

A Model to Simulate Web Users' Eye Movements

Myriam Chanceaux¹, Anne Guérin-Dugué², Benoît Lemaire¹ and Thierry Baccino³

¹ University of Grenoble, France
<first name>.<last name>@imag.fr

² University of Grenoble, France
Anne.guerin@gipsa-lab.inpg.fr

³ Laboratoire LUTIN, Paris, France
baccino@lutin-userlab.fr

Abstract. One of the most important tasks on the Web is foraging information. In this study, we present a computational model which simulates human eye movements during information seeking in Web pages. Human visual scanpaths are guided by their information needs, by the visual features of the stimuli and by what they previously processed. Our model takes into account both semantic (top-down) and visual (bottom-up) information, as well as a memory model in order to predict the focus of attention. Our model operates at the block level, but also at the word level. To validate this model, both participants and model were asked to seek information in a pseudo online newspaper. We find good correspondence between simulated and empirically observed scanpaths. Knowing where the user is looking at while searching for information is crucial for evaluating the usability of Web sites and contribute to the design of Web pages.

Keywords: Computational model, Information foraging, Web page, Usability, Semantic, Visual saliency, Memory

1 Introduction

A central concern of HCI research is how to determine the *usability* of user interfaces during both the design and final evaluation phases [1]. This usability assessment can be done by qualitative, criteria oriented and model-based methods (e.g., interview study, checklist, and cognitive model). Most of the time, cognitive models are missing and ergonomists have to rely on qualitative data to assess a user interface. The problem with such qualitative evaluation is that it is both expensive and time consuming. Model-based methods enable more detailed predictions of quantitative parameters, for example error rates, times and sequences of actions compared to qualitative and criteria based methods [2]. This allows to apply model-based methods in early phases of system-development processes, to detect usability problems and to change the examined user interface. Integrating these aspects, cognitive models additionally take into account the cognitive abilities and characteristics of human beings [3]. There are a number of cognitive models that have been used for usability

assessment with the earliest GOMS [4] or CCT [5] and more recent work using architecture like ACT-R [6] to test usability by means of simulated users [7]. These models may be outlines of tasks written on paper or computer programs which enable us to predict the time it takes for people to perform tasks, the kinds of errors they make, the decisions they make, or what buttons and menu items they choose. To sum up, cognitive models can be used in several ways: to determine ways of improving the user interface so that a person's task has fewer errors or takes less time, to build into the user interface to make software that reacts more effectively to help people use the system by anticipating their behavior or inferring their mental state, or as a means of testing current psychological theory. The goal is to radically reduce the time and cost of designing and evaluating usable systems through developing analytic engineering models for usability based on validated computational models of human cognition and performance.

Following the same approach of model-based methods for measuring the usability of a user interface, we present a computational cognitive model of Information Foraging on the Web. This model simulates the search for information on the Web (by combining the visual, semantic and memory processing) starting from realistic tasks and ill-defined goals. It is intended to explain how the visual exploration is made on Web pages containing both texts and images and as a consequence to provide fruitful information to any interface designers. The information seeking process implemented in our model is currently mainly based on textual information processing on one Web page without navigation.

2 Task

Our model is simulating a user searching for information on a pseudo online newspaper. Before explaining our model in details, we now describe the task, which will be performed by both humans and model.

This task consists in searching a piece of information in a simulated Web page representing an online newspaper. The goal is to find the block of text that best corresponds to an expression given beforehand. In this article we will refer to this expression as the *theme*. Examples of themes are *global warming*, *rise in the stock market*, *soccer players' win...* Since the goal is to find the block which *best* corresponds to the theme, users are encouraged to scan all blocks although they already found an interesting one, because the next one may be better. In addition, this design also encourages users to revisit previously visited blocks, which makes the overall scanpath quite complex.

Pages are composed of seven texts, one of them being associated with a picture. Texts are divided into 3 categories: 2 texts are semantically very close to the theme and are therefore potential targets; 2 texts are related to the theme, and 3 texts have no semantic link with the theme.

Figure 1 shows an example of a page. The theme is "planet observation". One text deals with the giant planets rings, another with a study of the solar system; they are the Strong Association texts (SA). Another text is on the problems of a spacecraft's antenna and another one on a software for aerial photographs; these two texts have

Low Association (LA) with the theme. The three others do not have any association with the theme (Without Association, WA). They are about a complaint filled by a company; about the French swimmer Laure Manaudou and about the after-career of a basketball player.

