

DEPTH Vs BREATH IN THE ARRANGEMENT OF WEB LINKS

Panayiotis G. Zaphiris
Institute of Gerontology &
Industrial and Manufacturing Department
Wayne State University
Detroit, Michigan

The purpose of this study was to examine the effect of depth and breadth in the arrangement of web link hierarchies on the user preference, response time and errors. An experiment is reported, where previous menu depth/breath tradeoff procedures are applied to the web link domain. The variables evaluated were five different web page linking strategies with varying depth and breadth. The results indicate that task completion time increases as the depth of the web site structure increases.

INTRODUCTION

Recent years have witnessed an explosive increase in public interest in information access and communication technologies. This exponential use, coupled with the rapid growth in volume of electronically available information sources, such as those available on the Internet, has further exacerbated the need for effective and efficient tools for information designers and users. As Piroli and Card (1995) state, for researchers and developers in human-computer interaction, this increases the need for models and analysis techniques that allow us to determine the value added by particular information access, manipulation, and presentation techniques, and to reveal design elements that may yield further enhancements. One such element is the information architecture methodologies used for presenting information on the world wide web (web).

This paper describes the results of an experiment where procedures in depth/breath tradeoffs, previously applied to menu design, are applied to web link hierarchies.

The topic of menu selection and especially depth versus breath tradeoffs has been extensively examined. Menu panels usually consist of a list of options. The options may consist of words or icons. The word or icon, conveys some information about the consequences of selecting that option. Sometimes the options are elaborated upon with verbal descriptors. When one of the options is selected and executed a system action occurs that usually results in a visual change on the system. The total set of options is usually distributed over many different menu panels.

Web indexes are organized in a similar structure. Links (very often 2-3 words, sometimes elaborated with verbal descriptors) are arranged in various levels of homepages. These links convey information about the page (with information or further sub-categories) that will be displayed if that specific link is selected.

Empirical Studies

Depth versus breadth in hierarchical menu structures has been the topic of much empirical research. The trade-off between menu depth (levels in a hierarchical menu structure) and breadth (number of menu selections present on a given

menu in the hierarchy) is by some researchers the most important aspects that must be considered in the design of hierarchical menu systems (Jacko, Salvendy, & Koubek, 1995).

Miller (1981) found that short-term memory is a limitation of the increased depth of the hierarchy. His experiment examined four structures (64^1 , 2^6 , 4^3 , and 8^2) with a fixed number of target items (64). The 64 items were carefully chosen that “they form valid semantic hierarchies” in each of the 4 menu structures. As depth increased so did response time to select the desired item.

Snowberry, Parkinson & Sisson (1983) replicated Miller’s study by examining the same structures but this time including an initial screening session during which subjects took memory span and visual scanning tests. They found that instead of memory span, visual scanning was predictive of performance, especially in the deepest hierarchies. Their experiment compared performances on both a randomized and a categorized 64 item display to performances on structures of increasing depth and decreasing breadth. They used a between-subject design with four different experimental conditions (menu structures - 64^1 , 2^6 , 4^3 and 8^2). Their results provide a nice replication of Miller’s (1981) results. With category organization held constant across conditions, search time improved as depth decreased and breath increased. Accuracy was found to improve as breadth increased (and depth decreased) regardless of display format in the broadest display

Kiger (1984) extended Miller’s research by conducting an experiment that provided users with five modes of varying menu designs of 64 end nodes.

8^2 : 8 items on each of 2 levels.

4^3 : 4 items on each of 3 levels.

2^6 : 2 items on each of 6 levels.

$4^1 + 16^1$: A 4 item menu followed by a 16 item menu

$16^1 + 4^1$: A 16 item menu followed by a 4 item menu

Performance and preference data were collected. The results of the experiment showed that the time and number of errors increased with the depth of the menu structure. The 4x16 structure had the fastest response times and the fewest

errors. The participants ranked the menus with least depth as the most favorable (The 8^2 structure was favored).

Wallace et al. (1987) confirmed that broader, shallower trees (4x3 versus 2x6) produced superior performance, and showed that, when users were stressed, they made 96 percent more errors and took 16 percent longer. The stressor was simply an instruction to work quickly ("It is imperative that you finish the task just as quickly as possible"). The control group received mild instructions to avoid rushing ("Take your time; there is no rush").

Norman and Chin (1988) fixed the number of levels at four, with 256 target items, and varied the shape of the tree structures. They recommend greater breadth at the root and at the leaves, and added a further encouragement to minimize the total number of menu frames needed so as to increase familiarity.

