Recognizing the Intended Message of Line Graphs*

Peng Wu[1], Sandra Carberry[1], Stephanie Elzer[2], and Daniel Chester[1]

[1] Computer and Information Science department, University of Delaware
[2] Computer Science department, Millersville University

Abstract. Information graphics (line graphs, bar charts, etc.) that appear in popular media, such as newspapers and magazines, generally have a message that they are intended to convey. We contend that this message captures the high-level knowledge conveyed by the graphic and can serve as a brief summary of the graphic's content. This paper presents a system for recognizing the intended message of a line graph. Our methodology relies on 1)segmenting the line graph into visually distinguishable trends which are used to suggest possible messages, and 2)extracting communicative signals from the graphic and using them as evidence in a Bayesian Network to identify the best hypothesis about the graphic's intended message. Our system has been implemented and its performance has been evaluated on a corpus of line graphs.

1 Introduction

Information graphics are non-pictorial graphics such as bar charts and line graphs. Although some information graphics are only intended to convey data, the overwhelming majority of information graphics in popular media, such as newspapers and magazines, have a message that they are intended to convey. For example, the line graph in Figure 1 appeared in *USA Today* and ostensibly is intended to convey the message that there has been a recent decrease in box office gross revenue in contrast with the preceding rising trend. We contend that a graphic's intended message constitutes a brief summary of the graphic's high-level content and captures how the graphic should be "understood".

This paper presents our methodology for inferring the intended message of a line graph. In previous research[7], we developed a system for identifying the intended message of a simple bar chart. However, line graphs differ from bar charts in several ways that significantly impact the required processing. First, line graphs are the preferred medium for conveying trends in quantitative data over an ordinal independent axis[12]. Second, as our extensive corpus studies demonstrated, the kinds of messages conveyed by line graphs differ from those conveyed by simple bar charts. For example, the line graph in Figure 2 ostensibly is intended to

* This material is based upon work supported by the National Science Foundation under Grant No. IIS-0534948 and by the National Institute on Disability and Rehabilitation Research under Grant No. H133G080047.

A.K. Goel, M. Jamnik, and N.H. Narayanan (Eds.): Diagrams 2010, LNAI 6170, pp. 220–234, 2010.

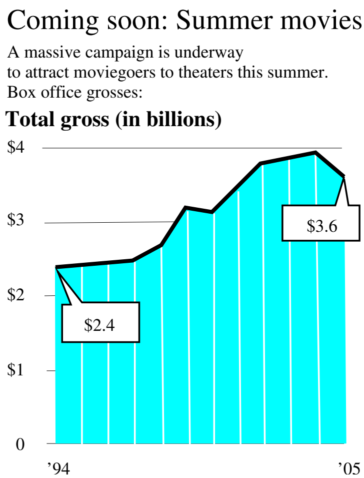

Fig. 1. Line Graph from USA Today

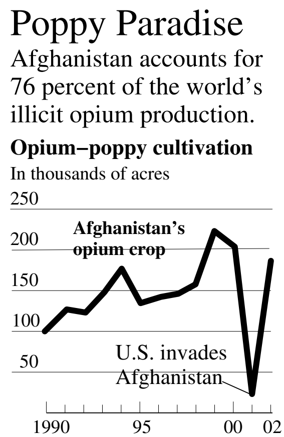

Fig. 2. Line Graph from Newsweek

convey a sudden big drop in Afghanistan's opium crop that is not sustained; in our research, we have not encountered a bar chart that conveys a message of this type. Third, although line graphs and bar charts share some of the same kinds of communicative signals, line graphs use other communicative signals that are not found in bar charts. Fourth, recognition of the message conveyed by a line graph relies on the viewer's ability to perceive it as a sequence of visually distinguishable trends rather than a set of discrete data points. Thus we need a method for identifying these trend segments. Moreover, these latter two factors necessitate a different structure and different processing for the message recognition system than was used for bar charts which relied heavily on perceptual task effort.

We are pursuing several projects that utilize a graphic's intended message. Our digital library project will use the graphic's intended message for indexing and retrieving graphics and for taking graphics into account when summarizing multimodal documents. Our assistive technology research is concerned with providing indiviuals with sight impairments with access to graphics in multimodal documents. Other work has tried to render graphics in an alternative form (such as musical tones or tactile images)[1,18] or as verbal descriptions of the appearance and data points in the graph[8]. We are taking a radically different approach. Rather than describing what the graphic looks like, we provide the user with a brief summary based on the graphic's intended message, along with a facility for responding to followup questions about the graphic.

Section 2 describes our overall architecture. Section 3 presents our approach to recognizing the intended message of a line graph. Section 4 presents the results of an evaluation of our implemented system, and Section 5 presents examples of graphics that have been processed by our system. Section 6 discusses related work, and Section 7 presents our conclusions and discusses future work. To our knowledge, our research group is the only effort aimed at automatically recognizing the communicative goal of an information graphic.

