Communicative Signals as the Key to Automated Understanding of Simple Bar Charts

Conference Paper *in* Lecture Notes in Computer Science · June 2006

DOI: 10.1007/11783183_5 · Source: DBLP

Stephanie Elzer Schwartz
Millersville University

Communicative Signals as the Key to Automated Understanding of Simple Bar Charts*

Stephanie Elzer[1], Sandra Carberry[2], and Seniz Demir[2]

[1] Dept of Computer Science, Millersville Univ., Millersville, PA 17551 USA
elzer@cs.millersville.edu
[2] Dept of Computer Science, Univ. of Delaware, Newark, DE 19716 USA
{carberry,demir}@cis.udel.edu

Abstract. This paper discusses the types of communicative signals that frequently appear in simple bar charts and how we exploit them as evidence in our system for inferring the intended message of an information graphic. Through a series of examples, we demonstrate the impact that various types of communicative signals, namely salience, captions and estimated perceptual task effort, have on the intended message inferred by our implemented system.

1 Introduction

Information graphics such as bar charts, line graphs and pie charts are an important component of many documents. As noted by [1], [7], and [24], among many others, a set of data can be presented in many different ways, and graphs are often used as a communication medium or rhetorical device for presenting a particular analysis of the data and enabling the viewer to better understand this analysis. Although some information graphics are only intended to display data, the majority of information graphics that appear in formal reports, newspapers, and magazines are intended to convey a message, thus leading us to consider information graphics as a form of language. As Clark noted, language is more than just words. It is any "signal" (or the lack of a signal when one is expected), where a signal is a deliberate action that is intended to convey a message [6]. Clark's expanded definition of language includes very diverse forms of communication (such as gesture and facial expression) and the common factors among these varied forms of expression is the communicative intention underlying them and the presence of deliberate signals to aid in recognizing these intentions.

The design choices made by the designer when constructing an information graphic provide the communicative signals necessary for understanding the graphic. The design choices include selection of graphic type (bar chart, line graph, pie chart, etc.), organization of information in the graphic (for example,

* This material is based upon work supported by the National Science Foundation under Grant No IIS-0534948.

aggregation of bars in a bar chart), and attention-getting devices that highlight certain aspects of a graphic (such as coloring one bar of a bar chart different from the others, mentioning data elements in the caption, etc.). This paper focuses on the communicative signals that result from these design choices. It begins by presenting a very brief overview of our system for inferring the intended message of a bar chart. It then describes the types of communicative signals that our system extracts from simple bar charts, and provides several examples that illustrate how different communicative signals impact the message that is recognized by our implemented and evaluated system. It concludes by discussing the applications that we envision for our system.

2 A Bayesian Approach to Graphic Understanding

Figure 1 shows the architecture of our system for inferring the intended message of an information graphic. The visual extraction module (VEM) analyzes the graphic and produces an XML representation containing information about the components of the information graphic including the graphic type (bar chart, pie chart, etc.) and the caption of the graphic. For a bar chart, the representation includes the number of bars in the graph, the labels of the axes, and information for each bar such as the label, the height of the bar, the color of the bar, and so forth [5]. The XML representation is then passed to the caption tagging module (CTM) which extracts information from the caption (see Section 3.2) and passes the augmented XML representation to the intention recognition module (IRM). The IRM is responsible for recognizing the intended message of the information graphic, which we hypothesize can serve as the basis for an effective summary of the graphic. The scope of the work currently implemented is limited to the processing of simple bar charts. By simple bar charts, we mean bar charts that display the values of a single independent attribute and the corresponding values for a single dependent attribute. Although the type of information graphics is limited, we believe that the concepts, mechanisms and framework of our methodology is broadly applicable and extensible to other types of graphics.

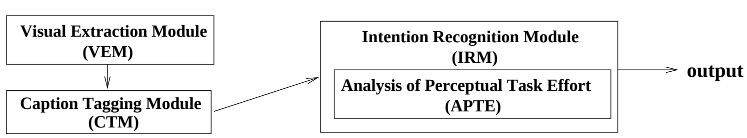

Fig. 1. System Architecture

2.1 Extending Plan Inference to Information Graphics

In our research, we have extended plan inference techniques that have been used successfully in identifying the intended meaning of an utterance to inferring the message conveyed by an information graphic [10]. Our goal is to identify the message that the graphic designer intended to convey, by recognizing his plan for the viewer — i.e., by recognizing the perceptual and cognitive tasks that the viewer is intended to perform in deciphering the graphic's intended message. By *perceptual tasks* we mean tasks that can be performed by simply viewing the graphic, such as finding the top of a bar in a bar chart; by *cognitive tasks* we mean tasks that are done via mental computations, such as computing the difference between two numbers [16]. Of course, we realize that not all graphics are well designed, and in the case of a poorly-designed graphic, we may not be able to infer the message that the graphic designer intended; in this case, our goal is to infer the same message that a human would get by viewing the graphic.

