

On the Vague Modelling of Web Page Characteristics Regarding Usability

Elena García Barriocanal¹, Miguel A. Sicilia Urbán², and J. Antonio Gutiérrez¹

¹ Computer Science Department. Polytechnic School.
University of Alcalá. Ctra. Barcelona km. 33.6
28871. Alcalá de Henares, Madrid. Spain
{elena.garciab, jantonio.gutierrez}@uah.es

² Computer Science Department. Polytechnic School.
Carlos III University. Avd. de la Universidad, 30.
28911. Leganés, Madrid. Spain
msicilia@inf.uc3m.es

Abstract. Vagueness is an inherent property of man-machine systems associated with some perceptual and cognitive characteristics of human information processing, as pointed out by Karwowski and other researchers. More concretely, some perceivable characteristics of interface designs, including sizes and quantities, are usually perceived by humans as vague categories that result in imprecise guidelines for interface usability. In this paper, we describe how such categories — in the specific case of the Web — can be modelled as fuzzy sets by using conventional membership function elicitation procedures, using Web page length and number of links as case studies. The resulting fuzzy sets can then be used for automated usability analysis processes using fuzzy rules to formalize vague guidelines.

Keywords. Fuzzy sets, automated usability analysis, membership function elicitation.

1 Introduction

Usability is a multifaceted concept that encompasses several attributes of the interaction of humans with software systems, in many cases summarized in three interdependent aspects named efficiency, effectiveness and satisfaction [8]. In consequence, human cognitive and perceptual abilities are relevant to interface design [24]. As described in [14], fuzziness is an inherent characteristic of human-machine systems and can't be overlooked in human information processing models and applications. Fuzzy set theory and related imperfect information handling mathematical frameworks [13] have already been applied to the field of human computer interaction [14]. Concrete applications include the fuzzy modelling of GOMS selection rules [15], the use of fuzzy set to model usability scores in usability evaluation [5], the extraction of personality indicators from textual contents

and images [23] and the use of fuzzy aggregation operators in the aggregation of heuristic usability evaluation questionnaires [9].

In the specific context of usability evaluation [6], guidelines and patterns often refer to the *structure* of the design of the software system, i.e. to the perceivable features of the interface. Examples — for the specific case of Web systems [21]— are *sizes* (of sections, headers, navigation bars and the like), *quantities* (of links, images) and *navigational patterns* (depth of hierarchies, frames). Many of these kinds of guidelines, when formulated by humans, are vague and intended to be approximate orientations. For example, in [3], it is said that “headers should not take more than 25% of a letter size page” or “pages should not be overcrowded with links”. The first guideline is clearly an arbitrary selection of a limit for a blurred size-relation, and the second one makes reference to a vague amount of links per page. The question that follows from the just described observations is: how can we make machines work with these vague categories?. The solution for the representation of many of these guidelines is that of mathematically characterizing the concepts that underlie them, and fuzzy set theory is a candidate for that purpose.

In this paper, an approach based on fuzzy sets for *automated usability analysis* (according to the definition in [12]) of Web sites is illustrated by focusing on two quantitative attributes of Web pages: page length and number of links, which are known to affect design quality and usability [11]. Concretely, it is described how existing membership function elicitation techniques [16] can be applied to the fuzzy characterization of the attributes, and how the resulting fuzzy sets, combined with fuzzy rules, can be used to develop automated usability analyzers.

The rest of this paper is structured as follows. In Section 2, the experimental design, procedure and results of the membership function elicitation is described. Section 3 briefly describes the design of a Web page analyzer based on vague page attribute characterizations. Finally, conclusions and future research directions are provided in Section 4.

2 Membership Function Elicitation

A variety of methods for the elicitation of membership functions have been described in the literature [16]. The focus of our study are two highly specific psychological categories with blurred edges, that must be characterized by data analysis on subjective perceptions. The selection of the elicitation technique was driven by a first phase of examination of the problem, which allowed for the establishment of the objectives and assumptions about the interpretation of membership values:

- The expected result was the obtention of a number of fuzzy sets representing granulated page lengths (short pages, medium pages, long pages and the like) and granulated quantities of links per page (few # of links, excessive # of links and the like).