2 System Architecture

Figure 3 shows our overall architecture. A Visual Extraction Module[3] is responsible for analyzing the graphic and providing an XML representation that captures a sampling of the data points (thereby discretizing a continuous line graph into a set of sampled data points), the axis labels, any annotations on the graphic, the caption, etc. The Caption Tagging Module[6] is responsible for extracting evidence from the caption (see Section 3.3) and producing an augmented XML representation that includes it. The Intention Recognition Module takes as input the augmented XML representation of a graphic and uses a Bayesian Network to identify its intended message. The remainder of this paper focuses on the Intention Recognition Module, which is enclosed by a dashed box in Figure 3.

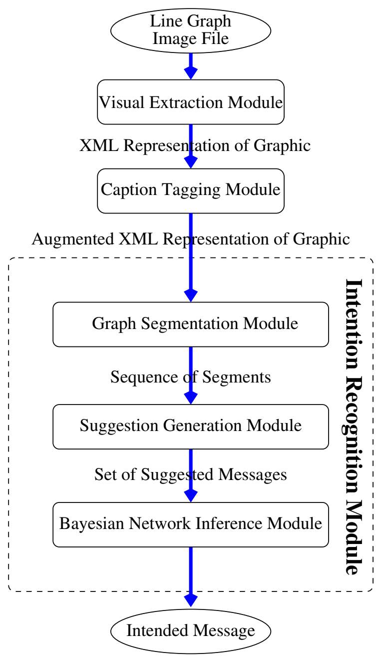

Fig. 3. Overall Architecture

3 Intended Message Recognition

Clark[4] has noted that language is more than just words; it is any deliberate action (or lack of action when one is expected) that is intended to convey a message. Under this definition of language, information graphics in popular media are a form of language with a communicative goal or intended message. In language understanding research, listeners use the communicative signals present in an utterance (such as cue words, intonation, mood, etc.) to deduce the speaker's intended meaning. We are extending this to the understanding of information graphics. Our methodology relies on extracting communicative signals from the graphic and using them as evidence in a Bayesian Network that hypothesizes the graph designer's communicative goal — i.e., the graphic's intended message. Of course, a graphic might be poorly designed, in which case the message that the graphic conveys might not be what the graph designer intended. But this is true of language in general; for example, if a speaker chooses the wrong words or uses the wrong intonation, his utterance will be misunderstood.

Our methodology for recognizing the intended message of a line graph consists of three steps: 1) segment the line graph into visually distinguishable trends, 2) use this segmentation to suggest possible messages for consideration by a Bayesian Network, and 3) extract communicative signals from the line graph and use them as evidence in the Bayesian Network to identify the graphic's intended message. The following sections discuss each of these steps.

3.1 Segmenting a Line Graph into Trends

In recognizing a line graph's intended message, human viewers do not reason about the set of individual data points connected by small line segments. Instead, they appear to treat the line graph as capturing a sequence of visually distinguishable trends. For example, the line graph in Figure 4 consists of many short rises and falls, but a viewer summarizing it would be likely to regard it as consisting of a short overall stable trend from 1900 to 1930, followed by a long rising trend (both with high variance). As observed by Zacks and Tversky[24], this tendency to associate lines with trends exists in part because of cognitive naturalness and in part because of ease of perceptual processing. In comparing bar charts and line graphs, they claim that people "should more readily associate

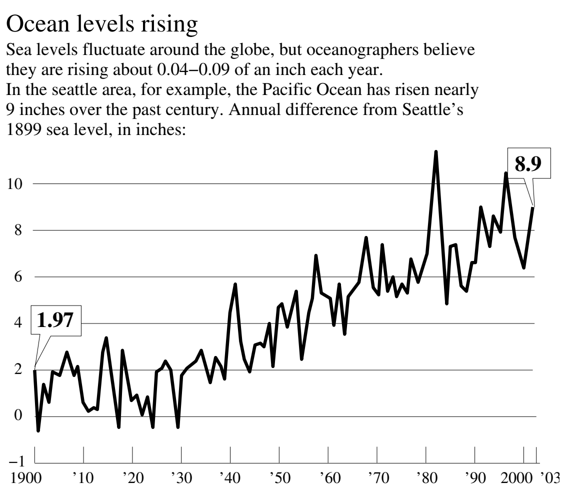

Fig. 4. Line graph from USA Today which consists of many short rises and falls

lines with trends because lines connect discrete entities and directly represent slope" and their experiments uphold this prediction. In fact, the cognitive "fit" of the line graph for representing trends is upheld by multiple findings from the basic Gestalt principles to the Wickens and Carswells' Proximity Compatibility Principle (grouping objects that are meant to be processed together) [22] to Pinker's model of graph comprehension [16]. As Zacks and Tversky noted[24], once this cognitive bias is utilized consistently by graph designers, viewers may come to rely on it. This is consistent with our view of graphics as a form of language with communicative signals. The graph designer is attempting to convey a trend, and is trying to make this message as easy as possible for the viewer to extract from the graph. Over time, the choice of graph type itself may become a type of communicative signal.

