Saying It in Graphics: from Intentions to Visualizations

Stephan Kerpedjiev♦, Giuseppe Carenini♣, Nancy Green♦, Johanna Moore♣, Steven Roth♦

♦ Carnegie Mellon University
Pittsburgh, PA 15213
kerpedjiev, ngreen, roth@cs.cmu.edu

♣ University of Pittsburgh
Pittsburgh, PA 15213
carenini, jmoore@cs.pitt.edu

Abstract

We propose a methodology for automatically realizing communicative goals in graphics. It features a task model that mediates the communicative intent and the selection of graphical techniques. The methodology supports the following functions: isolating assertions presentable in graphics; mapping such assertions into tasks for the potential reader, and selecting graphical techniques that support those tasks. We illustrate the methodology by redesigning a textual argument into a multimedia one with the same rhetorical and content structures but employing graphics to achieve some of the intentions.

1. Introduction

Information visualizations have been studied mainly as a tool for *data exploration*. Their role in a data exploration problem is to provide insight into features of the data that are hard to reveal using other representations. Visualizations however are also used for *communication* [1], a function that has received little attention from the computational point of view. In a communicative situation, such as an advertisement or a testimony, the goal is to make certain points believable, possibly supporting them with data. This is where graphics can be useful. It is a powerful vehicle for presenting the data in such a way that certain relationships are clear to the audience. To do this, however, the graphic designer must be able to understand the intent of the author and to apply techniques that convey this intent to the intended audience.

We propose that a task model mediate the intentional representation and the selection of graphical techniques. Prior work [5, 6] has shown that viewers interpret graphics by performing certain perceptual and cognitive tasks (operations) such as search, recognition and inference. The task model spells out procedures of such operations that viewers should perform on a graphic to achieve given communicative goals. Our methodology of *conceptual tasks* as an abstraction of perceptual and cognitive tasks is analogous to Casner's [2] methodology of designing visualizations that support data exploration. Using an intermediate task representation is advantageous in the following respects:

- It allows us to base our generation techniques on empirical, graphics interpretation research [6].
- It supports the reuse of task-based graphic design knowledge accumulated by other researchers [2, 9].
- It enables generation systems to respond intelligently when the user fails to achieve a communicative goal [8] by providing help in the form of a procedure for interpreting the graphic.

We begin by describing a media-independent representation of what needs to be conveyed (content) and why (intention). Next, we consider the problem of *media allocation*, i.e. the decision whether to use text and/or graphics to achieve a given intention. Then we describe how to transform the intentions to a *task representation* that can guide a graphic designer. Finally, we outline *graphical techniques* that support conceptual tasks.

To illustrate our work, we redesign the following paragraph from Bill Gates' testimony at a Senate hearing on March 3, 1998[1].

"Microsoft is often referred to as a "software giant." The facts, however, tell a different story. While Microsoft is clearly a leader in the computer software industry, our revenues account for less than 5% of total worldwide software revenues of $253 billion and only 1% of the information technology industry's collective revenues of $1 trillion. More than a dozen companies, including industry leaders such as IBM, Hitachi, Computer Associates, Oracle, Digital Equipment, Novell, Sybase and Sun Microsystems have more than $1 billion in annual software revenues alone. IBM's software revenues

---

[1] The testimony as distributed by the news agencies had no visualizations.

of $13 billion in 1997, are about the same as Microsoft's. And revenues for many of these companies have soared in recent years. (For example, Oracle's revenues rose from $1.2 billion in 1993 to $5.7 billion in 1997; over the same period, Sun's revenues rose from $4.3 billion to $8.6 billion.)"

2. Intention

According to the theory of speech acts [11], speakers produce utterances to affect the mental state of the hearers. This implies that before producing an utterance, the speaker has specific intentions such as to recommend an action or to convince the hearer that some proposition is true. The means of achieving those intentions in text are provided by the linguistic system: words to activate concepts and predicates, grammars to express propositions by clauses and sentences, and rhetorical relations to build arguments. Prior research in natural language generation has shown how text can be generated to achieve specific intentions [8]. We use a representation of the speaker's intentions as a starting point in the process of generating communicative visualizations.

