Information Seeking as Explorative Learning

Giorgio Brajnik
Dipartimento di Matematica e Informatica
University of Udine
giorgio@dimi.uniud.it

8th June 1999

Abstract

The paper starts with an overview of major problems hindering effective interactions with information retrieval systems. A brief review of models of the interaction taking place during information seeking is then provided with the aim of lying the ground for yet another model based on two levels of interaction: presentation and navigation. The former encompasses interactions with the user interface of the retrieval system, whereas the latter deals with the interaction that users experience with the information resource. It is claimed that this second level of interaction can be framed as an exploratory learning process and that its analysis provides adequate support for designing and evaluating information access systems.

1 Introduction

Task based evaluations of an information retrieval system are crucial steps in its development, as they can lead to appropriate levels of usability and gain thus user acceptance. This is especially true nowadays that more and more information resources are being searched not only by professional searchers, but also by a very heterogeneous user population.

Task based evaluations, founded on models describing how given activities are carried out, are potentially able to yield a usage picture that is more realistic than results produced by other system–oriented evaluation methods. Furthermore, evaluations focussing on usability, that is the property characterizing how well a system supports its users in performing their job, yield a picture that is richer than that produced by evaluations of effectiveness only. Several usability factors could concur in the definition of such a picture; for a generic interactive system these factors include: ease of learning, ease of retaining what has been learned, error rate, consequences of errors, ease of recovery, efficiency of use, flexibility, pleasure.

In the case of Information Retrieval, for many years practitioners have focussed on system–oriented evaluations of effectiveness, paying little attention to the role played by interactive processes taking place during information seeking. In order to move to a user–centered evaluation and promote comparison of results obtained under different experimental conditions, appropriate frameworks need to be defined. These frameworks have to support the formulation of tasks describing how the information seeking process is articulated. They have to provide also a grid of properties and observation methods in order to draw and justify conclusions about usability. So far, no such framework has emerged.

In this paper I shall outline essential features of tasks models widely discussed in the literature. The aim is to propose yet another framework that, in my view, is sufficiently general to cover the wide spectrum of models so far proposed, but is also showing a close relationship with usability factors. I will claim that this relationship makes usability inspection methods available to information retrieval researchers. Methods that, so far, have had no role in information retrieval while having the potential of being more economical and complementary to user–testing.

2 Interacting with information access systems

For this paper an *Information Access System* (IAS) encompasses both an information resource and the system that supports its organization, update and access to users. An IAS is more general than an information retrieval system as it includes the information resource and no assumption is made on the search model it supports (may be browsing, or retrieval, or both). The information resource comprises basic information items (like multi–media documents, web pages, encyclopedia chapters, terms, ...), relationships between them (like explicit hypertextual links, citation links, term co–occurrence links, ...) and auxiliary information providing additional relationships (like thesauri, classification schemas, citation indexes, ...).

End–users purposefully interact with an IAS in order to change their state of mind; a process called *information seeking* [**?**]. Research has demonstrated that information seeking is a very complex process [**?**, **?**, **?**, **?**, **?**, **?**]. Such a complexity is due to the extremely high variability of human behavior, a consequence of the cultural and cognitive state of the user when seeking information. Variability of user behavior is shown by the relatively large portfolio of actions that users take during the information seeking process (e.g. refining the query with new terms, performing relevance feedback, zooming on a set of documents, following an explicit link, scanning some results, ...) and decisions criteria that can be adopted (e.g. for viewing documents from the hit list, for selecting search terms, ...). Such a variability poses a big challenge to designers of IAS's, as the design problem is very little constrained; issues range from placing the human–machine boundary [**?**] to providing informal and flexible searching environments [**?**].

