COGNITIVE PERSPECTIVES OF INFORMATION RETRIEVAL INTERACTION: ELEMENTS OF A COGNITIVE IR THEORY

PETER INGWERSEN

(pi@db.dk)

Royal School of Librarianship
Copenhagen, Denmark

The objective of the paper is to amalgamate theories of text retrieval from various research traditions into a cognitive theory for information retrieval interaction. Set in a cognitive framework, the paper outlines the concept of polyrepresentation applied to both the user's cognitive space and the information space of IR systems. The concept seeks to represent the current user's information need, problem state, and domain work task or interest in a structure of causality. Further, it implies that we should apply different methods of representation and a variety of IR techniques of different cognitive and functional origin simultaneously to each semantic full-text entity in the information space. The cognitive differences imply that by applying cognitive overlaps of information objects, originating from different interpretations of such objects through time and by type, the degree of uncertainty inherent in IR is decreased. Polyrepresentation and the use of cognitive overlaps are associated with, but not identical to, data fusion in IR. By explicitly incorporating all the cognitive structures participating in the interactive communication processes during IR, the cognitive theory provides a comprehensive view of these processes. It encompasses the *ad hoc* theories of text retrieval and IR techniques hitherto developed *in* mainstream retrieval research. It has elements in common with van Rijsbergen and Lalmas' logical uncertainty theory [1] and may be regarded as compatible with that conception of IR. Epistemologically speaking, the theory views IR interaction as processes of cognition, potentially occurring in all the information processing components of IR, that may be applied, in particular, to the user in a situational context. The theory draws upon basic empirical results from information seeking investigations in the operational online environment, and from mainstream IR research on partial matching techniques and relevance feedback.

By viewing users, source systems, intermediary mechanisms and information in a global context, the cognitive perspective attempts a comprehensive understanding of essential IR phenomena and concepts, such as the nature of information needs, cognitive inconsistency and retrieval overlaps, logical uncertainty, the concept of 'document', relevance measures and experimental settings. An inescapable consequence of this approach is to rely more on sociological and psychological investigative methods when evaluating systems and to view relevance in IR as situational, relative, partial, differentiated

Journal of Documentation, vol. 52, no. 1, March 1996, pp. 3—50

and non-linear. The lack of consistency among authors, indexers, evaluators or users is of an identical cognitive nature. It is unavoidable, and indeed favourable to IR. In particular, for full-text retrieval, alternative semantic entities, including Salton *et al.'s* `passage retrieval' [2], are proposed to replace the traditional document record as the basic retrieval entity. These empirically observed phenomena of inconsistency and of semantic entities and values associated with data interpretation support strongly a cognitive approach to IR and the logical use of polyrepresentation, cognitive overlaps, and both data fusion and data diffusion.

1. INTRODUCTION

THIS PAPER DISCUSSES THE BASIC ELEMENTS of a global cognitive theory for information retrieval [IR] interaction from a cognitive point of view. Within this framework are outlined the principles underlying the concept of polyrepresentation applied simultaneously to the user's cognitive space *and* the information space of IR systems. When seen from a cognitive perspective all of the interactive communication activities in IR and information seeking can result in processes of cognition, which may occur in *all* the information processing components involved. The reader is referred to Ingwersen [3, pp. 123–146] for a further discussion and empirical foundation of this cognitive view of IR.

The concept of polyrepresentation seeks to represent the current user's information need, problem and knowledge states and domain work task or interest in the form of contextual structures of causality. At the same time, it implies that we should apply different methods of representation and a variety of IR techniques of different cognitive and functional origin to the information objects in the information space. The goals are to improve the intellectual access to information sources and, simultaneously, to provide the IR system with an enriched contextual platform that can support the user's information seeking. The theory aims to reach these goals by means of cognitive overlaps of objects in information space. The bigger the difference in time and by type between the original cognitive interpretations pointing to the same information objects, the more plausible the usability of such objects in a given IR situation. Since the initial conditions are vital for the ensuing retrieval process, the formation of the information need is re-analysed by viewing it in the context of the searcher's cognitive space. The theory attempts to demonstrate that by making adequate and combined use of empirical and analytic evidence, from mainstream R research as well as from operational and user-oriented investigations, it is possible to equate text retrieval with retrieval of information in the proper sense of this concept.

The paper is organised in the following way. The first section presents and discusses briefly the concepts required for the establishment of a cognitive framework for R, that is, the properties of the cognitive viewpoint and its understanding of the concept of information. This is followed by an overview of the different cognitive structures of significance for IR interaction and a brief outline of the development of the dominant IR research approaches, each one focusing on specific cognitive structures. The next section views existing empirical evidence for analysis of the formation of an information need during IR interaction and the extent to which polyrepresentation can be incorporated in the cognitive space of users. These analyses lead to a distinction between four basic types of information needs, which are consequential for the understanding of relevance and the experimental situations when re-introducing the user into the partial match research tradition. The following section focuses on the application of polyrepresentation of the data structures in information space, in which the notion of 'cognitive overlap' and its relation to data fusion play a vital role. The section includes an analysis of the use of smaller semantic entities or passages for full-text information retrieval. For the first time in IR the 'logical uncertainty' principle for IR [1, 4, 5] is seriously discussed and is considered a valid approach compatible with the cognitive view. This view enables polyrepresentation, controlled diffusion and fusion of structures, within both the user's cognitive space and the information space, to provide a dynamic framework for information retrieval interaction. The concluding discussion summarises important conceptual and experimental consequences.

2. THE COGNITIVE APPROACH TO IR THEORY

The cognitive point of view in information science implies that each act of information processing – whether perceptual or symbolic - is mediated by a system of categories and concepts which, for the information processing device, constitute a world model [6]. In relation to information science and IR this epistemological view is discussed in detail and historically analysed by Ingwersen [3, pp. 15-48], separating it from the cognitivistic and rationalistic traditions by associating it with more socio-hermeneutic approaches to information transfer and knowledge communication.

Essentially, the kernel of the viewpoint is that both the reception *and* the generation of information are acts of information processing. The way this processing is carried out is dependent on the world model of the actor - whether human or machine. However, the cognitive viewpoint has for long been (mis)interpreted as a user and interface-centred approach in which the human searcher acts as the recipient of something (data, signs, potential information) generated by a 'system' incorporating knowledge sources. For instance, Brookes' Equation model [7] has commonly been understood in this limited way. However, the viewpoint suggests that the machine or the system may *similarly* act as a recipient applying its own world model. This interchange of positions makes the viewpoint a forceful theoretical foundation for IR interaction and human-computer interaction [HCI] in general. We may consequently analyse the properties of the world models involved on both sides of the communication channel as described below.

In human information processing the world model constitutes the individual cognitive space which, consisting of highly dynamic and interchangeable cognitive structures, controls the perception and further processing of external input, e.g. during communication and IR. The individual cognitive space and, in particular, the current cognitive/emotional structures and states are determined by the experiences gained through time in a social and historical context. In automatic (symbolic) information processing the world model is dynamic but *not* self-contained. It consists of the human cognitive structures embedded in the system prior to processing. Its individual cognitive structures, e.g. in the form of algorithms or textual symbolic strings, may interact with one another and with structures generated by human beings external to the system – when ordered and capable of doing so. However, the processing will only take place at a linguistic sign level of communication – never at a cognitive level; see Figure 1.

Only to human generators or recipients of communicated signs the perception may *also* take place at a cognitive/emotional level, transforming the current cognitive states into new states, conceivably providing information transformed into knowledge and cognition. During any act of human or computerised communication the viewpoint regards all communicated messages as signs transferred at a linguistic level, which may be transformed into information at a cognitive level only via perception and mediation by the individual recipient's current cognitive state.

Consequently, this view implies an immediate cognitive 'free fall' into the lower levels of semantic, morpho-syntactic and lexical linguistic nature. The consequence is that any intentionality and meaning underlying the communicated messages are immediately lost and have to be rebuilt, i.e., interpreted by the recipient by means of those presuppositions which enable participation in the communication act. Computers (or books for that matter) hold predefined and fixed presuppositions, whilst those of human beings are individually unpredictable, formed as they are by episodic, semantic and emotional experiences.

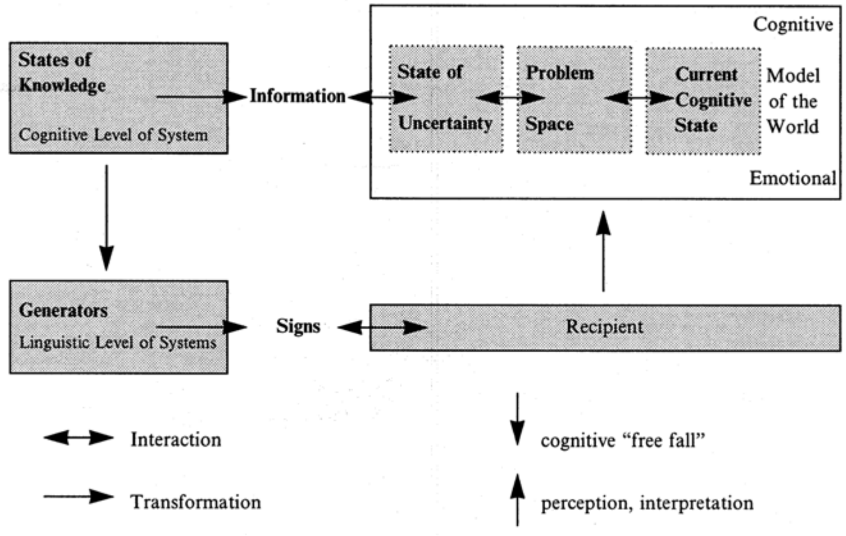

FIGURE 1. *The cognitive communication system for information science and* IR. *Transformation of Ingwersen [3, p. 33]*

Figure 1 illustrates these processes. At the linguistic level of communication we have the signs, e.g. the spoken words, images, or text strings*, which are generated by transformations of the generator's cognitive structures. As mentioned above, such generative acts of information processing may be done by human beings or machines alike. The communicated signs stay at this linguistic level up to that point at which a recipient's cognitive state may perceive and interpret them according to its world model and the actual problem space and state of uncertainty – the latter state being equal to an anomalous state of knowledge (ASK) [8]. For instance, if the recipient is human this interpretation of meaning *may* transfer into information in the real (rather than the metaphorical) sense of the concept, at the cognitive level. This is due to the fact that the cognitive structures are dynamic *and* self-contained. In the case of a machine as recipient the problem space and state of uncertainty may indeed exist, but its cognitive state is predefined and fundamentally static, and similar comments apply in the cases of neural networks or other seemingly dynamic electronic structures. The incoming signs are simply entirely, partly or not recognised, and *may be* interpreted (as in machine translation) by the predefined cognitive structures at the linguistic surface levels. No interpretation or transformation can take place by some self-contained structures because they do not exist, and no change in cognition can then occur.

Following the cognitive view the concept of information, from the perspective of information science, has to satisfy two conditions simultaneously [3, p. 33]: on the one hand information being the result of a transformation of a generator's knowledge structures (by intentionality, model of recipients' states of knowledge, and in the form of signs); and on the other hand being something which, when perceived, affects and transforms the recipient's state of knowledge.

The satisfaction of these two conditions means that information goes beyond meaning. A sentence, or indeed an image, may have several meanings or semantic values and each meaning may provide many different interpretations and transformations of state of knowledge, depending upon the situation of the recipient. In machine translation the attempt is to establish the one correct linguistic meaning (semantics) by means of context. In IR the goal is to give access to such plausible meanings or values as, given the situation, may entail transformations of a cognitive state, thus providing information in a pragmatic sense via context. The information is associated with the context, and *not* just with the individual semantic value. Operationally, for instance, the success of the printed KWOC indexes was mainly due to their associative capabilities.

Thus, the implication for the interactive communication process taking place between man and machine is that messages sent by machines or man to a human being can become information in the real sense; however, signs communicated by machine or man to machines can never become information, although they are perceived and affect the embedded cognitive structures. Such signs stay signs – or remain as ‘potential information’ – at a given linguistic surface level; see Figure 1.

*Throughout the paper ‘text’ signifies text and all other media-borne messages, such as images, video, films, spoken messages etc.

With the cognitive viewpoint in mind we may hence observe two fundamental characteristics of importance to IR:

- the uncertainties and unpredictabilities inherent in IR interaction;
- any presuppositions, meaning and intentionality underlying the communicated messages are vital but constantly lost.

Uncertainty and unpredictability are associated with the acts of interpretation during the effort to reach a cognitive communication level. These factors are, both in theory and in practice, individual and apply to man as well as system – as senders, recipients and processors. In the case of human interpretation this act requires more than just the ability to encode and decode. It requires a common semantics. From a different perspective Sparck Jones has likewise pointed to similar uncertainty factors [9]. The immediate loss of intentionality, presuppositions and meaning is crucial since these original properties are vital for the perception and understanding of generated messages.

The two fundamental characteristics are the major reasons for inter-indexer inconsistency, inconsistent assessments of 'relevance', searcher differences for identical queries, and the different results obtained by the variety of IR techniques. The latter lack of consistency is evidently founded in the fact that no interpretation takes place at all. Hence, direct and real information retrieval – as opposed to text retrieval performed at a low morpho-syntactic linguistic or single-word level – can only be carried out by the individual user himself. IR techniques will thus be most successful when acting as instruments that support the human retrieval of information.

The cognitive model, Figure 1, has serious consequences for the understanding of IR and the research in the field. The logic of the viewpoint points to the potential value of exploring the multifunctional and cognitive variety of representations inherently existing in, extracted or interpreted from information objects *and* from the cognitive space of a user. As far as possible, such explorable structures ought to demonstrate contextual properties in order to provide better clues for perception and interpretation, and to reduce the uncertainties as well as the unpredictabilities in IR on both sides of the communication channel. This exploration forms the initial elements of a cognitive theory for IR.

