CISM COURSES AND LECTURES

The series presents lecture notes, monographs, edited works and proceedings in the field of Mechanics, Engineering, Computer Science and Applied Mathematics.
Purpose of the series is to make known in the international scientific and technical community results obtained in some of the activities organized by CISM, the International Centre for Mechanical Sciences.

INTERNATIONAL CENTRE FOR MECHANICAL SCIENCES

COURSES AND LECTURES - No. 383

USER MODELING

PROCEEDINGS OF THE
SIXTH INTERNATIONAL CONFERENCE
UM97
CHIA LAGUNA, SARDINIA, ITALY
JUNE 2-5 1997

EDITED BY

A. JAMESON
UNIVERSITY OF SAARBRÜCKEN

C. PARIS
CSIRO AUSTRALIA

C. TASSO
UNIVERSITY OF UDINE

1997
SpringerWienNewYork

Users and Intermediaries in Information Retrieval: What Are They Talking About?

Tefko Saracevic[1], Amanda Spink[2], and Mei-Mei Wu[3]

[1] School of Communication, Information and Library Studies, Rutgers University, New Brunswick, NJ, U.S.A.
[2] School of Library and Information Sciences, University of North Texas, Denton, TX, USA
[3] Department of Social Education, National Taiwan Normal University, Taipei, Taiwan, R.O.C.

Abstract. Discourse between users and intermediaries (human agents), as they interact when searching large databases, serves the function of user modeling. Selected data from a real-life study are presented, categorizing the utterances and elicitations (questions) into seven categories. The results provide an empirical picture of constructing user models through discourse and searching. A stratified interaction model is used as a theoretical model and framework.

1 Information Retrieval and User Modeling

When, at the beginning of 1950's, Calvin Mooers, one of the information science pioneers, coined the term *information retrieval* (IR) he also defined the problems addressed by the activity: (1) How to represent and organize information intellectually? (2) How to specify a search intellectually? and (3) What systems and techniques to use for those processes? (Mooers, 1951). Since Mooers' time, IR has developed in sophistication in both theory and practice. Interaction evolved to become a hallmark of modern IR. However, on a basic level, the problems defined by Mooers are still with us. It is the second of these problems, *how to search intellectually*, that directly involves user modeling. Searching is based on queries derived from users. In turn, they represent a whole set of underlying variables related to users and use. Thus, from the very outset to this day, user modeling (although not necessarily under this name) has been an integral component of IR. There is no IR without user modeling of some sort.

Effectiveness in IR concerns retrieval not of any old information but of *relevant information*. Or rather objects—texts, images, sounds, or for short *texts*—conveying potentially relevant information. The term *relevant* refers to information that pertains to the user's problem or situation at hand, involving also the user's cognitive and affective states and beliefs. Thus, by choice *relevance* became the basic underlying notion in IR, with all its attendant ambiguities as a very human notion. Not that the notions of *uncertainty*, used in expert systems and knowledge bases, or *aboutness*, used in classification and indexing systems, are any less ambiguous.

Thus, user modeling in IR has to be understood in terms of (1) the notion of relevance, and (2) the process of interaction to sharpen the likelihood of relevance. The object of user modeling in IR is to bring about retrieval of relevant texts for (a) given user(s) or use(s) through a variety of interactive processes. A number of methods for user modeling have been developed in IR, as summarized later. The most powerful method in existence so far involves mediated interaction,

that is, an interaction involving a user, a human intermediary, and an IR system. The intermediary is an information professional (variously also called an information scientist, a reference or special librarian, an information broker, a searcher, an information officer, or the like) skillful in both user modeling and the subsequent searching of and retrieval from various IR systems and large databases. Intermediaries play various roles, among them to: assist in the diagnosis of the user's problem and in the (re)formulation of the question; suggest appropriate systems or databases for searching; translate the question into one or more queries and search strategies acceptable to the given system and database; conduct and modify searching; assist in the evaluation of results; provide the user with appropriate outputs; and/or counsel the user in follow-up activity. In other words, to use the AI parlance, an intermediary is a truly intelligent agent constructing, implementing, and modifying user models in all their complexity with considerable feedback.

