Helping People with Visual Impairments Gain Access to Graphical Information Through Natural Language: The *iGraph* System

Leo Ferres Avi Parush Shelley Roberts Gitte Lindgaard

Human-Oriented Technology Laboratory
Carleton University
Ottawa, Ontario, Canada
{leo_ferres,avi_parush,gitte_lindgaard}@carleton.ca
smroberts@connect.carleton.ca

Abstract. Much numerical information is visualized in graphs. However, this is a medium that is problematic for people with visual impairments. We have developed a system called iGraph which provides short verbal descriptions of the information usually depicted in graphs. This system was used as a preliminary solution that was validated through a process of User Needs Analysis (UNA). This process provided some basic data on the needs of people with visual impairments in terms of the components and the language to be used for graph comprehension and also validated our initial approach. The UNA provided important directions for the further development of iGraph particularly in terms of interactive querying of graphs.

1 Introduction

Graphs (line graphs, pie charts, etc.) are common and pervasive: they have been the accepted norm for the presentation of complex numerical data since at least the 18th century [1]. As [2, 3] point out, this manner of presenting numerical data takes advantage of the human ability to recognize certain visual patterns that the numbers by themselves (e.g. in tabular form) do not readily reveal. By quickly looking at a line graph, for example, we can determine, with somewhat less precision than knowing the real numbers, where the minimum and maximum values are, whether the values have an upwards or downwards general trend or whether it stays more or less the same, whether the line is smooth or "pointy", the scale, whether it looks like an "M" or like a semicircle or like a bell, etc. All these little pieces of information (and many more we have not mentioned here for lack of space) allow us to interpret the numbers and hypothesize about their relation on many dimensions, helping us draw conclusions about the nature of the data being represented.

All the cognitive mechanisms used to find these visual patterns [4–7] depend, rather obviously, on the human visual faculty [8]. If this faculty is not present, or considerably damaged, as is the case with blind people and people with visual impairments, then the problem resides in finding alternative ways to make (at least some of) the visual information encoded in graphs available to this audience. To this purpose, we implemented iGraph, a system that provides short descriptions and a query system to work with graphs. However, not any kind of description will do, and, on the other hand, describing everything there is to a picture (a graph in this case) may become, at times, unfeasible (e.g. when there are one hundred data points, or more than three lines in a line graph, etc.). Thus, on a first pass, our problem is not only an engineering or implementation problem, but a problem of verifying the need, building a better language model and validating the solution that iGraph provides. This paper reports the outcome of a User Needs Analysis (UNA) and Wizard-of-Oz (WOZ) study towards an improved version of iGraph. We first review the state of the art in R&D and some currently available applications. Second, we introduce our first implementation of iGraph. The third and main part of the paper is dedicated to assessing the needs and a validation of the concept of iGraph. Finally, we discuss future implementations of iGraph on the basis of the findings of the UNA and WOZ studies.

2 Related Work

The idea of presenting graphs multimodally is not new, although the ability to interact with it in natural language is. In the digital arena, there are currently two ways of helping people with visual impairments gain access to and work with the information presented by graphs: the “sonification” option and the “speech” or “reader” option. In the former, researchers have been exploring auditory graphs since at least the early 1990s [9, 10]. In these studies, several aspects of graph sonification [11, 12] and auditory graph comprehension [13–17] have been investigated, not only for the visually impaired, but also as an alternative means of data exploration [18, 19][1]. Typically, in these systems, the values of the ordinate of a graph are converted to the different notes and octaves of a musical instrument and time usually stands for the abscissa of the graph. Thus, one has access to the information by listening to the different notes along time.

For the NL (Natural Language) option, several studies investigating how to provide natural language descriptions of data graphs have been carried out and demo systems have been developed [20]. Worth noting are, for instance, PostGraphe and SelText, systems that generate a graph and its caption from an input of data in tabular form and some intentions from the users concerning the message the graph should convey [21–23]. In the same tradition of caption generation is SAGE [24], though this system aims at generating captions for 2D diagrams and maps as well as graphs. Other systems include the generation of summaries of potentially very large databases of time-series data using time-series data summarization algorithms. Examples of these systems include

[1] This list is nowhere near comprehensive, but it gives the interested reader an idea of the work carried out so far and where to go for more information.

weather report generation systems (e.g. SUM-METEO), summarizing climatological observations (such as temperature, wind speed, humidity, etc.) [25–29], and others such as SUMTIME-TURBINE [30] and NEONATE [31]. Even some relatively simple commercial applications that make use of NL methods are currently available (JAWS[2], a screen reader and ChartExplainer[3], a software environment for automatically generating natural-language summaries of charts and tables, developed by CoGenTex).

