

A Comparison of 2-D Visualizations of Hierarchies

Todd Barlow

Padraic Neville

SAS Institute Inc.

todd.barlow@sas.com

padraic.neville@sas.com

Abstract

This paper describes two experiments that compare four two-dimensional visualizations of hierarchies: organization chart, icicle plot, treemap, and tree ring. The visualizations are evaluated in the context of decision tree analyses prevalent in data mining applications. The results suggest that either the tree ring or icicle plot is equivalent to the organization chart.

1. Introduction

The visualization of hierarchical data has received a great deal of attention from the graph drawing community and the human-computer interaction communities. One of the most persistent problems is the display of large trees. A traditional view of a tree quickly exceeds the capacity of most conventional means of displaying the tree. This paper describes two studies that compare tree layouts for their suitability to display decision trees used in data mining software.

Decision trees are the most commonly used analysis tool in data mining. In most cases, a decision tree takes hundreds of variables as input and predicts the value of a target variable. The decision tree depicts rules for dividing data into groups. The first rule splits the entire data set into some number of pieces, and then another rule may be applied to a piece, different rules to different pieces, forming a second generation of pieces. In general, a piece can either be split or left alone to form a final group containing the leaves of the tree. In data mining, decision trees can have hundreds of leaves. The predictive model created by the decision tree suggests which variables and which relationships among the inputs and target are important.

This multivariate model is complex. Visualization is the primary means of communicating the structure of the model. Novice analysts rely on an application's visualization tools to help them understand the rules in the model. Expert analysts can make preliminary judgments about the suitability of a decision tree by evaluating the tree's topology without reference to details about the model's rules.

The traditional way of drawing a decision tree is to use a variation of the Reingold and Tilford [1] algorithm. (In this paper, to make reading easier, this way of

drawing a decision tree will be referred to as the organization chart.) The organization chart, as discussed below, has some limitations in its ability to adequately handle the display of these trees. Our research compares the organization chart to three alternative layouts. In two experiments we evaluate the ability of the layouts to communicate the decision tree's topology and to support common decision tree analysis tasks.

2. Compact Views of Trees

The size limitations of computer screens limit the usefulness of the organization chart for displaying decision trees. It is possible to draw a tree of several hundred leaves, as an organization chart, so that the entire tree is visible. However, the size of the nodes makes the tree useless for most purposes. Tree nodes typically contain text or graphics that describe the contents of the nodes. If a large tree is drawn so that all the nodes are visible, the nodes will be so small that the node contents are not legible.

In response to this problem, we decided to add a compact tree view to our data mining application. A compact view displays the entire tree in a small window. It is intended to give the user an understanding of the tree topology thereby facilitating data analysis. An example of how a compact view could be used in decision tree analysis appears in Figure 1. Figure 1 shows three linked views. Selecting a point in the assessment plot displays that subtree in the Compact View window and the Tree window. In this very simple example, the decision tree component of the data mining application has produced nine subtrees. The Assessment Plot window displays the criterion value for each subtree. The first step in the analysis is the assessment of the criterion value. The analyst looks at the relationship between the two curves. If the plot suggests that the model is suitable, he/she assesses the tree's topology in the Compact View window. In this example, the subtree with 6 leaves is selected in the assessment plot and appears in the Compact View window. The compact view allows the analyst to do a preliminary analysis of the subtree while the Tree window shows details on nodes and splitting rules. The Compact View window and Tree window are linked. Selecting a node in the compact view brings the selected node into view in the Tree window. This interaction supports analysis of the decision tree. For

