

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 meth-

ods 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,

similarity = 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, \dots, 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, \dots, 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 distinguishable 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

$$(R, \dashv)$$

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

1. $x \not\sim x$, for all $x \in R$ (Distinguishability Irreflexivity)
2. $x \dashv y$ implies $y \dashv x$ (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.

- Let R be the set of real numbers. Let $r_1 \dashv r_2$ when they differ on the integer part of their decimal representation.

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 $x \dashv y$ if and only if $d(x, y) \geq \epsilon$ 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 resolution space on the basis of a single distinguishability relation rather than a set of them, we do not impose a triangle inequality as we aim for a qualitative rather than a quantitative framework. Therefore, set of resolution spaces can be used to present a generalization of uniformities.

As we mentioned above we are interested in relative distinguishability. 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 $\dashv : D \rightarrow D$ is defined by

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

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

We have the following

Proposition 2 1. $X \subseteq Y$ implies $Y^\dashv \subseteq X^\dashv$ (that is the discriminant operator is antitone).

- $X \subseteq X^{\dashv\dashv}, X^\dashv = X^{\dashv\dashv}$

- $\dashv : D \rightarrow D$ 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 $X^{\dashv\dashv}$.

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

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

Apart from the conjunction and disjunction operation on the lattice of stable subsets, we can add an orthonegation unary operation that is simply the discriminant operator. Therefore, the negation of the stable subset will be A^\dashv . By the result above A^\dashv is stable because A is stable. The discriminant operator restricted in the stable subsets has power two, that is, $A = A^{\dashv\dashv}$. In addition $a \vee a^\dashv = \top$ and $a \wedge a^\dashv = \perp$. The above properties make the lattice of stable subsets an *ortholattice* (see [22]).

Our thesis is that the well-defined propositions describing 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 indistinguishable from some object contained in the extension of A . The object x can also be distinguished from all objects distinguished from A (in symbols, $x \in A^{\dashv\dashv} - A$ with $x \in A^\sim$). 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 ($x \notin A \cup A^\dashv$). 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 ($x \in A^\sim \cup (A^\dashv)^\sim$).

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 resolution 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 $v(t \wedge s) = v(t) \wedge v(s)$. We will call these maps *semi-lattice homomorphisms*. We shall move freely between meet-homomorphisms and ortholattice filters¹ because fil-

¹A filter F is a proper upper-closed subset closed under meets, i.e., $0 \notin F$, if $t \in F$ and $t \geq s$ then $s \in F$ and if $t, s \in F$ then $t \wedge s \in F$.

Table 1: Ortholattice Theory

$x \wedge x = x$	$x \vee x = x$	(Idempotency)
$x \wedge y = y \wedge x$	$x \vee y = y \vee x$	(Commutativity)
$x \wedge (y \wedge z) = (x \wedge y) \wedge z$	$x \vee (y \vee z) = (x \vee y) \vee z$	(Associativity)
$x \wedge (x \vee y) = x$	$x \vee (x \wedge y) = x$	(Absorption)
$x \wedge x^{-1} = \perp$	$x \vee x^{-1} = \top$	
$(x \wedge y)^{-1} = x^{-1} \vee y^{-1}$	$(x \vee y)^{-1} = x^{-1} \wedge y^{-1}$	(de Morgan)
$(x^{-1})^{-1} = x$		(Involution)
$x \vdash y$ is equivalent to either $x \wedge y = x$ or $x \vee y = y$		

ters arise as inverse images of 1. To see that let $v^{-1}(1)$ 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 $v(t \wedge s) = v(t) \wedge v(s) = 1 \wedge 1 = 1$, therefore, $t \wedge s \in v^{-1}(1)$. 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 $v(t \wedge s) \neq 1$. However, this case is impossible as $t \wedge s \in F$ because F is a filter.

Now that we established the above bijective correspondence, we shall define the desired resolution space. Let (R^T, \dashv) 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 $a \dashv b$ iff there exists $t \in T$ such that $a(t) = 1$ and $b(t^{-1}) = 1$. We will prove that \dashv is indeed a distinguishability relation, i.e. \dashv is irreflexive and symmetric. Irreflexivity follows from the fact that if $a \dashv a$ then there is $t \in T$ such that $a(t) = a(t^{-1}) = 1$. It follows that $0 = a(\dashv) = a(t \wedge t^{-1}) = a(t) \wedge a(t^{-1}) = 1 \wedge 1 = 1$, a contradiction. Symmetry is immediate since $t^{-1} = t$.