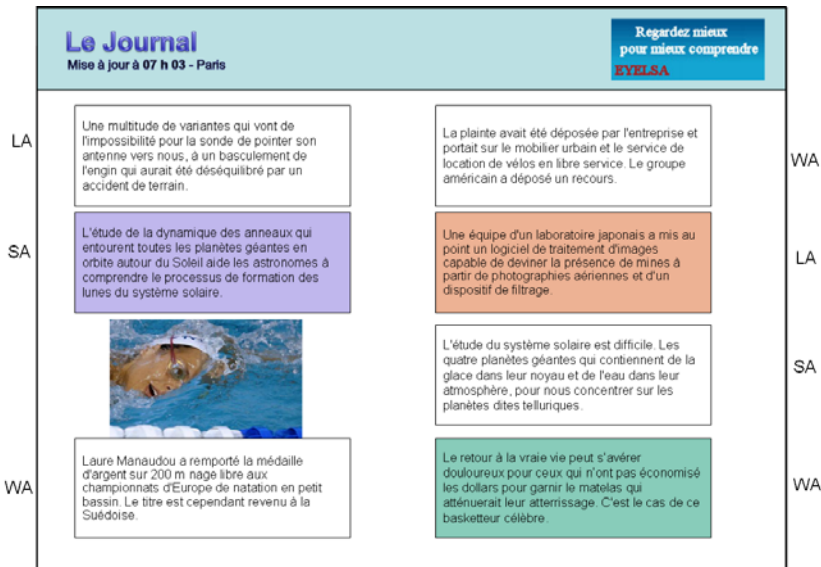


Fig. 1. Example of page, theme is *planet observation*. During experiment, block types are not visible (SA, LA and WA).

To design these pages, we used the semantic measures provided by Latent Semantic Analysis (LSA) [8]. A semantic space was constructed from a 24-million word corpus composed of articles from the French newspaper *Le Monde* (year 1999). Once this space was constructed, we chose 20 themes and represented all of them by a 300-dimensional vector. Then, we selected for each of them 2 SA texts, 2 LA texts and 3 WA texts according to their cosine with the theme vector. A high cosine corresponds to a high semantic similarity. We sometimes had to manually revise the texts in order to keep an approximately constant length. Here are the semantic similarity statistics. SA: mean 0.7, max 0.92, min 0.45; LA: mean 0.19, max 0.46, min 0.14; WA: mean 0.01, max 0.13, min -0.05.

Since our model is intended to simulate the user conflict between semantic and visual attraction, we added visual information to our pages in addition to their semantic organizations. Therefore, blocks can have a white or colored background. There are three possible colors per image. The colors were chosen in order (1) to show iso-luminance (same contrast with the white background and black text) so that they all have the same readability and (2) to have the same color contrast between them (they form the apex of a triangle in the Lab color space). In each page one SA, one LA and one WA have colored backgrounds.

In order to be closer to a real Web page and have more variability in the semantic and visual ways of eye attraction, we added a picture to each page. This picture can be

either associated with a SA block (4 out of 20 pages) or a WA block. There is no picture associated with a LA block. This avoids situations of reinforcement of the semantic link from the combination of textual and visual information. This link would not have been quantitatively controlled, with a risk that the semantic link between the theme and the LA block would have been greatly modified to the point of changing its semantic class (from LA to SA).

We designed 20 pages for 20 themes of information search (one theme per page). A randomized procedure (Latin square) on the colors and places created 6 lists (a participant only sees one list).

3 Model

Our model is intended to reproduce the average sequence of eye fixations a user is likely to do, given a page and a theme. The model output is therefore a sequence of fixations which are XY coordinates. We assumed that there are two distinct levels of navigation in the page. The first one consists in going from block to block (block level) whereas the second one operates within each paragraph (word level).

Basically at the block level this spatio-temporal model is based on the dynamic integration of visual and memory information. Each of these components is implemented by means of a conditional heat map of the current image, in which each of its elements is assigned a weight representing its relevance for the given component. The visual map is conditional on the location of the current fixation. It tends to reinforce local and visually salient regions. The memory map is conditional on the entire scanpath. It tends to penalize already seen regions. These maps are continuously updated during the simulated visual search. The two maps are integrated by a weighted sum and the simulated gaze is moved towards the best-weighted block. Once the new fixation has been selected, the second part of the model is required, and then maps are updated accordingly, a new block is chosen, and so on. From an initial fixation point, our model thus produces a path between block.