2 Rationale and Objectives of the Study

Even since the advent of user modeling by automatic or semiautomatic means in IR, or for that matter in AI, nothing has come close to matching the extent, complexity, and success of user modeling as done by skillful professional intermediaries in direct interactive contact with users. Thus, the observation and analysis of such activities involving users and human agents is of critical importance for understanding user modeling (Belkin et al., 1987). Moreover, unlike many other user modeling efforts, it is real and it has a rich context.

Intermediaries use various empirical methods for defining user models, which are also found in other diagnostic, interviewing, and counseling activities. This is accomplished through discourse, a complex dialog that takes place between users and intermediaries (dyadic dialog), or among users, intermediaries, and systems (triadic dialog). The modes of discourse may vary: oral, written or both; face-to-face or remote; with or without relevant or not relevant texts as models of what is or is not desired (relevance feedback); etc. But the basics of discourse remain.

This brings us to the central point of the study. The problem underlying all of the theoretical, experimental, and empirical activities in user modeling revolves around the classic and most difficult question (Belkin, 1993): What it is important to know about the user in order to support the user in interaction with the IR system? The answer has not been found by a long shot either in IR or in AI. In this study, we are trying to contribute some answers to that classic question. Discourse is the way in which users are modeled in IR. We may think that even users searching without intermediaries are engaged in a discourse with themselves, the system, and/or the retrieved texts. Users talk to themselves, carry on a dialog (written or oral) with a system, and converse with outputs. But the discourse between users and intermediaries is the most observable instance of user modeling. Thus, we suggest that analysis of such discourse can and does provide for a better understanding of what is involved in user modeling. In turn, a better understanding of this process is a necessary condition for improvements.

Data for analysis were selected from a large study whose aim was to contribute to formal understanding and characterization of IR interactions from the human perspective. The objective of this paper is to characterize discourse between users and professional intermediaries in interactive IR situations, in order to derive a discourse- and interaction-based user model for IR. The following questions are asked: What topics are covered in discourse and to what extent? What type of questions are asked by users and by intermediaries? How can we model these as interactions? The study is distinguished in that it derives data from observation of real users and intermediaries, in real IR interactions and settings. It is an empirical evaluation of what we believe is a typical behavior of users and intermediaries. It is a naturalistic study, one of the largest, if not the largest of its kind.

3 Approaches to User Modeling in Information Retrieval

System-centered approaches. Many formal methods related to IR systems have been developed and tested. In the *relevance feedback* approach, users are modeled through texts that are assessed as relevant (or not relevant)—they are used to retrieve or reject similar texts or clusters of texts (Spink and Losee, 1996). In the *query expansion* approach the initial or modified query is used as a basis for user modeling. Terms and logical connectives are expanded, contracted, added, switched etc. by automatic, semi-automatic, and manual techniques (Efthimidiadis, 1996). Another method is to build into the system ways and means by which users can on their own model their problem with the system's assistance. Examples are experimental systems by Oddy (1977) and Croft and Thompson (1987), but unfortunately such systems have remained in labs. Finally, the most widely used method combines the system- and user-centered approach. Originally, it was developed by Hans Peter Luhn, the most inventive information science pioneer, under the name of Selective Dissemination of Information (SDI) (Luhn, 1961). SDI, known also under a number of other names, involves the periodic retrieval of texts from recent updates, or filtering from streams of oncoming texts, based on what Luhn called *user profiles*. A profile is a search statement expressing user interests—it is a user model. Profiles are dynamic, they can and do change, and after a time they may stabilize. Elaborate ways have been developed for arriving at changing and testing user profiles. The Text Retrieval Conference (TREC), a current very large scale international IR evaluation effort, includes a routing track, which involves tests of profile optimization and learning. Profiles in IR largely precede and outperform "intelligent" agents and knowbots in AI.

Human-centered approaches. This perspective also involves several approaches. The area is rich, thus only a few examples are given. *Question analysis* and the reference interview are methods by which user modeling is accomplished through various interview and analysis techniques; they have received considerable attention over time (see, e.g., Taylor, 1968, Harter, 1992, Radford, 1996). Users' *cognitive aspects* have also been investigated to a great extent (see, e.g., Allen, 1991). In interactive studies, user modeling is treated as an integral part of the IR process (Belkin, 1993), a stance taken here. The main point in all of these studies is that user modeling is highly dynamic.

