1 TOWARD A BROADER LOGICAL MODEL FOR INFORMATION RETRIEVAL

Jian-Yun Nie*, François Lepage

Département d'informatique et de recherche opérationnelle,
Département de philosophie
Université de Montréal
C.P. 6128, succursale centre-ville
Montreal, Quebec, H3C 3J7 Canada

nie@iro.umontreal.ca, lepagef@ere.umontreal.ca

Abstract: The quality of an information retrieval (IR) model relies on its correct capturing of the notion of relevance. It is generally agreed that topicality is not the only relevance judgment criterion. More factors are concerned, in particular, the user's beliefs and his information state. However, no computational model capable of considering these factors has been defined until now. In this paper we suggest a new framework of logical modeling in order to make this possible. In this framework, all the factors other than topicality is encompassed in a retrieval situation. A retrieval situation is represented as a belief set. The relevance relation between a document and a query is represented as a counterfactual conditional whose evaluation is dependent on the given retrieval situation. This framework is a radical extension to the current computational IR models. This is a step forward to narrow the gap between computational approaches and cognitive approaches to IR.

1.1 INTRODUCTION

The ultimate goal of Information retrieval (IR) is to retrieve all and only the relevant documents for a user's information need. Consequently a good IR model is the one which gives each document a relevance estimation as close as possible to the user's own relevance judgment. The crucial problem in IR modeling is to correctly capture the notion of relevance within a computational model.

Since the beginning of IR, studies on relevance have been carried out from two different standpoints: computational and cognitive. In most computational models relevance is determined solely by document and query contents. If a document representation (about its contents) matches the query representation in some way, the document is considered to be relevant to some extent. This approach to IR and to relevance is commonly called the topical approach **?**, and relevance defined in this way is also referred to as system relevance **?**. A characteristic of this approach is that relevance judgments are static, context- and user-independent. That is, for any document-query pair, the judgment on their relationship is always the same.

Cognitive studies on relevance are usually carried out with a broader scope. Their goal is to find out as many factors as possible that affect the user's relevance judgments. According to these studies, relevance is no longer a topicality relationship between a document and a query in isolation. Rather it is their relationship within a certain context. The same document-query pair may receive different relevance evaluation in different contexts. We call such a context a *retrieval situation* (or *situation* for short) in this paper. A

*Partial funding provided by a research grant of the Natural Sciences and Engineering Research Council of Canada

retrieval situation is intended to encompass all the factors related to a particular user, other than topicality, that affect relevance judgments. The factors in a retrieval situation are called *situational factors*.

A cognitive IR model can be contrasted with a computational model as in Figure 1.1. Their difference precisely lies in the consideration of retrieval situation.

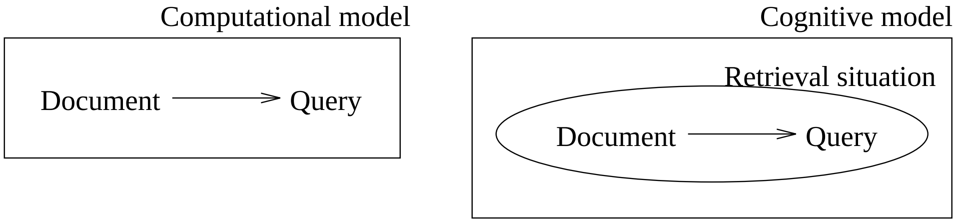


Figure 1.1 A view of computational models and cognitive models.

Although there are many computational models for topical estimation of relevance, no existing computational model is suitable to deal with situation-dependent relevance judgments. The difficulties of implementing such a broader computational model are certainly part of the reasons for this situation. However, we think that another reason lies in the incapacity for the theoretical tools (in particular, Boolean logic) used in traditional IR models to cope with situational factors. It is our belief that before building a computational model, one first has to identify an appropriate framework in which situational factors may be fully expressed and relevance judgments depend on them. This is the aim of the present paper: we intend to identify an extended logical framework which is capable of considering more situational factors than in current computational models. Two new elements are used in this framework: counterfactual conditional **?** is used to model the relevance relationship between document and query, and belief set **?**; **?** is used to represent a retrieval situation.

The paper is organized as follows. In section 1.2, we will first discuss the necessity to consider situational factors in IR, as well as the way they are represented in our model. In section 1.3, we will discuss logical modeling of the notion of relevance. Two particular logical connectives - material conditional and counterfactual conditional - will be contrasted. Counterfactual conditional will be suggested as the appropriate one. In section 1.7, we will deal with the modeling of retrieval situations. Finally some future work is pointed out.

1.2 THE NECESSITY TO CONSIDER SITUATIONAL FACTORS

How relevance is judged by a user? This is the central question for all the cognitive studies in IR. It is generally agreed that topicality is, although the most important, not the only factor that affects relevance judgments. There are a lot more. This fact has been recognized as early as the late 1950s **?**.

1.2.1 What situational factors to consider ?

In cognitive studies, although we still cannot answer precisely the questions such as what all the related factors are, and how they affect relevance judgments, there are common agreements on some most important factors. For example, in **?** Belkin identified the following important factors:

- the topic of the information the user is looking for (i.e. the topicality);
- the information already available;
- the user's state of knowledge or beliefs about the problem; and
- the goals and intentions of the user.

Belkin defines a state determined by all these factors a problem state. Similar factors have been identified in several other studies, and different names have been used to refer to them, for example: state of the user **?**, situationality of the user ?, the problem state ? and information need situation ?. Let us see how some of these factors may affect relevance judgments.

Topic. It is obvious that topic is the most important criterion for relevance judgment. In this paper, we assume that a query fully expresses the topic of an information need. A query is given a different status in our framework because of its prominent role in relevance judgment.

Knowledge state. The user's knowledge about the application area is very important. This is particularly true when the document under consideration does not directly concern the topic of the information need (query), but is related to it. Depending on the user's knowledge a connection may be or may not be established between the document and the query. For example, consider a query about "information retrieval" and a document about "document indexing". Depending on the user's knowledge, the document may be considered as relevant to the query or not. Suppose a user wants to find documents about "information retrieval". If he is an IR novice, and does not know that there is a connection between "document indexing" and "information retrieval", then a document talking about (and only about) "document indexing" will be seen as unrelated (thus irrelevant) to the query by the user. If later on, he becomes more familiar with IR, considers that "document indexing" is a necessary process for "IR", and wants more information about "IR", then the same document becomes relevant. We see that the user's knowledge has a great impact on his relevance judgments.

In order to make user-oriented relevance judgments, it is important for the IR system to have a knowledge store as similar as possible to that of the user. In the above example, to an IR expert, it is clear that there is a strong connection between "document indexing" and "information retrieval". However, before that the user (the first user) believes in (or consider as true) the connection, it is useless to provide him with the document. The point is that the system should not make inference beyond what the user can accept. In order to do this, the knowledge stored in a system should reflect the user's mind at the moment document retrieval even if it is incomplete and sometimes wrong from an expert's point of view. Only in this case can we talk about user-oriented relevance estimation.

As the user's knowledge is incomplete and even wrong, it is more exact to talk about "beliefs" rather than "knowledge". However, as "knowledge" is the term commonly used in cognitive studies in IR, we will use both terms interchangeably in our further discussions.

Information state. This aspect has been considered in the investigations of usefulness ?. As stated in ? usefulness is one of the multiple facets of relevance. In many cases users are only interested in new information. A piece of information already known, although strongly concerning the topic of the user's information need, may still be judged irrelevant because of its lack of novelty. This factor is particularly important in the case where there is a great deal of potentially relevant information, and the user is not interested by the known information. By keeping track of the user's information state, we can avoid flooding the user with the known information.

Moreover, the relevance of a piece of information may depend on the presence of another information in the information state. For example, suppose a user looks for documents providing information on "information retrieval" **and** information on "Internet". If in the information state there is already information about "information retrieval", then a document talking about only "Internet" may be considered as relevant (together with the information already known). On the other hand, if no information about "information retrieval" or "Internet" is already known, a document should contain information about both concepts to be considered as (completely) relevant. In order to deal with this dependence, an explicit representation of the known information is necessary.

Other situational factors. Many other factors have been objects of cognitive studies. For example, the goal for which a user searches documents has a great impact on relevance judgments. The goal of the user may be to learn about a particular topic, or only to find some references to put in his article. Relevance judgments in these two situations may be very different. Relevance judgments can also be influenced by the means used in the judgment process. For example, it turns out that judgments may depend on whether appropriate computer supports are available.