At the word level, the model is based on semantic information contained in the sequence of words fixated and compared with the theme of information foraging. This level enables the model to modulate the memory weight of the block, for further visits.

We now detail these two modeling levels.

3.1 Block Level

As we mentioned earlier, selecting the next block to process is based on two cognitive processes: a visual one and a memory one. Those two processes may be conflicting: the visual process may suggest visiting a close block whereas the memory process may indicate that it has already been seen recently. Our model therefore includes a mechanism to integrate those two sources of information. At this level, each process dynamically and independently determines a weight for each block.

Visual Process. The human eye is such that it has a local visual acuity which prevents it processing items that are far from the current location. Since long saccades are also physiologically costly, the human eye tends to operate locally. However, there might be visually salient items that could attract the attention. Sophisticated models of visual saliency have been proposed in the literature, mainly for natural scenes. These models are often based on the Feature Integration Theory [9]. Among them, the most popular is proposed by Itti and Koch [10]. It is based on a feature decomposition of the visual stimuli and generates a visual saliency map. The highest salient regions are then segmented and sorted according to their saliency value. For the first eye fixations on a picture, these models fit well the eye movements data when the visual stimuli have little semantic information and when the task is free without explicit task driving the scene exploration [11]. In the case of more demanding visual search, the visual saliency is progressively modulated over time by semantic and cognitive controls, depending of the type of the scene (a priori knowledge of the scene) and the task [12]. Our page is simple and here we just dynamically determined visual values for each block (B) by taking into account the spatial proximity with the current fixated block following the classical curve of visual acuity as a function of eccentricity [13]. In its current version, our model does not take into account visual features such as the various background colors of blocks.

Memory Process. The memory map also contains eight values representing the weight of each block. Human generally do not move back to locations previously visited. This mechanism is close to the Inhibition of Return (IOR) principle [14], which is often used in visual attention model. Here this process is used between blocks and not for each fixation, but the mechanism is quite similar. The IOR refers to an increased difficulty of orienting attention to a location to which attention has previously been directed.

Therefore, in our model the current block is given a weight of 0, meaning that it would be better not to go back to it. We also implemented a forgetting mechanism which enables to go back to previous location after a while. Actually, humans may visit again blocks that were considered interesting (see Results section). This level of interest is given to our model by the semantic similarity between the theme and the part of block that have been processed at the word level (see next section). Basically, blocks that have never been seen receive a maximum weight, blocks already seen but irrelevant are assigned a minimum weight and blocks already seen and interesting receive an intermediate weight according to their interest. All these weight are decreased at each step to simulate a forgetting mechanism. At each transition from one block to another one, the memory weight of each previously seen block B_i is calculated following this equation:

$$\text{MemoryWeight}(B_i) = \text{MemoryWeight}(B_{i-1}) + (1 - \text{MemoryWeight}(B_{i-1})) / \delta \quad (1)$$

δ was experimentally set to 6.

Integration. Each of these two processes can be viewed as computing a map, with exactly 8 distinct areas in our design. At each step, these two maps are summed to form the general map from which the best-weighted block is going to be fixated next. The sum is, for each block B_i :

$$\text{IntegrationMap}(B_i) = \alpha_M * \text{MemoryWeight}(B_i) + \alpha_V * \text{VisualWeight}(B_i) \quad (2)$$

These two processes do not have the same role. In fact in a previous version of this model we showed that the most important component in such a model is the memory process (see [15] for more details). In this previous experiment, the stimuli were very different (isolated words), but the involved processes were not. What was at the level of words in the previous experience is now at a block-level. α_M and α_V were set respectively to 0.65 and 0.35. In this new study we keep these parameter values. Although limited to 8 areas in the current task, this mechanism can be extended to the larger number of distinct items that are usually found in real Web pages. As we will see in the following sections, this simple combination of maps reproduces the typical participant course. In fact, memory will always encourage new blocks and the visual map will promote the blocks nearby.

