AUDIO REPRESENTATION OF GRAPHS: A QUICK LOOK

Andy Brown, Robert Stevens, Steve Pettifer

School of Computer Science
University of Manchester
Oxford Road, Manchester M13 9PL. UK
`andrew.brown@cs.man.ac.uk, robert.stevens@manchester.ac.uk,`
`steve.pettifer@manchester.ac.uk`

ABSTRACT

This paper examines the use of audio glances for giving an impression of the size, complexity and topology of abstract graphs. The first step in nearly all reading tasks, but particularly those of complex structured information such as equations, tables and diagrams, is a glance. This brief high-level overview of the information allows the reader to start to understand the nature of his task, and to develop strategies for reading. Yet a glance is currently unavailable to visually impaired readers. We describe an algorithm for generating earcons that present such glances through non-speech sound. An evaluation demonstrated that these were successful in conveying an impression of a graph to sighted volunteers. The success of this evaluation means visually impaired readers can now start their graph-based tasks with some of the benefits a glance can bring.

1. INTRODUCTION

This paper describes the design and evaluation of an audio glance at graph-based diagrams for visually disabled people. Such node-arc graphs (see figure 1) are commonplace and routinely used by many people in many disciplines to present and solve a variety of problems, from the simplest everyday tasks to complex abstract calculations.

The widespread use of graph structures in a variety of diagrams, coupled with their relatively constrained structured form, make them a good target for initial studies in presenting information more complex than simple text to visually disabled people. Simple text is essentially a linear presentation in speech. At its most complex it is a stack of lines forming a two-dimensional page. Such a presentation is easily dealt with by standard keyboard facilities to move up, down, left and right at a simple spectrum of granularities.

Moving away from such a presentation immediately increases the difficulty of reading. Algebraic notation, for instance, increases the complexity of the presentation by increasing the information density and through nesting, almost adding an extra dimension to the information which means that sighted readers do not read them in a simple linear fashion [1]. Similarly tables are truly two dimensional and cannot be read adequately in the same way as text. A reader needs to compare cells, rows, and columns whilst maintaining a knowledge of headings etc (for example [2, 3]).

Graph based diagrams are a increment on this spectrum, moving towards general diagrams without incurring the obvious difficulties in pictures. As Figure 1 demonstrates, a graph is a series of arcs and nodes. These may be labelled or unlabelled and span a range of complexity, determined by factors such as number of arcs and nodes and arcs per node. The use of a simple building block that can be increased to arbitrary complexity means basic presentation can be relatively simple, while bringing in several features of general diagrams. Finally, such diagrams include maps; UML and ER diagrams; circuit diagrams, trees (genealogy, evolutionary, etc.); and so on. Increasing access to such diagrams obviously opens up an enormous body of information to visually disabled people.

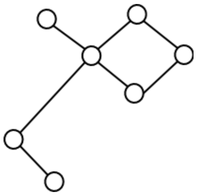


Figure 1: *An abstract graph of nodes and arcs.*

The first stage in many everyday graph reading tasks, from planning a journey to reading, is a glance. A quick glance at a graph to gain an impression of its size, content, complexity, etc. takes very little time but gives the reader a broad understanding of the nature of the graph and consequently the nature of his task. In the simplest of problems this glance may give sufficient information for an understanding to be reached with little further effort; in more difficult cases it serves merely to provide context – the reader has a rough idea of the problem complexity and can start to develop strategies for reaching an understanding, including strategies for reading the graph in more detail. Furthermore, glances are often used for orientation once more detailed exploration has begun.

What if the user is unable to see the graph? It is technically possible to enable a visually disabled user to move between nodes in some application (for example [2, 4]), reading the details as they go and *possibly* building an understanding of the topology. This is a bottom up strategy; how does such a reader get the head start sighted readers gain from their glance? One might argue that a glance is *more* important for the non-seeing user; if he or she has an approximate understanding of a graph's complexity and topology before commencing node to node exploration, he or she should be able to keep their orientation better and hopefully seek out the relevant parts of the graph more easily.

For such users the only channels available are haptic and audio. While the former offers potentially greater interactivity it requires sophisticated hardware; using audio glances, however, requires no more than a sound-card and headphones (or speakers). What is required from an audio glance if it is to achieve the same objectives as its visual counterpart? We propose the following characteristics.

1. It must be short. A glance is by definition a quick activity; more sophisticated overviews can be provided in the next stage of exploration. It is, however, unrealistic to expect them to be as short as visual glances; a few seconds is probably sufficient, and considerably shorter than 'bottom-up' exploration.
2. It must give an impression of the size and complexity. Even in the largest and most complex of graphs just knowing that it is large and complex is useful. For the simplest of graphs the listener may decide it is so trivial that reading each node and arc in turn is an appropriate strategy.
3. It could give an idea of topology, but this is not essential. The visual glance merely gives an *impression*; the audio glance should do likewise. Although it is impossible for a glance to convey the detailed topology of large and complex graphs, useful information may still be gleaned. For example it would be beneficial to identify that a graph has a simple linear section connected to a group of highly connected nodes
4. It is better to be vague than misleading: a misleading glance could cause great confusion, while a vague one merely demands further effort to develop the necessary overview.

