Validating Graph Drawing Aesthetics

Helen C. Purchase[1] and Robert F. Cohen[2] and Murray James[1]

[1] Department of Computer Science, University of Queensland,
St. Lucia, Queensland, 4072, Australia.
[2] Department of Computer Science, University of Newcastle,
Callaghan, New South Wales, 2308, Australia

Abstract. Designers of graph drawing algorithms and systems claim to illuminate application data by producing layouts that optimize measurable aesthetic qualities. Examples of these aesthetics include *symmetry* (where possible, a symmetrical view of the graph should be displayed), *minimize edge crossings* (the number of edge crossings in the display should be minimized), and *minimize bends* (the total number of bends in polyline edges should be minimized).
The aim of this paper is to describe our work to validate these claims by performing empirical studies of human understanding of graphs drawn using various layout aesthetics. This work is important since it helps indicate to algorithm and system designers what are the aesthetic qualities most important to aid understanding, and consequently to build more effective systems.

1 Introduction

The visualization produced by a graph drawing subsystem should *illuminate* application data. That is, it should help the user to understand and remember the information being visualized. A good layout can be a picture worth a thousand words; a poor layout can confuse or mislead. Designers of graph drawing algorithms and systems claim to illuminate application data by producing layouts that optimize measurable aesthetic qualities. Examples of these aesthetics include *symmetry* (where possible, a symmetrical view of the graph should be displayed, see e.g. [4, 7]), *minimize edge crossings* (the number of edge crossings in the display should be minimized, see e.g. [5]), and *minimize bends* (the total number of bends in polyline edges should be minimized, see e.g. [10, 11]).

The aim of this paper is to describe our work to validate these claims by performing empirical studies of human understanding of graphs drawn using various layout aesthetics. This work is important since it helps indicate to algorithm and system designers what are the aesthetic qualities most important to aid understanding, and consequently to build more effective systems.

There has been very little work in this area. This is the first study to address the effect of aesthetics on general understanding of graphs. Prior work has focused on specific applications [1, 3] and comparisons of two- and three-dimensional layouts [12].

Although many algorithms consider more than one aesthetic in their attempt to create an illuminating graph display, this initial study focuses on individual aesthetic principles, leaving the testing of the effectiveness of the various algorithms which embody them for a further project. We study the following popular graph drawing aesthetics: minimize edge crossings, minimize bends, and maximize symmetries. Our initial results are quite encouraging. Our results indicate that minimizing edge bends and minimizing edge crossings are both important aids to human understanding. However, our study is inconclusive as to the importance of maximizing symmetry. This final non-result is surprising to us, since we believed that maximizing symmetry would have a high impact on the understandablity of drawings.

The rest of the paper is organized as follows. Section 2 describes the scope of our experiment, section 3 explains our methodology, section 4 presents our results, and section 5 describes our analysis. Finally, in section 6 we give conclusions and directions for future work.

2 Scope

2.1 Hypotheses

Three common graph-drawing aesthetics were chosen for this study: *symmetry* (where possible, a symmetrical view of the graph should be displayed), *minimize edge crossings* (the number of edge crossings in the display should be minimized), and *minimize bends* (the total number of bends in polyline edges should be minimized). We restricted our study to undirected graphs. Other aesthetics that were considered were the direction of flow for directed graphs, and maximising the minimum angle between arcs from one node: these have been left for a later study.

Thus, our hypotheses were:

- Increasing the number of arc bends in a graph drawing decreases the understandability of the graph
- Increasing the number of arc crossings in a graph drawing decreases the understandability of the graph
- Increasing the local symmetry displayed in a graph drawing increases the understandability of the graph

To measure the understandability of a graph drawing, three graph-theoretic questions were chosen. The performance of a person in answering these questions about a graph drawing was considered an appropriate indicator of their understanding of the drawing. The three questions chosen were:

- How long is the shortest path between two given nodes?
- What is the minimum number of nodes that must be removed in order to disconnect two given nodes such that there is no path between them?
- What is the minimum number of arcs that must be removed in order to disconnect two given nodes such that there is no path between them?

2.2 Graphs

A dense graph and a sparse graph were defined. The sparse graph has 16 nodes and 18 arcs. The dense graph was created by adding 10 extra arcs onto the sparse graph.

