Automatically Recognizing Intended Messages in Grouped Bar Charts

Richard Burns[1], Sandra Carberry[1], Stephanie Elzer[2], and Daniel Chester[1]

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

Abstract. Information graphics (bar charts, line graphs, grouped bar charts, etc.) often appear in popular media such as newspapers and magazines. In most cases, the information graphic is intended to convey a high-level message; this message plays a role in understanding the document but is seldom repeated in the document's text. This paper presents our methodology for recognizing the intended message of a grouped bar chart. We discuss the types of messages communicated in grouped bar charts, the communicative signals that serve as evidence for the message, and the design and evaluation of our implemented system.

1 Introduction

Information graphics are non-pictoral graphics that display information, such as bar charts, line graphs, grouped bar charts, and pie charts. The purpose of an information graphic in popular media—national and local newspapers *(USA Today, Philadelphia Inquirer)* and magazines *(Time, Newsweek)*—is usually to communicate a high-level contextual message to the graph viewer, as opposed to merely displaying data for analysis.

Grouped bar charts are a type of information graphic. They are similar to *simple* bar charts in that they visually display quantifiable relationships of values; however they contain an additional *grouping* dimension. Despite this additional complexity, they still convey intended high-level messages. For example, the grouped bar chart in Figure 1 ostensibly conveys the high-level message that *"China has a greater rate of software piracy than the rest of the world."*

Clark [6] noted that language is more than just words, but rather is any "signal" or lack thereof, where a signal is a deliberate action that is intended to convey a message, such as gestures and facial expressions. We can view information graphics as a form of language. This paper presents a methodology for automatically reasoning about the most likely intended message of a grouped bar chart using their communicative signals as evidence; that is, we predict the graphic designer's high-level intention in designing the graphic.

Carberry [4] studied graphics from popular media and observed that the graphic's message was very often not repeated in the caption or headline, nor

P. Cox, P. Rodgers, and B. Plimmer (Eds.): Diagrams 2012, LNAI 7352, pp. 8–22, 2012.

in any article text. Thus, it is infeasible to perform only natural language processing techniques on the caption and headlines of the graphic and expect to consistently recognize high-level messages.

Some research has already considered the communicative intent of information graphics. Kerpedjiev et al. [14] proposed a methodology for automatically generating graphics that realize desired intentions. Fasciano [11], in the Postgraphe system, generated graphics based on the input of a communicative intention and a data set. Mittal [16], in the SAGE system, implemented a process which automatically generates captions which can be used to explain data in novel or creative graphics. Although the concept of generating good captions bears some similarity to identifying the intended message of a graphic, Mittal is given the data points that will be displayed and the communicative goal of the graphic. In our work, the communicative goal must be inferred by reasoning about the communicative signals in the graphic.

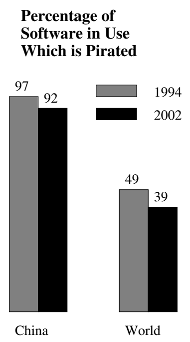

Fig. 1. From *NewsWeek*, “Microsoft Cozies Up to China”, June 28, 2004

Both Elzer [9] and Wu [21] have implemented systems which automatically recognize the most likely high-level message in simple bar charts and line graphs, respectively. However, grouped bar charts are much more complex than simple bar charts and line graphs; thus they convey a much richer and varied set of messages, the kinds of communicative signals are different, and inferring the intended message requires more complex reasoning.

At least three applications can greatly benefit from this research. The first is a system which provides sight-impaired individuals with alternative access to information graphics in multimodal documents by conveying the high-level content of its intended message via speech. The second is to use a graphic's intended message to index it for retrieval from a digital library. The third is to use the intended message of a graphic as its high-level content and take it into account during the summarization of a multimodal document. Our colleagues are actively investigating all three applications.

Section 2 describes our grouped bar chart collection, the identification of the types of high-level messages that graphic designers overwhelmingly convey in grouped bar charts, and the annotation of our corpus. Section 3 describes the communicative signals that appear in grouped bar charts. Section 4 presents our implemented Bayesian reasoning framework, describes how the extracted communicative signals are used as evidence in inferring the intended message of a grouped bar chart, and discusses the system's evaluation. Section 5 discusses future work motivated by the inherent additional complexity in grouped bar charts.

2 Messages

Our corpus is a collection of 222 grouped bar charts from popular media (mainstream newspapers and magazines).[1] We analyzed the corpus to identify the types of high-level messages that graphic designers communicate using grouped bar charts and generalized these into message categories. This section discusses our identified high-level message categories and presents examples from our grouped bar chart corpus; Table 1 lists all of the message categories, and their constraints and instantiated parameters.