3 Our Initial Approach: iGraph

The technologies mentioned in the previous section have proven useful in several areas, but they also have some important drawbacks. Auditory graphs, for example, do not work very well when the input graph is complex, containing more than three lines, or when it is not time-dependent, as in pie charts. In a pure auditory graph, legends and other linguistic pieces of information are not read, and dimensions such as color, depth, etc. are lost. Another way to represent graphs is natural language. This is so for several reasons: first, humans are more attuned to this system and we require less training in individuating different signs (words, in this case, as opposed to going from a C to a C#); second, NL is representationally a very rich way to describe much of the information shown in graphical representations. In short, a picture is worth a thousand words, but, when spoken words is all we have access to, those thousand words are enough.

This being said, the NL systems described above are far from ideal for our purposes here; namely, helping blind and visually-impaired people work with graphical representations of data. Most importantly, all the NL systems mentioned above are quite inflexible: they provide a single summary or a short caption of a graph, but not the ability for users to explore the graphs. The system we are developing offers this novel ability to interact by means of a quite flexible and scalable NL dialogue. Figure 1 below shows the main components of our system and some of the implementation techniques for the different sub-systems.

Our prototype (the *inspectGraph* system or iGraph) automatically provides descriptions of graphs and allows for natural language interaction with a given graph. The input to the system comprises the encoding of the necessary (and sufficient) properties of a graph in the EXtended Markup Language (XML) format (titles of the axes, values and the interval labels for the X-Axis). To allow for the description and the querying, iGraph implements the three following subsystems. The first takes the XML file, parses it and writes a logic version of the given graph, together with several mathematical properties of the input graph (minimums and maximums, slope of the increase or decrease between two points, etc). We call this the P-System. The second subsystem stores different kinds of rules for describing and querying the logic-mathematical version of the input graph generated by the P-System. We call this the C-System. The third is a language subsystem that takes the logic-like representations (and, in principle, all

[2] http://www.freedomscientific.com/fs_products/software_jaws.asp
[3] http://www.cogentex.com/products/chartex/

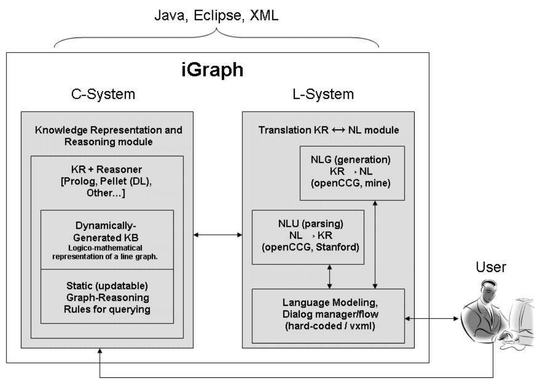

Fig. 1. The inspectGraph architecture

possible inferences) from the first and second subsystems and outputs a natural language text (both as a description and/or as the response to a query) plus the handling of a dialogue modelling algorithm for querying the graph. We call this the L-System.

As stated before, iGraph is more than an engineering problem. For the system to be useful, several issues must be considered. Two of the most important ones are *i*) what blind and visually-impaired people would ask and need to know about a graph, and *ii*) the language model needed to convey visual information successfully. To achieve this, we carried out a comprehensive UNA of people with visual impairments ranging from congenitally blind to recent blindness. The next section addresses our UNA methodology and some of the results.

4 The iGraph UNA: Method

A UNA is a family of methods used to uncover the needs and requirements a user has as well as the capabilities needed from the product to meet the user's goals. This involves understanding the target audience, their typical tasks, and their specific constraints, usually elucidated through a combination of observational techniques, including interviews, surveys and consulting with domain experts. The results provide user interface objectives, system requirements, and feature requirements[4]. The primary goal of the UNA and evaluation component of this project is to gain insight into the needs and requirements of InspectGraph users who are blind or visually impaired as well as maximize the usability of the existing InspectGraph product. This was accomplished by means of a semi-structured interview and a user-centered system evaluation.

