Towards a Cognitive-Based Approach to Distributed Requirement Elicitation Processes

Gabriela N. Aranda[1], Aurora Vizcaíno[2], Alejandra Cechich[1], Mario Piattini[2]

[1] Universidad Nacional del Comahue,
Departamento de Ciencias de la Computación,
Buenos Aires 1400, 8300 Neuquén, Argentina
{garanda, acechich)@uncoma.edu.ar

[2] Universidad de Castilla-La Mancha,
Departamento de Informática,
Paseo de la Universidad 4, 13071 Ciudad Real, España
{Aurora.Vizcaino, Mario.Piattini}@uclm.es

Abstract. *Distance between stakeholders working on a global software development project introduces problems in communication and control. Processes like requirements elicitation, where communication is crucial, have to be specially rethought in order to deal with these problems. As the requirement elicitation is a human-centred process, we propose using techniques from the field of cognitive psychology to define a new approach for it. Our goal is reducing problems in communication by proposing the most suitable elicitation techniques according to stakeholders' preferences. In this paper, we introduce our approach and illustrate how cognitive styles might be used to improve elicitation.*

1 Introduction

Global Software Development (GSD) is the development of software in scenarios where stakeholders are in many geographically distanced sites. This practise, that increases day by day, is done by virtual communities or teams that communicate using some kind of technology. Well-known are the problems during traditional requirement elicitation processes that have been extensively analysed in literature [8, 14]. Most of these problems are concerned with communication between stakeholders [18] which is critical during requirements elicitation. This communication becomes even more difficult when participants are distributed, since other issues, like time difference and cultural diversity also affect it [7].

There are two research areas that try to find solutions to such communication problems. One of them is Computer-Supported Cooperative Work (CSCW) which focuses on providing technologies to enable communication, as well as studying human behaviour when working in group. The other one is Cognitive Informatics, an interdisciplinary area that combines many science and engineering disciplines, such as informatics, computing, software engineering, and cognitive sciences [6, 20].

Since our main goal is to enhance interpersonal communication in geographically distributed teams, concepts from both areas come together. On the one hand, since people are distributed along many distanced sites, they must communicate each other using software, called *groupware*[1], which is part of the studies on CSCW. Some examples of groupware tools used during multi-site developments are e-mails, newsgroups, mailing lists, forums, bulletin boards, shared whiteboards, document sharing, chat, instant messaging, and videoconferencing [7, 13]. On the other hand, communication between people involves aspects of human processing mechanisms that are analyzed by the cognitive sciences. In our proposal, we are particularly interested in using some techniques from the field of psychology, called Learning Style Models, which may be useful to select groupware tools and elicitation techniques according to the cognitive styles of stakeholders.

Having this in mind, in section 2 we present basic concepts of cognitive informatics and learning styles models. In section 3 we introduce a general model of elicitation activities and our extension of such a model to use it in distributed environments, as well as a motivating example. The last sections of the paper present some related works and address some conclusions.

2 Cognitive Informatics and Learning Style Models

Cognitive Informatics relates cognitive sciences and informatics in a bi-directional way: (1) using computing techniques to investigate cognitive science problems like memory, learning, and thinking; and (2) using cognitive theories to investigate informatics, computing, and software engineering problems [20]. Our work is related to this second point of view, using concepts from cognitive psychology (concerning the way people attend and gain information and how these information processing mechanisms affect human behaviour [6]), to improve the requirement elicitation process.

---

[1] Since there is no total agreement whether groupware refers to simple technologies or more complicated systems, we prefer referring to every simple piece of communication technology as a groupware tool, and to the systems that combine them as groupware packages.

Part of cognitive psychology theories are cognitive styles, which are based on Jung's theory of psychological types published in 1921. Jung's theory classifies people preferences about perception, judgment and processing of information [16]. This classification has been used to analyse and understand differences in human behaviour. As an extension of it, different instruments have been designed to measure human characteristics and explain their differences.

Like cognitive styles, learning styles models (LSMs) are based on such observations. They classify people according to a set of behavioural characteristics pertaining to the ways they receive and process information and are used to improve the way people learn a given task.

So far, LSMs have been discussed in the context of analysing relationships between instructors and students. We take advantage of this kind of models and adapt it to virtual teams that deal with distributed elicitation processes. To do so, we consider an analogy between stakeholders and roles in LSMs, since during the elicitation process everybody "learns" from others. In this way stakeholders play the role of student or instructor alternatively, depending on the moment or the task they are carrying out [15]. Hickey and Davis [12] mention that elicitation is about learning the needs of the users. From our point of view, users and clients also learn from analysts and developers. For instance, they learn how to use a software prototype or a new vocabulary, etc.