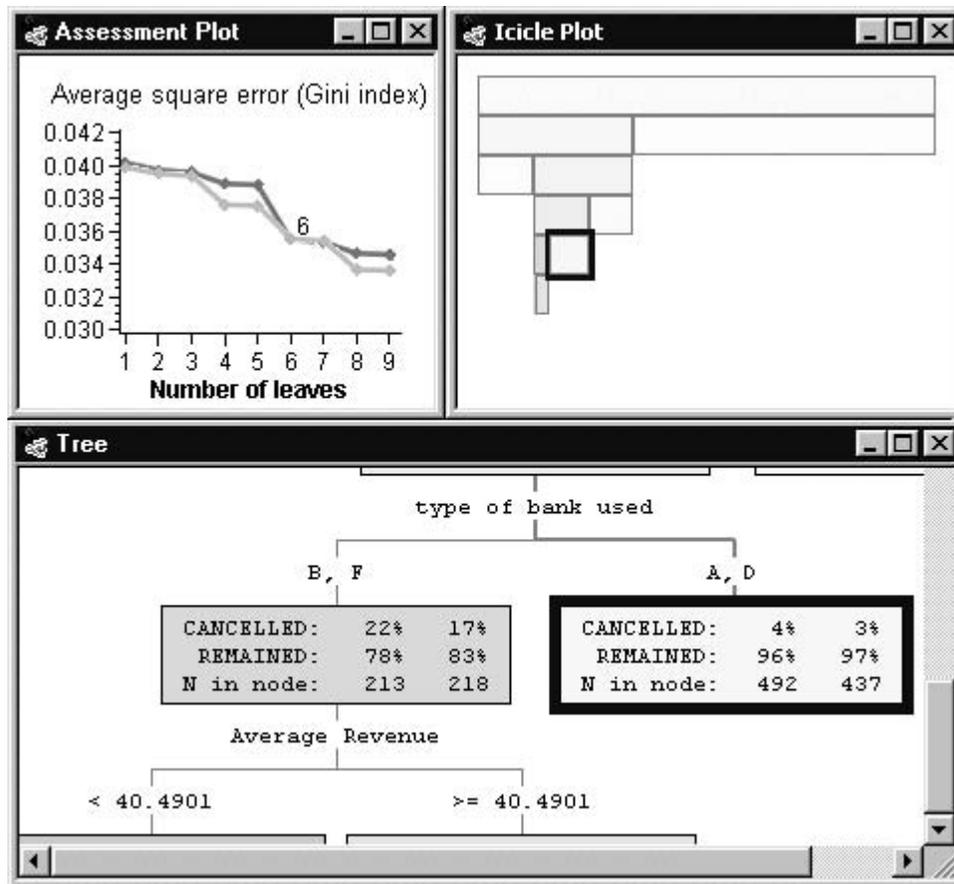


Figure 1. Example of compact view in data mining

example, the compact view can drive navigation in the Tree window.

Our goal was to create a compact view that would be as small as possible while providing useful information about the tree. The studies described in this paper were designed to help us choose a compact view for use in a data mining application. Figure 2 shows examples of the compact views that we compared.

Although an organization chart is the most common means of displaying a full-size decision tree, some data mining applications also use an organization chart as a compact view. These applications use the variation of the Reingold and Tilford algorithm. Our variation satisfies the four aesthetic rules stated in [1] and satisfies additional rules proposed by Herman, Delest, and Melancon [2]. The changes made to the algorithm minimize the overall width and height of the tree.

The tree ring is a space-filling visualization method. The tree ring displays tree topology and nodes size. Node size is proportional to the angle swept by a node. As with the organization chart, the tree ring has been used in data mining software to display decision trees. It is similar to the layouts used by Andrews [3] and Chuah [4].

The icicle plot is similar to the Kleiner and Hartigan's [5] concept of castles. Unlike the tree ring, the icicle plot

contains empty space. Like the tree ring, the icicle plot displays node size. Node size is proportional to the width of the node. At the time of these studies, this plot had not been used in a data mining application.

A treemap [6] is also a space-filling visualization method. It also displays tree topology and node size. It has been used to display hierarchical information in a number of domains including data mining.