2.1 Messages

Trend Messages. Trend messages convey a general trend (rising, falling, or steady) over a set of ordinal data points. For example, the grouped bar chart in Figure 2 ostensibly conveys the high-level message that *"China increased spending on education, social security, military, and rural support from 2004 to 2006."*, a *Rising-Trends-All* message category. Note that trends can be **within-groups** in which case each group of bars comprises a data series or **across-groups** (as in Figure 2 with the i^{th} bar in each group comprising the i^{th} data series). Table 1 shows that the *Rising-Trends-All* message category requires at least 3 data points for the trend and that data series be over a set of ordinal entities. The trends hold for each *series* and the overall message of Figure 2 can be represented as: *Rising-Trends-All(across-groups: {Education, Social security, Military, Rural support})*.

Relationship Messages. Relationship messages capture the consistency of the relative values for a set of entities, or the inconsistency of one set of relative values with respect to the other sets. For example, the grouped bar chart in Figure 8 ostensibly conveys the high-level message that *"The increased funding to Life Sciences is in contrast to the steady or decreased funding to the other research areas."*, a contrasting message that we can represent with the message category *Entity-Relationship-Contrast*. As with trend messages, the set of entities may be **within-groups**, as in Figure 8, or **across-groups** (the i^{th} bar from each group). The parameter <i> as listed in Table 1 is instantiated with the contrasting entity, in this case: *1st group (Life Sciences)*. Thus for Figure 8, the intended message is *Entity-Relationship-Contrast(within-groups:{Life Sciences, Psychology, ..., Other}, 1st group: Life Sciences)*.

Gap Messages. Gap messages recognize a high-level message involving either one *gap*, or a trend in the size of multiple *gaps*, where a *gap* is the approximate absolute difference between two values within the same entity. For example, the grouped bar chart in Figure 3 ostensibly conveys that *"There is an increasing gap between the number of patents filed and the number of patents issued, over the period from 1994 to 2003."*, and can be represented as *Gap-Increasing(across-groups:{'94,'95,...,'03})*. Figure 4 shows a *Gap-Crossover* message: *"The gap*

[1] The corpus is available online at http://www.cis.udel.edu/~burns/corpus

Table 1. Our grouped bar chart analysis identified numerous messages that graphic designers convey via grouped bar charts. Message categories have constraints (for example, a trend must exist over at least three entities) and one or more parameters which are instantiated. Constraints Key: O (ordinal entities are required), 3+ (at least three entities are required), 2 (two entity limit).

Message Category <Parameter(s)>	**Constraints**	**Gloss**
Rising-Trends-All <p>	O, 3+	There is the same trend (rising, falling, steady, or changing) for all entities in the <p> data series where <p> is "within groups" or "across groups".
Falling-Trends-All <p>		
Steady-Trends-All <p>		
Changed-Trends-All <p>		
Opposite-Trends <p>	O, 2	The two entities have opposite trends in the <p> data series where <p> is "within groups" or "across groups".
Contrast-Trend <p, i>	O, 3+	The <i>th trend is contrasting to all of the other trends in the <p> data series where <p> is "within groups" or "across groups".
Rising-Trends-Mostly <p>		There is some trend (rising, falling, steady) for a majority but not all entities in the <p> data series where <p> is "within groups" or "across groups".
Falling-Trends-Mostly <p>		
Steady-Trends-Mostly <p>		
Same-Relationship-All <p>		Each entity in the <p> data series has the same relative ordering of bar values where <p> is "within groups" or "across groups".
Opposite-Entity-Relationship <p>	2	The two entities in the <p> data series have a different relative ordering of bar values where <p> is "within groups" or "across groups".
Entity-Relationship-Contrast <p, i>	3+	The <i>th entity has a contrasting relative ordering of bar values compared to all of the other entities in the <p> data series where <p> is "within groups" or "across groups".
Same-Relationship-Mostly <p>		The majority but not all entities in the <p> data series have the same relative ordering of bar values where <p> is "within groups" or "across groups".
Gap-Increasing <p>	O, 3+	The gap between two entities is trending (increasing, decreasing) over the <p> data series where <p> is "within groups" or "across groups".
Gap-Decreasing <p>		
Gap-Crossover <p>		The gap between two entities decreased so much that one entity caught-up-to and then crossed the other where <p> (the sequence of gaps) is "within groups" or "across groups".
Gap-Comparison-Single <p, i>	3+	The gap in the <i>th entity in the <p> data series is being compared to the gaps of the other entities where <p> is "within groups" or "across groups".
Gap-Comparison-Pair <p>	2	The two gaps in the <p> data series are being compared with each other where <p> is "within groups" or "across groups".
Entity-Comparison <p, i>		The <i>th entity in the <p> data series is being compared to the other entities where <p> is "within groups" or "across groups".
Rising-Entities-All <p>	O, 2	All entities in the <p> data series have two data points and are (rising, falling, steady) where <p> is groups of series.
Falling-Entities-All <p>		
Steady-Entities-All <p>		
Rising-Entities-Mostly <p>		The majority but not all entities in the <p> data series have two data points and are (rising, falling, steady) where <p> is groups of series.
Falling-Entities-Mostly <p>		
Steady-Entities-Mostly <p>		

between the number of Internet users in the US and the number in China has steadily decreased until now China has more Internet users than the US." As we observe in Table 1, the *Gap-Increasing* and *Gap-Crossover* message categories require similar data constraints to the *Rising-Trend*: namely that there are at least three data points and that the trending is over a set of ordinal entities.

