



A Language for Specifying Informational Graphics from First Principles

Citation

Shieber, Stuart M., and Wendy Lucas. 2007. A language for specifying informational graphics from first principles. In Software and Data Technologies Second International Conference; Revised Selected Papers/ ICSOFT/ENASE 2007, Barcelona, Spain, July 22 - 25, 2007, ed. Joaquim Filipe, 33-45. Berlin: Springer-Verlag. <http://dx.doi.org/10.1007/978-3-540-88655-6>.

Published Version

doi:10.1007/978-3-540-88655-6

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:4745847>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

A LANGUAGE FOR SPECIFYING INFORMATIONAL GRAPHICS FROM FIRST PRINCIPLES

Stuart M. Shieber

Division of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA
shieber@deas.harvard.edu

Wendy Lucas

Computer Information Systems Department, Bentley College, Waltham, MA, USA
wlucas@bentley.edu

Keywords: Visualization language, information visualization, graph specification, charting.

Abstract: Information visualization tools, such as commercial charting packages, provide a standard set of visualizations for tabular data, including bar charts, scatter plots, pie charts, and the like. For some combinations of data and task, these are suitable visualizations. For others, however, novel visualizations over multiple variables would be preferred but are unavailable in the fixed list of standard options. To allow for these cases, we introduce a declarative language for specifying visualizations on the basis of the first principles on which (a subset of) informational graphics are built. The functionality we aim to provide with this language is presented by way of example, from simple scatter plots to versions of two quite famous visualizations: Minard's depiction of troop strength during Napoleon's march on Moscow and a map of the early ARPAnet from the ancient history of the Internet. Benefits of our approach include flexibility and expressiveness for specifying a range of visualizations that cannot be rendered with standard commercial systems.

1 INTRODUCTION

Standard information-visualization tools provide a predefined set of visualizations for representing tabular, primarily numeric, data. While representations such as bar charts and pie charts may be well-suited for certain data visualization needs, they are unable to represent more complex relationships involving multiple variables. Representations of network activity or business performance measures, for example, can benefit from novel visualizations over multiple variables, which are unavailable in commercial charting packages. Indeed, quite famous visualizations like Minard's depiction of troop strength during Napoleon's march on Moscow or the maps of the early ARPAnet from the ancient history of the Internet¹ are well out of the scope of standard visualization packages, despite the fact that they are built on the basis of the same first principles from which more standard visualizations are derived.

To allow for such visualizations, we introduce a language for specifying informational graphics from

¹See Figures 3 and 6 for our versions of these visualizations.

first principles. One can view the goal of the present research as doing for information visualization what spreadsheet software did for business applications. Prior to Bricklin's VisiCalc, business applications were built separately from scratch. By identifying the first principles on which many of these applications were built — namely, arithmetic calculations over geographically defined values — the spreadsheet program made it possible for end users to generate their own novel business applications. This flexibility was obtained at the cost of requiring a more sophisticated user, but the additional layer of complexity can be hidden from naive users through prepackaged spreadsheets.

Similarly, we propose to allow direct access to the first principles on which (a subset of) informational graphics are built through an appropriate specification language. The advantages are again flexibility and expressiveness, with the same cost in terms of user sophistication and mitigation of this cost through prepackaging.

For expository reasons, the functionality that we aim for is first presented by way of example. We hope that the reader can easily extrapolate from the pro-

In general, some scaling of the data might be useful. We define a 2-D Cartesian frame to provide this scaling, using it instead of the canvas for placing the points (Figure 1(b)).

```
let frame:twodcart with
  frame.map(x,y) = Canvas(x + 10, y/2 + 20)
in
  {make p:point with
    p.location = frame.map(record.f, record.g)
  | record in SQL("select f, g from Table1")}
  make a:axis with
    a.scale = frame.map,
    a.aorigin = (50,50),
    a.ll = (0, 0),
    a.ur = (200,200),
    a.tick = (50,50);
```

