Model of Making Decisions during an Information Search Task

Francisco López-Orozco, Luis D. Rodríguez-Vega

Universidad Autónoma de Ciudad Juárez,
División Multidisciplinaria de Ciudad Universitaria, Cd. Juárez, Chih.,
Mexico

Francisco.Orozco@uacj.mx

Abstract. This paper presents a cognitive computational model of the way people read a paragraph with the task of quickly deciding whether it is related or not to a given goal. In particular, the model attempts to predict the time at which participants would decide to stop reading the paragraph because they have enough information to make their decision. Our model makes predictions at the level of words that are likely to be fixated before the paragraph is abandoned. Human semantic judgments are mimicked by computing the semantic similarities between sets of words using Latent Semantic Analysis. A two-variable linear threshold is proposed to account for that decision, based on the rank of the fixation and the semantic similarity between the paragraph and the goal. Model performance is compared to eyetracking data of 19 participants.

Keywords: Eyetracking, information search, LSA.

1 Introduction

Knowing what web users are doing while they search for information is essential. Several cognitive models have been proposed to account for some of the processes involved in this activity. Pirolli & Fu ([8]) proposed a model of navigation. Brumby & Howes ([2]) describes how people process information partially in order to select links related to an information goal. Chanceaux et al. ([3]) show how visual, semantic and memory processes interact in search tasks.

Information search can be made on any kind of documents, but we are here interested in textual documents, composed of several paragraphs. Information search is different from pure reading because people have a goal in mind while processing the document. They have to constantly keep in memory this additional information. For example, if the task is to decide whether a paragraph is related or not to a given goal, the paragraph and the goal are the pieces of information involved and both have to be together managed in order to make a correct decision.

This paper attempts to model that particular decision making. It focuses on a behavior that is specific to information search, which is stopping processing a paragraph before it is completely read.
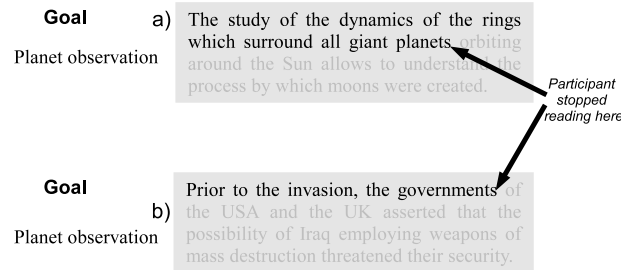
Fig. 1: Illustration of the 2 input data of the model: the goal and the paragraph. The paragraph is abandoned before its end because enough information has been gathered and maybe due to a) a high-relatedness b) a low-relatedness to the goal.

This particular problem has been studied by Lee and Corlett ([6]): participants were provided with a topic and a text, presented one word every second, and were asked to decide as quickly as possible if the text is about the given topic. However, we aim at studying a normal reading situation instead of presenting one word at a time. We will therefore rely on an eyetracker to identify the words processed. Figure 1 illustrates the situation we aim at modeling.

2 Experiment

In order to create and study a model, we designed an experiment to gather some data. This experiment was intended to emphasize the decision to stop reading a paragraph while one piece of information is stored in memory: the search goal. A set of 30 goals was created. Each one is expressed by a few words (e.g. *mountain tourism*). For each goal, 6 paragraphs were created (mean=30.1 words, σ=2.9), 2 of them being highly related to the goal, 2 of them being moderately related, and 2 of them being unrelated. We used Latent Semantic Analysis (LSA) (Landauer et al., [5]) to control the relatedness of a paragraph to the goal. We have not assured a uniform distribution among the frequency words. We sometimes had to manually revise the texts in order to keep an approximately constant length and a particular semantic similarity. The semantic similarity between two sequences of words such as a goal and a paragraph can be computed using the cosine function. The higher the cosine value, the more similar the two sequences of words. We trained LSA on a 24 million word general French corpus.