After analysing five LSMs [15] we found out that the model proposed by Felder-Silverman [9] was wide enough to build a complete reference framework choosing it as a foundation.

The Felder-Silverman (F-S) Model classifies people into four categories, each of them further decomposed into two subcategories as follows: Sensing / Intuitive; Visual / Verbal; Active / Reflective; Sequential / Global.

Characteristics of each subcategory are:

- *Sensing people* prefer learning facts. They like solving problems by well-established methods and dislike complications and surprises. Sensors tend to be patient with details and good at memorising facts and doing hands-on (laboratory) work.
- *Intuitive people* often prefer discovering possibilities and relationships. They like innovation and dislike repetition. They tend to work faster and to be more innovative than sensors. Intuitors do not like work that involves a lot of memorisation and routine calculations.

- *Visual people* remember best what they see (such as pictures, diagrams, flow charts, time lines, films, and demonstrations). They prefer visually presented information.
- *Verbal people* get more out of words, and written and spoken explanations. They prefer verbally presented information.
- *Active people* tend to retain and understand information by doing something active with it (discussing or applying it or explaining it to others). “Let’s try it out and see how it works” is an Active’s phrase.
- *Reflective people* prefer to think about information quietly first. “Let’s think it through first” is the Reflective’s response.
- *Sequential people* tend to gain understanding in linear steps, with each step following logically from the previous one. They tend to follow logical stepwise paths in finding solutions. They may not fully understand the material but they can nevertheless do something with it (like solve homework problems or pass a test) since the pieces are logically connected.
- *Global people* tend to work in large jumps, absorbing material almost randomly without seeing connections, and then suddenly "getting it". They may be able to solve complex problems quickly or put things together in novel ways once they have grasped the big picture, but they may have difficulty explaining how they did it.

People are classified by a multiple-choice test (available on the WWW[2]) that gives them a rank for each subcategory. Depending on the circumstances people may fit into one category or the other, being for instance, sometimes active and sometimes reflective; so preference for one category is measured as *strong*, *moderate*, or *mild*. Only when there is a strong preference, a person can be classified as a member of a certain group.

3 The Selection Process of Requirements Elicitation Techniques

Requirement elicitation techniques are used to capture and understand the needs of clients and end-users in order to build systems that fit their expectations. Since techniques employed during the requirement elicitation process influence the quality of the requirements, it is possible to improve the success of the products by improving the way techniques are selected [12].

[2] http://www.engr.ncsu.edu/learningstyles/ilsweb.html

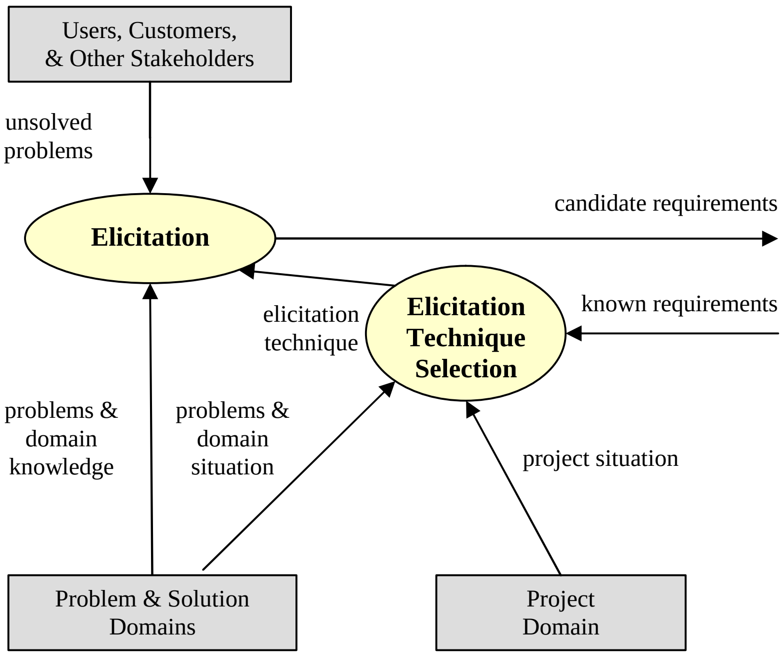

Figure 1: Details of elicitation activities (Hickey and Davis, 2002)

Hickey and Davis have proposed a general model for elicitation activities [11] which is shown in figure 1. This model represents a generalization for all requirements elicitation methodologies and techniques, and highlights the role that knowledge plays in the selection process.