3. COGNITIVE STRUCTURES IN IR INTERACTION

In the suggested epistemological framework cognitive structures are manifestations of human cognition, reflection or ideas. In IR they take the form of transformations generated by a variety of human actors, i.e. belonging to a variety of different cognitive origins; see Figure 2: systems designers and producers, IR technique developers, indexing rule constructors, indexers, authors of texts and images, intermediary mechanism designers, and users in a domain-related societal or organisational context. Consequently, in the system setting an IR system designer's cognitive structures may be represented by specific database architectures and one or several matching algorithms or logics, e.g. the vector space and probabilistic models. Human indexers' cognitive structures are represented

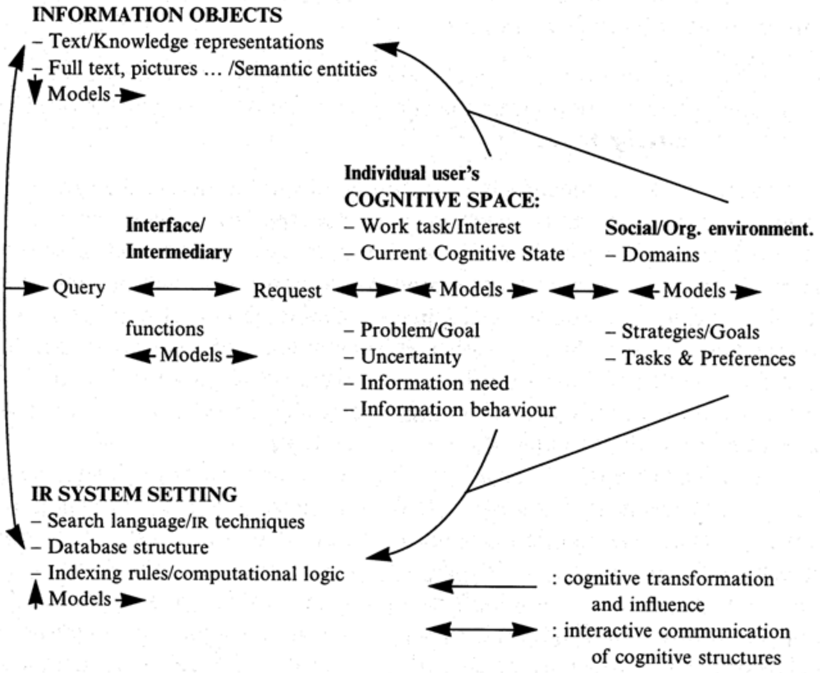

FIGURE 2. *Cognitive model of IR interaction. Extension of Ingwersen [3, p. 16]*

by the index terms added to the original information objects, these essentially being the result of an intellectual interpretation of an author's text or images, often guided by predefined rules and a thesaurus the semantic relations and knowledge representation for which other human beings are responsible. In automatic text representation any different weighting function or partial match technique that is applied represents a transformed cognitive structure. The author's text, including titles, captions, headings or cited works are representations of cognitive structures intended to be communicated as information objects. Later citations to that particular text imply a different cognitive origin, viz. interpretations by other human beings. Specific portions of the texts, e.g. titles, abstracts, figures, the introduction, or the full-text sections demonstrate different functional origins. Each document type displays different functional styles, as does each domain, and should be treated differently. In the model of Figure 2 the vertical interactive communication between System Setting and Objects takes place solely at the linguistic sign level. The interaction may momentarily reach a cognitive level only in the case of intervention by a human indexer. The intentionality, however, underlying the added index terms is immediately lost, according to this viewpoint.

The manipulation of user requests into query formulations during request model building and retrieval by an intermediary (whether human or computerised) represents additional cognitive structures, as shown in the centre of Figure 2. The cognitive structures embedded might be, for example, the Monstrat functions [10], based on the extraction of underlying problem descriptions, or the more comprehensive Mediator Model functionalities [3], such as the Domain and System Models, the Feedback Generator, User Model, and the Request Model Builder; the latter are also found in the I3R system [11]. However, at present the so-called 'intelligent retrieval' approach seems to have come to a standstill, being replaced by a more modest supportive attitude [3, 12].

At the right-hand side of Figure 2 are outlined the major different cognitive structures of individual users which may be represented to intermediary mechanisms and IR systems, e.g. the actual work task or interest leading to a current cognitive state which may end up in a problem or uncertainty state for the actual user. These mental activities take place in the context of a domain(s) of epistemic, social or organisational nature which, in addition to influencing the current searcher in a 'historical' socio-semantic sense (Figure 2, the extreme right-hand side) also constantly influences the authors of texts and systems design attitudes. The most simple form of a domain is an academic subject field, which is essentially a social construct represented by the collective cognitive structures of the individuals [13] forming the field. Other forms of domains are, for example, industrial sectors, individual firms or organisations, or professional groupings, such as journalists. The horizontal interactive processes shown in Figure 2 display communicative properties that are different from the vertical ones because of human interference, in accordance with the cognitive view discussed above.

Messages communicated to the system, also via an intermediate mechanism, remain constantly at a linguistic surface level. The intermediary mechanism may interfere at a very simplistic level, e.g. as done in partial match retrieval by 'analysing' the incoming request and assigning weights. The degree of uncertainty will thus be high. The interference may take more elaborate forms, for instance by means of request model building that may produce contextual structures, but these will still be at a linguistic level of communication: only the first condition of the information concept defined above is satisfied. Messages communicated from the system, including to and from an intermediary mechanism, remain at a linguistic level until they conceivably transform a human cognitive state by turning into information. At the individual level the uncertainty may be high, depending on the degree and adequacy of context provided by the system (and the mechanism). In the case of human intermediaries the uncertainty may also be high but for different reasons: the double act of interpretation performed by the human mediator and the final one performed by the end user. This may often give ca se to increased uncertainty within the cognitive space of the user during the process.

Our approach is in line with Kuhlthau's empirically based results of, and model for, information seeking [14] and Rasmussen *et al.*'s cognitive engineering approach to systems development by means of cognitive task analysis [15].

3.1. IR research – traditions in transition

Since the mid-seventies two research approaches to IR have been dominant: the early experimental mainstream tradition, which is deeply rooted in the natural sciences' rationalistic views of information transfer, and the younger user-oriented tradition, which is rooted in communication theory and socio-psychological methodologies.

These paradigmatic research traditions have often been criticised, most recently by Ellis [16, 17] who calls the former a 'rationalistic' paradigm because of its allegiance to computer science and its data-driven characteristics. The research tradition is mainly criticised for its lack of theoretical coherence and substantial results. Similarly, the user-oriented or, as Ellis puts it, the 'cognitive' paradigm demonstrates identical problems of *ad hoc* theory building and fragmentation. In a Kuhnean sense, we agree upon this pre-paradigmatic nature of IR research and theory building. However, as stated previously in connection with Figure 1, it seems a current misconception of the cognitive viewpoint to postulate that the user-oriented tradition mainly relies on such a view [18]. Although this R&D tradition sometimes adheres to cognitive psychological theories and methods, very few user-centred studies actually adopt a comprehensive and global cognitive perspective as their theoretical foundation. The great bulk of the published research insists on the user-driven perspective alone, or totally omits any perspective. The pragmatic research attitude is rather pre-dominant in IR.

3.1.1 The mainstream IR research tradition

By means of laboratory experiments taking place over more than three decades, the mainstream IR research community has continuously attempted to maximise and compare the performance of a variety of *ad hoc* retrieval techniques, essentially by keeping constant the information objects and a set of predefined queries. Fundamentally, the research is limited to the structures and vertical interaction at the extreme left-hand side of Figure 2. The experimental settings generally force the researchers to apply well and predefined requests and presuppose static information needs in order to establish a known ratio of relevant items per defined query. Any assumption that the mentally intrinsic need might indeed be variable during IR sessions would destroy the idea of known recall. This assessment procedure is most often carried out by one domain expert or the researcher himself – not by a panel. The remaining cognitive structures shown in the model are hence not taken into account: indeed, they do not exist in the setting. With the exception of van Rijsbergen's logical uncertainty principle [1, 4, 5], and Blair's suggestion of applying the later Wittgenstein's language philosophy as a foundation for representation in IR [19], all the statistical, linguistic or network founded IR theories are content and data-driven at a linguistic surface level (Figure 1) and often at the lowest lexical level. In other words, although in general mathematically sound, the theories are basically just for text retrieval – not for information retrieval. Besides, they are isolated from and competing with one another. Another common assumption seems to be that words in texts that are retrieved correspond to information. This is the fundamental reason for counting and weighting words, that is, signs. However, the best performing weighting and similarity/probability techniques could easily be applied in adequate (intentional) combinations and data fusions to retrieve data which, by *human interpretation,* may eventually be transformed into real information of value in a seeking situation.

The recent, large-scale TREC experiments actually allow individual users to participate in the IR process in order to perform ‘topical relevance’ assessments, provide relevance feedback and query modification, but the experiments still employ fixed, well-defined and invariable requests [20]. Also the Okapi experiments [21] move into the right-hand side of the model by allowing for real-life, often ill-defined, requests to be processed by the probability ranking technique. It is not the exclusion of real-life information needs from the closed settings that poses problems, but the inclusion of such needs, the uncontrolled uncertainty, and the ensuing variety of user interpretation. By introducing the cognitive space of the individual user, such experiments increasingly must rely on combinations of quantitative and qualitative investigative sociological methods, analogous to those in the user-oriented research approach. In particular, if comparison of performance between different IR techniques is to be carried out, the new experimental situation poses severe problems, since the number of variables increases and the situation becomes quite complex.

3.1.2 The user-oriented tradition

This tradition encompasses two main sub-traditions: the operational R&D approach; and the information seeking research tradition. In both traditions real information is understood as a result of human interpretation of data sources during communication and information interaction. Examples of these traditions are Belkin and Vickery’s contribution [22] and the continuous discussion of user-centred ‘relevance’ going as far back as Cuadra and Katter [23] or Saracevic [24] and lately re-opened empirically by Su [5], Bruce [26] and Barry [27], as well as theoretically and thematically by Schamber, Nilan and Eisenberg [28] and Harter [29]. The operational research is mainly limited to the individual cognitive space of users and its interaction with online systems through some form of intermediary or interface (Figure 2, centre and towards the left-hand side). Information seeking research is concentrated on the right-hand side of the model by including the social impact of (scientific) domains in the behavioural processes. With the exception of the operational Boolean systems, the variety of retrieval techniques shown at the left-hand side of Figure 2 are generally ignored.

However, the operational environment has produced profound empirically-based studies of how people may benefit from controlled indexing vocabularies, natural language representations (NLR), or their most probable combinations, e.g. Katzer *et al. [30],* Tenopir [31] or Lancaster [32]. Also, the community has contributed valuable results concerning user performance and behaviour during online retrieval sessions, involving both human intermediaries and end-user searching [25, 33, 34]. These investigations bridge the operational and the information seeking environments. However, mainstream R&D has not made use of these investigative results nor their methodology. Ignorance of online techniques cannot be the explanation. Similarly, a conscious rejection of the Boolean exact match technique, because of its ‘lower’ IR performance compared to other matching principles, must be regarded as a rather arrogant attitude that is not appropriate for a scientific community. A more logical reason is merely the fact that mainstream R&D has deliberately eliminated the disturbing influence of human indexers, intermediaries and users for the sake of experimental control. The impartial – but human – assessor alone is allowed to participate.

The information seeking community dates back to Waples [35]. In this environment we find domain-associated investigations of (scientific) users' information behaviour, such as the empirical studies of Ellis [36] and Brittain [37]. Early use and user studies are seminal [38]. One may observe the domain analysis approach, put forward recently by Hjørland and Albrechtsen, which is associated with methods of representation [39] (Figure 2 upper arrow). Further, there exists a link, which is currently being strengthened, to bibliometrics and scientometrics (Figure 2 upper circle of communication). The main approach in this tradition, however, is to produce general analytic and behavioural studies or models of information seeking and transfer. Bates, among others, represents this school, e.g. by producing her idea and search tactics [40, 41] and the exploratory paradigm [42], leading to the browsing and berry-picking information seeking behaviour (or technique) [43]. Kuhlthau's recent investigations dealing with cognitive uncertainty during the seeking process marks a step forward towards a deeper understanding of information behaviour [14].

3.1.3 Towards a cognitive turn

Evidently, there exists a mutual understanding between the two user-centred communities, whereas the mainstream, system-driven environment generally seems to float in splendid isolation so that the two predominant research communities do not really explore the ideas, methods and results of each other. With some exceptions, at least there exists an inherent and silent agreement that information behaviour, seeking and retrieval mainly take place among academics.

This current situation of mutual ignorance between the central mainstream and the user-oriented research communities is quite unsatisfactory. Meanwhile, the technological achievements made in other fields of IT will supposedly provide a second chance for the research communities to explore their common interests, e.g. the full intrusion into IR of world-wide networks combined with full-text publications online, including non-scientific and complex grey sources of information, as well as all sorts of multi- and hypermedia applications. Also the introduction of simplistic partial match techniques allows the ranking of items in commercial online systems, for instance, the Target facility or Status IQ, which will force a mutual relationship to develop. With the aim of re-introducing human searchers into the mainstream IR R&D in mind, and of exploring advanced IR systems within the user-oriented community [44], the comprehensive model, shown in Figure 2, is an attempt to provide a workable framework.

4. IR PHENOMENA AND CONCEPTS IN A GLOBAL COGNITIVE PERSPECTIVE

Based on the cognitive view, the conception of information and substantial empirical evidence, our approach is first to examine the mental formation of the information need. As the starting point for IR interaction the understanding of this formation is seen as a central issue in IR. This phenomenon leads to an analysis of the polyrepresentative nature in terms of causality of the user's cognitive space. The information need itself is divided into four distinctive forms based on the degree of its intrinsic stability and mode of definition. Formation and the polyrepresentative properties have consequences for the experimental settings applied in R research as well as for the conceptions of relevance.

On the system side a variety of evidence demonstrates a different pattern of polyrepresentation in the information space which may be explored and utilised in the form of cognitive overlaps. Within this framework the notion of 'document' is discussed in the light of smaller semantic entities for full-text (and multimedia) systems, and polyrepresentation of such information objects is exemplified.

4.1. The formation of the information need

The development of a desire for information is a result of communication, sensing or thinking processes. The outcome of such phenomena is a situation which makes a person's cognitive state process the circumstance in such a way that it recognises a current 'incompleteness or inadequacy', as stated by Mackay [45] or a knowledge gap as suggested by Dervin and Nilan [46]. With Hjørland one may state that the formation is a result of cognition processes [47].

A fundamental question is whether or not we should regard the notion of the information need as an outdated construct that should be replaced by the notion of a dynamic cognitive state, as argued by Harter [29, pp. 606-607). This replacement seems at first glance a reasonable solution, but it is inconsistent with most problem solving and information seeking theory and empirical results [14, 22] since it implies that in all circumstances *all* the searcher's cognitive structures are simultaneously dynamic and variable. Evidently, this position excludes the possibility of having some mental structures remaining stable for a while. A certain degree of cognitive stability in some mental structures is present since we in IR clearly are dealing with a typical breakdown situation' in the socio-hermeneutic sense of Winograd and Flores [48] which may be solved by a mental step-by-step mode. In our view we define the current cognitive state as the currently known which, drawing upon emotions, tacit knowledge, and external factors influencing the individual searcher, is indeed regarded as a variable state; but we allow for a periodic stability of mental sub-structures concerned with the domain work task, the problem state and the intrinsic information need, see Figure 2. For instance, the information need is quite stable (at least for a while) for information requirements of a verificative or conscious topical nature [3, 49], in which the data or topic searched for are known. Some of the current cognitive structures will change according to the outcome from the system. Then, the problem and uncertainty states may also become unstable, e.g. if the unexpected or nothing is retrieved. According to the degree of exploratory intentionality, or open mindedness, the original desire for information may shift position into a new periodically stable one [13] (Figure 3). However, the same stable work task/interest may be present during the entire retrieval session.