Nine drawings of each graph were created, with the number of bends, crossings and the amount of perceived symmetry varied appropriately. Metrics for the number of crossings and bends were easy: we merely counted them; determining a metric to measure the amount of symmetry in a graph drawing was more difficult. We needed a metric that could give a numerical value to perceived symmetry. The method used in this study is described in the appendix (see also [2]).

For each aesthetic, three drawings of each graph were produced, one with a small aesthetic measurement (*few*), one with an interim aesthetic measurement (*some*), and one with a large aesthetic measurement (*many*). The values of these measurements for the dense graph drawings were as follows:

dense graph drawing		bends	crossings	symmetry metric
dense-bends-few	**(dbf)**	6	0	0
dense-bends-some	**(dbs)**	18	0	0
dense-bends-many	**(dbm)**	30	0	0
dense-crossings-few	**(dcf)**	0	6	0
dense-crossings-some	**(dcs)**	0	24	0
dense-crossings-many	**(dcm)**	0	42	0
dense-symmetry-few	**(dsf)**	0	0	4.57
dense-symmetry-some	**(dss)**	0	0	25.7
dense-symmetry-many	**(dsm)**	0	0	51.0

The nine sparse graph drawings had a similar variation in the values of the aesthetics, maintaining the same *few*:*some*:*many* ratios as used in the dense graph drawings as much as possible (see Figs. 1 and 2).

3 Methodology

Experiments were conducted: 49 subjects worked on the dense graph, and 35 of these same subjects worked on the sparse graph a week later. The subjects were all second year computer science students, as representatives of a more general target population who use graph drawings in their work. They were each given an experimental booklet of the following form:

1. A brief description of graphs, and definitions of the terms *node, arc, path, path length.*
2. Explanation of the three graph-theoretic questions that they would be required to answer about the experimental graphs.

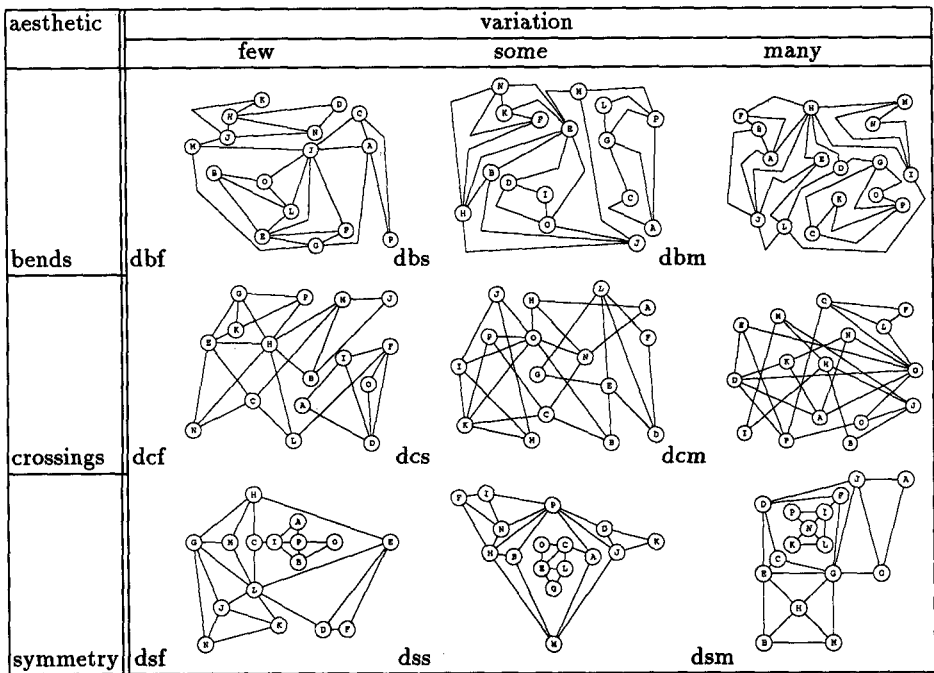| aesthetic	variation		
	few	some	many
bends	dbf	dbs	dbm
crossings	dcf	dcs	dcm
symmetry	dsf	dss	dsm