3.2 Word Level

At the level of the words inside a block, the model operates differently. The process by which humans gather information in a block of text is simulated by considering a word by word scanning. Some users do not actually read paragraphs linearly and prefer to pick words here and there, but most of our participants process a paragraph from its first word. However, what is very important to simulate is the fact that users do not read completely a paragraph if they realize after a while that it does not correspond to what they are looking for. All users stop reading uninteresting paragraphs and the challenge is to model that decision appropriately.

There are three conditions which seem to govern the decision to stop or to continue reading a paragraph:

- If the words processed so far are *unrelated* to the theme (using the LSA semantic measures), then the paragraph is abandoned with a zero value, so that the block is not likely to be visited again by the block level process;
- If the words processed so far are semantically *highly similar* to the theme, then the block is abandoned, but it is memorized as a potential solution by weighting it for the block level process with its semantic similarity with the theme;
- If the words processed so far are neither highly related nor semantically unrelated to the theme, then the next word is processed: the model keeps reading until it can take a decision.

Two thresholds of semantic similarity were defined for that purpose, one for defining what is a high similarity and one for the unrelatedness. These thresholds depend on the number of words already processed in the block: the more words the model sees, the more confident it can be in its decision. For example, if only two words are processed, the highly-similar-threshold is very high because the decision to abandon the block should only be done if the similarity between those two words and the theme is very very high. However, if 25 words have been processed, the threshold can be lowered because there are enough words to have a high confidence in the semantic measure. Thresholds are defined in the following way (P is a paragraph, W is the sequence of words processed so far and T is the current theme):

$$\text{Abandon (P) if similarity (W,T) < } \log(\text{nbWords(W)}) / 20 . \quad (3)$$

$$\text{Abandon (P) if similarity (W,T) > } 1 - \log(\text{nbWords(W)}) / 5 . \quad (4)$$

These functions were defined experimentally and provided an appropriate behavior. Figure 2 presents these thresholds as the function of the number of words processed so far. If the similarity with the theme is in-between the two lines, the model keeps reading and checks the next word. In the other cases, the block is abandoned and the block level part of the model determines which block will be fixated next.

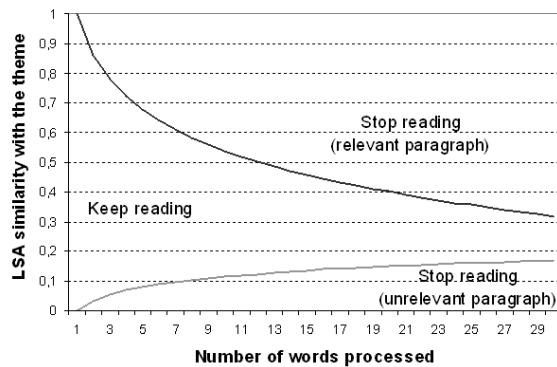


Fig. 2. Similarity thresholds as a function of the number of words

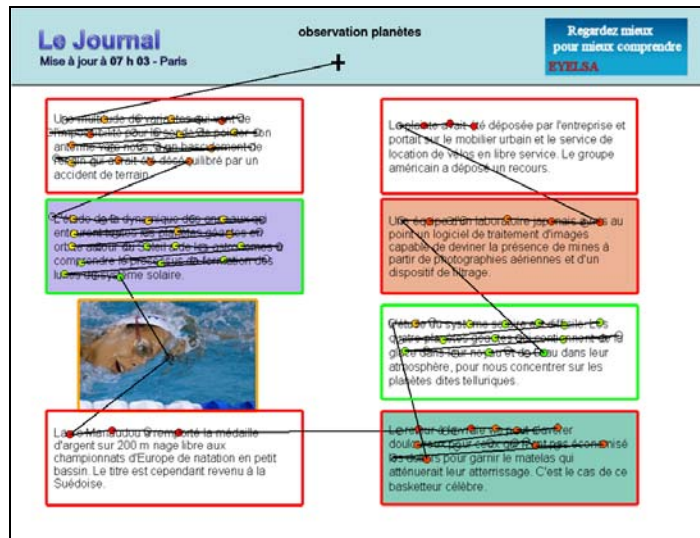


Fig. 3. Example of a scanpath generated by the model. Blocks abandoned because of high similarity with the theme are encircled with a green (gray)rectangle. Blocks abandoned because of unrelatedness are encircled with a red (black) rectangle.

Our model does not analyze the picture, nor does it take into account its visual saliency. Therefore, it considers it as a block and just makes one fixation on the middle of the image.