The experiment is composed of 30 trials, each one corresponding to a goal, in random order. Participants were asked to decide as quick as possible if a paragraph is related or not to a goal given beforehand. In each trial, one paragraph is presented to the participant. The participant should indicate when he is ready to make a decision. Then the paragraph disappears and the participant is asked to keep or reject the paragraph according its relatedness or not with the goal. A new paragraph is displayed and the participant should again decide if the paragraph it is related or not to the goal. This procedure is repeated until all

Le traitement des données est
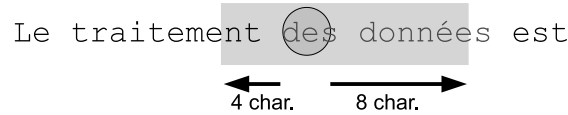
4 char.   8 char.

Fig. 2: Words processed during a fixation using our window-based approach.

6 paragraphs of the current goal were displayed. Each participant was therefore exposed to 30*6=180 paragraphs, and asked to decide if they are related or not to a goal. 19 students participated in the experiment. Eye movements were recorded using a SR Research EyeLink II eye tracker. From these coordinates, saccades and fixations were determined, leading to an experimental scanpath. The stimuli pages were generated with a software that stored the precise coordinates of each word on the screen. We wrote our experiment in Matlab, using the Psychophysics Toolbox (Brainard, [1]).

Before trying to mimic eye movements, we had to predict which words were actually processed by participants in each fixation. It is known that the area from which information can be extracted during a single fixation extends from about 3-4 characters to the left of fixation to 14-15 characters to the right of fixation (Rayner, [10]). This area is asymmetric to the right and corresponds to the global perceptual span. Therefore, more than one word may be processed for a given fixation. In order to determine which ones were processed for each fixation, we used a window, sized according to Rayner ([10]). He showed that the area from which a word can be *identified* extends to no more than 4 characters to the left and no more than 7-8 characters to the right of fixation and corresponds to the word identification span. Moreover, Pollatsek et al ([9]) show that even if information of the next line is processed during a reading task, participants are not capable of getting some semantic information. Therefore, the size of our window is 4 x 1 characters to the left plus 8x1 characters to the right of the fixation point. Since the initial fixations in the beginning part of a word facilitate its recognition more than initial fixations toward the end of the word (Farid & Grainger, [4]), we considered that a word is processed if at least the first third of it or the last two-thirds is inside the window. In the example of Fig. 2, two words were supposed to have been processed: *"des"* and *"données"*. The second one is considered because at least the first third of it is inside the window. The word *"traitement"* is not supposed to have been processed because at least the last two-thirds of the word are not inside the window.

3 Modeling

The model should be able to predict the way an average user would process a paragraph given a goal. Our method is therefore to consider the experimental scanpaths and for each participant's fixation to predict whether the paragraph would be abandoned or not. A very good model would predict an abandon at

the same time the participant stopped reading. A bad model would abandon too early or too late.

Such a decision making model on paragraphs needs to be based on a model of semantic memory that would be able to mimic human judgments of semantic associations. We used LSA to dynamically compute the semantic similarities between the goal and each set of words that are supposed to have been fixated.

3.1   Relatedness Effect of the Paragraph

The relatedness of the paragraph to the goal may play a role in the way it is processed. We suspected that if the paragraph is not related to the goal at all, the paragraph would be processed just to verify that is not relevant for the given goal since that all its words are unrelated (with a low semantic association) with the goal. The number of fixations needed to confirm its not relatedness with the goal may depend of each participant and in this paper we are not interested in this particular case. However, if the whole paragraph is highly related to the goal, the paragraph is composed of unrelated, moderately and highly related words to the goal. Due to this richness of words, the paragraph should be processed with the idea of investigate whether it is relevant or not. Here the relatedness of the paragraph to the goal plays a role in the way it is processed. Our analysis is restricted when the paragraph is strongly related to the goal.