For these situational factors, there does not seem to have a means to model them correctly at the current stage. So our further study in this paper will only consider knowledge state and information state. They together form a *retrieval situation*. So, a model of retrieval situation (or a situation model) S in our paper is defined by a model of knowledge state and a model of information state as follows:

$$S =< modelofknowledge\ state, modelofinformation\ state > .$$

An important property of retrieval situation is its dynamism. Recall that an IR process is embedded in a larger learning process. The goal for a user to retrieve documents is to acquire information from them, i.e. to learn new things. Once new information is learned, the user's knowledge state and information state change. Their counterparts in the situation model have to change accordingly. So a suitable modeling framework should be able to cope with changes in retrieval situations.

In the remaining of this paper, most of our discussion will be concerning models of retrieval situation rather than the real retrieval situation. So the word "model" in "model of retrieval situation" will be omitted when this does not arise ambiguity.

1.2.2 How to represent a retrieval situation, a document and a query?

There are several possible ways to represent a (model of) retrieval situation. One of them is to represent it as a set of sentences or as a deductive closure of such a set of sentences. These two representations are called *belief base* and *belief set* respectively **?**. For the present study the difference between the two notions is not important. So we will use belief set. The reason for us to take this representation form is mainly due to its simplicity and the existence of several operations for its changes. These operations make it easier for us to cope with changes in retrieval situations. However, different from counterfactual conditional, belief set is not essential in our approach. One may well replace it by another suitable form of belief representation.

Knowledge state. What kind of sentences do we put in a knowledge state? The answer is not obvious. In fact, there are two different levels of knowledge: conceptual level and informational level. The former is what one usually considers in AI studies: a concept is related to another one in a certain manner. For example, one may consider that any "IR system" is a kind of "information system". This relationship between the two concepts may be expressed as the following logical implication: $(IR\_system \rightarrow information\_system)$ where IR_system and $information\_system$ represent respectively the two concepts.

The other kind of knowledge expresses relationships between pieces of information. For instance, one may consider that a piece of information about a concept is related to another concept. In other words, a piece of information about the first concept is also a piece of information about the second concept to certain extent. This kind of knowledge is particularly important in IR. In fact, if any inference is be made in IR when judging relevance, it should be made at informational level. This is because the goal of IR is to answer the following question: given some information, what is its relation to the information one is looking for? The core of the question lies at the information level. For example, in order to infer that a document about "IR system" is relevant to a query on "information system", one has to possess a knowledge indicating that a piece of information on "IR system" is also a piece of information on "information system". Let us define the presence of a piece of information about the concept A as $I(A)$, then the above knowledge may be expressed as $I(IR\_system) \rightarrow I(information\_system)$.

The two kinds of knowledge are strongly related. On one hand, if two concepts are related, a piece of information about one concept is often related to another concept. On the other hand, if two pieces of information are related, the concept they describe are necessarily related. Despite this strong relationship, they are different. For example, one may accept that there is a strong relationship between two concepts (e.g. "cement" and "building"), but reject any connection between pieces of information about them (i.e. it is quite reasonable to consider an information about "cement" as being not related to "building", or vice versa). The reverse relationship seems more difficult to reject. One may hardly imagine that two pieces of information are considered as being related while the concepts they describe are not.

Knowledge bases used in AI and IR usually contain conceptual knowledge. In IR, thesauri are often used as reasonable knowledge bases. Relationships in thesauri are set at conceptual level: one concept is a super- (or sub-) concept of another. Experiments showed that using thesauri in IR leads to variable results. The problem may lie in the difference between conceptual knowledge and informational knowledge. One is often tempted to do a direct mapping from conceptual knowledge to informational knowledge as follows: if there is a relationship between two concepts, then there is a relationship between pieces of information about the concepts. Unfortunately, this direct transformation often fails. The reason is that informational knowledge is much more subjective than conceptual knowledge. From the same conceptual knowledge, different users may create different informational knowledge. In our previous studies **?**; **?**, we tried to make a transformation according to the user's relevance feedback, i.e. relevance feedback is used as a means to capture the user's subjectivity. Each conceptual relationship (from a thesaurus) between A and B is mapped into an uncertain information knowledge between $I(A)$ and $I(B)$. The degree of certainty of the informational knowledge is revised gradually according to relevance feedback. It is shown that the obtained informational knowledge (which is quite different from the conceptual knowledge) has a great positive impact on IR effectiveness.

In this paper we do not intend to deal with the problem of mapping from conceptual knowledge to informational knowledge. We only consider informational knowledge in our retrieval situation. In addition, we do not deal with the uncertainty problem in this paper. So our model of a knowledge state is a set of (logical) relationships between pieces of information. For example, $I(A) \rightarrow I(B)$ means: a piece of information about A is also a piece of information about B.

Information state, document and query. An information state, as well as a document, is considered as a set of information. Thus, they may be represented as a set of $I(A)$-form sentences. In traditional IR approaches, the information anchors A, i.e. the items described by pieces of information, are assumed to be simple concepts. It is known that this representation is not sufficient to describe the contents of documents and queries. For example, a document may contain not only information about "computer" and about "communication", but also about their relationship. For instance, "computer" *used in* "communication", or "communication" *between* "computers". Traditional representations only allow to express the first part. Several studies **?**; **?** have been carried out in order to create more refined representations. They have proven great benefits of such refined representations in precision-oriented IR. In this study we do not set any restriction on A. It may express a concept as well as a relationship between concepts. However, we will not deal with these different information anchors explicitly.

At this stage, it is worthwhile to emphasize the importance to distinguish concept from information. They were often not well distinguished in previous studies on logical models. In previous models, a document, as well as a query, is often expressed ambiguously as a logical expression of "atoms" as follows (where x, y, ... are atoms):

$$d = x \wedge y \wedge ...$$

$$q = x \wedge (y \vee z).$$

Atoms in such expressions have had two different interpretations: an atom x stands for a concept, or for a piece of information about a concept. Under the first interpretation, the above document is understood as providing information about a set of concept, i.e. $I(x \wedge y \wedge ...)$ in our expression. Under the second interpretation, the document representation means a set of information about individual concepts, i.e. $I(a) \wedge I(b) \wedge ...$ where a, b, ... are respectively the concepts described by x, y, ... The confusion between the two interpretations may be best illustrated with the negation operator $\neg$ (not). It was usually assumed that if an atom x is absent in a document representation, then $\neg x$ is in that representation (the closed-world assumption). Under the first interpretation $\neg x$ would mean $I(\neg x)$ in our notation, i.e. the presence of an information about $\neg x$. So the assumption says: if there is no information about a concept, then there is information about the negation of that concept. In our notation, it corresponds to: $\neg I(x) \vdash I(\neg x)$. This assumption is incorrect. Indeed, when there is no information about "computer" in a document, it is not meant that there necessarily is information about "$\neg$*computer*". "$\neg$*computer*" means any other concepts than "computer", for example, "mechanism", "medicine", and so on. Using the assumption under the first interpretation of atoms is dangerous.

The assumption can only be made under the second interpretation. It becomes trivially true: if there is not an information, there is not this information, i.e. $\neg I(a) \vdash \neg I(a)$.

Our distinction between concepts and pieces of information about concepts solves the confusion. Documents and queries are expressed at the information level. The unreasonable assumption is thus avoided.

1.2.3 How to estimate relevance ?

The classical understanding of relevance evaluation may be expressed as follows: If a document is related to a query, then it is considered as relevant (by the system). This idea has been implemented as a term matching between document and query representations, or as an implication relationship between the document and the query in classical logic. In these implementations, there is no consideration of retrieval situations.

In **?**; **?**, van Rijsbergen proposed a new logical framework for relevance estimation. The main idea may be expressed as the following: The relevance relation between a document d and a query q may be expressed as a non-classical logical implication $d \Rightarrow q$, and its evaluation is governed by the following *uncertainty principle*:

Uncertainty principle:

Given any two sentences α and β; a measure of the uncertainty of $\beta \Rightarrow \alpha$ relative to a given data set, is determined by the minimal extent to which we have to add information to the data set, to establish the truth of $\beta \Rightarrow \alpha$.

This is the first time that relevance estimation becomes dependent of a context (the data set in the uncertainty principle). We also note that relevance estimation becomes dependent on the process of minimal extension of the data set. According to how this process of extension is carried out, the relevance measure differs. However, the uncertainty principle does not specify how the minimal extension should be made. This leaves the meaning of the implication $\Rightarrow$ insufficiently defined. As we will point out in Section 1.3.1, this implication may be well interpreted as the classical implication, although not intended by van Rijsbergen. This makes the principle not precise enough to guide the construction of a computational model for IR. Nevertheless, van Rijsbergen's proposal marked a clear departure from the traditional IR approaches, and opened a new research direction in IR.