	Well-defined	(Internal) Ill-defined
	(External behaviour: Search loops)	
Stable	1 **Rich, variable, cognitive state** Conceptual 'throwness' Limited uncertainty Topical relevance assessm.: yes Curiosity: low Confined navigation	4 **Weak, variable, cognitive state** Conceptual 'breakdown' High uncertainty Topical relevance assessm.: no Curiosity: low Marching on the spot (dead ends)
Variable	2 **Rich, variable, cognitive state** Conceptual 'throwness' Controlled uncertainty Topical relevance assessm.: yes Curiosity: high Exploratory navigation	3 **Weak, variable, cognitive state** Conceptual 'breakdown' High uncertainty Topical relevance assessm.: no Curiosity: high Browsing
	(External behaviour: Berry-picking)	

FIGURE 3. *Matrix of four distinct forms of human intrinsic information needs*

In addition, we may expect information needs that are constantly variable in nature. In our framework the current cognitive state (e.g. the little known about what is desired), and the problem and uncertainty states would similarly be variable, invoking learning processes. The assumption is, however, that the internal perception of the work task still remains quite stable. In 1960, Mackay pointed to a reason or causality further underlying the current incompleteness, uncertainty state, ASK, knowledge gap, or problem situation: 'a particular area of interest' [45]. We regard this area of interest as identical to the individual's current cognitive work task or interest. Thus, this task may deal directly with a person's actual work within a specific domain, e.g. in computer science or within a software house, to carry out the design of knowledge bases for AI purposes. It may also simply be concerned with an 'interest' not directly related to working conditions but, for instance, be associated with culture, entertainment or hobby domains, or simply with curiosity. It is the task/interest which, strongly influenced by the domain and dominated by the individual intentionality and cognitive state, causes a problem state and an information need to emerge.

This understanding of the nature of the information need formation in IR interaction provides an evolutionary extension or specification of Belkin's ASK hypothesis [8, 50]. It explicitly incorporates additional cognitive and social structures underlying and functionally causing the formation to happen: the 'work task/interest' associated with a work 'domain' set in a social context (Figure 2). Hence, in our cognitive view a 'dynamic cognitive state', a 'problem state' or an ASK cannot replace the notion of the 'information need'; and a 'problem' is not sufficient as an intrinsic trigger for information seeking to occur. Obviously, the problem state should be included since it provides some additional clues to the *reasons* for acquiring information. Further, the notion of information need should be kept since it plays a vital role as an experimental constant in the mainstream research (as discussed in section 4.2.1 below).

4.1.2 Empirical investigations supporting the information need formation

Belkin *et al.'s* early investigations (protocol analyses) of the 'problem state formation' provide us with insights into two basic underlying phenomena: (1) the vagueness or uncertainty associated with people's formulations of their information need and, interestingly enough, also associated with the problem statements made during pre-search interviewing. Since Belkin's experiments were limited to the pre-search stage and did not include system interrogation [10, 51] he did not obtain evidence of the influence of systems feedback on the development of the problem state; (2) the realistic possibility of acquiring data concerning the 'current cognitive state', 'domains' and 'tasks' during IR interaction. Saracevic *et al.'s* large-scale investigations of online system-intermediary-user interaction incorporate exactly the dynamic influence of continuous feedback from the system [52]. In these studies one may directly observe the structures mentioned previously surfacing in richer forms and evolving through the interactive IR processes. Also Su's recent empirical study of kinds of relevance assessments exemplified by the user comments [25] provides strong evidence for the existence and dynamic nature of the cognitive state and other underlying cognitive structures. Bruce's experimental methodology makes it possible to trace directly the impact of different cognitive foci of attention during such assessments, that is, the dynamic change from topical relevance judgements into different situationally dependent assessments. Both investigations have in common the fact that such judgements are carried out by the cognitive state in co-operation with the domain, task and problem related structures, and the current state of the information need, and not by the need alone.

The matter is epitomised by a re-analysis of Ingwersen's sets of original thinking-aloud protocols [13] where identical patterns emerge. The protocols contain a great number of assessments and extensive relevance feedback based on the retrieved and presented full-text material in printed form. By means of the experimental setting, which initially made the end users browse the shelves with their own needs while thinking aloud prior to a system-librarian-user interaction, we may observe the cognitive formation of the information need as well as related uncertainty and problem states. In some cases the starting point is an intrinsically very vaguely formed information need which develops into a requirement for information well known to the user (and the experimenters) when finally encountering the intermediary. However, an internally well-defined and seemingly stable need often becomes a twisted, reduced or ill-defined formulation when interviewing and searching commerce. This problem of verbal reduction (the label effect) is mainly caused by the incompatibility of the models and expectations that the participants have of each other. The protocols, in addition, demonstrate vague and uncertain types of the internal information need, which remain intrinsically ill-defined during the end-user's independent seeking behaviour (Figure 3). Consequently, the communicated information request is similarly ill-defined and vague, and resembles the former labels. What appears to be known to the user are just the reasons for and problems connected with acquiring information.

The vital activity during interaction occurs when passages of information sources, indexes, tables of contents, etc. are presented to the user (and the intermediary). These tripartite interactions force the user either to reformulate an already conscious need and the problem state, or to formulate his underlying reasons for obtaining information in the case of an intrinsically vague need [53]. In addition, the feedback from the system encourages the intermediary to ask certain types of questions, particularly when the sources do not provide the requested information. The protocols demonstrate clear-cut statements and assessments associated with the underlying elements of ‘cognitive state’, ‘work task’ and ‘domain’, as exemplified by the user comments from Su’s investigation [25] and by the relevance categories most frequently applied according to Barry [27, p. 157]. For instance, in a protocol [13] concerned with searching for ‘an appropriate form of Boolean logic formulae’ (the user-defined information need) the user initially put up with request labels consisting of ‘Boolean logic’ including synonyms. Via feedback from the system the user decides to move on to other possible subject areas or domains containing ‘Boolean logic’ from the philosophical area to the mathematical domain. These areas are known to the user and form part of his current cognitive state. During this process many text entities are *explicitly* judged relevant to the ‘topic’ of ‘Boolean logic’, but only partially or non-relevant to the actual need/problem or the work task; for example, ‘this kind of formula is not what I want’. Since the preferred mode of formula obviously is not displayed in a suitable form in these full-text sources, the intermediary finally asks for the user’s current ‘field of (work) interest’. The answer, ‘applied computer science’, signifies the domain and triggers the user further to explain his or her underlying work task: the appropriate form of the formula is associated with ‘electronic circuits when repairing and designing computer systems’. Since a question like ‘what are you going to use the information for?’ had never occurred during the search interview, the user had probably not previously been able to supply any details of his existing and quite stable intrinsic knowledge about his task. Cognitively speaking, we may assume that the variety of relevance assessment categories empirically extracted from users [24, 27, 28] corresponds to the variety of the formation states or foci of the information need evolution.

The above empirical investigations of mental events taking place during information seeking demonstrate that the initial ‘state of uncertainty’ concerned with an information need momentarily becomes even more uncertain and variable during the initial phases of retrieval, mainly due to the nature of the feedback from the system and its interpretation [14]. Also in relation to complex work task operations, Byström and Järvelin have found that the types of information need may be very different, ranging from well-defined ones in less complex tasks to vague or non-defined needs (as well as problems/goals) in the case of very complex tasks. What then seems defined are only the task itself and the domain concerned with the task [54]. This tentative result coincides with investigations of why Executive Information Systems (EIS) and Strategic Information Systems (SIS) systems do not work properly when traditional solutions to information provision are applied: traditional information requirement analyses were carried out which did not investigate the underlying *ad hoc* executive tasks and decision making.

4.2 Polyrepresentation of the cognitive space

The different investigations of operational information seeking demonstrate that, aside from the variable information need and current cognitive state (the conceptually known), additional underlying cognitive structures concerned with domains, tasks/interest, and problems/goals or uncertainty are indeed present. Some of these appear in well-defined forms. They follow the bottom-up order of causality as shown in Figure 4.

The cognitive theory suggests that we should explore the phenomena related to, and the development of, the cognitive structures supplementing the internal information need on the user side of the interaction process. The means by which this is done is request model building (RMB) and the extraction of verbal contextual representations of such structures (Figure 4). Ideally, this polyrepresentative extraction may produce three functionally different contexts for further processing by the interface and the system at each given point in time [55]: (1) a ‘what’, i.e. a request version which includes what is currently known about the unknown (the wish or desire for information); (2) the ‘why’, i.e. a problem statement as well as (3) a work task and domain description. Some representations may easily appear identical, e.g. the problem and work task descriptions. But because of its nature a desire for information cannot exist without a reason or a purpose, that is, the underlying intentionality, as defined by Searle [56]. This implies that at least two functionally different representations of the what and the why should be extractable. Further, we may assume that at least the intrinsic work task perception may be the most stable representative structure of all the cognitive structures involved during the IR process. As with human intermediaries, the question mode of the RMB is vital for the means and outcome of this extraction. The functionally different contexts are thus controlled, implying that one is free to merge them, i.e. to perform request fusion (or real and simultaneous ‘query combination’), or to apply them separately at will to the system. The application of these polyrepresentations of the cognitive space to IR is demonstrated below in section 5.2.

The quality of such representations, in particular the request formulations and problem statements (Figure 4) depends on the degree of variability of the information need as well as the current status of the cognitive state, and the constant influence from the information environment.

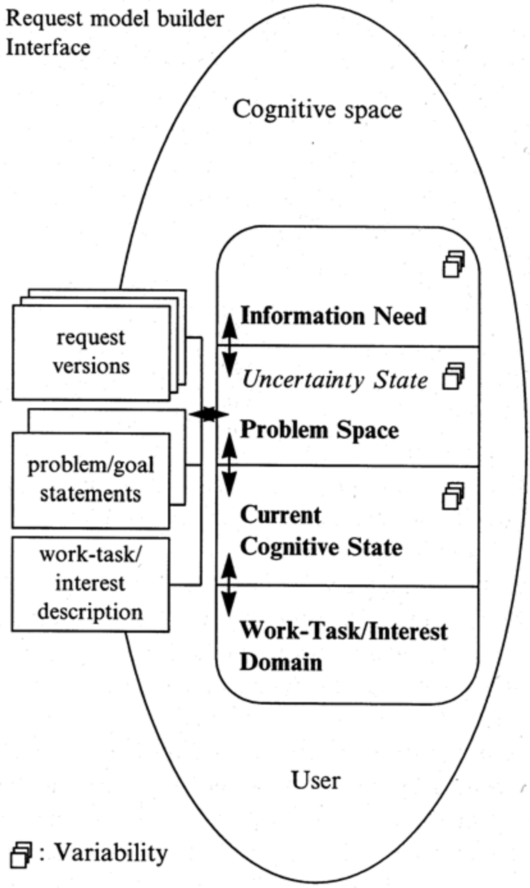

FIGURE 4. *The polyrepresentative nature of the cognitive space based on a bottom-up order of causality*

The empirical studies that have been described previously enable us to conclude that real-life intrinsic information needs are commonly variable, or shift focus, and may be intrinsically well or ill-defined at each new moment of the IR interaction (Figure 3). By nature they are exploratory, as suggested by Bates [42]. On the other hand, we cannot exclude the possibility that intrinsic information needs may indeed already be, or become stable in realistic terms during extended sequences of the IR process (focus stabilisation). The matrix (Figure 3) demonstrates four extreme cases of relations between variability and magnitude of definition of the intrinsic information need at a given point in time. More complex intermediate forms may obviously occur during real-life retrieval.

At each point in time, and indeed over a period, a searcher may occupy one form only. There are more or less variable problem states and stable domain/work task perceptions, that is, the supplementary polyrepresentative structures underlying each form of the current intrinsic information need. These structures may be intrinsically more or less elaborate. The current cognitive state, whether rich or weak, is assumed to be constantly transformable. In many ways information retrieval parallels the scientific processes of hypothesis testing and discovery.

Horizontally, stable information needs (boxes 1 and 4, Figure 3) are characterised by low curiosity, often resulting in search loops or spiral performance and encounters with dead-end retrieval (box 4). Regardless of box 1 or 4 the needs can be predefined, induced and kept stable during a session, as in SDI or certain patent search cases (commonly in box 1). Concerning all other properties the two types of information needs contrast each other. Because of its rich cognitive state and the stable nature of the information need, the box-1 searcher possesses a strong, but narrow intentionality. In contrast, a box-4 person's intent is vague or non-directed. In line with Winograd and Flores' conception [48] the mental 'breakdown situation' may permit a box-4 searcher to move into other boxes by hard-won step-by-step learning processes. Because of its beneficial properties the box-1 type is the preferred type of information need in mainstream IR research, and is discussed further below. The box-4 searching behaviour has been demonstrated in Ingwersen [13] by librarians searching on behalf of patrons not present during the retrieval process and by patrons during presearch interviewing. Similarly, experiments involving generalist library students as end users, who were induced with vaguely defined information needs, produce dead-end searching. The reasons seem to be that they lack personal domain knowledge as well as motivation for or clues to elaborating on the induced artificial need.

The variable information needs (boxes 2 and 3) are characterised by open mindedness, i.e. strong intent and a high degree of curiosity or imagination from the searcher. These needs can be experimentally predefined from the start of a session but cannot, by their nature, be kept stable during an investigation. In general, an externally observed searcher would be seen to perform berry-picking behaviour. Processes of learning as well as cognition would allow a searcher to pass from box 3 to box 2, and even to box 1, at a given moment during an IR session, although conceptual drawbacks from systems feedback may reverse the position. It ought to be possible to assess from external observation whether the person navigates by purpose (exploratory navigation) or performs more random non-directed navigation, i.e. browsing. The conceptual means for navigation, i.e. the cognitive states, are different with the well-defined need form in box 2 as the best suited for this conscious activity. Both navigation and browsing imply associative linking and interpretation, and signify that, for instance, query modification is possible for box-2 searchers. Box-3 persons may have difficulty in modifying their requests because of their intrinsically ill-defined form and the weak cognitive state. Instead, they may be able to apply a thesaurus, thereby recognising concepts and enhancing the query [57]. In both cases the searcher is open to novel clues provided by feedback displays from a system. Recent empirical studies by Efthimiades [58] and Allen [59] demonstrate that the structure and order of the display, whether in the form of term lists or records, plays a significant role in a person's perception and interpretation. This phenomenon is also known from ranked output experiments in which the relevance assessment time clearly influences the judgements made by several searchers on identical documents [60]: the more extended the evaluation period the more cognitive inconsistency and uncertainty associated with the binary judgements are displayed among the test persons. Cognitively speaking, one may assume that this increased inconsistency is caused by the additional involvement of the underlying situational problem and task structures which introduce a kind of ‘differentiated’ as well as ‘partial’ relevance assessments.

The differences between the boxes 1 to 3 concern the richness of the need and the nature of uncertainty as well as the possibilities for assessing relevance. Searchers in boxes 1 and 2 have conscious topical or verificative information needs whilst box-3 searchers demonstrate muddled needs [3, 49]. Only box-1 needs do not alter during the retrieval process. It is impossible to ask for topical relevance evaluations from box-3 searchers since the cognitive state is weak and the uncertainty as to the outcome is high. If required, the assessments will indeed be uncertain. Situation-based relevance judgements are achievable in all three cases since the underlying intentionality is strong.

Confined navigation, as in box 1, is defined as the conscious use of associations in a narrow area of information space, for instance, as defined by an exact or partial match technique. Exploratory navigation, box 2, displays consciously used associations in the totality of information space, while browsing signifies a random navigation procedure inhibited by intrinsic lack of knowledge, rather than by the dimensions of information space. Plausible inference networking [61] or spreading activation may provide means by which such retrieval activities can be initiated.

