TSINGHUA SCIENCE AND TECHNOLOGY
ISSN 1007-0214 02/17 pp7-12
Volume 16, Number 1, February 2011

Causal Inference in Graph-Text Constellations: Designing Verbally Annotated Graphs*

Christopher Habel**, Cengiz Acartürk†

Department of Informatics, University of Hamburg, 22527 Hamburg, Germany;
† Informatics Institute, Middle East Technical University, Ankara 06531, Turkey

Abstract: Multimodal documents combining language and graphs are wide-spread in print media as well as in electronic media. One of the most important tasks to be solved in comprehending graph-text combinations is construction of causal chains among the meaning entities provided by modalities. In this study we focus on the role of annotation position and shape of graph lines in simple line graphs on causal attributions concerning the event presented by the annotation and the processes (i.e. increases and decreases) and states (no-changes) in the domain value of the graphs presented by the process-lines and state-lines. Based on the experimental investigation of readers' inferences under different conditions, guidelines for the design of multimodal documents including text and statistical information graphics are suggested. One suggestion is that the position and the number of verbal annotations should be selected appropriately, another is that the graph line smoothing should be done cautiously.

Key words: causal inference; multimodal comprehension; human computer interaction

Introduction

Multimodal documents, composed of representations in different modalities (e.g., text, pictorial illustrations, graphs), are widely used both in human-computer interaction environments and in printed media. Graph-text constellations—the combination of statistical information graphics (e.g., line graphs, bar graphs, etc.; henceforth, "graphs") and text—constitute a sub-type of multimodal documents. Due to their communicative nature they are frequently used as a cognitive means in communication tasks, as well as in high-level problem solving. The pivot of multimodal comprehension, as in language comprehension, is the construction of "causal relations" among semantic entities[1], a process kindred to "causal inference" (also termed "causal attribution"). Graph-text constellations can influence causal inference by explicit argumentation or implicit argumentative hints in text comprehension[2] and in diagrams of different types[3,4]. For instance, verbal annotations in line graphs are used to annotate processes and states presented by graph lines. The annotations also serve for constructing the causal inferences that temporally coincide with graphically presented processes and states (Fig. 1). We use the term "causal inference", "causal attribution", and "causality" in a broader sense, to mean psychological phenomenon of causal inference rather than what causality is and its underlying mechanisms.

From the perspective of computer science and human computer interaction, generation of verbally annotated graphs—in particular positioning of verbal entities in a graph frame—presents a challenge for automatic diagram parsing and pattern recognition[5]. Nevertheless, the research on causal inference in graph-text

Received: 2010-11-29
* Supported in part by DFG (German Science Foundation) in ITRG 1247 'Cross-modal Interaction in Natural and Artificial Cognitive Systems' (CI-NACS)
** To whom correspondence should be addressed.
E-mail: habel@informatik.uni-hamburg.de; Tel: 49-40-428832417

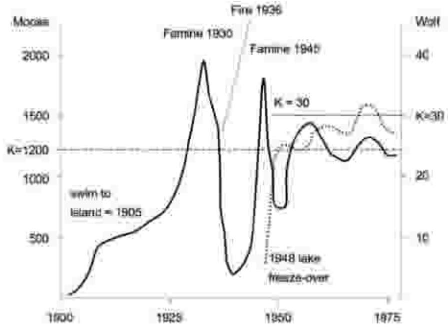


Fig. 1 Isle royale moose/wolf progression. Redrawn based on N. C. Heywood's Geog358 course material on Biogeography, UWSP (http://www.uwsp.edu/, retrieved on November 6, 2010)

constellations has been scarce except for a few studies[6]. In this study, we investigate causal inference in verbally annotated graphs from two perspectives. On the one hand, we empirically investigate human judgments of causal inference by taking humans as natural cognitive systems. Then we employ the empirical findings for developing design guidelines for verbally annotated graphs, as well as for providing the background for design and implementation of artificial cognitive systems that are capable of producing verbally annotated graphs.

1 Multimodal Comprehension of Graph-Text Constellations

Graphs comprise a distinct class of quasi-symbolic depictive representation, compared to pictorial illustrations, exhibiting depictive characteristics by using the space in a non-arbitrary manner in addition to having syntactic characteristics[7,8]. Pictorial illustrations are "representational pictures"[9] that can be characterized by their visual resemblance to the objects they stand for, whereas graphs are human artifacts that are designed to represent conceptual information about domain entities by means of graphical entities. For instance, the line segments in the line graph in Fig. 1 represent the time course of history in the population of wolves and moose.