3. Criteria for Evaluating Layouts

The tasks used to evaluate the compact views were based on the requirements of the user in a data mining context. The important requirements of a compact view are: (1) *Ease of Interpretation* - A user should be able to understand the parent-child and sibling relationships inherent in any tree. Analysts can look at a tree and determine if it is well-suited for their needs. The compact view should communicate the tree topology accurately and quickly. (2) *Comparison of Node Size* - One of the most important node characteristics is node size. Node size helps the analyst determine if the decision rules create meaningful or spurious groups. Small nodes are less interesting than large nodes because it is more cost effective and statistically reliable to target a few large

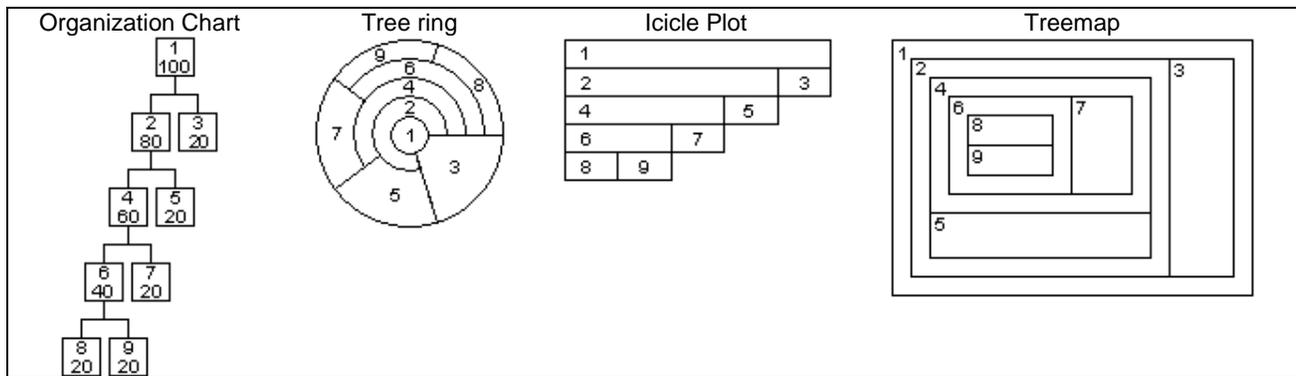


Figure 2. Different views of the same tree

groups than many small groups. (3) *User Preference* - Users sometimes have strong preferences for user interface elements. Although users may not always like what works best for them, their subjective evaluation of the compact views is an important part of the comparison.

4. Experiment 1

The first experiment compared the four views in their ability to communicate the topology of the tree and support comparisons of node size.

4.1. Design

The experimental design was a variation of a 4 (view) x 5 (task) repeated-measures design with tree type as a random factor. The design was not a full 4 x 5 factorial design because the organization chart was not used in one of the tasks.

4.2. Participants

Fifteen coworkers participated in the experiment. Eight of the participants were male and seven were female. None of the participants were familiar with the treemap, tree ring, or icicle plot. The participants were software developers, testers, or technical writers.

4.3. Tasks

We designed 5 tasks to test the ability of the view to communicate the tree topology:

- *Binary or N-Ary* - Participants had to decide if the tree was binary or n-ary. In the binary trees, all of the parents had two children. In the n-ary trees, most of the parents had more than two children. Participants pressed the *b* or the *n* key to select the binary or n-ary response, respectively.
- *Balanced Or Unbalanced* - Participants decided if the tree was balanced or unbalanced. Participants were told that balanced trees had leaves on the same level

or two consecutive levels. Otherwise, the tree was unbalanced. Participants pressed the *b* or *u* key to select the balanced or unbalanced response, respectively.

- *Deepest Common Ancestor* - When the tree appeared, two nodes were highlighted. The participant clicked on the deepest common ancestor of the highlighted nodes. Clicking on a node highlighted the node in a different color from the two highlighted nodes. Participants could select any node except the two highlighted nodes.
- *Number Of Levels* - Participants counted the number of levels in the tree. Participants entered the number of levels in the tree in the response window.
- *Three Largest Leaves* - Participants selected the three largest leaves by clicking on the leaves. Participants could select any node in the tree even if the node was not a leaf. No more than three nodes could be selected at a time. A response was scored as correct if the three largest leaves were selected. If the response included only 0 to 2 of the largest leaves the response was scored as incorrect. We decided that all of the views should be as small as possible because they were intended to be compact views. They would be used in addition to an organization chart containing text and graphics. This meant that none of the views could contain text. Since this prevented the organization chart from displaying node size information, it was not used in this task.

4.4. Procedure

We limited the overall size of the window containing the views to 25% of the area of the screen because they are intended to be compact views. They are also intended to be used along with other windows. Some trees did not fill the window while others used the entire window. The following paragraphs describe how we set the dimensions of each view.

