Exploration of Non-seen Diagrams

David J. Bennett
Department of Computer Science, University of York
York, U.K.

Alistair D. N. Edwards
Department of Computer Science, University of York
York, U.K.

Abstract

This paper describes an exploratory experiment investigating access to non-seen diagrams with a view to presenting such diagrams through an auditory interface. Sighted individuals asked questions of a human experimenter about diagrams they could not see, in order to learn about them. The dialogue was recorded and analysed. The analysis resulted in an insight into the strategies used by the participants and a handle on the information requirements of the participants. Results showed that participants could understand and internalise the simpler diagrams, though not with complete success, but faltered on the more complex diagram. Several strategies and points for further investigation emerged.

1 Introduction and Objectives

Diagrams are a very powerful means of communication – for those who can see them. For those who cannot, because of a visual impairment, it is very hard to produce an alternative, non-visual means of conveying the same information. This paper describes part of a project which is investigating ways of producing auditory alternatives to diagrams.

There have been a number of previous attempts to translate diagrams into auditory forms. The particular form of the line graph has been translated into tones of varying pitch by a number of researchers (e.g. [2]; [7]). Blenkhorn [1] has produced a non-visual form of software engineering diagrams, while Stevens et al [9] produced and audio version of syntax trees.

As a starting point for this project, it was necessary to ascertain the kind of information that a person needs when exploring a diagram through an auditory interface. To that end we undertook an investigation whereby a human interpreter took the part of the interactive information system, answering questions about a diagram posed by subjects who were not allowed to see any visual representation of it.

Having said that diagrams can be very effective for sighted people, it is not invariably true that they are more effective that sentential representations (Green and Petre [3]). In order to tell when they are Larkin and Simon [6] use the terminology of *information efficiency* and *computational efficiency*. Put briefly, this states that two representations have information equivalence if all the information in one representation can be inferred in the other, and vice versa. Computational equivalence is defined to be that the time taken and the effort required to draw an inference from two different representations is the the same. A diagram will be more effective than a sentential representation with information equivalent if it is more computationally efficient for the set of tasks performed by a user. Similarly one diagrammatic representation may be better than another of the same information efficiency if it has a greater computational efficiency, again for the set of tasks performed by a user.

We believe that successful representations tend to be so because they offer the information required by a user at a reasonable computational cost. The representation will tend towards being optimised for commonly performed tasks. In seeking representations for blind users to replace visual ones, it would be an oversight to change the task-computational cost profile of the original representation in the alternative. By this, we mean that easy tasks in the original representation should be comparatively easy in the alternative. What affects this computational cost profile?

Zhang and Norman [11] state that external representations (of which diagrams are one type) can provide memory aids; anchor and structure cognitive behaviour, by for example enforcing laws and rules; change the nature of a task, by converting rules and laws in the head into constraints of the diagram and also say that external representations are

an indispensable part of any distributed cognitive system. A fundamental problem of auditory representations is that they do not provide that external representation on which the person can operate; sounds are transient and cannot be directly manipulated.

Localisation [6] is the principle of grouping all information used for a task in a spatial manner. This can be done for both elements and attributes of elements in diagrams, avoiding large amounts of search and symbol matching to make a problem solving inference. The spatial arrangement facilitates this; "...once the first piece of information has been found, the next useful piece of information will be in an adjacent location. Search becomes nothing more than moving among adjacent locations, and nothing needs to be stored in memory" [10, page 170]. Larkin and Simon also state that "Diagrams automatically support a large number of perceptual inferences, which are extremely easy for humans" (p. 98). To this we would add the proviso that the human in question can see the diagram.

The task of presenting diagrams in sound is a complex one. It was therefore decided that certain simplifications should be applied, whereby we could initially investigate the feasibility of the idea. Thus we decided to use diagrams with well defined syntax and semantics and further to limit these to 'node and link' diagrams. This is a class of diagram found in many technical environments, with examples such as: Data-Flow Diagrams, flow charts, PERT charts, class diagrams and electronic and electrical circuit diagrams.

