What is a Good Diagram? (Revisited)

(Invited Keynote)

Peter Eades
School of Information Technologies
University of Sydney
Sydney, Australia
peter.eades@sydney.edu.au

ABSTRACT

Graphs have been broadly used to model binary relations since the beginning of Computer Science. Nodes represent entities, and edges represent relationships between entities. Such models become more useful when the graph model is represented as a diagram, because visualization of a graph enables humans to understand the underlying model.

A quality metric assigns a number $q(D)$ to each diagram $D$ such that $q(D)$ is larger than $q(D')$ when $D$ is a higher quality diagram than $D'$. Quality metrics for graph visualization have been discussed since the 1970s. Sugiyama et al. [3] wrote lists of quality metrics and his subsequent book [7] contains an extensive discussion. The seminal paper “What is a good diagram?” by Batini et al. [1] presented guidelines for database diagrams. These early works were entirely based on intuition and introspection; later Purchase et al. [4] began the scientific investigation of quality metrics with a series of experiments that validated some of the metrics.

Of course, the quality of a diagram is “a hopeless matter to define formally” (Batini et al. [1]): quality depends on specific users, specific tasks, and rather informal and subjective notions of aesthetics. Nevertheless, formal quality metrics are helpful if not essential in the design of automatic graph visualization methods, because such methods are optimisation algorithms with quality metrics as objective functions. As an example, it is well established that edge crossings in a diagram inhibit human understanding, and edge crossings form the basis of so-called “planarity-based” quality metrics. Methods that reduce the edge crossings have received considerable attention in the literature (see, for example, Jünger and Mutzel [2]).

In this talk we review the history of quality metrics for graph visualization, and suggest a new approach.

The new approach is motivated by two observations: (1) the size of data sets is much larger now than ever before, and it is not clear that established quality metrics are still relevant, and (2) there is a disparity between methods used in practice and methods used in academic research.

Using a pipeline model of graph visualization, we classify quality metrics into “readability” metrics and “faithfulness” metrics. Readability metrics measure how well the human user perceives the diagram; these metrics have been extensively investigated and they are (at least partially) understood. Faithfulness metrics (see Nguyen et al. [6]) measure how well the diagram represents the data; these metrics are not well developed and they are poorly understood.

We argue that faithfulness metrics become more relevant as the data size increases, and suggest that the commercial dominance of stress-based methods over of planarity-based methods is somewhat due to performance on faithfulness metrics.

We introduce some specific faithfulness metrics aimed at large graphs. In particular, we suggest that metrics based on proximity graphs (see Toussaint [5]) may play a key role.

Much of this talk is based on joint work and discussions with Karsten Klein, SeokHee Hong, and Quan Nguyen, among others.
BIOGRAPHY

Peter Eades is currently in transition to retirement, but still holds a position as Research Professor at the University of Sydney. Previous to this he held positions at the University of Queensland, the University of Newcastle, and NICTA, as well as visiting positions in Europe, Japan, the USA and Canada. He has current research projects in Graph Drawing, Interactive Visualization, Large Displays, and Algorithmics of Geometric Graphs. Some previous research projects have been in Software Visualization, Tabletop Displays and Remote Collaboration, Graph Algorithms, Combinatorial Optimization, and Haptic Shoes.

REFERENCES