In the light of the human mental behaviour observed and analysed previously, label effects may easily occur with respect to the verbal request formulations (and other polyrepresentative extractions) in the intrinsically well-defined forms of information needs, i.e. boxes 1 and 2. The label effect is ubiquitous in the ill-defined types of need forms. The problems for the IR system (and interface) are consequently twofold: (1) to assess the proper nature of the need and the underlying cognitive state, because all the labels may look alike; (2) to make adequate use of such small request contexts, in particular at the crucial initial phase of IR interaction. Not only does the system become uncertain as to what really should be retrieved and ranked, but it also demonstrates badly tuned informative support directed toward users. This is the rationale behind making use of the additional, even label-like, polyrepresentative contexts during the process. In a controlled and structured manner they provide more data than the request alone for the machine to manipulate and to make use of during retrieval and user support.

4.2.1 Consequences for the experimental settings

It is consequently necessary to establish improved methods and theory to support both the intrinsically stable as well as variable information needs, in particular the intrinsic ill-defined and hence vaguely formulated ones. One suggestion has been to make searchers navigate or browse via hypertext links prior to actually posing a request, in order to increase the learning and cognition processes.

Although practically feasible, e.g. as done presently in the World-Wide Web, this proposal seems unrealistic in an ever increasing pattern of possible links available in and among information objects. This alternative implies the acceptance of really randomised retrieval performance. Consequently, request-oriented retrieval based on polyrepresentative contexts and, for instance, probability and plausible inference networks are probably the most realistic means to enter the information space properly, i.e. at different access points simultaneously, as discussed in section 5.2 below.

It is hence appropriate to map the investigative settings on to the matrix (Figure 3) in order to illustrate what has been assumed and achieved in general terms.

The real-life, online Boolean and seeking experiments have mainly operated within the frame of the information need forms characterised by boxes 2 and 3, exploring real needs and with retrieval behaviour as the experimental variable. The classic mainstream research setting has constantly been geared to make available and test isolated techniques for the retrieval of *simulated* static and well-defined information needs, founded on an assumption of total cognitive stability regarding knowledge state, problem, work task and domains, i.e. a highly reduced simulation of the box-1 form of information requirements.

Consequently, the Cranfield and initial TREC experiments rely solely on pre-defined requests turned into query versions by the algorithms in question (the variable). The assumption of cognitive stability during the experimental session assures that 'objective' measurements can be made against a pre-established topical recall ratio. Hence, during this SDI-like routing exercise each algorithm, whether vector space, probabilistic or plausible inference, performs different expanding loops in information space. The paradox for this experimental setting is that it works nicely for machines and automatic algorithmic assessments, but is unrealistic exactly because of its underlying assumption of cognitive stability. Just as the mechanics in the pits prior to a Formula 1 motorcar race know which engines are theoretically the best designed and tuned we now know which retrieval engines theoretically perform the best. This picture changes somewhat when the driver puts his foot on the accelerator and makes his way through the race. Even in a closed circuit, metaphorically speaking, the driver is allowed a certain degree of freedom to make decisions. The satisfactory laboratory results may thus not hold in real environments [62]. Similarly, by introducing *any* human interpretation into the investigations of a stable need form, box 1, one will have to allow for active and variable cognitive states during relevance feedback operations. Then, even in the smaller and closed test collections, the predefined recall ratio in principle falls apart as an objective and absolute measure. This paradox evidently arises even at the time of pre-assessment of relevance by a human experimental assessor (the observer). Thus, in a cognitive sense this objective scenario is an illusion. It is relative and uncertain as is the common case in IR.

A strong indication for this argument is the inter-searcher experiments for identical pre-assessed and well-defined stable TREC requests reported by Belkin *et al.* [63] and Iivonen [64]: inter-searcher inconsistency is paramount, and the fusion of the different search performances produces results that are better than or equal to those of the best performing searcher. Also, and more crucially, the topical relevance ratio increases or becomes constantly modified (i.e. relative) throughout the experiments, since alternative 'topically relevant' items are discovered. With this introduction of human interpretation of the simulated box-1 type of stable needs, the misconceived underlying assumption of cognitive stability disappears on the searcher side. On the system side assessments ought to be viewed in a similar relative way.

With this relativity in mind one may then accept simulated rich and stable intrinsic information needs processed and compared by machines only; then each algorithm can be seen to simulate a highly reduced human mind with no interpretative capacity. The comparative measurements for two search engines are valid but not absolute; and we will not be able to estimate the exact nature of the outcome when the same engines are applied to searchers. The problem is due to the human subjective capability of intentional interpretation which cannot be realistically controlled – only normalised in simulated box-1 experiments.

At each moment, a given IR system contains many alternative roads to information out of which only one or a few are actually uncovered, depending on the ensuing encounter with an observer, this being an assessor, a researcher, an algorithm or indexer, or a searcher.

Effectively speaking one may therefore argue that what supposedly starts as a box-1 experiment with searchers involved, involuntarily or unexpectedly transforms itself into a simulated box-2 investigation which is characterised by intrinsically rich but variable information needs. The synergy generated via the current encounter may be expected to influence and slightly to alter the individual comprehension of the previous search conditions, that is, the information need and the interpretation of retrieval outcome as perceived a moment earlier.

The online Boolean retrieval experiments have long since operated within the framework of the forms of well- or ill-defined variable information needs, boxes 2 and 3, realistically including real-life inquiries and accepting unstable underlying cognitive structures. Not surprisingly, exactly this research environment has increasingly questioned the use of topical relevance as the common measurement method [27-29]. In order to promote alternative 'relevance' or performance measurements in both the real-life as well as the simulation environments associated with the need forms 1 to 3, we suggest introducing the aforementioned task-situational descriptions or frames of the intentionality underlying the information need. Such descriptions can be observed and extracted via RMB during real-life investigations. They could be assigned to the simulated requests in the TREC experiments; see section 5.2 and 5.2.1.

For example, in the 'Boolean logic' example analysed earlier [13] the user assessments were related to the underlying problem (i.e. to obtain the adequate form of the formula in a design process), the work task (i.e. design and repair of computer circuits), or the domain (i.e. applied computer science). Almost all the retrieved domain-related philosophical, mathematical, logical, and computer science texts concerned with this request on 'Boolean logic' would be regarded as topically relevant. Only a very few texts in the actual library collection were, in fact, *useful* to the user's work task and problem – but they were finally retrieved. With reference to the discussion by Salton [65] differentiated assessments would have measured the total IR interaction result over all iterations with a low performance ratio concerning the work task and problem fulfilment, a rather low ratio of topical recall and a very high traditional topical precision. Only the ultimate run provided both high task-situational and topical precision scores.

This phenomenon of differentiated relevance is related to and includes the 'situational' or psychological relevance conceptions originally introduced by Wilson [66] and carried further by Schamber *et al.* [28] and Harter [29]. Hence, the differentiation is associated with the eight or seven categories of relevance analysed by Saracevic [24] and Barry [27] respectively.

With the empirical research at hand one may strongly argue that only a small fraction of IR behaviour, IR systems performance, and solutions have been discovered so far. We know which IR techniques are the best tested and how they function in relation to simulated static and well-defined information needs in semi-closed experimental settings [67]. We may combine them in many ways, e.g. by data fusion, sequentially, or apply their overlaps of retrieved items. We are only beginning to observe what happens in the case of variable well- or ill-defined requirements, user participation and interpretation in best match environments. We know a good deal more of these research issues from the exact match scenarios. Their results and ideas ought to be explored further. As demonstrated by the TREC experiments scaling up is relevant – the complementary side-scaling into the searchers' cognitive spheres is necessary.

Since our assumptions and knowledge concerning searchers' mental models are rather more elaborate than taken into account by the mainstream experimental settings, alternative investigative scenarios should be designed. By involving the polyrepresentative structures of the cognitive space the complexity of models and designs evidently increases further. Hence, the experiments must necessarily involve several but controlled variables. This is obvious since information behaviour and transfer *are* complex. Investigative and statistical normalisations and analysis of variance should be performed by means of a more reliable number of test persons and a mixture of controlled as well as real-life information needs of all the different forms (Figure 3). Assessments of whatever relevance measure prior to experimentation in a large-scale information space will require a larger panel of *assessors independently* evaluating the same portion of the space. Evidently, this procedure will lead to inter-assessor inconsistencies of a similar nature as inter-searcher or inter-indexer inconsistency, and will not produce a final objective measure. We must thus expect, and should accept, very relative, variable and thus unreliable measures over time. The current TREC and Okapi investigations demonstrate this inherent relativity [68].

In the case of non-simulated studies pre-experimental assessments should be replaced by precision-related post-assessments by panels. Such evaluations will likewise be inconsistent and lead in principle directly to the re-introduction of partial relevance assessments [69]. The binary formal retrieval algorithms may consequently have to be applied in more complex ways, i.e. be exercised on the different sets of relevance assessments individually: The relevance discussions will continue in vain if one single definition is attempted; what is necessary is to understand and make clear which kinds of ‘relevance’ are applied in an experiment (or a running system): topicality and/or situational problem, preference or task oriented assessments. This research landscape is fundamentally a sociological one, dirty and tricky. At present, we are testing selected features of this scenario based on the assumption of situational and partial relevance assessments by real users for both simulated and natural information demands, as well as polyrepresentation in simplistic forms based on the tentative ideas that were published in Ingwersen [55].

4.3 Polyrepresentation of information space

Information space consists of two major interactive components (Figure 2): the system setting and the information objects. In a cognitive sense, these components are directly influenced by several different human cognitive structures which, during the event of IR, intermingle in time and by type of origin: authors of texts and other media-borne entities. In certain document types, authors are additionally responsible for their citations to other information objects; human indexers who, for certain document types, interpret these generated information objects producing representations; thesaurus constructors, who produce their interpretations of the vocabulary and semantic relations of concepts within different domains; producers of the storage media and their selection criteria, e.g. the citation indexes’ restricted contents; database designers, responsible for the database structures and availability of data extracted from the information objects, but placed in the system setting; retrieval engine designers, representing their computational retrieval logic in the form of algorithms, such as Boolean logic, vector space functionality, probability functions, plausible inference networks or logical uncertainty. During the actual act of retrieval the searchers of information basically play an interpretative role within this framework. In the case of IR experimentation additional cognitive structures are involved: those of the experimenter/assessor or observer.

Information space is consequently seen as a polyrepresentative structure in which *time* plays an important role: the same issue or topic is treated in conceptually and philosophically different ways, both at a given moment and over time, by different authors. This lack of consistency among authors is trivial but, nevertheless, the fundamental reason why social, psychological and scientific attitudes change over time, why knowledge, cognition and new questions emerge – and why IR is difficult to perform. This inconsistency is understood. Which other – and perhaps similar – human inconsistencies exist within the polyrepresentative structures, associated with the retrieval of information?

4.3.1 Cognitive inconsistencies and overlaps

Empirically speaking we have previously seen that a lack of consistency exists between searchers attempting to obtain information from the same source, by means of the same retrieval engine and having identical predefined requests [63, 64, 69]. The fact is that by observing each type of cognitive structure involved in IR one finds not only the expected discrepancy between these structures of different cognitive origin but one also finds inconsistencies within each type of structure.

Consequently, our hypotheses are:

1. all the inconsistencies are self-evident, inescapable, formally unpredictable, and of similar cognitive nature;
2. the more remote in cognitive origin and in time, the less the consistency;
3. the inconsistencies can be applied favourably to improve retrieval because:
4. if different cognitive structures, in defiance of the inconsistency, do, in fact, retrieve overlapping information objects, this cognitive overlap presents more ‘relevant/useful/...’ information objects than each independent structure;
5. the more different the cognitive structures producing an overlap are in time and by cognitive or functional type, the higher the probability of its ‘relevance/usefulness...’.

Empirical investigations from the IR field as well as from the bibliometric and scientometric subdisciplines of information science demonstrate clearly the validity of these assumptions.

Different indexing methods applied to the same text collection yield different sets of text or document entities for the same query, e.g. natural language vs. controlled vocabulary methods. Several empirical investigations of the matter [30-32, 70] carried out in the operational exact match environment demonstrate clearly that the intersection of controlled index term phrases and natural language representations from the basic index yields the best retrieval results, better than either separately. The intersection constitutes the cognitive overlap of items retrieved. The greater the variety in cognitive origin of method, the greater the difference in the results. Human indexing produces sets of terms that are rather different from automatically indexed collections of texts. This result is not surprising since the human indexer makes intentional interpretations, producing keywords that are not necessarily present in the object, whilst the automatic indexing algorithm is reduced to manipulating exactly the words present. Inter-indexer consistency tests are well-known. More recently, inter-linker inconsistency for hypertext link-creation is similarly found to be quite high [71]. Overlaps of indexing terms added by different indexers for identical documents, and hence of objects in a collection, range from 10 to 80%, depending on the degree of reinforced indexing rules and vocabulary control [72]. In Ellis *et al.* [71] no rules nor control were applied and, in line with the cognitive view, the expected low consistency among the linkers was measured to be between 15 and 40%, the length of the documents seeming to influence the degree of overlap.

One of the very few investigations of the alternative overlap between index term retrieval of journal articles and searching via citation indexing (texts citing the articles) has been carried out by McCain [73]. For the same queries in *Medline* and the ISI databases the result was as expected from a cognitive viewpoint: the two *very* different cognitive structures of indexers and works cited by other authors, the latter group interpreting the texts with completely different intentionality from the former, yielded an overlap of only 11% on average. Similar small overlaps are reported by Pao in a different field study of term and citation searching which also included relevance assessments of the retrieved overlap [74]. As above, the two representation methods are complementary to one another and the overlapping items are important. Therefore, the degree of overlap between natural language representations of full-text and citation indexing (as well as the NLR of such citations) ought to be explored further. The ISI databases are inadequate as stand-alone instruments for such investigations since they presently do not contain full-text items. However, such overlap experiments, including 'relevance' assessments, could be carried out in selected TREC files that are now available. The prediction is that the two sets of retrieved text entities are similarly complementary with a small cognitive overlap. Also overlapping are sets of texts represented by index terms assigned to source articles and to the references cited by the authors of these articles. This has been studied by Harter *et al.* within the domain of information science [75]. They found that the term overlap was below 20% in the same database. In a cognitive sense this small overlap is not surprising, given the different cognitive origins and intentionalities behind the sets of documents involved. Also the time difference between indexing the source and cited articles influences the way in which the indexer(s) interpret the documents. Similarly, quite small overlaps of items can be found when isolating data sets for bibliometric analyses from domain-dependent as well as ISI files for the same topic and time period [76]. The overlap, i.e. the duplicates from the Dialog One-Search, range from 18% upward depending on the specificity or speciality of the topic searched for. However, these duplicates are vital for retrieval, since they are retrieved by an overlap of the cognitive indexer structures originating from the rich basic index fields in the domain files and the means of representative structures found in the selected items of the ISI databases, that is, the author-derived titles/abstracts and simple keywords. In these two experiments [75, 76], no assessments of relevance were carried out, only overlap measures.

