8-2014

Towards a Formal Description of Understandability (Causality, Pre-Requisites): From Prosorov's Phonocentric Topology to More General Interior (Closure) Structures

Olga Kosheleva
University of Texas at El Paso, olgak@utep.edu

Vladik Kreinovich
University of Texas at El Paso, vladik@utep.edu

Comments:
Technical Report: UTEP-CS-14-56
Published in *Mathematical Structures and Modeling*, 2014, Vol. 31, pp. 18-26.

Towards a Formal Description of Understandability (Causality, Pre-Requisites): From Prosorov's Phonocentric Topology to More General Interior (Closure) Structures

Olga Kosheleva and Vladik Kreinovich
University of Texas at El Paso
500 W. University
El Paso, TX 79968, USA
olgak@utep.edu, vladik@utep.edu

Abstract

In many real life situations, a text consists of related parts; so, to understand a part, we need to first understand some (or all) preceding parts: e.g., to understand Chapter 3, we first need to understand Chapters 1 and 2. In many cases, this dependence is described by a partial order. For this case, O. Prosorov proposed a natural description of the dependence structure as a topology (satisfying the separation axiom T_0).

In some practical situations, dependence is more general than partial order: e.g., to understand Chapter 3, we may need to understand either Chapter 1 or Chapter 2, but it is not necessary to understand both. We show that such a general dependence can be naturally described by a known generalization of topology: the notion of an interior (or, equivalently, closure) structure (provided, of course, that this structure satisfies a natural analog of T_0-separability).

1 Prosorov's Topology-Based Description of Understandability: Reminder

Understandability: formulation of the general problem. A sufficiently long text usually consists of several parts. For example, a technical book consists of chapters. The parts are arranged in such a way that to understand a part, one needs to first understand some (or even all) previous parts. For example, to understand Chapter 3, one needs to understand Chapter 1.

In some cases, in order to understand a certain part, one needs to first understand all the preceding parts. However, in many other cases, not all preceding parts are needed. For example, in a book, Chapter 1 may contain preliminaries which are needed to understand all the chapters, but the following chapters do not depend on each other; in this example, to understand Chapter 3, it is sufficient to know the material from Chapter 1, and it is not necessary to know the material from Chapter 2.

The question is how to describe this relation between understandability of different parts of the text in precise terms.

Similar problems: describing causality and describing pre-requisites. Similar problems occur when we try to describe causality or pre-requisites. When describing causality, we have several events – e.g., in the increasing order of their time. However, the fact that an event A precedes event B does not necessarily mean that A causally influences B. How can we describe causality in precise terms?

Similarly, to graduate, a student needs to take a certain number of classes. There is usually a recommended order in which classes should be taken, but the fact that a class A precedes class B in this order does not necessarily mean that a student has to take class A before class B. For each class B, there is usually a specific list of pre-requisite classes, i.e., classes that need to be taken so that the student will be able to understand the material taught in class B.

Case of partial order. In many practical situations, the relation between the understandability of different parts of the text is a partial order: $A \le B$ means that to be able to understand part B, we first need to understand part A.

This is how relation between chapters is described in many textbooks: by explicitly listing this order. For example, if Chapter 1 is needed to understand Chapters 2 and 3, we get a diagram

$$\begin{tikzcd}
 & \mathrm{Ch1} \arrow[dl] \arrow[dr] & \\
\mathrm{Ch2} & & \mathrm{Ch3}
\end{tikzcd}$$

A general description of understandability. One possible way to describe the understandability relation is to describe possible *states of understanding*, i.e., the class C of possible sets of understood parts. For example, in the above diagram:

- it is possible that the reader does not understand anything; in this case, the set of understood parts is the empty set $\emptyset$;
- it is possible that the reader understand only Chapter 1; in this case, the set of understood parts consists of a single chapter: $\{1\}$;
- it is possible that the reader understands Chapter 1 and 2; in this case, the set of understood parts is $\{1, 2\}$;
- it is possible that the reader understands Chapters 1 and 3; in this case, the set of understood parts is $\{1, 3\}$; and

- it is possible that the reader understands all the chapters; in this case, the set of understood parts is the whole set $\{1,2,3\}$.