Other graphical properties of the point can be tied to other fields. Here, we use a colorscale, a primitive for building linearly interpolated color scales. A legend for the color scale is positioned on the canvas as well. (See Figure 1(c).)

```
let frame:twodcart with
  frame.map(x,y) = Canvas(x + 10, y/2 + 20)
in
  let color:colorscale with
    color.min = ColorMap("red"),
    color.max = ColorMap("black"),
    color.minval = 0,
    color.maxval = 10
  in
    {make p:point with
      p.location = frame.map(rec.f, rec.g),
      p.color = color.scale(rec.h)
    | rec in SQL("select f, g, h from Table1")}
    make a:axis with
      a.scale = frame.map,
      a.aorigin = (50,50),
      a.ll = (0,0),
      a.ur = (200,200),
      a.tick = (50,50),
    make c:legend with
      c.scale = color,
      c.location = frame.map(200, 250);
```

The ability to add other types of constraints dramatically increases the flexibility of the system. For instance, stating that two values are approximately equal (\sim), instead of strictly equal ($=$), allows for approximate satisfaction of equality constraints. Further, constraints of non-overlapping (no) force values apart. Together, these constraints allow dither to be added to scatter plots.

```
let frame:twodcart with
  frame.map(x,y) = Canvas(x/2+1, y)
in
  let points = {make p:point with
    p.location ~ frame.map(record.f, record.g)
  | record in SQL("select f, g from Table1")}
  in
    no(points),
  make a:axis with
    a.scale = frame.map,
```

```
a.aorigin = (50,50),
a.ll = (0,0),
a.ur = (200,200),
a.tick = (50,50);
```

(Although this functionality is not available in the currently implemented system, the basis for its implementation has been provided in earlier work on network diagram layout (Ryall et al., 1997).)

A line chart can be generated using a line object instead of a point object. Suppose we take the records in Table 1(a) to be ordered, so that the lines should connect the points from the table in that order. Then a simple self-join provides the start points (x_1, y_1) and end points (x_2, y_2) for the lines. By specifying the appropriate query in SQL, we can build a line plot. The specification language also allows definitions of more abstract notions such as complex objects or groupings. We can use this facility to define line charts as an object that can be instantiated with the given data and frame:

```
define lchart:linechart with
  let frame:twodcart with
    frame.map = lchart.map
  in
    {make l:line with
      l.start = frame.map(record.x1, record.y1),
      l.end = frame.map(record.x2, record.y2)
    | record in lchart.recs}
  in
    make l:linechart with
      l.map(x,y) = Canvas(x + 10, y/2 + 20),
      l.recs = SQL("select tab1.f as x1,
        tab1.g as y1, tab2.f as x2,
        tab2.g as y2
      from Table1 as tab1, Table1 as tab2
      where tab2.recno = tab1.recno+1"),
    make a:axis with
      a.scale = l.map,
      a.aorigin = (50,50),
      a.ll = (0, 0),
      a.ur = (200,200),
      a.tick = (50,50);
```

All objects share this ability to have their graphical properties instantiated with specified data. This allows generated visualizations themselves, such as scatter plots, to be manipulated as graphical objects. For instance, it is possible to form a scatter plot of scatter plots. Figure 2 depicts such a graphic, as generated by the following specification and using the data found in Table 1(b).

```
define s:spot with
  let frame:twodcart with frame.map = s.map
  in
    {make p:point with
      p.location = frame.map(rec.x, rec.y)
    | rec in s.recs},
  make a:axis with
    a.scale = frame.map,
    a.aorigin = s.aorigin,
    a.ll = s.ll,
```