- The presence of interpersonal disagreements in assessing what is a long page and what is a link overcrowded one led us to assume that fuzziness in these constructs is *subjective* and determined in part by device limitations.

As a consequence, we have adopted a *similarity* interpretation of membership (opposed to other views like *likelihood* or *random set* as described in [2]). In this view, the membership value $\mu_A(x)$ represents “the distance of x from a prototypical element y which fully belonging to A ($\mu_A(y) = 1$)”. Thus, users rate the belonging of a page into a concrete set by estimating a “distance” from an ideal, imagined prototypical instance which would be classified as crisply belonging to the given set. Although we have not selected concrete prototypes for the attributes being measured, they are easily determined by Web users (e.g. specific clearly “long” and “link overcrowded” pages are easily determined by them).

From the available techniques, we have selected two that are coherent with the interpretation adopted:

- First, the membership functions are directly determined by *membership exemplification* [2].
- And second, a variant of the *polling* technique [10] was used to validate the results obtained by membership function exemplification.

The details of the building of the case, the measurement procedure and its results are described in the rest of this section. It should be noted that the aim of the study was not that of achieving statistical significance, but to demonstrate the feasibility of an empirical approach.

2.1 Experiment Design

The experiment required a first step for the obtention of an appropriate Web page database, and after that, the procedure for the obtention of vague measures was defined, as described in this sub-section.

Obtention of Page Sample Database As a matter of fact, the majority of common Web users today use Microsoft Internet Explorer (IE) to navigate the Web³, and common screen resolutions are over 800x600 pixels. Due to these statistical measures, we have limited our study to the latest stable version of IE on that resolution⁴.

Two Web page URL sets were built, containing 50 pages each. They were obtained automatically from Google⁵ by a program in the Java language, and

³ According to OneStat.com metrics published December 2002, its share was about 95% of the market share

⁴ Nonetheless, the experiments could be extended in the future to other kinds of configurations according to more detailed browser and screen data.

⁵ One of the most used search engines, which provides programmable interfaces: [<http://www.google.com/apis/>]

stored in a Microsoft `Access` database. Pages were selected randomly from the most relevant results obtained from taking a random sample of query strings from the most used in the year 2002⁶. An approximately distribution of number of links an page length was obtained by manually filtering the larger set of pages obtained from `Google`.

After that, the size in pixels and number of links of the pages were obtained. The number of links was estimated by simply counting `href` tags⁷. Page length required a more elaborate procedure, since the actual length is dependant of the rendering of the page in the browser. Due to this fact, we developed a Windows program using the `WebBrowser` control in `VisualBasic.NET` — contained in the `SHDOCVIEW.DLL` library— which provides the same rendering engine than the one used directly by IE. The program load the collected locators, and measure its size on a default IE configuration in pixels. Technically, this is accomplished by using the `IHTMLDOMElement2` interface in the `MSHTML.DLL` library to address the body of the document and asking for the property `scrollHeight`.

Procedure The experiment take place on a Internet connected computer with the configuration mentioned above. Participants were given two list of clickable URLs to visit the corresponding pages. The first one was sorted in ascending order by length, and the second was sorted by number of links in the page. They were asked to put a grade in the [0..100] interval (the same scale used by Zysno in [25]) to the compatibility of each page with the linguistic labels in the label sets:

$$LS_{page-length} = \{very - short, short, medium, long, very - long\} \quad (1)$$

$$LS_{\#-links} = \{very - few, few, acceptable, excessive, crowded\} \quad (2)$$

Both label sets are considered to be ordered and to cover the whole semantic domain of the concept being measured. The usability-related interpretation of the terms was formulated by questions of the form “do you think this page is X ?” and “do you think that there are X links in this page?”. The number of labels suffices the purposes of the study, since guidelines and texts on usability rarely use linguistic hedges related to the concepts under study.

The evaluator observed them in the process of membership grade exemplification, taking notes on significant behavior. In our experimental design we have also taken into account Rosch observation about the speed of processing [22], according to which humans categorize faster the objects that are more prototypical, i.e. the objects perceived more clearly as typical cases of the category. Concretely, pairs (page, fuzzy set) that were found difficult to exemplify by participants were recorded.

