

08191 Abstracts Collection
Graph Drawing with Applications to
Bioinformatics and Social Sciences
— Dagstuhl Seminar —

Steve Borgatti¹, Stephen Kobourov², Oliver Kohlbacher³ and Petra Mutzel⁴

¹ University of Kentucky, USA
sborgatti@uky.edu

² University of Arizona, USA
kobourov@cs.arizona.edu

³ Universität Tübingen, D
kohlbach@informatik.uni-tuebingen.de

⁴ Universität Dortmund, D
petra.mutzel@tu-dortmund.de Country

Abstract. From May 4 to May 9, 2008, the Dagstuhl Seminar 08191 “Graph Drawing with Applications to Bioinformatics and Social Sciences” was held in the International Conference and Research Center (IBFI), Schloss Dagstuhl. During the seminar, several participants presented their current research, and ongoing work and open problems were discussed. Abstracts of the presentations given during the seminar as well as abstracts of seminar results and ideas are put together in this paper. The first section describes the seminar topics and goals in general. Links to extended abstracts or full papers are provided, if available.

Keywords. Graph drawing, visualization, social sciences, bioinformatics

08191 Executive Summary – Graph Drawing with Applications to Bioinformatics and Social Sciences

Graph drawing deals with the problem of communicating the structure of relational data through diagrams, or drawings. The ability to represent relational information in a graphical form is a powerful tool which allows to perform analysis through visual exploration to find important patterns, trends, and correlations. Real-world applications such as bioinformatics and sociology pose challenges to the relational visualization because, e.g., semantic information carried by the diagram has to be used for obtaining meaningful layouts and application-specific drawing conventions need to be fulfilled. Moreover, the underlying data often stems from huge data bases, but only a small fraction shall be displayed at a time; the user interactively selects the data to be displayed and explores the graph by expanding interesting and collapsing irrelevant parts. This requires

powerful graph exploration tools with navigation capabilities that allow dynamic adaption of the graph layout in real time. In this seminar we focused on the application of graph drawing in two important application domains: bioinformatics and social sciences.

We brought together theoreticians and practitioners from these areas and focused on problems concerning interaction with and navigation in large and dynamic networks arising in these application areas.

During the seminar, we identified and defined open graph drawing problems that are motivated by practical applications in the targeted application areas, tackled selected open problems, formulated the findings as a first step to the solution, and defined further research directions.

Keywords: Graph drawing, visualization, social sciences, bioinformatics

Joint work of: Borgatti, Steve; Kobourov, Stephen; Kohlbacher, Oliver; Mutzel, Petra

Extended Abstract: <http://drops.dagstuhl.de/opus/volltexte/2008/1552>

Overview on Biological Network Visualizations

Mario Albrecht (MPI für Informatik - Saarbrücken, DE)

This talk gives an overview about the application of various visualization methods to biological networks.

It also points to challenges for the use of visual analytics tools in network bioinformatics.

Keywords: Network, Visualization

Graph Drawing in the Social Sciences

Stephen P. Borgatti (University of Kentucky - Lexington, US)

A review of visualization practices in three social science areas that focus on relational data: multivariate/correlational analysis, cultural domain analysis, and social network analysis. It is suggested that an important area for development of algorithms is in visualizing trajectories through graphs, such as arise in the study of careers or the diffusion of innovation.

Keywords: Social science, social network analysis, cultural domain analysis, trajectories

Social Networks from a Graph Drawing Perspective

Ulrik Brandes (Universität Konstanz, DE)

Most notably in the work of J.L. Moreno (1934), visual representations have long been a means of both presentation and exploration of social structures. To create effective network visualizations, three main aspects have to be considered: the substance (data, interest, context, derived data), a design (mapping substance to graphical elements and variables), and algorithms (that are to realize a design without introducing artifacts obscuring the information). For selected substantive categories, I present historical and current examples of visualizations and point out open algorithmic issues associated with them. I conclude with more broadly defined challenges that are based on recurrent problems in social network analysis.

Keywords: Graph drawing, social network analysis

Network Biology - An Overview

Oliver Kohlbacher (Universität Tübingen, DE)

The introduction of networks as a model of integrating, analyzing, and interpreting data has led to a paradigm shift in biology over the last few years.