Not all combinations are possible. For example, it is *not* possible that the reader understands Chapter 2 but not Chapter 1.

In this case, the class of all possible sets if $C = \{\{1\},\{1,2\},\{1,3\},\{1,2,3\}\}$.

For the case of partial order, the class is a topology. In the case of partial order, the class C is easy to describe: a set S belongs to the class C if with each element $x \in S$, this set contains all preceding elements: if $y \in S$ and $x \leq y$, then $x \in S$. This makes perfect sense: the relation $x \leq y$ means that in order to understand part y, we need to first understand part x. Thus, if a reader understands part y, then we can conclude that this reader understands part x as well.

One can easily check that the class C of all such sets is closed under union and under intersection: if $S, S' \in C$, then $S \cup S' \in C$ and $S \cap S' \in C$. Since the set X of all parts is finite, we can conclude that the class C of all such sets is closed under finite intersection and general union – i.e., is a *topology*; in more precise terms, C can be viewed as the class of all open sets in an appropriate topology.

The idea of describing the understandability relation by a topology was first proposed by O. Prosorov [5, 6, 7, 8, 9, 10, 11]; he called the corresponding topology *phonocentric*.

The corresponding topology must be T_0. Not every topology corresponds to understandability: an additional restriction is that we need to be able to eventually understand all the parts (after studying them in some order). This means, in particular, that for every two parts $x, y \in X$, either we understand x first or we understand y first. In the first case, once we understand x, we have a set S of understood parts which contains x but not y. In the second case, once we understand y, we have a set S of understood parts which contains x but not y.

In other words, for every $x \neq y$, there exists an open set S that contains only one of the two elements x and y. This property of topological spaces is known as Kolmorogov's T_0-property; see, e.g., [13]. Thus, we conclude that the corresponding topology must be T_0.

Vice versa, if the topology is T_0, then there exists a sequential (linear) order in which we can study the parts and, at the end, gain the perfect understanding. First, let us start with the set $X = \{x_1, \ldots, x_n\}$ of all the parts, and let us show that there is a part x which can be studied right away, without the need to study any other part – i.e., a part for which the set $\{x\}$ is open. To find this part, we will prove that there exist open sets V of decreasing size – until we get an open set consisting of exactly one element. We start with the set X which is clearly open. Once we have an open set V which contains at least two different elements $x \neq y$, we can use the T_0-property to come up with an open set U which contains only one of them. The intersection $V \cap U$ is then non-empty and has strictly fewer elements than V. Since we started with finitely many elements, this procedure has to stop – and the only way for it to stop is to have an open set V that consists of exactly one element.

Once we found this starting element x, we can repeat the same argument and find the next element x', for which the set $\{x, x'\}$ is open, etc.

Vice versa, every T_0-topology can be thus interpreted. Indeed, let C be a T_0-topology on a finite set X. For each element $x \in X$, we can consider the intersection S_x of all open sets containing x. Since C is a topology, this intersection is also open – so it is the smallest open set containing the element x.

Let us show that the set $S_x - \{x\}$ is also open. Indeed, due to the T_0-property, for every $y \in S_x - \{x\} \subset S_x$, there exists an open set V_y which contains only one of the two elements x and y. The intersection $I_y \stackrel{\text{def}}{=} V_y \cap S_x \subseteq S_x$ is also an open set that contains only one of these two elements. This intersection I_y cannot contain x since S_x is the smallest of the open sets containing x, and I_y is the proper subset of S_x. Thus, the intersection I_y contains y and does not contain x. The (open) union of all the (open) intersections $I_y \subseteq S_x$ contains all elements $y \in S_x$ which are different from x and does not contain x – so this open union is equal to $S_x - \{x\}$.