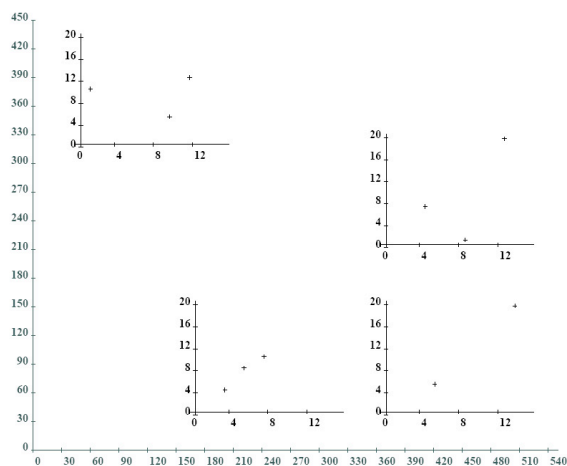


Figure 2: A scatter plot of scatter plots.

```

a.ur = s.ur,
a.tick = (4,4)
in
let outer:twodcart with
  outer.map(x,y) = Canvas(40*x-430, 35*y-380)
in
  let FrameRecs = SQL("select distinct
                        v1, v2 from matrix")
  in
    {make sp:spplot with
      sp.map(x,y) = outer.map(0.26*x +
        framerec.v1, 0.167*y + framerec.v2),
      sp.recs =
        SQL("select v3 as x,v4 as y from matrix
            where v1=framerec.v1
              and v2=framerec.v2"),
      sp.aorigin = (0,0),
      sp.ll = (-.5,-.5),
      sp.ur = (15,20)
    | framerec in FrameRecs};

```

3 A DETAILED EXAMPLE: THE MINARD GRAPHIC

As evidence of the flexibility of this language, we describe its use for specifying Minard's well known depiction of troop strength during Napoleon's march on Moscow. This graphic uses approximate geography to show the route of troop movements with line segments for the legs of the journey. Width of the lines is used for troop strength and color depicts direction. A parallel graphic shows temperature during the inbound portion of the route, again as a line chart. Our version of the graph is provided in Figure 3.

To generate this graphic, we require appropriate data tables: `marchNap` includes latitude and longitude at each way point, along with location name, direction, and troop strength; `marchTemp` includes latitude, longitude, and temperature for a subset of the inbound journey points, and `marchCity` provides latitude, lon-

gitude, and name for the traversed cities.

The main portion of the graphic is essentially a set of line plots, one for each branch of the march, where a branch is defined as a route taken in a single direction by a single division. Additional graphical properties (width and color) are tied to appropriate data fields. Textual labels for the cities are added using a text graphic object. A longitudinally aligned graph presents temperature on the main return branch.

The specification for this graphic (sans the temperature portion) is provided in Figure 4. After specifying some constants (lines 1–2), we define the depiction of a single branch of the march (6–16): a mapping (`m.map`) specifies a Cartesian frame (7–8) in which a line plot composed of a set of line segments (10–16) is placed, with one segment for each set of records (16). These records provide start and end points, along with widths at start and end (to be interpolated between), and color (11–15).

Thus, to depict a march leg, all that must be provided is the mapping and the set of records. These are constructed in lines 19–34. For each distinct direction and division (19), a separate march leg depiction is constructed (22–34). The mapping is a scaling of the Canvas frame (23), with records for the appropriate division and direction extracted from the `marchNap` database (24–34). Finally, the cities are labeled using the information in the `marchCity` database by tying coordinates of text labels to an offset of the latitude and longitude and setting a fixed color (37–42).

4 THE LANGUAGE

The examples from this paper were all generated from an implementation of our language. The implementation techniques are outlined in Section 5. The underlying ideas could, however, be implemented in other ways, for instance as a library of functions in a suitable functional language such as Haskell or ML.

The language is built around the specification of objects. The main construct is *objspec*, with a program defined as one or more *objspecs*. *Objspecs* are used for instantiating a graphical object of a predefined type, specifying relationships between the set of actual data values and their graphical representations, and defining new graphic types.