We have built an audio glance at graphs as part of the Kekulé project [4]. This is a general graph reading tool to enable investigation into how various forms of annotation can benefit readers. Providing an audio glance summarising the graph can be considered as a form of annotation on the graph itself. In this paper, we set the scene with an overview of some previous work (Section 2). We then describe the development of this glance (Section 3) and its Web based evaluation (Section 4). Section 5 gives the results, while in Section 6 we discuss the general success of our glance and its uses in non-visual reading of graph based information. We also look at avenues for future research.

2. BACKGROUND

The audio summaries described in this paper are descended most directly from the work of Stevens *et al.* [5]. Theirs was probably the first sonification of complex structured data; they demonstrated that it was possible to generate audio glances of mathematical equations. Despite the obvious differences in the structure of the information being presented, this current work is in many respects an extension of theirs: the intention is identical – to provide a glance that facilitates planning prior to reading. Stevens used different timbres to represent different types of element within the expression (operands, fractions, sub-expressions, etc) and manipulated the timing, pitch and amplitude according to a set of rules depending on the structure of the equation. Complex items were presented using a continuous tone. They found that listeners were able to derive sufficient syntactic information from the earcons to enable them to select the correct equation from among four alternatives.

Probably the only previous research explicitly aimed at sonifying graphs is Brewster's work with hierarchies. They examined if earcons could be used to present to users their position in a hierarchy [6, 7]. They tested both compound earcons and hierarchical earcons. In the more successful, compound earcons, different timbres were used to represent different numbers, and the hierarchy position was generated by concatenating. For example 3.2 would represent the second child of the third child of the root node. Using a tree (hierarchy) of 27 nodes in 4 levels, they found participants were able to identify their location by listening to a compound earcon with an accuracy of 97%. This work was aimed at giving specific information about position in a graph rather than an impression of its overall structure and as such is comparable not with the overview glances described in the bulk of this paper but with the highlighting in phase four (Section 4.3). Although successful, it is a much more constrained approach, that could not easily be generalised beyond trees.

Hermann and Ritter [8] describe an approach for 'interactive exploratory sonification of high-dimensional data' which is quite closely related to creating audio glances of graphs. Firstly their technique is designed to give an *impression* of the data as a first step in data analysis, and potentially allow one to spot features that would otherwise be hidden within the high dimensionality of the data. Secondly the sonification is actually performed on a graph – the network graph formed by employing a growing neural gas network on the data; here the neurons form the nodes while the edges are the connections between them. They sonify this graph by treating it as masses connected by springs; when a mass is excited, either directly by the listener or by another node via its connecting springs, it emits a noise. The entire network can be excited by 'shaking' it or the data can be explored by exciting a particular node and listening to the sound spread. This approach is very general and could well be used for summarising graphs rather than data. The interactive aspect is particularly appealing.

Another method for presenting spatial information is to scan over the data in a known pattern, sonifying each data point as it is passed over. This can be combined with two-dimensional or pseudo two-dimensional sound – the approach taken by Zhao *et al.* for presenting geo-spatial data [9]. Their 'spatial chloropleth map' presented an audio glance of data from US states by scanning the map, sonifying the data while speaking the state names. The stereo position gave the longitude, while the latitude was given by the pitch of a piano note. Volunteers were able to identify the class of data pattern (e.g., vertical strips) with 56% accuracy after a single listen (c.f. 20% for random guess). This type of sonification appears to require a spatial arrangement of the data that will fit a simple scanning pattern; it would be interesting to repeat the experiment replacing the states of the US with the counties of the UK, for example. More generally research into sonification of numeric data has much to offer in terms of use of different aspects of sound to present different aspects of data (parameter mapping; Barrass and Kramer discuss this briefly in their overview of sonification [10]).

3. AUDIO GLANCES OF GRAPHS

Earcons are abstract, structured, non-speech sounds [11] and as such are suitable for presenting structured information (e.g., graph topology) in a rapid manner. In our prototype system we use earcons to present the audio glance (we use the terms 'earcon', 'audio glance' and 'audio summary' interchangeably in this paper). These are generated by an algorithm that uses both the topology and spatial layout of the graph.

Our algorithm generates earcons from graph representations that describe both the topology (i.e., to which other nodes each node is connected) and layout (i.e., co-ordinates for each node).

The algorithm is essentially breadth-first, and works by following the connections from the left-most node. This node, identified by its co-ordinates, is played first – a tone with a duration of 100ms. After a pause of 300 ms, all nodes connected to it are played, with 50 ms between each. After another pause, all nodes connected to these nodes are played. This continues until all nodes have been played. In addition to representing the distance of a node from the left most node (as measured by number of arcs, rather than space) by time, the stereo axis is used to represent the vertical spatial dimension. This is done by detecting the centre of gravity along the y-axis, then the further above this line a node is, the further left its sound is played, and vice versa.

