# The Ostensive Model of developing information needs

**Iain Campbell** and **Keith van Rijsbergen**

Computing Science
University of Glasgow
Glasgow G12 8QQ

{iain, keith}@dcs.gla.ac.uk

Accepted to *CoLIS-2*, Copenhagen, October 1996.

**Abstract**

We present a model of the progressive development of information needs. It is a model that recognises the changing uncertainty inherent in a user's cognition of his information need. The approach centres around the collection and combination of ostensive evidence. We present uncertainty profiles associated with the discounting of ostensive evidence with respect to its age, and relate that to a new notion of relevance - Ostensive Relevance. This notion recognises the transient, inaccessible, spatio-temporal nature of relevance. We describe how these components come together to allow the Ostensive Model to be integrated with the traditional Binary Probabilistic Model. We describe the integration and show that it reveals an implicit assumption in the conventional estimation procedure for a particular conditional probability. The temporal aspects of the Ostensive Model allow a weakening of that assumption. The integration allows direct implementation of the Ostensive Model. Finally, we present an example that demonstrates the intuitive appeal of the approach over existing approaches to Relevance Feedback.

## 1 Introduction

This paper details a model of information needs that forms the central component of a new approach to the provision of support for information seeking. The approach is one where the seeking takes place in a dynamic graphically-presented information space. It is an environment free of textual requests or queries, where the user is never exposed to internal representation techniques or retrieval methods - instead, the user merely browses across information objects and links between them.

The retrieval system is driven by observational evidence collected from a browsing user who is assumed to have a developing information need - ie a need that changes over time as a result of exposure to information.

For a deeper understanding of that overall programme, of which the work of this paper is a component, the reader is directed to (Campbell, 95) - we shall make occasional reference to it as 'the previous report'. Complementing that report, the scope of this paper is restricted to the description of the model, its implementation, and its advantages.

The structure of this paper: Section 2 presents the Ostensive Model for developing information needs, and relates that development to observational evidence and uncertainty attached to that evidence. Section 3 integrates the Ostensive Model with the Binary Probabilistic Model to allow direct implementation, and introduces a new definition of relevance based upon uncertainty attached to ostensive evidence. Section 4 provides a worked example that displays the intuitive appeal of the model.

## 2 A model for developing information needs

In this section we first introduce terminology. We then describe our model of cognition, ostensive evidence, and uncertainty attached to such evidence.

MacKay (Mackay, 69) proposed that one may consider the brain as a black-box. This means that although one cannot understand the exact working mechanisms involved in cognition one can still hypothesise about the effects of its operation indirectly through the inputs provided to it, and through the outputs resulting from it. We adopt this view, and therefore, in this report we avoid absolute statements or statements regarding procedures involved in the internal workings of the brain.

### 2.1 Ontology

Following the ideas of MacKay, *Information* is defined as merely a sequence of bits, symbols, signs, etc. It is unchanging and absolute. If a modification is made to it, it is a different piece of information. This is quite different from the definitions used in fields such as library and information science where such matter is regarded as *data*, or *potential information* - that only becomes information when perceived by a cognitive agent (Ingwersen, 84). In the MacKay view, no such distinction is made, and no such implicit importance is given to a cognitive agent - information is information regardless of the existence of agents, and it will remain so after any activities of such agents.

The brain of a human is regarded as a black-box pattern-response mechanism. It is a probabilistic mechanism that has a *state of conditional readines* to output certain responses conditional on certain inputs. These inputs and outputs are information, as is the instantaneous state of the brain itself. We shall refer to that state as a *knowledge state*.

The brain has inbuilt mechanisms that not only produce outputs as a result of inputs, but that also modify internal probabilities related to the particular input received - ie it recognises and adapts to patterns in inputs. This is termed an *internal matching response*. It is by that mechanism that the brain learns.

Given the input of information, *interpretation* is the process of the internal matching response, and *meaning* is the resultant change in the state of conditional readiness. Here, the brain (ie its state and matching behaviour) is the *context* within which the information is being interpreted to produce a meaning. This process may or may not result in the output of information from the black box.

For notational convenience, we will refer to this interpretation process as the *exposure* of information to a brain.

In this ontology, information itself has no intrinsic meaning, but can be said to have as many potential meanings as there are contexts in which it could be interpreted. This point becomes important later where the lack of any objective meaning *in* a piece of information is central to a new definition of relevance

### 2.2 The components of the Ostensive Model

The model relates changes in the knowledge state of a user in response to information encountered during information seeking activities. We present the model diagrammatically with associated propositions and assumptions. The core components of the diagram are shown in Fig 1.

