A Cognitive Analysis of Information Retrieval

Stefano Mizzaro

The lackness of a formal account is probably one of the most evident of the shortcomings of information retrieval: concepts like information, information need, and relevance are neither well understood nor formally defined. This paper sketches a cognitive framework that permits to analyze these three central concepts of the information retrieval scenario.
The framework consists of concepts as cognitive agents acting in the world, knowledge states possessed by the cognitive agents, transitions among knowledge states, and inferences. On the basis of the framework, information is formally defined as a pair representing the difference between two knowledge states; this definition permits to clarify the distinction among data, knowledge, and information and to discuss the subjectiveness of information. On this ground, the concept of information need is examined: it is defined, it is studied in the context of the interaction between an information retrieval system and a user, and the well known classification in verificative, conscious topical, and muddled needs is analyzed. On the basis of the above definitions of information and information need, relevance is formally defined, and some critical features of this concept are discussed.
Keywords: Information retrieval, cognitive view, theoretical foundations, information, information need, relevance.

1 Introduction

Still today, the concepts of information, information need, and relevance, central ones in information retrieval (Salton, 1989; van Rijsbergen, 1979) and information science in general, are not well understood. This is perhaps the most evident manifestation of the lackness of a comprehensive theory of information retrieval, looked for by many researchers. In this paper I propose a cognitive framework on which basis such a comprehensive theory of information retrieval might be developed. The adequacy of the framework is then evaluated by using it for better understanding the three above mentioned central concepts of information retrieval (information, information need, and relevance).

The paper is structured as follows: in Section 2 a cognitive scenario is presented; on this basis, information (Section 3), information need (Section 4), and relevance (Section 5) are defined and analyzed. The last section summarizes the work done so far and sketches the future developments of this line of research.

2 A cognitive scenario

The importance of the cognitive view in information retrieval, and the powerful and adequacy of using cognitive instruments in the information retrieval field are largely recognized, and witnessed,

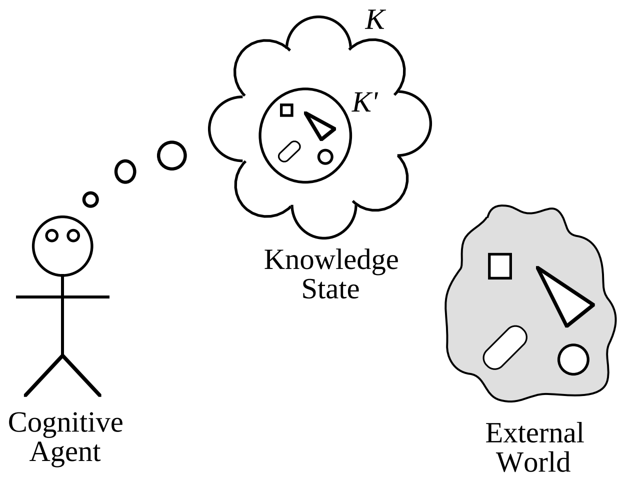

Figure 1: An agent, his KS, a part of the world, and its representation.

for instance, by accomplishments like the MEDIATOR and MONSTRAT models. I have not enough space to discuss here these issues, that are anyway well analyzed in (Ingwersen, 1992).

This section describes in an intuitive way some cognitive concepts: cognitive agents, knowledge states, knowledge items, transitions among knowledge states, and inferences.

2.1 Agents, knowledge states, and knowledge items

I assume that the world is populated by *Cognitive Agents* (henceforth simply *agents*), that each agent possesses a *Knowledge State* (KS), and that it is possible to separate an agent and his (her, its) KS from the *'External' World* (henceforth simply *world*), whatever it may be. Through his perception system, an agent *perceives* (a portion of) the world and *represents* it into his KS. The portion of the KS that corresponds to a portion of the world is said the *representation* of that portion of the world into the KS, and it can be more or less correct (i.e. corresponding to the world) and complete (i.e. taking into account every aspect of the world). An agent, on the basis of his KS, can *act* in the world.

Each KS is a collection of "atomic" components, that I call *Knowledge Items* (KI). Each KS is thus a set of KIs, and in the following I use some of the usual symbology of set theory, as $\in$ (belong), $\subseteq$ (subset), \ (set difference), $\varnothing$ (empty set), $\cup$ (union), $\cap$ (intersection), and so on, with the usual meaning extended to KSs and KIs.

Figure 1 illustrates intuitively the scenario presented so far: an agent perceives a portion of the world and represents it in *K'*, a subKS of his whole KS *K*.

The reader can imagine many alternatives for having a more concrete picture of KSs (and KIs), for instance: logical theories (i.e. sets of logical formulas) (Genesereth & Nilsson, 1987); semantic nets (Sowa, 1991); sets of beliefs (Genesereth & Nilsson, 1987) situations (Barwise & Perry, 1983; Devlin, 1991; Dretske, 1981), see also (Bruza, 1993; Lalmas & van Rijsbergen, 1996); recursive models (Mizzaro, 1994; Mizzaro, 1996a); minds and ideas (Bateson, 1979); and so on. I do not take position among these (and many others) alternatives in this paper, and I try to remain at a level of abstraction enough general for comprising all of them.

Let us go more deeply inside a KS. Whatever KSs and KIs are, I suppose that some *links* exist among the KIs and the KSs, similarly to what happens in Truth Maintenance Systems (Doyle,

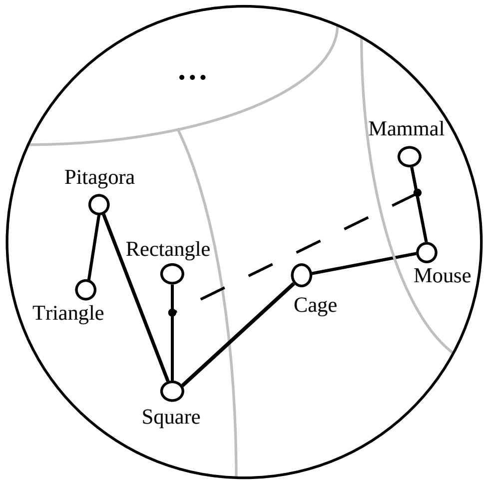

Figure 2: Links among and partitions of KSs.

1979) or in Recursive Models (Mizzaro, 1994; 1996a). This means that a KS can be partitioned into subKS, each partition containing the KIs more strongly linked. See for instance Figure 2: the KIs regarding, say, Euclidean Geometry (as the concepts of triangle, square, Pitagora's Theorem, and so on) belong to one partition, while the KIs regarding, say, mammals belong to another partition. These partitions, besides being subjective, are neither absolute nor clear-cut: it is (almost?) always possible to find a link chain between two KIs or KSs. It is a fuzzy, or perhaps fractal, situation. For instance, it is possible to link, say, a mammal with a square through the KSs about mice and cages. Furthermore, the links themselves are a kind of KIs, in order to, for instance, have links among links (and so on), as the dashed one in Figure 2 (that links two 'IsA' links).