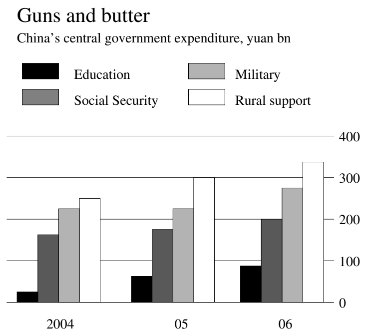

Fig. 2. Graphic from *The Economist*, "Planning the new socialist countryside", March 9, 2006

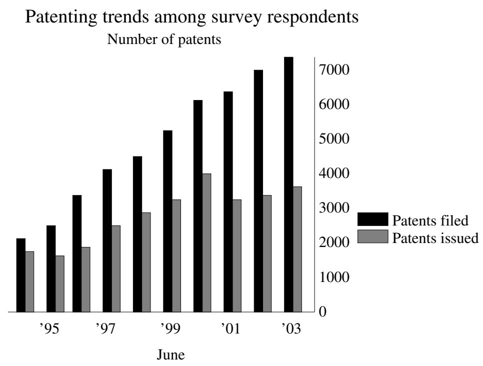

Fig. 3. Graphic from *Technology Review*, "A Mixed Bag of U.S. Institutions", July 2005

Some gap messages compare the gaps in one entity with the gaps of the other entities. For example, Figure 5 ostensibly conveys 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 can be represented as *Gap-Comparison-Single(within-groups: {N America, Europe, Latin America, Asia Pacific}, 1st group: <N America>).*

Entity Comparison Messages. Entity comparison messages compare one entity against the other entities. For example, the grouped bar chart in Figure 1 ostensibly conveys that *"China has a greater rate of piracy than the rest of the world."*. This is captured by the *Entity-Comparison* message category and is represented as *Entity-Comparison(within-groups:{China, World}, 1st group: China).*

Additional Messages. Space limitations preclude us from describing all of our 25 message categories listed in Table 1.

2.2 Annotation

Coders individually annotated each graphic in the corpus with the high-level message that it conveyed by determining its message category and the instantiation of its parameter.[2] Where there was disagreement, the coders discussed the graphic until a consensus was reached.

[2] Only a small number of grouped bar charts did not contain an intended message.

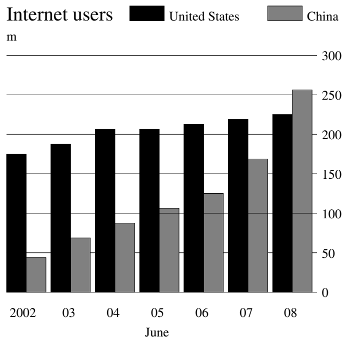

Fig. 4. From *The Economist Daily Chart*, July 31, 2008

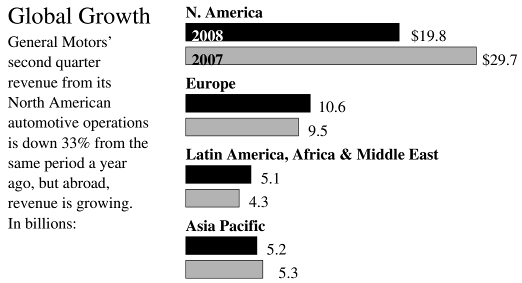

Fig. 5. Graphic from *Wall Street Journal*, "GM to Build Diesel Engines in Thailand", August 14, 2008

3 Communicative Signals

Graph designers use communicative signals in grouped bar charts in order to help convey their intended message. This section describers the kinds of communicative signals found in grouped bar charts. These signals are exploited as communicative evidence in the intention recognition system presented in Section 4.

3.1 Salience via Visual Signals

Visual signals are often used by graphic designers to make an entity (a bar or set of bars) salient. This suggests that the salient entity is an important part of the graphic's intended message—in our terminology, it should be an instantiation of the <i> parameter in the message categories of Table 1.

An entity can be made salient by coloring it differently from the other entities in the graphic. Figure 6 shows a graphic from *Time*, where coloring creates salience. Here 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.

Sets of bars can become salient based on their position in the graphic. For example, in Figure 8 the group "Life Sciences" is salient by virtue of its leading position which is not part of a natural (such as alphabetical) ordering of the groups.

A dramatic difference in height between one entity and the other entities can make an entity salient. The "Life Sciences" entity in Figure 8 also jumps out because it is so much taller than the other groups.