The `make` statement is used for instantiating a new instance of an existing object type (either built-in, such as `point` or `line`, or user-defined, such as `march`) with one or more conditions. There are two types of predefined objects: *generic* and *scale*. Generic objects provide visual representations of data values and include the primitive types of `point`, `cir-`

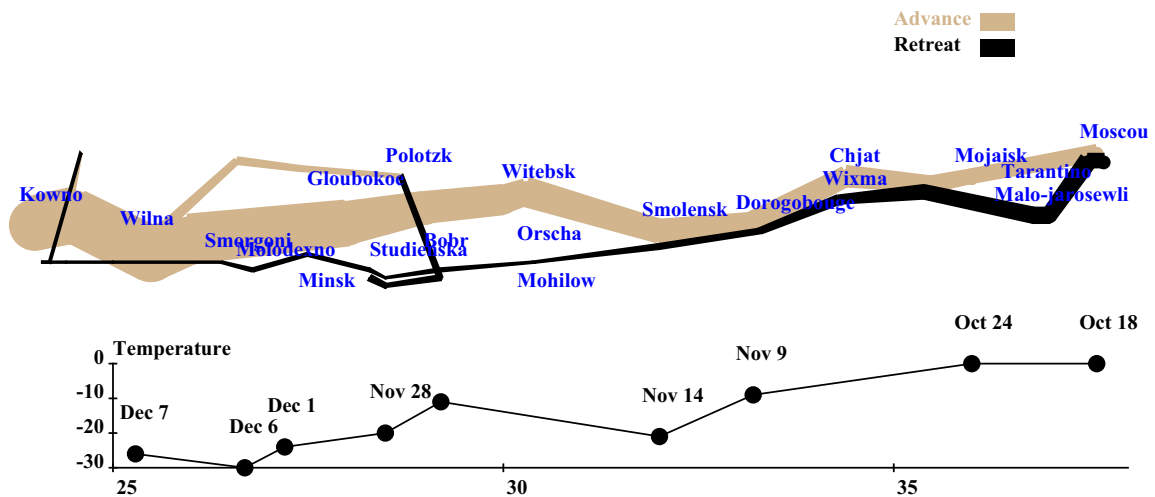


Figure 3: The graphic generated by a specification extended from that of Figure 4, depicting the troop strength during Napoleon's march on Moscow.

```

1  let scalefactor = 45 in
    let weight = .000002 in

    % Define the depiction of one (multi-leg) branch of the march identified
5  % by a distinct direction and division
    define m:march with
        let frame:twodcart with
            frame.map = m.map
        in
10     {make l:line with
            l.start = frame.map(rec.cx1, rec.cy1),
            l.end = frame.map(rec.cx2, rec.cy2),
            l.startWidth = scalefactor * weight * rec.r1,
            l.endWidth = scalefactor * weight * rec.r2,
15     l.color = ColorMap(rec.color)
        | rec in m.recs}
    in
    % Extract the set of branches that make up the full march
    let FrameRecs = SQL("select distinct direct, division from marchNap")
20  in
    % For each branch, depict it
    {make mp:march with
        mp.map(x,y) = Canvas(scalefactor*x - 1075, scalefactor*y - 2250),
        mp.recs = SQL("select marchNap1.lonp as cx1, marchNap1.latp as cy1,
25     marchNap2.lonp as cx2, marchNap2.latp as cy2,
        marchNap1.surviv as r1, marchNap2.surviv as r2,
        marchNap1.color as color
        from marchNap as marchNap1, marchNap as marchNap2
        where marchNap1.direct = framerec.direct
30     and marchNap2.direct = marchNap1.direct
        and marchNap1.division = framerec.division
        and marchNap1.division = marchNap2.division
        and marchNap2.recno = marchNap1.recno+1")
        | framerec in FrameRecs}
35
    % Label the cities along the march
    let labeloffset = .2 in
    {make d2:label with
        d2.location = frameTemp.map(recc.lonc - labeloffset, recc.latc + labeloffset),
40     d2.label = recc.city,
        d2.color = ColorMap("blue")
    | recc in SQL("select lonc, latc, city from marchCity")}

```

Figure 4: Partial specification of the Minard graphic depicted in Figure 3. This specification does not include the temperature plot.

cle, line, oval, rectangle, polygon, polar segment, and labels. Scales are used for mapping data values to their visual representations and are graphically represented by axes and legends. A scale is associated with a coordinate system object and defines a transformation from the default layout canvas to a frame of type `twodcart` or `twodpolar` (at this time).