The above assumptions are widely spread in many fields, for instance: artificial intelligence (under the label "logicism", see Genesereth & Nilsson, 1987; Nilsson, 1991), situation semantics (Devlin, 1991), cognitive science (Gardner, 1987), and human-computer interaction (Dix et. al., 1993). They are criticizable from many points of view (see for instance Birnbaum, 1991; Maturana & Varela, 1992; Mizzaro, 1995a), but they will be useful in the following of this paper for describing the interaction between a user and an information retrieval system. Thus, I do not take them as established truths, but as useful work hypotheses: for the sake of brevity, I avoid to analyze the (many) philosophical implications of these issues. In the same way, being the KS the only component of an agent that is analyzed here, I assume that the perceptual systems of different agents are similar, though this is obviously a rough abstraction.

2.2 Transitions among knowledge states

The KS of an agent may change as time goes on: when this happens, I say that a *transition* between an initial KS K^I and a final KS K^F takes place. A transition can take place for two different reasons:

- by (internal) inference: the agent reasons, reflects, and modifies his KS without any input from the world. This will be called *inferential transition*; it is the only kind of transition that can take place for an agent without a perception system;
- by receiving information: through his perception system, the agent perceives something (a

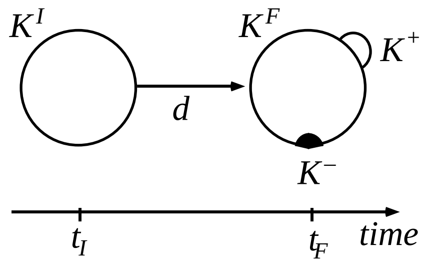

Figure 3: Initial and final KS of a transition.

datum) from the world and this leads to the modification of his KS (a transition into another KS). This will be called *noninferential transition*. If the datum leads to a change of the KS then the datum is said to *carry* information. Note that everything can be a datum, also nothing (i.e. no receiving from the world), because nothing is different from something, and so it can carry information (Bateson, 1979). This is the reason for distinguishing between an inference and a null datum.

The modification of a KS that takes place during a transition can be specified by what is added to the KS (a subKS here indicated by K^+) and what is removed from the KS (K^-). In Figure 3 the two KSs K^I (the initial one) and K^F (the final one) are represented by circles, the subKS K^+ added to the KS by the little white semicircle on the border of the final KS, the removed subKS K^- by the little black semicircle, and the transition between the two KSs by an arrow labelled by the corresponding datum (or by 'infer' if it is an inferential transition). Finally, at each KS can be associated a time instant (in figure, t_I and t_F are the time instants of K^I and K^F, respectively). Note that, besides adding new KIs (K^+), a datum can also lead to the removal of some subKS (K^-). This happens, for instance, when a fact is believed true in the KS before a transition and false later: the KIs representing the truthness are removed and the KIs representing the falsehood are added.

Let me emphasize that a noninferential transition between two KSs is not a mere accumulation of knowledge, but involves a restructuring of the KS. The research in the field of *belief revision* (Alchóurron et. al., 1985; Gärdenfors, 1988; Katsuno & Mendelzon, 1991) is on this topic; three kinds of transitions are defined: *expansion* (just adding something to a KS), *contraction* (removing something), and *revision* (modifying something, i.e. a contraction followed by an expansion). All three of these have to preserve some conditions on the KS: *consistency* (an agent cannot believe both one thing and its negation) and *logical omniscience* (an agent believes all the logical consequences of anything he believes). There are many considerations that could be done on this issue (e.g., are logical omniscience and consistency too strong requirements?), but they would lead us too far. What is relevant here is the need of restructuring the KS after receiving a datum. On the basis of the above sketched framework, this is explained through the links: the KIs linked to the added or removed ones are affected too, in a recursive way.

Therefore, the noninferential transition caused by a datum d can be divided into two parts, as shown in Figure 4: a first *perception* transition (labelled with the datum d in figure) in which the datum is perceived and something is immediately added to or removed from K^I, obtaining K'; and a second *restructuring* transition (labelled with *r(d)* for 'restructuring because of datum *d*') in which the restructuring operation takes place, and the KIs linked to the added (K'^+) or removed

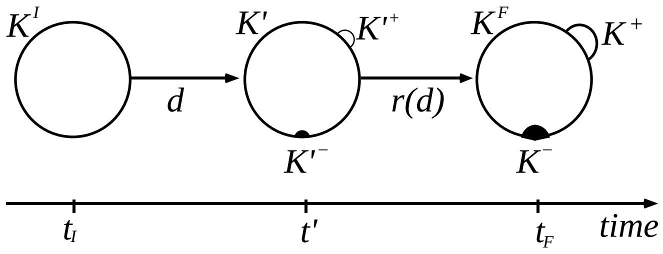

Figure 4: Perception and restructuring transitions.

(K'^-) KIs are affected.[1] Anyway, the resulting modification of the KS is fully represented by K^+ and K^-, thus in the following I will sometimes treat a noninferential transition as an atomic one.

Using the KSs and the transitions, it is possible to imagine a *network* (*à la* Kripke, see Hughes & Cresswell, 1968) of possible KSs of an agent: the nodes of the network are KSs, among which some are 'real' KSs (i.e. KSs before or later actually possessed by the agent) while other ones are possible KSs that do not become 'real' ones (i.e. the agent does not possess them, though he could); the arcs of the network are the transitions from one KS to another one. In Figure 5 some transitions among plausible KSs are represented. The KSs and transitions in the figure are the possible ones, but only one path from K_1 to K_9 is followed in the reality, for instance the one with the thickest lines, while the other KSs remain only plausible ones.

In the next three sections, the above introduced concepts (KS, KI, link, transition, K^+, K^-, and network of KSs) are applied to the *Information Retrieval* (IR) field, in order to analyze and better understand three central concepts of IR: information, information need, and relevance.

3 Information

In this section I propose a definition of information on the basis of the above introduced concepts. This definition leads to distinguish among data, knowledge, and information, to analyze the subjectiveness of information, and to go more deeply inside the KSs.

3.1 Data, knowledge, and information

I assume that *knowledge* exists only inside the agents' KSs (thus a book does not contain knowledge). A *datum* is an entity of the physical world that, once perceived by an agent, leads to a noninferential transition of a KS that changes, say, from K^I to K^F. When this happens, the datum is said to carry *information*.