Fig. 1. The Experimental Graph Drawings: dense graph

3. A simple example graph drawing (including some bends, crossings and symmetry), with three representative questions and their correct answers.
   At this stage in the experiment, the subjects were asked if they had any questions about graphs in general, or about the experiment. It was important to ensure that all the subjects knew what was going to be expected of them.
4. Six drawings of two "practise" graphs (sparse: 20 nodes and 21 arcs; dense: 20 nodes and 32 arcs). Each graph was drawn three times: once with crossings, once with bends and once with some symmetry. The three graph theoretic questions were asked of each of these drawings. The purpose of this section was to get the subjects used to the nature of graph drawings and the time constraints, and to ensure that by the end of this section they were comfortable with the task, before tackling the nine experimental graphs. The subjects were unaware that these graph drawings were not experimental.
5. A "filler" task which engaged the subjects' mind on a small problem unrelated to graphs. This was to ensure that their performance on the subsequent experimental graphs was not affected by any follow-on effect from the practise graphs. We used a simple word puzzle.
6. The nine experimental graph drawings (all sparse, or all dense), in a random order for each subject.
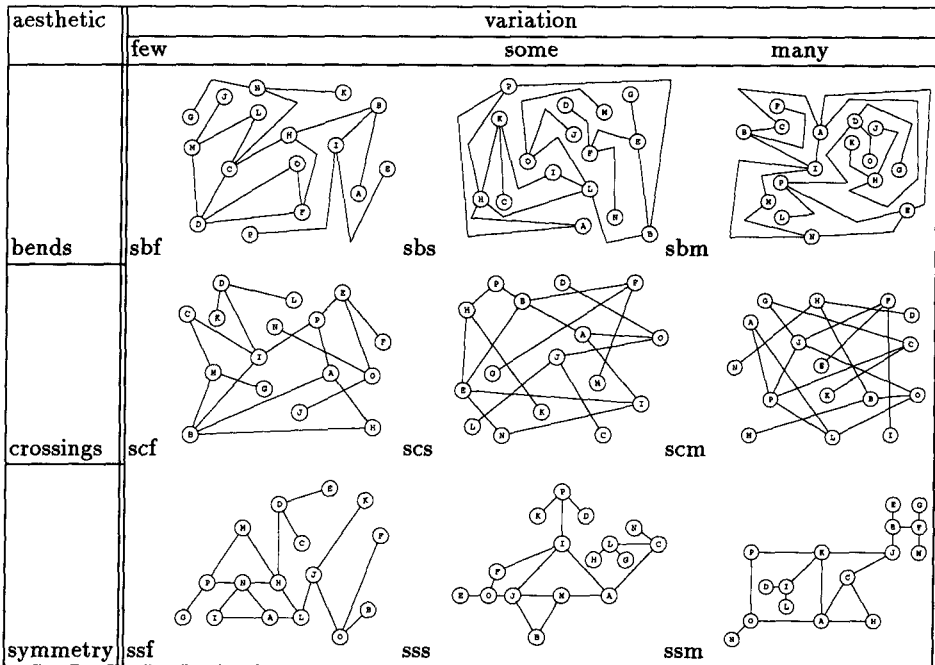
Fig. 2. The Experimental Graph Drawings: sparse graph

7. A questionnaire, asking the subjects about their current degree, how far they were into their degree, any previous experience of reading graphs, and if there were any particular features about the graphs that they could identify as assisting or hindering their performance (this information was for our own interest only: it was not used in the quantitative analysis of the results).

The time allowed for each component of the experiment was as follows:

Task	Time: dense graph	Time: sparse graph
Introduction and example	As necessary	As necessary
Practise graph 1	60 seconds	45 seconds
Practise graph 2	60 seconds	45 seconds
Practise graph 3	55 seconds	40 seconds
Practise graph 4	55 seconds	40 seconds
Practise graph 5	50 seconds	35 seconds
Practise graph 6	50 seconds	35 seconds
Filler task	60 seconds	60 seconds
Experimental graphs 1–9	45 seconds	30 seconds

The experiments were thus controlled for time, the graph structure and the questions; the independent variables were the number of bends, crossings and

the value of the symmetry metric, and the dependent variable was the number of errors each subject made in answering the questions for each graph drawing.

3.1 Graph Drawing and Question Variation

To ensure that it was not obvious that the underlying graph was the same for each graph drawing, the nodes in each drawing were randomly labeled with the letters A – P, thus preserving the graph structure, but ensuring that the relationships between particular named nodes were different.