The experiment described in this paper was intended to assist in the design of a computer interface to accomplish presentation of diagrams in sound. The experiment used electronic circuit diagrams, as there were a good number of locally available users of these. The use of experienced subjects (albeit not blind) avoided the need to learn the domain by the participants. Abilities of the participants varied with respect to domain knowledge, varying between novice and expert. Sighted subjects were used since it was not practical to find sufficient blind electronic engineers, but it was appreciated that there may be differences between the strategies and abilities of blind and sighted people in the tasks performed.

2 Methodology

The experiment was based on a prototype exploratory aural interface to diagrams. The experimenter took on the (eventual) role of a computer by answering questions posed by the participants.

2.1 Subjects

There were five sighted undergraduate (deemed novice and intermediate) and one sighted post-doctoral researcher (deemed expert) who took part in the experiment. Although the subjects were sighted, no attempt was made to "simulate" vision loss in the subjects except by not enabling them to see the diagrams.

2.2 Materials

Three different circuit diagrams were used in the experiment: a flip-flop, a half adder and part of a display driver for a seven-segment display. Each electronic component of the diagram had a label which roughly indicated the component's function. (E.g. 'AND' gates were labelled A1, A2 etc.) The connections between the components were not labelled explicitly, but distinguished by the components at each end and the position of the connections onto those components. (e.g. the "top input to A2".)

The diagrams were selected because of their varying numbers of components, connections, inputs and outputs. The flip-flop circuit should have been familiar to all participants, the undergraduates having been taught about it some weeks previously. The half-adder had not yet been taught to the undergraduates in their digital electronics course, though some knew of it from previous experience. Both of these two diagrams were expected to be 'idiomatic' in the eyes of the expert, in that they would be known by the user and be standard component combinations. The last diagram was not expected to be idiomatic to any of the participants.

The presentation of information to the participants was made by spoken word, live, by the experimenter. The experimenter did not have a written script or standard way to answer each question, simply because it was not possible to anticipate what questions would be asked.

2.3 Procedure

The procedure involved the participants asking questions about diagrams which were only visible to the experimenter. The participants were not allowed to make external representations (e.g. sketches) during the experiment. Each question was answered by the experimenter, unless it was either impossible to do so, or was felt by the experimenter to be beyond the capabilities of the eventual computer system.

The participants asked questions until they felt they knew, or understood, the diagram, or until they felt that the could not learn the diagram. After each question-and-answer session for a diagram, comprehension questions were posed to the participants. These tested knowledge and understanding of the diagram. Further questions about the diagram, by the participant, were allowed in order to answer these comprehension questions.

After the final diagram, a questionnaire covering the participants' opinion on ease of understanding and memory of the diagrams and salience of information given was completed.

3 Results

The tape recording failed for the interview with the expert, so results regarding the frequency of questions could not be obtained.

3.1 Frequency and Type of Questions asked

A summary table of the questions asked by the novice participants was produced. The questions entered into the table were paraphrased versions of the questions asked by the participants so that similar questions could be grouped together.

3.2 Overall strategies employed by participants

¿From the tape recording, it was found that participants used a two stage process to 'read' the diagrams. The process was roughly as follows:

1. Orientation
2. Reading
   (a) Ask about Node
   (b) Ask about Connections
   (c) Remember Node and follow a connection

Interestingly, this process shows some similarities to the model of audio-tactile exploration presented by Kennel [5, Figure 5, page 54]. This result tends to support the suggestion that using sighted participants was not inappropriate.

The Orientation part of the process was a short session of questions used to familiarise the participants with the overall nature of the diagram. One quarter of the questions asked fell into this category. All participants asked questions about the inputs and outputs of the circuits. Some asked additional questions regarding the nature of the 'stages' of the circuits and various other miscellaneous questions.