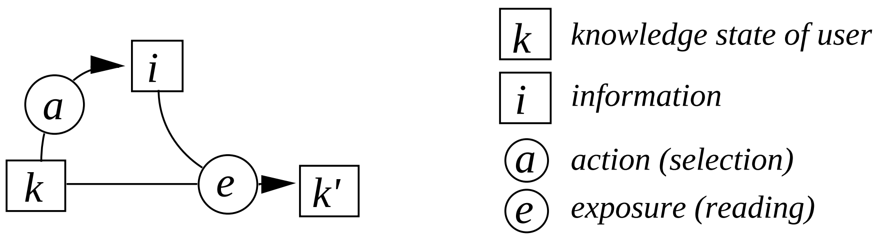


Figure 1: *The updating of a knowledge state through the selection of, and subsequent exposure to, information.*

Let *k* denote a knowledge state (ie an instantaneous state of conditional readiness). Within the context of information seeking, we make the following proposition regarding *k*:

> ***Proposition 1*** *Of all the motivating factors present in a user's knowledge state, those influencing the immediate actions of the user to the greatest extent are those pertaining to the information need most directly.*

Let *a* denote the actions referred to in P1. With respect to those actions, we propose the following:

> ***Proposition 2*** *The actions motivated by an information need are most likely to be those that will obtain information from the environment that is regarded by the user to be the most likely to satisfy their information need.*

Let us restrict the environment within which this information seeking behaviour is taking place to that of a range of information items, each with an attached highly abstracted form - eg documents showing only titles. In such an environment, the possible actions available to the user are limited to selecting and reading documents. Here, P2 would indicate the selection of the object that appears to the user to be the most likely to be relevant the information need - eg a document whose title or abstract suggested the highest relevance. Let *i* denote the information in that selected object.

The user would then expose himself to the information that made up that object - eg the user would read the document. This exposure of *k* to *i* would result (in all but the most trivial of cases) in an internal matching response. This internal update would therefore result in a new knowledge state *k'*. Let *e* denote that process.

In the Mackay view: *e* is the process of interpreting *i* with respect to, or within, a context *k*. That process of interpretation results in a change from *k* to *k'*. That change is the meaning of *i* with respect to *k*. The meaning is therefore clearly dependent upon the *k* within which it is formed. Were it a different *k*, for example a different user, or the same user on a different occasion, then the meaning would be different. This dependence of meaning upon interpretive context, although from a different perspective, is consistent with the 'cognitive viewpoint' of, for example, de Mey (deMey, 80).

The Ostensive Model does not say anything about how the processes involved in *e* proceed - it says only that a change from *k* to *k'* results from its exposure to *i*. Regarding the *nature* of the change, we propose the following:

> ***Proposition 3*** *Given P1 and P2, the majority of the changes in k, resulting from exposure to i, will be in those areas of k pertaining to the information need most directly.*

This new knowledge state *k'* may itself provoke an action *a'*, selecting a new information item *i'*, which through exposure *e'*, results in the further updated knowledge state *k"*. This process may iterate (Fig 2).

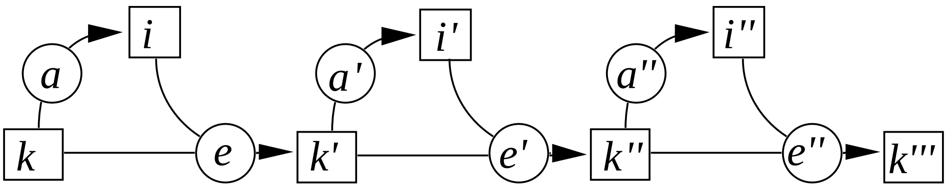


Figure 2: *The iterative updating of a knowledge state.*

## 2.3 Observables and non-observables

As already stated, the nature of *k* and the process *e* of it internally updating in response to incoming information is something that we cannot access directly, and of which we have little understanding. That nature and process are things about which we can currently only theorise on the basis of actions resulting from them - ie the observed behaviour of a user.

With the above in mind, the components of the model can be classified into those that can, and those that cannot be observed directly (Fig 3).

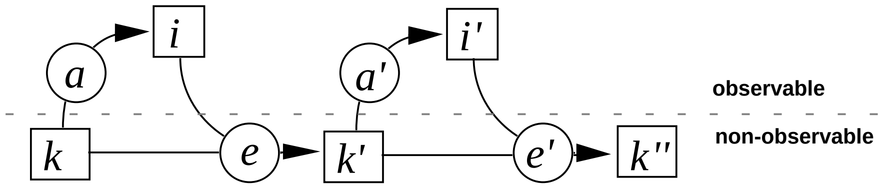


Figure 3: *Observable and non-observable components.*

One may consider the line separating the two classifications as representing the interface between the user and the outside world. Those components below the line are internal to the user and those above are external.

The action *a* of choosing from a selection (particularly a limited selection) of information objects can be observed, for example, in terms of the number and the nature of the objects rejected. The information *i* chosen (whether electronic or otherwise) can be observed trivially.

The other components (*k* and *e*), although not directly observable, may not be completely opaque. We intend to show that a limited grasp of certain qualitative aspects of *k* and *e* (that are of use in supporting information seeking) may be discerned from their observable manifestations.

P1 states that the action *a* is most strongly influenced by the information need aspects of *k*. This means that *a* is indicative, in some way, of those aspects of *k*. Similarly, P2 states that the information *i* chosen, is that regarded as most relevant to the information need. Thus, *i* is also indicative of the information need. P3 states that the change from *k* to *k'*, resulting from the exposure to *i*, will affect the information need most. This means that *i*, not only being indicative of *k* (by P2), will also be definitive of *k'*- ie it will be the only factor apart from *k* that directly affects the nature of *k'*.