The model focuses on the activities that analysts do to reach a common understanding of the requirements, which derives in an iterative process of selection and application of elicitation techniques. Each iteration *i* of elicitation is defined by the formula:

$$elicit_i (R_i, S_i, t_i) \text{ ® } R_{i+1}, S_{i+1}$$

where t_i is the elicitation technique that is applied in step *i* when R_i is the current state of knowledge about requirements and S_i is the current situation. After applying t_i, a new state of knowledge R_{i+1} and a new situation exist.

Then, the elicitation technique selection process, is modelled as a selector function *s*:

s (R_i, S_i, *c*(*T*)) ® {$t \in T$ | *t is applicable in situation S_i when the current state of requirements is R_i*}

In that way, the selector function *s* identifies the best possible techniques given a current state of knowledge and a particular situation.

The model also defines a personal selector function ***p***, that analyse personal preferences of analysts:

p ({*t*}, *P*) ® $t_i \in$ {*t*}

That means, given a particular situation and a particular knowledge state about requirements, the selector function σ will suggest a set of best techniques {*t*}. Then, the selector *p* will look for an intersection between the techniques suggested for σ and the analyst's personal preferences (*P*).

The generality of this model gives us the possibility of redefining it to express the aspects that distinguish the requirements elicitation process in virtual teams, and also the cognitive aspects expressed in our model as learning styles preferences.

As a first attempt [3] we have proposed an extension of this notation by means of analysing the preferences of each stakeholder from a cognitive point of view, and choosing the one that have more adherents We assumed that the selected option is suitable for the project itself too (for example, by analyzing relationships between elicitation and particular life cycles). The proposed extension of π function, called π*, has been defined as follows:

*p** ({*t*}, {PS_1}, {PS_2}, …, {PS_n}) ® $t_i \in$ {*t*}

where PS_i represents a set of techniques that fit the *i-th* stakeholder's preferences (which have been defined by mechanisms based on fuzzy logic and fuzzy sets and described in [1, 2, 3]), and $t_i \in$ {*t*} is the technique that appears in most of the PS_i.

Analysing the function *p** in an environment where stakeholders preferences are well-known, we found out that it is important to take into account the relative importance of stakeholders' preferences. A necessary improvement of the previous notation is weighting stakeholders' preferences. For instance, if some stakeholders' preferences are strong and the rest of the stakeholders are moderate or mild, the preferences that should be primarily considered are those of the first group of stakeholders. Then, different weights might differentiate the type of preferences we have, but also might be used to prioritize them according to different stakeholder's roles. In order to do so, we extend the function π*, called π**, as follows:

$$p^{**}(\{t\}, (\{PS_1\}, ws_1), (\{PS_2\}, ws_2), \ldots, (\{PS_n\}, ws_n)) \rightarrow t_i \in \{t\} \wedge t_i \in \{PS_j\} \wedge ws_j = \max(ws_1, ws_2, \ldots, ws_n)$$

where PS_i represents a set of techniques that fit the *i-th* stakeholder's preferences and ws_i is the weight –that means how strong its preferences are-, and the resulting t_i is a technique that is appropriate for the current situation and it is also appropriate for the stakeholder whose personal preferences are the strongest.

When analysing the selector function σ from a point of view where stakeholders are distributed along many geographically distanced sites, we think that new considerations have to be done.

Aspects that may affect the selection of requirement elicitation techniques in such a virtual environment are, for instance, time difference and the level of knowledge of a common language. When the time difference is wide enough to not allow synchronic communication, or do it for very short periods of time, the selection process should prioritize those techniques that work better on asynchronous basis. Similarly, when stakeholders do not share the same language, some of them would need more time to read, think quietly, look for some vocabulary in the dictionary, etc., so the selector function should also prioritize asynchronous-based and may be graphical-based requirement elicitation techniques.

With those considerations in mind, we extend the selector function σ as follows:

$$s^*(R_i, S_i, c(T), T_i, L_i) \rightarrow \{t \in T \mid t \textit{ is applicable in situation } S_i \textit{ when the current state of requirements is } R_i\textit{, according to restrictions } T_i \textit{ and } L_i\}$$

where T_i (time difference) indicates in which degree synchronous communication is possible between the sites that needs to interact; and L_i (degree of knowledge of a common language) indicates the level of fluency of communication.