Participants. Participants were recruited through recruitment notices posted on local mailing lists at Carleton University's Centre for Disabilities, and through personal

[4] http://www.usabilityfirst.com/glossary/term_576.txl

contacts. 11 participants whose ages ran from 19 to 57 were paid $20 for approximately an hour of their time. The interviews took place in our usability lab on campus or another location that was convenient for participants (i.e. home).

Materials. Pen and paper were used to record participants' responses. A Panasonic Leica Dicomar video camera and tripod was also used to record the interview.

Design. The overall method included a) demographic questionnaire b) informal interview c) information extraction and d) a wizard-of-oz study:

a) **Demographics**: this questionnaire included information regarding their background in terms age, gender, education, and information regarding their blindness.

b) **Informal Interview**: a semi-structured interview was conducted to gain insight into the users previous interaction with graphs, for example, where (work, school etc), how (tactile, descriptions etc), and the type of interaction (description from others, tactile graphs, other technology etc). This interview was also used to ask participants how they would like to be able to interact with graphs and what goals they would like to accomplish if they could.

c) **Information Elicitation**: participants were told that the experimenter had a graph in front of them. They were required to ask the experimenter any questions that would help them obtain enough information in order to understand what the graph looked like. This information helped to understand the type of language these participants used so that it could be also used within the system.

d) **Wizard-of-Oz study**: The wizard-of-oz (WOZ) technique enables unimplemented technology to be evaluated by using a human to simulate the response of a system and receive immediate feedback with regards to the systems' function and usability. This technique is used to test concepts and techniques during design and provides suggestions for functionality before it is designed and implemented. In this section participants were told that they are interacting with a computer system through a natural-language interface, though in fact they are not. Instead the interaction is mediated by a human operator, the wizard, allowing the participant more freedom of expression [32]. Natural dialogue researchers conclude that empirical studies of this unique communication situation are required for the development of user-friendly interactive natural dialogue systems (Dahlbck et al., 1993). In this particular section the experimenter read out what the iGraph system would currently say to describe a graph. To investigate the participants' understanding of the spoken graph they were then asked recall questions such as "what is the highest point in the graph?"

Overall, this experiment took a freeform approach which included observations and informal questions posed to 11 participants and took approximately one hour per subject. Again, audio and video of the entire interview were recorded for the purposes of evaluation and to show the results of our testing.

Results. The demographics of the participants interviewed in terms of their vision ranged from congenitally blind to non-congenitally visually impaired. However, the differences between these groups in terms of their needs did not differ enough to look at the data separately. Therefore the data listed include all participants. Education of the participants ranged from current undergraduate (n=2), Bachelor's degree (n=8), and Graduate student (n=1). Participants' previous experience with graphs differed in terms of how long they were or were not with sight. 5 participants were sighted at least through high school, and 6 were either born blind or lost their vision prior to high school. High school was the time when all participants indicated that they had the most interaction with graphs. Overall, the informal interview revealed that all participants attempt to avoid graphs. There are no technical aids that help them extract information from graphs available at the present time, which any of them knew of. However when they do need to read graphs they rely on the support of people around them to describe the graph(s). The information extraction section of the UNA revealed the most common words used to interact with the interviewer in order to extract information from the graph. Table 1 lists the words that were most frequently used to ask questions during the information extraction section of the interview. There were no differences in terms of the words used or the requests by those who had experienced graphs visually compared to those who had not.

Table 1. The 14 most frequently used words used by participants to extract graph information.

Case	Word	Frequency
1	X, Y Axes	41
2	line up/down	32
3	drastic/gradual shift/drop/jump	27
4	describe point by point	22
5	slope up/down	22
6	lowest/highest point	22
7	purpose of graph	21
8	range of axis	19
9	scale of axis	18
10	Main title	17
11	Type/kind of graph?	17
12	starting point	14
13	axis title	13
14	top/bottom right/left	13

The wizard-of-oz study revealed that overall most participants were able to visualize what the graph looked like following the reading of iGraph descriptions. 9 of 11 participants explicitly reported that they liked how the graphs were described, and all participants reported that they could find iGraph useful. Table 2 below shows the top 5 most requested requirements. All participants requested that the purpose, type of graph, main title and axes titles are read first, followed by the general trends, and then the details. Most participants requested that they iGraph describe the data point by point, however the most common remark was that this type of description may only be useful with simple, single line graphs.