3.2 Linguistic Signals

Although Elzer [7] observed that captions in popular media very often do not capture a graphic's intended message, captions often contain linguistic signals that help convey the message. We observed two kinds of linguistic signals in grouped bar charts: verb signals and linguistic structure signals.

Certain kinds of verbs can signal one or more high-level message categories. For example, in the caption for Figure 7, "Shrinking Giants", the verb *shrinking* suggests the message categories: *Falling-Trends-All*, *Falling-Trends-Mostly*, *Falling-Entities-All*, or *Falling-Entities-Mostly*.

The linguistic structure of the caption can signal the salience of a *specific entity*. For example, in the caption for Figure 9, "Obama captured first-time voters, but Clinton was strong among older voters", three of the four graphed entities (Obama, Clinton, older) are mentioned. They are each mentioned once in the caption in independent clauses; Obama and Clinton are both in subject position; older is in object position; however, Clinton is in a contrastive clause introduced by "but". This suggests that *Clinton* is a salient entity that is to be compared.

3.3 Relative Perceptual Effort as a Communicative Signal

Green et al. [12] 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. Thus, following Elzer [10] we view relative perceptual task effort as a communicative signal: messages that require more perceptual effort than others are less likely to be the message that the graph designer intended to convey. This correlates with Larkin and Simon [15] who observe that informationally equivalent graphics are not necessarily computationally equivalent, and Peebles and Cheng [17] who note that seemingly minor design changes can greatly affect performance on graph reading tasks.

For example in Figure 10, although both graphics contain the same data, individually they convey two different messages. The high-level message conveyed by the left graphic is ostensibly that male salaries are greater than female salaries in all of the subject areas, while the message conveyed by the right graphic is ostensibly a message of rank: that engineering and the physical sciences have the greatest salaries for both men and women. While this information can be obtained from either graphic, the design of the graphic affects the perceptual effort required and thus the intended message of the graphic.

We have built a cognitive model which produces a relative estimate of the perceptual effort required given a message and graphic, which is also considered as communicative evidence in the intention recognition system. This model is built in the ACT-R [2] cognitive framework, following the ACT-R theory as well as graph comprehension work from the psychological literature. Pilot eye-tracking experiments, in which we asked human subjects to perform specific graph tasks on grouped bar charts and subsequently analyzed their eye scan patterns and fixation and attention locations, also helped us identify the factors in grouped bar charts which affect the required recognition effort.

Pinker [18] identified high-level visual patterns such as linear lines and quadratic curves which are easily identifiable for most graph viewers. In our pilot experiments, we found that subjects fixated less on sets of entities whose

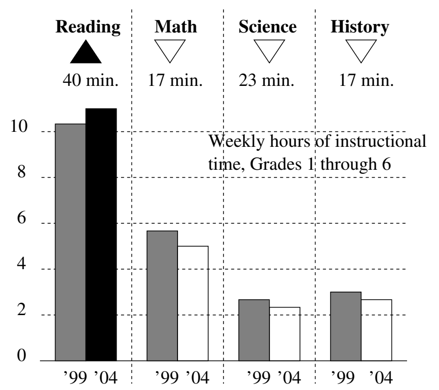

Fig. 6. Graphic from *Time*, "How to Fix No Child Left Behind", June 4, 2007

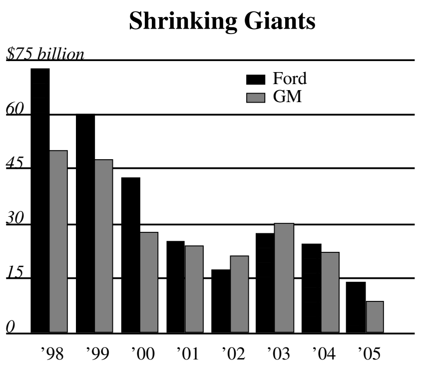

Fig. 7. Graphic from Graphic from Wall Street Journal, "Auto Industry, at a Crossroads, Finds Itself Stalled by History", January 2, 2006

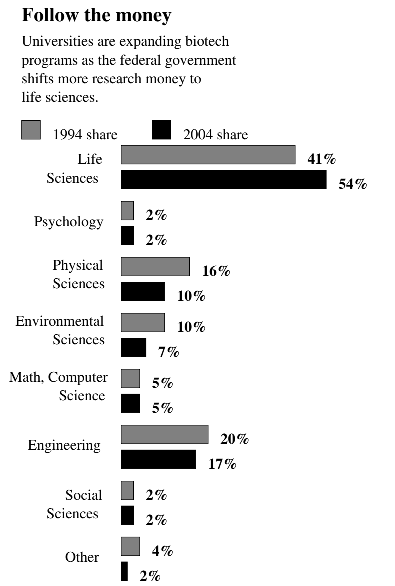

Fig. 8. Graphic from *USA Today*, "Universities grid for battle for biosciences supremacy", June 24, 2005

