Evaluating the probabilistic side of proactive computing using cognitive modelling methodology and the statistical inference of user's cognitive states

Denis Shirnin and Denis Zampunieris
Computer Science and Communications Research Unit,
University of Luxembourg,
6, rue Richard Coudenhove-Kalergi, L-1359, Luxembourg.
Email: {denis.shirnin, denis.zampunieris}@uni.lu

***Abstract*—In the following paper we present the work in progress, which aims to validate through theoretical and empirical study the probabilistic side of proactive computing. We choose the cognitive modelling methodology as an approach to implement the probabilistic inference of user's cognitive states during online activity. We create four cognitive models, which simulate various cognitive states of a user during the task of online search of medical terminology. Additionally, we elaborate several potential scenarios, which may take place during the given task, and we associate these scenarios with our models as the integral elements of the proactive system behaviour. Ultimately, by applying the rules of the Bayesian statistics, we envision to test the principles of proactive computing in the frame of probabilistic approach.**

***Index Terms*—Cognitive modelling, Proactive Computing, Model evaluation, Cognitive inference engines.**

I. Introduction

Throughout the last decade Proactive computing became one of the paradigms entailing the expertise of various research domains, which resulted in its gradual shift towards interdisciplinary field [1], [2]. In order to find the effective strategies to relate the aspects of human cognition to the specifics of software systems, we explore the potentials to include the cognitive science expertise into proactive computing research. Such position prompts us to adapt the attributes of human cognition to the algorithmic level.

Our current theoretical framework includes different theories and approaches among which, the proactive computing, and the cognitive modelling occupy the main positions. Each of the fields provides a valuable contribution of the constructive and applicable concepts and methods.

The paper represents a work in progress, which takes place in the frame of the graduate student's research.

A. Proactive computing

Being the main conceptual framework of our system, proactive computing provides us with its unique characteristics and potentials. It enables the system to generate the context aware capabilities and proactive actions, if needed, in relation to contextual situation of a user and his or her online activity.

In our research we understand the notion of *Proactivity* as it was defined by Tennenhouse. In his efforts to explore a new paradigm of computer science research, he proposes to re-examine the relationships between physical and abstract domains of a context in order to build more intuitive software systems, which will be enhanced by proactive type of behaviour [3], [4]. The chief features of a proactive system are characterised by its capacity to be aware of a surrounding context and associated developing situations. Moreover, the system has to be able to act accordingly, to the needs of such evolving situations on its own initiative by providing the adapted services [2], [5]. Perhaps, the most important attribute of proactive computing is the capability to translate a data associated with the situation development, and use it as an evidence in order to act accordingly, with respect to the future situations [6].

The current state of research in the given domain demonstrates that this field is still in its prime stage of development [4], [7]. Such position, allows and encourages us to explore the proactive computing in different directions. Throughout our past research efforts, we investigated the deterministic side of proactive computing, where we studied the possibility of implementing the preprogrammed set of context aware-capable rules in order to enhance the system with proactive type of behaviour [8], [9]. However, it represents only a segment of the theoretical framework applied in development of the proactive system. In order to enhance, and to elaborate the deterministic part of theoretical structure, we currently focus our research on studying and investigating the probabilistic side of proactive computing. Different approaches are used for re-examining the initial conceptual framework. One of the strategies to be considered is the *Cognitive modelling* approach.

B. Cognitive modelling

Currently, cognitive modelling methodology and Bayesian statistics are our main approaches in overall objective to augment proactive system with the probabilistic features. Given, that a user is the main agent of a context, we decided to use the cognitive modelling approach as the main technique to incorporate the statistical analysis of user's cognitive states into proactive system. The chosen approach reflects thoroughly

Proc. 2014 IEEE 13th Int'l Conf. on Cognitive Informatics & Cognitive Computing (ICCI*CC'14)
S. Patel, Y. Wang, W. Kinsner, D. Patel, G. Fariello, & L.A. Zadeh (Eds.)
978-1-4799-6081-1/14/$31.00 ©2014 IEEE

the aspects of human-computer interactions, which allow us to extend the capabilities of proactive system with more intuitive and user-oriented features.