Figure 1 shows a simplified version of the rhetorical structure [7] of the sample text. This structure can also be regarded as a communicative plan the nodes of which contain propositions that the reader should believe after reading the corresponding spans of text. The italic labels attached to the links represent the rhetorical relations used to achieve the effects at the parent nodes. The propositions at the leaves of the tree are directly asserted in the text.

The main point of the text is that Microsoft is not a software giant. To increase the credibility of this point,

the author supplies *evidence* that can be summarized by the following propositions:

- the percentage of the revenues of Microsoft in the relevant industries is small;
- there are other companies whose software revenues are of the same order as the revenues of Microsoft;
- many of these companies are experiencing rapid growth.

The content of the intentions at all levels of the communicative plan is represented by propositions [3]. Each proposition consists of a predicate and one or more arguments, which are constants, individual entities, or descriptions of objects.

In summary, the communicative plan spans several levels - from rhetorical relations to assertions to descriptions of individual elements, and each of these levels will contribute to the decision-making process of graphics generation.

3. Media allocation

Logically the first decision that needs to be made with respect to a given part of the plan is whether to realize it in text and/or in graphics. Graphics are often superior to text for showing quantitative relationships, coordinates (e.g., time and location), and multiple homogeneous facts [10]. If some part of the communicative plan relies on these types of assertions, the system might consider using graphics. This way the facts are likely to be assimilated more directly, bypassing cognitively slow memory-search operations [5].
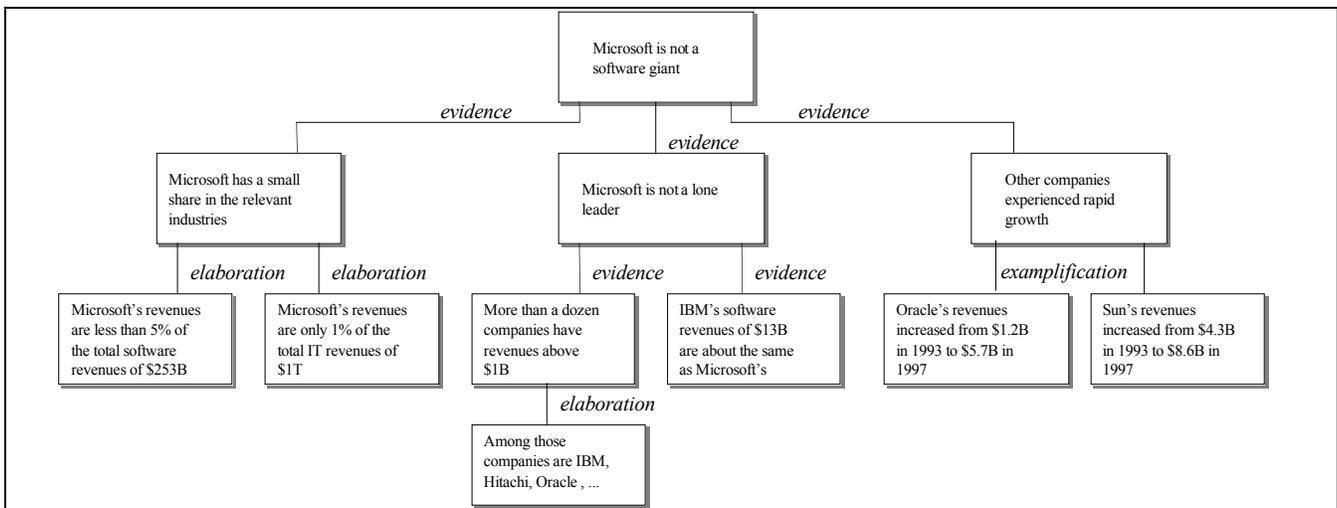

Figure 1. The communicative plan of the text

In our example, primary candidates for graphical presentation are the assertions near the bottom of the tree. Most of them are quantitative (e.g., the percentages of Microsoft's revenues, and the actual revenues of IBM, Oracle, and Sun). Some represent complex relationships (e.g., revenue trends). Others represent multiple homogeneous facts (e.g., the names of the companies with high revenues). Each assertion satisfying the above criteria is considered for presentation in graphics.

4. Mapping intent to tasks

Casner [2] demonstrated that the design of effective graphics is greatly influenced by the tasks that the user needs to perform. In other words, the same data can be visualized in many different ways but these visualizations will support different tasks. We hypothesize that tasks are an adequate representation for translating a media-independent intentional representation into a specification that guides the graphic designer. Communicative effects are achieved in graphics by *demonstrating* the relationships specified in the media-independent goal and letting the user *detect* them on their own by performing certain perceptual and cognitive tasks. Hence, for achieving each intention we need to find out what tasks the user will need to perform in order to get the message.