In most communicational settings, such as printed or electronic media in daily life as well as scientific journals, graphs are used in combination with text, thus leading to a combination of two representational formats, namely graph-text constellations. Verbally annotated graphs constitute a specific subcase of graph-text constellations (viz., graph-internal multimodality)[10]. The term "verbal annotation" refers to spatially contiguous textual entities that annotate parts of graphical entities within the graph frame, as "Famine 1930" and "Fire 1936" in Fig. 1. Verbally annotated graphs are frequently used for providing interaction between events, processes, and states presented by graphical and textual entities. They have a similar role, though in written form, to gestures in multimodal communication. Verbally annotated graphs have various functions other than annotating events, processes, and states represented by graphical entities. Those functions, such as labeling the numerical values, are beyond the scope of this study.

Habel, Acartürk, and colleagues[4,10-12] investigated comprehension of graph-text constellations, including verbally annotated graphs, from an information processing perspective. In a coarse-grained cognitive architecture for multimodal comprehension of graph-text constellations, Acartürk[4] analyzed the structural aspects of multimodal comprehension. Acartürk's architecture is composed of two modality-specific processing modules, namely a language comprehension module and a graph comprehension module. The language comprehension module is grounded on conceptual representations and processes, which make use of the representation format CRIL[13] (conceptual route instruction language). Tschander et al. focuses on the interpretation of verbal route descriptions based on the knowledge of spatial language and temporal and spatial concepts, which is a language comprehension task (viz., verbally instructed navigation)[13]. The graph comprehension module is based on the graph comprehension architecture proposed by Pinker[14]. According to this architecture, grounded in the common conceptual basis for graph-text comprehension the modality-specific processing of linguistic entities and graphical entities builds up a coherent model of a graph-text constellation, thus leading to integrated conceptual representations that are accessible by both modalities. Accordingly, events, processes, and states that are referred to by graphical entities in line graphs (e.g., increases and decreases that are presented by graph line segments) are also referred to by linguistic entities such as verbal annotations (see Fig. 1). Causal inference is one of the means for the construction of coherence relations between entities with the two modalities, thus contributing to a coherent interpretation of verbally annotated graphs. In particular, a coherent interpretation of events, processes, and states in verbally annotated graphs requires the integration of information provided by both linguistic entities and the relevant parts of graphical entities, which in turn leads to the construction of integrated conceptual representations.

In this study, we present our analysis of causal inference in verbally annotated graphs in two parts. First, we present an empirical investigation of humans' construction of causal inference during their course of interpretation of verbally annotated graphs. In the second part, we use the empirical findings to develop design guidelines for verbally annotated graphs, as well as for the design and implementation of artificial cognitive systems that have the capability of generating verbally annotated graphs.

2 Experiment

In an experimental investigation with twenty-nine participants (mean age of 23.4, SD = 3.5), readers' judgments concerning causal relationships between processes and states presented by graph lines and the events presented by verbal annotations were investigated, and participants' eye movement behavior during the course of interpretations was analyzed.

Verbally annotated line graphs were prepared by the experimenter for the purposes aimed at the experimental design conditions. Each graph represented a change of a domain value in time. Three domains (chemistry, finance, and biology) were used in the experimental material. The design of the material involved one independent between-subject parameter and two independent within-subject parameters, as described below.

As for the between-subject parameter, specified by the smoothness condition (SC), the participants were divided into two groups: one group (namely, the SC-straight group) was presented line graphs with straight graph lines, the other (namely, the SC-smooth group) with smooth graph lines. The first within-subject parameter, specified by the graph condition (GC), was concerned with the type and the order of graph line segments that form binary constellations: the condition GC-1 involved graphs that had a line segment representing an increase in the canonical reading direction of graphs (namely, an increase process-line), followed by a horizontal line (namely, a state-line). In GC-2, an increase process line followed a state-line; and in GC-3 a decrease process-line followed an increase process line. Finally, the second within-subject parameter, specified by the annotation condition (AC) was concerned with the number and position of verbal annotations. The left-AC condition involved a verbal annotation at the leftmost end of the graph line. The middle-AC involved a verbal annotation at the middle, at the intersection of the two segments of the graph lines. The double-AC involved two verbal annotations, one at the left and the other at the middle. Sample stimuli are shown in Fig. 2. The language of the experiment was Turkish. The experimental material presented in this paper was translated into English by the second author.