Our Graph Segmentation Module takes a top-down approach[11] to identifying sequences of rising, falling, and stable segments in a line graph. It starts with the original graphic as a single segment and decides whether it should be split into two subsegments; if the decision is to split the segment, then the split is made at the point which is the greatest distance from a straight line between the two end points of the segment. This process is repeated on each subsegment until no further splits are identified. The Graph Segmentation Module returns a sequence of straight lines representing a linear regression of the points in each subsegment, where each straight line is presumed to capture a visually distinguishable trend in the original graphic.

We used SMO (Sequential Minimal Optimization)[17] for training a support vector machine that makes a decision about whether a segment should be split. 18 attributes, falling into two categories, are used in building the model. The first category captures statistical tests computed from the sampled data points in the XML representation of the graphic. Two examples of such attributes are:

- Correlation Coefficient: The Pearson product-moment correlation coefficient [19] measures the tendency of the dependent variable to have a linearly rising or falling relationship with the independent variable. We hypothesized that the correlation coefficient might be helpful in determining whether a long set of short jagged segments, such as those between 1930 and the end of the graph in Figure 4, should be captured as a single rising trend and thus not be split further.
- Runs Test: The Runs Test estimates whether a regression is a good fit for the data points[2]. A run is a sequence of consecutive sampled points that all fall above the regression line or all fall below the regression line. The number of runs is then compared with an estimate of the expected number of runs R_{mean} and its standard deviation R_{SD}; if the actual number of runs exceeds $(R_{mean} - R_{SD})$, then the Runs Test suggests that the regression is a good fit and the segment should not be split. We hypothesize that the Runs Test might be helpful when a segment consists of more than two trends.

The second category of attributes captures explicit features of the segment and the graphic. The following is an example of such an attribute:

– Segment Fraction: This attribute measures the proportion of the total graph that comprises this segment. We hypothesize that segments that comprise more of the total graph may be stronger candidates for splitting than segments that comprise only a small portion of the graph.

We trained our graph segmentation model on a set of 649 instances that required a split/no-split decision. These instances were recursively constructed from a corpus of line graphs: each line graph constituted a training instance and, if that instance should be split, then each of the segments produced by splitting represented other training instances. Using leave-one-out cross validation, in which one instance is used for testing and the other 648 instances are used for training, our model achieved an average success rate of 88.29%.

The output of this Graph Segmentation Module is a sequence of segments that are hypothesized to represent visually distinguishable trends. For example, after the Visual Extraction Module converted Figure 4 from GIF format into an XML representation and the data points were sampled, the Graph Segmentation Module then segmented the data series into two visually distinguishable trends: a relatively stable trend from 1900 to 1930 and a rising trend from 1930 to 2003. As another example, the Graph Segmentation Module segmented the data series produced by the VEM for Figure 7 into two visually distinguishable trends: a rising trend from 1997 to 1999 and a falling trend from 1999 to 2006.

3.2 Suggesting Possible Messages

We analyzed a set of simple line graphs collected from various popular media, including magazines such as *Newsweek*, *Time*, and *BusinessWeek* as well as local and national newspapers. We identified a set of 10 high-level message categories that we believe capture the kinds of messages that are conveyed by a simple line graph. Table 1 presents these message categories.

To utilize a Bayesian Network for identifying the intended message of an information graphic, we need a means for suggesting the set of possible messages that should be considered in the network. The Suggestion Generation Module uses the 10 high-level message categories to construct all possible messages from the sequence of segments produced by the Graph Segmentation Module. In addition, we hypothesize that small changes at the end of a line graph, as in Figure 1, may be particularly salient to a viewer, especially if they represent the value of an entity near the current time. However, the Graph Segmentation Module will most likely smooth such small changes into an overall longer smoothed trend. Thus, a short routine using a statistical test is run that examines the end of the line graph and if it represents a change in slope from the preceding points, that short portion is treated as a separate segment. This short segment (if any) is merged with the result produced by the Graph Segmentation Module, and a Contrast-Trend-Last-Segment (CSCT) and a Contrast-Segment-Change-Trend