In addition to a unique identifier, each object type has predefined attributes, with conditions expressed as constraints on these attributes. Constraints can be of type equality (`'='`) or approximate equality (`'~'`). Constraints enforcing visual organization features (Kosak et al., 1994) such as non-overlap, alignment, or symmetry, can be applied to a set of graphical objects. These types of constraints are particularly useful in specifying network diagrams (Ryall et al., 1997). While such constraints can be specified in the language, our current implementation only supports constraints in the form of equalities.

All of the generic objects have a `color` attribute and at least one `location` attribute, represented by a coordinate. For 2-D objects, a Boolean `fill` attribute defaults to `true`, indicating that the interior of that object be filled in with the specified color. This attribute also applies to line objects, which have start and end `widths`. When `widths` are specified, lines are rendered as four-sided polygons with rounded ends.

Each type of scale object has its own set of attributes. A coordinate system object's `parent` attribute can reference another coordinate system object or the canvas itself. The mapping from data values to graphical values can be specified by conditions on its `origin` and `unit` attributes. Alternatively, a condition can be applied to its `map` attribute in the form of a user-defined mapping function that denotes both the parent object and the scale. Thus, a `twodcart` object whose `parent` is `Canvas`, `origin` is (5, 10), and `unit` is (30, 40) can be defined by the mapping function: `Canvas.map(30x + 5, 40y + 10)`.

Axes are defined in terms of their `origin`, `ll` (lower left) and `ur` (upper right) coordinates, and `tick` marks. Legends can be used to graphically associate a color gradient with a range of data values by assigning a color scale object to the `scale` attribute and a coordinate to the `location` attribute. Discrete colors can also be associated with data values, as in the Minard Graph example, where tan represents "Advance" and black represents "Retreat."

Attributes can also be defined by the user within a `make` statement. For example, it is often helpful to declare temporary variables for storing data retrieved from a database. These user-defined attributes are ignored when the specified visualization is rendered.

Another construct for an *objspec* is the `type` state-

ment. This is used for defining a new object *type* that satisfies a set of conditions. These conditions are either constraints on attribute properties or other *objspecs*. An example of the former would be a condition requiring that the color attribute for a new chart type be "blue." An example of the latter is to require that a chart include a 2-D Cartesian coordinate frame for use in rendering its display and that this frame contain a set of lines and points corresponding to data retrieved from a database.

A third type of *objspec* is a set of objects, a *setspec*. This takes the form of either a query string or a set of *objspecs* to be instantiated for each record retrieved by a query string. These two constructs are often used in conjunction with one another to associate data with generic graph objects. For example, a *setspec* may define a query that retrieves two values, `x` and `y`, from a database table. A second *setspec* can then specify a `make` command for rendering a set of points located at the `x` and `y` values retrieved by the query. Alternatively, the query can be contained within the `make` statement itself, as shown in the code for the scatter plots in Section 2.

It is often useful to associate a variable name with an *objspec* so that it can be referenced later. This is accomplished with the `let` statement, which provides access to that variable within the body of the statement. Because this construct is commonly used in conjunction with a `make` or `type` statement, two shorthand expressions have been provided. The first of these is an abbreviated form of the `let` statement in which the `make` clause is not explicitly stated, with: `let var:type` with *conditions* in *body* corresponding to `let var = make var:type` with *conditions* in *body*. Similarly, the `define` statement associates a new object definition with a variable name, where: `define var:type` with *conditions* in *body* expands to `let var = type var:type` with *conditions* in *body*.

Figure 5 demonstrates the usage of the above *objspec* constructs for defining a new type of object called a `netplot`. The code specifies the creation of a set of circles, referred to as nodes, with radius of 5, color of "black," and center locations corresponding to `x` and `y` values queried from a database and mapped to a frame. (Data values were estimated from a depiction of the ARPAnet circa 1971 available at <http://www.cybergeography.org/atlas/historical.html>.) A set of red lines are also specified, with the endpoint locations of the lines equal to the centers of the circles.