We must note that by using the term "time difference" (T_i) we refer not to real time difference between two sites, but the overlap referring to work time.

To express the values T_i and L_i we suggest using a fuzzy logic approach so as to count with scales like:

T_i = *{no-overlap, little-overlap, half-overlap, much-overlap, full-overlap}*

L_i = *{low, low-intermediate, intermediate, high- intermediate, high}*

A graphical representation of the requirement elicitation process in a distributed environment is shown in Figure 2.

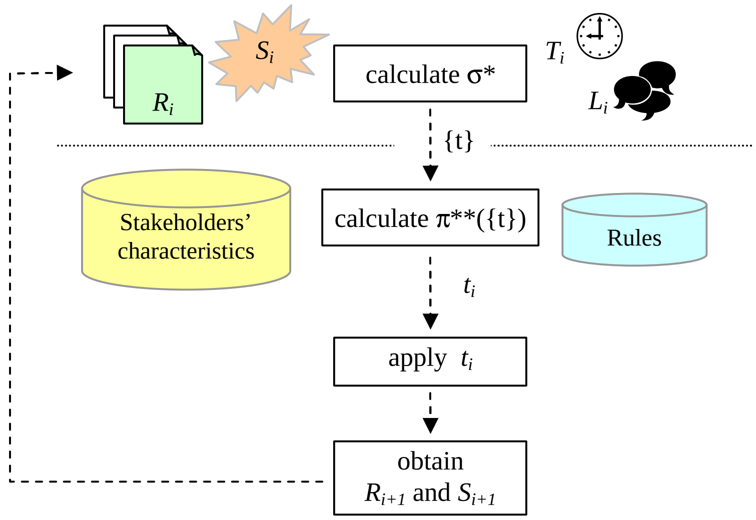

Figure 2: Requirement elicitation in distributed environments, as an iterative process of techniques selection and application.

4 Discussion: a motivating example

As we have done in [1], we prefer expressing personal preferences, gathered by Felder-Soloman test, by using the adverbs (and its correspondent abbreviations): Very (V), Moderately (M) and Slightly (S). They correspond to *strong*, *moderate* and *mild*, respectively, in the F-S model, but we have changed them to avoid confusion respect to the use of the first letter and have a more compact notation. For instance, the preferences for the category Sensing-Intuitive are: Very sensing (VSe), Moderately sensing (MSe), Slightly sensing (SSe), Slightly intuitive (SIn), Moderately intuitive (MIn), and Very intuitive (VIn). In such a way, information about a person is modelled as a quadruple, which represents his/her preferences about categories Active-Reflective, Sensing-Intuitive, Visual-Verbal and Sequential-Global.

Let us consider the results of the test applied to three stakeholders:

S_1 = (*MAc, SSe, MVe, SGl*); S_2 = (*SRe, VSe, VVi, VSq*); S_3 = (*SRe, SSe, SVe, SGl*)

Let us suppose that, by the application of the rules we have previously determined according to personal preferences, we obtain a set of "appropriate" elicitation techniques for each one E_1, E_2, and E_3.

Assuming that S_1 is the analyst, S_2 and S_3 are users, and {t} is the set of the appropriate requirement elicitation techniques obtained by the function σ*, three scenarios are possible:

Case 1: *Application of the personal function selector p.*

In this case, only the personal preferences of the analyst (S_1) would be taken into account. Suppose that S_1 prefers using some technique mainly based on words. That would make stakeholder S_2, who has a strong preference for the visual category, not feeling comfortable and therefore not committed enough with the collaborative task.

Case 2: *Application of personal function selector p**

In this case, techniques would be chosen from those that are repeated in most of the sets. Taking into account that S_1 and S_3 have some moderated and slight preferences for the verbal category, this would give the chance of a situation like the one described in case 1, where stakeholder S_2, who has a strong preference for the visual category, would not be comfortable.

Case 3: *Application of personal function selector p***

Since stakeholder S_2 has the strongest preferences, techniques would be chosen from the set E_2. Considering that preferences of stakeholders S_1 and S_3 are moderated or slight, the selection of techniques would not affect negatively their performance, but could improve significantly S_2's.

