Volume 4, No. 3, March 2013
Journal of Global Research in Computer Science
ISSN-2229-371X
REVIEW ARTICLE
Available Online at www.jgrcs.info

A REVIEW OF THE COGNITIVE INFORMATION RETRIEVAL CONCEPT, PROCESS AND TECHNIQUES

Parul Kalra Bhatia[*1], Chetna Choudhary[2], Dr. Deepti Mehtrotra[3] and Dr Abdul Wahid [4]
*[1]Department of IT, Amity University, Noida, Uttar Pradesh, India
pkbhatia@amity.edu [1]
[2]Department of IT, Amity University, Noida, Uttar Pradesh, India
cchoudhary@amity.edu[2]
[3]Director, ASCS, I- Block, Sector-44, Amity University, Noida, Uttar Pradesh, India
dmehrota@amity.edu[3]
[4] HOD CS &IT, MAMUU, Central University, Hyderabad, Andhra Pradesh, India
wahidabdul76@yahoo.com[4]

Abstract: The word "cognitive" refers to the thought process toward awareness or knowledge. In terms of Cognitive Science, it provides bridge between information processing, conceptualization of the resources, perceptual skills and topics related to the cognitive psychology. By retrieving information that based on cognitive concepts, process and techniques one can represent the current user's information need, their problem state and domain work or area of interest in the outline of structure and casualties. This poly-representational approach leads to cognitive process which is multitasking in the way of perception, attention, interpretation, understanding and remembrance of human behaviour interaction. With the help of implementation techniques of relevance feedback which validate and provide reliability metrics to calculate user behaviour using knowledge domain visualization, Training frameworks provide users how to proceed in searching and retrieving information.

Keywords - Cognitive Information Retrieval, Human Behavior Interaction, Multitasking, Poly-representation.

COGNITIVE SCIENCE

Cognitive Science is the multidisciplinary scientific research area where neuroscience, artificial Intelligence, philosophy, psychology, linguistics, anthropology, sociology, and education [1, 16] emphasis on human brain emotions, intelligence and behavior and focuses on the information. It represents the processed information in terms of perception, language and memory reasoning within the human nervous system and machines like computers, robots, inference engine. It helps in analyzing from low level learning to high level decision making and planning.

COGNITIVE INFORMATICS

Cognitive Informatics (CI) inherited form the field of cognitive science and informatics. CI provides the conceptual theories and mathematical computations which lays the foundation for knowledge based science and engineering like mechanical, electrical, electronics and computer engineering [3]. It forms computational systems which are capable enough to unify the processed information and patterns that explain the structure and organization of communicative information. CI objective is to implement engineering solutions like text, data mining and web, communicative robots, multimedia, multimodal interactive systems, real time or virtual environment for the purpose of distributed collaborative work [14]

COGNITIVE INFORMATION RETRIEVAL (CIR)

CIR is an interdisciplinary area of study that includes research from information science to cognitive science which helps to interact with the human computer on the basis of human factors. CIR is conceptualized as complex human information related human computer interaction processes that are embedded within an individual's everyday social and life context. CIR is an important part of the human information condition and critical to the development of new approaches to the design of Web and IR systems. Presently, Information Retrieval (IR) research originates from three separate fields: information science, computer science and contributions from the field of social science. [3] However, the three communities do not really communicate with each other, and use different methodologies. The authors examine these different methodologies in terms of dependent, independent and controlled for variables. As IR serves the goals of seeking information which primarily serves the goals of the user's work task (or other interest), and the authors suggest a possible area of reconciliation in the context or task for the user's information search [3].