An experiment by Jacko and Salvendy (1996) tested six structures (2^2 , 2^3 , 2^6 , 8^2 , 8^3 , and 8^6) for reaction time, error rates, and subjective preference. They demonstrated that as depth of a computerized, hierarchical menu increased, perceived complexity of the menu increased significantly.

EXPERIMENT

Hypothesis

This experiment tests the following hypothesis: The time, number of errors and complexity of retrieving information from the internet is directly proportional to the depth of the tree structure of the arrangement of the links of the web page. The three variables, that is, time, errors and complexity will decrease as the breadth of the tree structure increases and depth decrease.

Participants

A total of 17 people (6 females and 11 males) participated in this experiment. The participants' age ranged from 19 to 45 (mean=28.3, SD=6.0). Most of them were University of Maryland students. All subjects had good knowledge of the Internet (have been using computers for at least two years and are using the internet on at least a weekly basis) and all had used some kind of a web browser before our experiment.

Materials

Link Hierarchies. The web link hierarchies were made up of 64 end nodes with general information about the island of Cyprus (for example in the case of breadth 4 hierarchy the top level categories were Geography, Economy, People and History). The data was pulled out from the Almanac of Cyprus (1995).

Five different link hierarchies were designed using a standard web editor and hosted on a local web server. Each web page consisted of a set of links and a small logo at the top of the page.

The five models used are:

- Model 1: 2x6, breadth = 2, depth = 6
- Model 2: 4x3, breadth = 4, depth = 3

Model 3: 8x2, breadth = 8, depth = 2

Model 4: 4x1 16x1, first page breadth = 4, second page breadth = 16

Model 5: 16x1 4x1, first page breadth = 16, second page breadth = 4.

Following Kiger's (1984) methodology, initially the database was constructed using the 2x6 tree structure. Each of the other four structures was based on the organization of information in the 2x6 structure. For example, the 4x3 database was constructed by reducing the 2x6 tree to three levels and four choices at each level. This involved using the 2nd, 4th, and 6th levels of the 2x6 tree, with some elaboration of the link selections to overcome the information loss resulting from reducing the tree structure from six to three levels in depth (Kiger, 1984).

Figure 1 shows a screen shot of the top level page of one of the link hierarchies

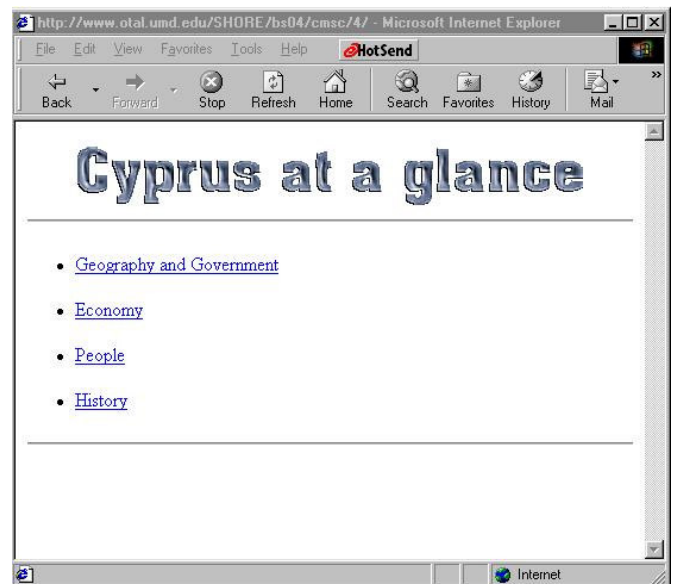


Figure 1: Screen shot of a link hierarchy used in this experiment.

Procedure

Each participant was first presented with general instructions about the experiment. This included a brief purpose of the experiment and a set of guidelines (Participants had to search for the information on the web site even if they knew the answer to the question, they had to use the back button on the browser and not the right mouse button or any other shortcuts to go back a level, they shouldn't browse for any other information other than the one they had been asked to find).

Participants were then given time to ask questions and finally after signing a consent form was asked to perform a total of ten search tasks (all of equal difficulty). He/She had to perform two tasks on each of the hierarchies presented to them. The following are two examples of the kind of questions participants were asked :

- What is the total population of Cyprus?
- What is the name of the capital city of Cyprus?

A random ordering of the five tree structures was determined prior to the experiment for each participant.

The total time to perform each task and the number of errors in completing each task were recorded through the web server logs.

After completing his/her last search task, the participant was asked to rank the five hierarchies in level of difficulty of use.

