Analyzing the Effect of Communicative Evidence in a Bayesian System for Grouped Bar Chart Message Recognition

Richard Burns
Dept. of Computer Science
West Chester University
West Chester, PA 19383

Sandra Carberry
Dept. of Computer Science
University of Delaware
Newark, DE 19716

Stephanie Elzer Schwartz
Dept. of Computer Science
Millersville University
Millersville, PA 17551

Abstract

Information graphics, such as bar charts and line graphs, generally have a communicative message that they are intended to convey when they appear in popular media. The communicative signals present or absent in the graphic help convey these underlying intended messages. We have designed a Bayesian system that automatically extracts communicative signals from a grouped bar chart information graphic to hypothesize its intended messages. This paper presents experiments that analyze the effect that each type of communicative evidence has on our system.

Introduction

Information graphics, such as simple and grouped bar charts, pie charts, and line graphs are non-pictorial visual devices that are often incorporated into multimodal documents to convey one or more communicative goals (Iverson and Gergen 1997; Green et al. 2004). When information graphics appear in popular media (as opposed to scientific text), they generally have a high-level message that they are intended to convey. For example the grouped bar chart in Figure 1, ostensibly conveys the message that *"China increased spending on education, social security, military and rural support from 2004 to 2006."* A study by Carberry et al. (Carberry, Elzer, and Demir 2006) found that a graphic's message is often not repeated in the graphic's caption or in the article accompanying the graphic.

Our current research involves developing systems that automatically recognize the intended message of information graphics. This work contributes to several important applications: (1) providing alternative access to information graphics for sight-impaired individuals by conveying their high-level content to the user via speech (Demir et al. 2010); and (2) supporting the retrieval of information graphics from a digital library where the graphic's message is used to capture its high-level content (Carberry, Elzer, and Demir 2006).

Our previous work has focused on message recognition for simple bar charts (Elzer, Carberry, and Zukerman 2011) and line graphs (Wu et al. 2010). In this paper, we present

our recent research on grouped bar charts. Like simple bar charts which display bars that depict quantifiable relationships among the values of entities, grouped bar charts also contain a *grouping* dimension.

The overall objective of this work is a system for recognizing the high-level messages conveyed by a grouped bar chart (Burns 2012). This system has several important pieces of input. First, a vision system processes the graphic and produces an xml representation of the graphic, including its caption and other textual information that is present. Then, any present (or absent) communicative signals in the graphic and captions is extracted and entered into a Bayesian network. Finally, the Bayesian network probabilistically reasons about this evidence and ultimately outputs the system's hypothesis about the most likely intended messages of the graphic.

This paper presents an analysis of how different types of communicative evidence affect the accuracy of our Bayesian system. Our results show the importance of linguistic structural evidence, which is a novel component of the grouped bar chart recognition system as well as the kinds of evidence that have the greatest impact on message recognition, which should be considered in future work on graph understanding (such as for pie charts and multiple line graphs).
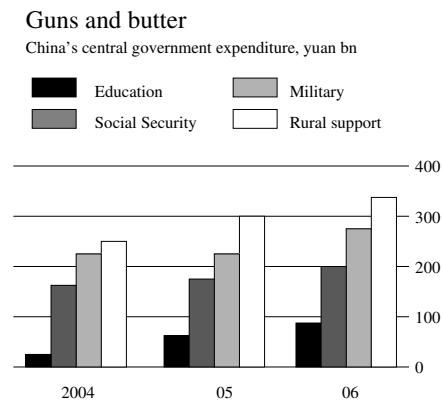

Figure 1: Graphic from *The Economist*, "Planning the new socialist countryside", March 9, 2006.
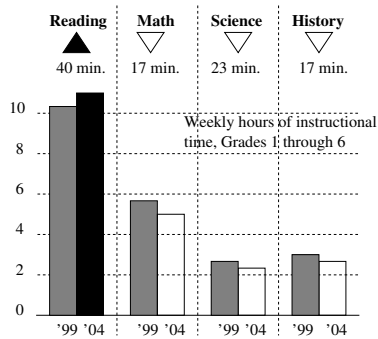
Figure 2: Graphic from *Time*, "How to Fix No Child Left Behind", June 4, 2007.