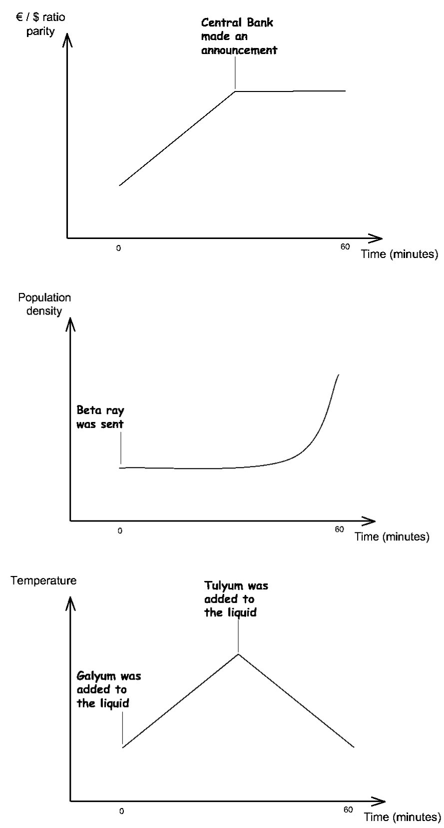


Fig. 2 Sample stimuli from the experiment

Following the presentation of a practice session and relevant context information about the domains, the participants were presented the graphs and a set of questions concerning the causal relationships. The questions had the form as exemplified in Table 1.

Table 1 Sample questions from the experiment. E is used as an abbreviation for the verbal annotation text. V is used as an abbreviation for the domain value of the graph.

Did E cause V to increase?
Did E cause V to stop increasing?
Did E cause V to remain constant?

Each participant answered a total of 30 questions by clicking on a visual analog scale. The scale had the end anchors "No with high probability" and "–100" at the left end, the end anchors "Yes with high probability" and "–100" at the right end, and the anchors "I am in the middle" and "0" at the middle. There was no time limitation and the experiment took about 30 min to complete. Participants' eye movements were recorded by a 50 Hz non-intrusive eye tracker.

3 Results

Participants' interpretations of causal attribution were analyzed in terms of the analysis of reported judgment scores. Participants' interpretation of the stimuli in the left-AC revealed that the position of the annotation resulted in ambiguities in most of the cases, as evidenced by a high variability in the judgment scores. The ambiguities were observed both in the SC-straight group and in the SC-smooth group. As for the middle-AC, the position of the annotation resulted in less ambiguities compared to the left-AC condition. Moreover, less ambiguity was observed in the SC-straight group compared to the SC-smooth group in the middle-AC. Finally, the use of double annotations in the double-AC resulted in much less ambiguity (in both the SC-straight group and the SC-smooth group) compared to both the left-AC and the middle-AC.

Participants' eye movement recordings were analyzed in terms of the mean gaze times on the relevant regions in the experiment stimuli. The analysis revealed supporting evidence for the difficulties concerning the interpretation of the causal relationships presented above. In particular, the analysis of the mean gaze times revealed that, in the left-AC, the participants spent more time for inspecting the graph (as well as for inspecting the question and the visual analog scale), compared to both the middle-AC and the double-AC conditions. In addition, the longer mean gaze times in the SC-smooth group, compared to the SC-straight group, suggests that the SC-smooth group had more difficulties in their interpretations compared to the SC-straight group. In other words, interpretations of the causal relationships were easier when the graph lines were composed of straight line segments rather than smooth line segments.

These findings show that both the position of the verbal annotation and the smoothness of the graph lines result in systematic differences in humans' causal inferences in verbally annotated graphs.

4 Design Guidelines and Implications for Software Development

The facilitating role of verbal annotations in multimodal comprehension of graph-text constellations has already been shown in previous studies[10,11]. However, the available software tools provide limited end-user support for designing and producing verbally annotated graphs. In this section, the empirical findings obtained from the presented experimental investigation will be used to develop two design guidelines for verbally annotated graphs at the user-level[4].