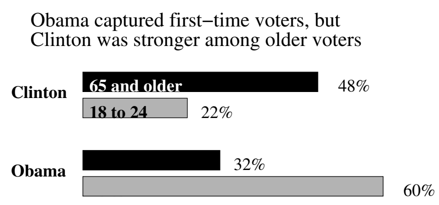

Fig. 9. Graphic from Time Magazine, January 21, 2008

bar heights resembled an easily identifiable visual pattern. Shah [19] noted that the grouping of data points will influence the perceived pattern recognition of trends and in our pilot experiments, subjects could still perceive trends despite the presence of *exceptions* (a data point which does not follow the trend). Peripheral vision—the ability for multiple objects to be processed in parallel in a guided search [1]—was present in our experiments: subjects showed the ability to perform some graph tasks without fixating on the first or last groups. Wickens and Carswell [20] defined the *proximity compatibility principle* and showed how close perceptual proximity is advised (the perceptual similarity of two elements) if and only if closeness in processing proximity is intended (the extent to which elements are used as part of the same task). We observed an increase in the time for subjects to perform graph tasks for grouped bar charts with increased noise

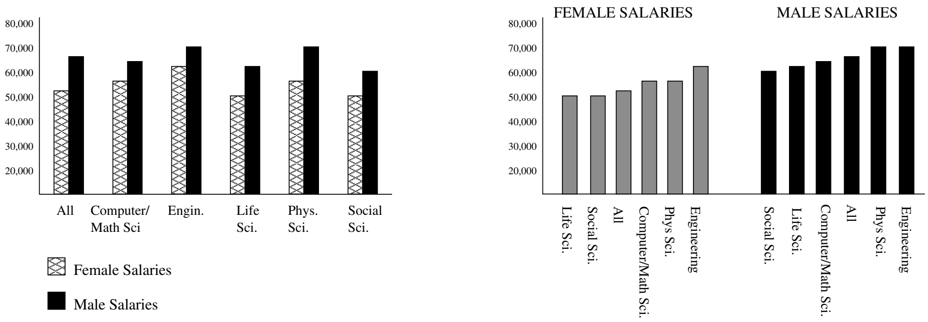

Fig. 10. Two bar charts designed from the same data

and visual clutter, where visual clutter is the close spatial proximity of two perceptually or semantically contrasting elements which should not be compared.

The design of our model of relative perceptual effort incorporates these factors so that the presence of high-level visual patterns in a graphic enables the model to process a graph quicker while the presence of visual clutter and exceptions cause an increase in processing time.

Our model was validated for a subset of our message categories in an initial experiment [3] where the relative time required for our model to perform graph tasks on a range of graphics was compared with the relative average for human subjects to perform the same tasks with the same graphics. Our current work includes validating our complete model for all message categories.

4 Recognizing the Intended Message

4.1 System Architecture

Our system for recognizing the intended message of a grouped bar chart requires an XML representation of a graphic which specifies each bar, each series and group of bars, their heights, colors, annotations, the axes labels, caption, etc. This is the responsibility of a visual extraction module [5]. The generated output is similar to Huang [13] who in addition to representing the graphic also considers vision issues such as identifying an information graphic, locating an information graphic within a noisy pdf document, and performing OCR on the text within the graphic.

Our intention recognition system for grouped bar charts is modeled with a Bayesian reasoning framework which captures the relationship between communicative signals and intended messages. Figure 11 shows the general structure of our Bayesian network. There is an *Intended Message* node at the top whose states are either message categories that still need to be further instantiated or messages that can only have *within-groups* or *across-groups* as their possible instantiation.[3] The five most prevalent states in our corpus are shown along

[3] Thus, we can effectively treat these as message categories in our design.

with their a priori probabilities before any evidence is entered into the network. The communicative signals for a graphic are the evidence for or against possible messages. These are represented in the leaves of the network as evidence nodes.

