In Proceedings of the Sixth International Conference on Cognitive Modeling, 378-379.
Mahwah, NJ: Lawrence Earlbaum.

Navigating within a web site: the WebStep model

Sara Rigutti (rigutti@psico.units.it)
Walter Gerbino (gerbino@units.it)

Department of Psychology, University of Trieste
Via Sant' Anastasio 12, 34134 Trieste, Italy

Introduction

Web navigation is a complex activity in which users interact with a system to achieve their goals. Current web-navigation models (Kitajima, Blackmon, & Polson, 2000; Miller & Remington, 2000; Pirolli & Fu, 2003) are based on the label-following strategy, which takes the semantic aspects of navigation elements as fundamental. We propose that goal-directed actions of users are supported by *semantic* and *pragmatic* cues embodied in all navigation elements. Semantic and pragmatic cues are related to context-independent and context-dependent aspects of element meaning, respectively. The intrinsic contents of icons and verbal labels provide semantic cues. Spatial features (position and distribution in the page), as well as structural features of verbal labels (shape and color), provide pragmatic cues.

Here, we focus on the pragmatic cues that differentiate embedded links from navigation-bar options. As regards spatial features, embedded links are placed in the central area of the page while the navigation bar is placed at the top. As regards structural features, embedded links are intermixed with the embedding text and can be perceptually segregated because of shape dissimilarity (underline vs. normal) and color dissimilarity (blue vs. black); while the navigation bar includes only homologous elements, grouped by shape and color similarity, good continuation of button contours, and proximity.

WebStep Model

We developed a new model of web navigation, called WebStep. The model evaluates spatial and structural aspects of embedded links and navigation-bar options to generate predictions about the probability and speed of element selection. WebStep refers to paradigmatic navigation cases in which selection cannot be explained by semantic cues only. In a simple scenario including embedded links and navigation-bar options, navigation elements are choice items with variable utilities. Element utility $U_E(d)$ depends on the distance from the target page, estimated by the user within a hierarchical representation of the web site (Farris, Jones, & Elgin, 2002). In such a representation, generic information connected to thematic macro-categories is at the top of the hierarchy while specific information is at the bottom.

The utility of navigation-bar options is constant over distance (squares in Fig. 1a), because the bar leads to the highest information level consistent with the target, independent of user's position. Bar utility depends on its position in the web page, becoming lower as the distance from the conventional top position increases. This is achieved by modulating bar utility by R, an attenuation factor proportional to the ratio between actual and conventional positions.

The utility of embedded links is an inverse function of distance (diamonds in Fig. 1a). Since embedded links lead to specific information, their utility increases as the estimated distance from the target becomes smaller. The utility function of embedded links has been chosen as the positive half of a Gaussian distribution centered in d= 0.

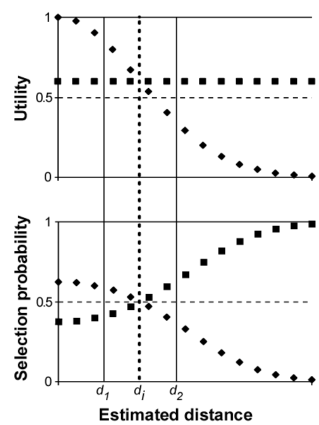

Figure 1: From utility (a) to selection probability (b) of navigation-bar options (squares) and embedded links (diamonds). Both function pairs intersect at d_i, where $U_B(d)= U_L(d)$ and $P_B(d)= P_L(d)= 0.5$.

The probability of element selection (Fig. 1b) results from the following equation:

$$P_E(d) = \frac{U_{En}(d)}{\sum_{i=1}^{N} U_{Ei}(d)}$$

where $P_E(d)$ is the selection probability of element E and N is the number of navigation elements.

Three predictions follow from WebStep assumptions.

1. The greater the estimated distance, the greater is the selection probability of a bar option with respect to the selection probability of a semantically equivalent embedded link. In particular, at the ordinal level $[d> d_i] \Rightarrow [P_B(d)> P_L(d)]$ and $[d< d_i] \Rightarrow [P_B(d)< P_L(d)]$.

2. At small estimated distances, as the bar approaches the conventional top position the difference between selection probabilities of the embedded link and bar options

decreases (and vice versa at large estimated distances). This formalizes the idea that the bar at the top is more easily identified than the bar at the bottom.

3. The difference between selection probabilities of the two navigation elements increases when the page heading is present. Such a regularity is achieved by implementing a distributed extraction of the estimated distance by means of a MonteCarlo method formalizing the non deterministic nature of the selection behavior.