We now have *a* and *i* being indicative of their associated *k*, and *i* being strongly indicative of the associated *k'*.

The same principles could be applied to the relationship between the observables and *e*. Nevertheless, as *e* immediately produces a new *k'*, and it is *k'* (the instantiation of the information need that we are trying to capture) that then determines the next *a'* and *i'*, a determination of *e* itself appears less interesting. Therefore, its presence is acknowledged but not investigated.

## 2.4 Ambiguity in the observables

Even upon unreserved acceptance of the above arguments regarding the relationships between observables and non-observables, it would seem unrealistic to expect to be able to glean a large amount of unambiguous information about *k* from single instances of *a* and *i*. Countless meanings could result from the tuple *(a, i)* and the certainty that could be attached to any single such possible meaning would therefore be low.

After several iterations of the *k* to *k'* process, several instances of *a* and *i* will have resulted and be available for observation. Individually, the possible meanings of each tuple would carry with them the same ambiguity. Nevertheless, when taken together, the tuples (unless made up of indistinguishable or identical instances of *a* and of *i*) would naturally demonstrate ambiguity resolution characteristics.

Similarly, taking together several documents indicated as relevant by a user, one can obtain a clearer idea of their interest than would be possible with only one. This principle is borne out in the efficacy of the Relevance Feedback process with which we are familiar in IR.

## 2.5 Ostensive definitions from observable evidence

The Ostensive Model and Relevance Feedback are processes that rely upon ostension - ie they implicitly exploit ostensive definitions:

> ***ostensive definition*** *(Philos.), the explanation of a word by presenting, pointing at, or otherwise indicating one or more objects to which it applies.* (OED, 93)

For our purposes, the term ‘word’ is taken as a denotation of the abstract notion ‘relevance to an information need’. In Relevance Feedback, a user explicitly indicates to the system documents regarded by him as relevant. From those indications, the system identifies attributes whose prevalence in the set of relevant-indicated documents differs from that in the collection as a whole. Those attributes are then used by the system as the criteria for selecting other possibly relevant documents.

Implicit in the above dictionary wording, and common in fuller expositions of ostension (eg (Quine, 53) and (Quine, 69)), is the assumption that evidence contributing to an ostensive definition are purposeful acts of communication. Such acts are the basis of conventional Relevance Feedback approaches.

We suggest that the restriction be lifted, or at least relaxed, to include other actions not intended by the actor as communicative. Evidence gathered merely from observation of a user (ideally unaware of that observation) engaged in information seeking would be an example.

This relaxation is not proposed in order to increase the amount of evidence available, but instead to change the manner by which evidence is generated and collected. In the previous report a desire was expressed to remove the requirement of explanation of an information need that currently burdens a user interacting with a query based IR system - ie switch from a reliance upon explicit communicative acts to only observed actions. Despite that desire, it was conceded that such evidence is useful in driving a process of system adaptation to individual users and information needs - ie any replacement would also need to be able to drive such a process.

Ostensive definitions of information needs could be built from observational evidence that has been collected, for example, from a user interacting with a browse based IR system. Such evidence could be collected without interference with the user; without requiring description; and without exposing the user to any internal representation methods used by the system. It would allow the user to concentrate on the task in hand - ie identifying relevant information. A user may not be good at describing his information need but is, by definition, able to identify something that is relevant to him - in fact, he is the only agent capable of doing so.

The idea proposed in the previous report was that such ostensive definitions of the information need be used to select appropriate information objects for the ‘candidate next steps’ in a browse path. In that way, the system could adapt to individual users and information needs throughout a session, but without the intrusive extraction of descriptions, and without even the (admittedly less intrusive) need to explicitly indicate the relevance or otherwise of individual objects. The act of selecting a next step would be taken to be an implicit assessment of relevance. We therefore intend to replace the communicative acts completely with observational evidence.

The candidate next steps, would be the restricted environment, within which the action *a* of selecting an information object *i* would take place. The collected ostensive evidence would then lead to an ostensive definition of *k*’, which would be used to predict the information objects most likely to be relevant to that *k’*. Those objects would be the next set of candidates and hence the next environment from which information is selected.

## 2.6 Uncertainty in ostensive evidence

The degree of uncertainty that we can attach to inferences about *k* and *k’* from *a* and *i* in each tuple *(k,a,i,k’)* appears impossible to determine absolutely. Nevertheless, we believe that, under certain conditions, comparisons can be made between the degrees of uncertainty associated with individual tuples.

Taking Fig 2 as an example, we have three iterations, each of which can be represented by a tuple: *t1=(k,a,i,k’) , t2=(k’,a’, i’, k”) , t3=(k”, a”, i”, k”’)*. We also have the time ordered sequence: *t1,t2,t3* - ie *t1* occurred before (and led to) *t2*, which occurred before (and led to) *t3*. It is the central importance of the time sequence that sets the Ostensive Model apart from the models implicit in current approaches. The time ordering of evidence will provide the key to relative degrees of uncertainty attached to individual pieces of ostensive evidence.