In this paper, we also consider the relevance relation between the document d and the query q as a logical implication $d \Rightarrow q$. However, we would like to give a more precise meaning to the implication $\Rightarrow$. To start, let us borrow some idea on relevance evaluation from cognitive studies of communication.

Information retrieval may be seen as a communication process between a system and a user. This point of view has already been taken in the very early studies in IR **?**. Let us consider the following definition of relevance in communication (**?**, pp.122):

> An assumption is relevant in a context if and only if it has some contextual effect in that context.

This definition may be mapped to IR as follows: An "assumption" in IR is that a document is retrieved and the information contained in it becomes known to the user. The "context" in our case is exactly the retrieval situation. The "contextual" effect in IR is that the user's information need is satisfied; or if we take uncertainty into account, the information need is better satisfied. The information need is better satisfied if there is more information relevant to it. Similarly to this definition, relevance in IR may be expressed as the following relevance principle:

Relevance principle:

A document is relevant to a query in a given retrieval situation if and only if it contributes to a better satisfaction of the query than in the given retrieval situation.

Let us analyze this principle in the following two cases:

1. the query is not satisfied in the given retrieval situation.

   This is the usual case in IR: the user has an unsatisfied information need, and more relevant documents are sought. The question underlying relevance estimation is the following: if the user acquires the information contained in the document, would his information need be satisfied?

   In order to answer this question, we have to *imagine* a new retrieval situation in which the information contained in the document is supposed to be known to the user. That is, the information should be added into the information state. This last operation leads the retrieval situation to a new one (as more information is added) and some old items in it may be retracted. As it is only imagined, we call it a *hypothetical situation*. There may be many ways to add new information into a situation. It is required that the hypothetical retrieval situation should only change the given retrieval situation *minimally*. This requirement confines our consideration only to the most likely hypothetical situation. If the user's information need is satisfied in the hypothetical situation, then the document is said to be relevant. Otherwise, it is irrelevant.

2. the query is already satisfied in the given retrieval situation.

   This is the case where no more document is needed because the user has the necessary information in the given retrieval situation. No document is to be judged relevant (useful).

Putting the above relevance analysis into symbols and omitting the uncertainty problem, the relevance principle may be expressed as follows:

> A document d is relevant to a query q in a retrieval situation S iff $S \not\vdash q$, but $S' \vdash q$.
> where S' is the hypothetical situation in which the document is supposed to have been obtained.

In comparison with the uncertainty principle, the minimal change is not done in order to satisfy the entire conditional $d \Rightarrow q$, but only the document which is the antecedent of the conditional. This difference gives a more precise meaning (counterfactual conditional) to the conditional $d \Rightarrow q$ as we will show in Section 1.4.

1.3 EXISTING APPROACHES TO RELEVANCE MODELING

In this section two representative aproaches will be reviewed: Boolean logic model and Bayesian probabilistic model.

1.3.1 Boolean model

In this model a document is represented as a conjunction d of pieces of information such as $d = I(a) \wedge I(b) \wedge ....$ A query is represented as a Boolean expression q of information criteria. A Boolean retrieval system is modeled as a set S of sentences in classical logic. These sentences are intended to represent the domain knowledge of the given application. We assume here that the domain knowledge is expressed at the information level.

Given such d, q and S, the estimation of the document's relevance is based on the following deduction of material conditional :

$$S \vdash d \rightarrow q$$

or equivalently:

$$S \cup \{d\} \vdash q.$$

In the simplest application cases, no domain knowledge is used. Relevance estimation is solely based on a direct matching between document and query. In this case S is assumed to be empty, i.e. $S = \emptyset$. In more "intelligent" systems, S contains some domain knowledge such as $I(a) \rightarrow I(b)$. In this case, a document may still be considered relevant if it indirectly satisfies the query through logical inference. With respect to the relevance principle described earlier, the hypothetical situation S' is $S \cup \{d\}$.

So far, Boolean model seems to work well. However, problems may arise if we intend to model situation-dependent relevance estimation, i.e. to use S to model a retrieval situation. Observe that the condition for the Boolean model to work correctly is that S and d be consistent, i.e. $S \cup \{d\} \not\vdash \perp$. If they are inconsistent, we will have $S \vdash d \rightarrow q$ for any query q. Its wrong interpretation in IR is that the document d is relevant to any query. This inconsistent case is abnormal in previous uses of Boolean model, and it is assumed to never occur. However, within our broadened context, they may occur. Let us now illustrate this as follows.

d may be inconsistent with the information state. Suppose in the user's current information state, no information about "IR" is known. So $\neg I(IR) \in S$. Now if our database contains a document talking about IR: $d = I(IR)$, it is clear that S and this d are inconsistent. This inconsistence case is not abnormal in our broadened context where we keep track of the user's known information.

One may suggest that the problem may be solved by separating the explicit representation of information state from the logical model. In such a way, the model is only responsible for judging if a document is topically related to the query. Then the selected documents are compared with the store of known information to filter out the known information. This solution is, however, partial: We cannot consider the case in which the relevance of a document depends on the presence of some other information, as we mentioned in section 1.2.1. In order to cope with the problem properly we have to integrate an explicit representation of information state in the model.

d may be inconsistent with the knowledge state. Recall that the user's knowledge (belief) is partial and even wrong. This means that a conclusion he makes may contradict the contents of some documents (let us assume that these latter are correct). For example, if a user observes through his prior experience that whenever he obtains information about "word stemming", he also obtains information about "IR". From this observation, he could conclude that $I(word\_stemming) \rightarrow I(IR)$, and put it into his knowledge stock. This piece of knowledge is not always valid because "word stemming" is not only used in "IR" but also in other domains such as "linguistics". Given a document describing "word stemming in linguistics", its content may be represented as $d = I(word\_stemming) \wedge I(linguistics) \wedge \neg I(IR)$. This document is inconsistent with the user's belief. Then we fall again into the inconsistent case.

The inconsistent cases just described correspond to the following deduction in classical logic (where α, β, ... denote pieces of information):

$$\neg\alpha \vdash \alpha \rightarrow \beta.$$

i.e. any material conditional is valid in a context which contradicts the antecedent of the conditional. Studies in AI and philosophical logic have found a number of fallacies for material conditional $\rightarrow$ when it is used to represent "if-then" sentences in reasoning. For IR, relevance relationship may be seen as a special case of "if-then" sentence: if d then q. The same fallacies equally apply to IR. As an example, let us consider the following fallacy.

$$\neg(\alpha \rightarrow \beta) \vdash \alpha.$$

If $\rightarrow$ is taken as a model of relevance, the above deduction may be interpreted as follows: If an information α is irrelevant to a query β, then the information is already know. Using this property in practice, it means: whenever a user judges a document "irrelevant", the system may deduce that the document only contains known information, and then put the contained information into the information state. It is not difficult to imagine the possible disaster consequences of this conclusion in later reasoning.

More elaborated discussions about fallacies of material conditional may be found in AI and logic literature (for example **?**). Most of them also apply to IR. They show the inappropriateness to represent relevance by material conditional.

Material conditional has also been rejected by van Rijsbergen in **?**; **?** for other reasons. However, it is not incompatible with the uncertainty principle. Indeed, one can interpret the symbol $\Rightarrow$ in the uncertainty principle as material conditional, and the uncertainty principle becomes the following:

Given a query q, and suppose the information contained in a document d is not already in the current retrieval situation (data set) S. So $\neg d$ is in S. Hence $d \rightarrow q$ is true in S. Then there is no need to add information in the retrieval situation. The conditional $d \rightarrow q$ is completely certain. However, from this, one cannot consider d as relevant to q in this case. This example shows that the connective $\Rightarrow$ in the uncertainty principle lacks a precise meaning. In the following section, we suggest to model relevance relation as a counterfactual conditional which has a well-defined semantics.

1.3.2 Bayesian probabilistic model

In probabilistic models a document may still be represented by a set of (possibly weighted) terms. A document representation d corresponds to the statement that information about some concepts is available, or equivalently the event that the corresponding document is retrieved. A query description q stands for the information need about some concepts. We can also use q to denote the set of the relevant documents to this query.

Given such d and q, most probabilistic IR models aim to estimate relevance according to the following conditional probability [21] : $P(q|d)$, i.e. the probability that the query is satisfied by relevant information given that the information contained in the document is provided to the user.