Table 1. Categories of High Level Messages for Line Graphs

Intention Category	Description
RT: Rising-Trend	There is a rising trend from $<\mathrm{param}_1>$ to $<\mathrm{param}_2>$
FT: Falling-Trend	There is a falling trend from $<\mathrm{param}_1>$ to $<\mathrm{param}_2>$
ST: Stable-Trend	There is a stable trend from $<\mathrm{param}_1>$ to $<\mathrm{param}_2>$
CT: Change-Trend	There is a $<\mathrm{slope}_2>$ trend from $<\mathrm{param}_2>$ to $<\mathrm{param}_3>$ that is significantly different from the $<\mathrm{slope}_1>$ trend from $<\mathrm{param}_1>$ to $<\mathrm{param}_2>$
CTLS: Contrast-Trend-Last-Segment	There is a $<\mathrm{slope}_2>$ segment from $<\mathrm{param}_2>$ to $<\mathrm{param}_3>$ that is not long enough to be viewed as a trend but which is different from the $<\mathrm{slope}_1>$ trend from $<\mathrm{param}_1>$ to $<\mathrm{param}_2>$
CTR: Change-Trend-Return	There is a $<\mathrm{slope}_3>$ trend from $<\mathrm{param}_3>$ to $<\mathrm{param}_4>$ that is different from the $<\mathrm{slope}_2>$ trend between $<\mathrm{param}_2>$ and $<\mathrm{param}_3>$ and reflects a return to the kind of $<\mathrm{slope}_1>$ trend from $<\mathrm{param}_1>$ to $<\mathrm{param}_2>$
CSCT: Contrast-Segment-Change-Trend	There is a $<\mathrm{slope}_3>$ segment from $<\mathrm{param}_3>$ to $<\mathrm{param}_4>$ that is not long enough to be viewed as a trend but which suggests a possible return to the kind of $<\mathrm{slope}_1>$ trend from $<\mathrm{param}_1>$ to $<\mathrm{param}_2>$ which was different from the $<\mathrm{slope}_2>$ trend from $<\mathrm{param}_2>$ to $<\mathrm{param}_3>$
BJ: Big-Jump	There was a very significant sudden jump in value between $<\mathrm{param}_1>$ and $<\mathrm{param}_2>$ which may or may not be sustained
BF: Big-Fall	There was a very significant sudden fall in value between $<\mathrm{param}_1>$ and $<\mathrm{param}_2>$ which may or may not be sustained
PC: Point-Correlation	There is a correlation between changes at $\{<\mathrm{param}_1>, \ldots, <\mathrm{param}_n>\}$ and the text annotations $\{<\mathrm{annot}_1>, \ldots, <\mathrm{annot}_n>\}$ that are associated with these points.

(CSCT) message are proposed for the last two or three segments of the graphic respectively.

Consider, for example, the graphic in Figure 5. The Graph Segmentation Module produces a sequence of three visually distinguishable segments. The Suggestion Generation Module proposes the following 11 possible messages[1]:

RT (5-21-05, 9-1-05)
RT (12-1-05, 4-25-06)
FT (9-1-05, 12-1-05)
BJ (5-21-05, 9-1-05)
BF (9-1-05, 12-1-05)

CT (5-21-05, 9-1-05, 12-1-05)
CT (9-1-05, 12-1-05, 4-25-06)
CTR (5-21-05, 9-1-05, 12-1-05, 4-25-06)
CTLS (9-1-05, 12-1-05, 4-25-06)
CSCT (5-21-05, 9-1-05, 12-1-05, 4-25-06)
PC (5-21-05, 9-1-05, 12-1-05, 4-25-06)

[1] Our system works with the actual points in the graph; for clarity of presentation, we only show the x-values for the points corresponding to $<\mathrm{param}_i>$ in Table 1.

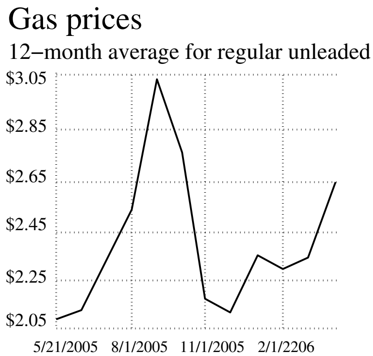

Fig. 5. Line graph from a local newspaper

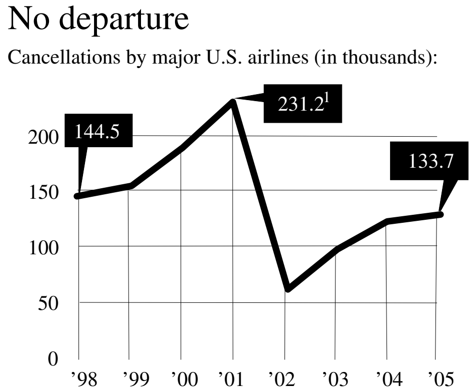

Fig. 6. Line graph from USA Today with multiple annotations. (This graphic appeared on a slant in its original form.)

3.3 Communicative Signals as Evidence

Just as listeners use evidence to identify the intended meaning of a speaker's utterance, so also must a viewer use evidence to recognize a graphic's intended message. We hypothesize that if the graphic designer goes to the effort of entering attention-getting devices into a graphic to make one or more of the entities in the graphic particularly salient, then the designer probably intends for these entities to be part of the graphic's intended message. There are several ways in which a graphic designer explicitly makes an entity in a line graph salient.