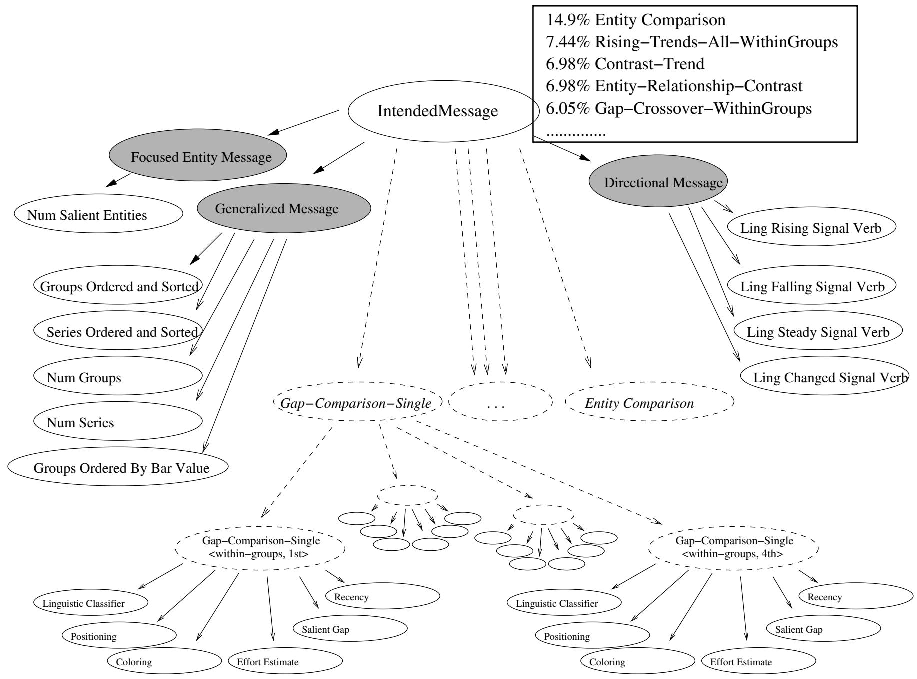

Fig. 11. Network structure which captures the probabilistic relationship between intended messages and communicative evidence

The grey nodes in Figure 11 between the top-level and leaves are deterministic, generalizing nodes which group together message categories that have a general feature in common. This alleviates the data sparseness problem that arises due to the limited size of our corpus. For example, the *Directional Message* node groups together message categories which convey the same direction, such as *Rising-Trends-All*, *Rising-Trends-Mostly*, *Rising-Entities-All*, *Rising-Entities-Mostly*, into the *Rising Messages* category. Then, the child evidence nodes (*Ling Rising*, *Ling Falling*, etc.) capture the probabilistic relationships between a *Rising Message* and the "signal verb" communicative signals. Similarly, the *Focused Entity Message* node generalizes those message categories that focus on a *specific entity (<i>)* (such as *Contrast-Trend* and *Entity-Comparison*) and captures the probabilistic relationship between the evidence node *Num Salient Entities* which represents how many entities in a graphic are salient. Finally, the *Generalized Message* node categorizes those message categories together (such as *Rising-*

Trends-All and *Falling-Trends-All*) for which we expect naive evidence (such as the number of bars per group) to have the same affect on each.

Message categories at the top level that have a parameter instantiation besides *within-groups* or *across-groups* are instantiated with a *specific entity* lower in the network. In this case, evidence nodes appear as children for each possible instantiation and are able to capture evidence that is only relevant for a specific parameter instantiation. For example, in Figure 11 the *Gap-Comparison-Single* message category node is instantiated for every possible group entity instantiation in Figure 5. The value for the *Salient Gap* evidence node will be positive only for the instantiated node *Gap-Comparison-Single(within-groups, 1st).*

4.2 Extracting Evidence

The evidence provided by visual communicative signals (such as whether a group of bars is colored differently from the other bars) is automatically extracted from the XML representation of a graphic and entered into evidence nodes. The text of accompanying captions and headlines is extracted and parsed to identify the presence of any signal verbs.

The extraction of linguistic structure signals is more complex. Headlines and captions are parsed into their clausal structures. A support vector machine was trained on our corpus of captions to produce a learned model that decides which of several mentioned entities is most linguistically salient. It uses features such as the frequency with which the entity is mentioned, the source of the mentioning (main headline, caption, etc.), the ordering of mentions (is one entity preceding), subject position, object position, main/subordinate clausal structure. The decision of the model is entered into the *Linguistic Classifier* evidence nodes.

The relative perceptual effort model takes the XML representation and outputs a relative time for the expected recognition of some message. This relative time is discretized and is entered into the *Effort* evidence nodes.

4.3 Training

Associated with each node in the Bayesian network is a conditional probability table that captures the probability of each value for the node given the values for its parent nodes. The conditional probability tables are learned from our corpus of graphics. The Bayesian network applies Bayes' rule to the network constructed for a new graphic to propagate the evidence through the network and compute the posterior probability for each node. Table 2 shows the conditional probability table that captures the probabilistic relationship between the high-level *Gap-Comparison-Single*(within-groups) message category and any visual gap salience of the instantiated <i> entity or the other *within-group* entities.[4]

[4] The *Gap-Comparison-Single*(within-groups) message category generalizes the *Gap-Comparison-Single(within-groups,<i>) messages where <i> is a group entity.*

Table 2. Learned conditional probability table for the Gap Salience evidence node under the *Gap-Comparison-Single(within-groups)* message category

Gap Comparison-Single <WithinGroups, i>	This Entity Only	This Entity Plus Others	Other Entities Only	No Entities
Intended	58.3%	24.9%	8.3%	8.3%
Not Intended	6.9%	14.8%	65.3%	12.9%