This design of algorithm ensures that some of the requirements outlined above are met. The earcons it generates are short for typical graphs (the glance generated for the graph shown in 2 is approximately two seconds long), and have the potential for conveying size, complexity and topology. The evaluation described in Section 4 is designed to determine if this potential is realised and all the requirements met.

It might be argued that mapping connective distance to time and spatial y distance to left-right is confusing, but the authors preferred this algorithm to a purely spatial one (where the two spatial axes are mapped to time and stereo) for its better representation of the graph topology. In order to keep the size of the experiment reasonable it was decided to test one algorithm; if this showed promise for the principle of summarising graphs with audio glances further experiments will be performed to identify effective algorithms.

The second algorithm used in this experiment was identical but for the addition of some 'ambient' sound. In this case the beeps representing the nodes were annotated with a sound to signify the presence of a ring. This sound was continuous, playing from just before the first node in the ring until just after the last. If there was more than one ring in a graph they were played at slightly different pitches.

4. THE EVALUATION

The aim of the evaluation was to determine whether the earcons generated by our algorithm meet the requirements outlined in Section 1. One experiment was performed to test this in two ways. The first tested this indirectly, by seeing if a graph could be recognised after using an audio glance; the second tested it more directly by asking participants to listen to a glance, then write a description of their impression. A range of graphs with different sizes, complexities and topologies was used.

The experiment itself was divided into four phases, which are described in more detail below. In summary, the first two phases presented a glance to the participants, then tested if they could recognise it from among three other graphs. These two phases differed only in a small change to the algorithm used to generate the earcons. The third phase required participants to listen to the glance then write down their impression of the graph, the quality of which was determined by asking them to match graphs to their descriptions. A fourth phase examined a different question, testing if summaries could be used as an orientation aid, highlighting particular nodes in a graph. Although not a specific requirement for a glance, the ability to orientate oneself with a glance is one of their possible uses. Each phase consisted of five questions. All graphs were abstract undirected graphs, with between 3 and 15 nodes, which resulted in audio summaries between 1.3 and 4.5 seconds long.

Achieving a success rate for matching graphs that is significantly greater than that expected by chance would support our hypothesis that the earcons successfully give an impression of the graph. To test the final requirement, that it is better to be vague than misleading, participants were asked to give the confidence they had in the correctness of each answer. A significantly lower confidence for wrong answers would indicate that, in general, participants were aware of when they were guessing, and would therefore be less likely to be misled in practice.

Evaluations were performed over the web using pages generated using PHP and Java applets [1]. These were designed to make cheating difficult, and each user action (listening to a glance or loading a page) was recorded. Participants were free to do the evaluation where and when they wished and were not supervised. Volunteers were sighted computer users recruited from the School of Computer Science at Manchester University. Volunteers were encouraged to use headphones for the evaluation, although stereo speakers were acceptable. They were asked which they were using and the answer recorded.

4.1. Phases 1 and 2

These were identical apart from the algorithm used to generate the audio summaries. Their aim was to see if the algorithms were successful in conveying an impression of the graph. This was tested by playing an audio glance, then asking the participants to match it to one of four graphs. Pilot studies indicated that presenting the graphs and sound together resulted in participants playing the sound multiple times, each time attempting to match the beeps directly to the nodes on one of the graphs. This was clearly not testing if they had formed a valid impression, so their actions were restricted and the procedure altered to the following:

1. Present an audio summary of a graph. This was played when the participant pressed a button and could only be heard once.
2. Ask the participant to select the graph they think the summary represented, from a choice of four. The options were not visible when the sound was initially heard. Moving from the listening phase to the selection phase required action by the user – it was hoped that this would limit their ability to remember the precise pattern of the glance and concentrate instead on the impression it gave.
3. Allow the participant to listen again to the summary, as many times as they wish, while viewing the four choices, and ask them to make a second selection. This would allow us to see if multiple glances were of any benefit. Each listen to the sound was recorded.

This was repeated on five graphs for each of the algorithms (see Figures 3 and 4). Distractors were designed to test participant's ability to use the glances to distinguish certain characteristics of the graphs. Questions 3 and 5 included graphs of different sizes. For the remainder all choices were eight node graphs – this

---

[1] Although results are no longer being collected, the evaluation may still be completed at `http://aig.cs.man.ac.uk/people/andybrown/audioeval/intro.php`.