After the first iteration, the current knowledge state is *k’*. *a* and *i* are our best, and only, ostensive evidence from which to make inferences about *k’*. After the second iteration we will be trying to make inferences about *k”* based upon the accumulated evidence *a, i, a’, i’*. After the third, inferences about *k”’* will be made based upon *a, i, a’, i’, a”, i”*. Therefore, as the amount of ostensive evidence increases, our uncertainty of having an accurate representation of the information need will reduce.

We see no grounds to say that, for example, the degrees of uncertainty associated with inferences within *t1* are any greater or less than those for the inferences within *t2* or within *t3*. For example, *a* and *i* say as much about *k’* as *a”* and *i”* say about *k”’* - ie the uncertainty of intra-tuple inferences are not comparable.

The model shows the causal path of all the ostensive evidence leading to the current knowledge state *k"*. We believe it to be self evident that there is a more direct path from *a'* and *i'* to *k"* than there is from *a* and *i* to *k"*. Note, that we are now talking exclusively with respect to *k"* - ie the current knowledge state. *k'* is now of little importance as it has disappeared.

With the directness of the causal links in mind, it seems reasonable to assume that the uncertainty attached to inferences about *k"* based upon *a* and/or *i* will be higher than the uncertainty attached to similar inferences based upon *a'* and/or *i'*. This can be generalised as:

> ***Proposition 4*** *As the age of an item of ostensive evidence increases, the uncertainty attached to inferences made upon it about the current knowledge state will increase.*

## 2.7 Uncertainty profiles

The model tells us that the uncertainty attached to evidence increases with the 'age' of that evidence, the question arises of the nature or form of that aging. In an effort to support further the suggested relationship and to give intuitions as to the meaning of such relationships in general, we shall not restrict the discussion to only increasing 'profiles'. We will discuss a decreasing, a flat, and finally the preferred increasing profile.

Consider first a profile where the uncertainty decreases with the age of the evidence (Fig 4). In all the profiles we shall present, the 'age' axis will increase to the left in order to retain the normal presentation of the passage of time as left to right. The zero point on the age axis is the current time, with the axis growing into the past.

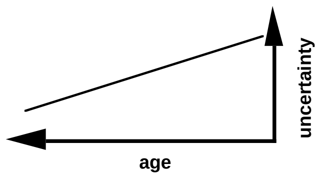


Figure 4: *A decreasing profile of uncertainty w.r.t. age.*

This profile implies that old evidence is more indicative of the current knowledge state than more recent evidence. This means that the early evidence has the most influence on the ostensive definition, and that subsequently observed evidence becomes of less and less importance. In practice, the accumulation of additional evidence quickly becomes insignificant, with the ostensive definition changing little, if at all - ie the ostensive definition will effectively converge. This has the disturbing implication that the user knew what he wanted at the start, got it, and is now simply wasting time by continuing to expose himself to information. In effect, the precise antithesis of the Ostensive Model.

Conventional Relevance Feedback approaches assume that all the evidence (in the case of that model evidence can only be documents) has an identical degree of uncertainty. This equates to a flat profile of uncertainty with age (Fig 5).

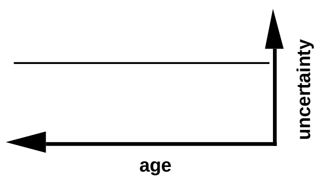


Figure 5: *A flat profile of uncertainty w.r.t. age.*

Thus, Relevance Feedback implicitly assumes that either: all evidence is generated by the same knowledge state, or that all the generating knowledge states are identical. In the Ostensive Model, such a situation is impossible, as exposure to one document changes the knowledge state and thus increases the uncertainty associated with it before another document can be observed as evidence.

In practice, the combination of the evidence in Relevance Feedback is an accumulating process with each observation contributing the same amount of evidence. This results in the ostensive definition converging steadily as more evidence is collected. This is a similar process to that of the decreasing uncertainty profile, although less extreme.

The profile suggested by the Ostensive Model is one of increasing uncertainty with age. Such a profile is shown in Fig 6.

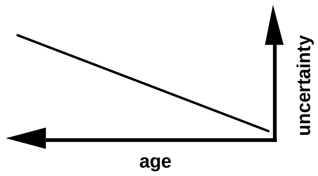


Figure 6: *An increasing profile of uncertainty w.r.t. age.*

This corresponds to the model in that the most recent evidence has the lowest attached uncertainty and will thus have the most influence on the ostensive definition. Here, all ostensive evidence plays a part in the ostensive definition, nevertheless, the most recent will play the greatest. This means that the ostensive definition will follow recent trends in the ostensive evidence, but will always have a component of the historical evidence.

The rate at which the uncertainty attached to evidence increases with age can be altered whilst still retaining the general increasing nature required by the Ostensive Model. Further, the rate of change of uncertainty can also be modified within a particular profile - Fig 7 shows two such profiles.