3.2   Relatedness Evolution of the Paragraph

All the paragraphs are composed of coherent text and with an approximately constant length. However the evolution of its similarity with each goal was not controlled. In some of the paragraphs, the similarity value could be going up or down from the beginning to the end of the paragraph and in other cases an alternating pattern of increasing/decreasing values could be observed. After a clustering procedure according to the cosine evolution of only the strongly related paragraphs we learned 3 different classes: *step*, *ramp* and *saw*.

The cosine evolution of step paragraphs is characterised by a constant value during some fixations followed by an abrupt change (positive jump) to a higher value which is set till to the *signal* end. In the case of the *ramp* paragraphs, a cosine value is maintained during some fixations but here it is followed by a progressive increasing of the cosine to a higher value which is set till to the *signal* end. The cosine evolution in the *saw* paragraphs is still more complicated. Here, the evolution of the cosine value follows a increasing/decreasing pattern along the duration of the *signal*. We called them *saw* because of its resemblance with a non-sinusoidal waveform. Some paragraphs were not clustered in any of these classes due its high complexity of its cosine evolution.

In order to simplify the posterior analyses, only the *ramp* paragraphs were chosen because they have a higher cosine variability than the *step* paragraphs but their variability is more predictable than the case of *saw* paragraphs.
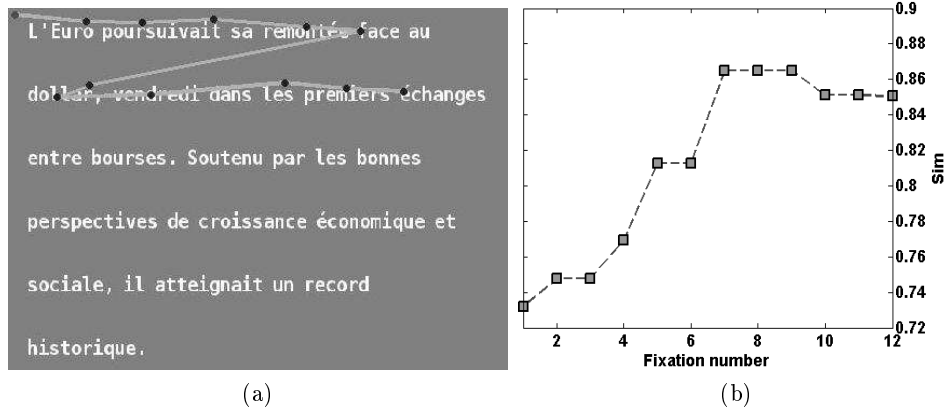
(a)   (b)

Fig. 3: a) Example of scanpath in the C|S condition. b) Its *Cos* evolution.

3.3   Modeling the Decision

Two variables involved. We first looked for the variables which could play a role in the decision to stop reading a paragraph p. Such a decision is made when the participant's perception about the relatedness of the paragraph with the goal is completed. The association to the goal g is obviously involved in that perception. Therefore, we defined a variable called $Cos = |sim(\text{words of } p, g)|$ in which sim is the LSA cosine between the two vectors. *Cos* changes constantly while a paragraph is processed since it depends on the words actually processed. When the words seen are highly related with the goal that variable has a high value and it has a low value (close to zero) when the words are unrelated with the goal. It can be easily calculated dynamically, after each word of the paragraph has been processed. Consider for example Fig. 3a. The goal is *"faiblesse du dollar"* (*weak dollar*). In the first fixation on the paragraph, only the word *"L'Euro"* is supposed to has been processed according to our window-based prediction. This word is highly related with the goal. Therefore $Cos = |sim(\text{"L'Euro", "faiblesse du dollar"})| = 0.731$.

During fixation 2, two words are processed, the word *"L'Euro"* and a new word *"poursuivait"* leading to a new value of $Cos = |sim(\text{"L'Euro poursuivait", "faiblesse du dollar"})| = 0.747$.