In the organization chart, the node size and minimum distances between nodes was kept constant for all of the trees. The overall width and height of the tree varied with the number of leaves and nodes in the tree. The widest tree determined the node size and separation between nodes. This tree spanned the width of the monitor. Node spacing and node size were manipulated until a reasonable balance between the two dimensions was achieved. Node width and height was 6 pixels. The minimum horizontal distance between siblings was one pixel. The minimum horizontal distance between nodes that were not siblings was 2 pixels. The vertical distance between levels was 6 pixels.

All of the treemaps covered the same area and used the same offset between levels. We set the size of the treemap so that it would display the deepest tree with the most leaves and provide a large enough offset between levels so that participants could click on a level. To draw that tree within the window, the offset between levels was 4 pixels. A pilot study showed that it was not a problem for participants to interact with the tree at that size.

The icicle plot width was increased until the smallest leaves in the widest tree were the same width as the nodes in the organization chart. The overall width of the icicle plots remained constant for all of the trees. The height of levels remained constant for all of the trees. The height of the levels was determined by the overall height of the deepest tree. The height of the deepest tree was set so that the area of the plot was approximately the same as the area of the same tree using when it was drawn as an organization chart.

For the tree ring, the root node radius and distance between levels was kept constant for all the trees. The root node radius and distance between levels were equal to the height of the levels in the icicle plot. The area covered by the tree ring varied with the number of levels in the tree.

We created eight trees that combined three tree characteristics: number of children (2, more than 2), number of leaves (16, 100), degree of balance (balanced or unbalanced). All eight trees were used in every task.

Each participant saw each of the four compact views. The compact views appeared on a 20" monitor with a screen resolution of 1024 pixels x 768 pixels. All of the views were drawn with 1 pixel-width black lines on a white background.

Participants were seated approximately 24" from the screen. Each participant completed each task four times (with the exception of the Three Largest Leaves task), i.e., once for each compact view. The order of tasks and views was randomized across participants. All the tasks had the same sequence. The task began with two practice trials. During the practice trials, the experimenter explained the task and answered the participants' questions. Eight trials followed the practice trials. The

participant was not allowed to ask questions during these trials. The data from these trials was used in the analysis. All of the tasks and trials were self-paced. No feedback was provided on the accuracy or speed of their responses. Response time was the time elapsed between the appearance of the tree and the pressing of the spacebar indicating the end of the trial. After completing all the tasks with all the views, participants ranked the compact views according to their preference for them.

4.5. Results

Response times (RT) were normalized using a log transformation before they were analyzed. RT was analyzed using a repeated measures analysis of variance with post-hoc (Tukey) comparisons of the mean RT. Response accuracy (RA) for each task was analyzed using a logistic regression with generalized estimating equations approach to fitting the logistic regression model. Comparisons were made between the percent correct for each of the views within each task. The top part of Figure 3 shows RT for each view in each task. The bottom part of Figure 3 shows percent correct for each view in each task. The results of the post-hoc comparisons appear below the horizontal axes. The mean response times for views with the same shade of gray are not significantly different ($p < 0.001$). For example, in the Balance task, the mean IP time is significantly less than the mean OC time, which is significantly less than the mean TR and TM times. The TR and TM times are not significantly different. The percent correct for views with the same shade of gray are not significantly different ($p < 0.001$). For example, in the Largest task, the percent correct for IP and TR are not significantly different. The percent correct for TM is significantly less than that of IP and TR.

Analysis of RT and RA suggests that there is not much difference in participant performance when using the icicle plot, tree ring, and organization chart. RT for the treemap was always slower than RT for at least one of the other plots. In four of the tasks RT for the treemap was the slowest of the four plots. Similarly, RA did not show much difference between three of the plots. RA for the treemap was less than RA for the other three plots in three of the tasks.