Connections. Although they address very similar processes, there is little crossover between the studies of user modeling in IR on the one hand and AI and cognitive science on the other hand. For example, user modeling studies such as the ones in the compilation by Kobsa and Wahlster (1989) or in articles appearing in the journal *User Modeling and User-Adapted Interaction* by and large do not cite or reflect on user modeling in IR. Conversely, only a few IR user modeling studies use concepts and techniques developed in AI. Notable exceptions in IR are works by Brajnik et al. (1987), who used analytic cognitive models developed in expert systems research for user modeling in "intelligent" IR, and by Logan et al. (1994), who applied Galliers' theory of belief revision, developed in AI, for use in IR. This relative isolation is unfortunate. Wheels are widely reinvented on both sides.

4 Modeling IR Interaction

The traditional IR model, represented in two prongs, a system and user one, in reality concentrates on the system side only. Moreover, it does not reflect and incorporate interaction at all (Belkin, 1993). Thus, a number of interactive IR models have been developed, and these models, unlike the traditional model, are suitable as a context for user modeling. Ingwersen (1996) took a broad approach and suggested cognitive representations by all participants in interaction users, texts, intermediaries and systems to serve as the base for a cognitive model of IR. Belkin et al. (1995) took a more specific approach and treated IR interactions as a series of episodes or frames, each of which supports different types of interactions and tasks. Finally, Saracevic (1996a, 1996b) proposed a stratified model of IR interactions, used in this study.

Space allows for only a brief description of the stratified model. Interaction is taken as a discourse between a user and "computer" through an interface. The "computer" involves much more than hardware (hence the quotation marks). It includes, among other things, computational capacities and procedures, and information resources or content. Both users and the "computer" are decomposed into different strata (levels), comprising distinctly identifiable elements or variables affecting the process in different ways. Interaction is then treated as an interplay between different user and "computer" strata or levels realized on the *surface level* through the interface. On the user side we can model *surface, cognitive, affective,* and *situational* levels. On the "computer" side we can also model levels: *surface, engineering, processing,* and *content,* as shown in Figure 1.

Interaction is a series of dynamic interplays and adaptations between levels. As the interaction progresses things change. For instance, on the surface level a query may be changed, terms added or deleted, different tactics employed, and so on, reflecting and affecting changes at other levels. Situational and cognitive states may be re-interpreted, new texts sought, etc.

We applied the stratified model to consider other IR notions also in strata. Relevance inferences are made in connection with different strata; thus, we suggested that in IR we have a dynamic interdependent *system of relevances* (note the plural) (Saracevic, 1996b). The model was also used in a study of identification and effectiveness of sources of search terms in queries (Spink and Saracevic, 1997). We are using here the stratified model as a basis for the explication of user modeling in IR. We are suggesting that user modeling is (i) an interactive process that (ii) proceeds in a dynamic way at different levels trying (iii) to capture user's cognitive, situational, affective and possibly other elements (variables) that bear upon effectiveness of retrieval, (iv) with an influence of intermediary interface capabilities, and (v) with an interplay with "computer" levels. It is an interactive diagnostic and counseling process.

5 Data Corpus

Details are presented in a number of papers, among them those by Saracevic et al. (1990) and Spink and Saracevic (1997); thus only a sketch is presented here. Forty self-selected academic users (faculty and doctoral students) with real information problems provided one question each for online searching on DIALOG. Four professional search intermediaries were involved, each handling ten questions. Questions concerned topics in medicine, the social sciences, the physical sciences and the humanities. A mean of 3 databases were searched per question.