During fixation 3, only the word *"poursuivait"* is processed, leading to the same value of $Cos = 0.747$.

In fixation number 4, $Cos = |sim(\text{"poursuivait sa remontée", "faiblesse du dollar"})| = 0.769$. In fixation 5, the *Cos* value goes up to 0.812 because of the words *"remontée face"* which makes the LSA vector much more similar to the goal vector. This value is maintained for the fixation 6. In fixation 7, the *Cos* value goes up to 0.865 because of the word *"dollar"* makes still the LSA vector much more similar to the goal vector. During fixation number 9, the words *"vendredi dans"* makes the LSA vector less similar to the goal vector and this effect is showed by the decreasing value of $Cos = 0.850$. This value is maintained till
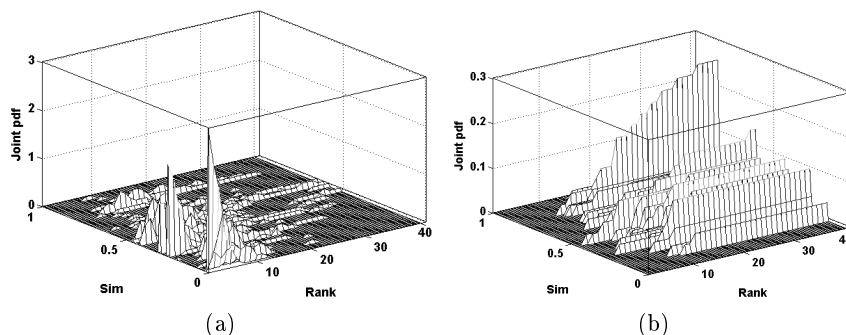
Francisco López-Orozco, Luis D. Rodríguez-Vega

(a)   (b)

Fig. 4: a) Empirical "no-abandon" distribution $\hat{p}_{CR}(c, r|\overline{Ab})$ and b) "abandon" distribution $\hat{p}_{CR}(c, r|Ab)$ in the $Cos \times Rank$ space.

to the 12th fixation (stop reading). Figure 3b shows the evolution of the Cos value along the fixations in the scanpath. This example illustrates that a high value of Cos may not directly induce the decision, in particular if it appears too early in the scanpath. We assume that the decision also depends on the number of words processed so far in the paragraph. The more words processed, the higher the confidence in the perception of the relatedness between the paragraph and the goal. If only two or three words have been processed, it is less likely that Cos is accurate. Therefore, we assume that there should be a relationship between Cos and the number of words processed. The second variable is then $Rank$ = number of words processed so far.

Abandon and no-abandon distributions. In order to study how the decision depends on these two variables, we computed two distributions in the $Cos \times Rank$ space of participant data: the distribution of the no-abandon cases and the distribution of the abandon cases. The goal is to learn the frontier between both cases in order to be able to predict if a sequence of words already processed is likely to lead to the abandon or the pursuance of the reading task. This work was done on two thirds of the data, in order to leave one third to test the model. Each participant fixation was associated to a point in the $Cos \times Rank$ space. $Rank$ is a discrete measure between 1 and the maximum number of fixations in the data (60 in our case). Cos has been computed according to the previous formula, taking into account the words already processed in each paragraph as well as the goal and discretized into one of 100 bins, from 0 to 1.

The no-abandon distribution was computed by simply counting the number of fixations that did not lead to an abandon for each cell of the $Cos \times Rank$ grid. It concerns all fixations except the last one of each scanpath.