We can thus describe the original topology by saying that to understand x, we need to first understand all the elements of $S_x - \{x\}$.

2 Need to Go Beyond Partial Orders: We Get Interior Structures

Need to go beyond partial order. In some texts, the structure is more complex. For example, if some result is needed to understand Chapter 3, we can have two different versions of this result: in Chapter 1 and in Chapter 2, so that understanding one of these two chapters is enough to understand Chapter 3.

In this case, we have the following possible states of understanding:

- it is possible that the reader does not understand anything; in this case, the set of understood parts is the empty set $\emptyset$;
- it is possible that the reader understand only Chapter 1; in this case, the set of understood parts consists of a single chapter: $\{1\}$;
- it is possible that the reader understand only Chapter 2; in this case, the set of understood parts consists of a single chapter: $\{2\}$;
- it is possible that the reader understands Chapter 1 and 3; in this case, the set of understood parts is $\{1, 3\}$;
- it is possible that the reader understands Chapters 2 and 3; in this case, the set of understood parts is $\{2, 3\}$; and
- it is possible that the reader understands all the chapters; in this case, the set of understood parts is the whole set $\{1, 2, 3\}$.

In this case, the class of all possible sets if $C = \{\{1\}, \{2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$.

In contrast to the previous case, here the intersection of two sets $S, S' \in C$ does not necessarily belong to the class C: for example, $\{1, 3\} \in C$ and $\{2, 3\} \in C$, but $\{1, 3\} \cap \{2, 3\} = \{3\} \notin C$.

How can we describe such more general situations?

Closure structures: reminder. A natural generalization of topological spaces are spaces with *closure structures*. To be more precise, closure structures generalize not the usual topology – i.e., the class of all open sets – but the class of all closed sets (i.e., complements to open sets).

A topology is usually defined as a class of sets (called *open*) which is closed under finite intersection and general union. Thus, the corresponding class of all *closed* sets in a topological space can be defined as a family of sets which is closed under finite unions and general intersections.

A *closure structure* on a set X is defined as a family K of sets $S \subseteq X$ which contains the empty set and which is closed under general intersections; see, e.g., [1, 4]. When the underlying set X is finite, the intersection property means two things:

- that the intersection of an empty family is closed – i.e., that the set X itself is closed, and
- that the intersection of every two closed set is closed.

In terms of the family of the complements $C = \{X - S : S \in K\}$, this means that the corresponding family C must have the following properties:

- the family C must contain the underlying set X and empty set: $\emptyset, X \in C$, and
- the family C must be closed under union: if $S, S' \in C$, then $S \cup S' \in C$.

Such families C are called *interior structures* [1, 2].

Definition 1. *Let X be a finite set. A class C of subsets of X is called an* interior structure *if this class contains the empty set and the set X and is closed under union. Elements of C are called* open sets.

Comment. Once we have a closure structure, we can define a closure $\overline{S}$ of a set S as the intersection of all the closed sets that contain S. The corresponding notion of a closure operator $S \to \overline{S}$ is mathematically equivalent to Tarski's notion of a consequence operator – that assigns, to every set of formulas S, the set $\overline{S}$ of all the formulas which can be deduced from S; see, e.g. [12].

The class of possible states of understanding is an interior structure. Indeed, we usually start reading in a state in which we do not yet understand the material described in any part of the text, which means that $\emptyset$ should be a possible state of understanding: $\emptyset \in C$. We should be able to end up in a state in which we understand everything, so we should have $X \in C$.

Similarly, if we can eventually learn all the parts from the set S, and we can also learn all the parts from the set S', then, by first learning S and then learning S', we can learn all the parts from both sets. In this case, our state of knowledge is the union of these sets $S \cup S'$, so $S \cup S' \in C$.