Emerging frameworks, models and theories are providing a more complex view of CIR that includes multitasking, relevance feedback, Human Information Behavior (HIB), longitudinal process models and visualization techniques [3] Researchers and students in the fields of information science, computer science, cognitive science, human factors and related disciplines, and scholars investigating CIR for their respective research work and researchers are thinking more

broadly about information seekers rather than just as users of IR systems and the technical problems that may arise during the use. [3]

COGNITIVE INFORMATION RETRIEVAL CONCEPT, PROCESS AND TECHNIQUES COGNITIVE INFORMATION RETRIEVAL: (CIR-CPT)

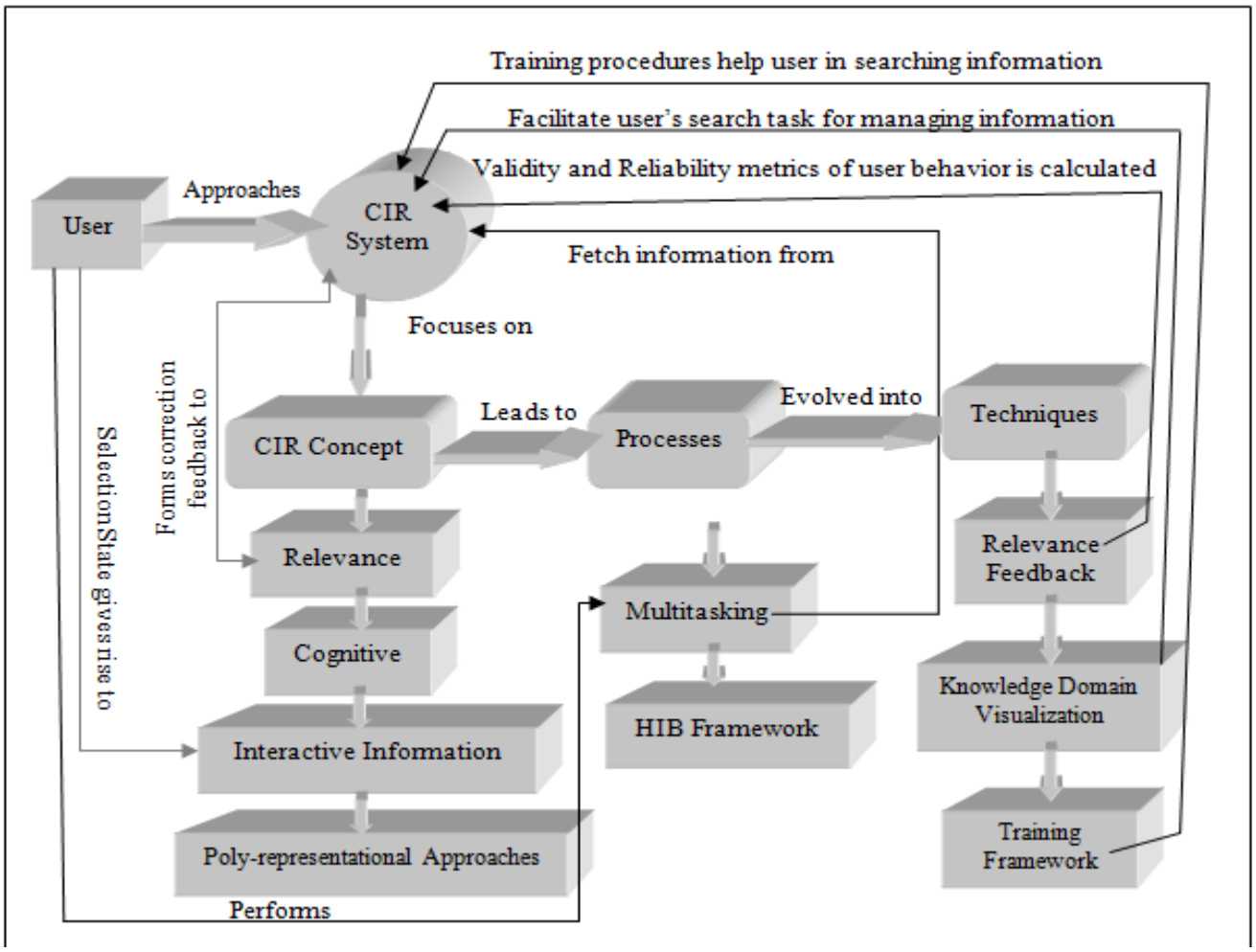

Figure 1:- A block diagram of CIR-CPT

Concept

CIR research concept based on relevance feedback with the help of user interaction with IR system. Studying the cognitive behavior of the user which provides valuable inputs in the form of documents to analyze design evaluate the complete picture of the user profile. [3]

Key concepts for conceptualizing CIR

(a). Relevance
(b). Cognitive and Interactive
(c). Poly-representational approaches.