The principles of cognitive modelling approach make part of a more global and generic theory of *Cognitive architecture*, which objective is to simulate and represent the structures of various human cognitive processes on computational level [10], [11]. The substantial efforts have been made to simulate the different phenomena of human cognition, where one of the notable examples is ACT-R system and its integrated theory of the mind [12].

Among general characteristics of the cognitive modelling methodology, several aspects can be highlighted. Essentially, the cognitive modelling approach may help to define the relationships between perception and representation and representation and overt behaviour [10]. Such approach allows in its turn to provide the detailed description of cognitive processes with algorithmic specificity and thus to implement it on computational level [11], [13].

Using cognitive modelling as an approach that takes into account the characteristics of user's cognitive processes, is one of the main objectives of our study. Furthermore, we have chosen this strategy as the technique for implementing the probabilistic aspects into proactive system. In the current state of research, we employ the probability theory in order to estimate the future model matching of user's cognitive states against empirical data.

In following sections, we show in details the specifics of cognitive modelling methodology. These sections are divided into description of the statistical structure, and the probabilistic attributes of the algorithm. In the last section, we show some examples of the proactive system behaviour based on principles of the Bayesian statistics.

II. Statistical components of cognitive modelling methodology

The main objective of current experiments is to explore, develop and validate the probabilistic side of our proactive system. Here we use, previously defined in our study the deterministic framework as the key disposition, for building around it the new probabilistic principles.

We use the local search engine and its medical database as the environment for our experiments. The search engine is the selection-based type, divided into two perspectives, search page and the results page. The search page contains query information arranged into categories, which are represented and accessible by the combo boxes (see Figure 1). Additionally, the search page includes an optional keyword field. The search can be achieved through the mouse clicking, by selecting sequentially the corresponding categories in the combo boxes. In order to open the next combo box, the user has to open the preceding one first. An exception is the keyword field, where the user may optionally type a sequence of characters to elaborate the category selection. The results page contains two fields, the result list and the comments box. The result list displays eight elements at most, where the rest is accessible through the scrollbar. The comments are displayed only when one of the result's elements is selected. The user may access a website, which is displayed at the moment of selection, and which is associated to a chosen result's element.

The proactive system will be tested locally by members of the University staff, where the participants will be asked to perform a search of medical related topics on the search engine. The capabilities of engine will be enhanced by our proactive system. If needed, the system will be tested on the larger scale later, by deploying the testing module into a search engine of the partner's medical website.