The transformation from intentions to tasks is not straightforward. A major difficulty arises from the fact that intentions and tasks have different presuppositions about the presentation environment. Perceptual tasks require that there is some external representation[2] of the domain that the user will be able to perceive. Intentions, on the other hand, do not. This discrepancy between intentions and tasks requires that, before or in the process of defining the tasks, the system also determine which objects will be represented externally. We refer to the set of objects that will be presented in the visualization as *scope*.

To illustrate the process of mapping intentions to tasks, consider the proposition "Sun's revenues increased from 1993 to 1997," intended to increase the user's belief that Microsoft is not the only company experiencing rapid growth. This goal can be achieved in a graphic that enables the user to effectively *perceive the trend* of Sun's yearly revenues. Such a characterization of the graphic is clearly task-oriented. To complete the characterization, we also need to define the scope, which is the set of objects to be included in the graphic. Given this proposition, the scope clearly is the set of yearly revenues from 1993 till 1997. The connected plot chart in Figure 2 is one, and possibly the most effective, technique to support this task and hence to achieve the original goal.

This goal was achieved by a one-shot task – *lookup the trend* of a variable (amount of revenues) with respect to another variable (year). Other goals lead to tasks with a more complex structure. Consider the proposition "IBM's software revenues are about the same as Microsoft's" used as evidence for the assertion that Microsoft is not a lone leader. This proposition can be conveyed effectively in a graphic that enables the user to *compare* the software revenues of the two companies. For this comparison to happen, however, the user first needs to *search* the space of graphical objects depicted in the graphic for the two graphemes that correspond to the two companies of interest. A *sequence* operator represents the dependency of the compare task on the search tasks. On the other hand, the search tasks for the two companies can be performed in any order; therefore, they are grouped by a *disjoint* operator. The bar chart in Figure 2 realizes this more complex, procedure-like task.

Space does not permit detailed description of the rules for mapping goals to tasks, but the most important ones are briefly summarized below:

- each proposition is mapped to a SEQUENCE of optional SEARCH tasks that identify the objects participating in the proposition, and either a COMPARE or a LOOKUP task that asserts something about these objects;
- depending on the type of the main predicate, the assert task is COMPARE if the main predicate is some kind of relation (e.g., $>$, $=$, $\approx$), or LOOKUP if the main predicate is an attribute;
- SEARCH tasks for more than one argument of the same proposition are grouped by DISJOINT if no dependencies exist between them;
- if one of the objects is described using some properties of the other, the SEARCH tasks are grouped by SEQUENCE;
- SEARCH tasks for the same object using different attributes are grouped by CONJOIN operator.

5. Task aggregation

Before extracting any useful information, the viewers need to spend some time and effort to understand the graphic. However, once they understand it, they can easily "read" more and more facts with great efficiency, in some cases even in parallel. Also, graphics can convey not only isolated assertions but also rhetorical relations between propositions, which makes them even more valuable. For example, in the financial markets domain, it is common to *contrast* the returns of two stocks by showing their trends on the same graphic. For these two reasons, it is often beneficial to pack several propositions into the same graphic. The realization of several intentions with one set of tasks is called *task aggregation*.

---

[2] "External representation" means outside user's mind.

Task aggregation occurs only if the corresponding intentions are related rhetorically and if their scopes are compatible. The condition for *rhetorical relatedness* ensures that the facts to be conveyed by the same graphic contribute to the same higher-level goal. For example, the assertion about the companies with revenues above $1 billion would rather aggregate with the comparison of the revenues of IBM and Microsoft than with the Microsoft's share in the software industry. The condition for *compatibility of scopes* means that the tasks refer to the same types of objects and either the descriptions of these objects use the same attributes or the scopes of the tasks intersect. If two tasks satisfy the conditions for aggregation, a new task is considered which combines them into a DISJOINT group.