2.1 Fundamental problems

There are at least three problems that are responsible for complex user behaviors. The first one is described by the Anomalous State of Knowledge (ASK) hypothesis [**?**], according to which users performing some activity feel that they have a knowledge gap that cannot be filled directly, and consequently they engage into an information seeking process without knowing precisely what to look for. Second, the Vocabulary Problem hypothesis [**?**], stating that there is an inherent vocabulary mismatch between the user seeking information and the persons organizing the information resource (i.e. writing the documents, indexing and classifying them, defining referral aids). Third, the Relevance problem [**?**, **?**]. Relevance is a complex multifaceted notion: relevance may be a relation between document surrogates, actual documents, or their contents and user's information need, a verbal request, or a formal query; it may be determined by the information seeker, by indexers, or by some external observer; it may take into account the reasons why the user is seeking information (i.e. utility vs. topicality); it depends on time (as the information seeking process evolves, the information need may change and so does relevance) and on context (e.g. a topical document that is already known to the seeker is useless).

2.2 Strategic difficulties

These problems have a strong impact on the success of an information seeking session. Several studies show that inexperienced users often get in trouble because of their poor ability to design and follow effective search strategies.

Novices seem to apply only simple strategies: they build their queries using only terms they have in mind, without consulting available thesauri and not taking advantage of the interaction with the database; they perform only little query modifications and refinements [**?**, **?**, **?**, **?**, **?**]. After retrieving documents, users do not look carefully at them, missing in such a way an opportunity to improve the search process. Even when users read retrieved documents, they are not always able to exploit the information they gathered: for instance, they judge a document as relevant without realizing that they can also extract some terms to better reformulate their queries [**?**, **?**, **?**]. Many users adopt a single search strategy for the whole session [**?**], with negative consequences when it is not well suited for the particular situation at hand (information need, data base and information retrieval system). Even when a valid but unfamiliar strategy is adopted, minor errors (like typing errors) can cause its abandonment [**?**].

Some users react in ineffective ways to difficult situations: for instance, with a boolean system where no documents have been retrieved, users may add new AND–ed terms, an operation which can only decrease the number of retrieved documents [**?**].

Many of the previous observations were also confirmed by an experiment carried out on a user interface for a boolean information retrieval system [**?**]. Analysis of the behavior of participants showed that users lack an overall strategic view of the search. Less successful users are bound to a wrong and fixed conceptualization of the information need and use a limited portfolio of actions. For instance, they keep adding/removing/modifying terms that don't derive from the information resource but are made up by the users when thinking on the query. They never do field restrictions to search on authors, journals, classification codes or controlled terms. When these users get stuck in some difficult situation (for instance when they keep getting no retrieved documents), they apply previously tried actions even though these were clearly ineffective.

The experiment showed that besides terminological knowledge and knowledge about basic query manipulation, a successful support to end users has to be based also on strategic help, aiming at enlarging the portfolio of actions that users might want to try.

3 Evaluation of information access systems

The large and increasing amount of available information resources has the consequence that more and more of them are being used by people that have never been exposed to any information retrieval training. Therefore it is crucial for acceptance of IASs that these reach certain threshold levels of usability, and be able, for example, to solve or avoid the strategic problems discussed above. Effective information retrieval systems are no more sufficient: it is the compound system "user+computing system+information resource" that has to be effective.

3.1 System centered vs. user centered evaluation

Evaluation of information retrieval systems has been performed so far either in a user–centered or a system–centered fashion. In the former, the user is a central actor and her behavior determines the outcome of the evaluation. In the latter, the evaluation is aimed at measuring some property of the system and very little attention is payed to the person actually driving the system, if any. With respect to acceptability of the evaluated system, a user–centered evaluation tends to be more valid, as it will yield results that better represent the way the system will be used in the real world. It is also more complete, in the sense that the picture that one can get about the system will be richer, potentially taking into account not only system related factors like effectiveness and efficiency, but also those related to user satisfaction, robustness, flexibility, and learnability. On the other hand, however, user–centered evaluations are more demanding, as the erratic behavior of a user is difficult to control in an experimental setting and important output variables (like satisfaction) are difficult to observe and measure in a reliable way.

3.2 Usability evaluation methods

User–centered evaluations of information access systems are performed, almost always, via empirical user–testing, where a number of subjects are asked to solve information problems using the system. Their behavior is observed, certain parameters are measured, others are asked, and then conclusions are drawn. Notice that even though many parameters related with the outcomes of the experiment have an objective nature (like number of queries per session, number of relevant documents per query, etc.) they are still strongly dependent on user's behavior (for example on the specific queries being entered, on the individual selection of documents, etc).