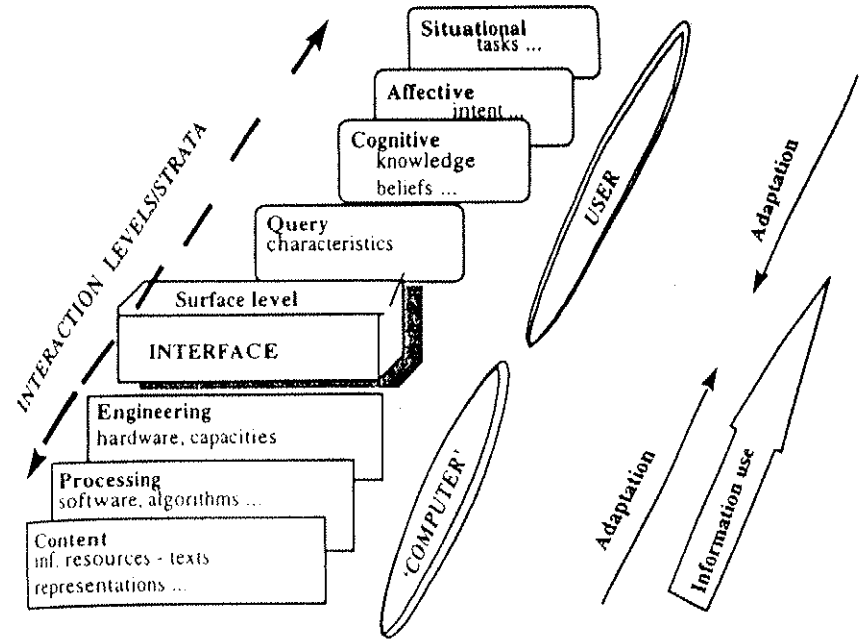

Figure 1. Elements in the stratified model of IR interaction.

Prior to the search users provided a written question statement. The interaction between users and intermediaries was videotaped during a pre-online search interview and during the actual online search. There were 46.05 hours of videotape (with a mean of 69.08 min. per question). The mean time for the pre-search interview was 13.04 min. and for online search 56.04 min. The transaction logs of searches (computer commands and responses) were also recorded. The discourse between users and intermediaries was transcribed from videos. Utterances in discourse were analyzed to develop a grounded-theory set of categories, as defined below, and then categorized accordingly. Similarly, computer commands and responses were categorized according to our own categories. The utterances in discourse from the transcripts and commands/responses from the transaction logs were synchronized with respect to time.

Users were given a printout of all items retrieved. They judged retrieved items as *relevant, partially relevant,* or *not relevant.* Altogether 6225 items were retrieved, of which users judged 3565 as relevant or partially relevant, and 2660 as not relevant, for a mean precision of 57%. However, there was considerable variation in retrieval from question to question: the maximum/minimum number of *total items* retrieved was 427/13 with a standard deviation of 85.9; the

max./min. number of *relevant or partially relevant* items was 348/1 with an SD of 71.43; and the max./min no. of *not relevant* items was 180/0 with an SD of 47.2. A number of other variables were identified through questionnaires and other analyses, as reported in the cited papers.

6 Categories of Utterances

The utterance is a basic unit of discourse analysis in user-intermediary interaction, thus also a basic unit in user modeling As a first step we derived a set of categories, focusing on the major theme or topic of an utterance. The categories were derived in a grounded theory approach and refined in several experiments, following other studies, such as that of Belkin et al. (1987). While the categories as defined can clearly be split further, we concentrated on a limited number of general categories to increase the reliability of coding and to increase the potential of detecting broad differences. We aimed at macro rather than micro analysis. The categories are shown in Table 1.

We suggested a stratified model of IR interaction as a framework for user modeling. The categories we defined can be mapped into the suggested user and "computer" strata or levels in the interaction model. However, with the exception of the category *context*, which maps into the *situational* level, the mapping is not one to one. The categories actually represent interplays between different user and "computer" levels. The category *terminology and restrictions* is realized on the *surface* level, but it is also affected by how it is represented on the *content* level in texts. The categories *system explanation* and *action* relate the *cognitive* level of the user with any level on the "computer" side. *Search tactics* relates any of the user levels with the *processing* "computer" level. The *review and relevance* category relates any and all user levels with the *content* "computer" level. *Backchanneling* helps the whole interplay process. The categories by and large characterize the interplay between levels in the interaction process. Together, the categories of utterances and the levels in interaction describe the process of user modeling in IR.

7 Results: Distribution of Utterances and Elicitations

Table 2 presents the number of utterances according to the defined categories, excluding the last category *extraneous* because those utterances do not contribute to user modeling. The arrangement is from the highest to the lowest number of utterances in the "Total" column.