⁶ Available at [<http://www.google.com/press/zeitgeist2002.html>]

⁷ Although this method is not completely accurate, due to “hidden” dynamically estimated links, it suffices for the purposes of this study.

2.2 Results and Discussion

The procedure was carried out with ten adult subjects. Only experienced Web users were included. The mean of the membership grade examples obtained was used to build discrete fuzzy sets with the pages ranked as domain. From them, a function for the final, continuous fuzzy sets were estimated by nonlinear regression.

Figure 1 shows the (non-normalized) resulting membership functions that describe page length.

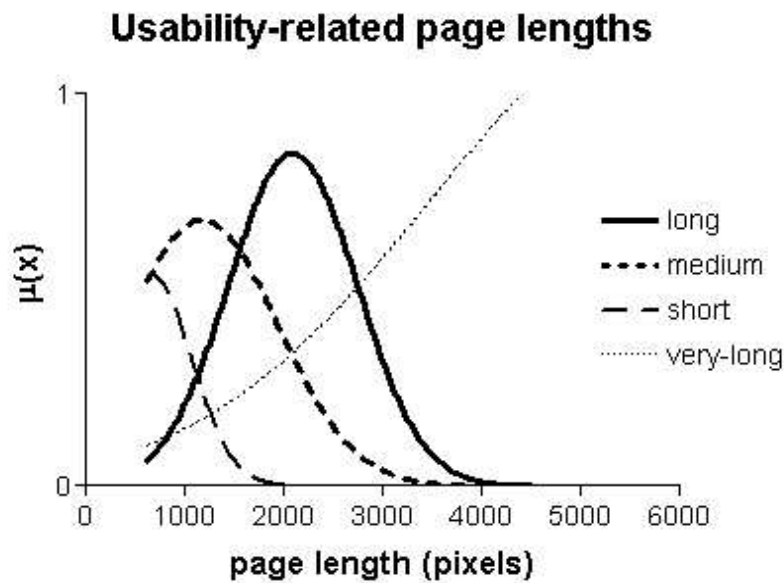


Fig. 1. Non-normalizing membership functions to describe page length.

The “very-short” label was fitted to a polynomial function. It has been omitted from the analysis since it resulted in a very sharp function around the minimum page lengths that can be found, of about 600 pixels, which makes it seldom useful for usability analysis. The rest of the labels were fitted using a Gaussian model. It should be noted that pages begin to be considered as rather “long” for values above two thousands pixels length, which in practice corresponds to more than three 800x600 areas. Since most pages used for the study were home pages, this result is mainly consistent with existing approximate guidelines like [17]. For example, if we take the following guideline: “short pages, those containing *one or two* screens of text, work well for the home page and menu pages when

users are scanning for link choices”⁸. Then, our observations would extend the number of “tolerable” screens at least to three. Another interesting fact is the large degree of overlap between long and very long pages, reflecting diverse user’s preferences.

Figure 2 shows the non-normalized resulting membership functions that describe quantities of links. The “very-few” label has been omitted since its number of non-zero points was non significant.

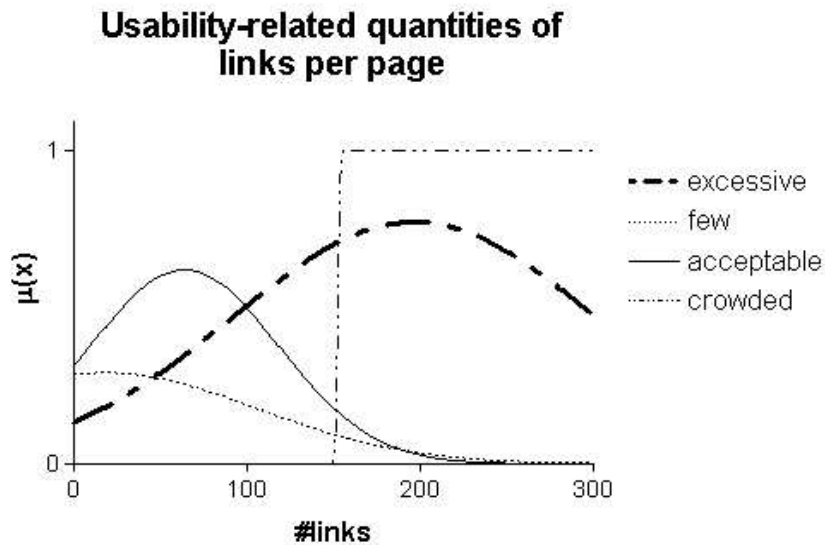


Fig. 2. Non-normalizing membership functions to describe quantity of links.