For a different class of information access systems, namely web sites, empirical user–testing is not the only evaluation method that is available. In fact, usability engineering methods like *heuristic evaluation* or *cognitive walkthrough* [**?**, **?**], have been applied to evaluate usability of web sites. The former method, based on a panel of expert reviewers, is centered on a set of usability heuristics guiding the analysis (like "strive for consistency", "ensure flexibility"). The experts, at the end of the analysis, report on usability problems with associated severity judgments. The latter method, cognitive walkthrough, is based on the underlying assumption that users learn to use a system while doing their job (explorative learning). To carry on a cognitive walkthrough evaluation, a specific scenario of usage must be defined in terms of user goals and knowledge levels. Then designers, analysts and possibly a sample of users investigate if a certain interface feature can be perceived and if it can be correctly interpreted by end–users; if correct actions are visible and if they can be linked to users' goals.

For generic interactive systems, the three evaluation methods of user–testing, heuristic evaluation and cognitive walkthrough lead to possibly overlapping but different sets of usability problems: thus each method may contribute significantly to the evaluation outcomes [**?**]. So far, IAS have only been evaluated via costly user–testing.

3.3 Task–based evaluations

The goal of an evaluation is to determine if, or the degree to which, a system has certain properties in order to assess whether those properties are satisfied at all, or to compare the system with similar ones.

To measure how well an interactive system supports users in performing their jobs, evaluations have to be embedded in scenarios that are, or at least approximate, actual real–world situations. In order to carry out such an evaluation, a task model needs to be defined. *Tasks models* are descriptions of activities that are carried out in order to achieve certain goals [**?**, **?**]. These models are generated by *task analysis*, which is the process of analyzing the way people perform their jobs: what they do, the objects they operate on and the required knowledge.

Tasks that have been used in actual experiments span the range from very short requests to longer descriptions of information problems providing also utility criteria (i.e. how the information sought will be used). In addition to information problems that are made up by experimenters and then "induced" to actual users, also real information needs have been used. Particularly relevant is here the study by Borlund and Ingwersen [**?**, **?**] which points out that there is no significant difference in the outcomes of experiments between induced and genuine information problems.

Task models play an important role in evaluation because they constrain the usage scenario of the system under analysis and because they have an impact on the outcome. Selected tasks need to be representative of the activities deemed important; tasks mixes have to ensure that outcomes are informative; and finally tasks have to be matched to user categories (such as novices and experts).

Task models are abstractions of actual activities and, as such, they are described at certain levels of detail. The top level goal (say, for example, to prepare a slide presentation) may be decomposed into simpler goals, that need to be achieved in a given order (i.e. a task). When to stop this decomposition is an important modeling decision, for it will affect the granularity of the model and therefore determine how well the model will support design and evaluation of the system. Task models are used, during system development, also to define where the boundary between human and machine lies. Tasks being below such a level are performed automatically by the machine. As pointed out by Bates [**?**], this is an important and difficult decision when developing an IAS.

4 Task analysis for information access systems

A central problem for evaluation of information retrieval systems is to find an abstraction level that captures phenomena that are central to usability.

In the information retrieval literature several task models have been discussed, spanning many abstraction levels. Often the aim of these models is to shed light on cognitive processes taking place in the minds of end–users or professional intermediaries performing searches. More rarely, models have been used to drive the development process, including evaluation, of information access systems.

In order to be useful at that, models have to support usability evaluation and hence include, among the tasks they describe, those that lie near and on both sides of the human–machine boundary. Besides being accurate (i.e. faithful to reality), appropriate models have to be able to explain human behaviors (e.g. strategic problems mentioned in section **??**.). Furthermore, task models should help evaluators understand what kind of models users develop about the system they are using. Tasks should also be general enough to abstract away the aspects that are specific to a search session to enable a transfer of findings among different evaluations.

In the following, a brief and non exhaustive review of some task models discussed in the literature is illustrated.

4.1 Information retrieval and information seeking models

The traditional model of information retrieval (e.g. [**?**]) centers on the processes of representing a document, formulating a query, matching a query with a document representation and refining the query. Though variations of the basic model exist, they are all system oriented: they don't say almost anything about why a user engages in the process, how relevant information is detected, in which ways queries are modified, what are the cognitive processes occurring in user's mind.