An identical picture is obtained with respect to overlaps between sets of texts retrieved via different partial (and exact) match techniques [77] for the same query. Formal retrieval models of similar nature produce substantial overlaps while quite different search logics decrease it. The pattern is the same when applying different weighting schemes. A more detailed outline of the characteristic retrieval power of the range of available matching techniques is provided by Belkin and Croft [78] and by Croft in relation to 'intelligent' IR [12]. From a cognitive perspective this evidence of variety is obvious. In a simplistic way such cognitive overlaps between different retrieval engines can be observed and directly tested in current operational systems. One may point to comparing the novel online technique which attempts to rank the retrieved items by 'topical relevance' – Dialog's Target command – with traditional Boolean online searching, as done by Tenopir and Kahn [44] and by Keen [79]. Similarly, but in a more sophisticated manner one may operationally search the functionally different author conceptions of information objects (titles, headings, captions, abstract, full text) and the cognitively different indexer interpretations of the same objects (keyword phrases) by means of cognitive online searching involving quorum logic. The outcome is a ranked list of small sets which can be directly compared to a Target output (see Appendix). Online data fusion as well as overlap investigations of the separate sets of objects from the two different retrieval logics can be performed, and simulated 'probability ranking' of topical relevance is possible by means of the Target feature for the overlap set, the individual sets and the fused sets. The principle of applying cognitive overlaps, both utilising the indexing structures and author-generated data structures based on different functional origin as well as the use of different retrieval engines to identical information objects is modelled in Figures 5 and 6.

Our hypotheses seem justified by the available empirical results. However, they can easily be further tested in experiments including or simulating real users and involving different types of precision-related assessments. Simplistic similarity measure investigations can provide relevant indicators and conditions for the magnitude and detailed characteristic properties involved in the creation and use of cognitive overlaps, for example, the writing or message style factor discussed below.

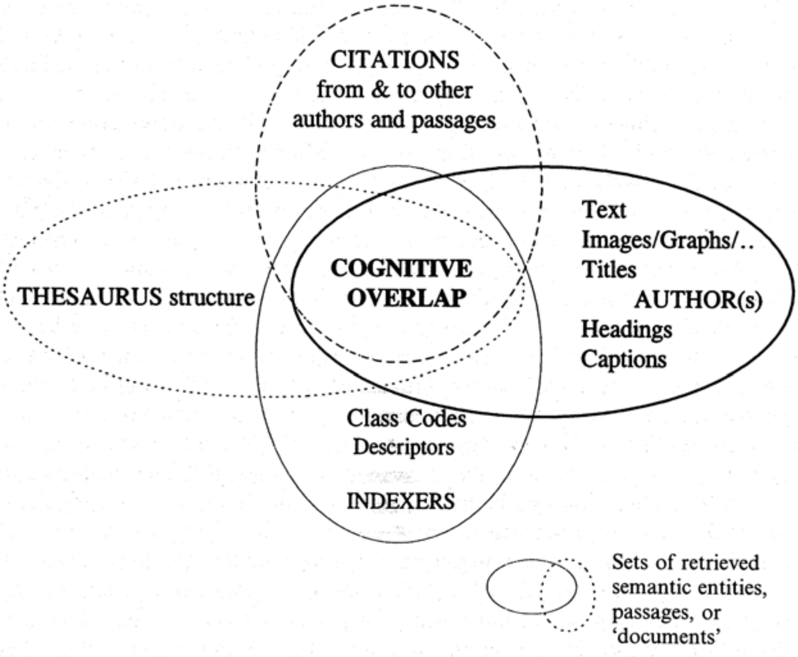

FIGURE 5. *Overlaps of cognitively different access points provided by one retrieval engine in information space, and associated with one searcher statement*

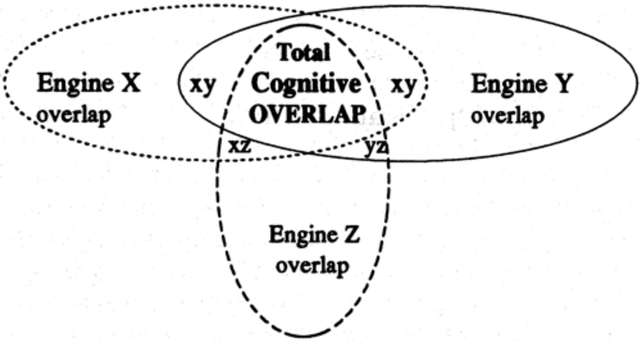

FIGURE 6. *Cognitive overlaps of semantic entities provided by the intersection of the polyrepresentative retrieval results from different retrieval engines in information space, associated with one searcher statement and generating a set of total cognitive overlap*

4.3.2 Polyrepresentation of semantic entities in full-text

With the increasing possibility of accessing full-text and graphical records, images, etc. the representation of such items would at first glance seem easy and simple: the automatic application of the signs and symbols inherent in the items themselves. This is the underlying rationale for the mainstream IR research carried out during the last three decades. In the previous sections we have demonstrated the limitations of this scenario. In addition, full-text objects present linguistic and psychological retrieval problems because they are larger and more heterogeneous than the abstracts used in the experimental settings of the past. One way used to ease the application of various weighting schemes, for instance the *tf* function, has been simply to cut at random the full text into smaller entities, without thinking of the semantics of the text. Another, more elaborate method is to introduce logical and linguistically confined passages, as done by Salton *et al.* [2, 80]. In this respect the passage, e.g. the paragraph in a news item or encyclopedic entry, is well chosen since the total information object of these types of documents may often contain several different content-bearing units. Only the heading holds them together. A completely different reason for breaking down larger information objects into smaller retrieval entities would be to observe their utilisation by searchers during retrieval activities, e.g. during situational or other relevance assessments. This was not necessarily the original intention in the work cited nor in the mainstream research community in general. The basic idea here was (and is) to assure normalisation of word counting and weighting. Our proposal for introducing what we, in a generic sense, call semantic entities is caused by the empirically founded and intuitive assumption that searchers in general obtain their information from such smaller entities – but in the context of physically adjacent or conceptually related entities. Ingwersen's full-text investigations [13] as well as the Japanese relevance assessment results [60] provide strong support for this assumption.

Adequate semantic entities of full text are defined as the smallest unit of a given text which provides a recipient with a semantically confined context. Depending on the document type, such entities may thus range from sentences constituting titles or figure captions, via paragraphs, e.g. in news release units, to sub-sub-sections in journal articles and books. An image or a graphical object are in themselves semantic entities. If embedded in or associated with a text, corresponding text entities exist that may act as textual representations of such objects. Similarly, the image object may represent the corresponding text entity; see the ‘Mark Twain Painting Case’ [13, pp. 24-25]. The suggestion that such smaller parts of traditional documents should be used for representation and retrieval purposes is not novel. For monographs the principle of using chapter headings and table captions was proposed and tested in relation to the Subject Access Projects (SAP) for library online catalogues more than a decade ago [81, 82]. Paragraphs are used in hypertext link-creation experiments [71] as in Salton’s passage retrieval of encyclopedic entries. In linguistically based IR the sentence is the preferred semantic entity for analysis and retrieval [83] and it has also been proposed by van Rijsbergen to form the kernel to which to apply the principle of logical uncertainty [1, 4, 5]. The SIMPR indexing machine makes use of chapter and section headings as well as full-text sentences in order to produce indexing concepts automatically by linguistic analysis and to point to potentially relevant sub-sections and paragraphs in large monographic sources [84, 85]. With the exception of Salton *et al.* [2, 80] and similar partial match attempts at a lexical linguistic level, one may thus observe a structural or syntactic approach to representation as well as to retrieval which tends to approximate to a contextual level in R research. The *context* surrounding or related to the entity is all-important in this method of advanced retrieval.

4.3.3 Paragraphs or sections as preferred semantic entity? – the stylistic influence on retrieval

Indeed, it seems adequate to assume that, in general, text paragraphs and single figures or tables provide the smallest semantically confined retrieval passage of documents that may produce the information searched for. If we, for instance, observe how authors directly cite other text units in their own texts, such units are commonly paragraphs or sentences. Evidently, one or other of the adjacent paragraphs provides the vital context embodying and giving meaning to the cited unit. On the other hand, one may note that the writing (or message) style apparent in the various forms of entities influences the retrieval parameters. Paragraphs from introductions are functionally different in style from the corresponding abstracts, and both styles are rather different from that applied to paragraphs embedded in the main body of the text. Kwok provides a further analysis of the document components and their retrieval potentials [86]. Every document type has its own style which may vary from domain to domain. Sociologists write differently from physicists, who again convey their academic messages in styles that are very different from those of journalists. Also citation styles vary from field to field. The style in spoken messages differs from that in written communication. Empirically there exist indications of this phenomenon when applying Zipf's 'law' to different domains. More recently, IR techniques have been used to trace term patterns in literature [87] and retrieval experiments of real-life (non-simulated) spoken 'documents' are under development in Cambridge by a team headed by Sparck Jones [88]. We may then still apply the paragraph as the basic retrieval unit but, depending on the media type and the domain in question, larger semantically confined portions of texts (and other kinds of messages) might in fact provide more adequate retrieval. For example, in academic articles there exists a conscious intent behind beginning a new section or entering a subsection heading; and article readers may prefer sections because they provide an adequate and confined context for perception. At present, however, we do not possess sufficient evidence to justify a proper semantic limitation. The access options for sections from articles and books are demonstrated in Figure 7.

Associated with the notion of semantic entities smaller than the traditional document is Bookstein's recent suggestion to review Bradford's bibliometric law of distribution in a multidisciplinary way [89]. A scaling analogy exists between the impact of journal articles on a journal in terms of its subject coverage, and the impact of sections (or paragraphs) as semantic entities forming part of a full document. Bookstein's basic idea is that journals of more general and multi-disciplinary subject coverage of a domain may contain individual articles, each of which points to a different specific subject area of that domain. Although the journal as a whole covers (is 'relevant' to or even core for) the entire domain each article contributes little to the total domain impact of that journal, compared to specialised journals which solely contain articles on a specific subject area. In Bookstein's critical analysis the common Bradford distribution function, quite disturbingly, ranks the specific journals much higher for that domain than the general core journal. Similarly, a full-text article may seem only marginally useful, simply because only one of its semantic entities actually covers the area asked for. The total document is represented by means of human or automatic indexing, but only by a limited number of keywords so that the useful specialised paragraphs or sections of the document are generally hidden. They are retrievable only by means of guess-work by a user, based on a reinterpretation of the selective interpretation made by the indexer (or the algorithm). A section on a specialised subject in a more general article should be regarded as being just as useful as the sections in an article which as a whole covers the specific and useful subject area. We may thus regard smaller semantic entities as productive for IR.

However, negative consequences exist for IR when using smaller units of full text, since the lack of data may reduce the range of representation and access methods that can be applied.

No 'paragraph header' normally exists in addition to the full-text terms and concepts. However, it is still functionally valid, for instance, to claim that an article section or a news item header logically may cover several sequential paragraphs of text. Words and concepts in such headings ought then to be weighted higher than those occurring elsewhere. Intersected with other NLR

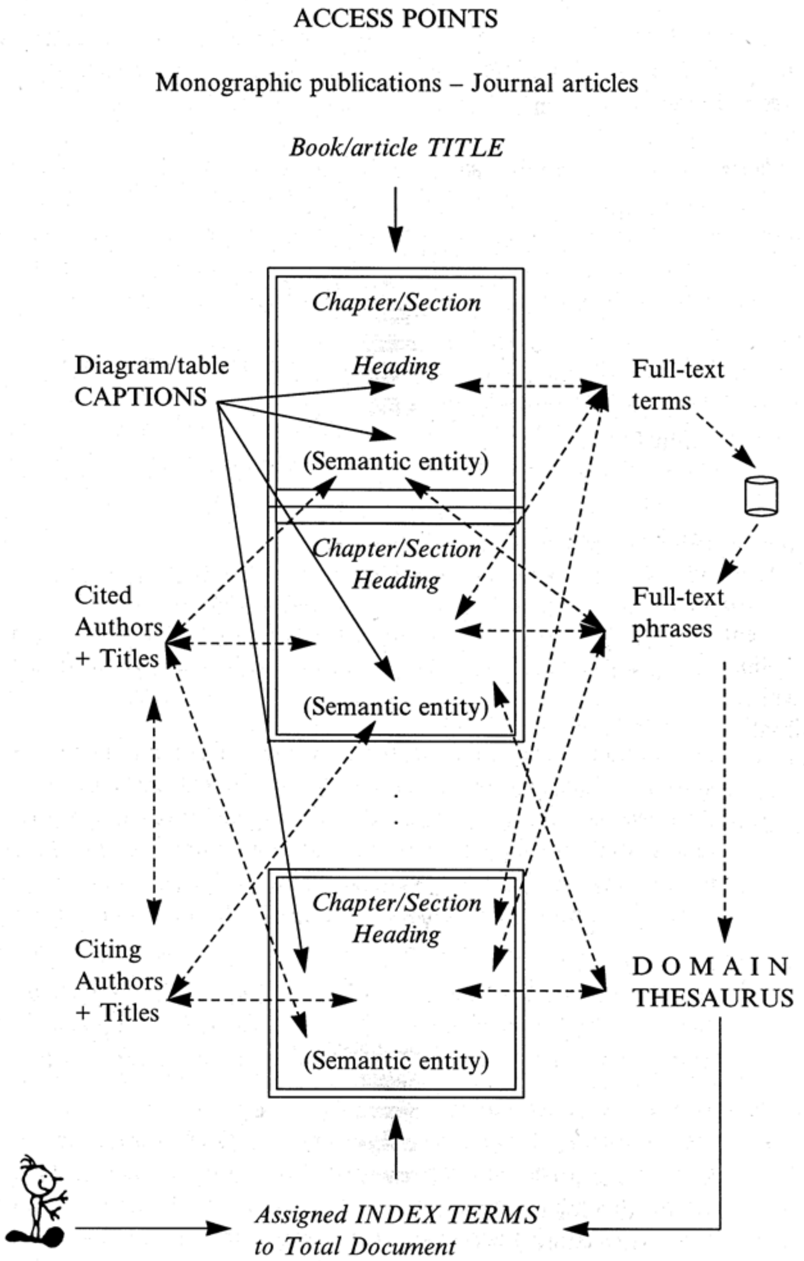

FIGURE 7. *Simultaneous access points associated with (sub)sections of journal articles or monographic material. Access via abstract words and successful retrieval profiles of previous searchers are not shown. (—►) signifies the optimal range of access points in bibliographic online systems; (-►) signifies the additional points of access to full-text material*

means from the text/message this, and other relationships of conceptual, and logical dependency between the whole and its entities, may be exploited in a structured manner, e.g. as attempted by Schmeltz Pedersen by application of lattice theory [90].

Another negative aspect of introducing semantic entities is the difficulty of applying human indexer interpretations to such entities in order to assign representations that are mainly concerned with cognitive elements *not mentioned* in the text or explicitly displayed in an image, but that are of potential value to retrieval and user. This is commonly the case for some document types, such as encyclopedic entries and online news items.