The abandon distribution was built from all very last fixations of all scanpaths, including also subsequent ranks. For example, if a given participant on a given stimulus made 13 fixations, the first 12 were counted in the no-abandon distribution and the 13th was counted in the abandon distribution. All virtual fixations from 14 to 60, with the same Cos value as the 13th were also counted in
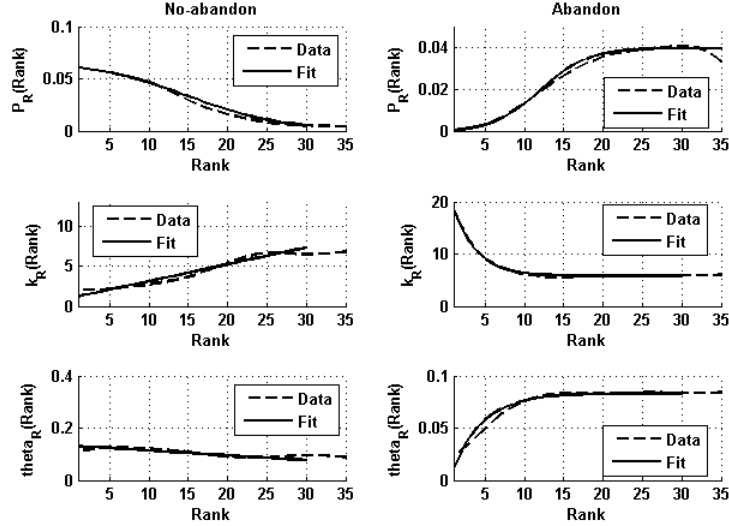
Fig. 5: Data and fitting of marginal distributions, shape and scale for the "no-abandon" and "abandon" distributions.

the abandon distribution, because if the participant stopped reading at fixation 13, he would have also stopped at fixation 14, 15, etc. The frontier between these two behaviors (continue or stop reading) is a curve in the $Cos \times Rank$ space. Depending on the location of any observation (c,r) above or under the curve, the reader's behavior can be predicted. To find this frontier, a methodology based on a Bayesian classifier is used. Let us consider a classification problem with two classes: Abandon (Ab) and No-abandon $(\overline{Ab})$. Given the posterior probabilities, which is the class of a two-dimensional observation (c,r) in the $Cos \times Rank$ space? The decision rule is then: $P(\overline{Ab}|c,r) \underset{Ab}{\overset{\overline{Ab}}{\gtrless}} P(Ab|c,r)$ with $P(\overline{Ab}|c,r) = \frac{P(\overline{Ab}) \times p_{CR}(c,r|\overline{Ab})}{p_{CR}(c,r)}$, and $P(Ab|c,r) = \frac{P(Ab) \times p_{CR}(c,r|Ab)}{p_{CR}(c,r)}$. Figures 4a and 4b represent the two empirical class-conditional probability density functions respectively $\hat{p}_{CR}(c,r|\overline{Ab})$ and $\hat{p}_{CR}(c,r|Ab)$. We adopt a statistical parametric approach. By this way, data will be regularized since they are obviously affected by the noise inherent to acquisition and pre-processing.

In the next sections, the statistical model to estimate the density functions and the prior probabilities are explained in order to use the Bayesian classifier:

$$P(\overline{Ab}) \times p_{CR}(c,r|\overline{Ab}) \underset{Ab}{\overset{\overline{Ab}}{\gtrless}} P(Ab) \times p_{CR}(c,r|Ab).$$

Parametric model for the "no-abandon" distribution. The class-conditional probability density function can be written as : $p_{CR}(g,r|\overline{Ab}) = p_{C|R}(c|R =$

$r, \overline{Ab}) \times p_R(r|\overline{Ab})$. Figure 5 (top, left) shows the empirical marginal distribution $\hat{p}_R(r|\overline{Ab})$. As the *Rank* increases, the probability of not abandoning the paragraph decreases. This evolution was modeled with a sigmoid function $\varphi(r) = \frac{P_{RMax} \times \left(1 + e^{-\alpha r_0}\right)}{1 + e^{\alpha(r-r_0)}}$. There are actually only two parameters to fit because the integral is 1.