The great variation in retrieval results, mentioned above, is mirrored in the variation in the number of utterances from question to question. The mean number of user utterances per question was 243, with a maximum/minimum of 665/62, and an SD of 126.34. The mean number of intermediary utterances was 283, with a max./min of 714/94 and an SD of 132.14. We also calculated the interaction times: The mean time per question for user utterances was 21.32 minutes (max./min. of 48.3/4.36 minutes with an SD of 644); the mean time for intermediary utterances was 47.69 minutes (max./min. of 102.5/16.3 minutes with an SD of 1251).

Let us now concentrate on the analysis of elicitations—a subset of utterances that refer to a verbal request for information by either party. They could be in a question or a non-question form, hence the term *elicitation* rather than *question*. Since elicitations reveal topics queried in interaction, they are a critical aspect in any and all user modeling. Data on elicitations, presented in Table 3, are culled from two studies that used the same data corpus: one that studied elicitations by users alone (Wu, 1993), and another that incorporated both users' and intermediaries' elicitations (Spink et al., 1996). However, we modified somewhat the elicitation categories from those two studies to conform to our categories of utterances. Wu had 10 categories for user elicitations and Spink et al. had 15 categories for intermediary elicitations; we combined some of the more detailed categories in each that obviously fall into categories of utterances as defined above. We also eliminated extraneous elicitations, for the reason given above. Again, the categories are ordered according to the number in the "Total" column.

Table 1. Categories of utterances distinguished in the empirical study.

Category number and name	Description
1. Context	User's problem or task at hand; information seeking stage; information, if any, collected so far; expectations and other aspects underlying the question; user domain and problem knowledge; user's plans.
2. Terminology and restrictions	Elaboration on and modification of concepts, terms, keywords and descriptors; generation of terms; specification of borderlines; restrictions such as with respect to language, years; technical term spelling.
3. Systems explanations	Workings and technical aspects of system used; technical explanation of searching; characteristics of databases and documents in system; other possible information sources; obtaining texts; costs involved.
4. Search tactics and procedures	Selection and variation of terms, fields, morphology, logic in search statements; commands; selection and variation in magnitude and output sizes, formats, order; output specification; correcting mistakes.
5. Review and relevance	Review of search statements with respect to the output; evaluation of output sources or content; relevance judgments of and feedback from outputs; decisions or questions on what is wanted based on search statements or output.
6. Action	Description of an ongoing or impending activity, e.g. thesaurus lookup, output formats, printing; explanation of what is happening.
7. Backchanneling—prompts, echoes	Communication prompts, fillers, acknowledgments, formulaic expressions, etc indicating listeners involvement, e.g. "O.K.", "Wow!", "Unhuh", "Right"; echoes and requests for repetitions e.g., "What?", "Pardon?", "Say that again"; pauses.
8. Extraneous	Utterances extraneous to the search interaction—greetings, formulaic courtesies, social comments and questions; personal matters.

Table 2. Distribution of utterances in user-intermediary interaction (each percentage refers to the column in question).

Category number and name	Intermediary		User		Total	
4. Search tactics and procedures	3360	30%	1680	17%	5040	24%
7. Backchanneling—echoes	1179	10%	3179	33%	4358	21%
5. Review and relevance	1996	18%	1825	19%	3821	18%
2. Terminology and restrictions	1265	10%	904	9%	2169	10%
3. Systems explanations	1399	12%	625	6%	2024	10%
6. Action	1554	14%	399	4%	1953	9%
1. Context	565	5%	1102	11%	1667	8%
Total	11318	100%	9714	100%	21032	100%
% of Total		54%		46%		100%

Table 3. Distribution of elicitations by intermediaries and users (each percentage refers to the column in question).

Category number and name	Intermediary		User		Total	
2. Terminology and restrictions	817	55%	288	32%	1105	46%
5. Review and relevance	251	17%	107	12%	358	15%
4. Search tactics and procedures	59	4%	240	27%	299	13%
3. Systems explanations	103	7%	174	20%	277	12%
1. Context	172	12%	1	0%	173	7%
7. Backchanneling—echoes	50	3%	47	5%	97	4%
6. Action	34	2%	34	4%	68	3%
Total	1486	100%	891	100%	2377	100%
% of Total		63%		37%		100%

8 Discussion: Implications for User Modeling

Several aspects are of interest to note from the distribution of utterances (Table 2). Users talk somewhat less than intermediaries (46% vs. 54% of all utterances), but the difference is not that large. Both take about the same number of turns. If we consider that the basic object of interaction is to model users, then both parties are participating in the process almost equally.

