

Comparing Interaction Strategies For Constructing Diagrams In An Audio-only Interface

Oussama Metatla, Nick Bryan-Kinns & Tony Stockman

Interaction, Media and Communication group

Department of Computer Science

Queen Mary, University of London, UK.

{oussama, nickbk, tonys}@dcs.qmul.ac.uk

ABSTRACT

Although research on non-visual access to visualisations is steadily growing, very little work has investigated strategies for constructing such forms of representation through non-visual means. This paper describes the design of two interaction strategies for constructing and manipulating relational diagrams in audio. We report on a study that compared the two strategies, and discuss their advantages and disadvantages in terms of how efficiently they support the activity of constructing diagrams in an audio-only interface.

Categories and Subject Descriptors

H.1.2 [Models and Principles]: User/Machine Systems - *human factors*. H.5.2 [Information Interfaces and Presentations e.g., HCI]: User Interfaces - *auditory (non-speech) feedback, interaction styles*

General Terms

Design, Experimentation, Human Factors.

Keywords

Diagram construction, Auditory display, Interaction strategies, Accessibility.

1. INTRODUCTION

Diagrams are a ubiquitous form of representing information. They are extensively used in a variety of contexts including transportation, advertising, education and technical reports. Furthermore, many of us construct diagrams for many purposes; such as to understand a procedure or to brainstorm ideas.

Current advancement in technology is changing the way we interact with information however. As everything around us is being computerised, we find ourselves creating and interacting with more and more digitally represented information through smaller and smaller devices. The gradual disappearance of the screen in these devices is certainly challenging traditional paradigms for representing and interacting with information. Moreover, without a visual display or the ability to visually access an interactive environment, we no longer can benefit from the advantages offered by inherently visual artefacts such as diagrams [2]. We therefore seem to be at a point where it is

not only appropriate but also crucial to explore novel ways for representing information in order to support efficient interaction in various contexts of activity. The tasks of constructing, exploring and retrieving information using diagrams provide a good context for investigating alternative interaction strategies. This is because designing effective means for presenting, navigating through and constructing such representations on digital devices is difficult without a visual display, yet important in contexts where the user's eyes are occupied or in the case of visual impairment.

Although research on non-visual access to visualisations is steadily growing, very little work has investigated strategies for constructing such forms of representation through non-visual means. There are of course manual solutions for doing so, using physical artefacts such as pins and cardboard [1], but these tend to be inadequate for handling complex graphics and for allowing flexible storage, editing and reproduction. The few existing computer-based solutions that support non-visual construction of graphical representations combine audio and haptic technologies. For instance, Kamel *et al.* designed an approach for supporting grid-based image sketching [3], and Yu *et al.* developed a Web-based application for creating virtual haptic graphs [9]. Recently, an audio-haptic application for constructing bar graphs was developed by McGookin *et al.* [4], and another for creating graphical images by Rassmus-Gröhn *et al.* [7]. Evaluations of these applications showed that users could effectively construct and manipulate graphical representations non-visually. In this paper, we describe the design of two interaction strategies for constructing relational diagram in an audio-only interface and report on a study that compares their efficiencies.

2. AUDIO-ONLY ACCESS TO DIAGRAMS

2.1 Relational Diagrams

We focused our investigations on relational diagrams. These are graphical representations that depict items in a particular domain and how they conceptually relate to each other. Examples of such diagrams include flowcharts, abstract maps, and modelling notations such as UML. Figure 1 is a general example of such form of diagrammatic representation.

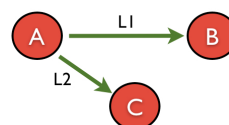


Figure 1. A simple relational diagram.

2.1.1 Multiple Perspective Audio Hierarchies

We designed a multiple perspective audio hierarchy to provide structured audio-only access to relational diagrams. The

hierarchy captures the represented relational information from a nodes perspective and a links perspective [5]. A user can browse through the hierarchy using the cursor keys on the keyboard in a similar way to typical file explorers and receives auditory feedback in the form of verbal and non-verbal sounds [6].