Concerning the probability density function $p_{C|R}(.)$, the natural model (Fig. 4a) is a Gamma one whose parameters depend on the *Rank* value. The shape $k(r)$ increases and the scale $\theta(r)$ decreases linearly (Fig. 5, left column). The linear regressions are only performed up to the *Rank*30 since that $\hat{p}_R(r > 30|\overline{Ab})$ is close to zero and there is no more enough data. Then we have:

$$p_{C|R}(c|R = r, \overline{Ab}) = \frac{A(r)c^{k(r)-1}}{(k(r)-1)!\theta(r)^{k(r)}}e^{-\frac{c}{\theta(r)}}, p_R(r, \overline{Ab}) = \varphi(r).$$

As the *Cos* value is between 0 and 1, $A(r)$ is a normalization function to ensure that $p_{C|R}(c|R = r, \overline{Ab})$ is a probability density function: $A(r) = F_{k,\theta}(1) - F_{k,\theta}(0)$, with $F_{k,\theta}(.)$ being the repartition function of a Gamma distribution with a shape k and a scale θ. We then obtained six independent parameters to model the complete "no-abandon" joint distribution (offset and slope for the sigmoid, and the coefficients for the two linear functions).

Parametric model for the "abandon" distribution. Following a similar approach the class-conditional pdf is written as : $p_{CR}(c, r|Ab) = p_{C|R}(c|R = r, Ab) \times p_R(r|Ab)$. The marginal pdf $\hat{p}_R(r|Ab)$ was modeled with another sigmoid function $\varphi'(r)$ (Fig. 5, top right). But here, it is an increasing function. At rank 0, there is no abandon and at the maximal *Rank* value, all scanpaths have shown an abandon. The conditional distribution $\hat{p}_{C|R}(c|R = r, Ab)$ is a Gamma distribution with a shape $k'(r)$ and a scale $\theta'(r)$. The shape $k'(r)$ exponentially decreases while the scale $\theta'(r)$ exponentially increases (Fig. 5, right column). Equations of the pdf are the same as the previous case, but with a different set of functions $\{\varphi'(r), k'(r), \theta'(r)\}$ which gives us eight parameters (2 for the $\varphi'(r)$, 3 for $k'(r)$ and 3 for $\theta'(r)$):

$$p_{C|R}(c|R = r, \overline{Ab}) = \frac{A'(r)c^{k'(r)-1}}{(k'(r)-1)!\theta'(r)^{k'(r)}}e^{-\frac{c}{\theta'(r)}}, p_R(r, \overline{Ab}) = \varphi'(r).$$

Modeling the decision as the function of Rank and Cos. As these two class-conditional probabilities were modeled, for each (*Rank*, *Cos*) values, the problem is to decide if there is enough information to stop reading ("abandon" class), or to continue reading ("no abandon" class). This binary problem is solved thanks to the Bayesian classifier. To find this decision rule, we have now to estimate the prior probabilities such as : $P(Ab) + P(\overline{Ab}) = 1$. $P(Ab)$ or $P(\overline{Ab})$ is another parameter to learn from the data. The total number of learning parameters is then 15 (6+8+1). The decision rule is then:

$$P(\overline{Ab}) \times p_{C|R}(c|R = r, \overline{Ab}) \times p_R(r|\overline{Ab}) \underset{Ab}{\overset{\overline{Ab}}{\gtrless}} P(Ab) \times p_{C|R}(c|R = r, Ab) \times p_R(r|Ab).$$

4   Model Learning

After learning the two posterior probabilities $P(\overline{Ab}|c,r)$ and $P(Ab|c,r)$ to represent the decision frontier between the two classes, the two prior probabilities are $P(Ab) = 0.64$ and $P(\overline{Ab}) = 0.36$. The intersection between the posterior probabilities is oblique which is what was expected, from a cognitive point of view. *Rank* and *Cos* are dependent on each other: at the beginning of processing the paragraph, there should be a high relatedness between the paragraph and the goal to make the decision. However, after more fixations have been made, that relatedness could be lower to decide to abandon the paragraph.

