Resolution Spaces: A Topological Approach to Similarity

Konstantinos Georgatos
Mathematics Department
John Jay College
City University of New York
445 West 59th Street
New York, New York 10019
e-mail: kgeorgatos@jjay.cuny.edu

Abstract

We argue that in order to reason with similarity we need to model the concept of discriminating power. We offer a simple topological notion called resolution space that provides a rich mathematical framework for reasoning with limited discriminating power avoiding the vagueness paradox.

1 Introduction

A central concept for Information Retrieval ([1, 2]) is that of similarity. Although an Information Retrieval system is expected to return a set of documents most relevant to the query word(s), it is often described as returning a set of documents most similar to the query. In particular, looking into the inner mechanism of the underlying model on which an Information Retrieval system is based, one often finds that relevance is based on (or derives from) a notion of similarity. This is true for the most significant Information Retrieval models such as the vector, probabilistic ([3]), and various logical models (e.g. boolean ([4]), non-classical ([1, 5]), information-theoretic ([6]), etc).

In the vector and probabilistic model, documents are judged similar according to various functions that depend on indexing terms and their frequency. In logical models, either a rank is presupposed ([5]) or is computed according to measures such as expected utility ([6]) or information content that in turn induce a measure of similarity. Recent approaches involve hyperlinks to judge similarity ([7, 8]). It is striking that there no single criterion for judging relevance and most probably more ingenious suggestions will emerge in the future.

This paper proposes a new approach to similarity using the topological notion of resolution space. Resolution spaces have been introduced in for modeling vague statements ([9]). The reason that standard topological methods are inappropriate to model similarity is that similarity in Topology is expressed with the notion of metric distance. Metrics however require complete separation between objects and quantitative data that often are unavailable. Therefore what we need is a general framework that can classify entities as similar or not similar (and eventually rank them) under any possible criterion. This is what we offer.

Our approach is based on the following thesis: two objects are similar when they are indistinguishable under an appropriate resolution. For example consider two documents containing the same text but formatted in a different way. In most approaches in Information Retrieval formatting is discounted and therefore those two documents are considered identical. This is an example of lowering our resolution in order to get rid of unnecessary detail. Further, suppose that we use an indexing mechanism. Two documents will be considered identical in case they contain the same indexing terms even if their text is different. So,

\[ \text{similarity} = \text{indistinguishability (under an appropriate resolution)}. \]

We formalize the notion of power of discrimination with the help of resolution spaces. We then show how one can define propositions on a resolution space. The algebra of the propositions is not boolean but as we will show forms an ortholattice.

The main thrust of this paper is a general need for modelling commonsense reasoning problems that have arisen in AI and, particularly, in the area of nonmonotonic logic as we expressed in [10] and [11]. We are convinced that an effective way of handling vagueness is central and necessary to Information Retrieval.
2 The Color Patches Problem Revisited

Perhaps the best way to illustrate the issues involved with similarity is the well-known red patches test. Suppose I have a large enough series of patches $p_1, p_2, \ldots, p_n$ changing gradually their color from red to white where it is impossible to distinguish a patch $p_i$ from its successor $p_{i+1}$. Where is the last red patch? If we choose, say $p_k$, as the last red patch then we run into troubles since $p_{k+1}$ seems to us of exactly the same color.

We will reformulate the above test in order to motivate the theory we will develop in the following section. Our version of the above test is the following: let $p_1, p_2, \ldots, p_n$ be any sequence of color patches. Now, we are handed a patch with some color called witt. Then, we are asked to identify the witt patches. This is our one (and, in our opinion, the only) way to go. We pick patches and compare them with the prototype witt we were given. Those we can distinguish from it form a set $F$ and those that do not will form another set $P$. The set $F$ is the set of those patches that are definitely not witt. The set $P$ is the set of those patches that could be witt as they are indistinguishable from the prototype but we have no way to find out. The sets $F$ and $P$ complement each other. We managed to create two well defined sets with no borderline cases. Those sets have been created through relative distinguishability. Relative distinguishability has a prominent role in philosophical approaches to vagueness (see [12, 13]). As pointed out in [14] relative distinguishability does not refine direct distinguishability. That is, two patches that belong to $F$ and $P$, respectively, can still be identical. However, we do know that the witt patch and the patches in $F$ are distinct. The formation of $F$ and $P$ resembles that of the recursive enumerable predicates. A predicate is r.e. just in case its extension function is r.e. In our case the discriminating function converges on all inputs but it can be partially (on the positive side) wrong!