More recently, Shneiderman [**?**] identified four unstructured tasks that describe some information seeking goals: *specific fact finding* (like finding a book record from its ISBN), *extended fact finding* (slightly more open questions, like finding what kind of music is published by a producer), *open–ended browsing* (like finding whether desertification and carbon–monoxide levels are related) and *exploration of availability* (like getting an overview of what kind of information about the Beatles is available on a web site). While giving some idea on possible goals of a user engaged in information seeking, these tasks are too general to support design and evaluation of IAS. Shneiderman in fact identifies, at a much lower level, four phases of information seeking which are closely related to the traditional model: *formulating* a query, *searching* the database, *viewing* the results and *refining* the query. He shows that even this level of analysis can be useful to explain *some* usability problems: for example the confusion that users of web search engines may face due to interfaces that don't tell which is the underlying query mechanism (whether boolean, statistical, quorum, etc.). What he doesn't do, unfortunately, is to fill in the gap between overall goals and low level actions.

Saracevic' stratified model [**?**] covers this gap. It describes information retrieval interactions as dialogues between users and machines through an interface at a surface level. Below the surface one can see descriptions of the system, at three different levels of abstraction: *engineering* (related to computer–specific aspects), *processing* (algorithms adopted in the information management system) and *content* (information items, their representations and relationships). Above the surface lie three strata used to describe the interaction from the user side: a *cognitive* level (where interpretation of information items applies), an *affective* level (dealing with beliefs, desires, motivations) and a *situational* level (dealing with the context that originated the information seeking process: why an information is sought and how it will be used). This model is fairly general but is generic as well: it doesn't have enough details to directly support evaluation. It can be used to formulate different notions of relevance, but using it to determine parameters to be observed, methods for measuring them, and hypotheses to explain or predict user behavior is very difficult.

A number of other models focus on an intermediate level which, in my opinion, is more appropriate for the usability evaluation problem. Bates [**?**] identified a level of analysis and synthesized a model that captures crucial aspects of successful information seeking sessions. She defined three concepts for discussing searchers behaviors: *tactic*, one or a handful of actions or thoughts made to further a search (e.g. moving to a broader term); *stratagem*, multiple tactics designed to exploit the structure of a search domain (e.g. "journal–run", i.e. browsing issues of a journal that is central to a topic); and *strategy*, a plan containing tactics and stratagems for an entire information search. She found that strategies employed by expert searchers are highly dependent on the specific information problem being solved and are determined by the tactics and stratagems that are opportunistically selected, applied and monitored to solve an information problem. Within a search session, queries are continuously shifting and successful users adopt varying patterns of search activities, switching among different stratagems and tactics. There is no single query that yields all the results, but these are gathered during the trajectory followed by the user. The importance of Bates' model stems from its concreteness. It describes specific actions that are taken by expert searchers and it also explains why novice users tend to face the problems discussed in section **??**.

Ellis [**?**] followed a different route when discussing information seeking patterns. On the basis of observations of the information seeking behavior of a sample of users (social scientists), he abstracted the behaviors into six categories, which are then used as a framework for a behavior model. The patterns are: *starting*, encompassing activities occurring at initial states of a search session, like focussing on already known references or authors; *chaining*, following references from documents at hand; *browsing*, based on scanning on–line material or shelves; *differentiating*, i.e. employing differences between information sources (like quality, rejection rates, frequency of citation) to filter the material being examined; *monitoring*, that is "... maintaining awareness of developments in a field through the monitoring of particular sources"; and *extracting*, i.e. "... systematically working through a particular source to locate material of interest" (more concentrated and directed than monitoring). Searchers' behavior can be described as a sequence of instances of those patterns.