The output generated from this specification is shown in Figure 6. This chart and the Minard graphic

```

define n:netplot with
  let frame:twodcart with frame.map = n.map
  in
    let nodes =
      {make c:circle with
        c.center = frame.map(rec.x, rec.y),
        c.radius = n.radius,
        c.color = n.ccolor
        | rec in n.netRecs}
    in
      {make l:line with
        l.start = nodes[rec.v1].center,
        l.end = nodes[rec.v2].center,
        l.color = n.lcolor
        | rec in n.edgeRecs},
      {make d:label with
        d.location = frame.map(rec.x + 2,
                               rec.y - 1),

        d.label = rec.node,
        d.color = n.ccolor
        | rec in n.netRecs}
    in
      make net:netplot with
        net.map(x,y) = Canvas(7*x, 7*y),
        net.netRecs = SQL("select node, x, y
                          from ArpaNodes"),
        net.radius = 5,
        net.edgeRecs = SQL("select v1, v2
                           from ArpaLinks"),
        net.lcolor = ColorMap("red"),
        net.ccolor = ColorMap("black");

```

Figure 5: Specification of the ARPAnet graphic depicted in Figure 6.

from the prior section demonstrate the flexibility and control for specifying visualizations that result from working with the right set of first principles. Further, the intricacies of the code required to generate these charts are hidden behind a relatively simple to use but powerful grammar.

5 IMPLEMENTATION

The output generation process that renders visualizations from code written in our language, such as those shown in the prior sections, involves four stages.

1. The code is parsed into an abstract syntax tree representation.
2. The tree is traversed to set variable bindings and to evaluate *objspects* to the objects and sets of objects that they specify. These objects are collected, and the constraints that are imposed on them are accumulated as well.
3. The constraints are solved. In the current implementation, the restriction to strict equality constraints means that solution is trivial, essentially by congruence closure. We plan to solve the more

expressive constraints as in related work on the GLIDE system (Ryall et al., 1997), by reduction to mass-spring systems and iterative relaxation, together with user advice to eliminate local minima.

4. The objects are rendered. Any primitive graphic object contained in the collection is drawn to the canvas using graphical values determined as a result of constraint satisfaction.

6 RELATED WORK

Standard charting software packages, such as Microsoft Chart or DeltaGraph, enable the generation of predefined graphic layouts selected from a “gallery” of options. As argued above, by providing pre-specified graph types, they provide simplicity at the cost of the expressivity that motivates our work.

More flexibility can be achieved by embedding such gallery-based systems inside of programming languages to enable program-specified data manipulations and complex graphical composites. Many systems have this capability: Mathematica, Matlab, Igor, and so forth. Another method for expanding expressiveness is to embed the graph generation inside of a full object-drawing package. Then, arbitrary additions and modifications can be made to the generated graphics by manual direct manipulation. Neither of these methods extend expressivity by deconstructing the graphics into their abstract information-bearing parts, that is, by allowing specification from first principles.

Towards this latter goal, discovery of the first principles of informational graphics was pioneered by Bertin, whose work on the semiology of graphics (Bertin, 1983) provides a deep analysis of the principles. It is not, however, formalized in such a way that computer implementation is possible.

Computer systems for automated generation of graphics necessarily require those graphics to be built from a set of components whose function can be formally reasoned about. The seminal work in this area was by Mackinlay (Mackinlay, 1986), whose system could generate a variety of standard informational graphics. The range of generable graphics was extended by Roth and Mattis (Roth and Mattis, 1990) in his SAGE system. These systems were designed for automated generation of appropriate informational graphics from raw data, rather than for user-specified visualization of the data. The emphasis is thus on the functional appropriateness of the generated graphic rather than the expressiveness of the range of generable graphics.