Table 2. The top five requests

Case	Request	Frequency
1	Purpose, Type and Titles first (main and axes)	11
2	Describe point by point	9
3	Describe general trends and then details	8
4	Ability to repeat/pause/rewind	8
5	Keep it simple	6

5 Conclusions and Future Work

Given this UNA study, we may conclude that iGraph is indeed a useful tool. Once iGraph is finished, it should provide immediate results to people with visual impairments when it comes to reading and interacting with graphs. The Wizard-of-Oz study, plus the quantitative analysis of the words and expressions used to refer to graphs allowed us to construct new, hopefully more descriptive descriptions of the input graphs. One important characteristic is that the main requirement is to provide the user with the general information first followed by the details. Below is an example of our current descriptions given the UNA and WOZ studies:

```
Description: The graph is about paper clip sales. It's a one- variable line graph.
The y axis is millions of dollars and the x axis is month. The maximum value of the
y axis is 93, the minimum is 10. The graph has a general upwards trend. Evolution:
It is quite irregular. It starts at 63. There is a moderate decrease at x1 to 23.
There is a steady increase from 23 to 91 from x1 to x3. There is a steady decrease
from 91 to 54 from x3 to x5. There is a small increase at x6 to 71. There is a
moderate decrease at x7 to 10. There is a dramatic increase at x8 to 93. There is
a moderate decrease at x9 to 38. Finally, There is a steady increase from 38 to 69
from x9 to x11.
```

These descriptions generated by iGraph, although also reporting the graph point by point, summarizes part of the graph: What the graph is about (paperclip sales), what kind of graph it is (a line graph), the title of the axes (the Y-axis is sales in millions of dollars and the X-axis is month), a general trend line given by interpolation (an upwards trend), some idea of how "steep" the lines look (small, moderate or sharp increase/decrease), minimum and maximum values and a summarization such that every time there is more than one edge increasing or decreasing, the application will output a "steady increase/decrease" from point A to point A+n: for instance, in the above example: "There is a steady decrease from 91 to 54 from x3 to x5". This provides with a more manageable load of information, even if it is at the expense of losing some of the details (how much x4 decreases wrt x3 and how much x5 decreases wrt x4, for example).

According to the studies reported in these pages, iGraph does indeed help blind and visually impaired individuals with their interaction with graphs. This will hopefully build independence by not relying on other people to provide description of data graphs and a less marked aversion to graphs as an important data exploration tool.

Acknowledgements

This paper was written with the generous support of HTX (Health Technology Exchange), Cognos Inc., the Ontario Research Network for E-Commerce and The Cybercartography and the New Economy project funded by the Social Sciences and Humanities Research Council (SSHRC) of Canada under the Initiative on the New Economy (INE) Collaborative Research Initiative Grant. The authors would also like to thank the members of the HOTLab, Yandu Oppacher and Zhihong Li for expert advice on coding parts of iGraph, Jing Liu for constant

support and, especially, the participants who took part in the UNA and WOZ studies.

References

1. Shah, P., Carpenter, P.A.: Conceptual limitations in comprehending line graphs. Journal of Experimental Psychology: General **124**(1) (1995) 43–61
2. Bertin, J.: Semiology of Graphics. The University of Wisconsin Press, Madison, WI (1983)
3. Carpenter, P., Shah, P.: A model of the perceptual and conceptual processes in graph comprehension. Journal of Experimental Psychology: Applied, **4** (1998) 75–100
4. Pinker, S.: A theory of graph comprehension. In Freedle, R., ed.: Artificial intelligence and the future of testing. Lawrence Erlbaum, Hillsdale, NJ (1990) 73–126
5. Trickett, S.B., Ratwani, R.M., Trafton, J.G.: Real-world graph comprehension: High-level questions, complex graphs, and spatial cognition. (under review)
6. Kosslyn, S.: Understanding charts and graphs. Applied Cognitive Psychology **3** (1989) 185–226
7. Lohse, J.: A cognitive model for the perception and understanding of graphs. In: Proceedings of the SIGCHI conference on Human factors in computing systems, ACM Press (1991) 137–144
8. Foster, M.: Evaluating models of visual comprehension. In: Proceedings of EuroCogsci 2003. (2003)
9. Flowers, J.H., Hauer, T.A.: The ear's versus the eye's potential to assess characteristics of numeric data: Are we too visuocentric? Behavior Research Methods, Instruments, & Computers **24**(2) (1992) 258–264
10. Flowers, J.H.: Thirteen years of reflection on auditory graphing: Promises, pitfalls, and potential new directions. In: Proceedings of ICAD 05-Eleventh Meeting of the International Conference on Auditory Display. (2005)
11. Brown, L., Brewster, S., Ramloll, R., Burton, M., Riedel, B.: Design guidelines for audio presentation of graphs and tables. In: Proceedings of ICAD 2003 Workshop on Auditory Displays In Assistive Technologies, University of Boston (2003)
12. Ramloll, R., Yu, W., Brewster, S., Riedel, B., Burton, M., Dimigen, G.: Constructing sonified haptic line graphs for the blind student: first steps. In: Proceedings of the fourth international ACM conference on Assistive technologies, ACM Press (2000) 17–25
13. Flowers, J., Buhman, D., Turnage, K.: Cross-modal equivalence of visual and auditory scatterplots for exploring bivariate data samples. Human Factors **39**(3) (1997) 341–51
14. Brown, L., Brewster, S.: Drawing by ear: Interpreting sonified line graphs. In: Proceedings of ICAD 2003, ICAD (2003) 152–156
15. Riedel, B., Burton, A.: Perception of gradient in haptic graphs: a comparison of virtual and physical stimuli. In: Eurohaptics 2001 Conference Proceedings. (2001) 90–92
16. Riedel, B., Burton, A.: Perception of gradient in haptic graphs by sighted and visually impaired subjects. In: Proceedings of Eurohaptics 2002, Edinburgh University (2002) 134–137
17. Flowers, J.H., Hauer, T.: Musical versus visual graphs: cross-modal equivalence in perception of time series data. Human Factors **37**(3) (1995) 553–69