The following section provides a further generalization of the above approach to the color patches problem.

3 The Two Values of Discrimination

In this section, we will elaborate on the terms of our language, called objects and the methods of distinguishing among them. How do we distinguish among objects? By a comparison test. For instance, $x$ and $y$ are indistinguishable because they look different. On the other hand $x$ and $x$ (as printed characters) are indistinguishable because they look the same even though they are not. A magnifying glass can certainly detect a difference among them.

Observing is not the only way of distinguishing. Measuring is another. Measuring means to compare an object with a prototype. For example, if Ann says “I am two meters tall” means that if we join two meters (meter as the physical object) the height of the resulting object will be indistinguishable from Ann’s height. This is the reason for the inherent value of prototypes. They become widely accepted points of reference and objects of comparison. During measurement we must also provide a method for doing so. When Ann says “I am two meters tall”, she means that we should also set the one edge of the two meter object on her feet and the other towards her head. Moreover, by edge she means the “intended” edges of the two-meter object’s length than the edges of its width. Although such explanations are usually cumbersome and implicit they are not always so. Height measurement is by now standard and everyone has been trained to decide distinguishability according to it. In contrast, an experiment is a measurement but we take great pains in reproducing it. Reproducing it means we know the method of the comparison we try to make rather than revealing some truth about the objects compared.

Intuitively, a resolution space is a set of objects representing the objects of our discourse. Objects can refer to entities, relations and functions among them, situations, etc. There is no preset bounds on what can be expressed by objects. If two objects are distinguishable then that means we have a method for telling them apart. Otherwise they are indistinguishable.

Definition 1 A resolution space is a pair

$$(R, \dagger)$$

where $R$ is a set of objects, $\dagger$ is a binary relation between members of $R$ called the distinguishability relation. We shall assume the following properties for $\dagger$:

1. $x \nott x$, for all $x \in R$ (Distinguishability Irreflexivity)
2. $x \dagger y$ implies $y \dagger x$ (Symmetry)

The complement of the distinguishability relation will be called indistinguishability and denoted by $\sim$.

Note here that the indistinguishability relation is reflexive and symmetric. Our idea of using a reflexive and symmetric relation in order to express indistinguishability is not new. Although many authors have argued that indistinguishability is better expressed through equivalence ([15, 16, 17]), many have also dropped transitivity as early as [18] (see also [19]).

Example 1

1. Let $D$ be the set of finite binary strings of finite length $n$. We can say that two strings are indistinguishable when they have the same length and differ in at most one digit. Otherwise, they are distinguishable.
2. Let $R$ be the set of real numbers. Let $r_1 \equiv r_2$ when
ty they differ on the integer part of their decimal represen-
tation.

Can a distinguishability relation be expressed in terms of a metric? It seems that a distinguishability relation is more basic than a metric. It seems plausible to think of metric as a way to generate distinguishability, for example, let $x \equiv y$ if and only if $d(x, y) \geq \varepsilon$ for some appropriate metric $d$. This concept has been long exploited in topology by the notion of uniformity (see for example [20]): a filter of reflexive and symmetric relations with an additional axiom expressing the triangle inequality and therefore the real-valued nature of the metric. Apart from defining a res-
olution space on the basis of a single distinguishability rela-
tion rather than a set of them, we do not impose a triangle
inquality as we aim for a qualitative rather than a quanti-
tative framework. Therefore, set of resolution spaces can
be used to present a generalization of uniformities.

As we mentioned above we are interested in relative distin-
guishability. This is so because we would like to express valuations. In particular, we would like to identify values
within the distinguishability domain. Our intention is to
build a richer structure that will reflect those intuitions. To
this end we will define the following operator

**Definition 2** Let $X \subseteq D$. The discriminant operator $\diamant : D \rightarrow D$ is defined by

$$X^\diamant = \{ y \mid \forall x \in X, x \equiv y \}.$$ 

The set $X^\diamant$ will be called the discriminant ball of $X$.

We have the following

**Proposition 2**

1. $X \subseteq Y$ implies $Y^\diamant \subseteq X^\diamant$ (that is the discriminant operator is antitone).