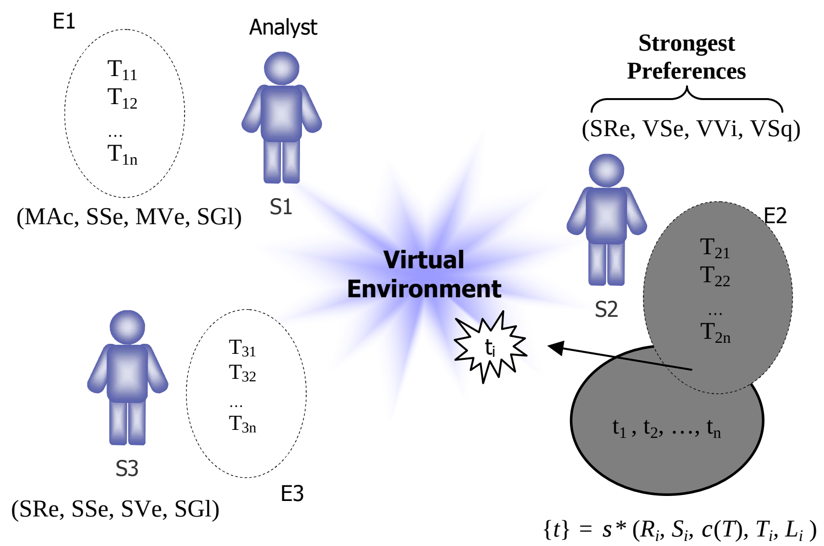

Figure 3: A selection of requirement elicitation techniques driven by personal preferences, based on learning styles of all stakeholders (Case 3)

5 Related works

Most of related works using learning styles in informatics concern educational purpose, such us its influence when learning recursion [21], End-User Computing tools [5], or programming [19], and also to define frameworks for designing multimedia for computer science courses [4] or web-based courses in general [17].

On the contrary, few related works use psychological techniques to solve problems in Software Engineering. One of them is the psychotherapeutic approach known as Neuro-Linguistic Programming (NLP), which has been applied by the Sophist group in requirements elicitation. They have developed a set of rules for analysing requirements linguistically, which can be applied within interviews, when writing requirements, and for checking written requirements. The main aspects these rules can help avoid are the underspecified process words (deletions); the use of universal and usually problematic quantors – all, each, never (generalizations) –; and finally the nominalizations (distortions) [10].

Another work on this direction is the use of cognitive styles as a mechanism for software inspection team construction [16]. This work describes an experiment that aims at proving that heterogeneous software inspection teams have better performance than homogeneous ones. The heterogeneity concept is analyzed according to the cognitive style of participants. During the experiment, 33 software inspectors were asked to analyse some code. Then they were classified using the MBTI method, an instrument similar to Felder and Silverman Model. Later, the quantity of detected errors and their types were compared individually and arranged into groups of different sizes and homogeneity/heterogeneity conditions (homogeneity = same cognitive style vs. heterogeneity = many different cognitive styles). Authors conclude that heterogeneous groups worked better than homogeneous ones.

Even when they also used the concept of cognitive styles to classify people, our approach is not the same. As we have explained previously, we aim at choosing the best strategies to improve communication for a given group of people, which means we do not try to set which people seem to be more suitable to work together, but to give the best solution for an already chosen group of people.

6 Conclusions

Virtual teams became a common way of developing software. To save costs, many organisations have adopted a distributed structure where members communicate through groupware tools.

The selection of appropriate technology and elicitation techniques in such environments is a subject of research, since when stakeholders feel comfortable with the technology and methodologies they use, information gathered during elicitation is expected to be more accurate. Stakeholders might feel more comfortable expressing their ideas and describing facts when using a tool closer to the way they perceive and reason about the world.

In this paper we have extended the Hickey and Davis model of requirements elicitation technique selection process, by adding features of distributed environments and knowledge about stakeholders' preferences when perceiving and processing information. Techniques selected by applying this model would improve the elicitation process in distributed teams.

We are aware that future work is needed to solve conflicts when stakeholders' preferences seem to be opposite. Also, because of the necessity of empirical results to validate our approach, we are planning an experiment that will involve computer science students from three different countries.