The graphic designer may annotate a point on a line graph with a value or a piece of text. This draws attention to that point in the line graph and serves as evidence that the point plays a role in the graphic's intended message. Consider the graphic in Figure 6. Three points in the graphic are annotated with their value. This suggests that these points are particularly important to the graphic's intended message — in terms of our representation, the points might serve as parameters of the graphic's intended message. This provides strong evidence for a Change-Trend-Return('98,'01,'02,'05) message since three of the four parameters of the message are salient in the graphic. Similarly, consider Figure 2. The low point in the graphic is annotated with text, suggesting that it is important to the graphic's message. This annotation might provide evidence for a Big-Fall(00,01) or for a Change-Trend-Return (where the annotation is on the point where the return begins), among others. The Visual Extraction Module is responsible for producing an XML representation of a graphic that indicates any annotated points and their annotations.

A point in the line graph can also become salient by virtue of its being referenced by a noun in the caption. This can occur by the caption referring to its x-axis value or even to its y-value, although the latter occurs less often. For example, if the caption on the graphic in Figure 2 were *"Poppies Missing in 01"*, the reference to the year "01" would lend salience to the low point in the graphic

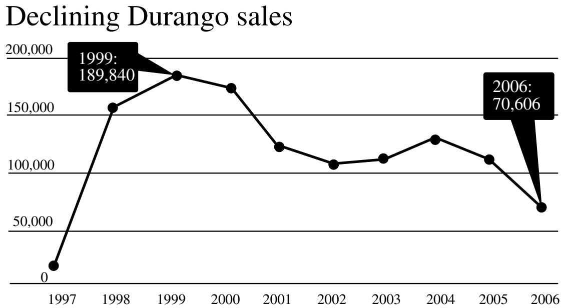

Fig. 7. Line graph from a local newspaper

even if it were not annotated. The Caption Tagging Module is responsible for augmenting the XML representation of a graphic so that it indicates any points that are referenced by nouns in the caption.

Certain parts of a graphic become salient without any effort on the part of the graphic designer. For example, a viewer's attention will be drawn to a sudden large rise or fall in a line graph. Similarly, a viewer will be interested in the segment at the end of a line graph since it captures the end of the quantitative changes being depicted. Although no specific effort is required by the graph designer, we posit that it is mutually believed by both graph designer and viewer that such pieces of the graphic will be salient. Our system extracts such evidence by analyzing the segments produced by the Graph Segmentation Module and using their slopes, their relative change with respect to the range of y-values in the graph, and their positions in the graphic as evidence.

Captions are often very general and do not capture a graphic's intended message[6]. For example, the caption on the graphic in Figure 2 fails to capture its message that there was a sudden big fall (that was not sustained) in Afghanistan opium production. Moreover, even when a caption conveys some of the graphic's message, it is often ill-formed or requires extensive world knowledge to understand. However, as in our work on simple bar charts, we have found that verbs in a caption often suggest the general category of the graphic's message. Adjectives and adverbs function similarly. For example, the adjective *"declining"* in the caption of Figure 7 suggests a Falling-Trend message or perhaps a Change-Trend message where the trends change from rising to falling.

Using WordNet, we identified potentially helpful verbs and organized them into classes of similar verbs. For example, the verbs *"jump"* and *"boom"* reside in one verb class, whereas the verbs *"resume"* and *"recover"* reside in a different verb class. The Caption Tagging Module is responsible for extracting such helpful words from the caption and augmenting the XML representation of the graphic to indicate the presence of any of our six identified verb classes.

Several features of the segments comprising a suggested message also provide evidence for or against that proposed message being the intended message of the graphic. The graphic designer presumably had a reason for including all of the points in a line graph. Thus the fraction of a line graph covered by the segments comprising a suggested message serves as evidence about whether that was the graphic designer's intended message — presumably, messages that cover much of the line graph are more likely to be the designer's intended message. (However, the intended message need not cover the entire graphic. For example, it appears that when conveying a Rising-Trend or a Falling-Trend, the graphic designer sometimes includes a small segment of points prior to the start of the trend in order to keep the viewer from inferring that the rise or fall might have started at earlier points not depicted in the graphic.)

3.4 The Bayesian Network

A Bayesian Network is a probabilistic reasoning system that can take into account the multiple pieces of evidence in a line graph in order to evaluate the various message candidates proposed by the Suggestion Generation Module. Rather than identifying an intended message with certainty, a Bayesian Network gives us the posterior probability of each candidate message, thereby reflecting any ambiguity in the graphic. In our project, a new Bayesian Network is built dynamically (using Netica[15]) each time a line graph is processed. The top-level node of our Bayesian Network represents each of the possible high-level message categories, such as Change-Trend or Big-Jump. Each of these high-level message categories appears as a child of the top-level node; this is purely for ease of representation. The children of each of these high-level message category nodes are the suggested messages (with instantiated parameters) produced by the Suggestion Generation Module.