On this ground, it is possible to define the information carried by the datum in (at least) two ways: (i) an *objective* way, in which information is in some way inherent in the datum; or (ii) a *subjective, contextual* way, in which information is not inherent in the datum, but depends also on the agent's KS. Here I follow the latter choice: referring to Figures 3 and 4, the information carried by the datum d in the transition from K^I to K^F is defined as the ordered pair

$$Information\ (d, K^I, K^F) = \langle K^+, K^- \rangle = \langle K^F \setminus K^I, K^I \setminus K^F \rangle,$$

[1] Of course, there are the problems of (i) understanding when a perceptual transition ends and a restructuring transition starts, and (ii) distinguishing between a restructuring transition and an inferential one. I do not further discuss these issues here.

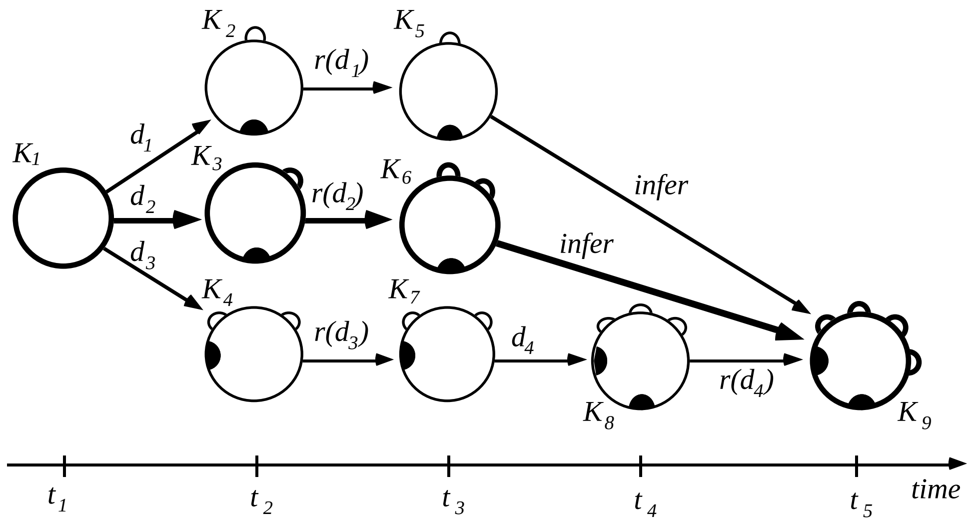

Figure 5: KS and transitions among them.

a way of expressing the *difference* between the two (final and initial) KSs.

Also Bateson (1972; 1979) and Brookes (1980) define information as a difference, but in ways that are different from the one proposed here. Following Bateson, an item of information is a difference (in the world) that makes a difference: on the basis of the above definitions, Bateson's difference is a perceived datum. Brookes proposes in his "fundamental equation"

$$K[S] + \Delta I = K[S + \Delta S]$$

that "information is a small bit of knowledge": a "knowledge structure" $K[S]$ is changed to a new knowledge structure $K[S + \Delta S]$ by the information ΔI. Brookes's view is more similar to the one proposed here than Bateson's one, but is anyway different: Brookes's knowledge and information are measured in the same units, and this does not hold for information as defined above.

With the subjective definition of information adopted here, the following two features hold (and they do not hold with the objective definition of information):

- The same datum can carry different information. For instance, if the datum is an utterance in some language, an agent understanding such a language can obtain information, while an agent not understanding the language cannot. A single 'bit' (i.e. an atomic datum, as 0/1, true/false, on/off) can carry a huge amount of information to an agent in an opportune KS, i.e. a KS with a high 'potential' (borrowing the term from physics) knowledge, in which a single bit triggers some transitions with a high difference between the initial KS and the final KS. The same datum can carry different information to two different agents or even to the same agent in different time instants: this can happen if the KSs of the two agents are different or if the KSs of the same agent in the two time instants are different.
- Two different data can carry the same information. For instance, an utterance uttered in two different languages carries the same information to an agent understanding both of the languages (and already knowing that the speaker knows both languages!). A number expressed through different 'formats' (8, VIII, 10_8, 1000_2, 20_4) carries the same information to an agent not 'sensible' to the difference of the base.

Hence, the KS plays a fundamental role in an agent receiving data: the information carried depends on the KS of the agent, and it should be said that a datum is 'interpreted' (not 'received') by an agent on the basis of his KS. It is possible to define an *interpretation* function

$$int: Data \times KS \to KS$$

that, given as argument a datum and a KS, assumes as value the KS resulting from the transition. On the basis of what said in Section 2.2 (see Figure 4), this function can be divided in the two components (*perception* and *restructuring*)

$$perc: Data \times KS \to KS$$
$$restr: Data \times KS \to KS,$$

in the following way:

$$int(d,K) = restr(d,perc(d,K)).$$

Despite of this evident subjectiveness of information, in everyday life the same datum sometimes (if not often) brings the same information to different agents. This may be explained assuming that the KSs of the agents that populate the real world (mostly human beings) are similar for genetical and social factors.[2] In this way, it becomes possible to speak of 'potential' information, and this is probably the reason for having an *Information Theory* (Shannon & Weaver, 1949), that should perhaps be called 'data' theory, in which the information is *objective*.

3.2 Prerequisite knowledge state

Supporting this view (that information is sometimes objective), note that the information received through a particular datum does not depend on the whole initial KS, but only on a *prerequisite* subKS (and thus the subjectiveness of information is less evident). This subKS, indicated by K^{pre}, must be such that the information received by the agent would not change if the initial KS of the agent were just K^{pre} instead of the whole KS. Let us see an example. An agent believes a wrong version of Pitagora's Theorem (for instance $a^2 + b^2 > c^2$, instead of the well known correct version $a^2 + b^2 = c^2$). When the agent receives the proof of the right version of the theorem (a datum), his KS changes accordingly. Referring to Figure 6 (in which the noninferential transition is taken as an atomic one), we have: K^I is the initial KS of the agent; d is the proof of the right version of the theorem; K^{pre} is the prerequisite KS and represents the notions of triangle, square, and so on, necessary for understanding the theorem; K^F is the final KS of the agent; K^+ is the subKS representing the right version of the theorem; and K^- is the subKS representing the wrong version of the theorem. Obviously, the KS of the agent may contain something more than K^+, K^-, and K^{pre}, but this is absolutely not relevant in this example.

Intuitively speaking, it is possible to characterize K^{pre} referring to what the user has to do with the received information (called 'work task or interest' in Ingwersen, 1996, and 'task' in Mizzaro, 1995b; Mizzaro, 1996b): the K^{pre} heavily depends on the intentions of the agent, on his aims and goals, that select a subKS of the initial KS bringing the K^{pre} to the agent's attention. In a more formal way, given an initial KS K^I, a final KS K^F, and a transition labelled with a datum d, K^{pre} is defined (using the *int* function) as a KS such that:

[2] This is less true if we consider people from different cultures, e.g. European vs. Asiatic, or different kinds of agent, e.g. human beings vs. computers. By the way, this might be an explanation of all the difficulties encountered in computer science, especially in artificial intelligence: a high difference between the KSs of the two kinds of agents, human beings and computers.