Figure 3 presents an example of scanpath on the image presented earlier. Obviously, the model does not run on an image: it knows exactly the XY location of each word of each paragraph.

4 Test of the Model

In order to test the validity of our model, we recorded the visual scanpaths of human participants and compared their behavior with the model.

4.1 Experiment

Participants. Thirty-eight participants took part in the experiment. All had normal or correct-to-normal vision.

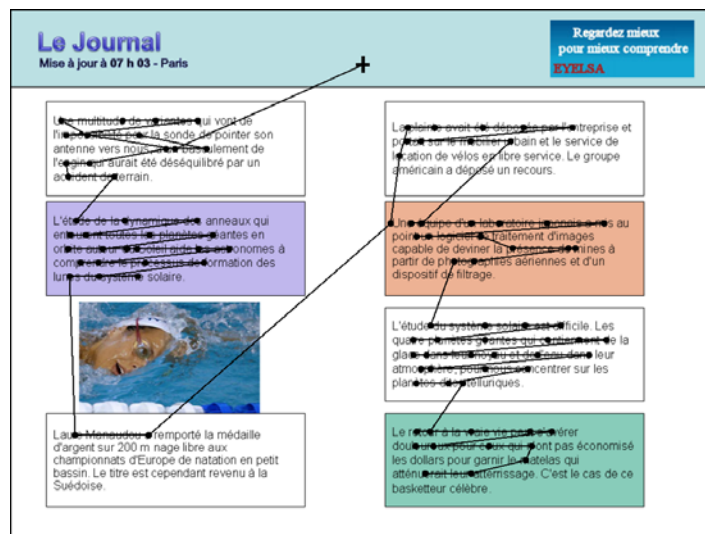


Fig. 4. Example of an experimental scanpath from a human being.

Apparatus. Eye movements were recorded using a SR research EyeLink II eye tracker, sampling pupil position at 500 Hz. Eye position data were collected only for the guiding eye. Search displays subtended 40 horizontal deg. of visual angle (1024 * 768 pixels). Head position and viewing distance were fixed to 50 cm. A chin-rest was also used to keep the subject's head stable. This apparatus records horizontal and vertical gaze position data (x and y) in screen coordinates (pixels) for the guiding eye.

From these coordinates, saccades and fixations are determined, leading to an experimental scanpath, as shown in figure 4.

Procedure. Each trial begins with an instruction followed by a cross fixation. This crosshair is positioned at the top of the page to be far from the text, in a non-informative area. After the participant gazed on this point an image appeared. Participants had to find the best block according to their opinion without maximum delay, then click on the selected block and give their confidence in their choice. Participants were asked to respond as quickly as possible without sacrificing accuracy and were told that many answers were possible. Figure 5 illustrates the process.

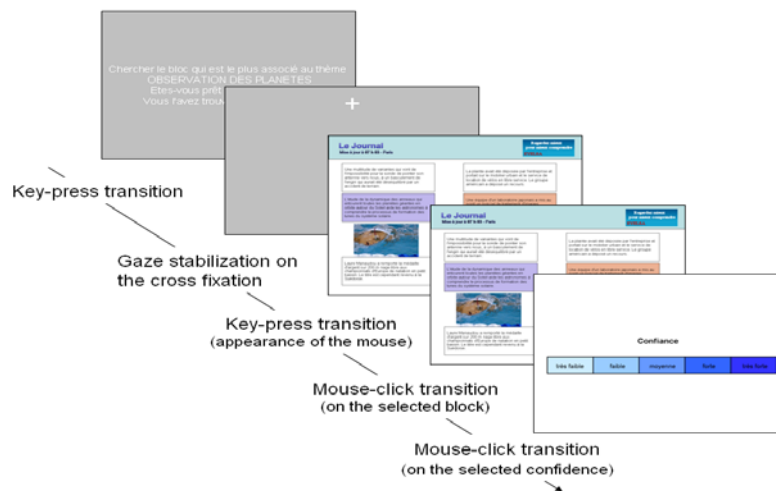


Fig. 5. Experimental process.

There are 20 different experimental pages, and each of the 38 users saw all the pages, leading to 760 scanpaths. A scan on one page takes on average 36,298 ms (SD 22,841), 89.33 fixations (SD 52.63), into on average 11.36 blocks (SD 6.06).