Following the work of Charniak [4] and others, we capture plan inference in a Bayesian network. In all plan inference systems, there is some explicit plan structure which defines the relationships among goals, subgoals and primitive actions. In Bayesian networks, the plan structure is captured by the network itself; each goal, subgoal, and primitive action is represented as a node in the network. If a goal can be decomposed into a particular set of subgoals or primitive actions, an arc from the goal to each subgoal (or primitive action) is used to represent this causal relationship. Our Bayesian network captures knowledge about how the graphic designer's goal of conveying a message can be achieved via the viewer performing certain perceptual and cognitive tasks, as well as knowledge about how perceptual and cognitive tasks decompose into sets of simpler tasks. For example, the graphic designer might intend for the viewer to find the relative difference between two data elements represented by bars in a bar chart — i.e., whether the value represented by the top of one bar is greater than, less than, or equal to the value represented by the top of the second bar. The viewer might achieve this goal by locating the bar that represents the first data element (by looking at the labels of the bars), locating the bar that represents the second data element, and then perceptually comparing the heights of the bars. This task decomposition is captured in our network, as shown in Figure 2.

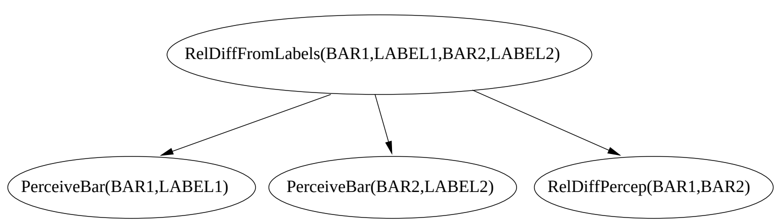

Fig. 2. A Task Decomposed Into Its Subgoals

In addition to nodes capturing the task structure, our Bayesian network for recognizing the message conveyed by an information graphic contains evidence nodes that reflect communicative signals in the graphic. Each evidence node captures the causal relationship between its parent node(s) being (or not being) part of the plan for the viewer to decipher the graphic's message and the presence of a particular kind of communicative signal in the graphic.

The arcs between the nodes in a Bayesian network capture causal dependencies using conditional probability distributions that represent the likelihood of a proposition given the various values of its parent node(s). Bayes' rule is then used to compute the probability of each proposition given causal evidence (from its parents) and diagnostic evidence (from its children) [9, 21]. In a Bayesian network used for plan inference, the root nodes represent various high-level hypotheses about an agent's plan, and the probabilities computed for those nodes represent the likelihood of the hypotheses given the available evidence [2]. In many domains, it is difficult to empirically determine the probabilities that should be used in the conditional probability tables; however, the probabilities used in our network have been obtained through an analysis of a corpus of 110 bar charts that were previously annotated with their intended messages. Since the focus of this paper is on identifying and exploiting the communicative signals present in simple bar charts, space precludes further discussion of the Bayesian network itself; however, greater detail can be found in [13].

3 Communicative Signals in Bar Charts

A key component of a plan inference system is the evidence that is used to guide the inference process, so that one hypothesis might eventually be preferred over another. Therefore, in extending plan inference techniques to the recognition of intentions from information graphics, we need to identify the types of evidence present in information graphics and exploit them in our Bayesian network. The evidence that we utilize in our system consists of communicative signals resulting from the design choices made by the graphic designer. Section 3.1 discusses the types of communicative signals that we extract from the graphic itself (such as highlighting, annotations, and perceptual task effort), while Section 3.2 discusses the communicative signals we extract from captions.

3.1 Communicative Signals in the Graphic Itself

Salience: Our contention is that if the graphic designer goes to the effort of employing attention-getting devices to make certain elements of the graphic particularly salient, then the salient elements serve as communicative signals — i.e., the designer probably intends for them to be part of the intended message of the graphic. By examining a corpus of bar charts gathered from magazines and newspapers, we have identified several of the most common design techniques that graphic designers employ to increase the salience of an element or elements in simple bar charts.

