View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by NASA Technical Reports Server

N94-24187

PROCESS AND REPRESENTATION IN GRAPHICAL DISPLAYS

Douglas J. Gillan
Lockheed Engineering and Sciences Company

Robert Lewis
Rice University

Marianne Rudisill
NASA Lyndon B. Johnson Space Center

INTRODUCTION

To survive and succeed in the world, people have to comprehend both diverse natural sources of information, such as landscapes, weather conditions, and animal sounds, and human-created information artifacts such as pictorial representations (i.e., graphics) and text. Researchers have developed theories and models that describe how people comprehend text (for example, see [8]), but have largely ignored graphics. However, an increasing amount of information is provided to people by means of graphics, as can be seen in any newspaper or news magazine, on television programs, in scientific journals and, especially, on computer displays.

Our initial model of graphic comprehension has focused on statistical graphs for three reasons: (1) recent work by statisticians which provides guidelines for producing statistical graphs (Bertin [2], Cleveland and McGill [4,5] and Tufte [10]) could be translated into preliminary versions of comprehension models, (2) statistical graphs play an important role in two key areas of the human-computer interface — direct manipulation interfaces (see [7] for a review) and task-specific tools for presenting information, e.g., statistical graphics packages, and (3) computer-displayed graphs will be crucial for a variety of tasks for the Space Station *Freedom* and future advanced spacecraft. Like other models of human-computer interaction (see [3] for example), models of graphical comprehension can be used by human-computer interface designers and developers to create interfaces that present information in an efficient and usable manner.

Our investigation of graph comprehension addresses two primary questions — how do people represent the information contained in a data graph and how do they process information from the graph? The topics of focus for graphic representation concern the features into which people decompose a graph and the representation of the graph in memory. The issue of processing can be further analyzed as two questions, what overall processing strategies do people use and what are the specific processing skills required?

GRAPHIC REPRESENTATION

FEATURES OF GRAPHIC DISPLAYS

Both Bertin [2] and Tufte [10] address the features underlying the perception and use of graphs. Bertin [2] focuses on three constructs, (1) "implantation," i.e., the variation in the spatial dimensions of the graphic plane as a point, line, or area; (2) "elevation," i.e., variation in the spatial dimensions of the graphical element's qualities — size, value, texture, color orientation, or shape; and (3) "imposition," i.e. how information is represented, as in a statistical graph, a network, a geographic map, or a symbol. Tufte [10] proposes two features as important for graphic construction, data ink and data density. Tufte describes data ink as "the nonerasable core of a graphic" [10, p. 93] and provides a measure, the data-ink ratio, which is the "proportion of a graphic's ink devoted to the nonredundant display of data information" [10, p. 93]. Data density is the ratio of the number of data points and the areas of the graphic. Tufte's guidelines call for maximizing both the data-ink ratio and, within reason, the data density, in other words, displaying graphics with as much information and as little ink as possible.

Both Bertin's and Tufte's ideas about the features of data graphs were derived from their experience as statisticians, rather than from experimental evidence. We decided to fill the empirical void concerning the features underlying graphic comprehension. In our first experiment, people simply judged the similarity in appearance and information displayed by all possible pairs of 17 different types of graphs (that is, 136 pairs of graphs). The graphs ranged from the familiar (line graphs, bar graphs, and scatter plots) to the more unusual (star graphs, ray graphs, and stick man graphs). The similarity judgments were analyzed with multivariate statistical techniques, including (1) cluster analysis, which shows the groupings or categories (clusters) that underlie people's judgments about a set of objects and (2) multidimensional scaling (MDS), which shows the linear dimensions underlying people's similarity judgments. The logic of these analyses was that people would cluster graphs and place graphs along dimensions based on the features of the graph [9].