We present different types of biological networks as well as some use cases and identify some areas where graph drawing can impact bioinformatics and biology.

Keywords: Networks, bioinformatics, graph drawing

Graph Drawing for Biological Networks

Falk Schreiber (IPK Gatersleben, DE)

Biological processes can be represented by different networks ranging from molecular networks (gene regulation, protein interaction, metabolism) to ecological and evolutionary networks. These networks are strongly linked to each other and embedded in space (e.g. cellular compartments, organs) and time (e.g. development, evolution). To help investigating and understanding biological processes intuitive network visualisation and navigation methods are required.

This talk presents an overview of types of biological networks, graph models to represent these networks, typical layout algorithms used to visualise such networks, and some tools used in systems biology. We will discuss ways to collaborate with users from the life sciences, to promote graph drawing methods in this application area, and possible research directions.

Keywords: Graph drawing, biological networks

Some comments ...

Vladimir Batagelj (University of Ljubljana, SI)

We presented a collection of ideas on interactive visualization of large networks in the standard visualization framework (zooming, glasses and lenses, ...).

We propose to develop a layout format for storing the visualization (graph drawing) results and use it as the input for network viewers.

Keywords: Large networks, visualization, layout, navigation, glasses, lenses, zoom

Radial Level Graphs and Recurrent Hierarchies

Franz J. Brandenburg (Universität Passau, DE)

The Sugiyama method a well-known concept for hierarchical graph drawings. It places the vertices on horizontal lines and attempts to avoid edge crossings and bends of edges. The method consists of four phases (1) decycling, (2) layering (3) crossing reduction, and (4) coordinate assignment or routing, which each encompass an NP-hard problem. For each phase there is a wide collection of algorithms.

We modify the Sugiyama method and place the vertices on concentric rings, or as recurrent hierarchies. Concentric rings are well-suited for the representation of centralities in social networks. Recurrent hierarchies were introduced by Sugiyama et al. in 1981, but have not been studied since then. Recurrent hierarchies can be drawn in 3D as a level graph on a cylinder, or in 2D with rays. They are well-suited to draw many cycles.

We investigate these modifications, and present algorithms for the four phases of the Sugiyama method.

Keywords: Graph drawing, Sugiyama method, radial drawings, recurrent hierarchies

Joint work of: Bachmaier, Christian; Brandenburg, Franz-J.; Brunner, Wolfgang; Lovácz, Gergő

Combining GD & InfoVis Techniques to Support Sensemaking

Carsten Goerg (Georgia Institute of Technology, US)

In this talk I present how techniques from the areas of graph drawing and information visualization can be combined in an interactive tool that supports a particular type of sensemaking: finding embedded narratives or plots across large collections of relatively brief text documents.

Our approach is decidedly human-centered and we have created the Jigsaw system as an embodiment of these principles. Jigsaw provides multiple visualizations depicting the documents, the entities within them, and connections between the two. Analysts interactively explore the entities and documents as a way to guide document selection for reading. Possible application domains of our approach are criminalistic investigations or the exploration of connections between biological entities, such as proteins and pathways, in scientific papers.

Keywords: Sensemaking, investigative analysis, information foraging, information visualization, multiple views

An InfoVis View on Graph Drawing

Andreas Kerren (Växjö University, SE)

Aim of this talk is to show the InfoVis view on the visualization of large graphs or networks. Firstly, I give a brief overview of the most important aspects and tasks of this field including a presentation of current challenges.

Many of the challenges can be easily transferred to the Graph Drawing field.

Furthermore, the talk discusses human-centered aspects and perspectives as these issues become more and more important in the development of InfoVis tools.

In a second part, a choice of several example tools and approaches is presented, especially regarding their interaction possibilities. Here, I focus on specific problems in the context of large graphs, for example, Focus & Context, relations between different (networked) data types, avoiding clutter within networks' representations, etc.

Keywords: Information Visualization, Graph Drawing, Visualization, Human-centered Information Visualization

Networks of European Science in the 18th Century

Lothar Krempel (MPI für Gesellschaftsforschung - Köln, DE)

Science historians have large and detailed databases on letters sent between researchers of the 18th century.

The talk presents results of collaboration with various research groups which allows to reconstructs the networks and overall topology of the trans-European "Gelehrtenrepublik" based on written communications.