In order to draw attention to a particular element (or elements) of a bar chart, the graphic designer may choose to *highlight* it. From our examination of a corpus of bar charts, we have found that graphic designers typically highlight an element or elements of a bar chart by drawing the viewer's attention to the bar itself or to an attribute of the bar, such as its label or its annotated value. For example, the designer could highlight a bar in the graphic by making it a different color, shade or texture than the other bars. This is a communicative signal, conveying to the viewer of the graphic that the bar (and thus the data element that it represents) is of significant import to the message of the graphic. Consider the graphic in Figure 3 which appeared in a local newspaper. The graphic appeared in shades of gray, as it is depicted here, with the bar representing June of 2001 in a darker shade than the other bars. The design choice to highlight this bar by making it a darker shade of gray than the other bars seems to signal the importance of the unemployment rate in June of 2001 to the message that the designer is attempting to convey with the graphic — ostensibly, the contrast between the unemployment rate a year ago and the most recent unemployment data. In addition, instead of highlighting the bar itself, the designer could highlight attributes of the bar such as its label or its annotated value in order to increase the salience of the represented element. Thus the highlighting of either 1) the bar itself or 2) attributes of the bar (such as the label or annotated value) serves as a communicative signal regarding the salience of the element to the message of the graphic.

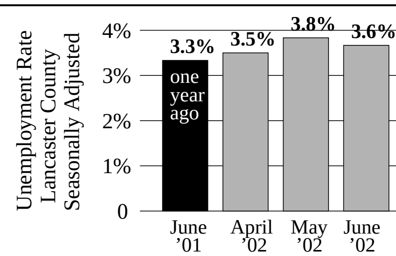

Fig. 3. Bar Chart Showing Unemployment Data[3]

A graphic designer can also convey the significance of an element in an information graphic by annotating the salient element in some way. The most common form of annotation in our corpus of information graphics is the annotation of an element with its exact value. Annotating individual elements with their exact values can signal salience if the annotations are *not* a general design feature of the graphic. If all of the elements are displayed with their exact values (as is the case in Figure 3), then we consider this to be a general design feature of the graphic since the annotations do not draw attention to a specific subset of elements. However, if only a subset of the elements are annotated with their values, the annotations signal the salience of those elements. This is the case in Figure 4 where only the first and last elements are annotated with their values. Annotations of bars in a bar chart are not limited to the exact value represented by the bar — they can also include content such as dates or other additional notes. In the graphic shown in Figure 3, the bar representing June 2001 has a

[3] This is based on a bar chart from a local newspaper (the Lancaster Intelligencer Journal on July 30, 2002).

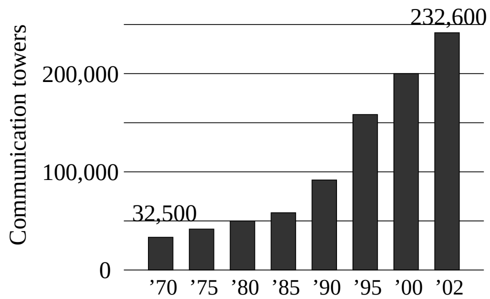

Fig. 4. Bar Chart Showing the Number of Communication Towers[4]

note annotation of "One year ago". The fact that this bar is annotated with additional information further indicates its importance to the message the graphic designer is trying to convey with the graphic.

We have also identified several factors that increase the salience of an element in a graphic without the application of any particular design techniques. Although no specific action is required on the part of the graphic designer to make these elements salient, we posit that it is mutually believed by both designer and viewer that such elements will be salient to the viewer. These elements include any element that is significantly taller than all of the other elements in the graphic and the most recent date on a time-line, since the viewer will certainly notice the height of a bar that is much taller than all of the others, and will naturally be interested in what has occurred most recently.

The Visual Extraction Module (VEM) of our system produces an XML representation of a bar chart that includes information about each bar such as bar height, annotations, color, and so forth [5]. By analyzing this XML representation, our system is able to identify any particularly salient elements of the graphic according to the aforementioned criteria for salience.

Perceptual Task Effort: Given a set of data, the graphic designer has many alternative ways of designing a graphic. As Larkin and Simon note, information graphics that are *informationally* equivalent (all of the information in one graphic can also be inferred from the other) are not necessarily *computationally* equivalent (enabling the same inferences to be drawn quickly and easily) [18]. Peebles and Cheng [22] further observe that even in graphics that are informationally equivalent, seemingly small changes in the design of the graphic can affect viewers' performance of graph reading tasks. Much of this can be attributed to the fact that design choices made while constructing an information graphic will facilitate some perceptual tasks more than others. Following the AutoBrief work on generating graphics to achieve communicative goals, we hypothesize that the designer chooses a design that best facilitates the tasks that are most important to conveying his intended message, subject to the constraints imposed by competing tasks [16, 14].