For example, consider the proposition "Oracle's revenues increased from 1993 to 1997," which leads to the same type of task as the analogous proposition about the increase of Sun's revenues. These two tasks share compatible scopes and contribute to the same high-level communicative goal. Therefore, they can be aggregated yielding the connected plot charts in Figure 2. A similar but more subtle aggregation occurs between the proposition about the similarity of the revenues of IBM and Microsoft and the companies whose revenues exceed $1 billion, which results in the bar chart in Figure 2.

6. Graphical techniques

Graphic design selects techniques that support given tasks. For instance, to help the user look up the trend of the yearly revenues from 1993 to 1997, we choose a connected plot chart (Fig. 2). Research in graph interpretation has clearly shown that linked plot charts are the superior technique for conveying trends although other techniques such as a bar chart are also feasible [6].

Another subtask, searching for yearly revenues by *year,* will be supported if the *year* attribute is encoded by position. The fact that *year* is an ordered data type ensures that the reader will easily find the value 1993 on the corresponding axis, identify the graphical object corresponding to this value, and associate it with the yearly revenue object. The next task, looking up the amount of revenue for the object just found, will be supported best by a label, even though this attribute may already be encoded by some other technique (e.g., position, color, saturation, or size).

A variety of graphical techniques supporting tasks have been collected in prior work [2, 9].

7. The example - revisited

The methodology described in the previous sections of the paper allows us to build systems that automatically design visualizations achieving communicative intentions.

Such visualizations are part of a more appealing multimedia presentation that takes advantage of both natural language and graphics.

One possible multimedia redesign of our sample text is shown in Figure 2, where the three intentions in support of the main point are achieved in separate sections of the presentation. Each section consists of one statement conveying the main intention and one or more graphics supporting it by conveying specific quantitative relationships. The pie charts emphasize the smallness of the share of Microsoft's revenues. The comparison between the revenues of IBM and Microsoft in the bar chart stands out against the background of the other high-revenue companies. The linked plot chart depicts the increasing yearly revenues of Oracle and Sun.

Future research is needed to determine whether multimedia representations like the one in Figure 2 are more effective than straight text.

8. Conclusion

We propose a methodology for designing visualizations that achieve given intentions. The main elements of this methodology are an intentional representation, media allocation, mapping intentions to tasks, and selecting graphical techniques. Other elements that were not presented in this paper but are related to the problem of visualization design are planning the presentation and coordinating the visualizations with natural language text. These we plan to tackle in our future work.

Compared to the work of others, ours is unique in the following aspects:

- We start with a principled representation of the intentions of the speaker created during media-independent presentation planning [8, 3].
- We employ graphics interpretation tasks as an intermediary representation that spells out the way intentions are achieved in graphics and guides the graphic designer.
- We employ the rich collection of graphical techniques of the Sage system applicable to a wide variety of graphics [9].

We are currently building a system, AutoBrief, which automatically plans and designs effective multimedia presentations of problems in transportation schedules [4]. The techniques described in this paper actually span several modules of this system. Our focus here, however, is on the problems and their solutions rather than on the system components and architecture.

Acknowledgments

References

1. Bertin, J. *Semiology of Graphics: Diagrams, Networks, Maps*. Madison, Wisconsin: The University of Wisconsin Press, 1983.

2. Casner, S.M. A Task-Analytic Approach to the Automated Design of Information Graphic Presentations. *ACM Transactions on Graphics,* 10(2), 1991, 111-151.

3. Green, N., Carenini, G., Kerpedjiev, S., Roth, S, and Moore, J. A Media-Independent Content Language for Integrated Text and Graphics Generation (to appear).

4. Kerpedjiev, S., Carenini, G., Roth, S. F., and Moore, J. D. AutoBrief: a Multimedia Presentation System for Assisting Data Analysis, *Computer Standards and Interfaces*, 18, 1997, 583-593.

5. Larkin, J., and Simon, H. Why a Diagram Is (Sometimes) Worth a Thousand Words. *Cognitive Science*, 11, 65-99.

6. Lohse, G.L. A Cognitive Model of Understanding Graphical Perception. *Human-Computer Interaction*, 8, 1993, 353-388.

7. Mann, W., and Thompson S. Rhetorical Structure Theory: Towards a Functional Theory of Text Organization. *Text,* 8(3), 1988, 243-281.

8. Moore J. D. *Participating in Explanatory Dialogues*. MIT Press, 1995.

9. Roth, S. F., and Mattis J. Data Characterization for Intelligent Graphics Presentation. *Proc. SIGCHI'90*, Seattle, WA, ACM, April, 1990, pp. 193-200.