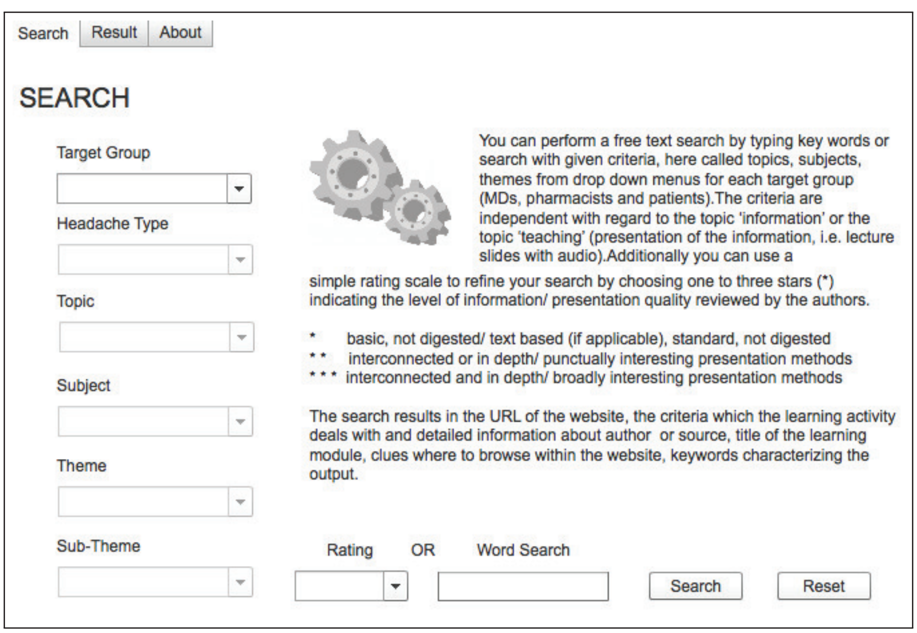

Fig. 1. Search page perspective

As we stated earlier, the initial objective of a study is to implement the user-oriented proactive behaviour into a target system and validate it with regard to the probabilistic aspects of data matching. Additionally, the objective of a study is to validate the concept with regard to the specifics of human cognitive processes, which are involved in certain types of the human-computer interactions [14], [15]. Therefore, by achieving the aforementioned goals, the system will be capable to detect the various cognitive states of a user through probabilistic matching of the cognitive models against a real-time empirical data [11], [16]. Subsequently, the implemented approach will allow the system to generate the proactive actions that aim to guide a user in his/her search attempts. The variety of proactive actions may include the detection of most relevant results, display of suggestions as well as prevention of potential issues, such as user's stumbling with the low-relevant results.

The main strategy for the mentioned above objectives consists of modelling and developing the various cognitive states of a user on algorithmic level, and its subsequent fusing into the system. Thus, for purposes of current study, we have specified two cognitive states to be tested, that is the user's state of satisfaction and the user's state of dissatisfaction. Once the models are developed and integrated, the probabilistic matching will be applied upon them. The last step will consist of defining and specifying the types of proactive behaviour.

A. *Statistical structure of the cognitive models*

Throughout the scientific research, Bayesian statistics has proved to be an efficient framework for representing and processing various information. It has been applied in different studies that aimed either to relate the models to a data or to build the probabilistic models of cognition [13], [17].

In our study, we use statistics as the main technique for the algorithm to connect and to relate to the specifics of user's cognitive states. Thus, if collected and allocated properly, the statistics may reveal the various types of information in relation to the user's contextual characteristics, such as the aspects of user's cognitive states, his/her objectives or interests during online activity. In such disposition, the statistical data represents a variation of user's cognitive manifestations.

In order to build the aforementioned framework, we employ several types of statistics, such as time detectors, timers, selection data, mouse clicks, time intervals, data interrelations, data associations and disassociations. The given, or the similar combinations of statistics allow to perform the algorithmic computations upon a data in order to reveal the hidden user's cognitive states and the various aspects of the context settings. Moreover, the additional types of statistics can be always added, if needed by the objectives of a model.

Essentially, the statistical data is used as a user-related input data, which necessitates an additional knowledge for an interpretation and its later manipulations. Thus, we use the cognitive science expertise in order to define the specifics of certain user's cognitive processes and their representations in overt behaviour. Later, we relate the defined cognitive characteristics to a statistical data. Therefore, the particular sequence of a statistical data will represent in the end a distinct aspect of the user's cognitive state. The main idea is to arrange and organise the statistics into logical, coherent sequences, chain of sequences, patterns and eventually models, which will represent in the end the overtly manifested instances of user's cognitive states. The objective of current experiments is to develop and to test four cognitive models, that is, the model of user's satisfaction and dissatisfaction expressed either by experienced or inexperienced user during an online search of medical terminology. Subsequently, the algorithm has to be able to detect a model and to distinguish its type, that is, to identify the user type and his/her current cognitive state.

In order for the system to be able to perform the statistical data collection, we elaborate and put into practice a mechanism, which consists of several steps. First, we define the relevant types of statistics that may help us to reveal the aspects of user's cognitive processes, and his/her context settings. The second important step is to define the statistical relationships, correlations, and disassociations between different types of input data. As we mentioned earlier, a sequence of predefined statistical data constitutes the complete model. Therefore, in order to detect a model, the algorithm has to detect progressively all its integral individual data inputs. For these purposes, we create a set of rules, where each rule has its own objective to detect the specific type of a data instance [18]. The rules themselves are part of the proactive system, which is connected to a target system. The rules are processed by the *Rules Engine*, which is responsible for storing, executing and iterating the rules [19], [20]. Initially, several rules are launched at the start of proactive system with the objective to be continuously activated in order to search for the first instance of a data input. Once the relevant type of statistics is detected, the initial rule launches the new rule, which is associated with the next type of statistics in a data sequence. The process of cognitive model detection continues this way until the decision point is reached, which is defined by the algorithm. According to proactive system configurations, all rules are launched sequentially, however, the certain rules can be activated non-sequentially, depending on model characteristics. Throughout the whole process, a model will start to form and to reveal itself from a real-time empirical data. Thus, as a next step, the detected segments of a model will allow us to apply the statistical inference of a potential further model matching against the user data.