7 References

1. Aranda, G., Cechich, A., Vizcaíno, A., and Castro-Schez, J.J. Using fuzzy sets to analyse personal preferences on groupware tools. In X Congreso Argentino de Ciencias de la Computación, CACIC 2004. San Justo, Argentina, October (2004) 549-560.
2. Aranda, G., Vizcaíno, A., Cechich, A., and Piattini, M. A Cognitive Perspective for Choosing Groupware Tools and Elicitation Techniques in Virtual Teams. In International Conference on Computational Science and its Applications (ICCSA 2005). Singapore: © Springer-Verlag Berlin Heidelberg, May (2005) 1064-1074.
3. Aranda, G., Vizcaíno, A., Cechich, A., and Piattini, M. A Cognitive-Based Approach to Improve Distributed Requirement Elicitation Processes. Accepted in 4th IEEE International Conference on Cognitive Informatics (ICCI'05). Irvine, USA, August (2005).
4. Blank, G.D., Roy, S., Sahasrabudhe, S., Pottenger, W.M., and Kessler, G.D., Adapting Multimedia for Diverse Student Learning Styles. The Journal of Computing in Small Colleges, 18(3), (2003) 45-58.
5. Bostrom, R.P., Olfman, L., and Sein, M.K. The Importance of Individual Differences in End-User Training: The Case for Learning Style. In 1988 ACM SIGCPR Conference. Maryland, April (1988) 133-141.
6. Chiew, V. and Wang, Y. From Cognitive Psychology to Cognitive Informatics. In Second IEEE International Conference on Cognitive Informatics, ICCI'03. London, UK, August (2003) 114-120.
7. Damian, D. and Zowghi, D. The impact of stakeholders geographical distribution on managing requirements in a multi-site organization. In IEEE Joint International Conference on Requirements Engineering, RE'02. Essen, Germany, September (2002) 319-328.

8. Davis, A., Software Requirements: Objects, Functions and States. Upper Saddle River. New Jersey: Prentice Hall, (1993).
9. Felder, R. and Silverman, L., Learning and Teaching Styles in Engineering Education. Engineering Education, 78(7), (1988) 674-681.
10. Goetz, R. and Rupp, C. Psychotherapy for System Requirements. In Second IEEE International Conference on Cognitive Informatics (ICCI'03). London, England, August (2003) 75-80.
11. Hickey, A.M. and Davis, A. Requirements Elicitation and Elicitation Technique Selection: A Model for Two Knowledge-Intensive Software Development Processes. In 36th Annual Hawaii International Conference on Systems Sciences (HICSS), January (2003) 96-105.
12. Hickey, A.M. and Davis, A.M. Elicitation Technique Selection: How do experts do it? In International Joint Conference on Requirements Engineering (RE03). Los Alamitos, California: IEEE Computer Society Press, September (2003) 169-178.
13. Lloyd, W., Rosson, M.B., and Arthur, J. Effectiveness of Elicitation Techniques in Distributed Requirements Engineering. In 10th Anniversary IEEE Joint International Conference on Requirements Engineering, RE'02. Essen, Germany, September (2002) 311-318.
14. Loucopoulos, P. and Karakostas, V., System Requirements Engineering. International series in Software Engineering, Mc Graw-Hill. New York, NY, USA, (1995).
15. Martín, A., Martinez, C., Martinez, N., Aranda, G., and Cechich, A. Classifying Groupware Tools to Improve Communication in Geographically Distributed Elicitation. In IX Congreso Argentino de Ciencias de la Computación, CACIC 2003. La Plata, Argentina, October (2003) 942-953.
16. Miller, J. and Yin, Z., A Cognitive-Based Mechanism for Constructing Software Inspection Teams. IEEE Transactions on Software Engineering, 30(11), (2004) 811-825.
17. Moallem, M. The Implications of Research Literature on Learning Styles for the Design and Development of a Web-Based Course. In International Conference on Computers in Education, ICCE 2002. Auckland, New Zealand, December (2002) 71-74.
18. SWEBOK, Guide to the Software Engineering Body of Knowledge, Software Engineering Coordinating Committee IEEE Computer Society, (2004).
19. Thomas, L., Ratcliffe, M., Woodbury, J., and Jarman, E. Learning styles and performance in the introductory programming sequence. In 33rd SIGCSE Technical Symposium on Computer Science Education. Cincinnati, Kentucky, USA, February (2002) 33-37.
20. Wang, Y. On Cognitive Informatics. In First IEEE International Conference on Cognitive Informatics, ICCI'02. Calgary, Alberta, Canada, August (2002) 34-42.
21. Wu, C.C., Dale, N.B., and Bethel, L.J. Conceptual Models and Cognitive Learning Styles in Teaching Recursion. In Twenty-ninth SIGCSE Technical Symposium on Computer Science Education. Atlanta, Georgia, United States, February (1998) 292 - 296.