This probability is not directly calculable. Bayes' rule is then used so that:

$$P(q|d) = P(d|q)P(q)/P(d)$$

in which $P(d|q)$ denotes the probability that the considered document is among the relevant documents to the query. In the Binary-independent model, this conditional probability is estimated as follows:

$$P(d|q) = \prod_{i=1}^{n} P(xi|q) \qquad t_i \in d.$$

in which x_i denotes the presence ($x_i = 1$) or absence ($x_i = 0$) of some term ti. The conditional probability $P(x_i|q)$ may be estimated according to the distribution of the term ti among the relevant and irrelevant document samples. In practice, the following measure the odd of relevance is used: $O(d, q) = P(q|d)P(\neg q|d)$

The fact remains that the model is based on the conditional probability $P(q|d)$.

The above model has often been criticized for its assumption of term-independence. Indeed, under the assumption, human knowledge in the domain of application can hardly be incorporated. Knowledge such as "programming is part of computer science" does express a kind of dependence between two terms. To solve this problem, a more recent probabilistic model [30] makes use of a Bayesian inference network [20] in which elements (e.g. terms) can be connected with each other by dependence relations. A Bayesian network in this approach is aimed to encode the domain knowledge as well as the composition of documents and queries. Relevance estimate is based on the calculation of $P(q|d)$ using probability spreading [20] .

There are also a number of other probabilistic models proposed for IR [5, 21, 24, 31, 32] . All these models are based on the conditional probability $P(q|d)$.

These models may face with the same problem as the Boolean model when they are used to consider situational factors. For example, if the availability of information is to be considered by one of the above models, then all the parameters should be set in accordance with this aspect. In particular, if an information A is already available, then the model should give $P(A) = 1$. On the other hand, if it is unavailable, it should gives $P(A) = 0$. Now, given a situation in which the information d and q are both unavailable and we want to evaluate the relevance of d to q. We need to evaluate the conditional probability P(q—d) in that model, but this probability is undetermined because both $P(q \wedge d)$ and $P(d)$ are equal to 0.

n a similar way, the conditional probability $P(q|d)$ also has difficulty to consider the user's prior knowledge and other situational factors which may contradict the document content under evaluation.

1.4 MODELING RELEVANCE AS COUNTERFACTUAL CONDITIONAL

Counterfactual conditional (denoted by the symbol >) has often been suggested as an appropriate replacement of material conditional in AI **?** and in philosophical logic?; **?** for reasoning. It solves most fallacies of material conditional. Let us first describe the principle of evaluation of counterfactual conditionals.

A counterfactual conditional $\alpha > \beta$, read as "if it were the case that α, then it would be the case that β", is evaluated according to the following principle known as the Ramsey test:

> To evaluate a conditional, first change the current situation minimally in order to accept the antecedent; then evaluate the consequent in the changed situation.

The interesting case is when the antecedent is false in the current situation. In this case a counterfactual conditional is not automatically assigned "true" as is a material conditional. Rather, a hypothetical situation is created which should meet the following two conditions:

- the antecedent is true in the hypothetical situation; and
- the hypothetical situation is minimally changed from the current situation.

The hypothetical situation is a simulation of what would happen if the antecedent were accepted. The second condition is to confine one's consideration to the most likely case(s) where the antecedent becomes true. If the consequent is true in the hypothetical situation, we say that the counterfactual conditional is true in the current situation.

Let us show two examples to illustrate the principle.

1. Suppose I'm driving a car, and say "if I braked, my car would stop".

   The current situation is that my car is running. So the antecedent of the conditional - I brake - is false. In this case, we have to consider the least changed hypothetical situation in which "I brake" becomes true. This situation is the one in which "I brake" and every other thing remains unchanged, including the fact that my car is in normal running conditions. In such a situation, my car would stop. So I can say that the conditional is true.

2. If Paris were in Spain, pigs could fly.

   We can imagine all kinds of situation in which the antecedent becomes true: the Spain conquers Paris, Paris is moved to Spain by some magic force, France concedes Paris to Spain, and so on. However, if we follow the rule of minimal change, all we changed from the current situation only concern the position of Paris. In none of the solutions is it necessary to change the incapacity for pigs of flying. So in all the minimally changed situations, pigs still cannot fly. We can thus conclude that the conditional is false.

In the two cases we just illustrated both the antecedent and the consequent are false in the current situation. If the consequent becomes true once we accept the antecedent, we can say that accepting the antecedent "forces" the consequent to be true. This case is often referred to as representing a "causal" relation between the antecedent and the consequent. There is still another case in which a counterfactual conditional is true: if the consequent is true in the current situation and remains true in the hypothetical situation. The conclusion drawn in the first case is stronger than that in the second case. If we are interested in causal analysis, it is possible for us to rule out the second case **?**. Nevertheless, in both cases counterfactual conditional expresses a stronger relationship than material conditional, i.e. $\vdash (\alpha > \beta) \rightarrow (\alpha \rightarrow \beta)$. Indeed, the truth of material conditional is solely determined by the truth values of the antecedent and consequent in the current situation. No inspection of the hypothetical situation is required. A "true" value for the material conditional may be established even if the antecedent has nothing to do with the consequent, for instance, in the second example above. In particular, when the antecedent is false or when the consequent is true, the material conditional is automatically true. This is the source of many fallacies of material conditional in reasoning.

The evaluation principle of counterfactual conditional may be best illustrated in Lewis' system **?**. Lewis' system is built using spheres of worlds. A world is a maximally consistent set of sentences, in which everything is determined. A system of spheres is built as follows: from a given world w, some other worlds are accessible but their distance to the given world may vary. According to this distance, a set of nested spheres around the given world may be constructed so as to create concentric spherical regions, as shown in Figure 1.2. Worlds (e.g. w_1 and w_2) having the same distance from the given world w are in the same spherical region around w. The larger the spherical region, the more distant the worlds in it from the given world.

A world is called an "A-world" if the sentence A is true in it. The A-worlds included in the minimal sphere where there is at least one A-world are called the "nearest A-worlds" to w. In any world w, the conditional $(A > B)$ is true if all the nearest A-worlds to w are also B-worlds. Figure 1.3 shows the case of $A > B$ being true and false in w (where the regions A and B represent respectively regions of A-worlds and B-worlds).

In the first case, there are three nearest A-worlds to w. These worlds correspond to what would most likely happen if A were accepted. In all of the three worlds, the consequent B is true. So, B would necessarily happen if A were accepted, i.e. $A > B$ is true in w. In the second case, the consequent B is false in one of the most likely worlds. So $A > B$ is false in w.

This evaluation is an implementation of the Ramsey test: The nearest A-worlds form the hypothetical situation to w. If all these A-worlds are also B-worlds, then the consequent B is satisfied in the hypothetical situation.

The system of Lewis is also called SOS-model: System-Of-Spheres). One can add the following constraint on this model: To each world anf fior any sentence A, there is exactly one nearst A-world. With this constraint, we obtain the system of Stalnaker (or world selection model - WS-model).

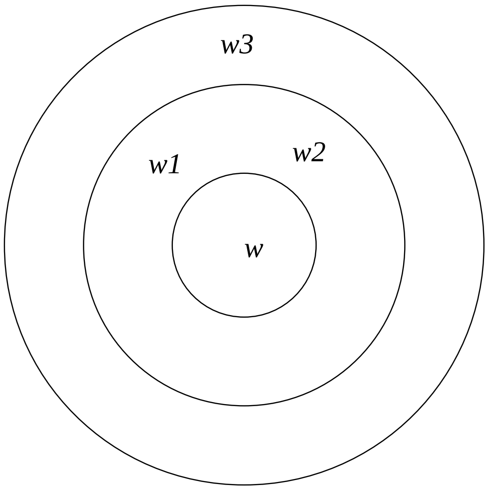
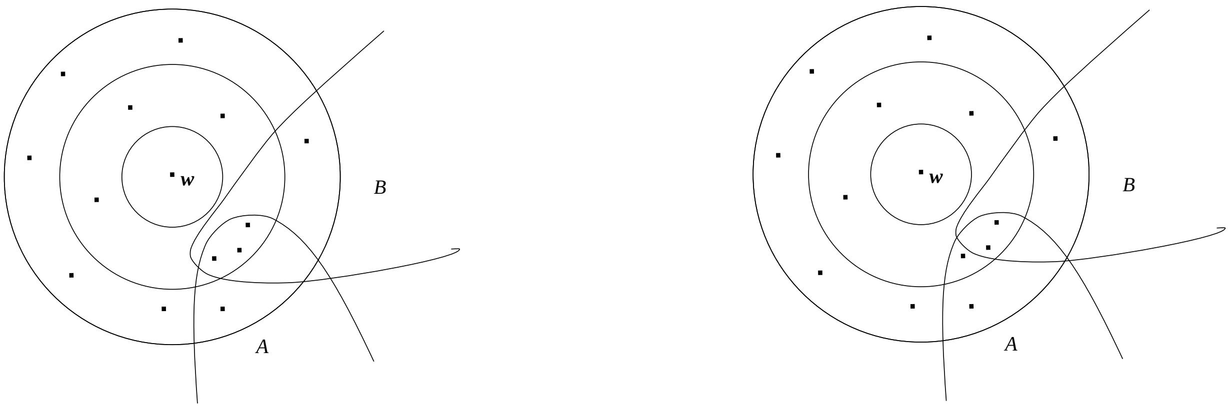


Figure 1.2 An example of spheres of worlds



$A > B$ $\neg(A > B)$

Figure 1.3 Examples of evaluation of counterfactual conditionals.

