Biologically Inspired Cognitive Architectures II:
Papers from the AAAI Fall Symposium (FS-09-01)

Analysis of the Web User Behavior with a Psychologically-Based Diffusion Model

Pablo E. Román and Juan D. Velásquez
University of Chile, Department of Industrial Engineering
República 701, Santiago, Chile.
proman@ing.uchile.cl, jvelasqu@dii.uchile.cl

Abstract

This work presents a new application of a mathematical theory of psychological behavior from Usher and McClelland and the random utility model from McFadden, to the web user behavior. The model describes the stochastic behavior of a general kind of web users, consisting of the probability of following a hyperlink for a specific length of time. The simulation experiment together with the artificial agent illustrates behavioral patterns characteristic of human subjects.

Introduction

Studying the web user behavior has an advantage over other cognitive fields. For instance, a medium sized website accumulates 1GB of data which is related to historic web user behavior per month. A different approach is presented, based on the neurophysiology representation of decision making in the brain of an agent (Usher and McClelland 2001), named LCA model (Equation 1). This model associates the neural activity levels (NAL) of certain brain regions with a discrete set of possible choices. Those NAL's (Y_i) evolves according to a stochastic equation while the agent is in the process of making their decision, until one of the NAL reaches a given threshold. In the previous case, the agent takes the decision that corresponds to the choice associated with the maximal NAL. The stochastic equation depends on the choice evidence levels (CEL's). A CEL (I_i) is a neural activity level of a brain region that is associated with a unique choice, which value computes a kind of likelihood for the choice before the decision is made.

The Model

Web users are considered stochastic agents according to the LCA model (Equation 1) with internal state Y_i (NAL's values) and incremental white noises dW_i. The links on a web page are the available choices $\{i\}$, and furthermore they also take into consideration the choice of leaving the web site. It is assumed that each agent's link preference is defined by a TF-IDF text vector μ (Manning and Schutze 1999). An agent prefers to follow similar links to the vector μ that represents its preferences. The utility values are given by the **dot** product between the **normalized** TF-IDF vector μ and L_i that represent the text associated with the link i. The CEL values I_i are approximated using the Logit model (Equation 2) (McFadden 1973).

$$dY_i = (-\kappa Y_i - \lambda \sum_{j \neq i} f(Y_j) + I_i)dt + \sigma dW_i \quad (1)$$

$$I_i = \omega e^{L_i \cdot \mu} / \sum_j e^{L_j \cdot \mu} \quad (2)$$

Web User are simulated by mean of artificial agents using the equation 1, resulting in the time it take to decide each options and the distribution of the sequences of pages.

Conclusions and further research

The experiment were performed using typical values (Usher and McClelland 2001) used for parameters $(\kappa, \lambda, \omega, \sigma)$. For μ, a vector was used with the TF-IDF value of the most important word from a real web site (http://www.dii.uchile.cl). The sequences of pages obtained by artificial agents are found to be closer to $\sim$5% of the average of the number of visits to each page. Nevertheless, better adjustment could be obtained if one considers μ to be a distribution of users to be fitted by non-parametric methods. The stochastic model seems to approximate the mechanism used by human agents for employing higher learning capacities (I_i) in order to make a decision. Others applications in Economy could be further explored with agents that made decision choices trading with goods.

Acknowledgments

This work has been partially supported by the Doctoral Grant from Conicyt Chile and by the Chilean Millennium Institute of Complex Engineering Systems.

References

Manning, C. D., and Schutze, H. 1999. *Fundation of Statistical Natural Language Processing*. The MIT Press.

McFadden, D. 1973. Is conditional logit analysis of qualitative choice behavior. *Zarembka (ed.), Frontiers in Econometrics, Academic Press*.

Usher, M., and McClelland, J. 2001. The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review* 2(1):550–592.