The cluster analyses indicated that people group graphs, at least in part, according to the physical elements of the graphs. Key clusters include graphs in which points were the dominant element (the two types of scatter plots, the range and density graphs), graphs consisting of straight lines (the surface, textured surface, and stacked bar graph), and those consisting of solid areas (the column and bar graphs). The categorization of the graphs according to physical elements agrees generally with Bertin's [2] construct of implantation.

The MDS analyses of the similarity judgments were combined with a factor analysis which resulted in three factors, each consisting of one informational dimension and one perceptual dimension, which accounted for 97% of the data. One factor differentiated perceptually simple graphs (e.g., the bar and line graphs) from perceptually complex graphs (the scatter plots, the 3-dimensional graph, and the surface graphs). A second factor separated graphs for which axes were unnecessary to read the graph (the pie, star, 3-dimensional, and stick man graphs) from those for which the axis contained information (especially the modified scatter plots — the range and density graphs [10]). Finally, the third factor tended to have informationally complex graphs (those with the most data) at one end and informationally simple graphs (those with the least data) at the other end. Accordingly, we hypothesize that people decompose a graph according to its perceptual complexity, figure-to-axes relation, and informational complexity. A subsequent experiment has shown that each of these factors relates to peoples' speed and accuracy in answering questions using these graphs [6].

REPRESENTATION IN MEMORY

The previous section of this paper addressed the features present when a user looks at a graphic. This section addresses the features that the user walks away with. Accordingly, the experiments looked at how a user represents the information from a graphic in memory.

Our research on memorial representation of graphics involved a simple experimental design: Our subjects worked with a set of graphs on one day, then we assessed what they retained about the graphic on a second day. The initial training day consisted of one trial with each of six different graphs during a 30 second trial. For three graphs, the subjects answered questions about the graphs, (e.g., What is the mean of the variables in the graph? and Which has the greater value, variable A or variable B?). For the other three graphs, they identified and drew the perceptual components of the graph, each component in a separate box. For example, in a line graph a subject might draw the points representing each variable, the lines connecting the points, the axes, verbal labels, and numerical labels.

Twenty-four hours after training, we tested the subjects using two different methods. We gave one group of 16 subjects a recognition test in which they looked at 24 different graphs and had to say whether they had seen precisely that graph during the training session. We constructed the 24 test graphs systematically. Each of the six graphs from the training session were presented during the test. Each training graph had three "offspring" that served as the distractors (or incorrect test stimuli) during the test. One type of distractor contained the same data as the training stimulus, but used a different graph type to display the data (New Graph-Same Data); a second distractor displayed the data using the same type of graph (Same Graph-New Data); the third distractor differed from the training graph in both graph type and data (New Graph-New Data). Perfect recognition would have resulted in 100% yes answers to the training graphs and 0% yes answers to the distractors. A second group of 14 subjects received a recall test in which they were asked to draw the graphs from Day 1 in as much detail as they could remember.

The results showed that people's recognition of the training graphs was very good. They correctly recognized the training graph 88% of the time, with little difference between the graphs used during training in the perceptual task (85% recognition) and those used in the informational task (90% recognition). Although false recognitions of the distractors were low overall (10% yes answers to distractors), the distribution of false recognitions was interesting. Of the 39 false recognitions by the 16 subjects, 29 (74%) were made to the Same Graph-New Data distractor. Friedman test chi-square (2 df) = 10.1, $p < .05$. The high false recognition rate when the same graph type was used (30% false recognitions to that distractor) suggest that the perceptual type of the graph has a strong representation in memory. We found that both training with an informational task and training with a perceptual task yielded similar high proportions of the total false recognitions for the Same Graph-New Data distractor, 77% and 70%, respectively.