Formally, these systems are defined as follows [19] .

Definition 1: WS-model A world-selection-function model (WS-model) is a model $< W, < \alpha, R, s >>$ such that:

1. α is the constant $\top$-function on the set of wffs;
2. R is a reflexive binary relation on $W \cup \alpha$
3. for all $w \in W$, $< w, \alpha > \notin R$ and $< \alpha, w > \notin R$
4. s assigns to each wff A and each $w \in W$ a member $s(A, w)$ of $W \cup \alpha$
5. if $s(A, w) \neq \alpha$, then $< w, s(A, w) > \in R$
6. if $s(A, w) = \alpha$, then for all $w' \in W$, either $w'(A) = 0$ or $< w, w' > \notin R$
7. if $w(A) = 1$ then $s(A, w) = w$
8. if $s(A, w)(B) = 1$ and $s(B, w)(A) = 1$, then $s(A, w) = s(B, w)$;
9. $W/(A > B) = w \in W : s(A, w)(B) = 1$

where $W/(A > B)$ selects the subset of the worlds in which $A > B$ is true.

Definition 2: SOS-model A system-of-spheres model (SOS-model) is a model $< W, \$ >$ such that:

1. \$ assigns to each $w \in W$ a nested set $\$w$ of subsets of W which is closed under unions and finite intersections;

2 $W/A > B = \{w \in W : \cup \$w \cap W/A = \emptyset$ or there exists an $S \in \$$ such that $\emptyset \neq S \cap W/A \subseteq W/B\}$.

The following axiomatic system (C2) is sound and complete for all the WS- models [28] :
- Axioms and rules of normal modal logic;
- Inference rules:

RCEC from $A \equiv B$, to infer $(C > A) \equiv (C > B)$
RCK from $(A_1 \wedge ... \wedge A_n) \rightarrow B$, to infer $[(C > A_1) \wedge ... \wedge (C > A_n)] \rightarrow (C > B)$, $n \geq 0$
ID $A > A$
MOD $(\neg A > A) \rightarrow (B > A)$ (or $A \rightarrow (B \rightarrow A)$)
CSO $[(A > B) \wedge (B > A)] \rightarrow [(A > C) \equiv (B > C)]$
CV $[(A > B) \wedge \neg(A > \neg C)] \rightarrow [(A \wedge C) > B]$
MP $(A > B) \rightarrow (A \rightarrow B)$
CEM $(A > B) \vee (A > \neg B)$

CEM expresses exactly the uniqueness of the nearest world.

The following axiomatic system (VC) is sound and complete [17] for all the SOS-models such that for each $w \in W$, $w \in \$w$ (which we will call Lewis' system) :

- The same rules and axioms as in C2 system except that CEM is replaced by the following CS:

CS $(A \wedge B) \rightarrow (A > B)$

From the above description one may notice a strong similarity between counterfactual conditional, on one hand, and the relevance principle and relevance evaluation in communication, on the other. This shows that counterfactual conditional captures the very principle of relevance evaluation.

Another way to justify the utilization of counterfactual conditional in IR is to examine its formal properties, and contrast them with the observations in cognitive studies of IR. This has been done in **?**. It turned out that the basic properties of counterfactual conditionals can be naturally adopted in relevance judgments. In other words, interpreting $>$ as relevance relation is reasonable for IR. As an example, let us show how the following property, which is rejected by all counterfactual conditionals, can also be reasonably rejected for IR (note that this property is satisfied by material conditional):

$$\text{(Weakening antecedent)} \qquad \vdash (\alpha > \beta) \rightarrow ((\alpha \wedge \gamma) > \beta).$$

To reject the property it suffices to build a counterexample. A strong counterexample is obtained when $\gamma \rightarrow \neg\beta$, i.e. the fact that γ (the presence of some information) provides counter-evidence for β. In this case we may have $\alpha > \beta$ in IR in some situation, but not $(\alpha \wedge \gamma) > \beta$. Thus the weakening antecedent property is not valid for IR.

A more concrete example is: $\alpha = I(skiing)$, $\beta = I(winter\_sport)$, and $\gamma = I(water)$. In general, a document providing information about "skiing" is relevant to the information need about "winter sport". But if the document also describes "water" (or better, that the kind of "skiing" is "water skiing"), then the user may know that the document is not at all related to "winter sport". The additional information gives a counter-evidence to the document's relevance to the query. In this case one may judge the document containing α and γ "irrelevant" to β. This is a case in which weakening antecedent is not verified. Thus the property may be rejected.

To conclude the section, our analysis suggests the following modeling approach:

Given a document expressed as d, a query as q in a certain retrieval situation S, the following expression denotes the main part of the relevance of d to q in S:

$$S \vdash d > q.$$

If we do not consider the uncertainty in relevance judgments, relevance may be evaluated according to the following principle:

A document d is relevant to a query q in situation S iff $S \not\vdash q$, but $S \vdash d > q$.

In the iff-part, we compare the evaluation of $d > q$ in S with respect to q. This comparison aims to avoid retrieving unnecessary documents when the query is already satisfied. The idea is very similar to causal analysis, i.e. one rules out the case where the consequent (q) is already true and remains true in the hypothetical situation. The remaining case represents a strong relationship between the antecedent and the consequent: accepting the antecedent forces the consequent to become true. This is the relevance case.

1.5 A NEW CONNECTIVE FOR RELEVANCE IN IR

The essence of relevance principle of the previous section can well be captured by the following expressions:

$$d \Rightarrow q =_{def} (d > q) \land \neg q$$

Based on the above axiomatic systems, we can easily prove the following valid rules and schemata for $\Rightarrow$.

Theorem 1: The following rules are valid in both Stalnaker's and Lewis' systems:

R1	From $A \equiv B$, to infer $(C \Rightarrow A) \equiv (C \Rightarrow B)$
R2	If $\vdash (B \rightarrow C)$ then $(A \Rightarrow C), \neg B \vdash (A \Rightarrow C)$
R3	From $(A \Rightarrow B)$, to infer $(\neg A \land \neg B)$
$\mathrm{IDw}_{\Rightarrow}$	$\neg A \rightarrow (A \Rightarrow A)$
	$\neg(A \Rightarrow \top)$
	$\neg(A \Rightarrow \bot)$
	$(A \rightarrow (A \Rightarrow A)) \rightarrow \neg A$
$\mathrm{MOD}_{\Rightarrow}$	$(\neg A \Rightarrow A) \rightarrow (B \Rightarrow A)$
$\mathrm{MP}_{\Rightarrow}$	$(A \Rightarrow B) \rightarrow (A \rightarrow B)$
$\mathrm{CC}_{\Rightarrow}$	$(((A \Rightarrow B) \land (A \Rightarrow C)) \rightarrow (A \Rightarrow (B \land C)))$
	$((A \Rightarrow C) \land (B \Rightarrow C)) \rightarrow ((A \lor B) \Rightarrow C)$
$\mathrm{CA}_{\Rightarrow}$	$(((A \Rightarrow B) \land (A \Rightarrow C)) \rightarrow (A \land B) \Rightarrow C))$
$\mathrm{CSO}_{\Rightarrow}$	$((A \Rightarrow B) \land (B \Rightarrow A)) \rightarrow ((A \Rightarrow C) \equiv (B \Rightarrow C))$.