4.4 Current Performance and Discussion

We evaluated our system using leave-one-out cross-validation[5], where the XML representation of each of the 222 graphics is in-turn used as a test graphic with the conditional probability tables computed from the other 221 graphs. Results are averaged over all the tests. Currently our system's accuracy rate is 65.6%: that is, the message category and instantiation that the system predicts matches exactly the consensus-based annotation. Table 3 shows our results. As a baseline, we use the message category that appears most often in our corpus; however, note that our system must recognize not only the correct message category but also the instantiated parameters. Our system more than triples the baseline success rate. Although our success rate is lower that that achieved by Elzer or Wu for bar charts and line graphs (78.2% and 73% respectively), grouped bar charts involve more than twice as many message categories and convey far richer messages, making recognition more complex. Note that our success rate improves to 78.6% if we use the top two system hypotheses; this results from grouped bar charts having secondary messages, where occasionally it is difficult to determine which message is primary and which is secondary. (See Future Work).

Table 3. Results of our system

Grouped Bar Chart System		
Average number of possible messages for a grouped bar chart: 20.2		
system	*criteria*	*accuracy*
Grouped Bar Chart System	top message matches annotation	65.6%
Grouped Bar Chart System	either of top 2 messages match annotation	78.6%
Baseline: predict most common possible message	top message matches annotation	20.2%
Other Systems		
Simple Bar Chart System (Elzer) [9]	top message matches annotation	78.2%
Line Graph System (Wu) [21]	top message matches annotation	73.0%

[5] Leave-one-out cross-validation, as opposed to 10-fold cross-validation, was used to mitigate some sparseness issues in the data set.

Table 4. Example showing that removing communicative evidence for Figure 8 affects the network's prediction that *Entity-Relationship-Contrast(within-groups:{Life Sciences, Psychology, ..., Other, 1st group:Life Sciences})* is the intended message

Likelihood Before	**Node**	**Evidence Before**	**Evidence After**	**Likelihood After**
only one piece of evidence removed:				
99.5%	Linguistic Classifier	only entity mentioned	no entities mentioned	94.1%
99.5%	Salient By Height	only entity that is salient by height	no entities salient by height	90.1%
99.5%	Positioning	first entity	neither first nor last	74.3%
evidence removed sequentially, one after another:				
99.5%	Linguistic Classifier	only entity mentioned	no entities mentioned	94.1%
94.1%	Salient By Height	only entity that is salient by height	no entities salient by height	44.2%
44.2%	Positioning	first entity	neither first nor last	1.25%

As an example of a graphic processed by our system, consider the grouped bar chart in Figure 8. The graph is processed by the Visual Extraction Module to produce an XML representation. Our system correctly predicts an intended message of *Entity-Relationship-Contrast (within-groups:{Life Sciences, Psychology,...,Other},1st group:Life Sciences)* with an almost certain probability of 99.5%. Three communicative signals are automatically entered in the evidence nodes for the *Entity-Relationship-Contrast* message category node instantiated with *<p>=within-groups, <i>=1st*, namely that it is the only entity mentioned in the caption, that it is the only entity that is visually salient by height, and that it is positioned first in a set of more than two entities. Table 4 shows how the network's prediction of this message decreases if the graphic were altered to eliminate some of the communicative signals and thus alter the evidence in the Bayesian Network. We see that the height salience and leading position of the entity are very important for the system's hypothesis of this message. The effect of removing *Life Sciences* from the text, and thus changing the evidence in the Linguistic Classifier evidence node, follows our intuition that an intended contrasting message may not always be linguistically salient in accompanying text. In each case, the presence of two other kinds of salience compensates when one kind of salience is removed as shown in the top half of Table 4. The bottom half of Table 4 shows the cumulative affect of removing several communicative signals. As we adjust the evidence entered into the network, the system's confidence in this message as the graph's intended message decreases and the likelihood of other possible messages increases.

5 Conclusion

5.1 Future Work

Sparseness of data and reference resolution issues (such as determining that *USA* refers to *United States*) are two major causes of our system's errors. We are working to address these problems.

Grouped bar charts are much more complex than simple bar charts because of their additional "grouping" dimension. This facilitates the communication of additional high-level messages, such as in Figure 1, *"the rate of Chinese piracy decreased less than the rest of the world"* which supplements our previous observed message that *"China has a greater rate of software piracy than the rest of the world."* Such secondary messages are novel to grouped bar charts and were not observed in the work of simple bar charts and line graphs by Elzer [8] and Wu [21], respectively. In general, secondary messages were not as apparent during the annotation of our corpus and their realization produced more disagreement among the coders. We are currently working on expanding our framework to automatically identify secondary messages in grouped bar charts.