A sigmoid function was used to model the “crowded” concept, showing a sharp boundary around the one hundred and a half amount of links. The observations of the experiments led us to conclude that this function is not appropriate to model the concept due to two main reasons: users took longer to make their assessments about this category, and many comments about its dependency on page-length were raised. Future studies should take the links–page length ratio as the underlying measure instead.

3 Use of the Vague Concepts in Automated Usability Analysis

Once the fuzzy sets that represent the measures are determined, they can be used to develop software that automatically diagnoses usability problems associated

⁸ [http://usability.gov/guidelines/page_length.html]

with them. To do so, rules encoding specific uses of usability *heuristics* [19] are combined with the fuzzy sets describing the vague measures. We have used the *Fuzzy Java Toolkit*⁹ to implement some of them in a simple usability analysis prototype.

For example, a (hypothetical) rule about page length can be stated as follows:

```
if page is long and page is homepage then
    reading-efficiency is low

if page is link-crowded then
    reading-efficiency is very low
```

These kinds of rules can be subsequently used to fire other rules that relate the concrete measures with more abstract and generic usability attributes [24]. For example, reading-efficiency contributes to the general efficiency attribute, which can be used to report or suggest corrective actions.

```
if reading-efficiency is low then
    efficiency is low
if efficiency is low then
    <<take the appropriate corrective or reporting action>>
```

Of course, the knowledge encoded in the formulation of the rules must come from usability experts or other empirical sources like [11].

4 Conclusions and Future Work

Existing membership elicitation techniques can be used for the measurement of linguistic values for Web page characteristics that are known to have influence in the usability of interfaces. As a case study, we have carried out an experiment to determine a number of fuzzy sets regarding Web page length and amount of links per page. It has been also described how these sets can be combined with fuzzy rules to come up with automated usability analysis tools.

Future studies should address a more rich repertory of usability-related quantifiable Web site attributes [11], and also more ambitious elicitation studies for membership functions. In addition, *diachronic* studies are worth being carried out, since some evidence has been found about the evolution of the user's attitudes, like willingness to scroll [20], which affects page length guidelines.

The experimental design followed provides a overall estimation of the studied parameters, but further stratification of the population would be desirable due to the highly heterogeneous nature of Web users. Particularly, age (due to the different needs of elderly or disabled people, see, for example [7, 18]) and navigation expertise (see, for example [1]) are good candidates for future studies. In addition, the *reverse rating* elicitation method could be used as an alternative

⁹ [http://www.iit.nrc.ca/IR_public/fuzzy/fuzzyJToolkit.html]

method or for the sake of verifying results obtained with other techniques [4], if only experienced Web users are considered. This method consists essentially in asking subjects to identify known Web pages that have a specific degree of membership in the concept being determined, so that expert users would be able to remember prototypical cases easily. In addition, the type of the pages should also be considered, since guidelines differ depending on it, e.g. home pages are “allowed” to be longer than normal reading pages [17], and the interaction between attributes should also be considered, for example, the #-links-page length ratio.