18. Brown, L., Brewster, S., Ramloll, R., Riedel, B., Yu, W.: Browsing modes for exploring sonified line graphs. In: Vol. II Proceedings of British HCI 2002. (2002)
19. Flowers, J.H., Buhman, D.C., Turnage, K.D.: Data sonification from the desktop: Should sound be part of standard data analysis software? ACM Trans. Appl. Percept. **2**(4) (2005) 467–472
20. McCoy, K.F., Carberry, M.S., Roper, T., Green, N.: Towards generating textual summaries of graphs. In: Proceedings of the 1st International Conference on Universal Access in Human-Computer Interaction, ACM Press (2001)
21. Fasciano, M., Lapalme, G.: Postgraphe: a system for the generation of statistical graphics and text. In: Proceedings of the 8th International Workshop on Natural Language Generation (INLG-96). (1996)
22. Corio, M., Lapalme, G.: Generation of texts for information graphics. In Saint-Dizier, P., ed.: Proceedings of the 7th. European Workshop on Natural Language Generation (EWNLG'99), Toulouse (1999) 49–58
23. Fasciano, M., Lapalme, G.: Intentions in the coordinated generation of graphics and text from tabular data. Knowl. Inf. Syst. **2**(3) (2000) 310–339
24. Mittal, V.O., Carenini, G., Moore, J.D., Roth, S.: Describing complex charts in natural language: a caption generation system. Computational Linguistics **24**(3) (1998) 431–467
25. Kerpedjiev, S.M.: Automatic generation of multimodal weather reports from datasets. In: Proceedings of the third conference on Applied natural language processing, Association for Computational Linguistics (1992) 48–55
26. Goldberg, E., Driedger, N., Kittredge, R.: Using natural-language processing to produce weather forecasts. IEEE Expert **9** (1994) 45–53
27. Reiter, E., Dale, R.: Building applied natural language generation systems. Journal of Natural Language Engineering **3**(1) (1997) 57–87
28. Reiter, E., Dale, R.: Building natural language generation systems. Cambridge University Press (2000)
29. Yu, J., Reiter, E., Hunter, J., Sripada, S.: A new architecture for summarising time series data. In: Proceedings of INLG-04 Poster Session. (2004) 47–50
30. Yu, J., Reiter, E., Hunter, J., Sripada, S.: Sumtime-turbine: A knowledge-based system to communicate time series data in the gas turbine domain. In: Proceedings of IEA/AIE-2003. (2003) 379–384
31. Sripada, S., Reiter, E., Hunter, J., Yu, J.: Summarising neonatal time series data. In: Proceedings of 2003 Conference of the European Chapter of the Association for Computation Linguistics. Companion Volume. (2003) 167–170
32. Dahlbäck, N., Jönsson, A., Ahrenberg, L.: Wizard of oz studies – why and how. Knowledge-Based Systems **6**(4) (1993)