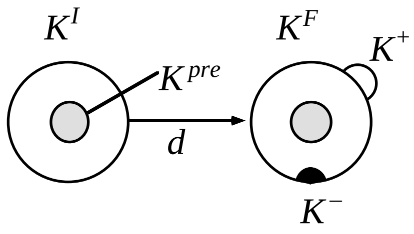

Figure 6: The interpretation of a datum does not depend on the whole KS.

(i) $K^{pre} \subseteq K^I$;

(ii) $\langle int(d,K^I) \setminus K^I, K^I \setminus int(d,K^I)\rangle = \langle int(d,K^{pre}) \setminus K^{pre}, K^{pre} \setminus int(d,K^{pre})\rangle$;

(iii) K^{pre} is minimal, i.e. $\neg\, \exists K \subseteq K^{pre}$ such that holds the previous property (ii).

Note that on the basis of this definition we obtain a restriction on K^- and K^{pre}: $K^- \subseteq K^{pre}$ (a particular case being $K^- = \varnothing$). This means that the subKS removed must be a part of the prerequisite KS, and this is quite reasonable. Figure 6 should be modified in this sense, and this will be taken into account in the following.

If we split the noninferential transition in the perception and restructuring ones, the situation becomes slightly more complex. The prerequisite KS changes after receiving *d*, as illustrated in Figure 7: K_1^{pre} is the prerequisite KS for the perception transition, K_2^{pre} for the restructuring transition, and the prerequisite KS for the whole noninferential transition is $K^{pre} = K_1^{pre} \cup K_2^{pre}$. The above conditions (i)–(iii) hold for such a K^{pre}, and analogous ones can be defined (using the *perc* and *restr* functions) for each of the two transitions:

(i_1) $K_1^{pre} \subseteq K^I$;

(ii_1) $\langle perc(d,K^I) \setminus K^I, K^I \setminus perc(d,K^I)\rangle = \langle perc(d,K_1^{pre}) \setminus K_1^{pre}, K_1^{pre} \setminus perc(d,K_1^{pre})\rangle$;

(iii_1) K_1^{pre} is minimal, i.e. $\neg\, \exists K \subseteq K_1^{pre}$ such that holds the previous property (ii_1);

(i_2) $K_2^{pre} \subseteq K'$;

(ii_2) $\langle restr(d,K') \setminus K', K' \setminus restr(d,K')\rangle = \langle restr(d,K_2^{pre}) \setminus K_2^{pre}, K_2^{pre} \setminus restr(d,K_2^{pre})\rangle$;

(iii_2) K_2^{pre} is minimal, i.e. $\neg\, \exists K \subseteq K_2^{pre}$ such that holds the previous property (ii_2).

4 Information need

The concept of *information need* has been studied for years by many researchers, among which:

- Mackay (1960) spoke of "incompleteness of the picture of the world", "inadequacy in what we may call his [the agent's] 'state of readiness' to interact purposefully with the world around him in a particular area of interest";
- Taylor (1968) spoke of visceral, conscious, formalized, and compromised information need, individuating four levels of question formation;
- O'Connor (1968) noted the ambiguous nature of the concept of information need;
- Belkin, Oddy, and Brooks (1982a; 1982b) spoke of "Anomalous State of Knowledge", the well known 'ASK';

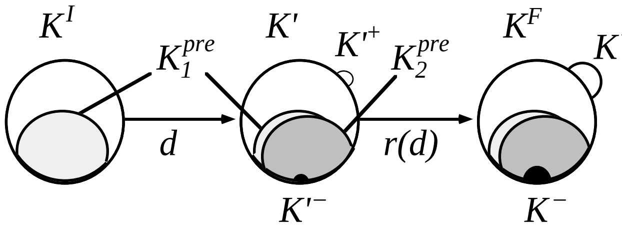

Figure 7: K^{pre} when considering perception and restructuring transitions.

- Ingwersen (1986; 1992) coined the ASK-like acronyms ISK (Incomplete State of Knowledge) and USK (Uncertain State of Knowledge) unifying these three acronyms in a common concept. He also proposed three fundamental types of information need (verificative, conscious topical, and muddled).

Notwithstanding these and many other studies, and the common usage of the term "information need" in the field of IR, still today the concept is neither understood nor defined. In this section it is analyzed on the basis of the above introduced cognitive scenario and definition of information: I propose a definition of information need, illustrate the activity of the user of an IR system, and describe the differences among the above mentioned three types of information need proposed by Ingwersen.

4.1 Definition of information need

What is an information need? Well, it is (obviously!) a *need* of *information*, and information is the "difference" between two KSs. A first attempt of graphically representing the situation is illustrated in Figure 8: an agent with an initial KS K^I does not posses the 'right' knowledge for solving a problem (or reaching an aim), and thus needs some additional information for reaching an adequate KS K^F, perhaps through some intermediate KSs (here and in the following figures an arrow containing three dots stands for a chain of transitions). The pair

$$\langle K^F \setminus K^I, K^I \setminus K^F \rangle = \langle \bigcup_i K_i^+, \bigcup_i K_i^- \rangle$$

is the information needed by the agent, in which the union of more K^+ and K^- indicates that the information may be obtained through subsequent steps.

But this is an uncomplete representation, because an information need can, in general, be satisfied in different ways: there is not a unique KS K^F in which the problem is solved, but there can exist different such KSs (named *final* KSs in the following) and different paths for reaching each of them. Figure 9 represents this more complete view: the dashed circles represent the final KSs.

Some of the final KSs of Figure 9 may be redundant: they do satisfy the need, but they contain also unuseful, or not used, additional knowledge. Only the minimal KSs among the dashed ones should be taken into account. In a more formal way, given the set

$$\boldsymbol{K} = \{K_1^F, K_2^F, K_3^F, \ldots\}$$

of all the final KSs, the redundant KSs can be eliminated normalizing $\boldsymbol{K}$ in $\boldsymbol{K^*}$ (the set of *minimal final* KSs) in the following way:

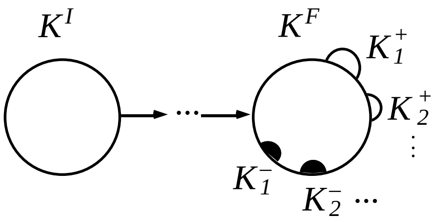

Figure 8: An information need as a difference of KSs.