A similar approach has been followed by Chen and Dahr [**?**]. Based on a protocol analysis of actual behaviors shown by a sample of users interacting either with librarians or with an information retrieval system, the authors framed users actions into a problem–solving context. Each behavior may thus be seen as a sequence of actions performed to move from one solution state to the next one. They abstracted out a number of "... approaches adopted to traverse the problem space": *known–item–instantiation,* i.e. retrieval of some known item (via author, title, call number) to get index terms to use in subsequent queries; *search–option–heuristics*, i.e. specifying queries by employing different search options, like controlled subject search, keyword title search, keyword subject search; *thesaurus–browsing*, to identify index terms; *screen–browsing*, i.e. browsing a list of terms alphabetically close to the user's search terms; *trial–and–error*, i.e. when searchers use whatever terms they have in their minds.

The latter two models provide a different perspective on how to classify and group users behaviors. Each pattern of behavior corresponds to a bunch of Bates' actions: for example, Ellis' "extracting" can be implemented by a journal run, an author search or a citation search; similarly, "thesaurus browsing" by Bates' term tactics.

5 Information seeking as explorative learning

It is not clear how the frameworks illustrated above could be used to support design and evaluation. Some are too general; others provide interesting clues on how user activity is organized, but cannot be directly used in evaluation. In this section I shall propose a level of analysis that, in my sense, has the potential to overcome such a limitation.

Consider the interaction between a user and an IAS that takes place during information seeking. For example, a user performing a zooming operation on an explicit set of documents in order to derive a sorted list of content–bearing terms that occur frequently in the set. Such an can be viewed at two levels of abstraction. The first level, called *presentation level,* refers to the direct interaction between the user and the IAS: the user perceives and interprets the state of the IAS from its presentation on output devices (screens, speakers, printers) and acts on the controls that the IAS offers. At this level of analysis one can decompose the zooming operation into a sequence of actions on the controls offered by the user interface of the system and interpretations on what is being displayed on the screen. The second level of abstraction, called *navigation level*, refers to the user interacting with the information resource. Via the IAS, the user explores a portion of the resource, moving in this virtual space of interconnected information items. It is as if the user were equipped with new sensors (enabling recognition among others of possible paths, of promising paths, of relevant information items, of current position) and with new actuators (enabling the move on a path, backtracking to a previous point, the close view on some information item). Using these sensors and actuators (offered by the user interface of the IAS) the user explores and experiences the virtual space defined by the information resource. The user needs to be able to interpret what is viewed in terms of useful, relevant, or topical documents, passages, terms, links or citations. The user needs to be able to spot and recognize different areas in the space, like clusters of useful, relevant, or topical information items, or like different result sets previously found; different routes, like the one followed to arrive to a certain point, or backtrack points of routes not completely followed so far, or dead–end routes leading to no useful results. Following the previous example on zooming, at the navigation level it can be decomposed into subactivities of selecting the starting documents and interpreting, in the context of the problem being solved, the resulting term list.

I claim that this second level interaction is the one that matters most to a user. The IAS and its user interface are simply the means through which the user sees and moves in the space. This situation is, in my view, rather peculiar to IASs. In other interactive settings, such as a spreadsheet used to manage a family budget, this second–level interaction is absent for the task itself is more or less known to the user. The user might explore the user interface to learn the program's features and underlying functionalities. But the user will not use the spreadsheet to learn how to manage a budget, or to explore different ways to manage it. *What–if* analyses, typical explorative activities performed with a spreadsheet, are not explorations of activities of the task. Rather, they aim at understanding how certain parameters are related.

Notice that, at any point in the information space, what can be seen and acted upon by the user depends on both the infrastructure underlying the information resource (i.e. available information items and their interconnections) and the tools supporting visualization and navigation. On a hypertext system, most interconnections are explicit and navigation is explicit as well: visualization is based on rendering of pages and links and navigation actions consist essentially of selecting links. With an information retrieval system, on the other hand, almost all the links are implicit, and need to be computed on the fly. For example, relevance feedback computes a virtual link relating a set of documents and a query to another set of documents; Ingwersen's ZOOM computes a link between a set of documents and a list of statistically significant terms; performing a journal–run can be viewed as following a link from a document to its journal and then other links to documents of other volumes of that journal. In order to support navigation, the IAS has to present (at least) the current position of the user in the space and it has to make explicit the routes passing through that position.