Theorem 2. The connective $\Rightarrow$ in both Stalnaker's and Lewis' systems does not have $\mathrm{ID}_{\Rightarrow}$, transitivity, triviality, contraposition, weakening antecedent, excluded-middle ($\mathrm{CEM}_{\Rightarrow}$), and $\mathrm{CS}_{\Rightarrow}$, that is:

$\mathrm{ID}_{\Rightarrow}$	$A \Rightarrow A$
	(or $(A \Rightarrow B) \rightarrow [\neg(A \rightarrow B) \rightarrow (A \rightarrow B)]$)
Transitivity	$(A \Rightarrow B) \land (B \Rightarrow C) \rightarrow (A \Rightarrow C)$
Contraposition	$(A \Rightarrow B) \rightarrow (\neg B \Rightarrow \neg A)$
Weakening antecedent	$(A \Rightarrow B) \rightarrow (A \land C \Rightarrow B)$
$\mathrm{CEM}_{\Rightarrow}$	$(A \Rightarrow B) \lor (A \Rightarrow \neg B)$
CS$\Rightarrow$	$(A \land B) \rightarrow (A \Rightarrow B)$
$\mathrm{CV}_{\Rightarrow}$	$[(A \Rightarrow B) \land \neg(A \Rightarrow \neg C)] \rightarrow [(A \land C) \Rightarrow B]$.

Some discussions. The difference between Stalnaker's system and Lewis system (i.e. CEM) does not seem to affect the connective $\Rightarrow$. For both systems, $\Rightarrow$ does not have $\mathrm{CEM}_{\Rightarrow}$. For all the properties we examined, they are either valid in both systems or in none of them.

In both cases, we have $\neg(A \Rightarrow \top)$ and $\neg(A \Rightarrow \bot)$. These sentences correspond to our intuition that no document should be judged relevant for an empty query ($\top$), or for an inconsistent query ($\bot$).

We also have $\mathrm{IDw}_{\Rightarrow}$: $\neg A \rightarrow (A \Rightarrow A)$ which means that a document is relevant only if it contains new information.

The property $\mathrm{MP}_{\Rightarrow}$: $(A \Rightarrow B) \rightarrow (A \rightarrow B)$ means that the connective $\Rightarrow$ is stronger than the material connective $\rightarrow$. This also allows us to conclude that a model based on one of the two connectives would be stronger than the Boolean model. The valid sentence $\mathrm{CC}_{\Rightarrow}$: $(((A \Rightarrow B) \land (A \Rightarrow C)) \rightarrow (A \Rightarrow (B \land C)))$ means that a compound query (conjunction) may be satisfied by a document if this latter satisfies the components of the query. This is also very intuitive.

However, $\mathrm{CS}_{\Rightarrow}$ is not valid. This corresponds to our earlier analysis that when a query is satisfied, we should not conclude that all the available information is relevant to it.

The property commonly rejected by non-monotonic logics, such as triviality, contraposition, weakening antecedent and transitivity, are not valid for our connective. This shows that the connectives may be suitable to capture the non-monotonic behavior of relevance.

From this analysis we can see that the connective defined corresponds strongly to what we desire to obtain in a broadened IR context. Thus it provides an interesting basis on which relevance may be further studied.

1.6 A PROBABILISTIC ACCOUNT FOR THE NEW LOGICAL CONNECTIVE

In our discussion, we did not consider the uncertainty in relevance judgments and evaluation. Although this is an important problem, it is not the focus of the present paper. If relevance uncertainty is to be considered, we have to define a function P_S that estimates the satisfiability of an expression in a situation S. With such a function, the degree $R_S(d, q)$ of relevance of d to q in S may be expressed as follows:

$$R_S(d, q) = P_S(d > q) - P_S(d).$$

The problem of uncertainty of counterfactual conditional has been studies in our previous paper **?**. A method for evaluating $P_S(d > q)$ has been proposed which is an extension to a previous calculation of probability of counterfactual conditional known as *imaging*. Interested readers may also refer to the chapter by Crestani and van Rijsbergen in this book. We do not go into more detail.

Yet a similar understanding is as follows:

A document is relevant to a query in a situation if providing the information contained in the document causes the query to be satisfied.

Here, we consider relevance as a causal relation. Although we cannot say that relevance and causality are the same notions, we can say that causality is a strong relevance. This may be verified with Sperber's definition of relevance: if the information contained in a document causes the query to be satisfied, then providing that information does have some "contextual effect in that context".

By confining ourselves within "causal" relevance, we do not hope to cover all the aspects of relevance, but to restrict ourselves to a strong relevance.

Many studies in philosophy and AI concern causal relation. Two main approaches have been used so far to estimate causal probability (or relevance): the conditional probability $P(q|d)$ or the probability of conditional $p(d > q)$.

For example, Jeffrey Jeffrey, 1965 suggested that relevance should be measured as follows:

$rel(q|d) = P(q|d) - P(q|\neg d)$ One of the problems with this definition is that it cannot eliminate the following accidental generalization problem [10] :

Suppose that, by coincidence, whenever there is a tornado near Topeka (d), there is a earthquake in Southern California (q); but tornadoes near Topeka does not cause, or is not relevant to, earthquake in Southern California. With Jeffrey's definition, however, we have $P(q|d) - P(q|\neg d) > 0$.

In order to solve the problem, Suppes defines causality as follows [29] :

Suppes' Theory

C causes E iff $P(E|C) > P(E)$ and there is no factor F such that $P(E|C, F) = P(E|F)$.

The element F is called the screening factor. A screening factor between C and E is such a factor that with its presence, C has no effect on E. The consideration of the screening factor in Suppes' theory allows to avoid the accidental generalization problem. For our example, the screening factor is the tectonic motion along the San Andreas fault. Thus tornadoes near Topeka does not cause earthquake in Southern California according to Suppes' theory.

However, this theory is not applicable when $P(C) = 0$, i.e. if the condition is contradictory to the given situation.

Causality has also been studied using conditional logic. It has been shown that the logic is well suited to consider the counterfactual cases, i.e. the antecedent contradicts the given situation. Typically, the conditional $A > B$ is read as "if A were the case, then B would be the case". Roughly speaking, the difference between the material conditional $A \to B$ and the counterfactual conditional $A > B$ lies in the counterfactual cases: when the antecedent A contradicts (or is false in) the given situation, the material conditional is evaluated to true systematically. This is not the case for the counterfactual conditional. In the same counterfactual case, the evaluation needs to consider the "nearest" alternative of the given situation in which the antecedent A becomes true. The truth of the counterfactual conditional is determined by the truth of the consequent B in that alternative situation. By considering the alternative situation, we can examine the possible link between the antecedent and the consequent.

The relationship expressed by a counterfactual conditional is also stronger than the probability dependence expressed by a conditional probability. To describe the difference between $P(B|A)$ and $P(A > B)$, let us

probability. To describe the difference between $P(B|A)$ and $P(A > B)$, let us first see how the probability of a proposition A is evaluated in general. Suppose we have a possible-world structure. Each world corresponds to a particular context in which every logical atoms has an evaluation. The correspondence of the possible world structure to the real world is established by distributing the total probability (i.e. 1) among all the worlds. Then the probability $P(A)$ is determined by the summation of the probabilities of all the worlds in which A is true. Having this evaluation method in mind, the difference between the two earlier probabilities may be explained as follows:

To evaluate $P(B|A)$, we confine our consideration to the set of the worlds in which A is true. We need to create a new distribution of the total probability within this set of worlds. This is done by redistributing the probabilities of the other worlds (in which A is false) to the worlds within that set, and the redistribution is proportional to the original probability of the worlds in the set. Then the conditional probability is the summation of the new probabilities of the worlds in which B is true. In terms of formula, it is equivalent to $P(B|A) = P(A \wedge B)/P(A)$. : the other worlds is done according to their closeness (to be determined by extra-logic criteria) to the worlds in the set. It is usually not proportional to the original probability of the worlds. Then $P(A > B)$ is the summation of the probabilities of the worlds in which B is true.

We can notice that even if the probability $P(A) = 0$, it does not prevent us to evaluate the probability $P(A > B)$ because the redistribution is no longer proportional to the original probabilities.

Conditional logic is one of the most used tool in AI to consider problems such as non-monotonic reasoning [15] and belief revision [13] , which may all encounter the counterfactual cases. In our earlier analysis, we described an IR operation as a potentially counterfactual problem, because when situational factors are considered, the information provided by a document under consideration may contradict the situational factors. It is thus suggested that conditional logic may be used as a general tool in order to avoid the unreasonable evaluation in such cases. Along with this line, we can approximate relevance using $P(d > q)$. This is the approach used in [18] . Although interesting properties have been obtained in this way, some properties of $>$ are undesirable for relevance in our broadened context. For example, the following property of $>$ is not appealing:

$$\vdash d \wedge q \rightarrow (d > q)$$

It would mean: if the content of a document is already known and a query is already satisfied, then the document is relevant to the query. In other words, any available information is "relevant" to any satisfied query. But in this case, it is useless to retrieve such an information.