2. $X \subseteq X^{\diamant}$, $X^\diamant = X^{\diamant\diamant}$

3. $\diamant : D \rightarrow D$ is a closure operator.

In fact, it is well known that an antitone operator on a
lattice induces a pair of Galois connection (see [21]) (here
the lattice is the lattice of the powerset of $D$). Closure
operators in this case can be generated as a composition of the maps that form the Galois connection.

The closed subsets of the closure operator will be called
stable. Stable subsets have the form $X^{\diamant}$.

The proposition below follows from a more general result
about the closed subsets of a closure operator (see [22])

**Proposition 3** The stable subsets of $D$ form a complete
lattice under $\subseteq$. If $\{A_i\}_{i \in I}$ is a family of stable sets then

$$\bigwedge_{i \in I} A_i = \bigcap_{i \in I} A_i \quad \bigvee_{i \in I} A_i = \left( \bigcup_{i \in I} A_i \right)^{\diamant}.$$ 

Apart from the conjunction and disjunction operation on
the lattice of stable subsets, we can add an orthonega-
tion unary operation that is simply the discriminant operator.
Therefore, the negation of the stable subset will be $A^{\diamant}$.

By the result above $A^{\diamant}$ is stable because $A$ is stable. The discriminant operator restricted in the stable subsets has
power two, that is, $A = A^{\diamant\diamant}$. In addition $a \cup a^{\diamant} = T$ and $a \land a^{\diamant} = \bot$. The above properties make the lattice of
stable subsets an ortholattice (see [22]).

Our thesis is that the well-defined propositions describ-

ign properties of objects in a resolution space correspond
to stable subsets. The reason is simple. If one allows a non-
stable subset to be an extension of a proposition $A$ then its
complement will contain a object $x$ that is indistinguish-
able from some object contained in the extension of $A$. The
object $x$ can also be distinguished from all objects distin-
guished from $A$ (in symbols, $x \in A^{\diamant\diamant} - A$ with $x \in A^\bot$).

Therefore, there is no apparent reason for excluding $x$ from
satisfying the property $a$. In addition, stability allow us to
handle affirmative and refutative assertions under a single
notion ([23]).

If a subset is stable then its negation is not necessarily
its complement. Therefore, there are objects that satisfy
neither the proposition nor its negation ($x \notin A \cup A^\bot$). There is a good reason for that, namely, such ambiguous
object is indistinguishable from every object satisfying $A$ as
well as indistinguishable from every object not satisfying
$A$ ($x \in A^\bot \cup (A^\bot)^\bot$).

Therefore, we suggest that the algebra of propositions on
a resolution space forms an ortholattice. The equational
theory of ortholattices appears in Table 1.

The above show

**Theorem 4** (Soundness) The set of propositions on a res-
olution space forms an ortholattice.

4 Completeness

We shall now give a representation of the ortholattice of
unambiguous propositions by providing a corresponding
resolution space whose lattice of stable subsets contains
the ortholattice. Our representation runs along the lines
of Goldblatt’s representation of ortholattice with its filters
([24]). Our contribution lies in turning the space of filters
into a resolution space of objects.

To this end, let $T$ be an ortholattice. Consider the maps $v$ from $T$ to $\{0, 1\}$ that preserve the meets of $T$ that
is $v(t \land s) = v(t) \land v(s)$. We will call these maps semi-
lattice homomorphisms. We shall move freely between
meet-homomorphisms and ortholattice filters\footnote{A filter $F$ is a proper upper-closed subset closed under meets, i.e.,
$0 \notin F$, if $t \in F$ and $t \leq s$ then $s \in F$ and if $t, s \in F$ then $t \land s \in F$.} because fil-
tters arise as inverse images of 1. To see that \( v^{-1}(1) \) be the inverse image of 1 under the semilattice homomorphism \( v \). If \( t, s \in v^{-1}(1) \) then \( v(t) = v(s) = 1 \) so \( v(t \land s) = v(t) \land v(s) = 1 \land 1 = 1 \), therefore, \( t \land s \in v^{-1}(1) \). Also, if \( t \in v^{-1}(1) \) then \( 1 = v(t) \leq v(s) \) so \( v(s) = 1 \). On the other hand, suppose that \( F \) is a filter of \( T \) and \( v \) is a map from \( T \) to \( \{0, 1\} \) such that \( v(t) = 1 \) if and only if \( t \in F \). The only case needs to be considered is \( v(t) = v(s) = 1 \) but \( v(t \land s) \neq 0 \). However, this case is impossible as \( t \land s \in F \) because \( F \) is a filter.