By our definitions, the Reading phase is that which follows the Orientation phase. That is to say that it is hard to characterize Reading phase questions, except that they are *not* orientation questions. Once a participant had moved on into the Reading phase he rarely asked questions from the Orientation phase set. The Reading stage provided the participant with the majority of the information about the functionality, or behaviour, of the diagram. Just under 75% of the questions came from this category.

It is interesting that users did not use either a regular depth-first or breadth-first search of the nodes on a consistent basis, but some kind of hybrid approach. The majority of participants worked from the inputs to the outputs in some fashion.

Participants asked questions which could be categorised as either 'Enquiries' or 'Confirmations'. Enquiries were questions such as "What is connected to element A1?", whereas Confirmations might comprise a question such as "So, A1 is connected to O1?". The participants generated 231 Enquiries and 86 Confirmations. Of the Confirmations five were incorrect (i.e. of the form "So, that implies *X*", when in fact *X* was untrue). Many Confirmation statements were made immediately after an Enquiry statement, perhaps indicating an attempt to internalise the information.

Termination of the question-and-answer sessions could occur for a number of different reasons. Participants were either happy that they understood the diagram, or felt that they were not going to be able to understand it, and were saturated with information that they could not process. The latter reason was typified by statements such as, "I don't think I'm going to understand this." Only one participant was interested in what proportion of the diagram had already been viewed.

3.3 Post Experiment Questionnaire

The post-experiment questionnaire covered three factors: ease of understanding, ability to remember and salience of information given by the experimenter for each diagram.

Using the rating supplied, Diagram 3 (the display driver) was seen by the participants to be significantly harder to understand ($F(2,8) = 32.44$, $p < 0.05$), but no diagram was significantly harder to remember than any other. Evidence from the tape recording showed that no one could remember all of Diagram 3, though the other two diagrams were remembered by many of the subjects.

Participants were asked to rate the salience of the information given to them by the experimenter in response to their questions. The scores on this question varied widely which appears to have been due to differing interpretations of the question, so a re-phrasing of the question to reflect this would have produced more informative results.

A last section of the questionnaire asked for suggestions on how to make the task easier. There were no suggestions.

4 Discussion

All of the subjects commenced by asking Orientation questions. This suggests that tools should start with a high level guided tour of the diagram. The tour might start at the highest level and descend under the user's control. For example, in a circuit diagram, the tour might start by stating the number of gates. At a request from the user, the number of gates of each type might then be listed, and so on. Whenever the user decides that enough detail has been given (or there is no more orientation information available) then the system would move on to Reading mode.

Nearly half the questions asked related to connectivity. That implies, with respect to this task, that a convenient means of accessing connectivity information would be a useful facility.

There was little interest in relatedness in the diagrams other than connectivity, perhaps because the first two diagrams were relatively simple. Only one participant asked about 'stages' of the circuit — an important concept in structuring an electronic circuit into smaller parts. We would have expected this to have been useful to other participants, if they had thought of it.

Neither was there any interest in the location of the diagram parts. This means that participants must have been *constructing* their own, functionally equivalent forms of the original diagrams in their head. According to Green and Petre [4] construction of structurally similar Visual Programming Language diagrams suffer from major difficulty in two Cognitive Dimensions: 'Imposed Guess-Ahead' and 'Viscosity'. These dimensions would also come into play in a mental construction process. Diagram constructors have to decide where to place each element before they know what else needs to fit into the diagram (Imposed Guess-Ahead). When change is required because a user has put a diagram element in the wrong place, a lot of effort is required to effect this change (Viscosity). This would not be necessary if users constructed facsimiles to the original form internally. To do this location would have to be described for each element.

It is likely that lack of experience with performing the task of reading non-seen diagrams meant that participants did not try to either, break the diagram into smaller sub-parts, or use the original designer's diagram form to aid them complete the task.