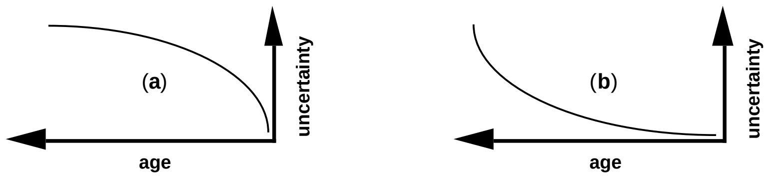


Figure 7: *A decelerating (a), and an accelerating (b) increase in uncertainty with age.*

If the bias of Fig 7(a) is taken to an extreme, we would have a step function where there would be zero uncertainty attached to the most recent evidence and maximum uncertainty attached to all other evidence. That would correspond to the models implicit in traditional browsing systems where the links available at any given object are based upon pair wise accessibility relationships to other objects in the space from the current object.

## 2.8 Paths and profiles

As described in the previous report, the intended application environment for the Ostensive Model is that of a browsing system where the user moves from node (ie information object) to node via links (ie accessibility relationships). These links will be generated dynamically based upon an ostensive definition of the information need. The definition will be made from the evidence collected along the path to that node. The path will be the sequence of nodes that was followed by the user.

The rapid/gentle bias mentioned above effectively means balancing the information relating to how the user got to a particular object (ie the path) against the information in the object itself.

There is clearly a range of ‘increasing with age’ profiles of evidential uncertainty. We cannot discern support from the Ostensive Model for any particular one - it tells only that the profile should be of that general form. It occurs to us that an appropriate bias reflects more the details of the process *e* of interpreting information than we are able either to measure or to understand. It seems a clear candidate for empirical determination.

Having determined empirically a generally suitable profile, and given that it becomes the basis of an IR system's efforts to support a user, it seems reasonable that it be the subject of optimisation during each information seeking session - ie a short-term learning approach be applied. For example, after each relevance indication by a user, the system could modify its uncertainty profile in such a way that it would have (more) correctly predicted the choice made by the user had it the chance again, given the same evidence. In this way the system could build completely tailored profiles reflecting, as best it could, the development of the users' information need.

### 2.9 The three elements of ostension

Quine (Quine, 53) proposes (almost as an aside) three components necessary for ostension to be used to capture spatio-temporal concepts (such as developing information needs). He first requires 'pure ostension' which equates to simple observed evidence, then he requires 'identification' which refers to the recognition of identity of the concepts being defined by the indivdual acts of pure ostension, and finally he requires 'induction' which is the process of combining the evidence.

The Ostensive Model has pure ostension in the form of objects indicated and not indicated as relevant. It has identification in the form of the assumptions inherent in the propositions P1, P2, and P3 - ie that all observed acts (within the restricted environments outlined above) are ostensive with respect to the information need. It has induction in the form of the uncertainty profiles placed across the evidence that guide its combination. The actual combination of evidence is not part of the model as described here. In the next section, we shall provide an example of integrating the Ostensive Model with a probabilistic retrieval model that will perform that final combination.

It is the induction through uncertainty profiles that sets the Ostensive Model apart from the conventional Relevance Feedback model. Instead of treating an information need (and hence relevance) as merely a spatial concept, the Ostensive Model regards it as spatio-temporal and performs its induction accordingly.

## 3 Applying and implementing the ostensive model

In this section we present the integration of the abstract Ostensive Model of information needs with a concrete retrieval model - ie the binary probabilistic model. Two key points facilitate this integration: Firstly, we describe a particular assumption implicit in that probabilistic model and show that the proposed method of integration actually adds knowledge and weakens the assumption. Secondly, we introduce the notion of ostensive relevance.

### 3.1 The conventional binary probabilistic model

Following the decision theoretic derivation given by van Rijsbergen (vanR, 79) we have for an information object $D_j$ and a set of desired features $x_1$ to $x_t$, a linear decision function relating the probabilities of observing relevance and of observing non-relevance given a particular object:

$$log \frac{P ( Rel \mid D_j )}{P ( NonRel \mid D_j )} = \sum_{i=1}^{t} w_i \cdot x_{ij} + \mathrm{C}$$

where

$$x_{ij} = \begin{cases} 0 \text{ ; absence} \\ 1 \text{ ; presence} \end{cases} \qquad w_i = log \frac{p_i ( 1 - q_i )}{q_i ( 1 - p_i )}$$

$C$ is constant for all objects with respect to a particular set of desired features. It therefore does not affect the resultant ordering of objects and is typically ignored in implementations.

There are the random variables:

$$p_i = P(x_i = 1 \mid Rel)$$

$$q_i = P(x_i = 1 \mid NonRel)$$

which are the probabilities of observing a particular feature given that we have observed relevance and non relevance respectively. The computation of the decision function can proceed once those probabilities have been provided. Traditionally, they have been obtained through estimates based upon counting within the set of objects indicated as relevant by the user:

$$P(x_i{=}1 \mid Rel) = \frac{P(x_i{=}1 \wedge Rel)}{P(Rel)}$$

$$\textit{Estimates:} \quad P(x_i{=}1 \wedge Rel) = \frac{r_i}{N} \quad , \quad P(Rel) = \frac{R}{N}$$

$$\textit{gives} \quad P(x_i{=}1 \mid Rel) = \frac{r_i}{R}$$

with respect to the standard contingency table:

| | *Relevant* | *Non-Relevant* | |
|---|---|---|---|
| $x_i{=}1$ | $r_i$ | $n_i - r_i$ | $n_i$ |
| $x_i{=}0$ | $R - r_i$ | $N - n_i - R + r_i$ | $N - n_i$ |
| | $R$ | $N - R$ | $N$ |

ie the probability of observing a feature given relevance is estimated as the proportion of the relevant-indicated objects that contain the feature. A similar process is followed for $q_i$ producing the proportion of objects *not* indicated as relevant (and hence assumed to be non relevant) that contain the feature:

$$P(x_i{=}1 \mid NonRel) = \frac{n_i - r_i}{N - R}$$

It is through an alternative, and we would argue more appropriate, estimation procedure for the conditional probability $p_i$ that the Ostensive Model will incorporate its uncertainty profiles.