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

The above implies that $A_t^{-1} = A_{t^{-1}} = A_t$ 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 $i(t \wedge s) = A_{t \wedge s} = A_t \cap A_s = i(t) \wedge i(s)$ because $a(t \wedge s) = 1$ if and only if $a(t) = 1$ and $a(s) = 1$ for all semilattice homomorphisms a . Also, $i(t \vee s) = i(t) \vee i(s)$, as we have $A_{t \vee s} = A_{(t^{-1} \wedge s^{-1})^{-1}} = A_{(t^{-1} \wedge s^{-1})}^{-1} = (A_{t^{-1}} \wedge A_{s^{-1}})^{-1} = (A_t^{-1} \wedge A_s^{-1})^{-1} = A_t \vee A_s$.

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 be 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 \dashv . 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^{-1} 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

- [1] C. J. van Rijsbergen, *Information Retrieval*. Sydney: Butterworths, 1979.
- [2] G. Salton, *Introduction to Modern Information Retrieval*. New York: McGraw-Hill, 1983.
- [3] G. Amati and C. J. van Rijsbergen, "Probability, Information and Information Retrieval," in *Working Notes of the Workshop on the treatment of Uncertainty in Logic-based Models of Information Retrieval Systems* (M. Lalmas, ed.), Department of Computing Science of the University of Glasgow, Glasgow University, September 1995.
- [4] G. Salton, E. A. Fox, and H. Wu, "Extended boolean information retrieval," *Communications of the ACM*, vol. 26, pp. 1022–1036, 1983.
- [5] G. Amati and K. Georgatos, "Relevance as Deduction: A Logical View of Information Retrieval," in *2nd International Workshop on Information Retrieval, Uncertainty and Logics*, (Glasgow, UK), pp. 21–27, July 1996.
- [6] G. Amati and K. van Rijsbergen, "Semantic Information Retrieval," in *Information Retrieval: Uncertainty and Logics* (F. Crestani, M. Lalmas, and C. J. van Rijsbergen, eds.), Information Retrieval, pp. 189–219, Boston: Kluwer Academic Publishers, 1998.
- [7] J. Carriere and R. Kazman, "Webquery: Searching and visualizing the web through connectivity," in *Proceedings of the Sixth International World Wide Web Conference [WWW6]*, pp. 701–711, 1997.
- [8] S. Chakrabarti, B. Dom, D. Gibson, S. Kumar, P. Raghavan, S. Rajagopalan, and A. Tomkins, "Experiments in topic distillation," in *ACM-SIGIR'98 Post-Conference Workshop on Hypertext Information Retrieval for the Web*, 1998.
- [9] K. Georgatos, "Resolution spaces and their logic," in *Proceedings of Workshop on Semantic Approximation, Granularity, and Vagueness*, 2000.
- [10] K. Georgatos, "Entrenchment relations: A uniform approach to nonmonotonic inference," in *Proceedings of the International Joint Conference on Qualitative and Quantitative Practical Reasoning (ES-CQARU/FAPR 97)*, no. 1244 in Lecture Notes in Computer Science, (Berlin), pp. 282–297, Springer-Verlag, 1997.
- [11] K. Georgatos, "To preference via entrenchment," *Annals of Pure and Applied Logic*, vol. 96, no. 1–3, pp. 141–155, 1999.
- [12] C. Peacocke, "Are vague predicates incoherent?," *Synthese*, vol. 46, pp. 121–141, 1981.
- [13] L. Linsky, "Phenomenal qualities and the identity of indistinguishables," *Synthese*, vol. 59, pp. 363–380, 1984.
- [14] T. Williamson, *Vagueness*. Routledge, 1996.
- [15] R. Aumann, "Agreeing to disagree," *The Annals of Statistics*, vol. 4, pp. 1236–9, 1976.
- [16] J. Hintikka, *Knowledge and Belief*. Ithaca, New York: Cornell University Press, 1962.
- [17] R. Fagin, J. Y. Halpern, and M. Y. Vardi, "A model-theoretic analysis of knowledge," *Journal of the Association for Computing Machinery*, vol. 38, no. 2, pp. 382–428, 1991.
- [18] H. Poincaré, *La Valeur de la Science*. Paris: Flammarion, 1905.
- [19] T. Williamson, *Identity and Discrimination*. Oxford University Press, 1990.
- [20] J. L. Kelley, *General Topology*. Princeton: Van Nostrand, 1955 (reprinted 1975).
- [21] B. A. Davey and H. A. Priestley, *Introduction to Lattices and Orders*. Cambridge University Press, 1990.
- [22] G. Birkhoff, *Lattice Theory*. No. 25 in Colloquium Publications, American Mathematical Society, 1948, reprinted 1995.
- [23] S. Vickers, *Topology via Logic*. Cambridge Studies in Advanced Computer Science, Cambridge: Cambridge University Press, 1989.
- [24] R. I. Goldblatt, "The stone space of an ortholattice," *Bulletin of London Mathematical Society*, vol. 7, pp. 45–48, 1975.