The results of the preference rankings showed participants preferred the icicle plot and organization chart. They did not like the treemap. Twelve of the participants ranked the treemap last among the four plots. The rankings were analyzed using Friedman's chi-square test for repeated measures which showed that there was an overall difference among ranks ($p < 0.05$). Comparisons between the views, using Nemenyi's Critical Difference, found significant differences between the

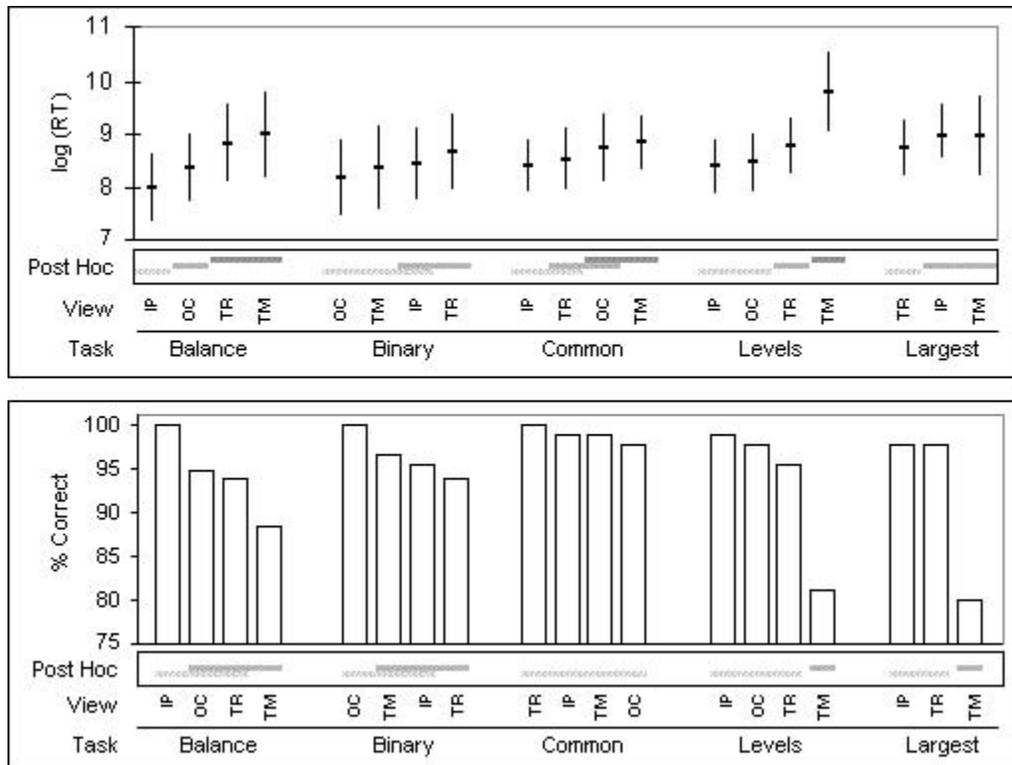


Figure 3. Mean response times \pm 1 std dev and percentage correct for Experiment 1. Response times include both correct and incorrect responses. NOTE: IP = Icicle Plot, OC = Organization Chart, TM = Treemap, TR = Tree ring

icicle plot and treemap, icicle plot and tree ring, and the organization chart and treemap.

5. Discussion

The organization chart was included in the study as our control condition in four of the tasks. The participants' performance with this view served as the standard against which performance with the other views was compared. The results from four of the tasks showed that the participants' performance with the icicle plot was equivalent to or better than their performance with the organization chart. We think that the icicle plot fared well in comparison to the organization chart because it retains the familiar top-to-bottom, left-to-right tree structure of the organization chart. In comparison, treemaps and tree rings require the user to learn how to interpret the structure.

The organization chart was not included in the three largest leaves task because it does not implicitly convey node size. The three remaining views use area, angle, or width to show the node size.