5.2 Summary

We have presented an implemented system which automatically hypothesizes the high-level intended message of a grouped bar chart. To our knowledge, no one has previously investigated the communicative signals in grouped bar charts, the wide variety of messages that grouped bar charts can convey, and a methodology for recognizing these messages. Our system automatically extracts communicative evidence from the graphic and incorporates it as evidence in a Bayesian network that hypothesizes the graphic's intended message. This work has several significant applications: (1) a system which provides sight-impaired individuals with alternative access to information graphics in multimodal documents, (2) indexing and retrieving grouped bar charts in digital libraries, (3) and the summarization of multimodal documents.

References

1. Anderson, J.R., Lebiere, C.: The Atomic Components of Thought. Lawrence Erlbaum Associates, Mahwah (1998)
2. Anderson, J.R., Matessa, M., Lebiere, C.: Act-r: A theory of higher level cognition and its relation to visual attenion. Human-Computer Interaction 12, 439–462 (1997)
3. Burns, R., Elzer, S., Carberry, S.: Modeling relative task effort for grouped bar charts. In: Taatgen, N., van Rijn, H. (eds.) Proceedings of the 31st Annual Conference of the Cognitive Science Society, pp. 2292–2297. Cognitive Science Society, Austin (2009)
4. Carberry, S., Elzer, S., Demir, S.: Information graphics: An untapped resource of digital libraries. In: Proceedings of 9th International ACM SigIR Conference on Research and Development on Information Retrieval, pp. 581–588. ACM, New York (2006)

5. Chester, D., Elzer, S.: Getting Computers to See Information Graphics So Users Do Not Have to. In: Hacid, M.-S., Murray, N.V., Raś, Z.W., Tsumoto, S. (eds.) ISMIS 2005. LNCS (LNAI), vol. 3488, pp. 660–668. Springer, Heidelberg (2005)
6. Clark, H.: Using Language. Cambridge University Press (1996)
7. Elzer, S., Carberry, S., Chester, D., Demir, S., Green, N., Zukerman, I., Trnka, K.: 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, pp. 223–230 (2005)
8. Elzer, S., Carberry, S., Demir, S.: Communicative signals as the key to automated understanding of bar charts. In: Proceedings of the International Conference on the Theory and Application of Diagrams (2006)
9. Elzer, S., Carberry, S., Zukerman, I., Chester, D., Green, N., Demir, S.: A probabilistic framework for recognizing intention in information graphics. In: Proceedings of the International Joint Conference on Artificial Intelligence, pp. 223–230. Association for Computational Linguistics, Morristown (2005)
10. Elzer, S., Green, N., Carberry, S., Hoffman, J.: A model of perceptual task effort for bar charts and its role in recognizing intention. International Journal on User Modeling and User-Adapted Interaction 16, 1–30 (2006)
11. Fasciano, M., Lapalme, G.: Intentions in the coordinated generation of graphics and text from tabular data. Knowledge and Information Systems 2(3) (August 2000)
12. Green, N.L., Carenini, G., Kerpedjiev, S., Mattis, J., Moore, J.D., Roth, S.F.: 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)
13. Huang, W., Tan, C.L.: A system for understanding imaged infographics and its applications. In: Proceedings of the 2007 ACM Symposium on Document Engineering, DocEng 2007, pp. 9–18. ACM, New York (2007)
14. Kerpedjiev, S., Green, N., Moore, J., Roth, S.: Saying it in graphics: from intentions to visualizations. In: Proceedings of the Symposium on Information Visualization (InfoVis 1998). IEEE Computer Society Technical Committee on Computer Graphics, pp. 97–101. IEEE (1998)
15. Larkin, J.H., Simon, H.A.: Why a diagram is (sometimes) worth a thousand words. Cognitive Science 11, 65–99 (1987)
16. Mittal, V.O., Carenini, G., Moore, J.D., Roth, S.: Describing complex charts in natural language: A caption generation system. Computational Linguistics 24(3), 431–467 (1998)
17. Peebles, D., Cheng, P.C.H.: Modeling the effect of task and graphical representation on response latency in a graph reading task. Human Factors 45, 28–45 (2003)
18. Pinker, S.: A theory of graph comprehension. In: Artificial Intelligence and the Future of Testing, pp. 73–126. Lawrence Erlbaum Associates, Hillsdale (1990)
19. Shah, P., Mayer, R.E., Hegarty, M.: Graphs as aids to knowledge construction: Signaling techniques for guiding the process of graph comprehension. Educational Psychology 91, 690–702 (1999)
20. Wickens, C.D., Carswell, C.M.: The proximity compatibility principle: Its psychological foundation and relevance to display design. Human Factors 37, 473–494 (1995)
21. Wu, P., Carberry, S., Elzer, S., Chester, D.: Recognizing the Intended Message of Line Graphs. In: Goel, A.K., Jamnik, M., Narayanan, N.H. (eds.) Diagrams 2010. LNCS, vol. 6170, pp. 220–234. Springer, Heidelberg (2010)