The questions themselves were randomised too: although the same three graph-theoretic questions were asked of each drawing, the pair of nodes chosen for each question was randomly selected from a list of node-pairs. This list of node-pairs was defined so that there was a range of possible correct answers to each question for each graph drawing.[3] This was to ensure that any variability in our data could not be explained away by the varying difficulty of the questions, which may have been the case if each graph drawing had had a fixed set of question-answer pairs.

4 Results

A within-subjects analysis method was used in order to reduce any variability that may have been attributable to the difference between the subjects (age, experience *etc*). Any learning effect was minimised by of the large number of graphs used in the experiment, the inclusion of the practise graphs, and the randomisation of the ordering of the graph drawings.

For each graph, the results for each aesthetic were collated and ranked. As an example, Table 1 shows the results for the *dense-bends* drawings, and the associated ranking for each subject.

The *Friedman analysis of variance by ranks* test was used to determine the χ^2 value for each aesthetic, and the probability that the ranks were produced by chance (p). The results were:

graph	aesthetic	χ^2	p
dense	bends	17.48	< 0.001
	crossings	23.09	< 0.001
	symmetry	2.3	< 0.3
sparse	bends	10.99	< 0.01
	crossings	10.56	< 0.01
	symmetry	0.27	< 0.9

---

[3] Thus, for each graph drawing, four node-pairs were defined that would give correct answers to the first question of either 2, 3, 4 or 5, two node-pairs were defined that would give correct answers to the second question of either 1 or 2, and three node-pairs were defined that would give correct answers to the third question of either 1, 2 or 3.

With a level of significance set at $\alpha = 0.05$, these results are therefore significant for bends and crossings in both the dense and sparse graphs, as $p < \alpha$ in both these cases.

Table 1. subject scores and rankings for *dense-bends* drawings

Subject	Scores			Rankings			Subject	Scores			Rankings		
	dbf	dbs	dbm	dbf	dbs	dbm		dbf	dbs	dbm	dbf	dbs	dbm
1	3.0	3.0	3.0	2.0	2.0	2.0	26	3.0	2.0	2.0	1.0	2.5	2.5
2	3.0	3.0	0.0	1.5	1.5	3.0	27	3.0	3.0	2.0	1.5	1.5	3.0
3	3.0	1.0	1.0	1.0	2.5	2.5	28	2.0	2.0	2.0	2.0	2.0	2.0
4	2.0	2.0	1.0	1.5	1.5	3.0	29	2.0	1.0	2.0	1.5	3.0	1.5
5	0.0	0.0	0.0	2.0	2.0	2.0	30	3.0	3.0	1.0	1.5	1.5	3.0
6	2.0	2.0	1.0	1.5	1.5	3.0	31	2.0	2.0	1.0	1.5	1.5	3.0
7	1.0	0.0	1.0	1.5	3.0	1.5	32	2.0	2.0	0.0	1.5	1.5	3.0
8	2.0	1.0	1.0	1.0	2.5	2.5	33	2.0	3.0	3.0	1.0	2.5	2.5
9	0.0	1.0	0.0	2.5	1.0	2.5	34	3.0	3.0	3.0	2.0	2.0	2.0
10	0.0	1.0	1.0	3.0	1.5	1.5	35	3.0	3.0	3.0	2.0	2.0	2.0
11	3.0	3.0	1.0	1.5	1.5	3.0	36	2.0	2.0	2.0	2.0	2.0	2.0
12	3.0	3.0	3.0	2.0	2.0	2.0	37	2.0	3.0	2.0	2.5	1.0	2.5
13	3.0	3.0	1.0	1.5	1.5	3.0	38	2.0	3.0	1.0	2.0	1.0	3.0
14	3.0	2.0	1.0	1.0	2.0	3.0	39	3.0	3.0	2.0	1.5	1.5	3.0
15	2.0	2.0	1.0	1.5	1.5	3.0	40	3.0	2.0	1.0	1.0	2.0	3.0
16	2.0	3.0	0.0	2.0	1.0	3.0	41	3.0	2.0	2.0	1.0	2.5	2.5
17	2.0	1.0	1.0	1.0	2.5	2.5	42	2.0	2.0	1.0	1.5	1.5	3.0
18	1.0	2.0	2.0	3.0	1.5	1.5	43	1.0	2.0	2.0	3.0	1.5	1.5
19	1.0	1.0	0.0	1.5	1.5	3.0	44	2.0	1.0	1.0	1.0	2.5	2.5
10	3.0	2.0	2.0	1.0	2.5	2.5	45	1.0	3.0	2.0	3.0	1.0	2.0
21	2.0	1.0	2.0	1.5	3.0	1.5	46	1.0	2.0	2.0	3.0	1.5	1.5
22	3.0	3.0	2.0	1.5	1.5	3.0	47	3.0	3.0	3.0	2.0	2.0	2.0
23	2.0	2.0	1.0	1.5	1.5	3.0	48	3.0	3.0	2.0	1.5	1.5	3.0
24	3.0	2.0	1.0	1.0	2.0	3.0	49	2.0	3.0	0.0	2.0	1.0	3.0
25	3.0	0.0	1.0	1.0	3.0	2.0							