### 3.2 Not all evidence is created equal

One can imagine a set of six objects indicated as relevant by the user (Fig 8), three of which contain a particular feature $x_i$ (ie $R{=}6$, $r_i{=}3$). Following the derivation of the conventional binary probabilistic model given above, the estimate of the probability of observing the feature given relevance will therefore be $p_i{=}0.5$

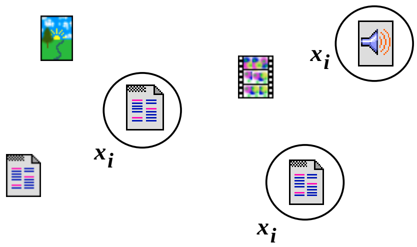


Figure 8: *A set of six information objects marked as relevant, three of which contain a particular feature.*

Implicit in the conventional approach is the assumption that all objects marked as relevant are equally useful and appropriate as sources of evidence for the estimation of $p_i$. There is no account taken of any property that could affect their individual appropriateness. Using the Ostensive Model, we can place a structure over the sources based upon their age.

If we take those same six objects and structure them according to their ages we have a *sequence* of relevance indications. Fig 9 might be the result of such an ordering.

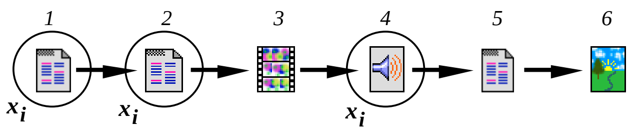


Figure 9: *A time-ordered sequence of six information objects marked as relevant, three of which contain a particular feature.*

This is a spatio-temporal record of information regarded as relevant. The Ostensive Model regards that record as an ostensive definition of the development of the information need. With this structure across the objects, we can associate degrees of uncertainty according to an 'increasing with age' profile as discussed above (Fig 10).

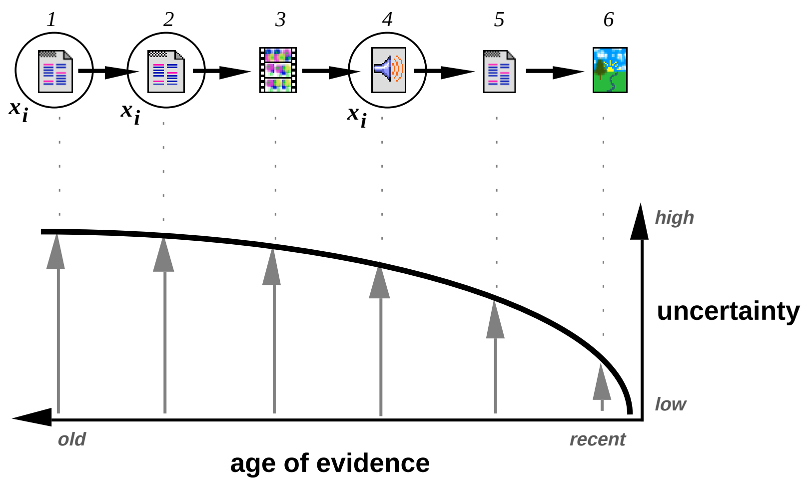


Figure 10: *Age-based degrees of uncertainty associated with six time-ordered pieces of evidence.*

From the diagram we see that the most recently indicated object has the lowest attached uncertainty, and therefore, we have the highest confidence in it being representative of the current information need.

In conventional systems, objects are indicated as relevant at various times and are not the only objects observed in detail by the user. In the Ostensive Model the assumption is that all documents observed are observed because the user regarded them as relevant. It is also assumed that they are observed immediately after each other (as a result of the restricted environment within which the information seeking takes place). The arrows in Figs 9 & 10 emphasise the resulting path-like nature of the sequence of objects.

## 3.3 Ostensive Relevance

In traditional approaches to IR the notion of relevance has had many and varied definitions. Common to these definitions is the necessary assumption that there exists a static and absolute relevance. There has been recognition made of the dynamic nature of information needs, and a case made for the adoption of it and an associated dynamic notion of relevance as the norm (Bates, 89). Nevertheless, to our knowledge, such work has not produced models that place a structure across the temporal stages of the information need or of relevance. As a result, applicable models or models that provide clear guidance for implementation have remained illusive.