Once nodes for each of the messages suggested by the Suggestion Generation Module have been added to the Bayesian Network, evidence nodes are entered into the network to reflect the evidence for or against the different suggested messages. Verb and adjective/adverb evidence suggest a general category of message, such as Rising-Trend or Change-Trend; thus they are attached as children of the top-level node in the Bayesian Network. Other evidence, such as whether there are annotations and whether they correspond with parameters of a message, serve as evidence for or against each suggested message; thus these evidence nodes are entered as children of each suggested message node.

Associated with each node in a Bayesian Network is a conditional probability table that reflects the probability of each of the values of that node given the value of the parent node. (The conditional probability table for the top-level node captures the prior probabilities of each of the message categories.) To construct the conditional probability tables, each line graph in our corpus of 215 line graphs was first annotated with its intended message as identified by consensus among three coders; we then analyzed each line graph to identify the evidence that was present, and computed the conditional probability tables from this analysis. One such conditional probability table is shown in Table 2. It gives the conditional probability that the endpoints $<\text{param}_1>$ and $<\text{param}_2>$ of a Rising-Trend($<\text{param}_1>$, $<\text{param}_2>$) message are annotated in the graphic, given that the intended message is (or is not) a Rising-Trend. For example, the

Table 2. A sample conditional probability table

Endpoints Annotated Table

Rising-Trend($<\text{param}_1>$, $<\text{param}_2>$)	InPlan	NotInPlan
Only one endpoint is annotated	12.3%	26.2%
Both endpoints are annotated	55.4%	3.6%
No endpoint is annotated	32.3%	70.2%

InPlan column of the conditional probability table shows that the probability that both endpoints are annotated is 55.4% if a Rising-Trend is the intended message, and the NotInPlan column shows that the probability is 3.6% if it is not the intended message.

4 Evaluation of the System

We evaluated the performance of our system for recognizing a line graph's intended message on our corpus of 215 line graphs that were collected from various magazines such as *Newsweek*, *BusinessWeek*, and from local and national newspapers. Input to the Intention Recognition Module is the augmented XML representation of a graphic. We used leave-one-out cross validation in which each of the graphics is used once as the test graphic, with the conditional probability tables computed from the other 214 graphics. Our system recognized the correct intended message with the correct parameters for 157 line graphs, which gave us a 73.0% overall success rate.

The system's errors are primarily due to sparseness of data. For example, if we have only one graphic where a particular verb class is used to indicate an intention category, then leave-one-out cross validation has no means to connect the verb class with that intention category and we are likely to get an incorrect result when hypothesizing the intended message of that graphic. In addition, if the Graph Segmentation Module does not produce the correct segmentation of a graphic, the Suggestion Generation Module is unlikely to produce a set of suggested messages that includes the graphic's intended message, and thus the Bayesian Network will not correctly hypothesize it. Therefore to improve the performance of our intention recognition system, we are working on identifying additional attributes that can produce a better learned model for graph segmentation, and we are collecting additional line graphs for training our Bayesian Network. However, even when our system does not produce the ideal result, the message hypothesized by our system still reflects the information in the graphic.

5 Examples

Consider the graphic in Figure 5. As described in Sections 3.2, our Graph Segmentation Module hypothesizes that the graphic consists of three visually distinguishable trends and our Suggestion Generation Module suggests a set of

11 possible messages for consideration by the Bayesian network. The Bayesian network hypothesizes that the graphic is conveying a Change-Trend-Return message — in particular, that the trend in gas prices changed in 9/1/05 (from rising to falling) but returned in 12/1/05 to its previous trend (rising). The system assigns this message a probability of 98.7% indicating that it is very confident of its hypothesis. Next consider the line graph in Figure 4 which illustrates the processing of a line graph consisting of a large number of short line segments. Our Graph Segmentation Module segments this line graph into two visually distinguishable trends, and the Bayesian network hypothesizes that the graphic conveys a changing trend from relatively stable between 1900 and 1930 to rising between 1930 and 2003 and assigns this hypothesis a probability of 99.9%

Now let us consider two graphics where the system is less certain about its hypothesized messages. In the case of the graphic in Figure 6, the system hypothesizes that the graphic is conveying a Change-Trend-Return in cancellations by major U.S. airlines (rising, then falling, then returning to a rising trend) and assigns the hypothesis a probability of 63.1%. However, the system also assigns a probability of 36.1% to the hypothsis that the graphic's intended message is that there was a big fall in cancellations between 2001 and 2002. The system prefers the change-trend-return hypothesis due to the stronger evidence — for example, there is no annotation on the low point at 2002 (thereby suggesting that the fall is not the primary message of the graphic), and there are annotations on other points in the graphic (thereby suggesting that those points should be parameters of the message).