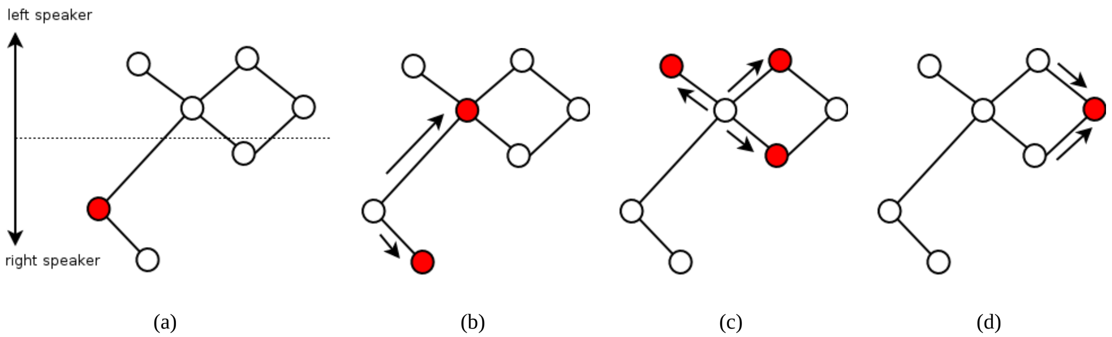


Figure 2: *The algorithm as applied to the graph in Figure 1. The horizontal line in (a) represents the y-value of the centre of gravity of the graph; nodes close to this line are played in the centre while nodes above and below are played towards the left or right respectively. The left-most node is played first. After a 300 ms pause the two nodes connected to it are played, one beep just left of centre, and another beep on the far right. Next all nodes connected to these two are played, and so on until all nodes have been played.*

was to properly test the ability to distinguish complexity and topology rather than just size. For both selections the user was asked to rate the confidence they had in their decision. This used a slider bar that was initially positioned at 50%; higher values (to the right) indicated greater confidence in the decision. As part of some related work participants were also asked which of the four graphs they thought were most and least complex, and which pairs most and least similar. These latter questions were asked after stage 3 above so as not to interfere with the main task.

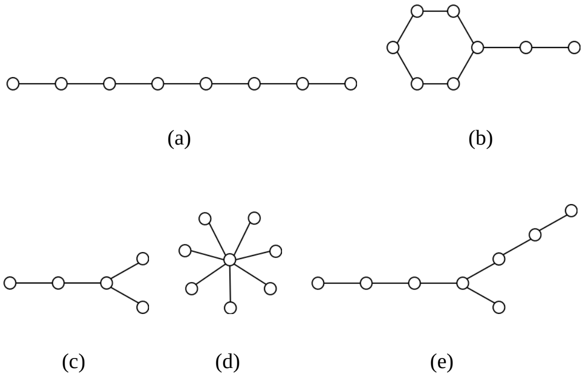


Figure 3: *The graphs used in phase one.*

4.2. Phase 3

The authors were concerned that phases one and two were testing the ability of participants to recognise graphs from their sounds, which was not necessarily the same as testing whether they gained a useful impression of the graph topology. It should be noted, however, that successful recognition would still suggest that the glances are giving useful information about the graph. Phase three aimed to test more directly whether people were able to identify anything about the graph (size, complexity, topology, etc.) from the summaries.

Participants were presented with five buttons for playing audio glances of five graphs of different size and complexity (see Figure 5). Each glance could be played up to two times. Next to the buttons were text fields for them to enter a description of the impression given by each summary. After submitting their descriptions, the volunteers were presented with 13 different abstract graphs. They were also presented with their five descriptions, although in a different order, and with each description a selection of check boxes. They were asked to select which graph they thought their description matched. Although encouraged to only select one, they could choose more graphs if they were undecided. This was intended to allow us to identify if, when a particular graph could not be identified, the type of graph could be recognised. As in the earlier phases they were given a slider to rate the confidence they had in their choice.

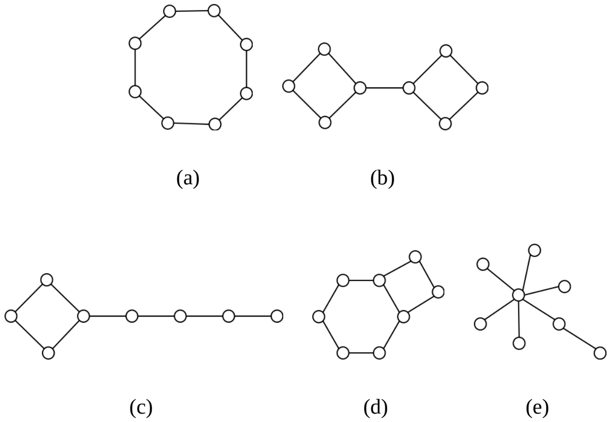


Figure 4: *The graphs used in phase two.*

4.3. Phase 4

The final phase of the experiment was to test whether audio glances could be used as an orientation aid during exploration, as well as providing context at the start of reading. To give orientation in-

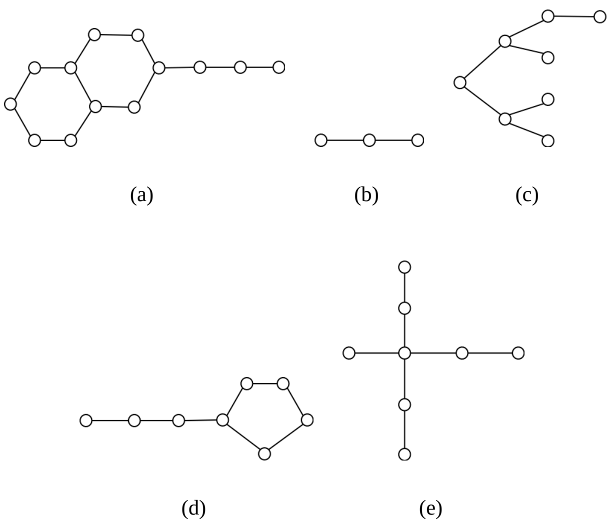


Figure 5: *The graphs used in phase three.*

formation one or more nodes of the graph were highlighted in the summary. This was achieved by playing them at a higher pitch, and was intended to give the listener an impression of their current position in the context of the whole graph. The ability of people to identify which graphs were highlighted was tested as follows.