The results from the recall test provide even greater support for the hypothesis that the representation of the graph type and certain perceptual features was exceptionally strong. Subjects had good recall for the graph type (71% of the graphs), the presence or absence of axes (71% correct recall of axes), and the perceptual elements (lines, areas, and points) in the graphs (53% correct recall of graph elements). In contrast, recall of information from the graphs was generally poor. For example, subjects had low recall rates for the number of data points in the graph (29% correct recall), the quantitative labels on the axes (10% of the labels), and the verbal labels of the axes and data points (12% of verbal labels). They recalled the correct spatial relations between data points only 22% of the time. In addition to showing the strength of the perceptual representation, these data suggest that the perceptual and informational representations of a graph are independent.

STRATEGIES FOR PROCESSING INFORMATION

Based on formal thinking-aloud protocols, as well as informal discussions with users, we have hypothesized that people use two different types of strategies when processing information from a data graph - an arithmetic, look-up strategy and a perceptual, spatial strategy. With the arithmetic strategy, a user treats a graph in much the same way as a table, using the graph to locate variables and look up their values, then performing the required arithmetic manipulations on those variables. In contrast, the perceptual strategy makes use of the unique spatial characteristics of the graph, comparing the relative location of data points.

We have hypothesized that users apply the strategies as a function of the task. Certain tasks appear to lend themselves better to one strategy than another. Answering a comparison question like "Which is greater, variable A or B?" would probably be answered rapidly and with high accuracy by comparing the spatial location of A and B. In contrast, a user answering the question "What is the difference between variables A and B?" about a line graph might be able to apply the perceptual strategy, but would be able to determine the answer more easily and accurately with the arithmetic strategy. In addition, we propose

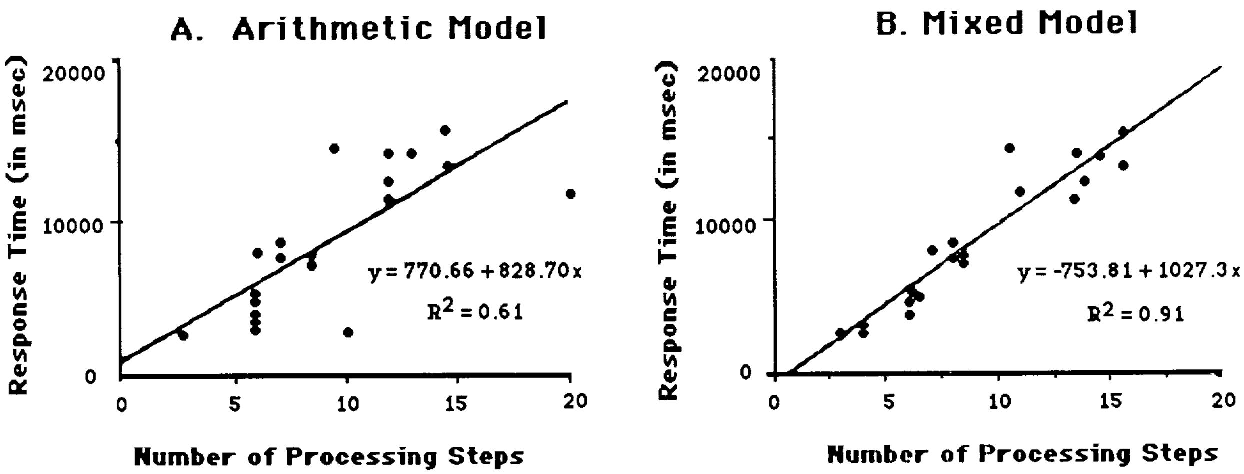

Figure 1. Response times for answering eight types of questions using three types of graphs as a function of the number of processing steps. A. Arithmetic strategy B. Mixed arithmetic-perceptual strategy.

that users vary their strategy according to the characteristics of the graph. For example, if a user were faced with a graph that had inadequate numerical labels on the axes, he or she would be forced to use the perceptual strategy to the greatest extent possible.