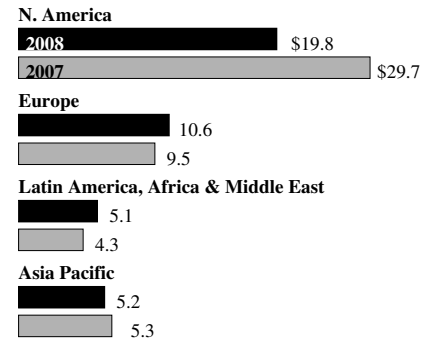

Figure 3: Graphic from *Wall Street Journal*, "GM to Build Diesel Engines in Thailand", August 14, 2008.

Messages Conveyed in Grouped Bar Charts

We collected 330 grouped bar charts from a variety of popular media sources, and analyzed this corpus to identify 25 different *message categories* that capture the kinds of messages that are overwhelmingly conveyed by grouped bar charts in popular media (Burns 2012). Human annotators recognized the intended messages in each graphic based on our generalized message categories. In this section we present a sampling of these message categories.

Trend Messages. Trend messages convey a general trend that is either rising, falling, or steady, over a set of ordinal data points. For example, the message conveyed by the grouped bar chart in Figure 1 is generalized into our *Rising-Trend* message category.

Relationship Messages. Relationship messages capture the consistency of relative values across a set of graphed entities, or the inconsistency of one set of relative values with respect to the other sets. For example, the grouped bar chart in Figure 2 ostensibly conveys that *"the amount of time spent on Reading increased while it decreased for the other subjects"*. This message is generalized into our *Entity-Relationship-Contrast* message category.

Gap Messages. The gap message category captures high-level messages that involve the *gap* between two bars, the approximate absolute difference between two values within the same entity. For example, the grouped bar chart in Figure 3 ostensibly conveys the message that *"the difference in North American revenue between 2007 and 2008 is much larger than the difference in revenue between 2007 and 2008 for the other areas listed"*. This message is generalized into our *Gap-Comparison* message category.

Secondary Messages

Grouped bar chart information graphics are more complex than simple bar charts and line graphs, because of the additional "grouping" dimension in grouped bar charts. We found during our corpus analysis that many grouped bar charts in popular media have *multiple* high-level messages, but often there is one single *primary* message and the possibility of a *secondary* message which though present, is not as easily recognizable as the *primary* message. Our message recognition system for grouped bar charts recognizes both primary and secondary messages.

Communicative Signals

Graph designers use communicative signals in grouped bar charts in order to help convey their intended messages. These design choices can make an entity or a set of entities salient. In this section, we present some of the common communicative signals that are found in grouped bar charts. Our automated system automatically extracts these signals from the graphic and probabilistically reasons about them to hypothesize the graphic's intended messages.

Coloring of a Bar Coloring or highlighting can make an entity salient compared to other entities in the graphic. Figure 2 shows a graphic from *Time*, where the '04 bar in the first group is colored differently from the '04 bars in the other groups, thereby drawing attention to the increased instruction on reading, in contrast with the decreased instruction for other subjects.

Salient Height A dramatic difference in height between one entity and the other entities can make an entity salient. For example in Figure 2, the "Reading" entity is salient because it is considerably taller than the other groups.

Salient Gap The difference in the size of a "gap" of an entity compared to all other entities can make an entity salient. In Figure 3, the difference in revenue for the group "N. America" is much larger than the other groups, drawing increased attention on it.

Positioning The positioning of a group or series of bars can make an entity salient. For example in Figure 2, the group "Reading" is salient by its position as the first group in the graphic. Reading is listed first and is not part of a natural ordering of the groups (such as an alphabetical ordering) .

Recency In grouped bar charts with a time dimension (such as Figure 1), the entity that is most recent may have increased salience to a graph viewer.

Textual Evidence Although Elzer (Elzer et al. 2005) observed that captions in popular media very often do not capture a graphic's intended message, captions often *do* contain linguistic signals that help convey the message. We observed two kinds of linguistic signals in grouped bar charts: (1) verb signals and (2) linguistic structure signals.

Obama captured first–time voters, but
Clinton was stronger among older voters

Clinton   **65 and older**   48%
          18 to 24   22%
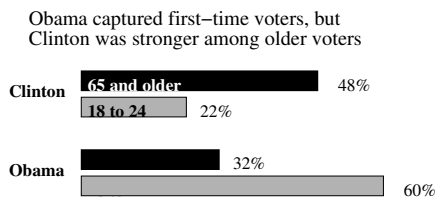
Obama   32%
          60%

Figure 4: Graphic from Time Magazine, January 21, 2008.

Verbs can signal one of more high-level message categories. In one graphic in our corpus, the verb "shrinking" in the caption "Shrinking Giants", suggests a *Falling-Trends-All* or *Falling-Trends-Mostly* message category, both of which are "Trend Messages".