1. An audio glance of a graph was presented without highlighting. Volunteers could listen up to five times.
2. On pressing the continue button they were shown the graph visually. This was intended to replicate detailed non-visual exploration, since orientation aids are only useful once one has started exploring.
3. On pressing continue button again the graph was hidden and they were shown a button that enabled them to listen to a summary where one or more nodes were highlighted. Hiding the graph while playing the glance avoided problems of direct matching of beeps to nodes identified in the pilots of phase one.
4. After the second listen the graph was shown again and they could listen no more. They were asked to select which nodes were highlighted; this was done by clicking on the appropriate nodes in the diagram.
5. On clicking the submit button participants were taken to stage one of the next graph. Five graphs were completed in this manner.

Completing the fifth graph of this phase represented the end of the experiment, and participants were asked for their comments. Figure 11 shows the graphs used and which nodes were highlighted.

5. RESULTS

Twenty-four participants completed the evaluation; the results are described in detail below. It is possible to make some general comments relating to all phases. Results showed that overall there was a high success rate, with the average participant identifying the correct graph in 13 of the 20 questions (65%). There was a wide variation in individual performance (standard deviation of 14%) with scores varying between 45% and 90%; these are shown in Figure 6.

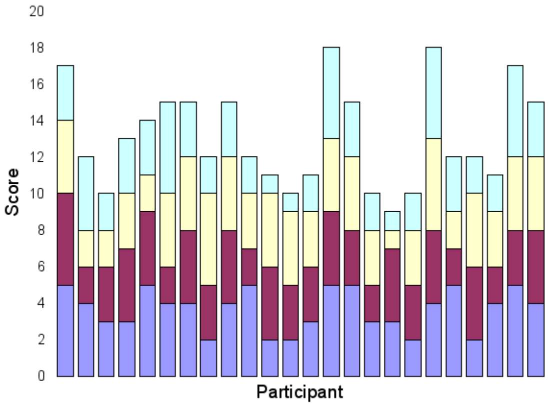


Figure 6: *The scores of individual participants for all phases. Each bar gives the score obtained by participant for each phase, with phase 1 at the bottom and 4 at the top. There was a maximum score of 5 for each phase.*

Several participants commented that they felt that they were improving as they went through the study – in the words of one:

> ‘Interesting learning as the test went along. I could definitely spot the structures (or something pretty close indeed) from the sounds by the end, although maybe only because a key dozen or so structures were used so I knew what I was listening for.’

Despite claims by those piloting the study that it was easier using headphones, the difference between those listening through headphones (67% correct) and those using speakers (59% correct) was not statistically significant.

5.1. Phases 1 and 2

Figure 7 shows the percentage of participants that selected the correct graph for each question, both after the first listen and after multiple listens. Questions 1 to 5 use the plain algorithm, for questions 6 to 10 the audio summary also highlighted rings. This shows that recognition was, on the whole, good, with an average of 70% of participants choosing the correct graph after just one listen. Interestingly there was no significant difference between this and the mean percentage of people correct after multiple listens (77%). Participants achieved an average of 7 questions correct out of 10; a binomial test for this, with a probability of success being 0.25, gives a probability of this result happening once by chance of 0.003.

Although there was no improvement in scores from annotating the sounds to indicate the presence of rings, analysis of the results shows there was a clear advantage in having it – only once was a graph not containing a ring selected when the audio glance contained this ring sound. This was on the first question for the algorithm and the mistake was rectified after listening again. There was more difficulty in identifying the topology of fused rings, as evidenced by questions 7 and 9. The correct graph and the close dis-

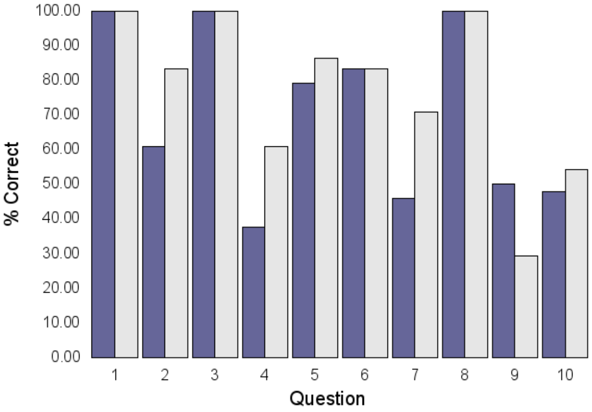
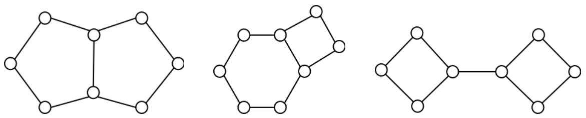