B. *Probabilistic character of the main algorithm*

The initial objective of a model is to simulate the particular cognitive state of a user, or to simulate a problem, which is related to user's online activity. Thus, when we develop a model we define all instances of statistical data and we relate them to the various aspects of user's cognitive states. Ultimately, a model will help to reveal some relevant data regarding the user's online activity. As we mentioned earlier, there are several types of models, where each model provides its unique perspective on a user's cognitive sate and his/her

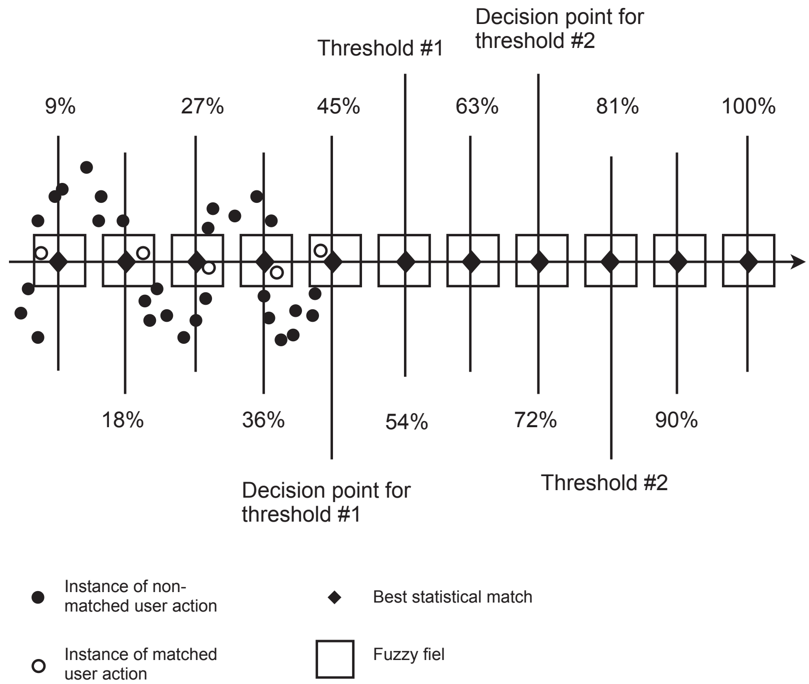

Fig. 2. Representation of the model matching process

objectives. Our goal is to enhance the target system with the proactive functions through such models. Therefore, in order to act proactively, the system has to possess some data in advance, which means it has to be able to make an inference about the needed data ahead. For these purposes, we apply the Bayesian statistics upon the models in order to estimate the probability of their future matching against the user's empirical data. It will allow the system to secure some time window for the proactive actions before an actual event happens.

In order to show the specifics of probabilistic inference, the structure of a model should be explained first. In overall schema, a model is the representation of a complete event or a problem. For the computing purposes, we divide the whole event into the smaller sub-events, represented by the instances of statistical data. Thus, a model is a static sequence of allocated micro patterns of the statistical data. Throughout its length, a model is marked by the milestones, or in other words the thresholds. The thresholds are the points on a model, which correspond to a percent representing the status of model progression. The objective of thresholds is to mark the absolute endmost moment of triggering the corresponding proactive actions, meaning that all proactive scenarios have to be launched before the actual threshold is reached. Several thresholds may be placed along the model depending on its complexity, meaning that numerous proactive scenarios may be applied throughout the whole model.

Figure 2 demonstrates an overall mechanism of a model matching and the data detection, which consists of the several steps. Primarily, the objective is to reveal the type of a model from the user's empirical data. For this purpose, the system has to approximate progressively the user's current actions towards one of the models. The user's activity is represented by the field of black dots, which are propagating throughout the whole model. The cognitive model itself is depicted as a horizontal line divided into eleven segments, represented by black rhombuses and dashed squares. Each black rhombus corresponds to the best match of user's input data that has to be detected in order to reveal a model. The dashed squares are the fuzzy fields, which mark the limits for statistical variables. Each data input that falls within the limits of a fuzzy field is evaluated by the algorithm as matched statistical instance (hollowed dots). Closer is the value of an input data towards a value of a model (black rhombus), stronger is the approximation of user's actions for further model estimation. All statistical instances outside of fuzzy fields are the samples of user's actions that are not taken into account by the algorithm.