4.2 Model Data

Our model performs the same task as humans, on the 120 (20 themes * 6 lists) pseudo Web pages. As it is deterministic, we obtained exactly 120 scanpaths we compared with empirical data.

For each page, we stopped the model when it reaches the average number of blocks seen by participants on this particular page.

4.3 Humans and Model Comparisons

Block Level. At this level, we looked at whether the participants and the model have systematic *block by block* paths.

The first column (the left one) is almost always visited from top to down, for both human (79.07%) and model (81.67%). Then, we noticed different behaviors: the matrix of transitions (Figure 6) between blocks showed us that the **W** path is the most common for humans, and the **U** path for our model. Indeed, humans are going down on the left and the right column. This might be due to the habit of reading from left to right and top to down, and the fact that the blocks are closer vertically than horizontally. On the contrary, the model goes up the second column (the right one) because its visual process prevents a large saccade from lower left to upper right.

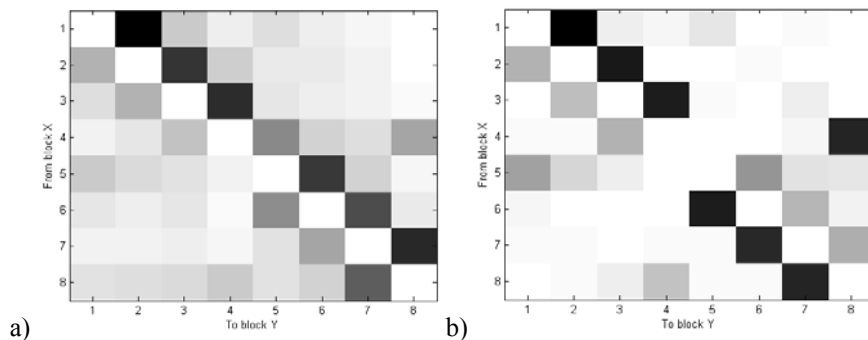


Fig. 6. Transitions matrices for human a) and model b). Blocks are numbered from upper left to lower left (1 to 4), then from upper right to lower right (5 to 8). The darker the cell, the more frequent the corresponding transition.

Once all blocks have been fixated, participants often returned to the most interesting blocks, those which are semantically closest to the theme, as did the model. Figure 7 shows these rates of refixations, which are, as expected, significantly higher for SA blocks than for others and especially WA ones. (F tests, all $p < 0.01$) This high rate of refixation for interesting blocks could be due to the fact that the goal in this information seeking task is ill-defined, and that two blocks are potential targets. There is therefore a choice to make between these two blocks, and to do it, reading them again could be necessary.

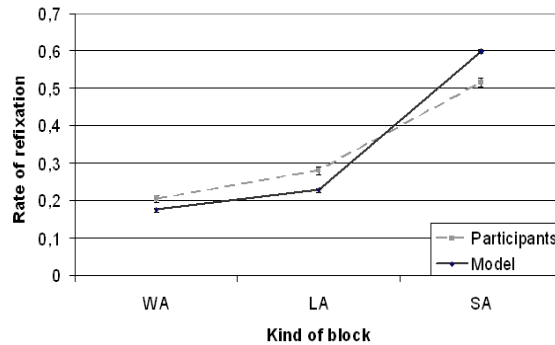


Fig. 7. Rate of blocks revisited for human and model according to the kind of block (Without Association, Law Association and Strong Association)

Last comparisons at this level will be for the background color of the block. Even if the model does not take into account this feature for the moment, it seems interesting to know how human deals with that.

What we observed on all data is that human gazed more often on colored block than others (Figure 8), $T(37)=12.92$, $p<0.01$. We plan to slightly improve our model in the future to account for that phenomenon.