Relevance:

The relevance concept currently forms the basis of correction feedback to the IR system, allowing the system to revise its matching algorithm so that it more accurately matches the real—as opposed to the initially expressed—information need of the user, represented in the interaction by the user's query. We currently evaluate system performance according to the system's ability to retrieve topically relevant documents; we therefore build systems for retrieving topically relevant documents. This strong view of relevance, Ruthven believes—with a single, objective reality to relevance based on topic—has been a major obstacle to the development of more naturalistic methods and systems.[3]

Cognitive and Interactive:

The cognitive aspect of users as they interact with and react to environmental stimuli coming at them forms the IR system. Within a cognitive perspective, information need as a conceptual basis for the user in the interaction is problematic. Cole, Beheshti, Leide and Large then specify the terms and definitions of the user-system interaction they wish to examine in detail. It seeks to reconceptualize information need as the

conceptual basis for user-system interaction. To this end, the interaction is represented as a series of interacting states: the user's task or problem state, and the user's cognitive state, the various need states for each task or problem that may arise during the user's interaction with the IR system. In turn, the IR system presents the user with information stimuli that have the potential to align the user's cognitive, task or information need states. The system stimulus message's role in the interaction is to stimulate the user to make in a selection state. When the user is in a Selection State, the user's knowledge structure understood and integrated an information process. [3]

Poly-representational approaches. :

Birger Larsen and Peter Ingwersen [4] describes a user and document representation methodology that gives the IR system a fuller picture of the user and the document set than current simple request-based systems allow, based on the principle of poly-representation. For the user, the poly-representation consists of the user's various concurrent information needs, emotional states, tasks, organizational constraint etc. — a multiprolonged representation of the user's cognitive space. For the document set, the poly-representation consists of the citation links, thesaurus terms, selectors (e.g., journal name, etc.), indexers terms and author's headings, captions, etc.—a multipronged representation of the document set. The poly-representations are then represented in the algorithm used by the system to match the user with the IR system database's document set. The result is a cognitive overlap—of the user's cognitive space and the information space of various configurations of document representations from various search systems. [5, 6, 3]

Cognitive Information Retrieval Process

Multitasking:

By Amanda Spink and Charles Cole [3] analyzes user multitasking information behavior while the user is interacting with an IR system. Currently, IR systems require users to search sequentially and are largely designed to support limited types of searching based on specifying queries that select documents or Web sites to fulfill a single information task. However, IR system users naturally engage in multitasking while they are accessing information from an IR system. The users may begin their IR system interaction with multiple topics, or they may begin with a single topic and then develop additional topics during the search pro-cess. Researchers define such behavior as natural, and examine possible mechanisms that enable the user to switch from one task to another while ostensibly engaged in accessing information from an IR system for one topic or task. The process of modeling multitasking within a cog-nitive IR framework, starting from Saracevic's (1997) Stratified model of user-IR system interaction. [7, 3, 8] They define the central problem as the user coordinating between various levels of problem identification, system problems, etc. for a single topic or task—on the level of searching the IR system—and the user's wider thinking while interacting with the IR system—on the level of seeking information, which involves coordinating multiple search tasks and multiple topic tasks [3,9]

Human Interaction Behavior (HIB):

Humans have used electronic information retrieval (IR) systems for more than 50 years as they evolved from experimental systems to full-scale Web search engines and digital libraries. The fields of library and information science (LIS), cognitive science, human factors and computer science have historically been the leading disciplines in conducting research that seeks to model human interaction with IR systems for all kinds of information related behaviors. As technology problems have been mastered, the theoretical and applied framework for studying human interaction with IR systems has evolved from systems-centered to more user-centered or cognitive-centered approaches. However, cognitive information retrieval (CIR) research that focuses on user interaction with IR systems is still largely under-funded and is often not included at computing and systems design oriented conferences. But CIR-focused research continues, and there are signs that some IR systems designers in academia and the Web search business are realizing that user behavior research can provide valuable insights into systems design and evaluation. [3]

