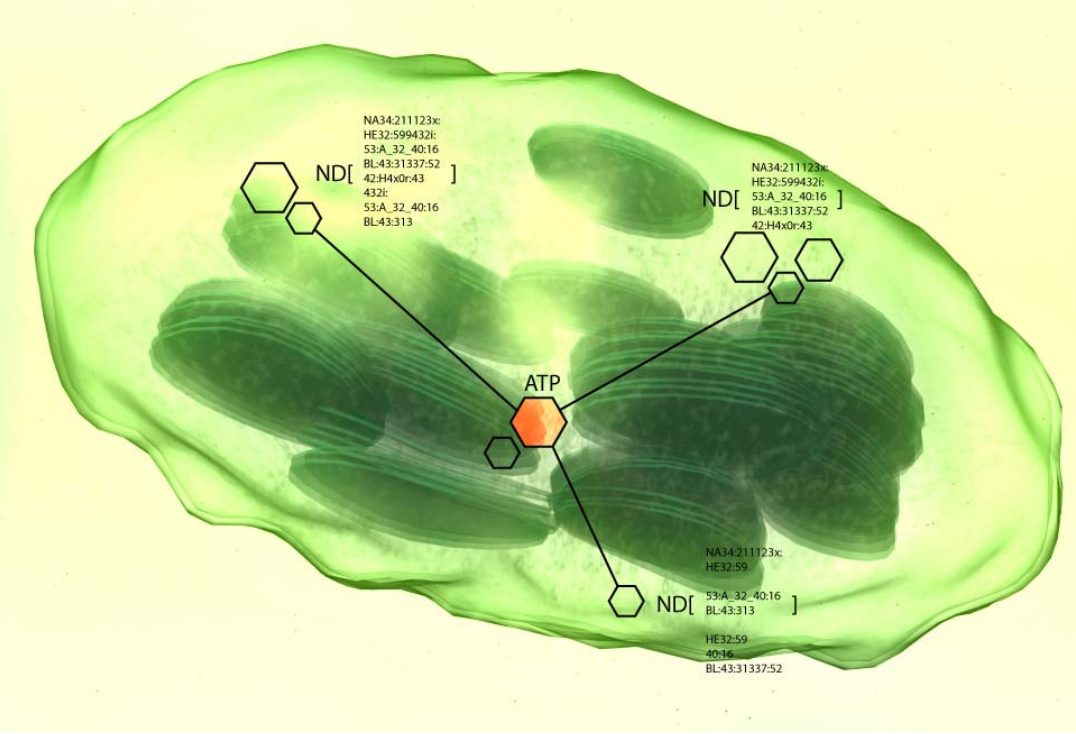

Figure 5: Three-dimensional model of the chloroplast overlayed with a network showing one of the metabolic processes in that organelle.