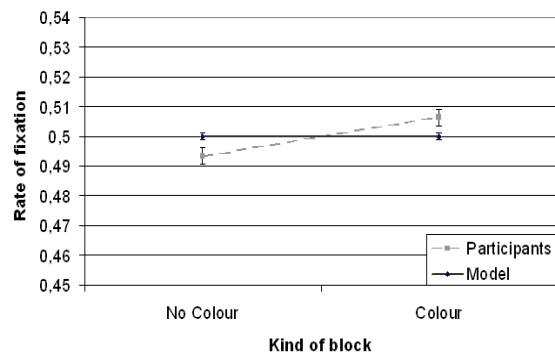


Fig. 8. Rate of blocks visited according to the kind of background, for model and participants

Word Level. At this level we looked at the number of fixations in a block for each block category (SA, LA or WA) during the first visit of the block. These data were normalized by the number of words in the paragraph. The results showed that the WA blocks are sooner left than the others. That means that people do not need to read a lot of words to decide that this block is unrelated to the goal. For instance, Figure 4 showed that only two fixations were necessary for that user to decide that the lower left block was unrelated to the theme. The others blocks (LA and SA) are read in more details, as shown in Figure 9. The rate of reading is significantly lower for WA

than SA and LA blocks (T tests, both $p < 0.01$) but there is no significant difference between SA and LA blocks ($T(37) = 1.44$, $p = 0.16$).

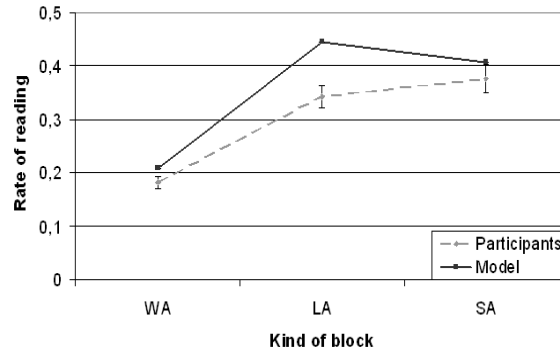


Fig. 9. Rate of reading (number of fixations divided by number of words) for human and model according to the kind of block

Our model obtains quite the same results, even if it reads more words before leaving the LA block, all differences are significant (F test, all $p < 0.01$). This is a bias in the model conception, indeed when a block is interesting but not very interesting the model does not leave it because it does not know how to consider it, and need to process more information.

5 Discussion

Our model of a user searching information on a Web page can be compared to two systems, SNIF-ACT and CoLiDeS, which are considered complementary in the literature [16]. SNIF-ACT [17-18] was implemented using the ACT-R architecture. It simulates the navigation from pages to pages, considering each page as a whole. Basically, SNIF-ACT computes the utility of staying in the current page compared to going back to the previous page, clicking a link to go to another page or leaving the website. This model is using environmental regularities as explained in [19]. When searching for information on one Web page, the model evaluates the probability for each link to lead to the desired information. It includes an adaptive link selection mechanism that sequentially evaluates links on a Web page according to their position. At each step SNIF-ACT is faced on the exploration/exploitation trade-off problem: how to balance exploration of new actions against exploitation of actions that are known to be good. There are three possible actions: attend to the next link (exploration), click on a link (exploitation) or return to the previous page.

Our model is only simulating a single Web page. It is therefore closer to CoLiDeS [20]. This system first segments the Web page into regions called patches. Then it uses user preferences or LSA semantic similarities between regions and the current user's goal to select the object which will be the focus of attention. Then it performs an action, generally clicking on an hyperlink.

Our model also takes as input a Web page and a user's goal. The main difference is that our approach is intended to be more cognitively plausible. For instance, instead of segmenting the whole page into relevant patches, our model operates more locally because it reproduces the human visual system which cannot process the entire page in a detailed way in one shot. In addition, our model attempts to predict at a lower level since it simulates eye fixations. However, the limit of our system is that it cannot for the moment deal with complex interface objects such as buttons, links or menus that CoLiDeS can manage. To apply the model to a larger number of Web pages, to deal with more realistic web pages, it must take into account these different objects. It could then be applied in Web design, enabling designers to have an idea of what users will read according to what they want, which areas of the screen will be visited, in what order, etc. Various information searches could be tested, helping designers to validate Web pages and information content of the pages.

Taking as input a simple Web page, our model is able to predict the course from block to block, and also the rate of reading in each paragraph according to an ill-defined goal.

Further steps in the development of the model will also focus on adding a visual saliency model considering that human tend to fixate more often the colored blocks, and to be able to deal with the picture. Another issue will be to diversify the visual features, and to change the distribution map of the blocks, to prevent the systematic block by block strategy, and add a more global strategy, with preferred directions and/or locations. Finally to extend the model we should add a navigation process between several Web pages instead of dealing with just one page.