As we mentioned earlier, the rules of the proactive system will be engaged in detection of an input data, that is, to compare each segment (mini pattern) of a model with the currently propagating path of user's activity. The process has a progressive character, meaning once the approximation has reached a certain percent of progression, it cannot drop below its current level. Furthermore, during model progression the algorithm is set to estimate the probability that a user with his/her current actions is moving towards a closest threshold before it is reached. Here we apply the Bayesian statistics, where given the evidence of already happened user's actions, and some prior data we estimate the probability if the model's data will keep matching the user's data until the end of the active model. If the algorithm estimates positively the user's future actions, the system will launch proactive actions corresponding to the needs of a current situation before the user reaches a threshold, and thus allowing to respond to a potential problem, or to provide a proactive suggestion. Such calculations are performed for each threshold until the end of current model.

The given above example demonstrates how the proactive system is capable to listen to a real-time empirical data through various models of user's cognitive states. Subsequently, the system launches the proactive actions by calculating a probability of occurrence of a modelled event. Due to the fact that all models are the representations of user's cognitive processes on an algorithmic level, the issued data can be used for various local calculations or for an outward data inference and the user's context interpretation. Therefore, the capabilities of proactive system allow to monitor the outside world and to interact intuitively with it through various types of proactive actions. The types of proactive behaviour may vary depending on a target environment, and the goals of a system.

III. Proactive system behaviour

The first important prerequisite for a proactive system is the type of environment where it is employed. This factor depends entirely on the type of a target system. Therefore, in order for proactive system to be aware of different aspects of a context, it has to possess, or to be able to infer along the way an information, which specifies the dispositions about a context and its actors. The second important prerequisite is the necessity for all information to be arranged, allocated and assembled coherently, in order to respond effectively to an event. Our role is to develop proactive scenarios, which take this aspect into account, and thus, represent the potential solutions or the way towards solutions for a particular problem or a context event. It is important to mention, that all proactive scenarios are always associated to the models, meaning that they are the potential solutions to a simulated problem represented by a cognitive model. While working on any proactive scenario, we always take into consideration the target environment, characteristics of user's cognitive activity, his/her objectives, and some related potential problems. In the end, all mentioned above factors allow us to design the numerous types of proactive scenarios.

In general, all scenarios are designed around the different aspects of user's context and his/her cognitive activities. In the present work in progress, we chose the medical search engine as the target environment, where the main type of user's activity is the querying of medical terminology. Here we show some of the possible scenarios, which could occur during user's online activity. In case of a scenario when a user has no prior experience in querying the medical terminology, or due to the different other reasons, he or she may encounter several problems during the search attempts. As an example, such problems could be specified as following: the results may have a low relevance because he or she uses very general terms, the list of relevant results may be too small due to the user's lack of knowledge of an exact terminology, or user may accidentally overlook the relevant results because a list is too big or he/she is inattentive or distracted. The proactive scenarios are designed around the aforementioned situations with regard to the user's search objectives, and the specifics of his/her cognitive activities.

By taking into consideration the user's overt behaviour, and the details of a context situation we build our hypotheses about user's current and future probable cognitive states. Such objective is based on the idea of cognitive reactions on a particular situation or an event. We understand the user's activity as continuous interrelations between his inner cognitive processes and his outer context actions, where the one is the reflection of another. By monitoring the user's current context actions, we may infer the certain aspects of his/her cognitive state, and therefore, the future possible actions. For the purposes of the present study, we take into consideration two types of users, that is, an experienced user and an inexperienced user in relation to the activity of querying a medical terminology.