In order to identify the perceptual tasks that the graphic designer has best enabled in the graphic, our methodology was to construct a set of rules that

[4] This is based on a bar chart from the newspaper USA Today.

estimate the effort required for different perceptual tasks within a given information graphic. To develop these rules, we applied the results of research from cognitive psychology. In doing this, we constructed a model representing the relative ease or difficulty with which the viewer of a graphic could complete various perceptual tasks. The component of our system that is responsible for estimating effort is called APTE (Analysis of Perceptual Task Effort). The goal of APTE is to determine whether a task is easy or hard to perform with respect to other perceptual tasks that could be performed on an information graphic.

In order to estimate the relative effort involved in performing a task, we adopted a GOMS-like approach [3], decomposing each task into a set of component tasks. Following other cognitive psychology research, we take the principal measure of the effort involved in performing a task to be the amount of time that it takes to perform the task, and our effort estimates are based on time estimates for the component tasks.[5] Wherever possible, we utilize existing time estimates (primarily those applied in Lohse's UCIE system) for the component tasks. For example, the rule shown in Figure 5 estimates the effort required to determine the exact value represented by the top of a bar in a bar chart, given that the viewer is already focused on the top of the bar.[6] In the case of condition-computation pair B1-1 (finding the exact value for a bar where the bar is annotated with the value), the effort is estimated as 150 units for discriminating the label (based on work by Lohse [19]) and 300 units for recognizing a 6-letter word [15]. In the case of B1-2 (finding the exact value for a bar where the top of the bar is aligned with a tick mark on the axis), the effort estimate includes scanning over to the dependent axis (measured in terms of distance in order to estimate the degrees of visual arc scanned [17]) in addition to the effort of discriminating and recognizing the label. Our eye tracking experiments showed that when the top of the bar is aligned with a tick mark, participants frequently repeat the task of scanning to the axis and reading the label (presumably to ensure accuracy), so our effort estimate also includes 230 units [23] to perform a saccade back to the top of the bar before repeating the task. Our set of APTE rules for estimating the effort of tasks in bar charts and our eye tracking experiments that validated those rules are described in [12, 10].

The evidence provided by perceptual task effort can sometimes be subsumed (probabilistically) by other communicative signals, such as a bar being highlighted or the presence of helpful words in a caption. However, this communicative signal plays an extremely important role in our system; in the absence of all other evidence, we always have communicative signals provided by perceptual task effort. For example, an information graphic might not have any salient elements and may be lacking a helpful caption, but our system can still reason about the relative ease or difficulty with which tasks can be performed on the

[5] The units of effort estimated by our rules roughly equate to milliseconds.

[6] *Rule-B1* does not estimate the effort required to get the value represented by the top of a bar in the case where the viewer must scan to the axis and interpolate an estimated value. This task is represented by a separate rule in our system.

Rule-B1:Estimate effort for task
PerceiveValue(<viewer>, <g>, <att>, <e>, <v>)

Graphic-type: bar-chart

Gloss: Compute effort for finding the exact value <v> for attribute <att>
represented by top <e> of a bar <b> in graph <g>

Conditions:
B1-1: IF the top <e> of bar <b> is annotated with a value,
THEN effort=150 + 300
B1-2: IF the top <e> of bar <b> aligns with a labelled tick mark on
the dependent axis, THEN effort=230 + (scan + 150 + 300) x 2

Fig. 5. APTE Rule for Estimating Effort for the Perceptual Task *PerceiveValue*

given graphic and may, therefore, be able to draw useful inferences about the intended message of the graphic designer.

3.2 Communicative Signals in Captions

When considering the communicative signals available in an information graphic, one might suggest relying on a graphic's caption to capture its primary message. However, the work of Corio and LaPalme [8], who studied captions with the objective of categorizing the kinds of information contained in captions in order to form rules for generating captions to accompany graphics, as well as our own corpus study of captions associated with bar charts, show that captions are often missing, ill-formed, or too general to be solely relied on for summarizing an information graphic. In our corpus study, almost half the captions (44%) failed to contribute at all to understanding the graphic's intended message, and only 34% of the captions were judged to convey most of the intended message. Details of our corpus study can be found in [11].