The participants' performance with the treemap was worse in the largest leaves task than with the other views. The problems encountered with the treemap are consistent with previous research. Turo and Johnson [6]

discuss the difficulty of comparing rectangular areas if both the vertical or horizontal dimension of the two rectangles are different. The difference in orientation makes comparisons slow and difficult. Cleveland [7] has said that the speed at which a graph is interpreted is closely associated with the accuracy of interpretation: fast interpretations tend to be accurate. The results from the treemap trials support him. In the three tasks in which the participants' accuracy with the treemap was significantly less than with the other views, their response times were significantly slower. A problem evident in Figure 2 also affected participant performance. Treemaps nest child nodes inside parent nodes with an offset between the edge of the parent and the edge of the child. This offset means that the area of a rectangle is not proportional to the size of the node. Rather, the area of a rectangle is proportional to the size of the node in relation to all of its ancestors. Nodes of the same size can be represented by different size rectangles if their depth in the hierarchy differs. The participants' less accurate performance in the largest leaves task is a result of this problem. Making the offset smaller can counteract this effect slightly. However, if the user has to interact with the treemap, the offset has to be large enough to support the interaction. In our case, the offset had to be large enough to make easy to select with a mouse. It is unlikely that the offset could be made smaller without compromising interaction.

Participants' performance with the tree ring was surprising — particularly when participants had to select the three largest leaves. Given the circular layout of the tree ring, the start and end angles of two leaves will be different. Gillan's [8] research suggests that comparing the leaf sizes involves mental translation and rotation of the leaves before making a judgment on relative size. Before comparing two angles a participant must mentally rotate one or both of the leaves so that they have the same orientation. Once the leaves have the same orientation, the participant translates one leaf onto the other to compare the size of the leaves. If participants use similar procedures to compare leaf sizes in the icicle plot, one would expect that the judgments would be quicker since all the leaves are oriented in the same direction. Since the orientation is the same the participant does not have to mentally rotate the leaf. One less step should result in faster response times.

We were also surprised at the results because we felt that the area of a leaf is its dominant perceptual feature but is indirectly related to leaf size. As shown in Figure 2, two leaves of equal size will have different areas if one of the leaves is higher in the tree than the other leaf. The leaf at the higher level will have more area and, to the user who is not familiar with the tree ring, appear to be bigger than the leaf at the lower level in the tree.

As Experiment 1 showed, participants' response time with the tree ring was faster (but not significantly different) than with the icicle plot. Neither the mental rotation and translation nor the potential confusion with leaf area appears to have affected the participants' ability to make quick and accurate comparisons. The most likely explanation of this result is that the tree ring is a space-filling algorithm in which all leaves touch the perimeter of the circle. The user only has to work around the perimeter of the tree ring to find and compare leaves. Finding the leaves in the icicle plot proved more difficult than in the tree ring. The uneven bottom edge of the icicle plot might have made finding the leaves harder than with the tree ring.

6. Experiment 2

The tasks in Experiment 1 tested the compact view's ability to communicate simple characteristics of the tree's topology. The tasks in Experiment 2 represented tasks involved in a more detailed analysis of a decision tree. Given the poor performance of the treemap in Experiment 1, Experiment 2 compared performance with the organization chart, icicle plot, and tree ring only.

6.1. Design

The experimental design was a 3 (view) x 2 (task) repeated-measures design with tree type as a random factor.

6.2. Participants

Fifteen coworkers participated in the experiment. Seven of the participants were male and eight were female. None of the participants participated in Experiment 1. None were familiar with the tree ring or icicle plot.

6.3. Tasks

There were two tasks: (1) node description task, and (2) node memory task. There were four versions of the node description task. The tasks were:

- Select the shallowest leaf in the tree.
- On level 3, select the ancestor of the deepest leaves in the tree.
- Select the leaf on level k where k was different for each tree.
- Select the deepest common ancestor of the leaves on the two deepest levels in the tree.

Each version of this task required the participant to read a verbal description and select a single node.

In the node memory task, participants were shown a highlighted node and told to memorize its location. At the beginning of each trial, the tree was drawn so that the target node appeared at the center of the window. The target node was highlighted. Once the participant believed he/she could remember the location of the target node, the participant clicked on the node. Then the tree was redrawn so that only the bottom edge of the tree was visible at the top of the window. The tree was shifted horizontally so that the horizontal center of the tree and the horizontal center of the window were aligned. The target node was no longer highlighted. The participant scrolled the window until he/she found the node that had been highlighted and clicked on it. The same target nodes were used for all three compact views. The target nodes were randomly selected with the restriction that half of the targets were terminal nodes and half were non-terminal nodes.