$\boldsymbol{K^*} = \{K \mid K \in \boldsymbol{K} \;\&\; \forall K' \in \boldsymbol{K}\, (\neg\, \exists (K' \rightarrow K))\}$

(where $K' \rightarrow K$ stands for a transition from K' to K). In Figure 9 we have $\boldsymbol{K} = \{K_1^F, K_2^F, K_3^F, K_4^F\}$ and $\boldsymbol{K^*} = \{K_1^F, K_2^F, K_3^F\}$.

Now it is possible to define the information need in a KS K as a set of pairs, the set of the information items needed:

$Need(K) = \{\langle K^* \setminus K, K \setminus K^* \rangle \mid K^* \in \boldsymbol{K^*}\},$

so that the information needed in a KS K is the set of information items that permit to change the KS in a minimal final KS. Thus, 'to satisfy the information need of an agent in a KS K' means 'to give him one of the information items of the set $Need(K)$'.

Obviously, Figure 9 represents (some of) the plausible KSs of an agent, but only a few of them are real ones: the agent follows a single path. Moreover, it should be evident that it is impossible to know 'a priori' which is the information needed.

This is not the whole story. An agent represents the world, and also himself in the world. Thus, an agent represents in his KS also his (perception of his) KSs: a part of each KS of Figure 9 represents the whole scenario illustrated in Figure 9. Therefore, we have two kinds of information

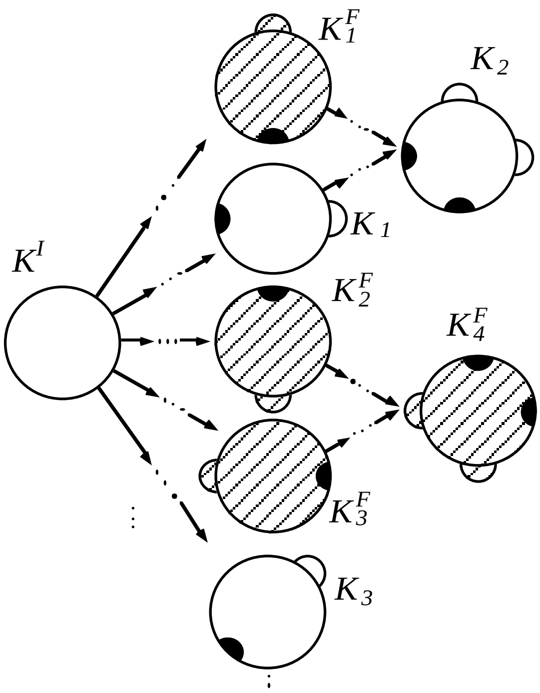

Figure 9: A more complete representation of an information need.

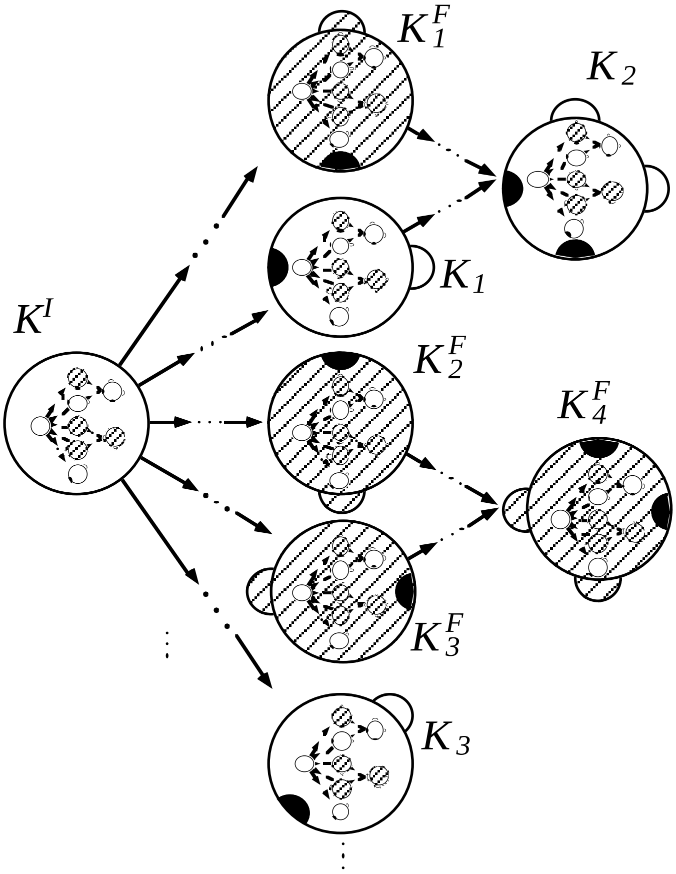

Figure 10: Representation of the information need inside the KSs.

need: the *Observer's Information Need* (ONeed), corresponding to Taylor's 'visceral need' (Taylor, 1968), about which I have discussed so far, and the *Agent's Information Need* (ANeed), corresponding to Taylor's 'conscious need', perceived by the agent.[3] The new situation is sketched in Figure 10, in which the representations inside the KSs and thus the ANeed are added.[4]

In order to formally define the ANeed, a *representation* function

$$repr: KS \times World \rightarrow KS$$

is needed. This function, given a KS and an object of the world, assumes as value the representation of the object in the KS: $repr_K(x)$ is the representation of the object x inside the KS K. Using this function, the ANeed in a KS K' as seen by an agent with a KS K can be defined as:

$$ANeed_K(K') = \{\langle repr_K(K^*) \setminus repr_K(K') , repr_K(K') \setminus repr_K(K^*)\rangle \mid K^* \in repr_K(\mathbf{K^*})\}.$$

Also $ANeed_K(K')$ is a set of pairs, and it is in general different from $ONeed(K')$. As the representation function is more correct and complete, $ANeed_K(K')$ becomes more similar to $ONeed(K')$. Anyway, correctness and completeness of the representation function $repr_K$ are sufficient but not necessary for having $ANeed_K(K') = ONeed(K')$.

The ONeed changes as time goes on: at time t_I, in KS K^I, the ONeed is the initial one, and

[3] I am implicitly assuming that the external observer is a sort of oracle, that has a correct image of the situation and can "see" inside the KS of an agent without affecting him.

[4] If inside the KS there is the representation of the KS, then there will be also the representation of the representation and so on. This leads to an infinite (if no fixed points are found) recursion, that should be handled in an opportune way.