RESULTS

Speed and Accuracy of Information retrieval

The average total reaction time (RT) for each model was calculated. Total reaction time refers to the total time it takes for a participant to locate the specific information. Table 1 shows a summary of the results, and graph 1 plots these data.

The analysis of variance (ANOVA) for total reaction time showed a significant effect ($F(4,16)=3.20, p<0.05$)

The accuracy of information retrieval was analyzed using the total number of errors (total number of errors was calculated as the number of clicks in excess of the optimum path) for each hierarchy. The analysis of variance for mean number of errors showed no significant effect ($F(4,16)=1.29, p>0.05$).

Model	2x6	4x3	8x2	4x1 16x1	16x1 4x1
RT	36.0 (14.9)	32.2 (27.9)	20.3 (11.7)	21.0 (11.3)	32.4 (19.8)

Table 1: Mean total reaction time in seconds (standard deviation in parenthesis)

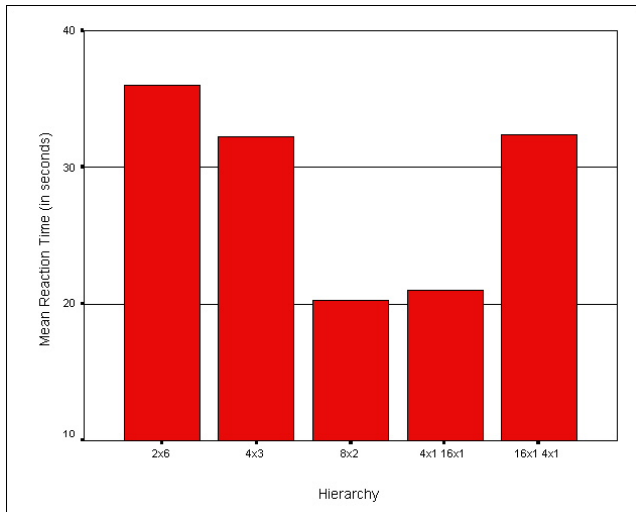


Figure 1: Average reaction time for each hierarchy

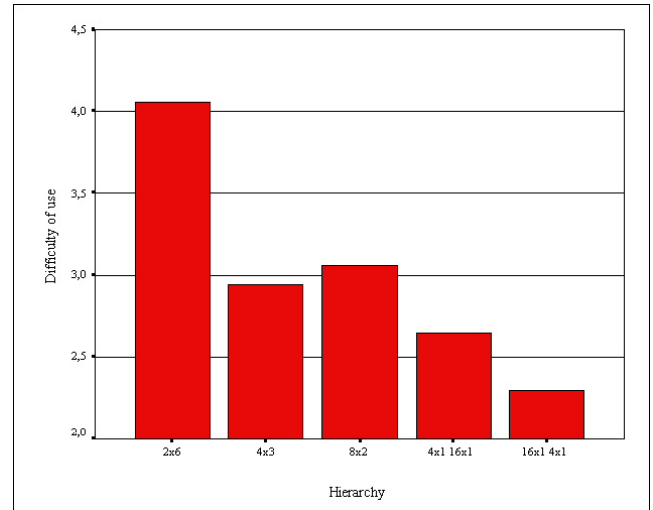
Difficulty of Use

The mean difficulty of use rankings are presented in Table 2 and plotted in Graph 2.

The ANOVA for the Ranking was significant ($F(4,16)=3.40, p<0.05$)

Model	2x6	4x3	8x2	4x1 16x1	16x1 4x1
Difficulty	4.1 (1.1)	2.9 (1.3)	3.1 (1.3)	2.7 (1.5)	2.3 (1.4)

Table 2: Difficulty of use rankings (SD in parenthesis)



Graph 2: Difficulty of use for each hierarchy

DISCUSSION

From Table 1 and 2 it can be seen that a general trend exists:

1. Hierarchy 2x6 (breadth=2, depth=6) is both considered by users as the most difficult to use and with the highest mean task time.
2. Although 8x2 (breadth=8, depth=2) has the lowest mean task time it is ranked 3rd in user preference.
3. Users seem to prefer a mixture of depth and breadth (16x1 4x1 is ranked first and 4x1 16x1 is ranked second in user preference).
4. Increase in breadth (in the range (2-16) investigated) does not seem to effect user speed in information retrieval.

From points 1 and 2 above our original hypothesis that access time is proportional to the depth of the tree structure of the arrangement of the links of the web page is justified. Point 4 above is supported by observing the fact that model 5 (16 links at level 1) has the 2nd biggest task time and model 4 (16 links at level 2) has the second smallest task time. This result shows that the number of links per page is not the driving factor when it comes to access speed. Instead depth seems to be driving the speed of information retrieval.