Figure 7: *Results for phases 1 and 2. The left hand bars show the percentage of participants that selected the correct graph after one listen, the right had bars after multiple listens.*

tractor for question 9 are shown in Figure 8; all participants recognised the presence of rings but 50% were unable to distinguish between the correct graph (8(b)) and the close distractor (8(a)) after the first listen. The pattern of beeps is identical for these graphs so distinguishing them requires careful identification of where the two ambient sounds denoting the rings start and finish. Similarly, twelve people incorrectly thought the glance for question 7 (which represented the graph shown in Figure 8(c)) was actually for 8(a). The recognition rate improved (with six people changing their selection) after further listens for this question, presumably because the different beep pattern helped identification. These two questions, particularly question 9, probably require close to the limit of the level of detail one would expect from a glance.



(a) (b) (c)

Figure 8: *Double-ring graphs were difficult to differentiate.*

The star topology of question 4 also presented some difficulty, although after further listens a recognition rate of 61% was achieved. Analysing the results with those from the same graph presented in phase 4 (Figure 11(c)) suggest that people might imagine the central node to be the one played first.

It is also possible to examine how accurate people's perceptions of correctness were, i.e., did they know when they might be wrong? If the mean confidence rating for correct answers (80%) is compared with the mean rating for incorrect answers (54%) it can be seen that participants usually had a reasonable idea when they were correct and when they were guessing. A two-tailed t-test (T=7.33) gave the probability of this difference being due to chance as less than 10^{-11}. There were still, however, cases where the participant was 100% confident, but wrong.

5.2. Phase 3

Figure 9 shows the percentage of participants that selected the correct graph for each question. The average score was 3.4 correct from the 5. With 13 graphs to choose from the probability of selecting 3 correct graphs by chance is 0.004.

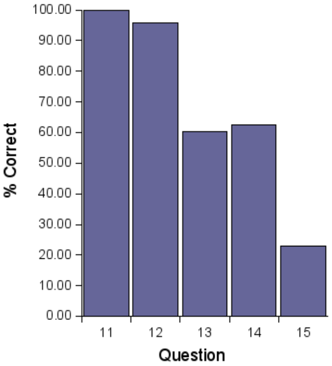


Figure 9: *Results for phase 3.*

Everyone got the first question right and all but one the second, and the descriptions matched the graphs well, indicating that the correct impression had been given. Nearly all descriptions for the first graph (Figure 5(a)) were variations on ‘Two rings then a chain’, and the second were ‘Short chain’ or ‘Three nodes in a chain’. The single error on graph 2 was where the description simply *‘A Chain’* and an 8-node linear chain was selected instead of the 3-node one.

There was more difficulty with graph 3 (Figure 5(c)); although some described it quite accurately (e.g., *‘binary tree on its side’*) most of the descriptions were more vague. Eight participants mentioned branching or forking, five used the term *‘tree’*, and five explicitly noted that there were no rings (e.g., *‘no rings just a lot of branches, medium complexity.’*); this was clearly noticed by most as only five participants selected graphs with rings (7 of the 13 choices contained one or more rings).

Graph four (Figure 5(d)) was identified by a little over half the participants. All but one participant identified that there was a chain followed by a ring, but seven erroneously thought the ring was followed by another node or two. Five of these selected a graph with two nodes then five nodes forming two joined rings, with one node hanging off the end.

The final graph (Figure 5(e)) caused the most difficulty. Again most people recognised that there were no rings in the structure – only three matches were made with ring structures. Fifteen (63%) recognised (and specifically mentioned) branching, tree-like structures or intersecting chains, but only four of these could match the graph with their description. One example is *‘a short chain then a branch to make a tree like structure’*; this is an accurate description of the graph, but its author was unable to make the correct choice.

5.3. Phase 4

Overall the results from the highlighting phase of the evaluation indicate that people were able to identify approximately which nodes were highlighted but not always the exact combination. Figures 10 and 11 summarise the results.