Although captions cannot be relied upon as the sole mechanism for recognizing a graphic's message, we do want to exploit any evidence that is contained in a caption. However, our corpus analysis also showed that full understanding of a caption through a general natural language understanding system would be problematic. This is due to the fact that many captions are ill-formed (often involving ellipsis or sentence fragments), or would require extensive domain knowledge or analogical reasoning to understand. Moreover, once the caption was understood, we would still need to relate it to the information extracted from the graphic itself, which appears to be a difficult problem. Thus we began investigating whether shallow processing of the caption might provide evidence that could be effectively combined with the other communicative signals gleaned from the graphic itself. Our analysis provided the following observations:

– Verbs in a caption often suggest the general category of message being conveyed by the graphic. An example from our corpus is *"American Express total billings still lag"*; the verb *lag* suggests that the graphic conveys that some entity (in this case *American Express*) falls behind some others.

– Adjectives in a caption also often suggest the general category of message being conveyed by the graphic. An example from our corpus is *"Soaring Demand for Servers"* which is the caption on a graphic that conveys the rapid increase in demand for servers. Here the adjective *soaring* is derived from the verb *soar*, and suggests that the graphic is conveying a strong increase.
– Words that usually appear as verbs, but are used in the caption as a noun, may function similarly to verbs. An example is *"Cable On The Rise"*; in this caption, *rise* is used as a noun, but suggests that the graphic is conveying an increase.
– Nouns in a caption sometimes refer to an entity that is a label on the independent axis of the graphic. When this occurs, the caption brings the entity into focus and suggests that it is part of the intended message of the graphic. An example from our corpus is *"Germans miss their marks"* where the graphic displays a bar chart in which Germans correlates with a label in the graphic and the graphic is intended to convey that Germans are the least happy with the Euro.

Based on these observations, we designed and implemented a type of shallow processing of captions to extract communicative signals consisting of 1) *helpful* verbs and adjectives (identified through our corpus study, WordNet [25] and a thesaurus [20]) and 2) nouns which match the label of a data element in the bar chart (this is actually a way for a designer to emphasize the salience of an element or elements through the caption.)

4 Exploiting Communicative Signals in an Automated System

The communicative signals that we have discussed (such as highlighting, perceptual task effort and caption evidence) are utilized as evidence nodes in our network; for each perceptual task node in the network (such as PerceiveBar in Figure 2), evidence nodes representing the various communicative signals are added as children of the perceptual task node. Evaluation of our system using leave-one-out cross validation[7] on a corpus of 110 bar charts showed that our system had a success rate of 79.1% at inferring the message conveyed by the graphic. The system was judged to fail if either its top-rated hypothesis did not match the intended message that was assigned to the graphic by the human coders or the probability rating of the system's top-rated hypothesis did not exceed 50%. The output of our system consists of a logical representation of the top-rated hypothesis and its relevant parameters from which a natural language gloss can be derived.

[7] Leave-one-out cross validation in the context of a Bayesian network means that for each test graphic, all conditional probability tables are computed anew with the data from the test graphic itself excluded.

In order to demonstrate how our Bayesian system utilizes the communicative signals, this section presents several examples which illustrate how different kinds of evidence impact our network's hypothesis as to the intended message of a bar chart. In order to clearly show the impact of the choices made by a graphic designer, we will consider multiple variations of bar charts displaying the same underlying data.[8]

4.1 Bar Chart with Minimal Communicative Signals

As our first example, consider the bar chart shown in Figure 6, which shows data regarding the average purchases by customers of various credit card companies. The input to the Intention Recognition Module (IRM) is a file containing an XML representation of the information graphic. Before being processed by

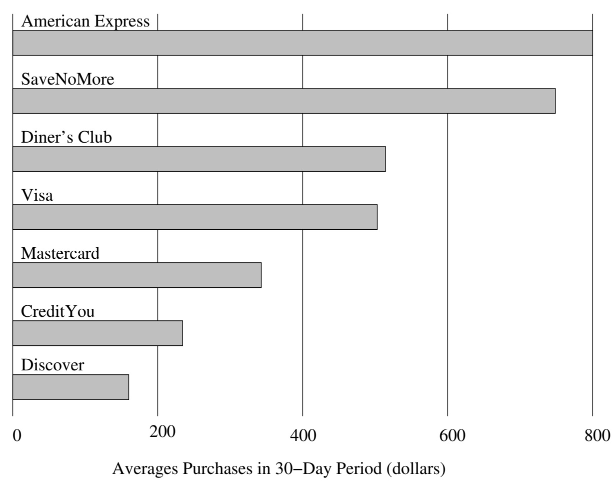

Fig. 6. Information Graphic Example

the Intention Recognition Module, the XML representation is augmented by the Caption Tagging Module (CTM) to include information about any relevant verbs, adjectives, and nouns. Since the bar chart shown in Figure 6 does not have a caption, no additional caption information will be recorded in the XML; that is, there are no communicative signals provided by a caption. Also, according to our criteria for salience of an element, there are no particularly salient elements in the bar chart shown in Figure 6. Thus the only communicative signals in this graphic are the relative effort of the different perceptual and cognitive tasks.