References

1. Benyon, D.: Adaptive systems: a solution to usability problems. *User Modeling and User-Adapted Interaction* 3 (1993):65-87
2. Bilgiç, T. and Türksen, T.: Measurement of Membership Functions: Theoretical and Empirical Work. In: D. Dubois and H. Prade (eds.) *Handbook of Fuzzy Sets and Systems Vol. 1, Chapter 3, Fundamentals of Fuzzy Sets*, Kluwer (1999):195-232
3. Borges, J.A., Morales, I. and Rodriguez, N.J.: Guidelines for designing usable World Wide Web pages. In: *Proceedings of the ACM Conference on Human factors in computing systems* (1996):277-278
4. Chameau, J.L. and Santamarina, J.C.: Membership Functions I: Comparing Methods of Measurement, *International Journal of Approximate Reasoning* 1 (1987):287-301
5. Chang, E., Dillon, T.S., Cook, D.: Measurement of Usability of Software Using a Fuzzy Systems Approach. In: *Proceedings of the 8th International Conference on Software Engineering and Knowledge Engineering, SEKE '96* (1996): 69-76
6. Dumas, J.S. and Redish, J.C.: *A practical guide to usability testing*. Intellect (2000)
7. Ellis, R. D. and Kurniawan, S. H.: Increasing the usability of Online information for older users: A case study in participatory design. *International Journal of Human-Computer Interaction*, 12(2) (2000):263-276
8. Frøkjær, E., Hertzum, M. and Hornbæk, K.: Measuring Usability: are effectiveness, efficiency and satisfaction really correlated?. In: *Proc. of Human Factors in Computing Systems* (2000):345-352
9. García, E., Sicilia, M.A., Hilerá, J.R., Gutierrez, J.A.: Computer-Aided Usability Evaluation: A Questionnaire Case Study. In: *Advances in Human Computer Interaction*. Typorama (2001) 85-91
10. Hersh, H.M. and Caramazza, A.: A fuzzy set approach to modifiers and vagueness in natural language. *Journal Experimental Psychology*, 105(3) (1976):254-276
11. Ivory, M.Y. and Hearst, M.A.: Improving Web Site Design. *IEEE Internet Computing*, Special Issue on Usability and the World Wide Web 6(2), March/April (2002): 56-63
12. Ivory, M.Y. and Hearst, M.A.: The State of the Art in Automated Usability Evaluation of User Interfaces. *ACM Computing Surveys*, 33(4) (2001):1-47
13. Klir, G.J. and Wiermann, M.J.: *Uncertainty-Based Information. Elements of Generalized Information Theory*. Physica-Verlag, Heidelberg (1998)
14. Karwowski, W. and Mital, A. (editors): *Applications of Fuzzy Set Theory in Human Factors*, Elsevier Science Publishers, Amsterdam (1986)

15. Karwowski, W., Kosiba, E., Benabdallah, S. and Salvendy, G.: A framework for development of fuzzy GOMS model for human-computer interaction. *International Journal of Human-Computer Interaction* 2(4) (1990):287–305
16. Krishnapuram, R.: Membership Function Elicitation. In: E. Ruspini, P. Bonissone and W. Pedrycz (editors): *Handbook of Fuzzy Computation*, Institute of Physics Publishing (1998):B3.2:1–11
17. Lynch, P.J. and Horton, S.: *Web Style Guide: Basic Design Principles for creating Web Sites*, Yale University Press (1999)
18. Meyer, B., Sit, R.A., Spaulding, V.A., Mead, S.E., Walker, N.: Age Group Differences in World Wide Web Navigation. In: *Proceedings of the ACM Conference on Human Factors in Computing Systems* (1997)
19. Nielsen, J.: Heuristic evaluation. In Nielsen, J., and Mack, R.L. (eds.): *Usability Inspection Methods*, John Wiley and Sons, New York (1994):25–61
20. Nielsen, J.: Changes in Web Usability Since 1994. Alertbox for December 1 (1997), available at <http://www.useit.com/alertbox/9712a.html>;
21. Nielsen, J.: *Designing Web Usability: The Practice of Simplicity*. New Riders (1999)
22. Rosch, E.: Principles of categorization. In Rosch, E. and Lloyd, B. (Eds.) *Cognition and Categorization*. Hillsdale, NJ, Lawrence Erlbaum Associates (1978)
23. Sanchez, E. and Santini, J.A.: Fuzzy Logic and the Internet: a Fuzzy Web Site Evaluator. In: *Proceedings of the North American Fuzzy Information Processing Society (NAFIPS'02)* (2002):344–347
24. Van Welie, M., van der Veer, G.C. and Elins, A. (1999), Breaking down usability. In: *Proc. of Interact'99* (1999):613–620
25. Zysno, P.: Modeling membership functions. In Rieger (ed.): *Empirical Semantics I*, Vol. 1 of *Quantitative Semantics*, vol. 12, Studienverlag Brockmeyer, Bochum (1981): 350-375