Another questionable property is the following:

$$\vdash d > \top$$

That is, any document is relevant to an empty information need (i.e. a requirement asking for nothing). In this case it is more reasonable to provide the user with no document rather than with all the documents.

Yet another questionable property is reflexivity:

$$\vdash d > d$$

This property may seem intuitive in IR. However, we cannot accept it as valid in the case where the user is only interested by new information. If the information contained in a document is already known to the user, this document is irrelevant for the user even if it concerns exactly the topic of the query. Thus this property is not always desirable for relevance.

All the above properties are in fact related to the state of available information. In order to capture the criterion concerning available information, we suggest to model relevance as follows:

Principle

Given document d and query q, d is relevant to q if $P(d > q) > P(q)$

It means: If in a hypothetical situation where a document is provided, the query is better satisfied than it is, then the document should be considered relevant.

This estimate corresponds closely to our earlier understanding of contributes to a better satisfaction of q than in the current situation. It also corresponds to the following theory of causality [10] :

Giere's Theory
C causes E iff $P(C > E) > P(E)$ and there is no factor F such that $P((C \wedge F) > E) = P(F > E)$.

The difference between them is as follows:

In Giere's theory, F is used to detect screening factors between C and E. In a very large perspective, the screening phenomenon may occur in IR: given a query, a first document (for example, better written) may be always preferred to a second document about the same subject. In this case, the first document may "screen" the second. That is, in the presence of the first document, the second document is no longer relevant. This is a phenomenon to be considered when we want to evaluate the relevance of a stream of documents to an information need, rather than a single document. In this paper, however, our scope is more restricted: we only consider a single document at a time. The screening effect among documents is not dealt within this paper.

In this section, we considered two alternative approaches to consider relevance in IR. We indicated that the approach based on conditional probability may face with problems when situational factors are considered. The approach based on conditional logic allows us to avoid these problems. However, the problem with conditional logic lies in the difficulty to evaluate it. Indeed, to evaluate a counterfactual conditional, we need extra-logical criteria to set the closeness relation among all the worlds. There have been extensive studies on this subject (for example, [13, 15, 16]).

1.7 AN OUTLINE FOR COPING WITH CHANGES IN RETRIEVAL SITUATIONS

The basic idea of our framework is the dependence of relevance evaluation on the retrieval situation. Different users have different retrieval situations, and the same user also has different retrieval situations at different moment. Therefore, it is crucial to have a way to construct different situation models for different users, and also to adapt a situation model to a particular user during the uses of the system. The construction of a situation model may be seen as a particular problem of adaptation: to construct is to adapt from an empty model. So the central problem for us is how to change a situation model according to the user.

There is an extensive literature on changes in belief sets (**?**; **?**; **?**). As our retrieval situation is modeled as a belief set, operations on belief sets make changes in retrieval situations easier to express. However, several problems specific to IR and to our modeling approach need to be dealt with.

1.7.1 When does a situation need to be changed?

One may imagine a possible, though maybe naive, way which is to give the whole control to the user: When the user thinks that the model of the situation is not or no longer suitable, she may ignite a change in it by giving new pieces of knowledge and information that she considers "true". The system then renews its model of situation in consequence. This involves two main tasks:

1. putting the new pieces of knowledge and information in the new situation model;
2. re-examining the old pieces of knowledge and information in the retrieval situation to see if one should keep them or throw them away.

This approach is not realistic as users usually are unable to determine which piece of knowledge or information they should add into the situation model. Another possible way to ignite a change in a model of retrieval situation is by taking the user's own relevance judgments as new beliefs. These new beliefs are then used to renew the situation model so as to accommodate these beliefs. Suppose, for example, that in a situation modeled as S the system judges that the document d_1 is relevant to the query q, and the document d_2 irrelevant. That is: $S \vdash d_1 > q$ and $S \vdash \neg(d_2 > q)$. However, the user gives different judgments, telling that d_1 is not relevant, whereas d_2 is. Then the sentences $\neg(d_1 > q)$ and $(d_2 > q)$ are considered as new beliefs. The situation model S is to be changed to S' so that $S' \vdash \neg(d_1 > q)$ and $S \vdash d_2 > q$.

This is a more realistic way to adapt the model of retrieval situation. It looks similar to the approaches using relevance feedback. However, there is the following difference between them: In the previous approaches relevance feedback is usually used to modify the query representation (query vector). However, the mechanism for document-query comparison (the calculation of vector similarity) remains the same. Relevance feedback only has effect on the that particular query. It may be considered as a transient learning from the user. The approach we suggest above uses relevance feedback to change the model of retrieval situation, as a consequence, the method of relevance evaluation. After the change, the same document-query pair no longer has the same evaluation. This approach may be seen as a persistent learning from the user. In our further discussion, we will focus on this approach.

1.7.2 *What kind of change in a situation?*

In AI studies two basic kinds of belief changes facing a new belief have been identified: belief revision and belief updating **?**; **?**. Both modify a belief set to accommodate the new belief in a minimal way. After the change, the new belief becomes completely true, whereas the old beliefs should be re-evaluated. The difference between them lies in the *preservation condition*. Revision is preservative whereas updating is not. Preservation condition for revision is expressed as follows:

$$\text{(P)} \qquad \text{If } \neg A \notin S \text{ and } B \in S, \text{ then } B \in S\dot{+}A$$

where S is a belief set and $S\dot{+}A$ is the result of revising S with respect to A. This preservation condition means: if a new belief (A) is consistent with the old belief set (S), then every belief (B) in the old belief set is preserved (i.e. it is still in the new belief set $S\dot{+}A$). This condition is justified by the principle of information economy, i.e. an acquired belief should not be lost unless we are forced to abandon it.

While this principle is intuitively appealing, it does not fit in all kinds of belief changes. In **?**, it is pointed out that revision and updating deal with two different kinds of belief change. Revision is appropriate for changes facing a new belief about an unchanged real world, i.e. one knows more about the same real world. On the other hand, updating is appropriate for belief changes facing a new belief due to a change in the real world. The following example is used in **?** to illustrate the difference between the two operations.

Suppose that all we know in S about a particular room is that there is a table, a book and a magazine in it, and that either (α) the book is on the table, or (β) the magazine is on the table, but not both. The belief set contains $(\alpha \wedge \neg\beta) \vee (\neg\alpha \wedge \beta)$.

Now if we learn (from somebody else) that (α) the book is on the table, then we would conclude $(\alpha \wedge \neg\beta)$. This is the result obtained by belief revision because of a new information about the unchanged real world.

On the other hand, if we do not learn that α, but order a robot to put the book on the table. As a consequence, we know that α becomes true. However, we do not know whether β or $\neg\beta$ because the robot did nothing on the magazine. So β or $\neg\beta$ remains as before. The changed situation may be represented as $\alpha \wedge (\beta \vee \neg\beta) = \alpha$. This second case is a belief updating due to a change in the real world.

In our case, the real world is the user's real retrieval situation. According to whether the user changes it, revision or updating should be used respectively. If the difference in relevance judgment is due to a change in the real retrieval situation, the old retrieval situation model should be updated. On the other hand, if there is no change in the real retrieval situation, the difference in relevance judgment is due to the fact that the situation model does not fit the real retrieval situation. Then a revision operation is appropriate to make the model better suited.

The distinction between the two operations in IR is difficult to detect in practice. It is unreasonable to ask the user to tell the system which operation is appropriate because changes in real retrieval situations are often smooth, and the user himself is not always aware of them.

The preference between the two operations can only be given from a pragmatic point of view. From the point of view of implementation, updating is more difficult to do than revision. In the case of revision, the revised belief set is the one which is minimally changed from the old belief set. In the case of updating, one needs to examine all the models (worlds) of the belief set, and to determine the closest model for each of them. This is a very expensive operation. With the great number of concepts and information involved in IR, an direct implementation of the updating operation would be inefficient. From this point of view, it is preferred to choose revision as the default operation when a disagreement in relevance judgment is detected between the system and the user. Updating may be called occasionally when the user indicates that a radical change takes place in his real retrieval situation.

Another reason in favor of revision is the fact that, at the current stage of situation modeling, there are much more chances for the system to disagree with the user because of the imprecision and errors in the situation model than because of changes in the real retrieval situation. So revision is an operation which should be implemented in priority.