6.4. Procedure

The equipment and general procedure was the same as in Experiment 1. All trials were self-paced. Timing began when the tree appeared. The task instructions appeared in the window border. Timing ended when the participant pressed the spacebar.

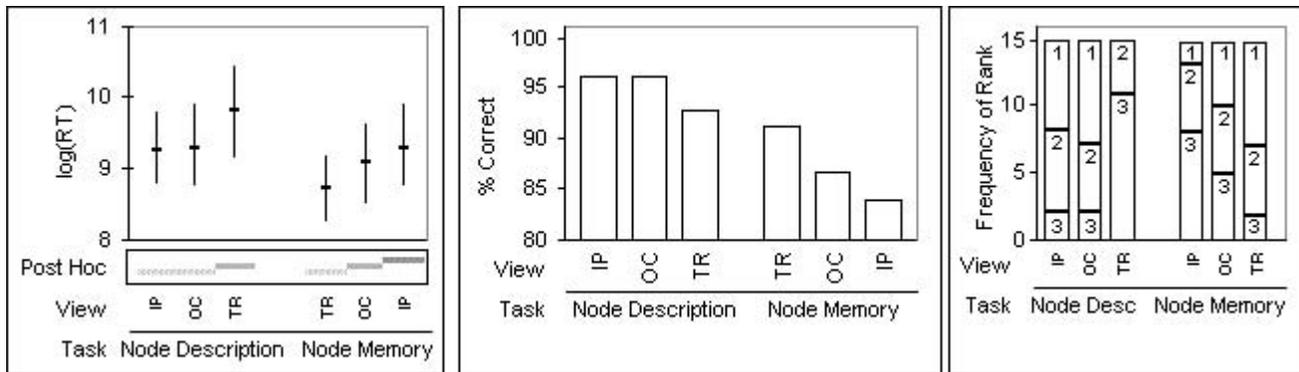


Figure 4. Mean response times +/- 1 std dev, percentage correct, and preference rankings for Experiment 2. Response times include both correct and incorrect responses. NOTE: IP = Icicle Plot, OC = Organization Chart, TR = Tree ring

As in Experiment 1, the overall size of the window containing the views was 25% of the screen. In the node description tasks the window height and width varied between views so that the entire tree was visible. In the node memory task, the window was square. This meant that the entire tree was not visible at any time.

We created four binary and five n-ary trees with 30 to 112 leaves for the node description tasks. We created five binary and five n-ary trees with 128 to 159 nodes and 65 to 113 leaves for the node memory task.

In the node description tasks, each participant saw a combination of the three views using three of the nine trees. They completed each task nine times, i.e., once for each compact view - tree combination, for a total of 36 trials. The order of tasks and views was randomized across participants. Each tree was used an equal number of times across all participants. The entire tree was visible in these tasks. The height and width of nodes and trees was determined as described in Experiment 1. After completing all the trials, participants ranked the compact views according to their preference for them.

In the node memory task, each participant saw every combination of trees and views for a total of 30 trials. Node size in the organization chart was increased to 10 pixels x 10 pixels. The icicle plot and tree ring were drawn so that the area of the smallest node in the tree was equal to the area of the nodes in the organization chart. After completing all the trials, participants ranked the compact views according to their preference for them.

6.5. Results

Response times (RT) were normalized using a log transformation before they were analyzed. RT was analyzed using a repeated measures analysis of variance with post-hoc (Tukey) comparisons of the mean RT.

Response accuracy (RA) for each task was analyzed using a logistic regression with generalized estimating equations approach to fitting the logistic regression

model. Comparisons were made between the percent correct for each of the views within each task. Figure 4 shows RT and percent correct for each view in each task. The results of the post-hoc comparisons appear below the horizontal axes. The mean response times for views with the same shade of gray are not significantly different ($p < 0.001$). The differences in percent correct between views was not significant in either task.

The results of the preference rankings also appear in Figure 4. The participants preferred those views with which they responded faster. The rankings were analyzed using Friedman's chi-square test for repeated measures. There was a significant difference in the rankings for the node description task but not for the node memory task ($p < 0.05$). Comparisons among the views, in the node description task, using Nemenyi's Critical Difference, showed significant differences between the icicle plot and tree ring, and the organization chart and tree ring.