There are some anomalies while comparing Table 1 with Table 2.

- Model 5 is ranked 1st in preference, but from Table 1 we can see that is the 2nd most slow system.
- Model 3 is ranked 2nd less preferred but is the fastest of all.

Both of these anomalies are difficult to be explained scientifically by this experiment. Both Model 3 and Model 5

are of two step depth but they seem to follow extreme opposite trends on the statistics. What seems to be an "anomaly" is the preference ranking given to these designs, which proves once again the diverse human nature. Our users tended to prefer heterogeneous designs (4x1 16x1 and 16x1 4x1) to homogeneous designs.

Our participants, made positive comments about their experience in testing this model. Comments ranged from "I learned more about Cyprus" to very precise suggestions (by one of our subjects) on how this test should be run and what should be improved.

This experiment couldn't find a significant effect of errors so the initial hypothesis that errors in information retrieval increase as the depth of the hierarchy increase could not be proven. This is mainly due to the fact that there were very few errors performed by our participants in performing their search tasks.

Overall the results are in agreement with those of Kiger (1984) where it has been proven that access time is proportional to depth in menu selection.

5. Conclusion

Impact for practitioners

Through this experiments it was achieved to prove the initial hypothesis that mean task time in accessing specific information on the WEB is proportional to the depth of the tree structure. The author believes that practitioners (e.g. WEB designers) should keep in mind these observations while constructing homepages on the internet. Low depth designs of a taste of variety (e.g. Model 4 and 5) seem to be preferred by users.

Suggestions for future researchers

Possible expansions of this projects might include:

1. Investigating the impact of the number of links per page combined with depth and breadth variations.
2. Investigating the impact of the complexity of questions on both access time and errors. Although in this experiment care has been taken to pick questions (tasks) that were of relatively the same complexity it will be interesting to observe what happens as the complexity of the tasks is varied.
3. An expansion of this test might be to run it both on site (as done) and on-line with remote subjects. This would get a bigger and more diverse group of users.

Also, although a number of theoretical/analytical studies (Lee & MacGregor, 1985; Paap & Roske-Hofstrand, 1986; Fisher et al, 1990) have tried to model menu selection, there has still been no attempt to apply those models to online information retrieval (as applied to hierarchical information structure) especially to take into account user expertise (expert Vs novice), age and other individual differences.

The author would like to express his thanks to Ms Lianaeli Mtei and Prof. Ben Shneiderman whose help in all stages of this study has been of great value. Data for this experiment was collected by the author and Ms Lianaeli Mtei as part of a Human-Computer Interaction class group project at the University of Maryland, College Park. Portions of this paper were first drafted as part of that class project.

REFERENCES

- Cyprus Press and Information Office (1995). The Almanac of Cyprus. Nicosia, Cyprus: Press and Information Office.
- Jacko, J., & Salvendy, G. (1996). Hierarchical menu design: Breadth, depth, and task complexity. Perceptual and Motor Skills, 82, 1187-1201.
- Jacko, J., Salvendy, G., and Koubek, R., "Modeling of Menu Design in Computerized Work," Interacting with Computers, 7, 3, pp. 304-330
- Kiger, J. I. (1984). The depth/breadth tradeoff in the design of menu-driven interfaces. International Journal of Man-Machine Studies, 20, 201-213.
- Lee, E. & MacGregor, J. (1985). Minimizing user search time in menu retrieval systems. Human Factors, 27, 157-162.
- Miller, D.P. (1981). The depth/breadth tradeoff in hierarchical computer menus. Proceedings of the Human Factors Society--25th Annual Meeting (pp. 296-300). Santa Monica: Human Factors Society.
- Norman, K. & Chin, J. (1988). The effect of tree structures on search in a hierarchical menu selection system. Behaviour and Information Technology 7, 51-65.
- Paap, K.R., & Roske-Hofstrand, R.J. (1986). The optimal number of menu options per panel. Human Factors, 28, 377-385.
- Pirolli, P., Card, S., (1995). Information Foraging in Information Access Environments. Proceedings of CHI 95 Human Factors in Computing Systems, Denver, CO, April, 1998, ACM Press, pp. 51-58
- Snowberry, K., Parkinson, S., & Sisson, N. (1983). Computer display menus. Ergonomics, 26, 699-712.
- Wallace, D., Anderson, N. & Shneiderman, B. (1987). Time stress effects on two menu selection systems. Proceedings of Human Factors Society, Thirty-First Annual Meeting, 727-731

ACKNOWLEDGEMENTS