Figs. 3 and 4 show the trends in the average number of errors made by each subject, when plotted against the variation in each aesthetic.

One of the questions we asked the subjects at the end of the questionnaire was: "Can you identify any characteristics of the graph drawings that made them more difficult to read?" Responses to this question included:

- "Many interwoven arcs like spiderwebs"
- "Lines crossing over"
- "Crooked lines"
- "The zig-zag lines and lack of efficient structure"
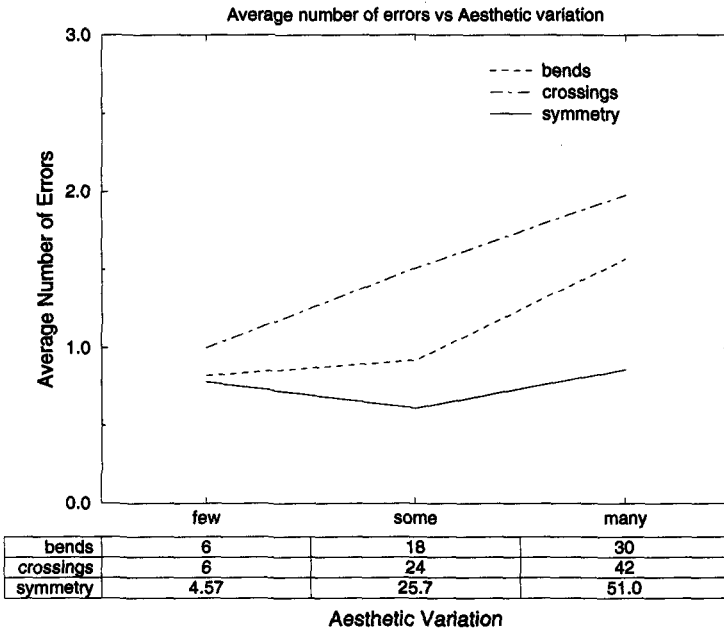- "Disorganised, unstructured"
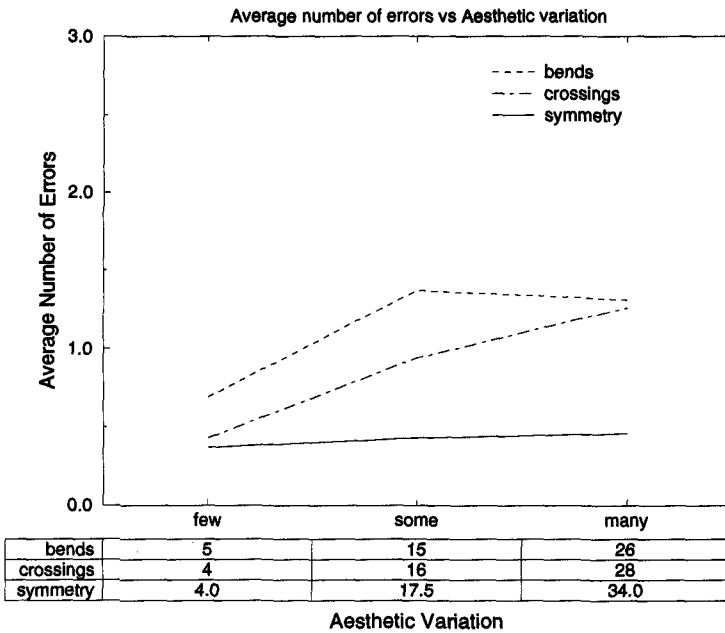
Fig. 3. Results for the dense graph