Second, the linguistic structure of caption text can also make specific entities salient in the graph. Consider the caption for Figure 4, "Obama captured first-time voters, but Clinton was stronger among older voters", which mentions three of the four graphed entities ("older", "Clinton", and "Obama"). Each of these entities is mentioned once in the caption in independent clauses, both "Obama" and "Clinton" are in subject position, and "older" is in object position. However, "Clinton" is also in a contrastive clause that is introduced by the conjunction "but", which suggests that "Clinton" is a salient entity that is to be compared.

Effort Another communicative signal is the relative cognitive effort required to recognize some message given the design and layout of the graphic. This signal follows Green et al. (Green et al. 2004) who hypothesized that graphic designers construct graphics that facilitate as much as possible the tasks that the graph viewer will need to perform to understand the graphic's message.

We have built a cognitive model (Burns, Carberry, and Schwartz 2013) in the ACT-R modeling framework that outputs a relative estimate of the perceptual effort required to recognize the message.

Role of Communicative Signals in Message Recognition

We have built a Bayesian network graphical model (Burns 2012)—with a single top level node (representing the possible intended message of the graphic), a set of communicative evidence node leaves, and intermittent levels—that automatically recognizes the intended messages of a grouped bar chart. The conditional probability tables in the Bayesian network are trained using our annotated corpus of graphs.

Using leave-one-out cross-validation, the overall system accuracy for recognizing the intended primary message of a grouped bar chart is 65.5%, which far exceeds a baseline accuracy of 18.8% that results from selecting the most commonly occurring possible message.

As noted earlier, it is common for grouped bar charts to have both a primary and a secondary message. The system's accuracy for correctly predicting either the primary or secondary message of a graphic within the top two messages that it hypothesizes is 88.5%, compared to a baseline

of 29.7% when no evidence is entered into the system.

We conducted two experiments to discover the effect that each type of of communicative evidence had on the accuracy of our system.

1. an experiment that evaluates the system while only one type of communicative evidence is included (Table 1)

2. an experiment that observes degradation in the system's performance when a particular source of communicative evidence is removed (Table 2)

Table 1 shows 14 separate cross-validation runs of our system. Each run is a system configuration with either no communicative evidence or only the listed evidence source. We used a one-tailed non-parametric McNemar test, which is often used on nominal, matched-pair data to show the statistical significance of change. Each run is compared to hypothesis H_1, which is the system configuration with no communicative evidence.

Some of our individual types of communicative evidence require more explanation. The *Effort-Inter* type of communicative evidence is our cognitive model's output of the relative effort required to recognize some particular message in a graphic, with respect to all possible messages (e.g. "does this graph have a trend that is very difficult to recognize?"); in contrast the *Effort-Intra* type of evidence is the cognitive model's output of relative effort with respect to only the given message, (e.g. "does this graph have a trend that is very difficult to recognize compared to most trends?"). The *All Text* evidence includes both evidence from verb signals and the linguistic structure. The *Generalized* type of communicative evidence captures evidence about the properties of the graphic's data rather than its visual features, such as the number of groups in the graph, the number of series in the graph, whether group entities are ordinal and presented in a sorted order, etc. Finally, the *Generalized Instantiated* type of evidence includes properties such as the number of contrasting groups in the graphic.

We were somewhat surprised at the results of our system when communicative signals are individually added. Despite our previous analysis (Carberry, Elzer, and Demir 2006) which observed that a graphic's message is often not contained in the graphic's caption or in the article accompanying the graphic, the communicative evidence of the *Linguistic Structure* was very helpful when we added it into the system. This suggests that there is a considerable amount of beneficial information that can be automatically extracted from text sources (such as if one entity is frequently mentioned). Previously, we also hypothesized that effort communicative evidence would be very valuable in the grouped bar chart system, because it had a large effect for the simple bar chart system (Carberry and Elzer 2007). This is also reflected in the results of our experiment.

The inclusion of some evidence sources actually degraded performance compared to the baseline for the *Top* criterion, which suggests that multiple types of evidence are required to hypothesize the primary message of a graphic.

In Table 2, we present another 14 runs, where in each run the listed evidence source has been removed. The experiment results suggest that although communicative signals

Table 1: Impact of evidence in the Bayesian System, measured by individually adding different sources as the sole piece of evidence. **Top**- system accuracy measured by correctly predicting primary message of a graphic; **Top2**- system accuracy measured by correctly predicting either the primary or secondary message of a graphic within the top two messages that it hypothesizes.