The top category, comprising one fourth of the utterances, deals with various *search tactics and procedures*, with intermediaries in this category talking substantially more than (twice as much as) users. In other words, users get and exchange a lot of information about the immediate processes concerned with the search itself. Searching itself, with associated changes, enters into user modeling in a major way. In contrast, the bottom category, comprising only 8% of all utterances, deals with the *context* of the question and user. This was a surprise—we expected that there would be much more talk about the various things that fall under context. Neither users nor intermediaries explicate on this topic a lot. This challenges the usual assumption that user modeling largely involves modeling of context.

Interestingly, *backchanneling*, with over a fifth of the utterances, is the second largest category. These are brief utterances (even grunts) that facilitate communication, indicate active participation, provide echoing questions, and the like. They seem to play a significant role in interaction. Users made considerably more than twice as many backchanneling utterances, which may reveal characteristic features of the nature of user participation, such as a lot of confirmation, indication of understanding, and posing of brief questions. Backchanneling represents a unique human device and tactic that speeds communication and increases mutual understanding, and it seems to be important in user modeling. Maybe communication with computers is so limited and unsatisfactory because it does not involve this important human communicative element.

With fewer than a fifth of the utterances, *review and relevance* is the third highest category. Here, the utterances are almost equally divided among users and intermediaries. These utterances reflect possible changes in or confirmations of what is going on in searches, outputs, and their relations. In turn, they may suggest new and other terminology and tactics, thus they may affect considerably the subsequent utterances in other categories. Thus, review of searches and relevance assessments also play an important role in user modeling.

Utterances about *terminology and restrictions*, *systems explanations*, and *action* comprise the other three categories, with roughly one tenth of the utterances each. Not surprisingly, intermediaries made about three times as many utterances in categories dealing with systems explanations and action as users—after all, that is their domain. These categories indicate the importance of providing users with explanations of what is in the systems, what capabilities exist, and what is going on, as an integral part of user modeling.

Elicitations present quite a different picture than utterances overall. Close to one tenth of all utterances were elicitations. This small proportion also surprised us. Intermediaries made about three fifths of all elicitations, showing a reliance on the question-answer process for user modeling by intermediaries. However, users also had their fair share of elicitations, with some two fifths of the total, showing a similar reliance by users on understanding what is going on. Elicitations are a leading edge of interaction, triggering other actions. The question is: What kind? User modeling and elicitations go hand in hand.

As to the categories, close to half of all elicitations are about *terminology and restrictions*, with intermediaries asking about one and a half times as many terminological questions as users, probably suggesting or confirming choices. There seems to be a terminological imperative governing elicitations. This is not surprising, because queries entered into the interface for searching are terminology to start with, and terminology presents a major problem and decision making aspect in IR and in user modeling. Thus, terminology plays a major role in question-answer sequences of user modeling and predominates over other questions.

The other half of elicitations cover the other six categories. About a seventh of all elicitations are devoted to *review and relevance*. This was a surprise, for we thought that there would be more elicitations about these topics. *System explanations* and *search tactics and procedures* had a bit more than a tenth of the elicitations each. Surprisingly, only 7% of the elicitations were about *context*. Finally, *backchanneling* and *action* are at the bottom with negligible percentages. The combined elicitations related to relevance, tactics and systems explanation seem to be another major component in user modeling, while context elicitations play a smaller role.

The difference in the number of utterances from question to question was very large. So was the difference in the time spent talking by users and intermediaries. In general, intermediaries talked much longer, over two times as long as users. By far the longest time was spent by intermediaries on the category *terminology* and related aspects, the same category where they had the most elicitations. This again indicates the power of the terminological imperative in user modeling. Clearly, all questions are not equal, indicating that the nature of the question itself and all that goes with it requires user modeling of different intensity, and even different types.