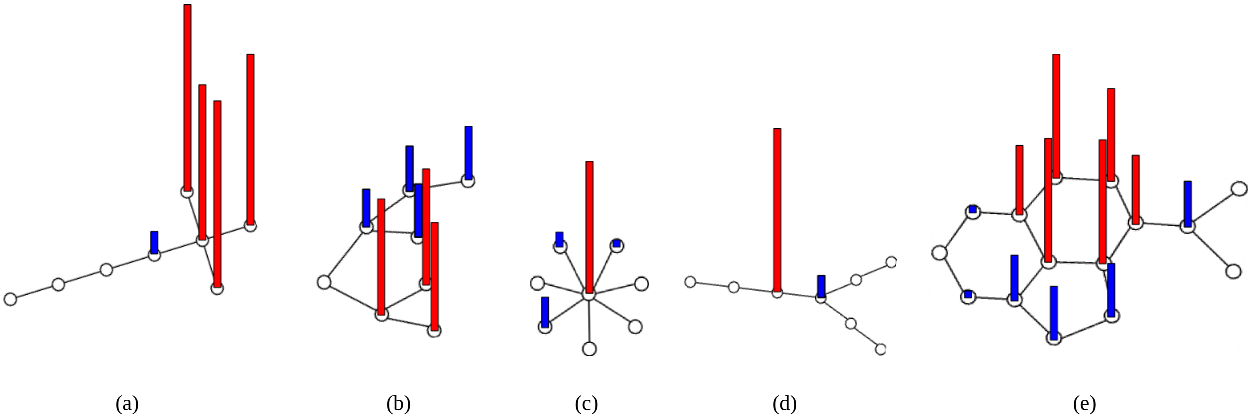


Figure 11: *The graphs presented in phase 4. The height of the bars represent the number of people who thought a node was highlighted; a red bar indicates it was actually highlighted, a blue bar that it was not. Some of the graphs have been rotated slightly for clarity (although this would have no effect on the sounds).*

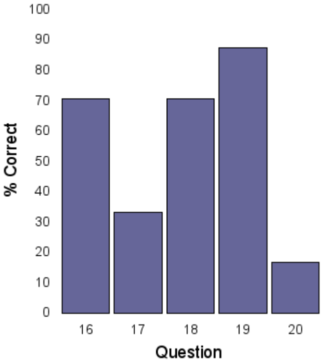


Figure 10: *Results for phase 4.*

These results show that there was little difficulty identifying the highlighted nodes in simple graphs where the beeps representing these nodes were not mixed amongst other nodes. For example, 88% of participants were correct in graph 4, and 71% in graphs 1 and 3, while the nodes in the lower branch of the tree of graph 2 were mixed amongst the nodes in the higher branch and it was difficult to identify exactly which were highlighted. The same was seen, to a greater extent, in graph 5. 11 shows, however, that it was possible to identify the approximate area of the highlighting. It was also easier to identify a single node than a cluster. In these cases all participants recognised that only a single node was highlighted even if they did not identify exactly which node(s) were selected. Where multiple nodes were highlighted, and particularly where they were mixed with other nodes, it was clearly more difficult – on graph 5, where 6 nodes were highlighted, participants selected between 2 and 9.

6. DISCUSSION

The results presented above confirm our hypotheses and indicate that the audio glances created fit the requirements outlined in Section 1. The algorithm ensures that glances are kept short, at least over the range of graphs used in this evaluation. Evidence that the size and complexity of a graph is conveyed by a glance is provided by the successful recognition over a range of graphs, supported by the descriptions given in phase three.

The third phase also demonstrates that successful recognition is through the formation of a mental impression of the graph rather than an ability to match glance to diagram. It is quite clear that some topology is conveyed by the glances, at least in simple cases. Evidence has also been presented that shows that wrong answers were associated with lower confidence, suggesting that the glances also fulfill the fourth requirement. Stevens use of a single tone to represent complex regions [5] might be worth investigating should misidentification prove a problem.

It is not too speculative to suppose that regular use of such glances would bring improved results. As described above, some participants described a learning effect through the evaluation. We can also examine the cause of some of the mis-identifications and attribute them to lack of familiarity. For example, although the extra sounds warning of the presence of rings were successful in preventing participants incorrectly identifying the sound as representing an acyclic graph, the converse was not true. It is reasonable to suggest that more practice would help users remember that the lack of a ring noise means that there are no rings present. Brewster's work [7] was essentially testing whether people could learn the earcons representing hierarchy position – their high recall rates suggest that patterns of sound can be learnt.

The use of these glances for highlighting has only been explored in a very limited field by Brewster ([6], described above), but this evaluation shows that particular nodes can be highlighted and their approximate location identified (or even the exact location in simple graphs). There are many possible applications for this type off highlighting in non-visual graph exploration. As mentioned in the introduction, it could be used as an orientation aid, reminding the user where in the graph they are, or as a marker to indicate where a graph has been edited. Search results could also be summarised in this way, giving readers a rapid idea of where in the graph the nodes of interest are located.

The evaluation has, however, highlighted some problems with the algorithm. There is a tension between topology, which is the real meaning of the graph, and layout. The knowledge of the layout of a graph is probably important to visually impaired users, particularly if they are to discuss it with a sighted colleague, and may be a useful orientation aid when exploring (e.g., to know that they are near the top right of the diagram). The algorithm, however, does not present both topology and layout. This tension is demonstrated in the final graph of phase three (Figure 5(e)); although descriptions were quite accurate, matching was not. We suggest that this was because the mental image of a branching tree involved branches directed to the right with the root on the left. The diagram, however, had two branches of nodes aligned vertically which might have been expected (incorrectly) to be played as one group. We would however argue that, in this case at least, the impression given was a good basis for further investigation. Approaches similar to that of Zhao *et al.* [9] may offer a solution – using 2D or pseudo-2D sound to provide further clues to the topology.