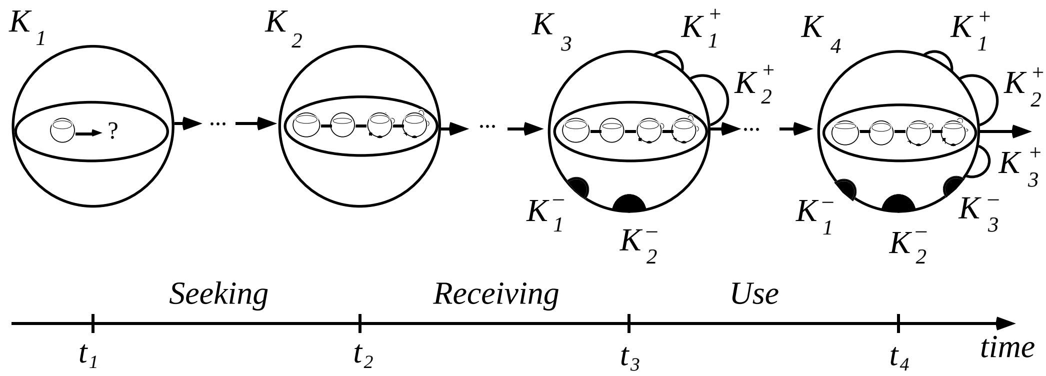

Figure 11: KSs in information seeking, retrieving, and use.

changes during the receiving of information. The ANeed is a representation of the ONeed, and thus it is time dependent too. Moreover, remember that a representation inside a KS is by no means correct and complete: an agent might need some information and not notice it, while another agent might not need information and believe that he needs it. Thus, it is likely that the ANeed is more adherent to the ONeed as the time goes on (more appropriately, as new data are being received), because the agent can perceive in a more correct and complete way his ONeed.

Note also that in order to define the ANeed, I have supposed that an agent represents his possible KSs (and the set ***K****). There are three very strong assumptions here: (i) the KSs (and ***K****) are objects of the world; (ii) an agent thinks in terms of KSs and transitions among them (while probably he does not); (iii) the agent represents all the possible KSs.

4.2 Information seeking, receiving, and use

Let us go more deeply into the details of the KS of an agent that tries to satisfy his information need using an IR system. His activity can be divided into three phases, graphically represented in Figure 11 (in which only the 'real' KSs, not the plausible ones, are reported):[5]

- *Information seeking* in which the agent tries to understand *how* (i.e. which are the steps to take) to satisfy his information need. At the end of this phase (K_2), the ANeed is very different from the initial one (in K_1), while the rest of the KS is practically unchanged. Let us suppose that in this phase the agent interacts with an IR system. Then, at the end of the phase, the user knows which documents to read and how he will try to satisfy his information need (obviously, only in an approximate way, as nobody can know the future with certainty).
  In this phase the user of the IR system is not interested in information, but rather in *metainformation*, i.e. in information about the information that he will obtain in the next phases. In the case of a bibliographic IR system, the metainformation is extracted from the surrogates of the documents.
- *Information receiving* in which the agent receives the data (in this case, documents) individuated in the previous phase, reads and studies the documents in order to reach the

[5] Note that the situation is not so linear, because not all the transitions are "good" ones: there might be some transitions that take the user away from the satisfaction of his information need (instead of bringing him near to it). Anyway, the user, before or later, perceives this fact, and accordingly modifies his behavior: there is a sort of *feedback*.

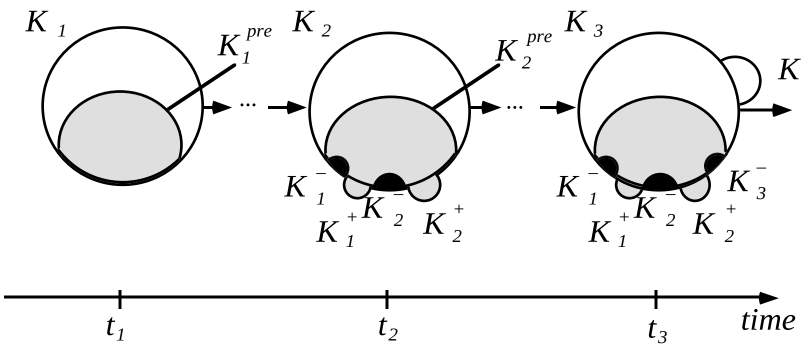

Figure 12: Modification of K^{pre} in a muddled need situation.

adequate KS K^F that permits him to start the following phase. In this phase, the KS of the agent is largely modified.

- *Information use* in which the agent uses the information received for acting (for instance, he writes something, or speaks about it to other people, or infers some other facts, or just do nothing). In this last phase, the KS of the agent is modified only by inferential transitions.

Usually, the IR researchers study only the first phase, guiltily neglecting the other two that should instead be taken into account. Summarizing, one should write:

IR = Information seeking + Information receiving + Information use.

4.3 Information need types

On the basis of empirical data, Ingwersen (1986; 1992) proposes three "fundamental types of information needs in IR": *Verificative need* (the user knows the bibliographic data of the needed documents. It is more a database problem than an IR problem); *Conscious topical need* (the user needs information about a topic that he knows well); and *Muddled topical need* (the user needs information about a topic that he does not know well. This is an ill-defined information need).

The information receiving phase of these three types of needs can be described on the basis of the above proposed cognitive scenario. The first two kinds of information need are nothing special: one or more transitions take place in order to modify the KS of the user.[6] In the more interesting case of a muddled need (Figure 12), some transitions take place, with the characteristics that the prerequisite KS (the gray area in figure) changes along the transitions chain: the user has to learn something (i.e. to receive some information) in order to be able to learn something else (i.e. to receive the information properly needed). Coming back to the Pitagora's Theorem example (Section 3.2), a muddled need would be one in which the user does not know the basics of Euclidean geometry (triangle, square, and so on): he has to learn those concepts before understanding the theorem.[7] Thus, in the case of a muddled need, the standard so called *Automatic Query Expansion* techniques (Magennis, 1995) (based on the assumption that the first expression of the user need is a correct one) seem not adequate, and *Interactive Query Expansion* ones (see for instance Brajnik et. al., 1995; Brajnik et. al., 1996) seem mandatory.

[6] Anyway, the conscious topical need may be interpreted as a muddled one by an observer (e.g. an intermediary), because a *label effect* may occur: the user do not externally express all the (well defined) KIs (Ingwersen, 1992).

[7] Note that the prerequisite KSs used in the chain will belong to the final KS satisfying the need.