We distinguish the type of a user as experienced (*user A*), or inexperienced (*user B*) due to the possibility that the search engine can be accessed either by a patient or a doctor. For the purpose of an example, we will show the specifics of cognitive modelling approach for each type of user. Thus, by working on a model, which involves the user type *A*, we have to take several things into consideration. Primarily, we define the general characteristics about a user type. For the user *A* they are as following: his or her means to obtain the objectives are methodical and logical, the results of his or her querying most probably will reflect the initial objectives, his or her context actions are consistent. Such disposition provides us with the stable ground for developing the proactive scenarios for this type of user. The characteristics of user *A* facilitate for us the development of a target model because the user's objectives are easy to detect as they are clear for the user himself. As an opposite example, in order to develop a cognitive model of the user type *B*, we define in the same manner the characteristics, which most probably will reflect the user's qualities. Such characteristics can be represented as following: the user's means to obtain an objective are chaotic, the results of his/her searching most probably will go astray from the initial objectives. Therefore, in order to help a user in his/her online activity we have to detect the certain consistency in his or her context actions. In comparison to the user *A*, it is more difficult for us to create a cognitive model of the user *B*, and subsequently to detect his or her search objectives as a user himself doesn't have a clear strategy. This may be observed in user's context actions, which are the reflection of the cognitive activity. It means if actions are inconsistent, so are the user's initial objectives. This is the primary reason why it is important for us to detect at least the smallest consistency in user's actions in order to do the further inference of his/her search objectives.

To make the aforementioned theoretical perspective clear, we present a practical example of its application for user *B*. In order to detect a model, that is, the type of a user, and his/her search objectives, the system has to identify all the information related to the user first. It means, the system has to approximate positively a model data towards the empirical user data. Therefore, on the search page the system has to collect the different statistics about a user, such as the user's query selections, keyword inputs, time the user spends on each selection, time intervals between selections, time of a query submission and so on. Such information will allow the algorithm to make the first hypothesis about a user type and his/her search intentions. The next step is the monitoring of user's behaviour on the result's page. Here the system performs the statistical data collection as in case of search page. More specifically, we detect the selected result's items, mouse hovering time over the result list, skipped and subsequently selected result's items, total time of viewing the result list and so on. Ultimately, all collected statistics will allow the algorithm to make an inference about the user's objectives, and thus, to specify the potential issues, which can take place during the user's later query attempts. For instance, if the algorithm collects enough of evidence, and starts to detect the approximate model of user's dissatisfaction with the given results, it will launch according proactive scenario to deal with a current situation. Subsequently, the proactive rules will extract the more relevant results, basing on the history of user's previous choices, and will present it to a user before he or she finishes the search. It is important to mention that all calculations are done within the seconds, meaning that system may provide proactive suggestions to a user, most probably even before he fully realises the potentially stumbling situation himself. Throughout the whole process, the algorithm will continue to approximate progressively a model towards the user data. Simultaneously, it will keep evaluating the probability of user's future cognitive states in order to detect in time the moments, where the proactive actions may be needed.

In the end, the statistical and probabilistic analysis of data, which represents the aspects of user's cognitive activity, will allow the system to interact intuitively with the outside world, and to act proactively with regard to the user's needs.

IV. Conclusions

In the current work in progress, we show the methodology, which aims to validate through theoretical and empirical study the probabilistic side of proactive computing. We describe the procedures of experiments, which aim to employ the cognitive modelling as a key approach for implementing the probabilistic model evaluation. Such disposition in our opinion allows to define the mechanisms of the probabilistic data inference based on a cognitive model approximation. In addition, such approach allows to reflect accurately the specifics of a proactive system behaviour. The elements of the aforementioned theoretical module, constitute an effective framework for the ongoing system development.

Our future work will consist of practical implementations, and empirical data analysis issued from experiments. We envision to present the upcoming results, their analysis, and related technical descriptions in the next publication in one of the associated scientific journals.

References

[1] T. Timpka, "Proactive health computing," *Artificial Intelligence in Medicine*, vol. 23, no. 1, pp. 13–24, 8 2001.

[2] G. Vlahakis and D. Apostolou, "Proactivity in service based applications," *Informatics (PCI), 2012 16th Panhellenic Conference on*, pp. 62–67, 2012.

[3] D. Tennenhouse, "Proactive computing," *Communications of the ACM*, vol. 43, no. 5, pp. 43–50, May 2000.