As a second example where the system is less certain about its hypothesized message, consider the graphic in Figure 7 and two of the suggestions proposed by the Suggestion Generation Module: a Change-Trend(1997,1999,2006) and a Falling-Trend(1999,2006). There are a number of communicative signals in the graphic that were deliberately entered by the graph designer: 1) the annotation giving the value for the year 1999, 2) the annotation giving the value for the year 2006, and 3) the adjective *"declining"* in the caption *"Declining Durango sales"*. Other evidence entered into the Bayesian Network includes (among others) the portion of the graphic covered by each suggested message, and the relative width of the last segment of each message. For the Change-Trend message, the message covers the whole graphic and the last segment covers more than half of the graphic; for the Falling-Trend message, the last (and only) segment covers much, but not all, of the graphic.

The system considers all of the suggested messages and the evidence entered into the Bayesian Network; it hypothesizes that the graphic's intended message is that there is a falling trend in Durango sales from 1999 to 2006 and assigns this hypothesis a probability of 54.06%. The hypothesis that the graphic is intended to convey a Change-Trend (rising from 1997 to 1999 and then falling from 1999 to 2006) is assigned a probability of 45.90%. All the other suggested messages share the remaining 0.04% probability. The probabilities assigned to the Falling-Trend and Change-Trend messages reflect the ambiguity about the intended message that is inherent in the graphic. The presence of the adjective *"declining"* and the annotations on both points that are parameters of the Falling-Trend message, but only annotations on two of the three points that are parameters of the Change-Trend message, caused the system to prefer the Falling-Trend message over the Change-Trend message. Notice that while the graphic in Figure 7 does show a short rising segment prior to the long falling trend from 1999 to 2006, the focus of the graphic is on the falling trend rather than on a change in trend. (Production of Durango cars only started in 1997, so the first part of the graph primarily reflects the "ramp up" in initial sales, not a changing trend.)

Now let us examine how the system's hypothesis changes as we vary the communicative signals in the graphic. Suppose that we add an extra annotation giving the value of Durango sales in 1997. Now the system's hypothesis changes dramatically — it identifies the Change-Trend as the intended message of the graphic and assigns it a probability of 99.5%, with the Falling-Trend message assigned a probability of 0.5%. Note that although the adjective *"declining"* is most associated with a Falling-Trend message, it can also be used with a Change-Trend message to draw attention to the falling portion of the changing trend.

Now let's return to the original graphic in Figure 7 with only two points annotated, but let's change the caption to *"Durango sales changed"*. Whereas *"declining"* might be used in the caption of a Change-Trend message, it is less likely that the verb *"changed"* would be used with a Falling-Trend message. Once sgain, the system hypothesizes that the graph is intended to convey the changing trend in Durango sales, rising from 1997 to 1999 and then falling from 1999 to 2006, but only assigns it a probability of 95.2% due to the ambiguity resulting from only two points being annotated.

6 Related Work

Shah et. al.[20] had people describe line graphs to examine how the graph design affects what people get as the message of the graphic. Our work used Bayesian network to reason about the messages of the graphic from the evidences, which implemented the automated recognition of line graph's messages.

Yu et. al.[23] developed a pattern recognition algorithm for summarizing interesting features of automatically generated time-series data such as from a gas turbine engine. However, they were analyzing automatically generated machine data, not graphs designed by a graphic designer whose intention was to convey a message to the viewer. Futrelle and Nikolakis[10] developed a constraint grammar for parsing vector-based visual displays and producing structured representations of the elements comprising the display. The goal of Futrelle's work is to produce a graphic that is a summary of one or several more complex graphics[9]. Note that the end result will again be a graphic, whereas our goal is to recognize a graphic's intended message.

A number of researchers have studied the problems of classifying time series data into a pattern category[14] or judging the similarity between time-series data[13,21]. Their main goal is to identify the pattern of a query time series by calculating its similarity with a predefined pattern. Dasgupta et.al.[5] identify anomalies or events in a data series. Our research differs from these efforts in that we are segmenting line graphs into visually distinguishable trends that can be used to suggest possible messages for consideration by a system that recognizes the graphic's intended message.

7 Conclusion and Future Work

Information graphics in popular media generally have a message that they are intended to convey, and this message is often not captured by the graphic's caption or given in the accompanying article's text. This paper has presented an implemented and evaluated methodology for identifying the intended message of a line graph. Our methodology involves segmenting the graphic into visually distinguishable trends, extracting communicative signals from the graphic, and using these in a Bayesian Network that hypothesizes the graphic's intended message. The evaluation of our system's performance demonstrates the effectiveness of our approach.