[8] The data displayed in the example graphics is based on a bar chart that appeared in the September 13, 2004 issue of Business Week. However, the original graphic had only five bars; we have added two bars to enrich the complexity of the graphics. To our knowledge, CreditYou and SaveNoMore are fictitious entities.

Given the XML representation of the bar chart, our system hypothesizes that the graphic is intended to convey the relative rank in amount of purchases for the various credit card companies displayed in the bar chart and assigns this intention a probability of 88.6%. Other possibilities also have some probability assigned to them. For example, the intention of conveying that American Express has the highest average customer purchases is assigned a probability of 9.9% because the bars are in sorted order according to height, thus making it relatively easy for the viewer to recognize the maximum, and because finding the entity in the graphic with the maximum value is a fairly common intention (occurring approximately 22.7% of the time in our corpus). However, there is no other evidence suggesting that the bar representing the maximum value is salient (such as that bar being highlighted), so the system hypothesizes that the primary intention of the graphic is to convey the relative rank of all of the companies listed. Several other hypotheses have nearly negligible probabilities (less than 1%), including the hypotheses that the graphic is intended to convey that Discover has the lowest average customer purchases, the relative difference between two elements, or the rank of a particular element.[9]

4.2 A Bar Chart with a Single Salient Element

Now suppose that the bar representing Diner's Club was darker than the other bars in the bar chart (as shown in Figure 7), thus making this element salient. The fact that the bar is highlighted provides strong evidence that it plays a role in the intended message of the graphic. Our system takes this into account, along with the relative effort of the different perceptual tasks, and the Bayesian network that is constructed for this information graphic hypothesizes that the intended message of the graphic shown in Figure 7 is to convey that Diner's Club ranks third among the companies shown in the bar chart with respect to average customer purchases over a 30-day period. This hypothesis is believed to be extremely likely, as reflected by its calculated probability of 97.26%.

4.3 A Bar Chart with Two Salient Elements

Elements of the graphic could also be made salient in other ways, such as through annotations. Suppose that the bar representing Diner's Club was still darker than the other bars, but that the bars representing Diner's Club and Mastercard (and only those bars) were annotated with their exact values, as shown in Figure 8. Here the evidence still suggests the salience of Diner's Club, as in the previous example, but also suggests that Mastercard is salient. The fact that two bars are now salient will provide evidence against intentions involving only Diner's Club and will favor hypotheses involving both bars. Thus it is not surprising that the system hypothesizes the intention of the graphic to be finding the relative difference (and the degree of that difference) between the average customer purchases of Diner's Club and Mastercard and assigns it a likelihood of 88.8%.

[9] For the latter two messages, there were several instantiated hypotheses with nearly equal probabilities.

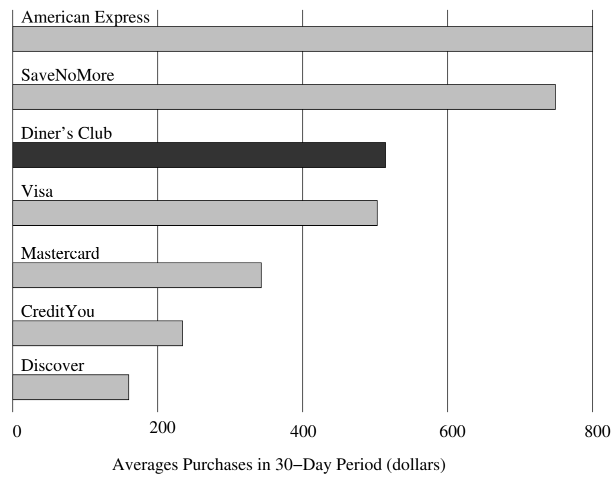

Fig. 7. Information Graphic Example with a Single Salient Element

4.4 A Bar Chart with a Helpful Caption

So far, none of our examples have included a caption. Suppose, however, that the caption of our previous example was "Diner's Club Beats Mastercard" In this case, the initial XML representation would be augmented by the Caption Tagging Module to include additional information about the caption. The verb "beat" (the root form of "beats") is considered to be a helpful verb, and the nouns "Diner's Club" and "Mastercard" match the labels of bars in the bar chart, so this information is included in the augmented XML representation that is input to the Intention Recognition Module (IRM). The inclusion of "Diner's Club" and "Mastercard" in the caption provides additional evidence of the importance of these elements to the intended message of the graphic, and the presence of the verb "beat" provides evidence regarding the general category of message of the graphic. These additional communicative signals strengthen the system's belief in its hypothesis of the relative difference and degree message, and it now assigns it a probability of 99.8% (as opposed to 88.8% without the caption).