Acknowledgements

We would like to thank Gelu IONESCU for providing us LisEyeLink software; and Olivier SZARZENSKI for his work during the experiment. We also thank participants who accepted to pass the experiment.

References

1. Baccino, T., Bellino, C., & Colombi, T. (2005). *Mesure de l'utilisabilité des Interfaces*. Paris: Hermès Science Publisher (Lavoisier).
2. Dix, A. J., Finlay, J. E., Abowd, G. D., & Beale, R. (1998). *Human-Computer Interaction*. London: Prentice Hall Europe.
3. Newell, A. (1990). *Unified Theories of Cognition*. Cambridge, MA: Harvard University Press.
4. Card, S. K., Moran, T. P., & Newell, A. (1983). *The Psychology of Human-Computer Interaction*. Hillsdale: New Jersey: Lawrence Erlbaum Associates.
5. Kieras, D. E., & Polson, P. (1985). An approach to the formal analysis of user complexity. *International Journal of Man-Machine Studies*, 22(4), 365-394.
6. Anderson, J. R., Matessa, M., & Lebiere, C. (1997). ACT-R: A theory of higher level cognition and its relation to visual attention. *Human-Computer Interaction [Special Issue: Cognitive Architectures and Human-Computer Interaction]*, 12(4), 439-462.

7. West, R. L. & Emond, B. (2001). Can cognitive modeling improve usability testing and rapid prototyping? In *Proceedings of the Fourth International Conference on Cognitive Modeling*, pp. 271-273. Mahwah, NJ: Lawrence Erlbaum Associates.
8. Landauer, T., McNamara, D., Dennis, S., Kintsch, W., (2007) eds.: *Handbook of Latent Semantic Analysis*. Lawrence Erlbaum Associates
9. Treisman, A., Gelade, G.: A feature-integration theory of attention. *Cognitive Psychology* 12(1) (1980) 97–13614
10. Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience*, 2(3), 194-203.
11. Mannan, S.K., Ruddock, K.H., Wooding, D.S. (1996). The relationship between the locations of spatial features and those of fixation made during visual examination of briefly presented images. *Spatial Vision* 10 165–188
12. Henderson, J., Brockmole, J., Castelano, M., Mack, M. (2007) Visual saliency does not account for eye movements during visual search in real-world scenes. In van Gompel, R., Fischer, M., Murray, W., Hill, R., eds.: *Eye movements: A window on mind and brain*. Oxford: Elsevier 537–562
13. Geisler, W. S., & Perry, J. S. (2002). Real-time simulation of arbitrary visual fields. In *Proceedings of the 2002 symposium on eye tracking research & applications* (pp. 83–87). ACM New York, NY, USA.
14. Klein, R. M. and MacInnes, J. W. (1999) Inhibition of return is a foraging facilitator in visual search. *Psychological Science*, 10(4):346–352.
15. Chanceaux, M., Guérin-Dugué, A., Lemaire, B., & Baccino, T. (2008). Towards a model of information seeking by integrating visual, semantic and memory maps. In B. Caputo & M. Vincze (Eds.), *ICVW 2008*. Heidelberg: Springer-Verlag.
16. Kitajima, M., Polson, P.G. & Blackmon, M.H. (2007). CoLiDeS and SNIF-ACT: Complementary Models for Searching and Sensemaking on the Web. *Human Computer Interaction Consortium (HCIC) 2007 Winter Workshop*.
17. Pirolli, P.L. & Fu, W.-T. (2003). SNIF-ACT: a model of information foraging on the World Wide Web. In *Proceedings of the 9th International Conference on User Modeling*, Lecture Notes in Artificial Intelligence, 2702, 45-54.
18. Fu, W. and Pirolli, P. (2007). SNIF-ACT: a cognitive model of user navigation on the World Wide Web. *Human-Computer Interaction*, 22, 4, 355-412.
19. Fu, W.-T. (2007), Adaptive Tradeoffs between Exploration and Exploitation: A Rational-ecological Approach. Exploitation: A Rational-Ecological Approach. In Gray, W.D. (Ed), *Integrated Models of Cognitive Systems*, Oxford: Oxford University Press.
20. Kitajima, M., Blackmon, M.H. & Polson, P.G. (2005). Cognitive architectures for website design and usability evaluation: comprehension and information scent in performing by exploration. In *Proceedings of HCI-International 2005*, vol. 4: Theories, Models and Processes in HCI (paper 1866).