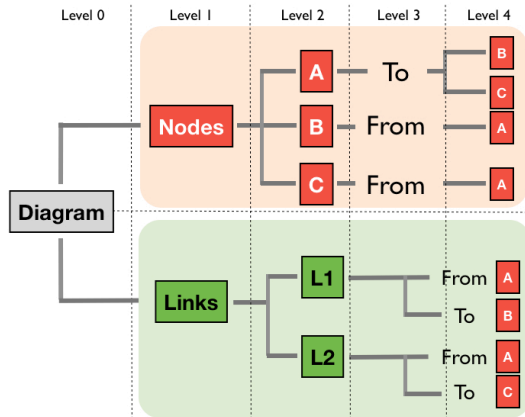


Figure 2. The multiple perspective hierarchical approach for accessing the relational diagram in Figure 1.

The hierarchy is also overlaid with special navigation capabilities that extend basic hierarchical navigation to allow more flexible access to its parts. For instance, the user can jump to designated parts of the hierarchy without having to browse to them, and can easily switch between the different perspectives.

3. CONSTRUCTING DIAGRAMS

To support non-visual construction of relational diagrams through the multiple perspective audio hierarchy, we designed two interaction strategies: a *Guided* and a *Non-Guided* strategy.

3.1 Guided Strategy

In a guided interaction strategy, the system plays the role of an agent that assists the user when executing an editing task. The system and the user engage in a conversation-like interaction, where they collaborate by exchanging information to realise the desired task. The user invokes the system’s assistance on a particular action by triggering a series of system prompts. By responding to each prompt, the user is guided through the necessary steps required to accomplish an editing action (Figure 3).

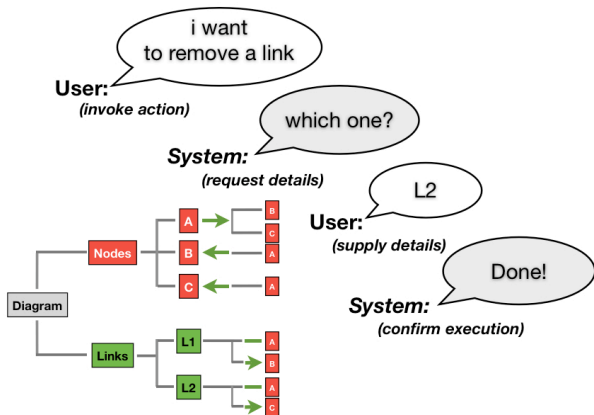


Figure 3. In a Guided interaction strategy the user responds to systems prompts independently from the hierarchy.

The series of system prompts that the user responds to depend primarily on the action they invoke to initiate an editing task. This strategy therefore allows the user to execute any editing action on any given item without having to locate it within the hierarchy. That is, when editing an item on a diagram, the user does so away and independently from the hierarchy. Thus, in a guided interaction strategy, exploring and editing the diagram are separated into two distinct modes of interaction; an exploratory and an editing mode.

3.2 Non-Guided Strategy

Conversely, the system in a non-guided interaction strategy does not provide any explicit assistance to the user. To construct or edit an item in a diagram, the user must first locate it within the hierarchy before editing it. Figure 4 exemplifies this concept; to remove a link from a diagram, the user must first explore the appropriate path to locate the link on the hierarchy. Once found, the user can issue the desired editing command. The system then interprets the current position combined with the issued command as one complete editing expression, and executes it appropriately.

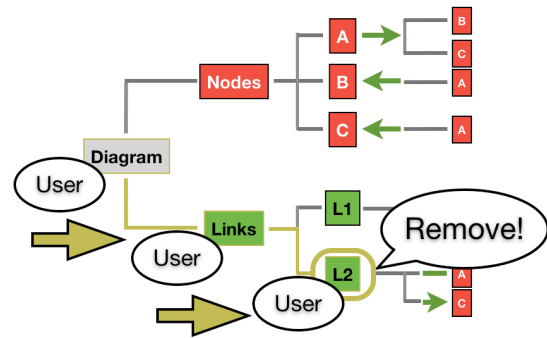


Figure 4. In a Non-Guided interaction strategy the user locates the item on the hierarchy in order to edit it.