4.5 Conflicting Communicative Signals

Now consider a significant variation of the graphic design. Suppose, again, that the bar representing Diner's Club was darker than the other bars. But now suppose that the bars were sorted by the alphabetical order of their labels, rather than by descending order of their height. This variation is shown in Figure 9. The perceptual task of determining the rank of Diner's Club is now estimated as being *hard* (or difficult) to perform. This new evidence results in the system assigning a probability of only 5% to the GetRank message. In fact, the system hypothesizes the most likely message (with a probability of 63.1%) to be that

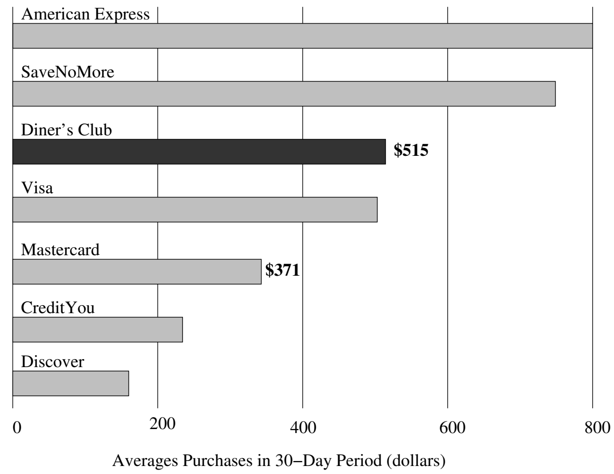

Fig. 8. Information Graphic Example with Two Salient Elements

American Express has the highest customer purchases. This is a dubious conclusion, but it illustrates the impact of the conflicting communicative signals on the inference process. Diner's Club is salient, but any tasks that involve this element (such as getting the rank) are difficult to perform or do not match the salience evidence (for example, we could compare Diner's Club with another element, but we have no evidence suggesting the salience of another element). In contrast, finding the maximum is a much easier task to perform on this graphic. However, this is clearly a poorly designed graphic, and the conflicting communicative signals regarding salience and perceptual task effort present in this graphic result both in our system having relatively little confidence in its conclusion, and also make it difficult even for humans to hypothesize an intended message of the graphic (as reflected by the probability of 63.1%).

5 Conclusion

This paper has demonstrated how we exploit communicative signals from bar charts in our automated system. We have presented illustrative examples of simple bar charts which display the same underlying data, but contain very different communicative signals regarding the message that the viewer is intended to recognize. The messages inferred by our system for these different bar charts show the impact that various forms of evidence, namely salience, captions and estimated perceptual task effort, can have on the inference process.

In the future, our intention recognition system will become part of several larger projects. We hypothesize that the core message of an information graphic (the primary overall message that the graphic conveys) can serve as the basis for an effective summary of the graphic. This summary could then be used in

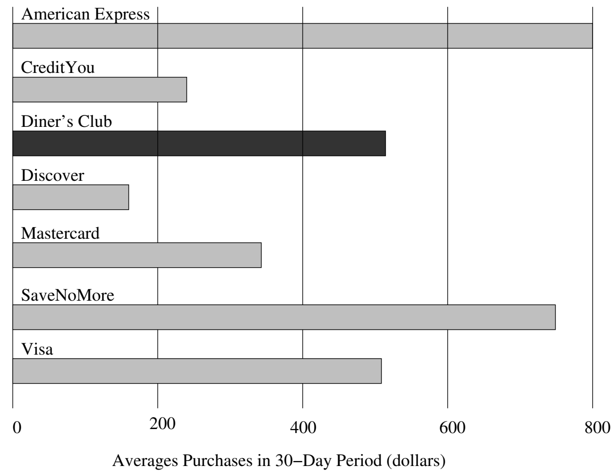

Fig. 9. Information Graphic Example with Conflicting Communicative Signals

a variety of ways. For digital libraries, the summary of the information graphic could be used to appropriately index the graphic and to enable intelligent retrieval. If there is accompanying text, the summary of the graphic can be used in conjunction with a summary of the document's text to provide a more complete representation of the document's content. For individuals who are sight-impaired or who are using low bandwidth devices, using the core message of the information graphic as the basis for a summary would provide access to the informational content of the graphic in an alternative modality. For graphic designers, the core message inferred by the system might indicate problems with the design of a graphic and thereby provide a tool for improving the graphic.