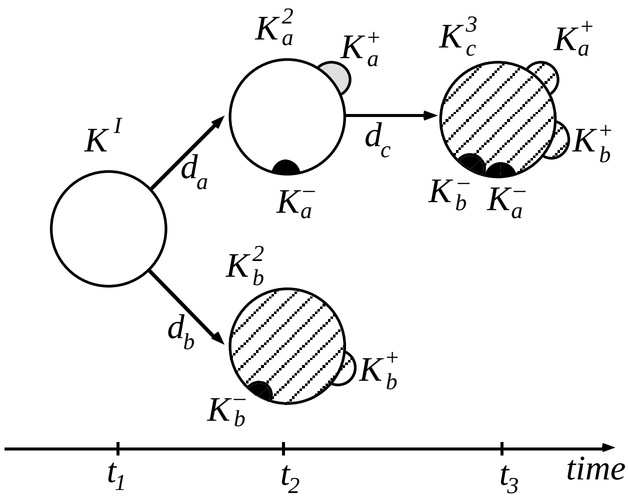

Figure 13: Relevance of information to information need.

5 Relevance

Relevance (Froehlich, 1994; Saracevic, 1975; Schamber et. al., 1990) is another concept, crucial in IR (and information science in general) and lacking a final definition, that can be defined and analyzed on the basis of the scenario presented in the previous sections.

The concet of relevance has already been used in Section 3.2, for emphasizing that a part of the initial KS of an agent receiving some data can be not relevant for the interpretation of the data. But this relevance is not the only one, and is not the one studied in IR. In IR, one speaks of information (or data, i.e. documents) relevant to an information need (or to one of its expressions, i.e. request or query; see Mizzaro, 1995b; Mizzaro, 1996b for a detailed discussion of this issue).

Figure 13 illustrates a network of KSs (the noninferential transitions are *not* split into perception and restructuring ones). From the initial KS K^I, two transitions can take place, through receiving one of two data (d_a or d_b). Let us consider by now only the two possible KSs at time t_2 (K_a^2 and K_b^2): if only K_b^2 is a final (i.e. satisfying the information need) KS, then the information $\langle K_b^+, K_b^- \rangle$ carried by d_b is relevant, while the information $\langle K_a^+, K_a^- \rangle$ carried by d_a is not relevant.

More generally and formally, given an information need

$$ONeed = \{\langle K_1^+, K_1^- \rangle, \langle K_2^+, K_2^- \rangle, \ldots\},$$

an item of information $\langle K^+, K^- \rangle$ is relevant to the information need if and only if there is an 'intersection' between the two:

$$relevant(\langle K^+, K^- \rangle, ONeed) \textbf{ iff } \exists i (K_i^+ \cap K^+ \neq \emptyset \vee K_i^- \cap K^- \neq \emptyset).$$

This means that an information item is relevant to an information need if the information item helps to reach a KS satisfying the need.[8]

The situation is anyway not so simple. A datum might seem not relevant to an ONeed at the end of a particular transition, and become relevant on the basis of a successive transition: relevance

8 Remember that (relevant) prerequisite KS have to belong to the final KS (see the Footnote 7 at the end of Section 4.3).

depends on the future. In order to illustrate this issue, let us consider the KS K_c^3 of Figure 13 and suppose that: (i) a successive transition from K_a^2, caused by another datum d_c, could take place; (ii) the KS K_a^+ of K_a^2 (obtained from d_a) is a part of the prerequisite subKS for this transition; and (iii) this transition leads to a KS K_c^3 that differs from K_a^2 because contains K_b^+ and does not contain K_b^- (and so is a sort of "union" of K_b^2 and K_a^2). Then, d_a is relevant: more specifically, d_a seems not relevant at time t_2 (after he has been perceived), while its relevance appears at time t_3 (an arbitrary time later). In a similar way, a datum might seem relevant before and become not relevant after. Note furthermore that the datum d_c carries the same information $\langle K_b^+, K_b^- \rangle$ of the (different) datum d_b, starting from different KSs.

Anyway, the above definition of relevance can be extended. The extension from information to data (i.e., in the IR case, documents or surrogates) is straightforward: a datum is relevant if and only if the information carried is relevant; the problem is that the information carried by a datum is not univocal, as illustrated in Section 3.1. Also the extension from ONeed to ANeed is simple. Being the ANeed a set of pairs analogously to the ONeed (see Section 4.1), the relevance of information to the ANeed is defined in a way similar to the relevance to the ONeed. Remember that the ONeed is different from the ANeed, so it is possible that an item of information is relevant to the ONeed and not relevant to the ANeed (or vice-versa). This difference is likely to decrease as the agent, receiving data, approaches to a KS in which his ONeed is satisfied.

6 Conclusions and future work

This paper proposes a cognitive framework for analyzing the information retrieval scenario in a formal way. In Section 2 some cognitive prerequisites are illustrated (KS, KI, link, K^+, and K^-). On this ground, three crucial concepts of information retrieval are analyzed: in Section 3 information is defined as a pair representing the difference between two KSs; in Section 4, information need is defined as a set of information items; in Section 5, relevance of information to information need is defined as a set intersection operation. Besides clarifying these three concepts, the adequacy of the above proposed cognitive framework is thus assessed.

This research is still at a preliminary stage, and there are many promising future developments. First of all, the cognitive scenario should be enriched in order include into the description:

- A more detailed analysis of what is inside the KSs. The dichotomies knowledge vs. metaknowledge (Genesereth & Nilsson, 1987), implicit vs. explicit knowledge (Nebel, 1990), actual vs. potential knowledge (if an agent knows the axioms of a theory, does he know all the theorems of the theory?), and other ones should be taken into account. Also the links should be analyzed more in depth, as they seem to play an important role in the transitions between KSs.
- A more dynamic vision of KSs. In this paper, I have preferred to define static KSs in order to avoid the problems related to logical omniscience (Genesereth & Nilsson, 1987). This is the reason for putting the inferences outside the KSs. The alternative way of including the inferences inside the KSs (and thus take into account the area of belief revision, see Alchóurron et. al., 1985; Gärdenfors, 1988; Katsuno & Mendelzon, 1991) should be considered.

- The intention of an agent (Devlin, 1991). The aims and goals play a crucial role in the interpretation of a datum, and they can change when a transition between KSs takes place.

When such an enriched cognitive scenario is available, the analysis of information need and relevance should be reconsidered, taking into account also the relevance to a request expressed by an agent: the relevance to a request is different from the relevance to ONeed and ANeed. Also other concepts of the IR field should be studied, as the *task* of the user of an IR system (i.e. what the user has to do with the retrieved documents) that should be included once introduced the aims and goals of the agent. Moreover, the topic-task-context triplet (Mizzaro, 1995b; Mizzaro, 1996b) should consequently be clarified.

Finally, it could be interesting to fully formalize this work, looking for an axiomatic theory.

Acknowledgements

Thanks to Iain Campbell, Paolo Giangrandi, Giuseppe O. Longo, Carlo Tasso, Keith van Rijsbergen, and an anonymous referee for many useful comments.

References

Alchóurron, C. E., Gärdenfors, P. & Makinson, D. (1985), On the logic of the theory of change: partial meet functions for contraction and revision, *Journal of Symbolic Logic* 50, 510-530.