Thus, there is no implied intermediary between the user and the hierarchy. The accomplishment of an editing action depends primarily on the status of the hierarchy at the moment it is executed, and the user can directly perceive the changes in its status as soon as it is completed. In this strategy the tasks of exploring and editing the diagram are therefore combined into one complementary mode.

4. EXPERIMENTAL STUDY

4.1 Overview

The two interaction strategies differed primarily in the way they structure the interaction in terms of combining or separating exploratory and editing modes. We were particularly interested to know whether this difference has any effect on users ability to build and retain a coherent understanding of the diagrams they construct. Because the system’s prompts that users must respond to are mainly in the form of speech output, users will require extra time to both receive and react to the extra spoken feedback in the guided strategy. We therefore hypothesised that:

- H1: Diagram construction in a non-guided strategy requires less time than in a guided strategy.

Because the non-guided strategy combines both exploratory and editing actions into one mode, users must simultaneously concentrate on two tasks to manipulate a diagram. Users in the guided strategy on the other hand can focus on editing without being distracted by exploring the hierarchy. We hypothesised that:

- H2: Diagram construction in a guided interaction strategy will promote more coherent understanding of the diagram.

To test these hypotheses, we manipulated interaction strategy as an independent variable in a within-subjects experimental design.

4.2 Experiment Design

4.2.1 Participants & Setup

12 males and 12 females volunteered for the study. They were a mixture of undergraduate and postgraduate students studying computer science, engineering and business. Experimental sessions were made up of three phases. In each phase, participants were given written descriptions of simulated systems and were asked to construct Entity-Relationship (ER) diagrams to model them. No time limit was imposed and participants were informed that they could take as long as they wished to construct their diagrams. All scenarios were of similar complexity. Participants controlled the interface through a computer keyboard and received auditory feedback through normal computer speakers. All participants were sighted and were not blindfolded.

4.2.2 Test Conditions

Given a system description, participants in a control condition were asked to identify and list potential entities and potential relations and to write these down on formatted sheets. Since most participants were familiar with such task, the control condition was used as a benchmark against which their performance using the auditory interface could be compared. In the second and third conditions, participants were given two further system descriptions, one at a time, and were asked to construct an ER diagram using an auditory interface that implements one of the strategies. Sufficient training was provided before the actual tests.

4.2.3 Post-Construction Tasks

After constructing each diagram, participants were asked to answer a series of questions about it. The questions were designed to test the users understanding of information about the diagrams that was not explicitly mentioned in the written text. Such information could only be revealed through the construction process. For examples the question: "Which Entity was involved in most of the relations?", could only be answered once the diagram has been constructed. They were asked a total of 14 questions.

4.2.4 Measurements & Data Gathering

We measured completion times for constructing the diagrams in the second and third condition as well as participants scores on post-construction questions in the three conditions as dependant variables. Informal interviews were also conducted at the end of each experimental session where we asked each participant to elaborate on their experience. All sessions were video recorded.

5. RESULTS

5.1 Completion Times

Participants spent an average of 23 minutes (SD=5.7) to complete a diagram when using the guided interaction strategy, and 19 minutes (SD=6.4) when using the non-guided strategy. A t test showed that these differences are significant ($t=3.039$) at $p<0.005$, which confirms hypothesis H1.

5.2 Scores

Participants scored an average of 76% on the post-construction tasks when using the guided interaction strategy (SD=8.5), 77% when using the non-guided strategy (SD=11.3), and an average

of 80% in the control condition (SD=10.8) (Figure 6). We again used a t test to calculate the levels of significance in the variance between individual scores across the three conditions. These were not significant for the differences between the control and the non-guided conditions ($t=1.138$), and not significant between the non-guided and the guided conditions ($t=-0.563$). Variances in scores were however significant between the control and the guided conditions ($t=1.927$) at $p<0.05$. Contrary to what was expected, these results do not support hypothesis H2.

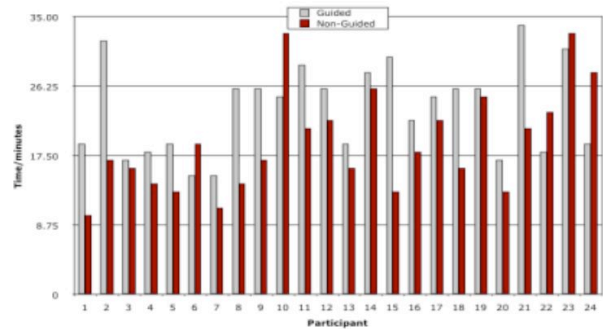


Figure 5. Total construction times for each participant.