Consider the models described in section **??** under this navigation perspective. Bates' concepts of tactics and stratagems provide the infrastructure of the information space: each tactic or stratagem is a link or a short chain of links. Ellis' behavioral model and the categories identified by Chen and Dahr can be viewed as criteria to classify and group the links departing from a given navigation point. These categories can be used as ways to present different alternatives to a user which is stuck in some place. For example, one suggestion given to a user could be to perform a "differentiation" operation, to filter out some material according to criteria not yet used. Such a suggestion could then be implemented by following appropriate links, to a citation index to get frequency of citations, to a journal description to get its rejection rate. In this way users would have the opportunity to learn how to navigate in the information space.

Secondly, with respect to the problems described in section **??**, the navigation metaphor appears to be fairly general. In fact, in cases where the ASK hypothesis holds, the information seeking process can be characterized as exploration of the information space with the aim of learning something about it (for example its structure over a certain area; whether specific items exist within an area; what kind of items exist in that area). Information seeking can thus be viewed as explorative learning. In the cases where the ASK hypothesis does not hold, like specific fact finding, information seeking process can nevertheless be viewed as a navigation in the space: through known routes towards known destinations.

5.1 Usability factors

In general, usability of an IAS is affected by three components: user interface, information resource and tools for visualization and navigation.

Usability certainly depends on factors related to the navigation level. A usable IAS would provide many means to perform a productive navigation/exploration and would support decision making steps during those processes. As a consequence, users would be able to effectively explore portions of the space and gather results, if there are any. Effectiveness would not be measured only with parameters like precision and recall, as they are system–oriented. But instead it would be much more closely related to usability.

In addition, since navigation is implemented by the user interface of the IAS, usability at the presentation level strongly affects user's ability to navigate. At the presentation level, all the factor groups traditionally associated to usability apply [**?**]:

- *learnability*, including *predictability* (support to determine the effect of future actions), *operation visibility* (how the availability of operations which can next be performed is shown to the user), *consistency* (likeness in input/output behavior arising from similar situations), *generalizability* (support to extend knowledge of specific interaction to similar situations), *familiarity* (the extent to which user's knowledge in other domains can be applied to the current interaction);

- *flexibility*, including *dialogue initiative* (allowing users freedom from artificial dialogue constraints), *multi–threading* (support of more than one interactions at a time), *task migratability* (ability to pass control for a given task, for example by defining macros), *customizability* (modifiability of the interface by the user or by the system), *equal opportunity* (ability to interchange input and output: "if you can see it, you can use it");

- *robustness*, including *observability* (ability of the user to understand the state of the interface from what can be perceived), *reachability* (possibility to move from any given state to any other state of the interface), *persistence* (duration of the effect of a communication act and the ability of the user to make use of it), *recoverability* (ability to take corrective actions once errors have been recognized), *error prevention* (for example via the principle of "commensurate effort": if something is difficult to undo, it should have been difficult to do in the first place), *responsiveness* (response times of the system should be consistent and match user expectations)

At the navigation level, usability is related with the navigation and exploration tasks. Usability problems will be problems affecting such tasks, hindering users capability to understand where in the space they are and how they can move within it. Users should always be able to respond to questions like: where am I? what am I viewing? how is it related to my goal? where can I go? how did I came here? am I getting closer to my destination? am I making progress? what has been the effect of my last navigation action(s)? what will be the effect of action X?

Factors related to usability at this level include:

- *observability*, that is the ability of the user to understand where she is in the space from what she perceives on the interface. A relevant checklist for determining observability is: how did she arrive at that point, how close to the destinations she is, what was the effect of her last action(s), if she made progress towards her goal, what are currently active backtrack points, why they were chosen, how easy is to scan the hit list or a full document. A crucial role here is played by visualization techniques that can be used to present to the user the information space and navigation state(s).

- *operation visibility*, that is how the availability of routes which can be chosen is shown to the user and how comprehensible they are to the user. Relevant questions are: where can she go? How can she have an idea of where a route will bring her? Facilities to group and classify retrieved, viewed and selected information items (documents, passages, terms) are essential as many operations actually depend on these sets (e.g. relevance feedback, query refinement, zooming). Furthermore, all necessary referral aids should be made available and understandable, like citation indexes, classification structures, etc.