[4] R. Want, T. Pering, and D. Tennenhouse, "Comparing autonomic and proactive computing," *IBM Systems Journal*, vol. 42, no. 1, pp. 129–135, 2003.

[5] M. G. C. A. Cimino, B. Lazzerini, F. Marcelloni, and A. Ciaramella, "An adaptive rule-based approach for managing situation-awareness," *Expert Systems with Applications*, vol. 39, no. 12, pp. 10 796–10 811, 9 2012.

[6] J. Hong, E.-H. Suh, J. Kim, and S. Kim, "Context-aware system for proactive personalized service based on context history," *Expert Systems with Applications*, vol. 36, no. 4, pp. 7448–7457, 5 2009.

[7] R. Helaoui, "Towards a proactive system based on activity recognition," *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2010 8th IEEE International Conference on*, pp. 849–850, 2010.

[8] S. Coronado and D. Zampunieris, "Continuous Proactivity in Learning Management Systems," in *IEEE EDUCON 2010 Conference*, 2010, pp. 199–204.

[9] D. Shirnin, S. Reis, and D. Zampunieris, "Experimentation of Proactive Computing in Context Aware Systems: Case Study of Human-Computer Interactions in e-Learning Environment," in *Cognitive Methods in Situation Awareness and Decision Support (CogSIMA), 2013 IEEE International Multi-Disciplinary Conference on*, San Diego, US, February 2013, pp. 269–276.

[10] B. Emond and R. L. West, "Cyberpsychology: a human-interaction perspective based on cognitive modeling." *Cyberpsychology & Behavior: The Impact Of The Internet, Multimedia And Virtual Reality On Behavior And Society*, vol. 6, no. 5, pp. 527–536, 2003.

[11] R. Sun, "Theoretical status of computational cognitive modeling," *Cognitive Systems Research*, vol. 10, no. 2, pp. 124–140, 6 2009.

[12] J. R. Anderson, D. Bothell, M. D. Byrne, S. Douglass, C. Lebiere, and Y. L. Qin, "An integrated theory of the mind," *Psychological review*, vol. 111, no. 4, pp. 1036–1060, 2004.

[13] T. L. Griffiths, N. Chater, C. Kemp, A. Perfors, and J. B. Tenenbaum, "Probabilistic models of cognition: exploring representations and inductive biases," *Trends in Cognitive Sciences*, vol. 14, no. 8, pp. 357–364, 8 2010.

[14] B. R. Newell and A. Bröder, "Cognitive processes, models and metaphors in decision research," *Judgment and Decision Making*, vol. 3, no. 3, pp. 195–204, 2008.

[15] D. E. Meyer and D. E. Kieras, "A computational theory of executive cognitive processes and multiple-task performance: Part 1. basic mechanisms," *Psychological review*, vol. 104, no. 1, p. 3, 1997.

[16] Z. Duric, W. D. Gray, R. Heishman, F. Li, A. Rosenfeld, M. J. Schoelles, C. Schunn, and H. Wechsler, "Integrating perceptual and cognitive modeling for adaptive and intelligent human-computer interaction," *Proceedings of the IEEE*, vol. 90, no. 7, pp. 1272–1289, Jul 2002.

[17] M. D. Lee, "How cognitive modeling can benefit from hierarchical bayesian models," *Journal of Mathematical Psychology*, vol. 55, no. 1, pp. 1–7, 2 2011.

[18] S. Reis, D. Shirnin, and D. Zampunieris, "Design of proactive scenarios and rules for enhanced e-learning," in *CSEDU 2012 - Proceedings of the 4th International Conference on Computer Supported Education*. Porto, Portugal: SciTePress, 2012, pp. 253–258.

[19] D. Zampunieris, "Implementation of a Proactive Learning Management System," in *E-Learn World Conference on E-Learning in Corporate, Government, Healthcare and Higher Education*, Hawaii, 2006, pp. 3145–3151.

[20] D. Zampunieris, "Implementation of Efficient Proactive Computing Using Lazy Evaluation in a Learning Management System," *International Journal of Web-Based Learning and Teaching Technologies*, vol. 3, no. 1, pp. 103–109, 2008.