Now that we established the above bijective correspondence, we shall define the desired resolution space. Let \( (R^T, \vdash) \) be the resolution space whose objects are the filters of the ortholattice \( T \). By the result above, each object can be considered a semilattice homomorphism. It only remains to define the distinguishability relation: let \( a, b \in R \) then \( a \vdash b \) if there exists \( t \in T \) such that \( a(t) = 1 \) and \( b(t) = 1 \). We will prove that \( \vdash \) is indeed a distinguishability relation, i.e. \( \vdash \) is irreflexive and symmetric. Irreflexivity follows from the fact that if \( a \vdash a \) then there is \( t \in T \) such that \( a(t) = a(t^\dagger) = 1 \). It follows that \( 0 = a(\neg t) = a(t \land t^\dagger) = a(t) \land a(t^\dagger) = 1 \land 1 = 1 \), a contradiction. Symmetry is immediate since \( t \vdash t^\dagger = t \).

Let \( S \) be the lattice of the stable subsets of \( R^T \) under \( \subseteq \). It remains to show that \( T \) can be embedded in \( S \): let \( i \) be the map from \( T \) to \( S \) defined by \( i(t) = A_t \), where \( A_t \) is the set of semilattice homomorphisms that map \( t \) to 1, i.e. if \( a \in A_t \) then \( a(t) = 1 \).

Observe that \( A_t^\dagger = A_{t^\dagger} \): for the left to right inclusion, if \( a \in A_t^\dagger \) then \( a \vdash b \) for all \( b \in A_t \). In particular, \( a \vdash \tilde{t} \), where \( \tilde{t} \) is the semilattice homomorphism such that \( \tilde{t}(s) = 1 \) if and only if \( t \leq s \). So there is \( s \in T \) such that \( \tilde{t}(s) = 1 \) while \( a(s^\dagger) = 1 \). The former implies that \( t \leq s \) so \( s^\dagger \leq t^\dagger \) and the latter implies \( a(t^\dagger) = 1 \). It follows that \( a \in A_{t^\dagger} \). The other inclusion is straightforward.

The above implies that \( A_t^{\dagger\dagger} = A_{t^\dagger\dagger} = A_t \) and therefore the map \( i \) is well-defined. We will show that this map is an isomorphism.

First, the map \( i \) is an ortholattice homomorphism. We have that \( i(t \land s) = A_t \land A_s = i(t) \land i(s) \) because \( a(t \land s) = 1 \) if and only if \( a(t) = 1 \) and \( a(s) = 1 \) for all semilattice homomorphisms \( a \). Also, \( i(t) \land i(s) = i(t) \lor i(s) \), as we have \( A_t \land A_s = A_{(t \lor s)^\dagger} = A_t^\dagger \land A_s^\dagger = (A_t^\dagger \land A_s^\dagger)^\dagger = A_t \lor A_s \).

Second, it is clear that the map \( i \) is a bijection proving the following.

**Theorem 5** (Completeness) For every ortholattice\( T \) there is a resolution space \( R_T \) such that \( T \) can embedded in the complete ortholattice of stable subsets of \( R_T \).

5 Conclusion

The notion of a resolution space provides a useful mathematical tool for dealing with similarity. Its main advantage is simplicity but at the same time it provides an expressive mathematical framework.

So far, our approach has been biased. We have relied solely on the distinguishability operator \( \vdash \). The operator \( \sim \) provides a considerable alternative. Define the similarity ball of \( X \) with

\[
X^\sim = \{y | \forall x \in X, x \sim y \}.
\]

It can be easily seen that \( x^\dagger \) and \( x^\sim \) are complements. However, the discriminant and similarity ball of \( X \) are not necessarily complements when \( X \) is not a singleton. An ortholattice of stable subsets under the similarity operator can be defined in a dual way we defined it above. This observation along with a space that can accommodate more...
than one distinguishability operator provides what we believe is the most promising research direction of resolution spaces.

References