CIR research is the integration of CIR within the broader human information behavior (HIB) framework which makes up the human information condition. The HIB perspective for CIR seeks to create a more holistic understanding of CIR that takes into account the HIB context in which human-IR system interaction takes place. [3, 15]

New Directions in research should also examine the interplay between information and non-information tasks, as well as beginning to conceptualize the so-called interrupted HIB behaviors users engage in during search session like "vanity searching." Such HIBs should be analyzed in terms of the role these behaviors may play in providing coordinating or switching mechanisms between individual search and seeking tasks during a multitasking search session. [3]

Cognitive Information Retrieval Techniques:

Relevance feedback

By Diane Kelly examines the techniques used by researchers to study the implicit relevance feedback behavior of users engaged in interacting with an IR system. Implicit relevance feedback techniques may offer some solution to this problem. Implicit relevance feedback techniques based on the users requirements gathering and document preferences which help in interacting with the system naturally. The author refers to her own previous work which observed user behavior then classified the behavior into five categories describing the underlying intent of the observed behavior. The resulting five categories are: examine, retain, reference, annotate and create. Kelly then reviews the implicit rele-vance feedback research to illustrate how such studies have been conducted and how feedback has typically been measured and used. The author concludes that by employing the validation and reliability checks in the form of metrics, the users actual interaction with the system can be represented (e.g. the length of time a document is displayed in the browser window is assumed by

these studies to be equivalent to how long the user reads the docu-ment) [3, 8]

Knowledge Domain Visualization (KDV):

By Peter Hook and Katy B'orner analyzes various visual representation techniques for the purpose of improving the usability of Educational Knowledge Domain Visualizations (KDVs). KDVs visualize information spaces for specific topics or a broad range of topics, indicating to the user some sort of structure in a given information space. They can be incorporated into the IR system either at the front-end, as an overview of a topic space to assist the user in formulating an access point to the system's database, or at the backend of the system in the results list. The visual representation of a topic or domain structure assists the user, especially the domain novice, in discerning the most relevant entry points into the information space, as well as facilitating user integration of the objective knowledge space represented by the KDV into the user's subjective image of it. KDVs facilitate the user's search task of organizing the information search, information access, and management of his or her interaction with the IR system, and the information found in the system's database. The authors are primarily interested in how visualizing knowledge domains can facilitate user internationalization of scholarly knowledge by taking advantage of the way human vision and human spatial cognition work. [10, 3]

Training Frameworks:

By Wendy Lucas and Heikki Topi describe techniques for training users in the information searching of IR systems. The authors set these search training techniques inside the broader perspective of information seeking. Different stages in the seeking process lead to different user cognitive processes, requiring different training models. Within each seeking stage, the user goes through a search process. Stage 1 in the Lucas and Topi model [11] of the search process is the articulation of the user's information need. Stage 2 is the conceptualization of the need as an executable query. Stage 3 is the formulation of the conceptual query for a given IR system. Stage 4 is entering the query into the search interface. And Stage 5 is the user understanding and interpreting the search results. Each stage requires different information search training procedures, which are, in turn, different, depending on the stages of the seeking processes. [11, 13] The second Lucas and Topi search training strategy is to gather user errors and then specifically train users on correcting these errors. Categories of user errors are conceptual knowledge errors, semantic knowledge errors, and technical knowledge errors (Sormunen and Pennanen, 2004). Examples of user errors from the Excite study of Spink, Wolfram, Jansen, and Saracevic (2001) are the infrequent use of Boolean operators and errors when the Boolean operators are used, such as errors in capitalization, the use of the ampersand instead of AND, etc.[12,3]

SUMMARY AND OUTLOOK

The goal of our paper is to provide an overview of CIR concepts, techniques and process. It shows how human information condition with regard to increasing needs, interacts with IR systems and this paper targets the researchers, educators, students and practitioner in the fields of information science, computer science, cognitive science, human factors and related disciplines, and scholars investigating CIR. The potentiality of cognitive concepts emerged into relevance interaction information through poly-representational approaches which are processed into multitasking human behavior. As a result, have to achieve above approaches we require techniques i.e. relevance feedback, KDV and training framework.