	few	some	many
bends	6	18	30
crossings	6	24	42
symmetry	4.57	25.7	51.0

Fig. 4. Results for the sparse graph

	few	some	many
bends	5	15	26
crossings	4	16	28
symmetry	4.0	17.5	34.0

5   Analysis

Two of our hypotheses are confirmed in both sparse and dense graphs (crossings and bends), but the symmetry hypothesis is inconclusive. We believe that this is due to a 'ceiling effect'. The length of time for the subjects to answer the questions (45 seconds for dense, 30 seconds for sparse), was too generous: most of the subjects managed to get all the questions right for the symmetrical graphs (even those which had a low measure of symmetry). This meant that there was no variability in the symmetry data, and thus the results of this analysis are not significant. In pilot tests, when we used times as high as 75 seconds for dense graphs and 60 seconds for sparse graphs, we observed a similar ceiling effect on the graph drawings for all three aesthetics.

Additionally, symmetry is very different in nature to the other two aesthetics considered here. Other significant points which may explain why the symmetry hypothesis was not as conclusive as the other two hypotheses are:

- The symmetry metric we defined may not adequately quantify perceptual symmetry. In future work we will examine weighting the metric by the proportion of nodes involved in symmetrical relations as well as the proportion of arcs. We also only considered local symmetry: it is possible that running similar experiments with global symmetries might produce better results.
- Secondly, these experiments tested the ability of the subjects to understand the structure of graph drawings. It is possible that maximising the amount of perceptual symmetry in a graph drawing is of more benefit when the symmetry is used for the perceptual organisation of semantic concepts instead of merely the organisation of the graph structure. Future experiments will test subjects' ability to understand the meaning of graph drawings from specific application areas.
- In controlling the experiment, we avoided introducing potential interaction between the aesthetics by having no variation of two aesthetics while varying the third. For example, all the graph drawings that varied the number of bends had no crossings and a zero value for symmetry. There is a potential conflict here between the nature of our three hypotheses: with bends and crossings, the hypothesis was that *increasing* their number would result in increased errors; with symmetry, the hypothesis was that *decreasing* the value of the metric would result in increased errors. By keeping the number of crossings and bends at zero for the symmetrical graph drawings, the drawings were being made simpler. This is a contrast to the bends and crossings diagrams, where keeping the symmetry metric at zero made the drawings more complex. Therefore, perhaps it is no surprise that a ceiling effect was observed with the reading of the symmetry drawings, when the same time period was used for reading all the graphs of the three aesthetics.

It is possible that, although we cannot state conclusively that the symmetry hypothesis is confirmed, it may indeed be the case that increasing the number of crossings and bends is more detrimental to the understandability of a graph drawing than a reduction in symmetry.

6   Conclusions and Further Work

In this paper, we have described our study to validate important graph drawing aesthetics. Our study confirms two of our hypotheses:

- Increasing the number of arc bends in a graph decreases the understandability of the graph.
- Increasing the number of arc crossings in a graph decreases the understandability of the graph.

These results indicate that the general intuition that edge bends and edge crossings should be minimized is correct.

Our third hypothesis remains unconfirmed:

- Increasing the local symmetry displayed in a graph increases the understandability of the graph.

We plan to conduct a further study to investigate this hypothesis further. We are building an interactive system to conduct experiments which will allow us to get more valid and conclusive results, and to expand our studies. It will enable us to:

- measure the time taken to answer the questions correctly as the dependent variable, rather than the number of errors made within a set time
- study the effect of aesthetics on human understanding of graphs drawn on a display device
- compare the understandability of drawings on paper and on a display
- design and perform other experiments

Future work includes the following:

- studying other graph drawing aesthetics (for example, orthogonality, maximizing the minimum angle, etc)
- studying the effectiveness of various graph drawing algorithms
- studying the effectiveness of graph drawing aesthetics when used in application domians (for example, software enginnering design diagrams)
- studying the effect of aesthetics in three-dimensional drawings.

Appendix: A Metric for Symmetry

A measurement of the amount of symmetry in a graph drawing was an essential part of the study: we needed to be able to vary the symmetry displayed in a graph drawing, and to have some metric by which we could measure it. For the purposes of this experiment, symmetry is quantified by [2]:

1. counting all the symmetrical relationships about each axis of symmetry.
2. adding the symmetry values from all the axes together.

3. multiplying this total by the proportion of the graph drawings' arcs which are symmetrical objects.