For instance, at rank 10, a *Cos* of 0.7 is necessary to stop reading, whereas at rank 15, a value of 0.3 is enough. The frontier is rather linear and can be approximated by the following equation in the $Cos \times Rank$ space: $Cos_0 = -0.0473 \times Rank + 0.9849$. That equation was included in the computational model. That model constantly computes the *Cos* value while it is moving forward in the text, increasing the *Rank* value. As soon as the current *Cos* value is greater than Cos_0, the decision is to stop reading the paragraph.

In order to test the model, we ran it on the remaining one third of the data. For each fixation in this testing set, the model decides either to leave or not to leave the paragraph. If the model did not leave at the time the participant stopped reading, simulation is pursued with the next rank and with the same value of the gap, and so on until the decision is made. The average difference between the ranks at which model and participant stopped reading was computed. We got a value of 4.2 (SE=0.5). To assess the significance of that value, we built a random model which stops reading after each fixation with probability p. The smallest average difference between participants' and model's ranks of abandoning was 9.9 (SE=0.9) and was obtained for $p = 0.22$. Our model therefore appears to be much better than the best random model.

5   Conclusions

We presented a binary model which predicts the sequence of words that are likely to be fixated before a paragraph is abandoned given a search goal. In spite of the drawbacks of LSA method, we got good model's performances but a further investigation should be done in order to know in more details the impact of these limits in our modeling work. For example, a natural comparison of performance would be by using LSA versus a simple Bag of words (BoW). In our model, two variables seem to play a role: the rank of the fixation and the semantic similarity between the paragraph and the search goal. We proposed a simple linear threshold to account for that binary decision. Our model will be improved in future work. In particular, we aim at considering a non linear way of scanning the paragraph, using another model of eye movements (Lemaire et al., [7]). We also plan to tackle more realistic stimuli as well to consider other decisions involved in Web search tasks.

References

1. Brainard, D. H.: The Psychophysics Toolbox. Spatial Vision, 10, 433–436 (1997)
2. Brumby, D. P., Howes, A.: Good enough but I'll just check: Web-paged search as attentional refocusing. In: Proc of the 6th ICCM Conference, 46–51 (2004)
3. Chanceaux, M., Guérin-Dugué, A., Lemaire, B., Baccino, T.: A model to simulate Web users' eye movements. In: Proc of the 12th INTERACT Conference, LNCS 5726, Berlin: Springer Verlag, 288–300 (2009)
4. Farid, M., Grainger, J. How initial fixation position influences word recognition: A comparison of French and Arabic. Brain & Language, 53, 351–368 (1996)
5. Landauer, T., McNamara, D., Dennis, S., Kintsch, W.: Handbook of Latent Semantic Analysis. Lawrence Erlbaum Associates (2007)
6. Lee, M.D., Corlett, E.Y. Sequential sampling models of human text classification. Cognitive Science, 27(2), 159–193 (2003)
7. Lemaire, B., Guérin-Dugué, A., Baccino, T., Chanceaux, M., Pasqualotti, L.: A cognitive computational model of eye movements investigating visual strategies on textual material. In: L. Carlson, C. Hölscher, T. Shipley (Eds.) Proc of the Annual Meeting of the Cognitive Science Society, pp. 1146–1151 (2011)
8. Pirolli, P., Fu, W.: SNIF-ACT: a model of information foraging on the world wide web. In: P. Brusilovsky, A. Corbett, F. de Rosis (Eds.), 9th ICUM, pp. 45–54 (2003)
9. Pollatsek, A., Raney, G. E., LaGasse, L., Rayner, K.: The use of information below fixation in reading and in visual search. Can J Psychol, 47, 179–200 (1993)
10. Rayner, K.: Eye movements in reading and information processing: 20 years of research. Psychological Bulletin, 124(3), 372–422 (1998)