Keywords: Social networks, scientific correspondences, botanics, 18th century

Joint work of: Krempel, Lothar; Hächler, Stefan; Ruisinger, Marion; Stuber, Martin

Software Demo: Constraint-based Exploration of Large Graphs in Dunnart

Tim Dwyer (Monash University - Clayton, AU)

In this talk we demonstrate Dunnart⁵, a constraint-based network diagram editor. In particular, we show Dunnart's ability to browse large graphs, such as those used by the bioinformatics community.

In Dunnart, constrained network layout is run continuously during interaction. It constantly improves the layout while the user moves or resizes nodes. In this model, constraints are used to maintain non-overlap and preserve the topology of the graph during interaction, preventing shapes from crossing shapes or connectors. The user is able to pause the layout and make manual changes to the topology by holding down the ALT key. Dunnart offers the user constraint-based placement tools for alignment, distribution or separation. These create persistent relationships that will be enforced throughout further editing. Such tools, are useful for drawing attention to important sections of the diagram, or can allow parts of the layout to be constrained to match recognised textbook biological layouts.

We demonstrate how the continuous network layout model can be interactively used for exploration of large graphs, showing examples from the biological sciences and software engineering domains. We use a overview+detail model where a fast mostly-unconstrained force-directed layout method is used to layout the entire large network in an overview window. A detailed view shows a dynamic subnetwork of around forty nodes.

This detailed view is arranged with our standard topology-preserving constrained graph layout. One of the nodes is considered the active node and we determine a surrounding network, based on graph theoretic distance. Nodes connected (via edges) to parts of the graph not shown in the detailed view are displayed with spikes to indicate the connections to unshown sections of the graph. The user can middle-click on a node to make it the focal node. This will centre the detailed view on it and bring into view the network around it.

Once nodes have been shown in the detailed view, their positions are subsequently locked in the overview window during relayout. Their topology and any placement relationships they are involved in are remembered and restored when the user brings them back into view. In this way the user is able to easily move throughout a large network and incrementally tidy and beautify it while exploring.

Keywords: Graph drawing, constraint-based layout, large graph exploration, visualization software

Joint work of: Dwyer, Tim; Wybrow, Michael

⁵ <http://www.csse.monash.edu.au/~mwybrow/dunnart/>

Software Demo: SPIKE - Signaling Pathway Integrated Knowledge Engine

Idan Zohar (Tel Aviv University, IL)

Biological signaling pathways that govern cellular physiology form an intricate web of tightly regulated interlocking processes. Data on these regulatory networks are accumulating at an unprecedented pace. The assimilation, visualization and interpretation of these data have become a major challenge in biological research, and once met, will greatly boost our ability to understand cell functioning on a systems level.

To cope with this challenge, we are developing the SPIKE knowledge-base of signaling pathways. SPIKE contains three main software components: 1) A database (DB) of biological signaling pathways. Carefully curated information from the literature and data from large public sources constitute distinct tiers of the DB. 2) A visualization package that allows interactive graphic representations of regulatory interactions stored in the DB and superposition of functional genomic and proteomic data on the maps. 3) An algorithmic inference engine that analyzes the networks for novel functional interplays between network components.

Keywords: Signaling pathways, visualization software

08191 Working Group Report: Visualization of Trajectories

We considered the following problem: Given a set of vertices V and a set of paths P , where each path is a sequence of vertices, represent these paths somehow.

We explored representations in different dimensions and with different conditions on the paths.

Keywords: Graph drawing, trajectories, paths

Joint work of: Borgatti, Stephen; Brandes, Ulrik; Kaufmann, Michael; Kobourov, Stephen; Lubiw, Anna; Wagner, Dorothea

Full Paper: <http://drops.dagstuhl.de/opus/volltexte/2008/1555>

08191 Working Group Report: Graph Drawing Techniques for Successive Augmentation of Multiple Layer Drawings

Joe Fowler (University of Arizona, US)

Overview presentation of laying out successive series of graphs in which the layout of the first graph is given but can be manipulated so long as orthogonal ordering, edge direction, and structures within the first graph are maintained.