10. Roth, S.F., and Hefley, W.E. Intelligent Multimedia Presentation Systems: Research and Principles. In M. Maybury (ed.) *Intelligent Multimedia Interfaces*, AAAI Press, 1993, 13-58.

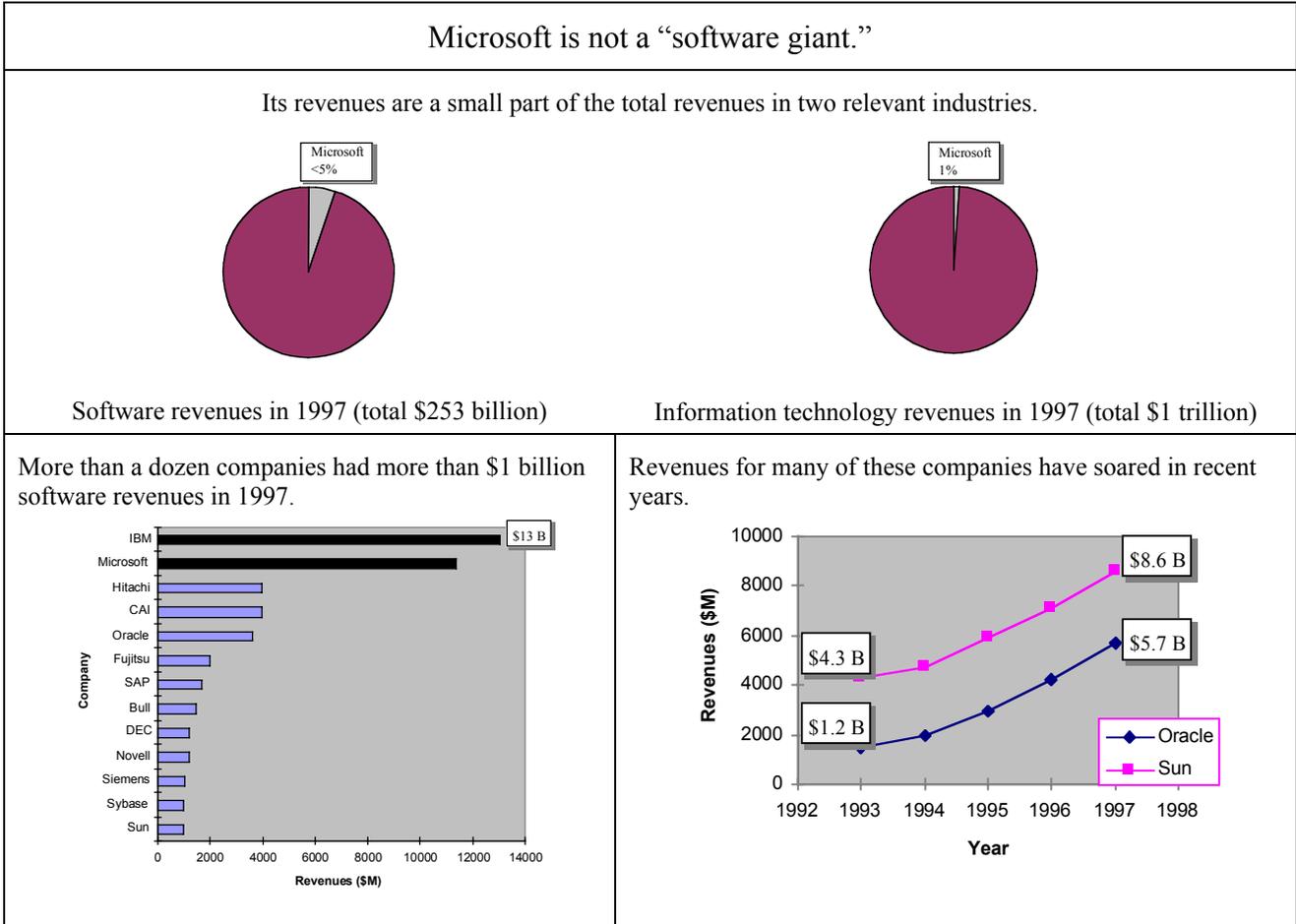11. Searle, J.R. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press, 1969.

Figure 2. A multimedia presentation of the argument