Another related criticism that might be levelled at this algorithm is its left-to-right reading (with up and down mapped to left-right stereo). The choice of starting node will have a big impact on the sound of the glance and how much information can be gleaned from it. One approach could be to sonify along the longest dimension of the graph, but if we are to use positional information in the later exploration, it is of no use having an apparently random orientation for the initial glance. It is not clear how to resolve the issue of topology versus layout, but the type of sonification used by Hermann and Ritter [8] might offer something, and its interactivity could be useful during later exploration – giving readers the ability to listen to the graph from different viewpoints. A further area for investigation is to examine how effective this algorithm is at representing more highly connected graphs; those used in this evaluation are on the whole quite sparse.

Less fundamental problems also arose. It was clear that, despite their usefulness overall, the noises presenting rings could be improved. The difficulties observed in distinguishing the graphs of Figure 8(b) and erroneous identification of an extra node or two to the right of the ring in the graph of Figure 5(d) indicate that participants found it difficult to identify the exact start and finish of these sounds. Further investigation could bring improvements. Some participants also complained that the speed was too great. Although we would argue that it is important that the glance is rapid, some facility for adjusting the speed of glances could also assist users, particularly in the early phases of learning.

Beyond the difficulties discussed above, there is also much scope for investigating other parameters of the algorithm such as the relative size of the pauses, the nature of the beep representing the nodes, and how best to distinguish multiple rings. Having demonstrated that such glances can be successful there is a need to apply it to real graphs. This would require investigation into how to present different node and arc types, directed arcs and various other features. A further avenue for investigation is to examine how audio glances like these may be used in an audio analogue of the coordinated multiple views of visualisation [12].

Nevertheless, these points reveal opportunities for investigating improvements to a successful audio glance at abstract graphs. These structures form an important part of much information and a glance at their form is an important aspect of improving access to many information sources.

7. ACKNOWLEDGMENTS

Andy Brown is funded by an EPSRC studentship.

8. REFERENCES

[1] R.D. Stevens, A.D.N. Edwards and P.A. Harling, “Access to Mathematics for Visually Disabled Students through Multimodal Interaction,” *Human-Computer Interaction (Special issue on Multimodal Interfaces)*, pp 47–92, 1997.

[2] L.M. Brown, S.A. Brewster, R. Ramloll, M. Burton and B. Riedel, “Design Guidelines for Audio Presentation of Graphs and Tables,” in *ICAD 2003 Workshop on Auditory Displays In Assistive Technologies*, 2003.

[3] Y. Yesilada, R. Stevens, C. Goble and S.Hussein, “Rendering Tables in Audio: The Interaction of Structure and Reading Styles,” in *Proceedings of the Sixth International ACM SIGACCESS Conference on Computers and Accessibility*, Atlanta, USA, October 2004, pp. 16–23.

[4] A. Brown, R.D. Stevens, and S. Pettifer, “Evaluation of a non-visual molecule browser,” in *Proceedings of the Sixth International ACM SIGACCESS Conference on Computers and Accessibility*, Atlanta, USA, October 2004, pp. 40-47.

[5] R.D. Stevens, S.A. Brewster, P.C. Wright and A.D.N. Edwards, “Design and evaluation of an auditory glance at algebra for blind readers,” in *Auditory Display: The Proceedings of the Second International Conference on Auditory Display*, Santa Fe, USA, 1994.

[6] S.A. Brewster, A. Capriotti and C.V. Hall, “Using compound earcons to represent hierarchies,” *HCI Letters*, vol. 1, no. 1, pp. 6–8, 1998.

[7] S.A. Brewster, V.-P. Raty and A. Kortekangas, “Earcons as a Method of Providing Navigational Cues in a Menu Hierarchy,” in *Proceedings of HCI’96*, London, UK, 1996, pp. 167–183.

[8] T. Hermann and H. Ritter, “Neural Gas Sonification – Growing Adaptive Interfaces for Interacting with Data,” in *Proceedings of the 8th International Conference on Information Visualisation*, London, UK, 2004, pp. 871–878.

[9] H. Zhao, C. Plaisant, B. Shneiderman and R. Duraiswami, “Sonification of Geo-Referenced Data for Auditory Information Seeking: Design Principle and Pilot Study,” in *Proceedings of ICAD 04 – Tenth Meeting of the International Conference on Auditory Display*, Sydney, Australia, 2004.

[10] S. Barrass and G. Kramer, “Using sonification,” *Multimedia Syst.*, vol. 7, no. 1, pp. 23–31, 1999.

[11] M. Blattner, D. Sumikawa and R. Greenberg, “Earcons and icons: Their structure and common design principles,” *Human Computer Interaction*, vol. 4, no. 1, pp. 11–44, 1989.

[12] M.Q.W. Baldonado, A. Woodruff and A. Kuchinsky, “Guidelines for using multiple views in information visualization,” in *Proceedings of Advanced Visual Interfaces, 2000*, Palermo, Italy, 2000, pp 110–119.