5 Conclusions and Further Work

This experiment and a theoretical task analysis of diagram usage form the basis of the interface of a research tool that enables individuals to explore Central Heating System diagrams. The tool is, in itself, not designed to solve in a single sweep the problems of blind diagram access, but to investigate various aspects of diagram access.

One major area of interest to us is the model of the diagram and method of navigation presented to a user of the tool. Different diagrams types have different ways of looking at them. The participants in this experiment used (predominantly) one view of the diagrams (the connection based view). There was little or no use of a decomposition view, breaking the diagrams into smaller sub-parts which should be easier to remember. Similarly they only tended to use one kind of navigation around the diagram: element to element. Diagrams, however are not normally read in one particular manner exclusively, by sighted people, but read and used in a number of different ways. We hypothesise that the match-mismatch hypothesis will be equally as valid for different views and navigation methods on non-seen diagrams as it is for diagrams and sentential forms [3]. By showing that different types of task are performed more easily with a representation that 'fits' the task, we hope that future work on alternative interfaces, not only of diagrams, but all representations, will pay more attention to the range of tasks to be performed and the models required to support those tasks.

As hinted in the discussion one aspect is the presentation of position of components within the diagram. Rigas successfully used scales of notes to present two dimensional position of items in his work [8]. The hypothesis is that presenting position of the components using this already proven method should reduce the time taken to complete tasks on the diagrams using the tool when compared to not presenting the position at all.

These two hypothesis will be investigated using the tool created with the help of this experiment.

References

[1] Blenkhorn, P. and Evans, D. G.: A Method to Access to Computer Aided Software Engineering (CASE) Tools for Blind Software Engineers. In W. L. Zagler et al (eds) Computers for Handicapped Persons: 4th International Conference ICCHP '94. Springer-Verlag 1994; pp. 321-328.

[2] Edwards, A. D. N. and Stevens, R. D.: Mathematical representations: Graphs, curves and formulas. in D. Burger and J.-C. Sperandio (ed.) Non-Visual Human-Computer Interactions: Prospects for the visually handicapped., 1993; Paris: John Libbey Eurotext. pp. 181-194.

[3] Green, T. R. G. and Petre, M.: When Visual Programs are Harder to Read than Textual Programs. In Human-Computer Interaction: Tasks and Organisation, Proceedings ECCE-6 (Eds van der Veer, G. C. et al), 1992.

[4] Green, T. R. G. and Petre, M.: Usability Analysis of Visual Programming Environments: A 'Cognitive Dimensions' Framework. Journal of Visual Languages and Computing 1996; 7:131-174.

[5] Kennel, Andrea R.: AudioGraf: A Diagram-Reader for the Blind. In: ASSETS 96: The Second Annual ACM Conference on Assistive Technologies. The Association for Computing Machinery, 1996, pp 51-56.

[6] Larkin, J. H. and Simon, H. A.: Why a Diagram is (Sometimes) Worth Ten Thousand Words. Cognitive Science 1987; 11:65–99.

[7] Mansur, D. L., Blattner, M. and Joy, K.: Sound-Graphs: A numerical data analysis method for the blind. Journal of Medical Systems, 1985; 9: pp. 163-174.

[8] Rigas, D. I.: Guidelines for Auditory Interface Design: An Empirical Investigation, PhD Thesis at Loughborough University, Leicestershire, UK, 1996.

[9] Stevens, R., Harling, P. and Edwards, A. (1996). Reading and writing syntax trees for phrase structured grammars with a speech-based interface. In D. Burger (ed.), New technologies in the education of the visually handicapped, (Paris), John Libbey Eurotext. pp. 271-276.

[10] Winn, W.: An Account of How Readers Search for Information in Diagrams. Contemporary Educational Psychology 1993; 18:162–185

[11] Zhang, J. and Norman, D. A.: Representations in Distributed Cognitive Tasks. Cognitive Science 1994; 18:87–122.