This resulted in the following formula for measuring symmetry:

$$\sigma = \left(\frac{A_o}{A_{total}}\right) \times \sum_{i=1}^{M} \left(\frac{A_{di}}{2} + \frac{N_i}{2} + A_{si}\right)$$

where:

- σ is the measure of symmetry; the higher the value of σ, the more symmetrical the drawing
- A_o is the number of arcs in the graph which are symmetrical objects
- A_{total} is the total number of arcs in the graph
- M is number of axes of symmetry
- A_{di} is the number of double symmetrical arcs about axis i
- N_i is number of symmetrical nodes about axis i
- A_{si} is the number of single symmetrical arcs about axis i

It appears that this metric gives an appropriate measure of the perceived symmetry of a graph drawing [2].

Definitions

An *axis of symmetry* is defined as a straight line that has a minimum of three symmetrical objects (nodes (N), singly symmetrical arcs (A_s) and doubly symmetrical arcs (A_d)) mirrored on either side. The minimum of three symmetrical objects prevents an axis of symmetry being drawn at right angles, midway through every arc in a graph. Thus preventing every arc and pair of directly connected nodes contributing to symmetry, regardless of the overall appearance of the graph.

Around any axis of symmetry is a group of symmetrical objects. Collectively these objects make a *locality of symmetry*. While a *locality of symmetry* is defined by one axis of symmetry, the same group of symmetrical objects can have numerous axes. There is no restriction on the direction of these axes, but for each *locality of symmetry*, no two axes can be parallel. This prevents the same symmetrical relationship being counted over multiple axes.

The following objects are considered *symmetrical objects*:

- A double symmetrical arc (A_d) is an arc that is mirrored by another arc about the axis of symmetry.
- A single symmetrical arc (A_s) is a single arc that is bisected at right angles by the axis of symmetry.
- An axially symmetrical arc (A_a) is an arc that runs along the axis of symmetry.
- A symmetrical node (N) is a node that is mirrored by another node about the axis of symmetry.

The following three relationships are considered *symmetrical relationships*:

- Two mirrored double symmetrical arcs (A_d).
- Both halves of a single symmetrical arc (A_s).
- Two mirrored symmetrical nodes (N).

The final metric value is obtained by multiplying the number of symmetrical relationships in the graph by the *proportion of arcs which are symmetrical*. Thus, axially symmetrical arcs (A_a) contribute to this proportion even though they are not involved in a symmetrical relationship. If all the arcs in a graph are symmetrical objects then σ is simply the total number of symmetrical relationships over all axes of symmetry.

Acknowledgments

References

1. C. Batini, L. Furlani, and E. Nardelli. What is a good diagram? a pragmatic approach. In *Proc. 4th Int. Conf. on the Entity Relationship Approach*, 1985.
2. S. Bhanji, H.C. Purchase, R.F. Cohen, and M. James. Validating graph drawing aesthetics: A pilot study. Technical Report 336, University of Queensland Department of Computer Science, 1995.
3. C. Ding and P. Mateti. A framework for the automated drawing of data structure diagrams. *IEEE Transactions on Software Engineering*, SE-16(5):543–557, 1990.
4. P. Eades. A heuristic for graph drawing. *Congressus Numerantium*, 42:149–160, 1984.
5. D. Ferrari and L. Mezzalira. On drawing a graph with the minimum number of crossings. Technical Report 69-11, Istituto di Elettrotecnica ed Elettronica, Politecnico di Milano, 1969.
6. R. Gottsdanker. *Experimenting in Psychology*. Prentice-Hall, 1978.
7. R. Lipton, S. North, and J. Sandberg. A method for drawing graphs. In *Proc. ACM Symp. on Computational Geometry*, pages 153–160, 1985.
8. G.L. Lohse, K. Biolsi, N. Walker, and H.H. Rueter. A classification of visual representations. *Communications of the ACM*, 37(12):36–49, December 1994.
9. S. Siegel. *Nonparametric Statistics for the Behavioral Sciences*. McGraw-Hill, 1956.
10. R. Tamassia. On embedding a graph in the grid with the minimum number of bends. *SIAM J. Computing*, 16(3):421–444, 1987.
11. H. Trickey. Drag: A graph drawing system. In *Proc. Int. Conf. on Electronic Publishing*, pages 171–182. Cambridge University Press, 1988.
12. C. Ware, D. Hui, and G. Franck. Visualizing object oriented software in three dimensions. In *CASCON 1993 Proceedings*, 1993.