To summarise the application of the paragraph as the selected semantic entity, polyrepresentation may take several different forms. At least three of the four cognitively different structures of representation and retrieval modes shown below may be involved simultaneously producing cognitive overlaps for the same database:

Author-generated representations:

- la. the text itself, including overlying headings and related captions and, in certain document types, citations to other texts;
- lb. structures by other authors actually citing the paragraph, or at least the page holding it, may be involved in certain document types;

Automatically generated representations:

- 2a. the method(s) applied to representing the paragraph via form(s) of automatic indexing, weighting, and linguistic text analysis;
- 2b. the retrieval technique(s) applied to the original text and the forms of representations derived from 2a. The representations and the techniques may be identical, since each best match technique may also serve as a means for automatic indexing at the single-term level;

Indexer/domain expertise:

- 3. a domain-dependent thesaurus structure (if existing) applied to and supporting the representations. As a constructive supplement thesaurus entry points might be user-generated associations [91];

Searcher-generated representations:

- 4. the successful retrospective search profiles which, stored for a given time period, retrieved the paragraph. Condensed term lists, like those produced by the frequency facilities Zoom or Rank based on selected sets, belong to this access category, since they are derived from such profiles and may provide access to the underlying set of items.

The author-generated context associated with objects of other types of media in multimedia documents may provide access to such objects. If sections or subsections are applied instead, a few additional access options are available (Figure 7): section and chapter headings for which the author is responsible. As for paragraphs the functional difference between the locations of sections in the text should be taken into account. However, the time issue is dealt with only partially. Only the access structures 1b+2a+2b and 4 are dynamically changeable.

For each different cognitive origin mentioned at least one version of a given representative structure may be present in relation to the same semantic entity – but several different simultaneous versions may actually be expected. For instance, several different methods of automatic representation may be applied simultaneously: (a) single term extraction and weighting, combined with (b) linguistic analysis producing concepts – as in SIMPR which may (c) be validated via a thesaurus structure, producing 'controlled' index terms not present in the analysed text (Figure 7). Several retrieval engines may be applied in parallel: vector space, probability, clustering, spreading activation, etc. – with each being applied to different representations if required. Each engine will retrieve different sets of paragraphs or sections with a varying degree of overlap in a calculated and intentional manner. Thus, one is free to manipulate the search outcome, for example by producing a variety of overlapping sets, to create conceptual network structures spreading in a variety of directions into information space (data diffusion)*, to fuse these sets, or simply to ignore some of the sets retrieved. This bottom-up approach to representation and retrieval of semantic entities makes it also possible to view larger parts of the entire document in the light of the retrieval power of its own entities. Commercially, this feature is presently incorporated in the Status IQ database software which is based on the Sparck Jones/Robertson probability ranking principle. In general, the mode of manipulation should depend on the intentionality of the current user, as perceived by the system.

5. POLYREPRESENTATION IN IR – THE GLOBAL MODEL

By applying polyrepresentation, both in the form of several simultaneous methods of representing objects and retrieval engine techniques as well as user-generated need representations from cognitive space, redundancy is paramount. In just the same way that one can profit directly from the inescapable cognitive inconsistencies, the introduced redundancy can be pursued and exploited intentionally.

This principle of intentional redundancy applied to polyrepresentation in IR originates from the arguments put forward by Sparck Jones who was concerned with the necessity of 'working through redundancy, that is, different ways of referring to the same concept and of linking different concepts in form of a conceptual network thrown over the underlying information objects' [9, p. 9]. The generic inference network proposed by Turtle and Croft [61] follows similar lines of thinking. We call this redundancy 'intentional', since general and non-estimated redundancy in relation to representations and retrieval of information objects may not always be productive in IR interaction. The degree, mode and function of redundancy should be determined by knowledge of the current user's information

*Data fusion thus is an implementation of a special case of polyrepresentation whilst data diffusion equals intentionality to keep separate the result sets.

behaviour, e.g. as inferred or believed by the intermediary mechanism, based on elaborate models of searcher behaviour (as discussed in sections 4.1–4.2).

In other words, from a cognitive perspective as many different cognitive structures as possible should be made available and applied during IR interaction, though in accordance with an estimation which allows for a calculated selection of exactly such overlaps that are regarded most appropriate to the current retrieval situation. This issue of estimation is not primarily seen as a mathematical one but rather as conceptual, logical, and psychological issues.

Simplified intentional redundancy has been performed for decades in the operational online environment by means of different forms of NLR, human indexing and thesaurus support applied to the same bibliographic record collection, but it has rarely been applied to full-text entries.

5.1 Logical uncertainty – a compatible logic for IR

The logical uncertainty principle (LUP) for IR proposed by van Rijsbergen [4, 5] and further elaborated in the information calculus (IC) for IR [1], is a logic which is directly aimed at providing a general theory for *information* retrieval. Similar to our own, the very fundamental assumption in this approach is that the understanding of information comes first and that logic and probability come second. Van Rijsbergen and Lalmas' notion of information 'starts from the position that given an ontology of objects individuated by a cognitive agent it makes sense to speak of *the information contained in one object about another*. Indeed we can talk of an information flow existing between two objects. ... information is seen to be dependent on a context and is taken to have a level of intentionality, that is, information is *about* something and, as we all know, aboutness is an intentional notion. Information also depends on there being a cognitive agent to classify it' (italics by the cited authors) [1]. Close examination of the understanding of information shows that essentially it contains the same properties *and* conceptual consequences as discussed in previous sections of this paper: the information concept in section 2, and the cognitive structures leading to the logic of polyrepresentation, in sections 3–4. Both approaches see information as going beyond meaning, being something which, depending on *context* as well as on *intentionality*, is a result of a cognitive situation and interpretation. Information is that contextual supplement which, added to an object in the form of a semantic entity or a cognitive structure, makes its aboutness perceived. The perception takes place at an individual intentional level and the aboutness is individualised. Operationally speaking, human hypertext linkers or authors act as such agents who make explicit the existence of the information flow or nearness between semantic entities. Searchers perform similar junctions between information objects and their cognitive space.

The published LUP/IC form of the theory seems to be directed towards the data structures distributed in information space, but the cognitive agent need not be only a human generator (Figure 1). The theory implies that the 'agent' can be taken as a notion for the entire range of cognitive structures actually involved in IR (section 4.3), that is, authors, human indexers, retrieval engine designers, searchers, etc., directly leading to the concept of cognitive polyrepresentation. In our view the transformation (interpretation) of data 'content' into aboutness and information can only be performed at a cognitive pragmatic level (Figure 1) for instance as shown above, by the authors themselves, human indexers, or the searchers; other cognitive agents involved are represented by quite simplistic structures mainly at the lexical or morpho-syntactic level, e.g. the traditional retrieval algorithms or formal logics. These may indeed add context, possess embedded intent, but only process data at a linguistic surface level, i.e. constantly at the content level. Hence, in our view 'the "information" contained in one object about another' holds only potential cognitive properties until it becomes interpreted at a cognitive level by an agent. This is my understanding of the LUP/IC theory and seems to be compatible with the stated underlying conception of the theory since: 'This interpretation [of the uncertainty principle in IC] is only syntactic; the strength of the logical approach is that it opens the door to incorporating semantic information through transformation [by adding information to initial objects or worlds]. The older models for IR do not allow the incorporation of such semantics' [1]. It could be inferred from this that by processing data content at the syntactic level, as discussed in section 2, explicit as well as implicit semantic values in objects can be detected and used for retrieval. The explicit values are the linguistic meanings directly observable (e.g. the sentence: 'time flies like an arrow' provides four to five syntactically feasible interpretations). There exist implicit values for each meaning, i.e. semantics which might be but which are not present in the object (e.g. the location-facet: 'time flies like an arrow in London, not Africa') [92]. The number of possible implicit values is finite, depending on the actual context and language practice, but any cognitive interpretation of such meanings may in principle produce an infinite number of information values over time. This phenomenon creates uncertainty, namely: which and how much context needs to be added to retrieve those semantic values which provide the information searched for? The first three Principles of the IC logic demonstrate ways of achieving this goal. In relation to the Fourth Principle it is suggested that: 'the more ways there are of satisfying a query [i.e. a request] the more relevant the initial document might be. It is like having many independent pieces of evidence for an hypothesis' [1].

Our suggestion of polyrepresentative cognitive overlaps hence conforms to the Fourth Principle of the IC logic presented in van Rijsbergen and Lalmas [1]. Consequently, our polyrepresentative approach set in the framework of the cognitive view is seen as a conception of IR compatible with and overlapping the theory of information calculus and the principle of logical uncertainty.

5.2 Polyrepresentation exemplified

We propose to extend the principle of intentional redundancy into the user's cognitive space, exploiting the full potential of its different cognitive structures. These constitute a powerful inference network with truly dynamic properties and based on human intentionality. Polyrepresentation on both the system side and the user side is assumed to provide improved access structures for all the actors during IR interaction.

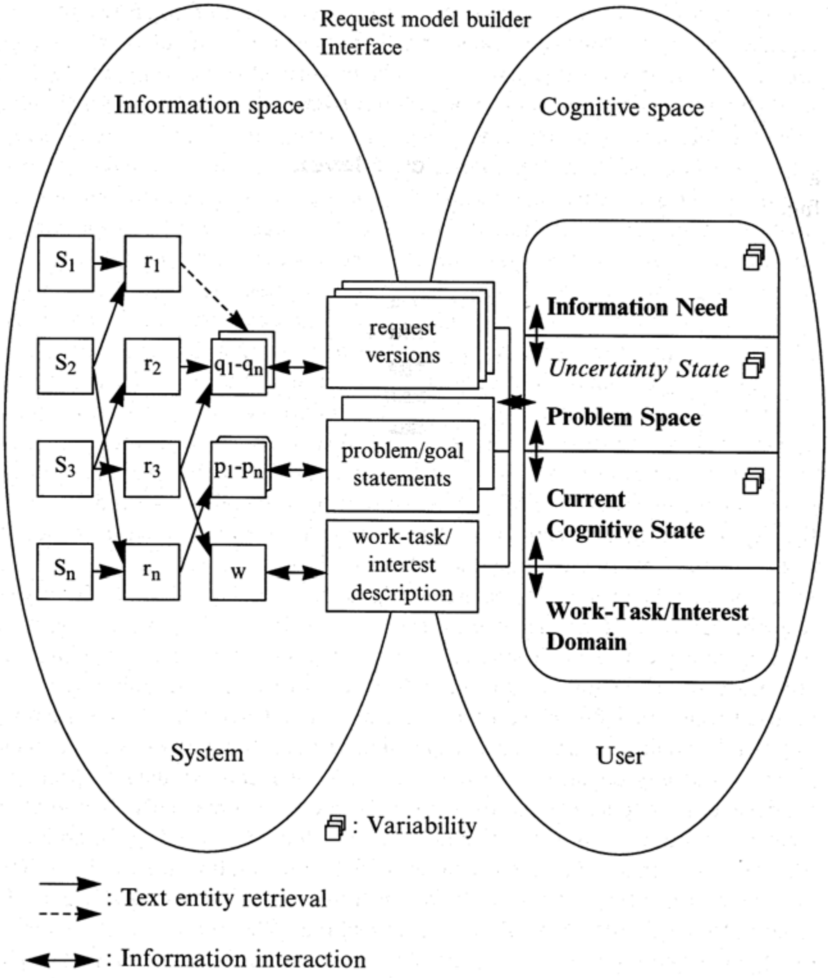

FIGURE 8. *The Global Model of Polyrepresentation in IR, extending the network model of information space by Turtle and Croft [61] to incorporate representations of the user's cognitive space and interface functionality.* r_1-r_n *signify different representative methods retrieving different (sets of) semantic entities,* S_1-S_n

Figure 8 displays the global model of polyrepresentation of semantic entities in information space *and* the user's intrinsic cognitive structures following the principle of intentional redundancy. On the right-hand side, the model tries to encompass the participating cognitive elements responsible for the formation of the individual intrinsic information need and the resulting request versions. In this respect, the elements form the polyrepresentative entities of the user's cognitive space in a vertical, bottom-up mode of causality, leading to the information need. The representations extracted by interface functionalities, such as the RMB and User Model Builder functions in the Mediator Model [3, pp. 209—216] or similar model builders suggested elsewhere [10, 11], provide a structure of statements and descriptions to be applied to or by the IR system. Each individual cognitive element may then form the basis for further elaboration in a horizontal dimension, if required, by structured questioning via an interface functionality, as suggested in Ingwersen and Wormell [92]. Both dimensions may hence provide the context necessary for applying the principles of logical uncertainty and information calculus.

A third dimension of context arises from the degree of changeable conceptual background knowledge of the user (i.e. his domain knowledge), associated with his dynamic ‘current cognitive state’. Again, various interface models have suggested functional solutions to extract this type of knowledge from the user, which is essential for estimating the nature and quality of the current intrinsic information need, Figure 3 [3, 11, 93]. A fourth dimension of time exists during any retrieval session, signifying the dynamic variability and versioning of intrinsic cognition (and learning) and verbal statements demonstrated by the model.

The interface at the centre plays a key role in this scenario, since its outlined functionalities ensure that the intentional redundancy principle may be enforced. The aforementioned estimations as to which representative methods and retrieval techniques to apply, and under which circumstances, are based on the interface’s perception of the retrieval situation of the user. The crucial point is the estimation of the cognitive nature and quality of the information need. For instance, if it is perceived that the user simply wishes to verify and locate items already explicitly known, the in-built intention of the interface should logically be to abandon redundancy and apply exact match retrieval on simplified representations of entities in information space. On the other hand, if the need seems to be of a topical nature, there are three basic alternatives (Figure 3): (1) the intrinsic need is well-defined and stable; (2) it is well-defined and variable; (3) the information need is intrinsically vague, ill-defined or non-existent, and consequently variable. According to the cognitive view it would be mandatory in the second and third cases to allow the user to explore the information space in a multidimensional way. We extend this ‘berry-picking’ metaphor also to encompass the conceivable (re)formulation by the underlying problem or goal which, as stated previously, may also be vaguely developed internally to the user.

At the left-hand side of the model (Figure 8) the (*r*...) notation signifies nodes of concepts derived from different representative methods (—▸ *r*...) applied to the semantic entity nodes (*S*…). The (*S*…) notation may also signify sets or clusters of semantic entities. As the figure stands it demonstrates the application of one retrieval technique, e.g. the probability-based method. In relation to each representation of the user's cognitive space, i.e. (*q*..., *p*..., *w*), including the domain, the retrieval of semantic entities is based on the probability that an entity satisfies these representations individually or in combination with one another (*r*... —▸ *q*…, *p*..., *w*), as discussed above. However, only real assessments can be carried out by the actual user in relation to his intrinsic cognitive representations.

A similar model would represent another retrieval technique in action (Figures 5 and 6), e.g. the vector-space model, in principle yielding different nodes or (sets of) semantic entities for the same (*q*...), (*p*...), etc., and for the same semantic entities (*S*…) embodied by the same methods of representation (*r*...) as before. (r_1 – r_n) might, for instance, be matching concepts validated by a domain thesaurus (r_1), simple concepts extracted by linguistic text analysis (r_2), single text terms (r_3), or personal names extracted from the entities (r_n).

In the case of an ill-defined or vague intrinsic information need, the interface should concentrate on the other remaining cognitive elements internal to the user by allowing him to follow and explore different conceptual paths in the information space. These paths are associated with the aforementioned different diffused structures of representation, as shown on the left-hand side of Figure 8. The theory suggests that for each cognitive element of the user space (the right-hand side, Figure 8), there should exist a corresponding network of concepts (*r*...) and semantic entities (*S*…) which may be explored by the user. For each such element, say the 'work task', a path is denoted by at least one of several different methods of text representation (r_3) matching the user's task description extracted by the interface (*w*). The resulting semantic entities are signified by (S_3). As stated above, each retrieval technique applied to the methods of representation will, theoretically, generate a slightly different set of entities (S_3')(Figure 6).