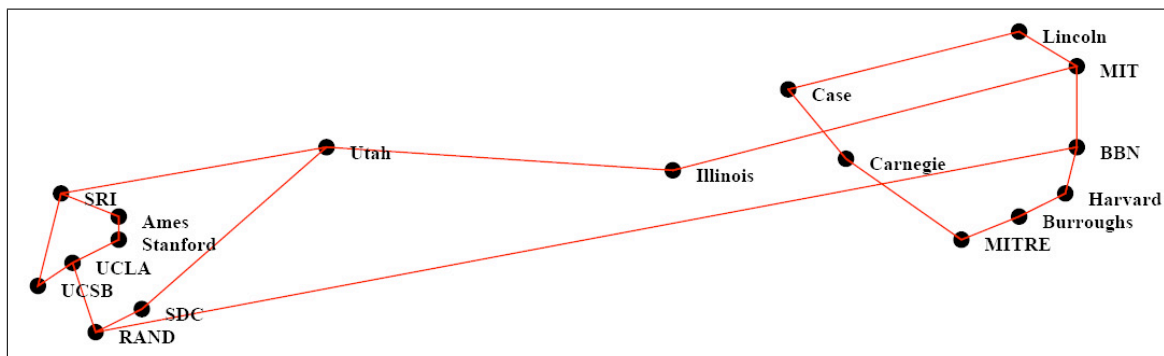


Figure 6: A network diagram depicting the ARPANet circa 1971.

The SAGE system serves as the basis for a user-manipulable set of tools for generating informational graphics, SageTools (Roth et al., 1995). This system shares with the present one the tying of graphical properties of objects to data values. Unlike SageTools, the present system relies solely on this idea, which is made possible by the embedding of this primitive principle in a specification language and the broadening of the set of object types to which the principle can be applied.

The effort most similar to the one described here is Wilkinson’s work on a specification language for informational graphics from first principles, a “grammar” of graphics (Wilkinson, 1999). Wilkinson’s system differs from the one proposed here in three ways. First, the level of the language primitives are considerably higher; notions such as Voronoi tessellation or vane glyphs serve as primitives in the system. Second, the goal of his language is to explicate the semantics of graphics, not to serve as a command language for generating the graphics. Thus, many of the details of rendering can be glossed over in the system. Lastly, and most importantly, the ability to embed constraints beyond those of equality will provide us the capacity to generate a range of informational graphics that use positional information in a much looser and more nuanced way. This latter advantage will distinguish the present work from all previous work in the area with the exception of the GLIDE system for network diagram drawing. We hold for future work the implementation of a GLIDE-like facility for satisfying the generated constraints.

7 CONCLUSIONS

We have presented a specification language for describing informational graphics from first principles, founded on the simple idea of instantiating the graph-

ical properties of generic graphical objects from constraints over the scaled interpretation of data values. This idea serves as a foundation for a wide variety of graphics, well beyond the typical sorts found in systems based on fixed galleries of charts or graphs. By making graphical first principles available to users, our approach provides flexibility and expressiveness for specifying innovative visualizations.

REFERENCES

- Bertin, J. (1983). *Semiology of Graphics*. University of Wisconsin Press.
- Kosak, C., Marks, J., and Shieber, S. (1994). Automating the layout of network diagrams with specified visual organization. *Transactions on Systems, Man and Cybernetics*, 24(3):440–454.
- Mackinlay, J. (1986). Automating the design of graphical presentations of relational information. *ACM Transactions on Graphics*, 5(2):110–141.
- Roth, S. F., Kolojchick, J., Mattis, J., and Chuah, M. C. (1995). Sagetools: An intelligent environment for sketching, browsing, and customizing data-graphics. In *CHI '95: Conference companion on Human factors in computing systems*, pages 409–410, New York, NY, USA. ACM Press.
- Roth, S. F. and Mattis, J. (1990). Data characterization for graphic presentation. In *Proceedings of the Computer-Human Interaction Conference (CHI '90)*.
- Ryall, K., Marks, J., and Shieber, S. M. (1997). An interactive constraint-based system for drawing graphs. In *Proceedings of the 10th Annual Symposium on User Interface Software and Technology (UIST)*.
- Wilkinson, L. (1999). *The Grammar of Graphics*. Springer-Verlag, New York, NY.