7. Discussion

None of the views was clearly superior to the other views in Experiment 2. Participants' performance depended on a combination of the task and the view. They responded faster in the node description tasks when using the icicle plot and organization chart. In contrast, when performing the node memory task, they were faster when using the tree ring and slowest when using the icicle plot. The differences in response accuracy were not significant. We think that the superiority of the tree ring in the node memory task probably is due to the uniqueness of each node. Each node has a unique combination of start angle, end angle, outer radius, and inner radius. In comparison, the icicle plot nodes all have a uniform height. Only their width and level change. The constant node height and orientation that helped the icicle plot in the other tasks was a disadvantage in the node memory task.

8. Conclusions

Of the four compact views compared in the two experiments only one view distinguished itself from the others. The treemap was uniformly disliked by the participants and their performance while using it was worse than with the other three views. Their response times were slower in four tasks. Their response accuracy was lower in three tasks.

Of the remaining three views, identifying the best view is more difficult. The results of Experiment 1 suggest that the icicle plot is equivalent to or better than the organization chart and tree ring on most tasks. Participants also preferred the icicle plot and organization chart over the tree ring. Given the need to display node size as part of the compact view, one could conclude that the icicle plot is the best compact view for use in analyzing decision trees. The results of Experiment 2 are equivocal. Participants' performance on the node description tasks favor the icicle plot and organization chart. In contrast, their performance on the node memory task showed that the tree ring does a better job of supporting rapid recognition of node location.

Even though Chuah [4] found problems with a view similar to the icicle plot, we found that the icicle plot compared favorably to the other compact views in both experiments. One reason may be that it retains the left-to-right and top-to-bottom orientation of the organization chart. The similarity in orientation to the organization chart probably makes it easy to understand and to use. It is clear that the icicle plot could be enhanced to support memory for node location. We are currently evaluating modifications to the view to improve performance during sequential node examination.

The results show that a compact view can be created that is equal to or better than the most familiar form of the organization chart. The addition of node size could detract from the other views. The additional information is as likely to obscure the tree topology as it is to enhance understanding of the tree. However, the addition of node size does not seem to hurt the icicle plot or tree ring.

We have implemented a version of both the organization chart and the icicle plot in our software. Our experience with the views, as part of data mining software, shows that users prefer a choice of the organization chart or icicle plot. Their task determines which view works best for them.

9. References

- [1] E. Reingold and J. Tilford, Tidier Drawings of Trees. *IEEE Transactions on Software Engineering*, Washington DC, 1981, 223-228.
- [2] I. Herman, M. Delest, and G. Melancon, *Tree Visualization and Navigation Clues for Information Visualization*, National

Research Institute for Mathematics and Computer Science INS-R9806, Amsterdam, Netherlands, 1998.

- [3] K. Andrews and H. Heidegger. Information Slices: Visualizing and Exploring Large Hierarchies using Cascading, Semi-circular Discs, in *IEEE Information Visualization Symposium 1998 Late Breaking Hot Topics* ACM Press, Research Triangle Park NC, 1998, 9 - 12.

- [4] M. Chuah, Dynamic Aggregation with Circular Visual Designs, in *Proceedings of the 1998 IEEE Information Visualization Symposium 1998*, ACM Press, Research Triangle Park, NC, 1998, 35-43.

- [5] B.K. Kleiner and J.A. Hartigan, "Representing Points in Many Dimensions by Trees and Castles", *Journal of the American Statistical Association*, American Statistical Association, Alexandria, VA, June 1981, 260-272.

- [6] D. Turo and B. Johnson, *Improving the Visualization of Hierarchies with Treemaps: Design Issues and Experimentation*. University of Maryland Human-Computer Interaction Laboratory Technical Report 92-06, College Park, MD 1992.

- [7] W. Cleveland, *Elements of Graphing Data*, Kluwer Academic Publishing, New York, NY, 1994.

- [8] D. Gillan, A Componential Model of Human Interaction with Graphs V: Using Pie Graphs to make Comparisons, in *Proceedings of the IEA 2000/HFES 2000 Congress*, HFES, San Diego, CA, 2000, 439-442.