This discussion leads to the following suggestion: Given that the the user's judgment $\neg(d > q)$ (or $d > q$) disagrees with the system's judgment $S \vdash d > q$ (or $S \vdash \neg(d > q)$), the user's judgment is used as the new belief to revise the situation model S. The revised situation model S' is $S \dot{+} \neg(d > q)$ (or $S \dot{+} (d > q)$ respectively).

1.7.3 What else do we need to change a situation model?

To change a belief set, it is necessary that there is an ordering among the beliefs. Gärdenfors calls it epistemic *entrenchment* **?**. According to the ordering, some beliefs are more entrenched (believed more firmly) than some others. To change a belief set which is contradictory with a new belief, one has to give up some old beliefs while adding the new belief in order to keep the whole belief set consistent. When there are several old beliefs so that if one of them is retracted the whole belief set will be consistent, the ordering is useful for determining the least entrenched one to give up first.

Epistemic entrenchment is not a characteristic determined by beliefs themselves. If we only look at the beliefs (logical sentences), there is no indication on which one is more entrenched than another. Rather some extra-logic criteria should be applied to determine the ordering.

There are several possible approaches to create an ordering among sentences (beliefs) in a model of retrieval situation. One of them is to ask the user to partially order her relevance judgments. This approach has already been taken in several other studies (e.g. **?**). It requires that, in addition to tell which document is relevant and which is irrelevant, the user is also asked to indicate how certain she is about her judgments. If a document d_1 is firmly judged relevant to the query q, and another one d_2 is judged relevant with a lesser degree of certainty, then the relation $d_1 > q$ is more entrenched than $d_2 > q$. One way to achieve this is to classify relevance judgments into several groups. For instance, "relevant", "relatively relevant", "not very relevant", "not relevant". Relations $(d > q)$ in "relevant" group are more entrenched than those in the "relatively relevant" group; and relations $(\neg(d > q))$ in the "not relevant" group are more entrenched than those in the "not very relevant" group.

Another possible way is to determine the ordering of beliefs according to their effects to relevance judgments. If a belief, when used to make a relevance judgment by the system, leads to an agreement with the user, then its degree of entrenchment is increased; if its use leads to a disagreement, its degree of entrenchment is decreased. In a long run, it may be expected that firmly believed sentences will be attributed higher degrees of entrenchment than more uncertain beliefs. This approach is similar to the adaptation approach used in **?**.

With a partial ordering of beliefs, it becomes possible to determine the weakest belief(s) that contradicts the new belief to give up. One may find descriptions of such operations in **?**.

1.7.4 Is the approach feasible?

We suggested to change a situation model (belief set) according to a belief which is expressed as a counterfactual conditional. This problem has never been studied in AI in which one is concerned with the revision or updating with respect to sentences of classical logic (no counterfactual is concerned). The question one may raise is whether such a change is feasible. This question is equivalent to the following one: is it possible to check the consistence of a counterfactual conditional with a set of sentences?

The results of de Swart **?** and Gent **?** give a positive answer to this question. They showed two sound and complete proof systems for two counterfactual conditional logics (VC and VCS). Their systems are based on the tableau technique, using a limited set of inference rules. Their results show that in the two logics, the validity or invalidity of any counterfactual sentence $A > B$ with respect to a set of sentences may be proven by applying the inference rules. So we are able to check the consistency of a set of sentences including counterfactual conditionals.

From a theoretical point of view, a general approach to revise a belief set S according to a counterfactual conditional $A > B$ may be as follows:

- Take $\{A > B\}$ alone as the initial new belief set S'.
- In the decreasing order of entrenchment, sentences in S are examined.

  - If a sentence C in S is consistent with S', then it is added in S'; otherwise, it is not added.

  - If there is a subset of sentences $C_1, C_2, ..., C_n$ of the same degree of entrenchment such that deleting one of them is enough to make the whole consistent with S', then the following disjunctive sentence is added in S': $(\neg C_1 \wedge C_2 \wedge ... \wedge C_n) \vee (C_1 \wedge \neg C_2 \wedge ... \wedge C_n) \vee ... \vee (C_1 \wedge C_2 \wedge ... \wedge \neg C_n)$.

Although theoretically feasible, from a practical point of view, the above approach is inefficient. In particular, in the last case, as the system is unable to determine the exact piece of belief to give up, the smallest change (i.e. the one which retains all the alternatives) is taken. Since we can only determine a very partial ordering for relevance judgments, this case often occurs. So, more and more long disjunctive sentences will be produced in the situation model. At the end, the reasoning process of the system will become completely inefficient. A possible solution to gain efficiency is to allow the user to further order the alternatives, or to choose one of them, whenever possible.

The conclusion is that although the approach is theoretically feasible, an efficient implementation method should be found before its practical use in IR.

1.8 COMPARISON WITH SOME OTHER STUDIES

There has been an attempt in constructing an axiomatic system for the notion of aboutness [6] . Aboutness and relevance are different: a document about some concept is not necessarily relevant to a query on that concept, but the reverse is true. The difference precisely lies in topicality and relevance. Aboutness deals with topicality only.

The present work is related to [22, 23] in which the use of conditional gic in IR is first suggested. However, no strict logic system has been established wherein. Some idea implied in conditional logic (imaging) has also been used in IR implementation [8] , but the corresponding model does not deal with situational factors. Imaging has been used as an alternative way to estimate relevance probability in a non-counterfactual case.

Although we have contrasted the probability of conditional with the conditional probability, the two probabilities are not completely independent. In fact, as shown in [14] , under the following two conditions, they are equal to each other:

condition 1: $d \rightarrow [(d > q) \equiv q]$ is a logical truth. condition 2: d$>$ q is stochastically independent of d, i.e.: $P(d > q|d) = P(d > q)$.

Assertion 1 [14]

Under the two conditions above, if $P(d) > 0$, then

$$P(d > q) = P(q|d)$$

.

The first condition is entailed by CEM and Modus Ponens for $>$: $A \wedge (A > B) \vdash B$, thus true in Stalnaker's system. If we further assume condition 2, then we may raise the question about the difference between the two probabilities in our cases. The most noticeable difference lies in the case where we have $P(d) = 0$. This is exactly the counterfactual case raised by situational factors.

Based on Gibbard's assertion, we can easily infer the following assertion:

Assertion 2

A probabilistic IR model based on $P(q|d)$ may be expressed as a particular case of a model using $P(d > q)$ in which the two mentioned conditions are verified.

This assertion shows that our approach may be a generalization of the existing probabilistic IR models.

1.9 CONCLUDING REMARKS AND FURTHER RESEARCH

Cognitive studies suggest that relevance judgments in IR systems should take more situational criteria into account, but the current computational models are unable to do it. The framework we suggested enlarges the scope of modeling by including situational factors. This study is a step toward to reduce the gap between computational models and cognitive models.

Our framework is based on 1) a representation of relevance relation by counterfactual conditional, and 2) a representation of retrieval situation by belief set. We have shown the advantages of using counterfactual conditional instead of material conditional. Then an approach is outlined to cope with changes in retrieval situations. This, however, was done only at a theoretical level.

To compare with some related works, let us point out our different utilization of relevance feedback. Relevance feedback has been used for, among others, 1) reformulating a query, and 2) establishing a preference order among a set of documents **?**. In the first case relevance feedback information only has a temporary effect on the system's behavior. Once the query has been evaluated the relevance feedback information is forgotten by the system. In the second utilization, relevance feedback has a long term effect on relevance judgments. The preference order established is used to choose the most preferred documents first. However, several problems may arise in this approach. 1) The study suggests to order documents without considering the query. The same order is thus created for all the queries. This is not realistic because queries express the most important criteria for relevance judgments. 2) Once an order has been created among documents, there is no mechanism provided to change it. Thus the approach cannot cope with the dynamism in relevance judgments. Our approach gives a possible solution to both problems.

In this paper we only dealt with the validity of a counterfactual conditional (relevance relationship) with respect to a belief set (retrieval situation). The uncertainty problem was not our focus point. However, uncertainty is an inherent characteristic of relevance judgment. It is important for a model to deal with it correctly. One way to cope with the uncertainty is to define a probabilistic function $P_S(d > q)$ with respect to a retrieval situation S. To do this, we have to apply the technique of imaging **?**. The utilization of imaging in IR has been investigated in **?**. For imaging in our case, things are more complicated: in **?** no situation is considered, whereas in our case imaging should be related to a situation. In our previous study **?** this problem has been investigated. The proposed approach can still be applied here.

We did not deal with the implementation problem of the model. The is a very complex problem. This same problem is also the subject of many studies in AI which aim to find efficient methods to revise and to update belief sets. Our future work will deal with the implementation issue of the proposed model.