What is not evident from the data as presented, but comes out clearly from a qualitative evaluation of the process, is the nature of the changes that occur during each interactive session. Utterances are here categorized with respect to their basic category, but as in a grammar they do not reflect semantics, even less pragmatics, that is they do not reflect the meaning of what is actually going on. Most evident are considerable *shifts* or *transitions* that happen as the interaction unfurls. Among these are: shifts in terminology and tactics; redefinition of the problem and refitting to the situation; illumination of some dimension of the situation; changes in rationale and expectations; changes in cognitive realization, understanding and learning, and the like. Thus, as a rule and not as an exception, a question as asked and subsequently processed is quite different at the end than at the beginning, undergoing a number of transformations in between. User models and modeling changes in the course of the interactive process. Unfortunately, in IR and elsewhere we have not as yet developed methodologies that may appropriately aggregate such changes, shifts, and transitions in interaction, and provide a sort of transitional grammar. We do not even have a good classification of these shifts. Beyond anecdotes we cannot document them well. For user modeling these shifts are crucial, critical incidents. However, intermediaries understand professionally that user modeling is an evolving transitional, shifting process, thus they direct the interaction, utterances, elicitations, decisions, suggestions etc., accordingly. User modeling in IR comprises shifts and transitions. It evolves as it goes along. It is not a static, one time deal. It does not rest on the initial query alone.

The data presented here are descriptive of the types of utterances and elicitations that go on in user-intermediary interaction. However, the data do not address the effectiveness of different elements or categories in interaction. In this analysis we have not addressed the critical questions related to effectiveness, such as: Are some interactions with different distributions of categories more effective than others in terms of retrieval of relevant texts? In terms of user utility and satisfaction ratings? Do they provide significant (positive or negative) correlation with a myriad of other variables that enter into IR? We have data that may answer some of these questions, e.g. correlations between the distribution of utterance categories in different questions with a variety of effectiveness measures and other variables. But these data await another analysis and another paper. However, even the picture we presented is useful in itself for further understanding of the content and diversity of the interactive user modeling process.

9 Conclusions

The central premise of this study is that analysis of discourse between users and intermediaries can provide a better understanding of user modeling. To do this, we have shown the distribution of categories of utterances, including elicitations, by users and intermediaries from a large corpus of interaction data, and we have discussed these categories at length as types involved in user modeling in IR. Furthermore, we have suggested a stratified IR interaction model to serve as a framework for user modeling. The categories of utterances for the most part represent interplays between user and "computer" levels in interaction. The strength of the study is that it is based on empirical, real-life evidence. It is also a unique study in many respects, particularly with respect to size. The weakness is the same as in all empirical studies: Conclusions pertain really only to the evidence at hand. That is, generalizations, while documented, should be taken as no more than hypotheses to be confirmed by further study.

Data show an extensive and dynamic interplay between users, intermediaries, and the "computer" as the third party in discourse. Many of the utterances and categories pertain to the actions or results related to the interplay with that third party. A large proportion of the utterances, in particular elicitations, were aimed at a dynamic (re)formulation of the user model, with an active involvement of "computer" levels. For instance, a lot of the utterances deal with the system itself, how users react and remodel themselves from what the system provides and how. This indicates that in user modeling there is much more going on than in the narrow perspective of user modeling encompassed by relevance feedback, query expansion, and related automatic or semi-automatic techniques in IR, or in AI. User modeling involves various interactive loops among various levels on both sides—a user model is evolving and changing as the process unfurls. We suggest that along with a system of relevances, we have a system of user models (plural) in IR involved in every interaction where users participate actively. The stratified model describes the elements in this system and categories of utterances describe the interplays between elements.

The concepts and findings elaborated here do suggest a different understanding of what is involved in user modeling in or outside of IR. We also question whether user modeling, as such a complex process, can be attempted with any degree of success by reducing it to one or just a few elements, whatever they may be. Yet this is a common approach in both IR and AI. Many approaches rest on assumptions that constitute a gross oversimplification of user modeling. Reductionism did not work elsewhere, and it does not work with user modeling either. It is not surprising then that attempts to automate user modeling have been quite limited. We are suggesting again, as we did at the outset, that detailed observation of what is actually going on in user modeling, particularly involving human agents, is a fruitful ground for further understanding of the process, which in turn is a prerequisite for any improvements and for the derivation of design criteria for automating this delightful and completely human process.