Key: $* \ p < .05 \quad ** \ p < .01 \quad *** \ p < .001$ (one-tailed, matched pair McNemar test)

	Top	Top2
No Evidence (baseline)	18.8%	29.7%
Only Effort	16.7%	40.6% ***
Only Effort-Inter	11.8%	33.9%
Only Effort-Intra	16.7%	46.4% ***
Only Text	60.0% ***	85.2% ***
Only Verb Signals	21.8%	34.8% *
Only Linguistic Structure	33.6% ***	66.1% ***
Only Coloring	20.9% *	32.4% **
Only Salient Height	18.2%	33.6% **
Only Salient Gap	14.8%	28.2%
Only Positioning	13.6%	31.8%
Only Recency	15.8%	30.9%
Only Generalized	49.1% ***	73.6% ***
Only Generalized Instantiated	12.4%	32.1%

Table 2: Impact of evidence measured by separately removing different pieces of evidence.

	Top	Top2
All Evidence (baseline)	65.5%	88.5%
Remove All Effort	66.7%	90.6%
Remove Effort-Inter	65.2%	87.6%
Remove Effort-Intra	65.5%	88.5%
Remove All Text	55.8% ***	82.4% ***
Remove Verb Signals	64.5%	87.3%
Remove Linguistic Structure	56.1% ***	83.0% **
Remove Coloring	63.0%	86.7%
Remove Salient Height	63.0%	87.6%
Remove Salient Gap	63.9%	86.7%
Remove Positioning	63.0%	87.3%
Remove Recency	64.5%	87.6%
Remove Generalized	59.1% ***	84.8% *
Remove Generalized Instantiated	63.6%	85.8% *

contribute to the recognition of a graphic's messages, it is possible that the absence of one kind of communicative signal is compensated by the presence of other communicative signals. The only communicative signals that do not appear to be compensated by other signals are the *Generalized* evidence signal and the *Linguistic Structure* signal. Interestingly, both of these types of evidence capture properties of the graph and its surrounding text, rather than aspects of the visual choices made by the graphic designer to make an entity or set of entities more salient.

Conclusion

This paper has presented our analysis of the effect that various types of communicative signals have in our system for correctly hypothesizing the intended messages of a grouped bar chart. Our current and future work is motivated by the results presented here. We are beginning to investigate other types of information graphics (pie charts, multiple line graphs), and posit that the kinds of communicative evidence that have the greatest effect on message recognition for grouped bar charts will also be important for these other types of information graphics.

References

Burns, R.; Carberry, S.; and Schwartz, S. 2013. Modeling a graph viewer's effort in recognizing messages conveyed by grouped bar charts. In *Proceedings of the 21st International Conference on User Modeling, Adaptation, and Personalization*, 114–126.

Burns, R. 2012. *Automated Intention Recognition of Grouped Bar Charts in Multimodal Documents*. Ph.D. Dissertation, University of Delaware.

Carberry, S., and Elzer, S. 2007. Exploiting evidence analysis in plan recognition. In *International Conference on User Modeling*, 7–16.

Carberry, S.; Elzer, S.; and Demir, S. 2006. Information graphics: An untapped resource of digital libraries. In *Proceedings of 9th International ACM SigIR Conference on Research and Development on Information Retrieval*, 581–588. New York, NY: ACM.

Demir, S.; Oliver, D.; Schwartz, E.; Elzer, S.; Carberry, S.; and McCoy, K. F. 2010. Interactive sight into information graphics. In *Proceedings of the 2010 International Cross Disciplinary Conference on Web Accessibility (W4A)*, 16:1–16:10. New York, NY, USA: ACM.

Elzer, S.; Carberry, S.; Chester, D.; Demir, S.; Green, N.; Zukerman, I.; and Trnka, K. 2005. 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*, 223–230.

Elzer, S.; Carberry, S.; and Zukerman, I. 2011. The automated understanding of simple bar charts. *Artificial Intelligence* 175(2):526–555.

Green, N. L.; Carenini, G.; Kerpedjiev, S.; Mattis, J.; Moore, J. D.; and Roth, S. F. 2004. 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.

Iverson, G., and Gergen, M. 1997. *Statistics: The Conceptual Approach*. New York: Springer-Verlag.

Wu, P.; Carberry, S.; Elzer, S.; and Chester, D. 2010. Recognizing the intended message of line graphs. In *Proceedings of the 6th International Conference on Diagrammatic Representation and Inference*, 220–234.