Keywords: Graph drawing, multiple layouts

Joint work of: Fowler, Joe; Gansner, Emden; Hong, Seok-Hee; Kaufmann, Michael; Krempel, Lothar

08191 Working Group Report: Navigation in large graphs

Georg Sander (ILOG - Bad Homburg, DE)

This is the report (slides) of the working group "Navigation in large graphs". Graphs from social sciences and bioinformatics are often so large that it is difficult to handle them. Given a large graph with multilevel clustering, we developed an algorithm to navigate through the graph by means of an overview and a detail view. The layout of the detail view is calculated on the fly when the corresponding region becomes visible the first time, and is kept to have a consistent view when the user moved the detail view back to a region already seen.

This leads to a partial layout algorithm that respects already known area placements of clusters.

Keywords: Graph drawing, large graph, navigation

Joint work of: Sander, Georg; Lauther, Ulrich; Nikolov, Nikola; Wong, Hoi-Ming; Nachmanson, Lev

08191 Working Group Report: X-graphs of Y-graphs and their Representations

We address graph decomposition problems that help the hybrid visualization of large graphs, where different graphic metaphors (node-link, matrix, etc.) are used in the same picture. We generalize the X -graphs of Y -graphs model introduced by Brandenburg (Brandenburg, F.J.: Graph clustering I: Cycles of cliques. In Di Battista, G., ed.: Graph Drawing (Proc. GD '97). Volume 1353 of Lecture Notes Comput. Sci., Springer-Verlag (1997) 158–168) to formalize the problem of automatically identifying dense subgraphs (Y -graphs, clusters) that are prone to be collapsed and shown with a matricial representation when needed. We show that (planar, K_5)-recognition, that is, the problem of identifying K_5 subgraphs such that the graph obtained by collapsing them is planar, is NP-hard. On the positive side, we show that it is possible to determine the highest value of k such that G is a (planar, k -core)-graph in $O(m + n \log(n))$ time.

Keywords: Graph drawing, X-graphs of Y-graphs, visualization of large graphs

Joint work of: Batagelj, Vladimir; Brandenburg, Franz J.; Didimo, Walter; Liotta, Guiseppe; Patrignani, Maurizio

Full Paper: <http://drops.dagstuhl.de/opus/volltexte/2008/1556>

08191 Working Group Report: Edge Thresholding

When working with very large networks it is typical for scientists to present a “thinned out” version of the network in order to avoid the clutter of the entire network. For example in the hypothetical case of illustrating trading patterns between groups of nations it might be appropriate to limit the inclusion of inter-nation edges to all those that are significant in terms of their weight but do not, say, associate with a country outside the grouping. Arising from a discussion during one of the introductory sessions we became interested in a problem relating to the discovery of “key events” in a network, in terms of an ordered addition of edges to the network.

Keywords: Graph drawing, edge thresholding

Joint work of: Healy, Patrick; Dwyer, Tim

Full Paper: <http://drops.dagstuhl.de/opus/volltexte/2008/1557>

08191 Working Group Report: Visually Comparing a Set of Graphs

Markus Geyer (Universität Tübingen, DE)

We consider methods to visually compare graphs, more to focus on the differences of the graphs than on the similarities. Our two-level approach constructs a meaningful overview of the given graphs combined with a detailed view focusing on a local area of change.

The actual layout of these graphs has to be evaluated depending on the specific type of biological network to be visualized in each case. We look into different variants and propose properties to be optimized in our visualizations.

Keywords: Graph drawing, visual graph comparison

Joint work of: Geyer, Markus; Estrella-Balderrama, Alex; Gutwenger, Carsten; Klein, Karsten; Schulz, Michael; Albrecht, Mario; Kohlbacher, Oliver

Full Paper: <http://drops.dagstuhl.de/opus/volltexte/2008/1553>

08191 Working Group Report: Some open problems in molecular biology from a graph drawing perspective

Falk Schreiber (IPK Gatersleben, DE)

Biological data generated and analyzed in the life sciences can be often interpreted and represented by graphs.

Although several general and special-purpose tools and libraries are available for drawing biological networks, there is a great demand for better approaches as existing tools are either not adequate for handling large graphs or do not adhere to the special drawing conventions and recognized layouts of biological networks. In this discussion group, we tried to identify some representative use cases that demonstrate the need for advanced algorithms for presenting, exploring, evaluating, and comparing biological network data. The result of the discussion has been published as technical report http://ls11-www.cs.uni-dortmund.de/downloads/papers/TR08_1_003.pdf.