- *predictability*, that is user ability to guess and predict the effects of a certain action. For example, predicting the effects of a zooming operation in full detail is practically impossible, but users could learn and then exploit the fact that it will return content–bearing terms occurring frequently in the given document set. Similarly, predicting the effects of a relevance feedback operation could simply lead to the conclusion that more documents similar to the selected ones will be retrieved.

- *reachability*, referring to user ability to navigate from any state in the information space to any other. For example an information retrieval system offering only an automatic query refinement mechanism would prevent its users to get to arbitrary states, as the only reachable ones would be those that can be reached via the feedback mechanism and automatic query refinement. Reachability is also affected by the capability to resume navigation from a backtrack point.

- *recoverability*, that is the ability to reach a desired state after recognition of some error in previous interaction. Navigation errors include getting into dead–ends (no hit situation, no appropriate terms), getting into a too–many hits situation, failing to get useful results, looping around a cycle of ineffective states (for example by repeatedly tweaking with addition/removal of query terms, one of the problems observed in [**?**]). Recoverability is improved, for example, by constructive strategic suggestions provided by the system to the user giving an indication of a tactic or stratagem that could be used to overcome the problem.

- *consistency*, that is likeness of system behavior in similar situations. For example, actions like journal runs should be executable in the same way on journals, conference proceedings or article collections. Zooming should be possible on any set of full–text documents, passages or document surrogates.

- *customizability*, that is the ability to define shortcuts to reach certain states or to follow a chain of links. For example, to speed up navigation through fixed routes, like when one monitors a subject area.

5.2 Usability evaluation methods

>From a preliminary analysis, still waiting for a strong evidence, it appears that usability inspection methods like heuristic evaluation and cognitive walkthrough can actually be used upon this framework. For example, heuristics used in the former method can be applied to the navigation level of an IAS as follows:

1. "visibility of system status" refers to observability and operation visibility;

2. "match between system and real world" refers to the way the navigation metaphor is presented to the user. From the perspective of this heuristics, a usable IAS would promote, in a natural way, the development of a user model of itself based on concepts of navigation, routes, interconnected information items.

3. "user control and freedom" refers to predictability, recoverability and reachability. Such a system would be easy to "steer" during navigation to suit user's desires;

4. "consistency and standards", refers to consistency and standards adopted in the representation and manipulation of information;

5. "error prevention" refers to predictability (to avoid selecting inappropriate routes) and operation visibility (to avoid selecting non applicable actions);

6. "recognition rather than recall" refers to observability (by making explicit relevant parts of the navigation state) and operation visibility (by making explicit applicable routes and links destinations, i.e. by computing the link in background and presenting the results);

7. "flexibility and efficiency of use" refers to customizability and reachability;

8. "Aesthetics and minimalist design" refers to visualization techniques;

9. "help users recognize, diagnose and recover from errors" refers to observability, operation visibility and recoverability;

10. "help and documentation" refers to reachability of information about purposes, coverage, timeliness, referrail aids of an IAS.

6 Conclusions

The navigation and presentation levels enable analysis of the information seeking process at two levels of abstraction and group usability factors into two disjoint sets. This should simplify both the design and the evaluation processes of interactive information retrieval systems. I have tried to provide a preliminary discussion of some of the benefits of this perspective. It is however necessary to perform a much more detailed analysis of this framework, at least along the following lines:

- to define specific tasks (dependent on a certain specific information resource) and determine the actual support that the navigation and presentation levels give to design and evaluation of the information access system;

- to refine and run usability inspection methods and validate them by comparing their findings across different systems and with respect to usability problems found via user testing.

References

[1] M. J. Bates. The design of browsing and berrypicking techniques for the online search interface. *Online Review*, 13(5):407–424, 1989.

[2] M. J. Bates. Where should the person stop and the information search interface start? *Information Processing & Management*, 26(5):575–591, 1990.

[3] N.J. Belkin. Anomalous states of knowledge as a basis for information retrieval. *The Canadian Journal of Information Science*, 5:133–143, 1980.

[4] P. Borlund and P. Ingwersen. The development of a method for the evaluation of interactive information retrieval systems. *Journal of Documentation*, 53(3):225–250, June 1997.