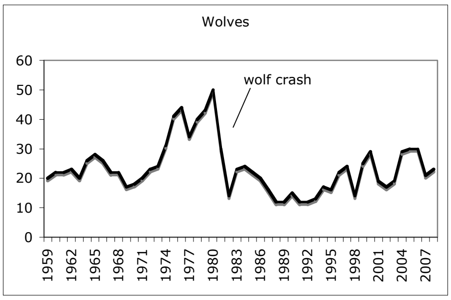
Guideline 1: The position and the number of verbal annotations should be selected appropriately.

The available software tools provide limited support for positioning verbal annotations in the graph space. In particular, the positioning of verbal annotations exhibits similar problems to the reflow problem in text editing software, i.e. the problem of appropriately rearranging the display space after changing the visual properties of a text. The reflow problems and their technical solutions have been discussed in text-only documents, such as freeform ink annotations[15,16]. Figure 3 exemplifies a reflow problem in verbally annotated graphs. The graph in Fig. 3a was produced by the recent version of a popular spreadsheet software. It was then resized and the one in Fig. 3b was obtained. Note the position of the verbal annotation "wolf crash".

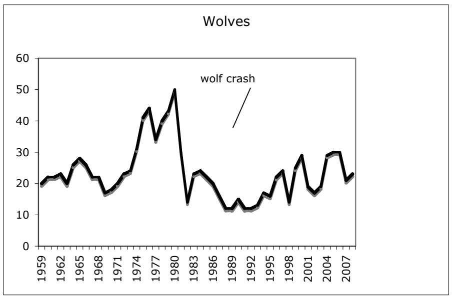
The second design guideline is concerned with using smoothed-line graphs, based on the experimental findings reported in the previous section.

Guideline 2: Graph line smoothing should be done cautiously.

Guideline 2 reveals its importance when it is applied to recently available software tools that provide end-users with smoothing functions. In most cases, the smoothing functions are used for making graph lines

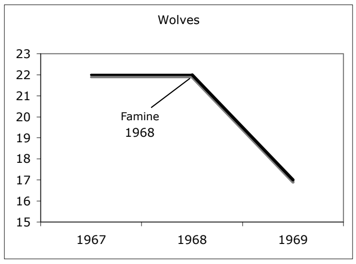


(a)

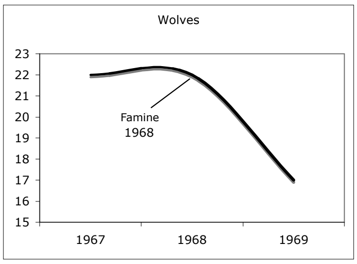


(b)

Fig. 3 The reflow problem in verbally annotated graphs



(a)



(b)

Fig. 4 The smoothing problem in verbally annotated graphs

look visually appealing. However, the smoothing functions may result in artifacts that lead to ambiguous (and erroneous) interpretations, as exemplified in Fig. 4. The verbally annotated graph in Fig. 4a was produced by the software tool that was used before to produce the graphs in Fig. 3. Applying the smoothing option provided by the software tool produced the graph in Fig. 4b. Note how smoothing resulted in artifacts that influenced causal inference. In addition to the erroneously generated increase process-line segment, the location of the end of the annotation icon (i.e., the straight line between the verbal annotation and the graph line) with respect to the graph line lead to misinterpretations concerning the causal relationship between the event presented by the verbal annotation and the processes and states presented by the graph lines.

From a technical point of view, specifying a set of anchor points in the graph space can solve the reflow problems in verbally annotated graphs. In particular, the anchor points specify the positioning of the annotation text, which is usually a text box, with respect to the relevant graphical entity, such as a graph line segment.

5 Conclusions and Future Work

Humans make inferences not purely on statistical data; rather, they interpret the data depending on characteristics of the communicative mode. For this reason, statistical information graphics have potential communicative roles as well as being the visual representation of data. In communicating statistical data by multimodal, graph-text documents, the common, integrated comprehension of the graph-text combination documents may result in different interpretations and even misinterpretations due to the limitations of the graphical means used to convey information. In this study, we have investigated the standard use of annotations, (i.e., annotation of text phrases via a line which on the one hand points the annotation text, on the other hand points the graphical elements) in simple line graphs. The results of our study showed that the position of the annotation and the shape of the graph line affected readers' judgments concerning the causal attributions between the event presented by the annotation and the process-lines and state-lines in the graph.