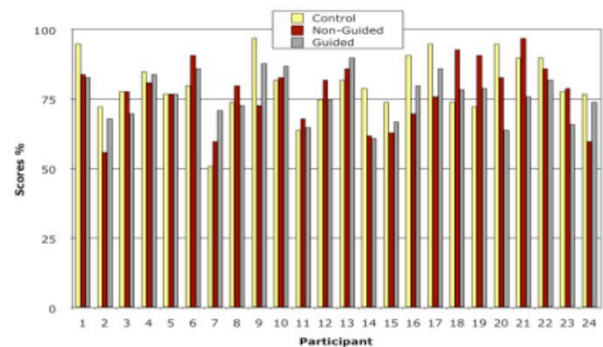


Figure 6. Total scores on post-construction tasks.

6. ANALYSIS & DISCUSSION

6.1 Completion Times

Even though editing actions in both strategies were carefully designed to involve the same number of operational steps, participants spent significantly more time executing such steps in the guided strategy than they did in the non-guided strategy. Therefore, having the system act as an intermediary agent between the user and the hierarchy significantly increases interaction times.

It is of course possible to decrease interaction times by decreasing the amount of speech and increasing the speed of the spoken output, but this must be properly handled as processing rapidly spoken information is not a trivial task without prior training. More concise speech output can be provided by using Spearcons [8] for example. These are brief speech-based audio cues that have been found to speed up performance on menu navigation while preserving the intelligibility of an interface.

6.2 Diagram Comprehension

Our second hypothesis predicted that the participants would demonstrate richer understanding of the constructed diagrams when using the guided strategy. This was not supported. In comparison to the control condition, the results showed that participants performed significantly better on the post-construction tasks when using the non-guided strategy. From

the questionnaire data and the interviews, the majority of participants asserted that they found it harder to build a fuller understanding of the constructed diagram and to keep track of what has been constructed when using the guided strategy because they felt detached from the representation.

Nonetheless, regardless of their overall preference, all participants had a positive response to the ability of executing editing actions without having to 'physically' locate items within the hierarchy; an ability only afforded by the guided strategy. We see this as an advantage of using the auditory medium to interactively access and manipulate a graphical artefact. Audio in this instance allowed a certain level of representational freedom, which would not have been available if the artefact was accessed through a more direct and representationally dependent medium. For example, it would not be possible to use a pen (or a PHANTOM Omni haptic device) to edit parts of a diagram drawn on physical or virtual paper unless the pen is directly situated on the part that one wishes to edit.

6.3 Editing Categories

We could identify two distinct categories of diagram editing activities along which participants' preferences were structured.

6.3.1 Global Editing

Global editing includes actions that alter the overall state of one perspective in the hierarchy, and typically results from adding new nodes or new relations to the diagram. This category of editing therefore only affects levels 1 and 2 of the hierarchy (Figure 2). The majority of participants did not have any particular preference between the two strategies when performing global editing.

6.3.2 Local Editing

Local editing includes actions that alter the state of an individual item within a particular perspective, thus affecting levels 3 and 4 of the hierarchy (Figure 2). When using the guided strategy, most participants found themselves unable to keep track of the progress of some local editing actions. This is because in this strategy, the deeper the item is located within the hierarchy the more system prompts a user must respond to in order to edit such item.

Furthermore, participants tended to execute local editing actions systematically. That is, they preferred to complete all local editing related to a particular item before moving on to the next. The guided strategy did not support such systematic editing in a way that the participants found useful. Progressing through the same set of prompts more than once for each item, in addition to the potential of losing track, affected performance times and increased users' frustration with the interface. The non-guided strategy on the other hand allowed for the systematic execution of local editing actions in a way that matched how the participants planned and structured the construction. A user would have to locate the desired item only once, then, while situated within that part of the hierarchy, they can issue all the necessary local editing commands.