In total, the model demonstrates that by introducing the concepts of 'problem space' and 'work task/domain', in addition to the traditional 'request'' formulation, supplementary separate paths for exploration and retrieval are made available: paths which otherwise might never be reached via the need-associated query versions alone. By means of the additional information from the system, more differentiated and useful situational relevance judgements and query modification are possible during the following part of the session ($r_1 \longrightarrow q_2$).

Not demonstrated on the model is the additional intentional possibility to combine, or fuse, all the retrieved sets of entities (e.g. S_2 – S_n), regardless of which cognitive structure they belong to in the user's cognitive space, and in defiance of which type of representation or IR technique by which they are retrieved. Fundamentally, this procedure implies that the request, problem and task statements, and the domain description should be viewed as one comprehensive manifestation (or fusion) of the cognitive space at a given point in time. This 'request fusion', which does not correspond to the notion of query combination [94], is the recommended strategy to follow if the initial request, problem statement and work task/domain descriptions are poor, conceivably signifying an ill-defined and variable intrinsic information need (box 3, Figure 3).

Irrespective of the precise structures in the cognitive space, the structured but different sets of retrieved semantic entities can be fused into one output set in which the ranking, according to the cognitive view, will primarily depend on the degree of overlap between the various sets (Figure 6). The highest rankings are given to semantic entities retrieved by all or most techniques as well as means of representation.

These intentional strategies allow the user to perform assessments directly associated with each of his own intrinsic cognitive structures. It is thus possible to separate ‘situational relevance’ from ‘topicality’.

Hence, the theory makes it conceivable alternatively to fuse (1) the statements from the user’s cognitive space, (2) the various retrieval results, (3) the statements and retrieval structures on both sides of the interface, or (4) to diffuse or keep separate the structures on both sides. These data fusion approaches (1 to 3) are precision-oriented. One may suggest that one might carry out the final ranking of entities within each of the defined overlaps by probability ranking principles, due to their cognitively appropriate and direct use of the notion of non-relevance. However, since no relevance assessments do exist during the first retrieval run associated with the different statements extracted from the cognitive space, any other ranking principle, for instance the vector space model, might also serve this kick-start purpose.

5.2.1 Suggested platforms for testing and implementation

It would be possible directly to implement the model of Figure 8 into several ongoing experiments, e.g. the Okapi project [21]. The user response to each element, extracted by the interface, will typically consist of few conceptual structures. The virtue of Okapi so far has been its conscious dedication to a realistic user environment in which searchers actually may carry variable and ill-defined information needs (boxes 2 and 3, Figure 3). The further advantage for Okapi would be the augmentation of potentially relevant data in the form of problem statements or work task description for (initial) probability retrieval. In addition, the data would be structured, that is, available for various forms of manipulation by the system in supporting the user's exploration and modification of statements.

The theory of cognitive polyrepresentation could be implemented in relation to the TREC experiments by enhancing the description of the predefined ‘information requests’, already supplied by the project. At present these contain a ‘Domain description’ and a ‘Request version’ of the need [20]. In addition, each request is explained in more depth by a ‘Description’ and a detailed essay is provided on the kind of material which is supposed to be relevant. This essay, or relevance assessment description, corresponds to the notion of ‘Preference’ (or could be interpreted as ‘Goals’), on the extreme right-hand side of Figure 2. Hence, we suggest adding to, or preferably replacing, the essay with a realistic ‘Work Task/Interest Description’. This enhancement into intentionality would make it possible during the relevance judgement and human request modification processes [95] to know *why* an entity ought to be relevant.

Essentially Figure 8 demonstrates how several statements can be extracted from the user's cognitive space and applied to several different methods of representation, but using one retrieval engine. Figure 5 attempts to show how *one* cognitive statement, e.g. a formulated request version, can be processed in information space during one run by polyrepresentative means. The Venn diagram demonstrates the request being retrieved by the overlap of structures of different cognitive (and functional) origin: author, indexer, and thesaurus generated contexts, by means of *one* retrieval logic, e.g. the Boolean logic or a probabilistic model (see also the Appendix).

Similar but different overlaps of semantic entities could be retrieved simultaneously during the same run by means of other retrieval functions. A total or partial intersection of the various IR engines could be performed, and the conceivable overlap between them would, in a cognitive sense, point to the most probably 'relevant' semantic entities (Total Cognitive Overlap, Figure 6). It will be seen that the searcher or the system is free to explore the individual sets, and the range of cognitive overlaps constituted by the different retrieval functions (xyz, xy, xz, yz) can be explored.

A different set of diagrams would display a similar pattern associated with another statement extracted from the searcher, e.g. a work task description. The new pattern could be intersected with the former one, and truly novel cognitive overlaps discovered. These overlaps would signify the probability that both request and task situation might be satisfied by some, very few, semantic entities simultaneously. The next run based on relevance feedback, i.e. interpretations in relation to both cognitive structures, would modify the search output and create new alternative patterns and paths in the information space to be overlapped, explored or fused. Conceptually and operationally, this exercise is prohibitive if carried out by human beings: but what do we have machines for?

6. CONCLUDING DISCUSSION

The cognitive approach to IR theory makes use of the major discoveries, experiences, investigations, and experimental tests made over more than half a century in information science. It is supported by substantial empirical evidence from both information seeking research and from studies made in the operational and informetric environments. In addition, it does not contradict the IR developments achieved so far with partial matching methods.

The concept of polyrepresentation of the user's cognitive space involves representing not only the current (often topical) information need, but also (and more importantly) the underlying problem space, actual work task or interest, and the dominant work domain(s). These elements are associated with the information need formation by following a principle of causality, that is, the intentionality underlying the fact of having such a need at all. Information needs internal to the user are basically regarded as variable, but may also exist in stable, rich or ill defined forms according to the nature of the current cognitive state of the user in a given seeking situation. The properties of the representations of all the elements in terms of request, cognitive state, problem and work task formulations may point to the actual cognitive condition of the intrinsic information need.

The same principle can be applied in a different manner to the objects of information space. By making calculated use of the variety of representations of such objects, or semantic entities, generated over time by agents with different cognitive and intentional origins one may exploit the cognitive overlaps of objects retrieved by such representations and IR techniques. It thus becomes possible to link the spaces of searchers and information in a complex but structured manner, providing exploratory alternatives to support the searcher's information seeking effort.

The ongoing large-scale laboratory experimental settings are found valid, but confined to simulated investigations of invariable information needs represented by well-formulated and predefined requests for information. This Cranfield-like approach conforms to a very limited case of retrieval in both realistic and cognitive terms. On the other hand, the research efforts have produced valid platforms which can be implemented in natural searcher environments for further development and testing. In this confined experimental scenario we suggest, as a minimum, that researchers should also monitor problem and work task relationships by adding simulated situational descriptions of the purpose of presenting each defined request to a system. It thus becomes possible to compare topical and situational relevance for the same request for different retrieval engines. Following the theory, however, simulated retrieval based on artificial requests (although increasingly including searcher participation) cannot replace all real-life information need situations. The initial as well as the ensuing cognitive conditions for the IR process are different, mainly due to the influence of natural intentionality and the dynamic properties of the cognitive space. Realistic situations ought to be investigated in full, since we do not know, for instance, the effects of applying several retrieval engines simultaneously to variable or ill-defined natural needs for information. In view of the theory, each formal retrieval technique processes signs, not information, at a lexical linguistic level. By parallel use of a variety of search techniques by searchers we may produce hitherto unknown synergetic advantages and novel paths to information. Certain network-based and linguistically-based retrieval engines work at a syntactic and structural level. Probably only the logical uncertainty and information calculus principles may provide openness towards a contextual and semantic level. This latter approach to IR is consequently regarded as being compatible with the cognitive theory. By incorporating polyrepresentative principles we come closer to achieving contextualisation in IR, as seen from the searcher's point of view. The goal of retrieval research is to provide means to stimulate not simulate the searcher's IR endeavour. Hence, in realistic terms, and in order to investigate the exploratory and cognitive quality of different principles of representation and best matching, the experimental settings should in addition attempt to apply initially non-existing or extremely ill-defined requests with rich problem or task definitions.

For 'relevance' assessment the theory suggests and allows for relative, but statistically reliable, inter-subjective assessments, regardless of the form of 'relevance' that is used. This relativity is found in the current TREC and Tipster investigations at the experimental level. The theory takes any pre-defined or post-defined relevance ratio per request as established in ways that are uncertain, similar to searcher assessments. In a statistical sense, all laboratory experiments based on a consensus decision by a panel or on one person's initial base-line judgement are uncertain and unreliable. The decade-old inter-indexer as well as the recent investigations of inter-linker and inter-searcher consistency investigations prove this. Thus, the theory suggests that all the inconsistencies that are found in IR are unavoidable manifestations of like cognitive and intentional dispositions. Since the very nature of IR is characterised by uncertainty and unpredictability, these problems should not be eliminated from the simulated laboratory experiments, since they cannot be avoided in real-life situations. Instead we believe that these inconsistency phenomena are highly valuable and worthy of exploration, favouring a more precise understanding of IR as a process of intentional interpretation and cognition.

Partial ‘relevance’ of information objects is proposed and effectively follows the presented approach. Information objects should be judged ‘relevant’ in a non-binary sense. A second dimension of partial ‘relevance’ arises, since the semantic entities, which may not be identical to entire documents, are regarded as the main objectives of assessment and retrieval. The question is: which objects are to form the base-line evaluation: bibliographic details, i.e. titles and/or abstracts; the full-text document; each single or several semantic entities or passages? The theory also encourages the inclusion of differentiated relevance, implying that ‘topicality’ relevance is supplied with work task and problem relevance, or rather, forms of situational usefulness: several objects may be judged topically relevant, but only be partially relevant or irrelevant for a given work task/interest and problem, and vice versa. A single semantic entity on its own may indeed shift in degree of usefulness in relation to a particular intentionality over time. Effectively, all factors controlling the information life-cycle also influence retrieval and relevance.

The enhanced simulated and real-life experimental situations are expected to provide a more detailed understanding of IR, although, of course, the number of experimental variables will increase further when compared to current research. The theoretical approach is testable, preferably in non-simulated scenarios, but requires extensive populations to be statistically valid. Because information *per se* is unpredictable the theory does not claim to make formal predictions of the search outcome in particular situations. However, its general prediction is that the more remote in time and by cognitive or functional type the access structures are to a given data set, the higher the probability that the objects retrieved by an overlap of such structures are useful to a given search situation. Because of the time dimension, the unpredictability, and the range of interpretation levels, IR can be thought of as a discontinuous, non-linear process.

The significant point is to recognise that for a system of information objects, there exist alternative cognitive potentials at each given point in time; only one or a few are uncovered, depending on the initial condition and the intentional interpretation that is made of them by an agent or observer.

ACKNOWLEDGEMENTS

I wish to thank the referees for their useful suggestions. In particular, I wish to thank Professor Peter Willett for his kind support during my struggle with his mother tongue.

APPENDIX: COGNITIVE ONLINE SEARCHING APPLYING POLYREPRESENTATION AND QUORUM PRINCIPLES

For the sake of simplicity the principle is shown for a search consisting of two terms or concepts. The search can be carried out for example in the Dialog or ESA-IRS databases where * represents logical AND and + represents logical OR.

Given terms: **A** and **B**, where **A** < **B** for values of document frequency (*df*) both in the title (/TI) and descriptor/identifier fields (/DE,ID) in the collection (= no. of online hits), i.e. *df***A**/TI < *df***B**/TI < *df***A**/DE,ID < *df***B**/DE,ID.

The cognitive quorum principle implies that we should constantly intersect terms originating from different cognitive agents for the same information objects. This requires intersecting the terms in such a way that they are present simultaneously as far as possible in the appropriate record fields, that is primarily in the /TI field for author[§] and the /DE,ID fields for indexer. Additionally, the terms with the lowest frequencies are combined before terms showing higher *df*-values until all possible combinations have been exhausted. Previous sets are subtracted to produce unique sets. (The term with the highest frequency in the collection (= no. of online term postings using the command 'set postings on' in Dialog) is given priority if terms are holding equal/similar *df*-values.)

ss A/TI * B/TI * A/DE,ID * B/DE,ID	set 1 (kernel set combination)
ss A/TI * B/TI * A/DE,ID NOT (S1)	set 2 (second priority set)
ss A/TI * B/TI * B/DE,ID NOT (S 1+S2)	set 3 (third priority set)
ss A/TI * A/DE,ID * B/DE,ID NOT (S1+S2+S3)	set 4 (etc.)
ss B/TI * A/DE,ID * B/DE,ID NOT (S1+S2+S3+S4)	set 5 (etc.)
ss A/TI * B/DE,ID NOT (S1+S2+S3+S4+S5)	set 6 (etc.)
ss B/TI * A/DE,ID NOT (S1+...+S6)	set 7 (etc.)

From here on, no more cognitive combinations by different terms are possible.

ss A/TI * B/TI NOT (S 1 +... +S7)	set 8 (both terms by author origin)
ss A/DE,ID * B/DE,ID NOT (S 1+...+S8)	set 9 (both terms by indexer origin)

§ By introducing functional difference within the same cognitive origin, i.e. including the terms (A and B) searched in the abstract and/or full-text fields, the sets with non-zero values can be further divided into specific groupings of objects. For instance, use the command:

ss A(10N)B/AB * A(S)B/TX

for text objects with both fields present; otherwise separately. (10N means at most ten words' distance between A and B in the abstract field; S implies that A and B are found within sentences/paragraphs in full-text fields.)

No more combinations by different terms are possible. (S 1 +..+S9)= ss **A** * **B** = the common online retrieval method using the basic index.

ss A/TI NOT (S1+ ...+S9)	set 10
ss B/TI NOT (S1+...+S10)	set 11
ss A/DE,ID NOT (S1+...+S11)	set 12
ss B/DE,ID NOT (S1+...+S12)	set 13
ss (S 1 +...S 13) = ss **A** + **B**	set 14 (the union set)

The sets 1 to 7 display a ranked list of sets derived from intersections of cognitively different structures of representation. The recently introduced Target command allows for simplistic ranking of documents retrieved in a given set [79]. Target is applying quorum logic (in the traditional way), document term frequencies and collection term frequencies as elements of its ranking algorithm. A Target search on the terms A and B can be intersected with each of the sets 1 to 7 and the overlapping retrieved documents can be isolated. These documents constitute cognitive overlaps of two different retrieval engines (Figure 6). Alternatively, Target can be applied to each set above in order to rank the documents in order of simplistic 'probability of relevance'.