References

Allen, B. L. (1991). Cognitive research in information science. In Williams, M. ed., *Annual Review of Information Science and Technology* 26:3–37.

Belkin, N. J., Brooks, H. M. and Daniels, P. J. (1987). Knowledge elicitation using discourse analysis. *International Journal of Man-Machine Studies* 27:127–144.

Belkin, N. J. (1993). Interaction with text: Information retrieval as information seeking behavior. *Information Retrieval: 10. Von der Modellierung zur Anwendung*. Konstanz, Germany: Universitätsverlag Konstanz, 55–66.

Belkin, N. J, Cool, C., Stein, A., and Thiel, U. (1995). Cases, scripts, and information seeking strategies: On the design of interactive information retrieval systems. *Expert Systems with Applications* 9:379–395.

Brajnik, G., Guida, G., and Tasso, C. (1987). User modeling in intelligent information retrieval. *Information Processing & Management* 23:305–320.

Brajnik, G., Mizzaro, S., and Tasso, C. (1996). Evaluating user interfaces to information retrieval systems: A case study of user support. In *Proceedings of the Nineteenth Annual ACM/SIGIR Conference on Research and Development in Information Retrieval*, 128–136.

Croft, W. B., and Thompson, R. (1987). I^3R: A new approach to the design of document retrieval systems. *Journal of the American Society for Information Science* 38:389–404.

Efthimidiadis, E. N. (1996). Query expansion. In Williams, M. ed., *Annual Review of Information Science and Technology* 31:121–187.

Harter, S. P. (1992). Psychological relevance and information science. *Journal of the American Society for Information Science* 43:602–615.

Ingwersen, P. (1996). Cognitive perspectives of information retrieval interaction: Elements of a cognitive IR theory. *Journal of Documentation* 52:3–50.

Kobsa, A., and Wahlster, W., eds. (1989). *User Models in Dialog Systems*. Berlin: Springer.

Logan, B., Reece, S., and Sparck Jones, K. (1994). Modeling information retrieval agents with belief revision. In *Proceedings of the Seventeenth Annual ACM/SIGIR Conference on Research and Development in Information Retrieval*, 91–100.

Luhn, H. P. (1961). Selective dissemination of new scientific information with the aid of electronic processing equipment. *American Documentation* 12:131–138.

Mooers, C. (1951). Zatocoding applied to mechanical organization of knowledge. *American Documentation* 2:20–32.

Oddy, R. N. (1977). Information retrieval through man-machine dialogue. *Journal of Documentation* 33:1–14.

Radford, M. L. (1996). Communication theory applied to the reference encounter: An analysis of critical incidents. *The Library Quarterly* 66:123–137.

Saracevic, T. (1996a). Modeling interaction in information retrieval. In *Proceedings of the 59th Annual Meeting of the American Society of Information Science* 33:3–9.

Saracevic, T (1996b) Relevance reconsidered. In Ingewersen, P., and Pors, N. O., eds., *Information Science: Integration in Perspectives*. Copenhagen: Royal School of Librarianship 201–218.

Saracevic, T., Mokros, H., and Su, L. (1990). Nature of interaction between users and intermediaries in online searching: A qualitative analysis. In *Proceedings of the 53rd Annual Meeting of the American Society for Information Science*, 27:47–54.

Spink, A., Goodrum, A., Robins, D., and Wu, M. M. (1996). Search intermediary elicitations during mediated online searching. In *Proceedings of the 19th Annual ACM/SIGIR Conference on Research and Development in Information Retrieval*, 120–127.

Spink, A., and Saracevic, T. (1997). Interactive information retrieval: Sources and effectiveness of search terms during mediated online searching. *Journal of the American Society for Information Science* 48. Forthcoming.

Spink, A., and Losee, R.M. (1996). Feedback in information retrieval. In Williams, M., ed., *Annual Review of Information Science and Technology* 31:33–78.

Taylor, R. S. (1968). Question negotiation and information seeking in libraries. *College & Research Libraries* 29:178–194.

Wu, M. M. (1992). *Information Interaction Dialogue: A Study of Patron Elicitation in IR Interaction*. Unpublished Ph.D. dissertation, Rutgers University, New Brunswick, NJ.