Barwise, J. & Perry, J. (1983), *Situations and attitudes*, MIT Press, Cambridge, MA.

Bateson, G. (1972), *Steps to an Ecology of Mind*, Chandler Publishing Company.

Bateson, G. (1979), *Mind and Nature*, E. P. Dutton.

Belkin, N. J., Oddy, R. N. & Brooks, H. M. (1982a), ASK for information retrieval: Part I. Background and theory, *Journal of Documentation* 38(2), 61-71.

Belkin, N. J., Oddy, R. N. & Brooks, H. M. (1982b), ASK for information retrieval: Part II. Results of a design study, *Journal of Documentation* 38(3), 145-164.

Birnbaum, L. (1991), Rigor mortis: a response to Nilsson's "Logic and artificial intelligence", *Artificial Intelligence* 47, 57-77.

Brajnik, G., Mizzaro, S. & Tasso, C. (1995), Interfacce intelligenti a banche di dati bibliografici, in D. Saccà, ed., *Sistemi evoluti per basi di dati*, Franco Angeli, Milano, pp. 95-128.

Brajnik, G., Mizzaro, S. & Tasso, C. (1996), Evaluating user interfaces to information retrieval systems: A case study on user support. To appear, accepted for the publication in *ACM SIGIR'96*.

Brookes, B. C. (1980), The foundations of information science. Part I. Philosophical aspects, *Journal of Information Science* 2, 125-133.

Bruza, P. D. (1993), Stratified Information Disclosure: A Synthesis between Hypermedia and Information Retrieval, Ph.D. Thesis, Katholieke Universiteit Nijmegen, The Netherlands.

Devlin, K. (1991), *Logic and Information*, Cambridge University Press, Cambridge, England.

Dix, A., Finlay, J., Abowd, G. & Beale, R. (1993), *Human-computer interaction*, Prentice-Hall, New York.

Doyle, J. (1979), A truth maintenance system, *Artificial Intelligence* 12, 231-272.

Dretske, F. (1981), *Knowledge and the Flow of Information*, Bradford Books, MIT Press.

Froehlich, T. J. (1994), Relevance reconsidered – Towards an agenda for the 21st century:

Introduction to special topic issue on relevance research, *Journal of the American Society for Information Science* 45(3), 124-133.

Gärdenfors, P. (1988), *Knowledge in Flux – Modeling the Dynamics of Epistemic States*, MIT Press, London,

Gardner, H. (1987), *The mind's new science: A history of the cognitive revolution. With a new epilogue by the author: Cognitive science after 1984*, Basic Books, New York.

Genesereth, M. R. & Nilsson, N. J. (1987), *Logical Foundations of Artificial Intelligence*, Morgan Kaufmann Publishers, Los Altos, California.

Hughes, G. E. & Cresswell, M. J. (1968), *An Introduction to Modal Logic*, Methuen and Co. Ltd., London.

Ingwersen, P. (1986), Cognitive analysis and the role of the intermediary in information retrieval, in R. Davies, ed., *Intelligent Information Systems*, Horwood, Chichester, England, pp. 206-237.

Ingwersen, P. (1992), *Information Retrieval Interaction*, Taylor Graham, London.

Ingwersen, P. (1996), Cognitive perspectives of information retrieval interaction: Elements of a cognitive IR theory, *Journal of Documentation* 52(1), 3-50.

Katsuno, H. & Mendelzon, A. O. (1991), Propositional knowledge base revision and minimal change, *Artificial Intelligence* 52, 263-294.

Lalmas, M. & van Rijsbergen, C. J. (1996), Information calculus for information retrieval, *Journal of the American Society for Information Science* 47(5), 385-398.

Mackay, D. M. (1960), What makes the question, *The Listener* 63(5), 789-790.

Magennis, M. (1995), Expert rule-based query expansion, in *British Computer Society Information Retrieval Specialist Interest Group Colloquium*, Crewe, England.

Maturana, H. R. & Varela, F. J. (1992), *The tree of knowledge*, Shambala, Boston and London. ISBN 0-87773-642-1.

Mizzaro, S. (1994), TOBI - An Ontological Based Interpreter for Temporal Presuppositions and Counterfactuals, in R. Trappl, ed., *Cybernetics and Systems '94*, Vol. 2, World Scientific, Singapore, pp. 1879-1886. Proceedings of the Twelfth European Meeting on Cybernetics and System Research, organized by the Austrian Society for Cybernetics Studies, held at the University of Vienna, Austria, 5-8 April 1994.

Mizzaro, S. (1995a), La conoscenza in intelligenza artificiale, in *Atti del Congresso Annuale AICA'95*, Vol. II, pp. 1066-1073.

Mizzaro, S. (1995b), Le differenti relevance in information retrieval: una classificazione, in *Atti del Congresso Annuale AICA'95*, Vol. I, pp. 361-368.

Mizzaro, S. (1996a), Recursive models: A computational tool for the semantics of temporal presuppositions and counterfactuals in natural language, Research Report UDMI/11/96/RR, University of Udine, Department of Mathematics and Computer Science, Via delle Scienze, 206 – Loc. Rizzi – Udine, Italy.

Mizzaro, S. (1996b), Relevance: The whole (hi)story, Research Report UDMI/12/96/RR, University of Udine, Department of Mathematics and Computer Science, Via delle Scienze, 206 – Loc. Rizzi – Udine, Italy. Submitted to Journal of the American Society for Information Science.

Nebel, B. (1990), Reasoning and Revision in Hybrid Representation Systems, Vol. 422 of *Lecture Notes in Artificial Intelligence*, Springer-Verlag, Berlin.

Nilsson, N. J. (1991), Logic and artificial intelligence, *Artificial Intelligence* 47, 31-56.

O'Connor, J. (1968), Some questions concerning "information need", *American Documentation* 19(2), 200-203.

Salton, G. (1989), *Automatic Text Processing – The transformation, Analysis, and Retrieval of Information by Computer*, Addison-Wesley, Reading, MA.

Saracevic, T. (1975), Relevance: A review of and a framework for the thinking on the notion in information science, *Journal of the American Society for Information Science* 26(6), 321-343.

Schamber, L., Eisenberg, M. B. & Nilan, M. S. (1990), A re-examination of relevance: Toward a dynamic, situational definition, *Information Processing & Management* 26(6), 755-776.

Shannon, C. E. & Weaver, W. (1949), *The Mathematical Theory of Communication*, University of Illinois Press, Urbana.

Sowa, J. F., ed. (1991), *Principles of semantic networks*, Morgan Kaufmann, San Mateo, CA.

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

van Rijsbergen, C. J. (1979), *Information Retrieval*, 2nd edition, Butterworths, London.