In contrast, the Ostensive Model assumes, in fact requires, a dynamic notion of relevance. Secondly, relevance is something that can be tied down for only an instant - ie it will change with time. Further, it cannot be determined in advance - its nature will only become apparent at the instant that it is formed in the brain of the person making the relevance assessment. Thirdly, the model recognises the inaccessibility of relevance - ie the inability to directly observe the knowledge state of the user and the actual perception of relevance within it. It can only be observed 'from a distance' through ostensive observable evidence.

With these thoughts in mind, we define a new notion of relevance from the point of view of an external observer (eg an IR system).

> ***Proposition 5*** *The Ostensive Relevance of an information object is the degree to which evidence from the object is representative/indicative of the current information need.*

Ostensive Relevance, in effect, captures how 'ostensive' a piece of observed evidence is. This notion is clearly only valid when evidence is observed from objects that the user selects under the conditions set out in the abstract description of the model (ie under propositions P 1 to P5).

We can express such confidence in terms of probability. If we take a binary view of this notion of relevance (ie 'representative' or 'not-representative') - we can talk of the probability of observing that relevance, ie $P(Rel)$.

That probability of relevance can be measured at each information object $D_j$ and therefore expressed as a conditional probability - $P(Rel \mid D_j)$. The value of $P(Rel \mid D_j)$ is inversely related to the ostensive uncertainty at $D_j$. This produces an opinionated function of $P(Rel \mid D_j)$. Fig 11 shows the relationship between evidential uncertainty and the object-conditional form of Ostensive Relevance as prescribed by the Ostensive Model.


Figure 11: *The inverse relationship between uncertainty and the object-conditional form of Ostensive Relevance.*

The total probability function relates the probability of Ostensive Relevance to the individual object-conditional probabilities over the structure of relevant-indicated information objects:

$$P(Rel) = \sum_{j=1}^{R} P(Rel \mid D_j) \cdot P(D_j)$$

## 3.4 Incorporating the ostensive evidence

The interpretation of the evidential uncertainty profile as an object-conditional probability function of Ostensive Relevance provides the key to incorporating the ideas of the Ostensive Model into the conventional Binary Probabilistic Model.

A new estimation procedure is developed for the random variable $p_i$:

$$p_i = P(x_i=1 \mid Rel) = \frac{P(x_i=1 \wedge Rel)}{P(Rel)}$$

Applying the joint probability function allows the above probabilities to be expressed in terms of their component probabilities at each of the $R$ relevant-indicated objects:

$$= \frac{\sum_{j=1}^{R} P(x_i=1 \wedge Rel \mid D_j) \bullet P(D_j)}{\sum_{j=1}^{R} P(Rel \mid D_j) \bullet P(D_j)}$$

Simplifying the notation through the definition of new probability functions at $D_j$, and cancelling out the $P(D_j)$ terms (as they are assumed to be constant for all $D_j$):

$$= \frac{\sum_{j=1}^{R} P_{D_j}(x_i=1 \wedge Rel)}{\sum_{j=1}^{R} P_{D_j}(Rel)}$$

Assuming the conditional independence of $x_i$ and $Rel$ with respect to $D_j$:

$$= \frac{\sum_{j=1}^{R} P_{D_j}(x_i=1) \bullet P_{D_j}(Rel)}{\sum_{j=1}^{R} P_{D_j}(Rel)}$$

This is a major assumption, particularly when one considers that it is the very observation of a feature that is used to predict the ultimate observation of Relevance. At this stage in the Ostensive Model's development, this simplifying assumption allows the integration of the ostensive and probabilistic models to go through. Nevertheless, this possibly sub-optimal integration has intuitively attractive properties that will be outlined in the next section.

The probability of observing $x_i$ given the observation of $D_j$, is trivially observable as equal to the value of $x_i$ at that object.

$$= \frac{\sum_{j=1}^{R} x_{ij} \bullet P_{D_j}(Rel)}{\sum_{j=1}^{R} P_{D_j}(Rel)}$$

This leaves the individual object-conditional probabilities of observing Ostensive Relevance to be inserted. Therefore, for each relevant-indicated information object $D_j$ we need only substitute the value of $x_{ij}$ at that object and the value of $P(Rel \mid D_j)$ from the uncertainty profile of the Ostensive Model.

In effect, the numerator works by 'switching on' the $P(Rel \mid D_j)$ component for each object containing the feature $x_i$, with the size of that component being determined by the Ostensive Model's opinion of the uncertainty attached to the source of that evidence. The denominator normalises those components into a *0..1* range. Thus, each $x_i$-containing relevant-indicated object contributes an amount towards the estimation of $p_i$ depending upon its probability of Ostensive Relevance.

# 4 An intuitive description and example

Here we show an example of how the incorporation of the uncertainty profiles might affect the final estimations given to the variable $p_i$. To do this we create an uncertainty profile and compare its effect when applied to two different sequences of relevant-indicated objects.

Let us adopt an uncertainty profile and resulting probability function of the form outlined in Figs 10 & 11. The probability function could be something like $2^{-a}$ where $a$ would be the age of the evidence (eg the number of steps since the object was indicated as relevant). This might give a sequence of six object-conditional probabilities such as *{1/64, 1/32, 1/16, 1/8, 1/4, 1/2}*. The actual values are not important for this example, only that they follow the general profile required by the Ostensive Model.