References

1. Jacques Bertin. *Semiology of Graphics.* The University of Wisconsin Press, Madison, Wisconsin, 1983.
2. Sandra Carberry. Techniques for plan recognition. *User Modeling and User-Adapted Interaction*, 11(1–2):31–48, 2001.
3. S. K. Card, T. P. Moran, and A. Newell. *The Psychology of Human-Computer Interaction.* Lawrence Erlbaum Associates, Inc., Hillsdale, NJ, 1983.
4. Eugene Charniak and Robert P. Goldman. A bayesian model of plan recognition. *Artificial Intelligence*, 64(1):53–79, November 1993.
5. Daniel Chester and Stephanie Elzer. Getting computers to see information graphics so users do not have to. In *Proceedings of the 15th International Symposium on Methodologies for Intelligent Systems (ISMIS), Lecture Notes in Artificial Intelligence 3488*, pages 660–668. Springer-Verlag, 2005.
6. Herbert Clark. *Using Language.* Cambridge University Press, 1996.
7. William S. Cleveland. *The Elements of Graphing Data.* Chapman and Hall, New York, 1985.

8. Marc Corio and Guy Lapalme. Generation of texts for information graphics. In *Proceedings of the 7th European Workshop on Natural Language Generation EWNLG'99*, pages 49–58, 1999.
9. Marek J. Druzdzel and Linda C. van der Gaag. Building probabilistic networks: 'where do the numbers come from?'. *IEEE Transactions on Knowledge and Data Engineering*, 12:481–486, 2000.
10. Stephanie Elzer. *A Probabilistic Framework for the Recognition of Intention in Information Graphics*. PhD thesis, University of Delaware, December 2005.
11. Stephanie Elzer, Sandra Carberry, Daniel Chester, Seniz Demir, Nancy Green, Ingrid Zukerman, and Keith Trnka. Exploring and exploiting the limited utility of captions in recognizing intention in information graphics. In *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 223–230, June 2005.
12. Stephanie Elzer, Sandra Carberry, Nancy Green, and James Hoffman. Incorporating perceptual task effort into the recognition of intention in information graphics. In *Diagrammatic Representation and Inference: Third International Conference, Diagrams 2004, LNAI 2980*, pages 255–270, March 2004.
13. Stephanie Elzer, Sandra Carberry, Ingrid Zukerman, Daniel Chester, Nancy Green, and Seniz Demir. A probabilistic framework for recognizing intention in information graphics. In *Proceedings of the Nineteenth International Joint Conference on Artificial Intelligence (IJCAI)*, pages 1042–1047, July 2005.
14. Nancy Green, Giuseppe Carenini, Stephan Kerpedjiev, Joe Mattis, Johanna Moore, and Steven Roth. Autobrief: an experimental system for the automatic generation of briefings in integrated text and information graphics. *International Journal of Human-Computer Studies*, 61(1):32–70, 2004.
15. B. E. John and A. Newell. Toward an engineering model of stimulus response compatibility. In R. W. Gilmore and T. G. Reeve, editors, *Stimulus-response compatibility: An integrated approach*, pages 107–115. North-Holland, New York, 1990.
16. Stephan Kerpedjiev and Steven F. Roth. Mapping communicative goals into conceptual tasks to generate graphics in discourse. In *Proceedings of Intelligent User Interfaces*, pages 157–164, 2000.
17. Stephen M. Kosslyn. Understanding charts and graphs. *Applied Cognitive Psychology*, 3:185–226, 1989.
18. Jill H. Larkin and Herbert A. Simon. Why a diagram is (sometimes) worth a thousand words. *Cognitive Science*, 11:65–99, 1987.
19. Gerald Lee Lohse. A cognitive model for understanding graphical perception. *Human-Computer Interaction*, 8:353–388, 1993.
20. Merriam-Webster On-Line Thesaurus, http://www.webster.com.
21. J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann Publishers, 1988.
22. David Peebles and P. C-H. Cheng. Modeling the effect of task and graphical representation on response latency in a graph reading task. *Human Factors*, 45:28–46, 2003.
23. J. E. Russo. Adaptation of cognitive processes to eye movement systems. In J. W. Senders, D. F. Fisher, and R. A. Monty, editors, *Eye movements and higher psychological functions*. Lawrence Erlbaum Associates, Inc., Hillsdale, NJ, 1978.
24. Edward R. Tufte. *The Visual Display of Quantitative Information*. Graphics Press, Cheshire, CT, 1983.
25. Wordnet, http://wordnet.princeton.edu/main/.