[5] P. Borlund and P. Ingwersen. The application of work tasks in connection with the evaluation of interactive information retrieval systems: empirical results. In S. Draper and K. van Rijsbergen, editors, *Final MIRA Conference*, Glasgow, UK, 1999.

[6] G. Brajnik, S. Mizzaro, and C. Tasso. Evaluating user interfaces to information retrieval systems: A case study on user support. In *SIGIR96, 19th International Conference on Research and Development in Information Retrieval*, pages 128–136, Zurich, Switzerland, 18–22 August 1996.

[7] H. Chen and V. Dhar. Cognitive process as a basis for intelligent retrieval system design. *Information Processing & Management*, 27(5):405–432, 1991.

[8] A. Dix, J. Finlay, G. Abowd, and R. Beale. *Human–computer interaction*. Prentice–Hall, 1993.

[9] D. Ellis. A behavioural approach to information retrieval system design. *Journal of Documentation*, 45(3):171–212, 1989.

[10] D. Ellis. A behavioural model for information retrieval system design. *Journal of Information Science*, 15:237–247, 1989.

[11] C. H. Fenichel. Online searching measures that discriminate among users with different type of experience. *Journal of the American Society for Information Science*, 32(1):23–32, 1981.

[12] G.W. Furnas, T.K. Landauer, L.M. Gomez, and S. T. Dumas. The vocabulary problem in human–system communication. *Communications of ACM*, 30(11):964–971, 1987.

[13] D.G. Hendry and D.J. Harper. An informal information–seeking environment. *Journal of the Americal Society for Information Science*, 48(11):1036–1048, 1997.

[14] I. Hsieh-Yee. Effects of search experience and subject knowledge on the search tactics of novice and experienced searchers. *Journal of the American Society for Information Science*, 44(3):161–174, 1993.

[15] R. R. Larson. The decline of subject searching: Long-term trends and patterns of index use in an online catalog. *Journal of the American Society for Information Science*, 42(3):197–215, 1991.

[16] V. Manglano, M. Beaulieu, and S. Robertson. Evaluation of interfaces for IRS: Modelling end–user searching behaviour. In *Colloquium on Information Retrieval (IRSG) — Draft*, pages 137–146, 1998.

[17] G. Marchionini. Information–seeking strategies of novices using a full–text electronic encyclopedia. *Journal of the American Society for Information Science*, 40(1):54–66, 1989.

[18] Gary Marchionini. *Information Seeking in Electronic Environments*. Cambridge University Press, Cambridge, UK, 1995. 224 pp., (ISBN 0–521–44372–5).

[19] S. Mizzaro. How many relevances in information retrieval? *Interacting with computers*, 10(3):305–322, 1998.

[20] W.M. Newman and M.G. Lamming. *Interactive system design*. Addison–Wesley, 1996.

[21] J. Nielsen. *Usability engineering*. Academic Press, 1993.

[22] T. Saracevic. Relevance: a review of and framework for the thinking on the notion in information science. *Journal of the Americal Society for Information Science*, 26(6):321–343, 1975.

[23] T. Saracevic. Modeling interaction in information retrieval (ir): a review and proposal. In *Proc. of the 59th ASIS Annual Meeting*, volume 33, Baltimore, Maryland, 1996.

[24] B. Shneiderman. *Designing the user interface*. Addison–Wesley, 1998.

[25] M. Sullivan, C. L. Borgman, and D. Wippern. End–users, mediated searches and front–end assistance programs on dialog: A comparison of learning, performance and satisfaction. *Journal of the American Society for Information Science*, 41(1):27–42, 1990.

[26] M. B. Twidale, D. M. Nichols, G. Smith, and J. Trevor. Supporting collaborative learning during information searching. In *CSCL'95 Proceedings of Computer Support for Collaborative Learning*, pages 367–374, Bloomington, Indiana, 17–20 October 1995.

[27] C. J. van Rijsbergen. *Information Retrieval*. Butterworths, London, 2nd edition, 1979.

[28] A. J. Vollaro and D. T. Hawkins. End–user searching in a large library network: A case study of patent attorneys. *Online*, 10(4):67–72, 1986.