We have run a series of experiments to test our hypotheses about graphic processing strategies. The response time data from these experiments are consistent with a model that suggests that users tend to apply the arithmetic strategy, but will shift to the perceptual strategy under certain conditions. In the basic experiment, subjects used three types of graphs — scatter plot, a line graph, and a stacked bar graph. They were asked eight types of questions about each graph type: (1) identification — what is the value of variable A? (2) comparison — which is greater A or B? (3) addition of two numbers — A+B. (4) subtraction — A-B, (5) division — A/B, (6) mean — (A+B+C+D+E)/5, (7) addition and division by 5 — (A+B)/5, and (8) addition of three numbers A+B+C. Subjects were instructed to be as fast and accurate as possible. We predicted that the subjects' time to answer the questions using a graph would be a function of the number of processing steps required by a given strategy. Accordingly, with the arithmetic strategy, determining the mean should take longer than adding three numbers, which should take longer than adding two numbers.

We began by fitting the data to a model based on the assumption that subjects used an arithmetic strategy for all questions with all graphs. Figure 1A shows the fit of that model to the response time data. The response time generally increases as the number of processing steps increases, so the model accounts for some of the variance, 61%, but many of the data points fall far from the regression line. This model is poorest at predicting performance on two trials with the stacked bar graph — the mean and the addition of two numbers — and for the comparison trials with all three types of graphs; subjects responded on the comparison trials and the mean trial more quickly than predicted.

As discussed above, a comparison appears to be a likely task for subjects to use a perceptual strategy. In addition, the stacked bar graph intrinsically lends itself to adding the five variables by a perceptual strategy. The total height of the stack represents the cumulative value of the five variables. Accordingly, for model 2, we assumed that subjects used a perceptual strategy to determine the cumulative value of the stacked bar graph (then looked up the value and divided by 5 arithmetically) and used only the perceptual strategy to make all comparisons. Figure 1B shows how a version of that model fits the data. This model captures a substantially greater amount of the variance, 91%, than did Model 1. In this version of the model, the regression function slope suggests that each processing step required about 1 second to complete, except for steps requiring subtraction or division (which the model assumes took 1.5 and 2 seconds, respectively).

The fit of the mixed arithmetic-perceptual model to the data, together with subjects' verbal protocols when answering questions using graphs, support our hypotheses: (1) that people use both arithmetic and perceptual strategies with graphics, (2) that for many typical questions, the bias appears to be for the arithmetic strategy (perhaps because of the greater accuracy with that strategy), and (3) subjects switch strategies as a function of the characteristics of the question and graph.

A THEORY OF GRAPHIC COMPREHENSION

The focus of the rest of this paper is on an overall theory of graphical comprehension designed to help in the development of graphic displays. The theory covers the entire process of graphic comprehension from the motivation to look at a graph, to the use of the graph, to remembering the graph.

In general, when I look at a graph, I have a particular purpose in mind — I am usually trying to answer a specific question. Thus, stage 1 in graphic comprehension would consist of either forming a representation of the question to be answered (if the question had to be remembered), or producing the question by inference or generalization. The final cognitive representation of the question would probably be much the same, regardless of whether I read it, remembered it, or generated it. The likely representational format for the question would be a semantic network (e.g., [1] and [8]). Determining the answer to the question would function as the goal of my graphic comprehension.

At the start of the second stage in graphic comprehension, I would look at the graph. On looking at the graph, I would encode the primary global features — the presence or absence of the axes and the type of graph. These would be encoded in a format that would permit reproduction of certain lower level features, such as the orientation of both the elements that make up the graph type and the axes. For example, subjects in our representation experiments generally recalled the horizontal orientation of the bars in a column graph, despite (or, perhaps, because of) their difference from the more typical vertical bar graphs. Interestingly, features that one might expect to be important to a graph user, such as the number of data points, appear not to be encoded as part of this global encoding stage. One hypothesis of this model is that features represented during the global encoding stage receive the bulk of the representational strength. That is to say, they will be the best remembered.