Our current work is using a graphic's recognized message as the basis for summarizing the high-level content of graphics from popular media, in order to provide alternative access for individuals with sight-impairments. We are also investigating the use of the intended message to index and retrieve information graphics, to produce summaries that take into account a multimodal document's information graphics as well as its text, and to extract information from multimodal documents. To our knowledge, our project is the only current research effort to identify an information graphic's intended message and utilize it in processing multimodal documents.

References

1. Alty, J., Rigas, D.: Communicating graphical information to blind users using music: The role of context. In: Proceedings of CHI 1998, pp. 574–581 (1998)
2. Bradley, D.C., Steil, G.M., Bergman, R.N.: OOPSEG: a data smoothing program for quantitation and isolation of random measurement error. Computer Methods and Programs in Biomedicine 46, 67–77 (1995)
3. Chester, D., Elzer, S.: Getting computers to see information graphics so users do not have to. In: Proc. of the International Symposium on Methodologies for Intelligent Systems, pp. 660–668 (2005)
4. Clark, H.: Using Language. Cambridge University Press, Cambridge (1996)
5. Dasgupta, D., Forrest, S.: Novelty detection in time series data using ideas from immunology. In: Neural Information Processing Systems Conference (1996)
6. 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: Proc. of the Association for Computational Linguistics, pp. 223–230 (2005)
7. Elzer, S., Carberry, S., Demir, S.: Communicative signals as the key to automated understanding of bar charts. In: Barker-Plummer, D., Cox, R., Swoboda, N. (eds.) Diagrams 2006. LNCS (LNAI), vol. 4045, pp. 25–39. Springer, Heidelberg (2006)

8. Ferres, L., Verkhogliad, P., Lindgaard, G., Boucher, L., Chretien, A., Lachance, M.: Improving accessibility to statistical graphs: the inspectgraph system. In: Proc. of ACM SIGACCESS Conference on Computers and Accessibility (2007)
9. Futrelle, R.: Summarization of diagrams in documents. In: Mani, I., Maybury, M. (eds.) Advances in Automated Text Summarization, pp. 403–421. MIT Press, Cambridge (1999)
10. Futrelle, R., Nikolakis, N.: Efficient analysis of complex diagrams using constraint-based parsing. In: Proc. of the International Conference on Document Analysis and Recognition, pp. 782–790 (1995)
11. Keogh, E., Kasetty, S.: On the need for time series data mining benchmarks: A survey and empirical demonstration. Data Mining and Knowledge Discovery 7, 349–371 (2003)
12. Levy, E., Zacks, J., Tversky, B., Schiano, D.: Gratuitous graphics? putting preferences in perspective. In: Proceedings of the ACM Conference on Human Factors in Computing Systems, pp. 42–49 (1996)
13. Lin, J., Keogh, E., Lonardi, S., Patel, P.: Finding motifs in time series. In: Proceedings of the Workshop on Temporal Data Mining, pp. 53–68 (2002)
14. Megalooikonomou, V., Wang, Q., Li, G., Faloutsos, C.: A multiresolution symbolic representation of time series. In: Proc. of the International Conference on Data Engineering, pp. 668–679 (2005)
15. N.S.C. Netica (2005)
16. Pinker, S.: A theory of graph comprehension. In: Freedle, R. (ed.) Artificial Intelligence and the Future of Testing, pp. 73–126. Lawrence Erlbaum, Mahwah (1990)
17. Platt, J.C.: Fast training of support vector machines using sequential minimal optimization. In: Advances in Kernel Methods: Support Vector Learning, pp. 185–208 (1999)
18. Ramloll, R., Yu, W., Brewster, S., Riedel, B., Murton, M., Dimigen, G.: Constructing sonified haptic line graphs for the blind student: First steps. In: Proceedings of Assets 2000, pp. 17–25 (2000)
19. Rodgers, J.L., Nicewander, W.A.: Thirteen ways to look at the correlation coefficient. The American Statistician 42, 59–66 (1988)
20. Shah, P., Mayer, R.E., Hegarty, M.: Graphs as aids to knowledge construction: Signaling techniques for guiding the process of graph comprehension. Journal of Educational Psychology 91(4), 690–702 (1999)
21. Toshniwal, D., Joshi, R.: Finding similarity in time series data by method of time weighted moments. In: Proc. of Australasian Database Conference, pp. 155–164 (2005)
22. Wickens, C.D., Carswell, C.M.: The proximity compatibility principle: Its psychological foundation and relevance to display design. Human Factors 37(3), 473–494 (1995)
23. Yu, J., Reiter, E., Hunter, J., Mellish, C.: Choosing the content of textual summaries of large time-series data sets. Natural Language Engineering 13, 25–49 (2007)
24. Zacks, J., Tversky, B.: Bars and lines: A study of graphic communication. Memory & Cognition 27(6), 1073–1079 (1999)