REFERENCES

1. VAN RIJSBERGEN, C. *and* LALMAS, M. An information calculus for information retrieval. *Journal of the American Society for Information Science* (pre-printed version; forthcoming).
2. SALTON, G., ALLAN, J. *and* BUCKLEY, C. Approaches to passage retrieval in full text information systems. *In: 16th ACM-SIGIR Conference Proceedings, Pittsburg, 1993.* New York, NY: ACM Press, 1993, 49-58.
3. INGWERSEN, P. *Information retrieval interaction.* London: Taylor Graham, 1992.
4. VAN RIJSBERGEN, C. A new theoretical framework for information retrieval. *ACM-SIGIR Conference Proceedings.* Pisa: IEI, 1986, 194-200.
5. VAN RIJSBERGEN, C. The science of information retrieval: its methodology and logic. *In: Conference Informatienvetenschap in Nederland.* Den Haag: Rabin, 1990, 20-38.
6. DE MEY, M. The relevance of the cognitive paradigm for information science. *In:* HARBO, O. *et al., eds. Theory and application of information research.* Mansell: London, 1980, 48-61.
7. BROOKES, B.C. The foundations of information science: Part 1: Philosophical aspects. *Journal of Information Science, 2, 1980,* 125-133.
8. BELKIN, N.J. Information concepts for information science. *Journal of Documentation, 34, 1978, 55-85.*
9. SPARCK JONES, K. *Retrieving information or answering questions.* The 8th British Annual Research Lecture. London: British Library, 1989. (BLRD/c/130)
10. BELKIN, N.J., BROOKS, H. *and* DANIELS, P.J. Knowledge elicitation using discourse analysis. *International Journal of Man-Machine Studies, 27, 1987,* 127-144.

11. CROFT, W.B. *and* THOMSON, R. I3R: A new approach to the design of document retrieval systems. *Journal of the American Society for Information Science, 38(6),1987, 389-404.*
12. CROFT, W.B. Approaches to intelligent information retrieval. *Information Processing and Management, 23(4), 1987, 249-254.*
13. INGWERSEN, P. Search procedures in the library analysed from the cognitive point of view. *Journal of Documentation, 38, 1982, 165-191.*
14. KUHLTHAU, C. *Seeking meaning.* New York: Ablex, *1993.*
15. RASMUSSEN, J., PEJTERSEN, A.M. *and* GOODSTEIN, L.P. *Cognitive engineering: concepts and applications.* Wiley, 1992.
16. ELLIS, D. *New horizons in information retrieval.* London: Library Association, 1990.
17. ELLIS, D. The physical and cognitive paradigms in information retrieval research. *Journal of Documentation, 48(1), 1992, 45-64.*
18. INGWERSEN, P. The cognitive viewpoint in IR. *Journal of Documentation, 49(*1), 1993, 60-64.
19. BLAIR, D.C. *Language and representation in information retrieval.* Oxford: Elsevier, 1990.
20. HARMAN, D. Overview of the first TREC Conference. *In: ACM-SIGIR Proceedings, Pittsburgh, 1993.* New York, NY: ACM Press, *1993, 36-47.*
21. HANCOCK-BEAULIEU, M. Query expansion: advances in research in online catalogues. *Journal of Information Science, 18,* 1992, 99-103.
22. BELKIN, N.J. *and* VICKERY, A. *Interaction in information systems.* British Library, London, *1985.* (LIR Report 35)
23. CUADRA, C.A. *and* KATTER, R.V. Opening the black box of ‘relevance’. *Journal of Documentation, 23,* 1967, 291-303.
24. SARACEVIC, T. Relevance: a review of and framework for thinking on the notion in information science. *Journal of the American Society for Information Science, 26,* 1975, 321-343.
25. SU, L. The relevance of recall and precision in user evaluation. *Journal of the American Society for Information Science, 45, 1994, 207-217.*
26. BRUCE, H. A cognitive view of the situational dynamism of user-centered relevance estimation. *Journal of the American Society for Information Science, 45(3),* 1994, 142-148.
27. BARRY, C.L. User-defined relevance criteria: an exploratory study. *Journal of the American Society for Information Science, 45(3),* 1994, 149-159.
28. SCHAMBER, L., EISENBERG, M. *and* NILAN, M. A re-examination of relevance: toward a dynamic, situational definition. *Information Processing and Management, 26(6),* 1990, 755-776.
29. HARTER, S.P. Psychological relevance and information science. *Journal of the American Society for Information Science, 43, 1992, 602-615.*
30. KATZER, J., MCGILL, M.J., TESSIER, J.A., FRAKES, W. and DAS GUPTA, P. A study of the overlap among document representations. *Information Technology: Research and Development, 2,* 1982, 261-274.
31. TENOPIR, C. *Issues in online database searching.* Englewood Cliffs, NJ: Libraries Unlimited, 1989.

32. LANCASTER, F.W. *Indexing in theory and practice.* London: Library Association, 1991.
33. BORGMAN, C. All users of information retrieval systems are not created equal: an exploration into individual differences. *Information Processing and Management, 25(3),* 1989, 237-252.
34. FENICHEL, C.H. An examination of the relationship between searching behaviour and searcher background. *Online Review, 4*(4), 1980, 341–347.
35. WAPLES, D. The relation of subject interest to actual reading. *Library Quarterly, 2,* 1932, *42-70*
36. ELLIS, D. A behavioural approach to information retrieval system design. *Journal of Documentation, 45*(3), 1989, 171–212.
37. BRITTAIN, J.M., *ed. The Social Sciences: the supply and demand for documentation and data.* London: Rossendale, *1982.*
38. MARTYN, J. Information needs and users. *ARIST, 9, 1974, 3–23.*
39. HJØRLAND, *B. and* ALBRECHTSEN, H. Toward a new horizon in information science: domain-analysis. *Journal of the American Society for Information Science, 46*(6), 1995, 400–425.
40. BATES, M.J. Information search tactics. *Journal of the American Society for Information Science, 30*(4), 1979, 205–214.
41. BATES, M.J. Idea tactics. *Journal of the American Society for Information Science, 30,* 1979, 280–289.
42. BATES, M.J. An exploratory paradigm for online information retrieval. *In:* BROOKES, B.C., *ed. Intelligent information systems for the information society.* London: North-Holland, 1986, 91–99.
43. BATES, M.J. The design of browsing and berry-picking techniques for the online search interface. *Online Review, 13(5),* 1989, 407–424.
44. TENOPR, C. *and* CAHN, P. Target & Frestyle: Dialog and Mead join the relevance ranks. *Online,* May, 1994, 31–47.
45. MACKAY, D.M. What makes the question? *The Listener, 63,* May 5, *1960,* pp. 789–790.
46. DERVIN, B. *and* NILAN, M. Information needs and uses. *ARIST, 21,* 1986, 3-33.
47. HJØRLAND, B. *Emnerepræsentation og informationssøgning. Bidrag til en teori på kundskabsteoretisk grundlag.* Göteborg: Valfried, 1993. (Doctoral dissertation)
48. WINOGRAD, T. *and* FLORES, C.F. *Understanding computers and cognition.* Norwood, NJ: Addison-Wesley, *1986.*
49. INGWERSEN, P. Cognitive analysis and the role of the intermediary in information retrieval. *In:* DAVIES, R., *ed. Intelligent information systems.* Chichester: Horwood, 1986, 206–237.
50. BELKIN, N.J., ODDY, R. *and* BROOKS, H. ASK for information retrieval. *Journal of Documentation, 38,* 1982, Part I: 61–71; Part II: 145–164.
51. BELKIN, N.J. Cognitive models and information transfer. *Social Science Information Studies, 4,* 1984, 111–129.
52. SARACEVIC, T., MOKROS, H. *and* SU, L.T. Nature of interaction between users and intermediaries in online searching: a qualitative analysis. *ASIS Proceedings,* 1990, 47–54.

53. FIDEL, R. Searchers' selection of search keys. *Journal of the American Society for Information Science, 42(7),* 1991; I. The selection routine, 490—500; II. Controlled vocabulary and free-text searching, 501-514; III. Searching styles, 515—527.

54. BYSTRÖM, K. *and* JÄRVELIN, K. Task complexity affects information seeking and use. *Information Processing and Management, 31(2),* 1995, 191—214.

55. INGWERSEN, P. Polyrepresentation of information needs and semantic entities. *In: 17th ACM-SIGIR Conference Proceedings, Dublin, 1994.* London: Springer, 1994, 101-110.

56. SEARLE, J.R. Intentionality and its place in nature. *Synthese,* 61, 1984, 3-16.

57. HANCOCK-BEAULIEU, M., FIELDHOUSE, M. *and* DO, T. An evaluation of inter-active query expansion in an online library catalogue with a graphical user interface. *Journal of Documentation, 51*(3), 1995, 225—243, and personal communication with first author, 1995.

58. EFTHIMIADES, E.N. A user-centered evaluation of ranking algorithms for interactive query expansion. *In: 16th ACM-SIGIR Conference Proceedings, Pittsburg, 1993.* New York: ACM Press, 1993, 146—159.

59. ALLEN, B. Perceptual speed, learning and information retrieval performance. *In: 17th ACM-SIGIR Conference Proceedings, Dublin, 1994.* London: Springer 1994, 71-80.

60. MORITA, M. *and* SHINODA, Y. Information filtering based on user behaviour analysis and best match text retrieval. *In: 17th ACM-SIGIR Conference Proceedings, Dublin, 1994.* London: Springer, 1994, 272-281.

61. TURTLE, H. *and* CROFT, W.B. Inference methods for document retrieval. *In:* VIDICK, J., *ed. ACM-SIGIR Conference Proceedings, Bruxelles, 1990.* Bruxelles University, 1990, 1-24.

62. BLAIR, D.C. *and* MARON, M.E. An evaluation of retrieval effectiveness for a full-text document retrieval system. *Communications of the ACM, 28(3),* 1985, 289-299.

63. BELKIN, N.J., COOL, C., CROFT, W.B. *and* CALLAN, J.P. The effect of multiple query representations on information retrieval performance. *In: 16th ACM-SIGIR Conference Proceedings, Pittsburg,* 1993. New York: ACM Press, 1993, 339-346.

64. IIVONEN, M. Searchers and searchers: differences between the most and least consistent searchers. *In: 18th ACM-SIGIR Conference Proceedings, Seattle,* 1995. New York: ACM Press, 1995.

65. SALTON, G. The state of retrieval system evaluation. *Information Processing and Management, 28(4),* 1992, 441-449.

66. WILSON, P. Situational relevance. *Information Storage and Retrieval, 9,* 1973, 457—471.

67. ROBERTSON, S.E. *and* SPARK JONES, K. *Simple, proven approaches to text retrieval.* Cambridge: University of Cambridge, Computer Laboratory, December 1994. (Technical Report, No. 356)

68. ROBERTSON, S.E. *and* HANCOCK-BEAULIEU, M. On the evaluation of R systems. *Information Processing and Management, 28(4),* 1992, 457-466.

69. SARACEVIC, T. *and* KANTOR, P.B. A study of information seeking and retrieving: III. Searchers, searches, and overlaps. *Journal of' the American Society for Information Science, 39(3),* 1988, 197-216.
70. PAO, M.L. Relevance odds of retrieval overlaps from seven search fields. *Information Processing and Management, 30*(3), 1994, 305-314.
71. ELLIS, D., FURNER-HINES, J. *and* WILLETT, P. On the creation of hypertext links in full-text documents: measurement of inter-linker consistency. *Journal of Documentation, 50*(2), 1994, 67-98.
72. TELL, B. Document representation and indexer consistency. *ASIS Proceedings, 6,* 1969, 285-291.
73. MCCAIN, K.W. Descriptor and citation retrieval in the medicine behavioural sciences literature: retrieval overlaps and novelty distribution. *Journal of the American Society for Information Science, 40, 1989, 110-114.*
74. PAO, M.L. Term and citation searching: a field study. *Information Processing and Management, 29*(1), 1993, 95-112.
75. HARTER, S.P., NISONGER, T.E. *and* WENG, A. Semantic relationships between cited and citing articles in library and information science journals. *Journal of the American Society for Information Science, 44*(9), 1993, 543-552.
76. HJORTGAARD CHRISTENSEN, F. *and* INGWERSEN, P. Fundamental methodological issues of data set creation online for the analyses of research publications. *In: 5th International Conference on Scientometrics and Informetrics,* Chicago, June 1995, 103-112.
77. VAN RIJSBERGEN, C. *and* AGOSTI, *M., eds.* Special issue: Information Retrieval. *Computer Journal, 35(3),* 1992, 193-298.
78. BELKIN, N.J. *and* CROFT, W.B. Retrieval techniques. ARIST, 22, 1987, 109-145.
79. KEEN, M.E. How does Dialog's Target work? *Online & CDROM Review, 18*(5), 1994, 285-288.
80. SALTON, G. *and* BUCKLEY, C. Automatic text structuring and retrieval. Experiments in automatic encyclopedia searching. *In: 14th ACM-SIGIR Conference Proceeding, Chicago, 1991.* New York: *ACM* Press, 1991, 21-30.
81. ATHERTON-COCHRANE, P. *Books are for use: final report of the Subject Access Project.* Syracuse, NY: Syracuse University, School of Information Studies, 1979.
82. WORMELL, I. *Subject Access Project - SAP: improved subject retrieval for monographic publications.* Lund: Lund University, 1985. (Doctoral dissertation)
83. SMEATON, A. Progress in the application of natural language processing to information retrieval tasks. *Computer Journal, 35,* 1992, 268-278.
84. SMEATON, A., VOUTILAINEN, A. *and* SHERIDAN, P. The application of morpho-syntactic language processing to effective text retrieval. *In: Esprit 90 Conference Proceedings.* Dordrecht: Klüver, 1990, 619-635.
85. GIBB, F. Knowledge-based indexing in SIMPR: integration of natural language processing and principles of subject analysis in an automated indexing system. *Journal of Document and Text Management, 2,* 1993, 131-154.

86. KWOK, K.L. Experiments with document components for indexing and retrieval. *Information Processing and Management, 24,* 1988, 405-417.
87. FRITH, A.R., ROBERTSON, A.M. *and* WILLETT, P. Effectiveness of similarity measures and of query expansion techniques for searching databases of 16th, 17th and 18th century English text. *Journal of Document and Text Management, 2,* 1993, *155-172*.
88. SPARCK JONES, K. Personal communication, January-April, 1995.
89. BOOKSTEIN, A. Towards a multi-disciplinary Bradford Law. *Scientometrics,* 30(1), 1994, 353-361.
90. SCHMELTZ PEDERSEN, G. A browser *for* bibliographic information retrieval, based on an application of lattice theory. *In: 16th ACM-SIGIR Conference Proceedings, Pittsburgh, 1993.* New York: ACM Press, 1993, *270-280*.
91. PEJTERSEN, A.M. *Interfaces based on associative semantics for browsing in information retrieval* Roskilde, Denmark: Risø National Laboratory, 1991.
92. INGWERSEN, *P. and* WORMELL, I. Ranganathan in the perspective of advanced information retrieval. *Libri, 42(3),* 1992, 184-201.
93. BRAJNIK, G., GUIDA, G. *and* TASSO, C. User modelling in intelligent information retrieval. *Information Processing and Management, 23*(4), 1987, 305-320.
94. KANTOR, P.B. Two heads are better than one: the potential of data fusion concepts for improvement of online searching. *In: Proceedings of the 13th National Online Meeting, New York,* 1992, 147-151.
95. CALLAN, J.P. *and* CROFT, W.B. An evaluation of query processing strategies using the TIPSTER collection. *In: 16th ACM-SIGIR Conference Proceedings, Pittsburgh, 1993.* New York: ACM Press, 1993, 347-355.

(Revised version received 9 October 1995)