REFERENCES

[1] Thagard, Paul, Cognitive Science, the Stanford Encyclopedia of Philosophy (Fall 2008 Edition), Edward N. Zalta (ed.).

[2] Thagard, Paul, Cognitive Science, The Stanford Encyclopedia of Philosophy (Fall 2008 Edition), Edward N. Zalta (ed.)

[3] I. Ruthven, "Integrating approaches to relevance," in New Directions in Cognitive Information Retrieval, (A. Spink and C. Cole, eds.), pp. 61–80, Netherlands: Springer, 2005.

[4] Larsen, B.; Ingwersen, P.; and Kek¨al¨ainen, J. 2006. The polyrepresentation continuum in ir. In IIiX: Proceedings of the 1st international conference on Information interaction in context, 88–96. New York, NY, USA: ACM.

[5] INGWERSEN, P. Information retrieval interaction. London: Taylor Graham, 1992.

[6] INGWERSEN, P. Cognitive Persepective of Information Retrieval Interaction: Elements of a Cognitive IR Theory , Journal of Documentation, vol. 52, no. 1, March 1996, pp. 3—50

[7] Saracevic, T. (1997). The stratified model of information retrieval interaction: Extension and applications. Proceedings of the American Society for Information Science, 34, 313-327.

[8] Diane Kelly, "Methods for Evaluating Interactive Information Retrieval Systems with Users" [2009] Foundations and Trends R_ in Information Retrieval Vol. 3, Nos. 1–2 (2009) 1–224.

[9] J. Allan and et al. Challenges in information retrieval and language modeling. Workshop at University of Amherst, 2002

[10] Hook, A. and Börner, K., "Educational knowledge domain visualizations: tools to navigate, understand and internalize the structure of scholarly knowledge and expertise", Dordrecht:Springer, 2005,187-208.

[11] Wendy Lucas, Heikki Topi ,"Need Help? The Effects of Training on Web Searches", Proceedings of the International Conference on Information Technology: Computers and Communications (ITCCí03) 0-7695-1916-4/03 $17.00 © 2003 IEEE

[12] Sormunen, E. & Pennanen, S. (2004). The challenge of automated tutoring in Web-based learning environments for IR instruction. Inf. Res, 9: paper 169 [Available at http://InformationR.net/ir/9-2/paper169.html].

[13] Wendy Lucas , Heikki Topi ,"Using a process graph to improve system-user knowledge sharing", CHIMIT'07 ,

March 30-31, 2007, Cambridge, MA, U.S.A.Copyright 2007 ACM 1-59593-635-6/07/0003

[14] http://www.cognitiveinformatics.com/en/what-is-cognitive-informatics

[15] Jackson, Erin W. BA; Olive, Kenneth E. MD, "Ethics Committees in Small, Rural Hospitals in East Tennessee" Southern Medical Journal: May 2009 - Volume 102 - Issue 5 - pp 481-485

[16] http://en.wikipedia.org/wiki/Cognitive_science

[17] http://www.iamcalm.com/category/neuroscience-research/

.

SHORT BIO DATA FOR THE AUTHOR

Parul Kalra Bhatia is working as an Assistant Professor in the Depatment of Information Technology Amity School of Engineering & Technology, Amity University, Noida. She has 8 years experience in the field of Academics and is actively involved in research & development activities. She is pursing Ph.D. (CSE) in the field of Information Retrieval. She has her M.Sc. degree in Computer Science in 2003 and M.Tech in Computer Science and Engineering in 2005 from Banasthali Vidyapith, Rajasthan. Her are of interest includes Text Mining, Information Retrieval- Cognitive and Personalized Copyrights, Intellectual Property Rights, Digital Rights Management, Privacy Rights Management, Advance Database Management System. She has successfully published papers in national and international journals.
research papers

