Graph Data Format Workshop Report

Ulrik Brandes[1], M. Scott Marshall[2], and Stephen C. North[3]

[1] University of Konstanz, Department of Computer & Information Science
Box D 188, 78457 Konstanz, Germany. `Ulrik.Brandes@uni-konstanz.de`
[2] Centrum voor Wiskunde en Informatica, P.O. Box 94079, 1090 GB Amsterdam, The Netherlands. `scott@cwi.nl`
[3] AT&T Research, 180 Park Ave., Bldg. 103, Florham Park, NJ 07932-0971, U.S.A.
`north@research.att.com`

Abstract. Prompted by the increasing demand for a standard exchange format for graph data, an informal workshop was held in conjunction with Graph Drawing 2000. The participants identified requirements for such a standard and formed a group to work out a proposal. The current status of this effort is publicly available at `http://www.graphdrawing.org/data/format/`.

1 Introduction

Graph drawing tools need to store and exchange graph data. Despite several earlier attempts to define a standard, no agreed-upon format is widely accepted, and indeed, many packages support only their own custom format.

Motivated by the goals of tool interoperability, access to benchmark data sets, and, most importantly, data exchange over the Web, the Steering Committee of the Graph Drawing Symposium started a new initiative. In a concerted effort, a standard format will be defined and proposed to developers of graph drawing and related software.

As a first step towards this end, an informal workshop was announced on the GD 2000 Web page[1] and held the day before the start of the symposium. The workshop featured two invited presentations and plenty of discussion.

Date and location:	20 September 2000, 14:00—17:30, Colonial Williamsburg
Local organization:	Joe Marks and Kathy Ryall
Minutes:	Helen Purchase
Chair:	Ulrik Brandes
14:00–14:20	Stephen North: *Introduction to XML*
14:20–14:40	Scott Marshall: *Survey of Graph Data Formats*
15:00–17:30	Discussion blocks on - general requirements - specific goals - future activities

[1] `http://www.cs.virginia.edu/~gd2000/`

The widespread interest in a standard was underscored by the unexpectedly large number of participants. With several others noting that they were, unfortunately, unable to attend, the following persons were present at the meeting:

- Ulrik Brandes (University of Konstanz)
- Stina Bridgeman (Brown University)
- Giuseppe Di Battista (University of Rome III)
- Peter Eades (University of Sydney)
- Ashim Garg (State Universty of New York)
- Carsten Gutwenger (Max Planck Institute for Computer Science)
- Michael Himsolt (DaimlerChrysler Research)
- Seokhee Hong (University of Sydney)
- Michael Jünger (University of Cologne)
- Michael Kaufmann (University of Tübingen)
- Sebastian Leipert (University of Cologne)
- Giuseppe Liotta (University of Perugia)
- Joe Marks (Mitsubishi Electric Research)
- Scott Marshall (CWI Amsterdam)
- Petra Mutzel (Vienna University of Technology)
- Stephen North (AT&T Research)
- Maurizio Patrignani (University of Rome III)
- Helen Purchase (University of Queensland)
- Kathy Ryall (University of Virginia)
- Galina Shubina (Brown University)
- Susan Sim (University of Toronto)
- Roberto Tamassia (Brown University)
- Luca Vismara (Brown University)
- Vance Waddle (IBM Research)

2 Summary of Discussion

The discussion was held in an open and constructive atmosphere. None of the topics required vote-taking, since each was discussed until either there was no noticeable dissent or a topic was identified as requiring more detailed investigation.

In a first round of discussion, the scope of the project was delineated by gathering potentially relevant features of a common exchange format. Among such requirements were the support of hypergraphs, clustering, hierachical graphs, graph properties and corresponding certificates, dynamic graphs, 3D graphics, animation, and even layout algorithms as part of the description of a graph drawing. In addition, issues such as scalability and human readability were addressed briefly.

Given this vast array of options, the unanimous decision was to focus on a simple format that should be sufficient for the most common cases, yet extensible to incorporate more advanced features. However, it became obvious that even for a simple format, several difficult decisions need to be made. At this point, Susan Sim was asked to report on the current state of these decisions in a similiar project called GXL,[2] aimed at defining an exchange format for graphs arising in software reengineering and graph transformation. Subsequently, several of these issues were discussed in some depth, for instance, wether a file should contain unique identifiers for nodes (yes) and edges (optionally), or whether a graph should be allowed to have both directed and undirected edges (yes), or parallel edges (yes) or loops (yes).

While opinions on some of these details differ from the path currently taken in GXL, it was agreed that the proposal should be defined in cooperation with the GXL consortium and other relevant groups.

3 Results

The consensus among the participants was that a graph data format should have a layered architecture that conceptually separates the following four aspects (in order of increasing level of detail):

1. structure (vertices, edges, incidences, attributes)
2. topology (ordering of edges, crossings, etc.)
3. shape (bends, feature shapes, etc.)
4. geometry (positions, feature sizes, etc.)
5. rendering (colors, styles, icons, textures, etc.)

The number of layers of the format need not necessarily match this list, but it should be possible to omit higher layers from a graph description. The design should be open to allow formats for applications other than graph drawing to specify other characteristics on top of the lower layers. In particular, the goal is to collaborate with groups working on a data exchange format for software reengineering and graph transformations.[3]

While extensions to incorporate layout algorithms, dynamic graphs, clustering, and so on may not be part of the initial proposal, the need for such extensions should be taken into account.

As a result of the discussion, a working group was formed, and several others have joined since, or expressed interest in contributing. The group will focus on defining the layer structure with extension points. It is understood that a draft proposal will be made available for public review before the 2001 Symposium on Graph Drawing (GD 2001), and that the proposal will be considered for publication in the GD 2001 proceedings. The current status of the project is maintained at `http://www.graphdrawing.org/data/format/`.

[2] Graph Exchange Language, see `http://www.gupro.de/GXL/`.
[3] See `http://www.cs.toronto.edu/~simsuz/wosef/` for more information.