Recent developments in user-interface technology—such as interactive software programs to design and generate graphs, as well as easy to use systems for generating hypertext, for publishing and for realizing web- pages and web-sites—provide users with the opportunity to design and produce graphs and graph-text constellations in myriad formats. For instance, most of the available word processing software tools now have spell checking and grammar checking capabilities that have been showing a steady progress over time. Software tools for annotating documents in collaborative writing have been developed and tested for usability[15,16]. Nevertheless, the development of assistance systems for design, production, and checking of multimodal constellations has not yet fully progressed to reach the designer (i.e., producer) as the end-user. Despite the significant progress in information visualization, the support for visualization-based communication has been limited[17]. Based on the empirical findings reported in this paper, we proposed a set of guidelines for the design of text-graph constellations in statistical information graphics. These guidelines are useful for the individual user who creates text-graph documents as well as for developers of design software tools for end users.

Acknowledgements

Special thanks go to Jianwei Zhang who is the spokesperson of CINACS, to Maosong Sun, and to other colleagues both at the University of Hamburg and at Tsinghua University for their useful comments and suggestions. In addition, we thank Human Computer Interaction Research and Application Laboratory of the Middle East Technical University for their generous technical support.

References

[1] Graesser A C, Millis K K, Zwaan R A. Discourse comprehension. *Annual Review of Psychology*, 1997, **48**: 163-189.

[2] Hilton D J. Conversational processes and causal explanation. *Psychological Bulletin,* 1990, **107**(1): 65-81.

[3] Oestermeier U, Hesse F W. Verbal and visual causal arguments. *Cognition*, 2000, **75**: 65-104.

[4] Acartürk C. Multimodal comprehension of graph-text constellations: An information processing perspective [Dissertation]. University of Hamburg, Hamburg, 2010.

[5] Futrelle R P. Summarization of diagrams in documents. In: Mani I, Maybury M, eds. Advances in Automated Text Summarization. Cambridge, MA: MIT Press, 1999: 403-421.

[6] Green N. Design of information graphics for causal arguments. In: Proceedings of the International Joint Conference on Artificial Intelligence, Workshop on Computer Models of Natural Argument. Edinburgh, UK, 2005.

[7] Kosslyn S M. Image and Mind. Cambridge, MA: Harvard University Press, 1980.

[8] Kosslyn S M. Understanding charts and graphs. *Applied Cognitive Psychology*, 1989, **3**(3): 185-226.

[9] Alesandrini K L. Pictures and adult learning. *Instructional Science,* 1984, **13**: 63-77.

[10] Acartürk C, Habel Ch, Cagiltay K. Multimodal comprehension of graphics with textual annotations: The role of graphical means relating annotations and graph lines. In: Howse J, Lee J, Stapleton G, eds. Diagrammatic Representation and Inference: Lecture Notes in Computer Science. Berlin/Heidelberg: Springer, 2008, **5223**: 335-343.

[11] Acartürk C, Habel Ch, Cagiltay K, et al. Multimodal comprehension of language and graphics: Graphs with and without annotations. *Journal of Eye Movement Research*, 2008, **1**(3): 2, 1-15.

[12] Habel Ch, Acartürk C. Eye-tracking evidence for multimodal language-graphics comprehension: The role of integrated conceptual representations. In: Navarretta C, Paggio P, Allwood J, eds. Proceedings of NODALIDA 2009. Odense, Denmark, 2009: 9-14.

[13] Tschander L, Schmidtke H, Habel Ch, et al. A geometric agent following route instructions. In: Freksa F, Brauer W, Habel C, et al., eds. Spatial Cognition III. Heidelberg: Springer, 2003: 89-111.

[14] Pinker S. A theory of graph comprehension. In: Freedle R, ed. Artificial Intelligence and the Future of Testing. Hillsdale, NJ: Erlbaum, 1990: 73-126.

[15] Brush A J, Bargeron D, Gupta A, et al. Robust annotation positioning in digital documents. In: Proceedings of the CHI 2001. New York, ACM Press, 2001: 285-292.

[16] Bargeron D, Moscovich T. Reflowing digital ink annotations. In: Proceedings of CHI 2003. New York, ACM Press, 2003: 385-393.

[17] Viégas F B, Wattenberg M. Communication-minded visualization: A call to action. *IBM Systems Journal*, 2006, **45**(4): 801-812.