6.4 Mode Confusion & Management

The guided strategy structures the interaction into two distinct modes, an exploratory and an editing mode. This made participants prone to mode confusion. By observing participants' interaction and examining the logged data, we were able to identify two instances where such confusions occurred. Both instances were mainly due to a combination of: 1) participants distraction from the construction task, and 2) system's failure to properly convey information about its current state.

6.4.1 Within-Mode Confusion

Participants would lose track of where they were within a given system-user dialogue of information exchange. For example, a user maybe distracted from the audio interface by shifting their attention to the written text in order to supply the system with the requested detail, and after returning back to the interface, they would forget which step they were at, and fall into a within-mode confusion.

6.4.2 Cross-Mode Confusion

In the second instance, participants would lose track of which mode they were currently in. For example, after initiating an editing action, shifting their attention to the written text and back to the interface again, a user would forget that they were in the middle of an editing sequence. They might carry on interacting with the system as if they were in the exploratory mode. Expecting it to behave accordingly and unable to properly interpret its feedback, they fall into a cross-mode confusion.

6.4.3 Preventing Mode Confusion

We initially used different gendered voices to distinguish between interaction modes. However, this was not enough to prevent the users from falling into the two described mode confusions. When interviewed about the speech feedback, most participants reported that they were indeed aware of the presence of two different voices in the system, but failed to recall where and what each voice was used for.

A more efficient solution to avoid within-mode confusions would be to include a 'request status' command that would provide the user with a detailed description of the current status of the system, including whether it is in a standby awaiting input. In a previous study [6], we found that using continuous ambient sounds in the hierarchy efficiently communicate contextual information. A possible solution to avoid cross-mode confusion, therefore, would be to introduce a special ambient sound that dominates the auditory interface when a particular mode is initiated.

7. CONCLUSION

We presented the design of two interaction strategies for supporting the construction of relational diagrams in an auditory interface and reported on a study that compared their efficiencies. A non-guided strategy was found to support faster interaction, while a guided strategy allowed a certain degree of representational freedom that reduces browsing and exploration steps. Both strategies promoted similar levels of diagram comprehension.

8. REFERENCES

- [1] Broockshire, G. R., Teaching UML Database Modelling to Visually Impaired Students. *Issues in Information Systems* (2006), Vol. VII no. 1, 98-101.
- [2] Cheng, P.C-H., Lowe, R. & Scaife, M. Cognitive Science Approaches to Understanding Diagrammatic Representations. *AI Review* (2001) 15, 5-27.
- [3] Kamel, H. M., Landay, J. A., Sketching Images Eyes-Free: a Grid-Based Dynamic Drawing Tool for the Blind. In *Proc. of the 5th International ACM Conference on Assistive Technology*, Edinburgh, Scotland, (2002), 33-40.
- [4] McGookin D. K., Brewster, S. A., Graph Builder: Constructing Non-Visual Visualisations. In *Proc. of the 20th Annual British HCI Group conference*. London, UK, (2006), 263-278.
- [5] Metatla, O., Bryan-Kinns, N., Stockman, T., Auditory External Representation: Exploring and Evaluating the Design and Learnability of an Auditory UML Diagram. In

Proc. of the International Conference on Auditory Display, Montréal, Canada, (2007), 411-418.

- [6] Metatla, O., Bryan-Kinns, N., Stockman, T., Using Hierarchies to Support Non-Visual Access to Relational Diagrams. In Proc. Of the 21st Annual British HCI Group Conference. Lancaster, UK, (2007), 215-225.
- [7] Rasmus-Gröhn, K., Magnusson, C., Efring, H., Iterative Design of an Audio-Haptic Drawing Application. Extended Abstract In Proc. of ACM Special Interest Group on Computer-Human Interaction, CHI 2007, Montreal, Canada, (2007).
- [8] Walker, B. N., Nance, A., Lindsay, J., Spearcons: Speech-Based Earcons Improve Navigation Performance in Auditory Menus. In Proc. of the 12th International Conference on Auditory Display, London, UK (2006).
- [9] Yu, W., Kangas, K., Brewster, S. A., Web-based Haptic Applications for Blind People to Create Virtual Graphs. In Proc. of the 11th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, (2003), 318-325.