Let us adopt the object sequence of Fig 9 as the first test sequence and invert the presence/absence of features to produce the second test sequence. These sequences are shown in Fig 12.

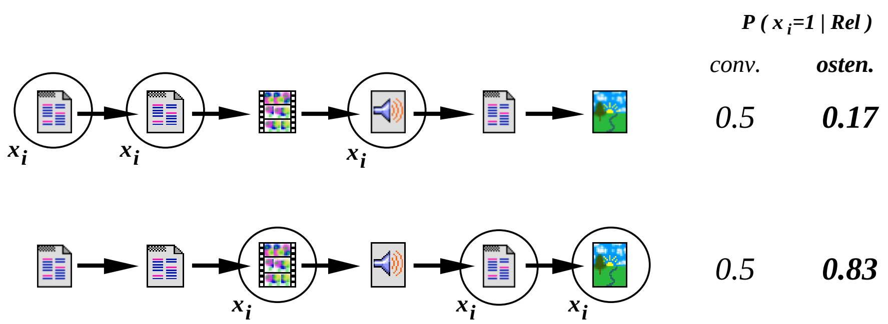


Figure 12: *Two sequences and their respective $p_i$ values from the conventional and from the ostensive estimation procedures.*

The conventional procedure would, as already shown above, produce an estimate for $P(x_i \mid Rel)$ of *0.5* for the first sequence. The Ostensive Model's estimation would be *0.17.*

For the second sequence, the conventional procedure would, again, produce an estimate of *0.5*. The Ostensive Model would produce *0.83*. These values are shown in Fig 12, alongside the sequences for comparison.

The conventional approach sees no difference between the two sequences. In fact it does not see the sequence, only an unordered set. It simply identifies half of the objects as having the feature and the other half as not having it, and so estimates the probability of observing the feature in any future set of relevant objects as *0.5* for both sequences.

Looking at the first sequence, we would suggest that the occurrence of the feature is more evident at the beginning, less evident in the middle and not evident at the end. Conversely, in the second sequence we would suggest that the feature is not evident at the start, but becomes more evident towards the end. We would therefore argue that intuitively one would expect to see the feature in the next, as yet unseen, object more highly in the second sequence than in the first. Further, one could argue that our expectation after the first sequence would be really quite low, and after the second sequence really quite high.

If one accepts such intuitions, then we believe the probabilities produced by the Ostensive Model appear more appropriate.

# 5 Summary

We have described a new model of information needs that recognises their developmental nature, identifies the major force driving that development, and accepts the inaccessibility of information needs.

We emphasised the transient and inaccessible nature of information needs as spatio-temporal entities and used this to motivate a structured view of the uncertainty attached to ostensive evidence resulting from their development. We outlined the conditions under which simple ostensive evidence might be reliably collected. We stated that a particular class of uncertainty structures was the most appropriate, that the actual form may require empirical determination, and further that it may then be the subject of short-term learning during an information seeking session

We proposed a new notion of relevance compatible with the model and showed how it provides the key to integrating the Ostensive Model with the conventional Binary Probabilistic Model. At a pragmatic level, the integration achieves an implementable retrieval function. At a theoretic level, the integration exposed an implicit assumption made in the conventional technique for making a particular necessary probability estimation (ie that of $p_i$). That technique was replaced with a more sophisticated procedure that incorporated the structured view of temporal evidential uncertainty.

The areas attracting our current attention are those of the assumption of object-conditional independence of a feature's occurrence and Ostensive Relevance, and the determination and optimisation of a suitable 'increasing with age' uncertainty profile.

We concluded by giving an example of the application of the Ostensive Model to two notional sets of relevant-indicated objects and suggested that, compared to the conventional approach, the Ostensive Model demonstrated itself to be an intuitively more appropriate model.

## 6 Acknowledgements

This work has been supported by the ESPRIT Basic Research Action "FERMI - Formalisation and Experimentation on the Retrieval of Multimedia Information" (Fermi, 94).

## 7 References

Marcia J. Bates. The design of browsing and berrypicking techniques for the online search interface. *Online Review*, 13(5):407--423, 1989.

Iain Campbell. Supporting information needs by ostensive definition in an adaptive information space. In Ian Ruthven, editor, *MIRO '95*, Workshops in Computing. Springer Verlag, 1995.

*The FERMI project*. ESPRIT Basic Research Action 81034. Documentation and technical reports Web server. "http://www.dcs.gla.ac.uk/fermi".

Peter Ingwersen. A cognitive view of three selected online search facilities. *Online Review*, 8(5):465--492, 1984.

Donald M. MacKay. *Information mechanisim and meaning*. M.I.T. Press, 1969.

Marc de Mey. The relevance of the cognitive paradigm for information science. In Ole Harbo, editor, *Theory and application of Information Research*, pages 48-61. Mansell, 1980.

Willard van Orman Quine. *From a logical point of view*. Harvard University Press, 1953.

Willard van Orman Quine. *Ontological relativity and other essays*. Columbia Unversity Press, 1969.

*Oxford English Dictionary* (CD-ROM Edition). Oxford University Press, 1993.

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