The third stage in graphic comprehension is to use the goal and the global features of the graph to select a processing strategy. If my goal were to compare the value of variables or (possibly) to compare a trend, I would select a perceptual strategy. If my goal were to determine the sum of four variables, and numbered axes were present and the graph type supported it (e.g., a line graph or a bar graph), then I would select the arithmetic strategy.

During the next stage, I would implement the processing steps called for in the strategy determined in the third stage. For example, adding variables A and B from a line graph would involve the following processing steps: (1) locate the name of variable A on the X axis, (2) locate variable A in the x-y coordinate space of the body of the graph, (3) locate the value of variable A on the Y axis and store in working memory, (4) locate the name of variable B on the X axis, (5) locate variable B in the x-y coordinate space of the body of the graph, (6) locate the value of variable B on the Y axis and store in working memory, and (7) add the value of variable A to the value of variable B to produce the value "sum."

Because the semantic and quantitative information (i.e., the variable names and values, respectively) are processed to some extent during this phase, some of that information will be represented, but, as our recall data suggest, not strongly. As a final stage in graphic comprehension, I would examine the result from processing step 7, the "sum," to determine if it plausibly met the goal set in comprehension stage 1. If the response was a plausible fit with the goal, I would incorporate the answer into the semantic network that represented the goal.

This theory directs both future research in graphics and the design of graphical computer interfaces. For example, future research will be needed to determine specific processing models for different questions using the perceptual strategy. In addition, predictions about the memory for quantitative and semantic information in a graph need to be tested. Finally, many of the design principles derived from the theory are concerned with the complex relations between the task (or goal), the characteristics of the graphical display, and the processing strategies. For example, if a subject is likely to use arithmetic strategy (e.g., with an addition or subtraction question), the axes should be numbered with sufficient numerical resolution. The graph type should allow the user to read a variable's value directly from the axis and should not require multiple computations to determine a variable's value (as a stacked bar graph does). One of our long-term goals is to produce a model of graphic comprehension that is sufficiently elaborate to allow us to build tools to aid in the design of graphical interfaces.

ACKNOWLEDGEMENTS

The authors would like to thank Sarah Breedin and Tim McKay for helpful comments on this paper and, along with Kritina Holden and Susan Adam, for helpful discussions about graphics.

REFERENCES

1. Anderson, J. R. *The Architecture of Cognition*, Harvard University Press, Cambridge, MA, 1983.

2. Bertin, J. *Seminology of Graphics* (translated by W. J. Berg), University of Wisconsin Press, Madison, WI, 1983.

3. Card, S. K., Moran, T. P., and Newell, A. *The Psychology of Human-Computer Interaction*, L. Erlbaum Associates, Hillsdale, NJ, 1983.

4. Cleveland, W. S., and McGill, R. "Graphical perception: Theory, experimentation, and application to the development of graphical methods," *Journal of the American Statistical Association*, 79, 1984, 531-554.

5. Cleveland, W. S., and McGill, R. "Graphical perception and graphical methods for analyzing and presenting scientific data," *Science*, 229, 1985, 828-833.

6. Gillan, D. J., Lewis, R., and Rudisill, M. "Models of user interaction with graphical interfaces: I. Statistical graphs," *Proceedings of CHI*, 1989, ACM, New York, 375-380.

7. Hutchins, E. L., Hollan, J. D., and Norman, D. A. "Direct manipulation interfaces," *User-System Centered Design*, L. Erlbaum Associates, Hillsdale, NJ, 1986, 87-124.

8. Kintsch, W., and van Dijk, T. A. "Toward a model of text comprehension and production," *Psychological Review*, 1978, 85, 363-394.

9. Shepherd, R. N., Romney, A. K., and Nerlove, S. B. *Multidimensional Scaling: Theory and Application in the Behavioral Sciences*, Seminar Press, New York, 1972.

10. Tufte, E. R. *The Visual Display of Quantitative Information, Graphics Press*, Cheshire, CT, 1983.