Simulation

WebStep selection probabilities were computed by a Matlab program using three input parameters:

bar position, from a maximum for the bar on top to a minimum for the bar on bottom;

d_{ideal}, ranging from zero (user in target position) to ∞;

heading: presence= 1, absence= 0.

The program extracted 1000 distance values from a Gaussian distribution of distances centered in d_{ideal}, with heading-dependent width $\sigma_{presence} > \sigma_{absence}$. For every distance, element utilities were computed and the final utility value was defined by taking the average. Selection probabilities were computed on the basis of final utility values. This process was iterated 50 times to reach higher precision in the selection probability estimation. Final probabilities were computed as averages from 50 iterations. Selection time was computed by means of the linear combination of a component inversely proportional to selection probability and a processing constant for each element: $T(E)= 1/P_E(d) + K_E$. By manipulating free parameters values the intersection between the two utility functions was set at (4.0, 0.6). As a consequence $U_B(d)$ was equal to 0.6 while $U_L(d)$ was the left half of a Gaussian centered in d= 0 with σ= 31.3.

Experiments

To test the empirical validity of WebStep we run two experiments, in which the semantic aspects of bar options and embedded links were equivalent. Every label appeared on both elements, which could be differentiated only on the basis of their spatial and structural features. To test the first prediction we manipulated the starting navigation page. In Exp. 1 the starting navigation page was in the no-target position, inducing a high estimated distance from target (content-target in-congruency). In Exp. 2 the starting-navigation page was in target position, inducing a low estimated distance from target (content-target congruency). In both experiments we tested prediction 2 by manipulating the bar position (top vs. bottom) and prediction 3 by manipulating the the heading (presence vs. absence). The selection rate and time of elements in the starting navigation page were recorded and used as dependent variables.

Results

Obtained results were consistent with WebStep predictions. In Exp. 1 we found a preference for the selection of the bar with respect to the embedded link, $\chi_1^2= 6.37$; $p< 0.05$: *users use the bar to move towards the macro-category congruent with the target*. Furthermore, consistently with second prediction, the bar selection rate increased when the bar was in the conventional top position, [$\chi_1^2 = 5.31$; $p< 0.05$]. Consistently with the third prediction the difference between the selection probability of the two elements increased in presence of the heading [$\chi_3^2 = 14.20$; $p< 0.01$]

The inverse relation between selection time and selection rate (slope= -1.9; r^2= 0.8) confirmed that the choice of an element with low selection probability requests a higher time than the choice of a navigation element with high selection probability.

In Exp. 2, consistently with the first prediction we found a preference for the embedded-link selection when the heading was present, [$\chi_1^2 = 4.263$; $p< 0.05$]: *users use the embedded link to move towards the sub-category congruent with the target*. Furthermore, consistently with second prediction, the difference between selection-rate of the two elements was lower when the bar was at the bottom than at the top. We also found a higher difference between elements selection probability when the heading was present [$\chi_3^2 = 14.94$; $p< 0.01$].

As a final analysis we compared selection times and selection probabilities derived from simulations with values obtained in Exps. 1 and 2. The fit was significant for both probability (r^2= 0.7, $F_{1,15}$= 29.7, $p< 0.001$) and selection time (r^2= 0.5; $F_{1,15}$= 22, $p< 0.001$).

Conclusions

Results are consistent with the idea that navigation depends also on pragmatic aspects of elements, as well as on contextual aspects like the page heading. Our WebStep model fits obtained effects.

References

Farris J.S., Jones K.S. & Elgin P.D. (2002). Users' schemata of hypermedia: what is so spatial about a web site? *Interacting with Computers*, 14, (5), 487-502.

Kitajima M., Blackmon M.H. & Polson P.G. (2000).. A Comprehension-based Model of Web Navigation and Its Application to Web Usability Analysis. In S. McDonald, Y. Waern & G. Cockton (eds.), *People and Computers XIV - Usability or Else! (Proceedings of HCI 2000)*, Springer, pp.357-373.

Miller C. S. & Remington R. W. (2000). A computational model of Web navigation: Exploring interactions between hierarchical depth and link ambiguity. In *The proceedings of The 6th Conference on Human Factors and the Web*.

Pirolli P. & Fu W. T. (2003). SNIFT-ACT: A Model of Information Foraging on the World Wide Web. *Ninth International Conference on User Modeling*, Johnstown, PA.