Analog of T_0-property for closure spaces. Similar to the fact that not all topologies represent classes of sets of understood parts, not all closure spaces have this property. An additional restriction is that we need to be able to understand all the parts (after studying them in some order). This means, in particular, that for every open set $S \subseteq X$ – i.e., a possible set of understood parts – and for every two elements $x, y \in S$, either we understand x first or we understand y first. In the first case, once we understand x, we have a set $S' \subseteq S$ of understood parts which contains x but not y. In the second case, once we understand y, we have a set $S' \subseteq S$ of understood parts which contains x but not y. Thus, we arrive at the following definition.

Definition 2. *We say that an interior structure C has* T_0-property *if for every open set $U \in C$ and for every two different elements $x, y \in U$, there exists an open set $U' \subset U$ that contains only one of the two elements x and y.*

Comment. Similarly to the topological case, one can easily check that under this property, it is possible to sequentially understand all the parts – and, vice versa, if it is possible to sequentially understand all the parts, then the corresponding class of sets of understood parts has the T_0-property.

Proposition. *For every interior structure C on a finite set X, the following two conditions are equivalent to each other:*

- *the structure C has the T_0-property;*
- *the elements of the set X can be ordered into a sequence $X = \{x_1, x_2, \ldots, x_n\}$ in such a way that all the sets $\emptyset$, $\{x_1\}$, $\{x_1, x_2\}$, ..., $\{x_1, \ldots, x_k\}$, ..., $\{x_1, \ldots, x_n\}$ are open.*

Comment. A similar description can be made for causality and for pre-requisites. In all three cases, the main difference between this more general case and the case of partial order is that:

- in the case of partial order, we had "and"-rules: to understand Chapter 3, we need to know Chapter 1 *and* Chapter 2;
- in the more general case, we may have "or"-rules as well: e.g., in the above example, to understand Chapter 3, we need to know Chapter 1 *or* Chapter 2.

Vice versa, every T_0-interior structure can be thus interpreted. Indeed, let C be a T_0-interior structure on a finite set X. For each element $x \in X$, we can consider all *minimal* open sets containing x – i.e., all open sets that contain x but for which no open proper subset contains x.

Similarly to the topological case, we can show that for each such minimal set s_x, the difference $s_x - \{x\}$ is also open. We can thus describe the original interior structure by saying that to understand x, we need to first understand all the elements of $s_x - \{x\}$ for one of the minimal sets s_x.

Computational aspect. The fact that we are ready, e.g., to understand Chapter 5 once we understood Chapters 1 and 2 can be described as a rule

$$5 \leftarrow 1, 2.$$

If, to understand Chapter 5, we need to either understand Chapters 1 and 2 or understand Chapters 3 and 4, we can describe this by saying that both pairs of Chapters (1 and 2, and 3 and 4) are sufficient to understand Chapter 5. To describe this, we need rules:

$$5 \leftarrow 1, 2.$$

$$5 \leftarrow 3, 4.$$

By combining these rules, we can represent the whole dependence structure as a corresponding logic program (without negation).

This representation enables us to easily check whether the given dependence structure allows us to learn all the material. At each step of the corresponding algorithm, we mark parts which can be understood. In the beginning, no parts are marked. Then:

- First, we look for all the rules of the type $a \leftarrow$, i.e., we look for the parts which can be learned at first. For each such rule, we mark a as potentially understandable.
- At each stage, we look for the rules $a \leftarrow b, \ldots, c$ for which all the parts in the right-hand side are already marked. For each such rule, we mark a as potentially understandable.

This way, we will either mark all the parts – in which case all the material can be understood – or we get stuck, meaning that the given dependence structure does not allow us to understand everything.

Comment. This idea is in perfect agreement with the use of closure structure in logic programming; see, e.g., [3].

Acknowledgments. This work was supported in part by the National Science Foundation grants HRD-0734825 and HRD-1242122 (Cyber-ShARE Center of Excellence) and DUE-0926721.

The authors are greatly thankful to Oleg Prosorov for valuable discussions.

References

[1] T. S. Blyth, *Lattices and Ordered Algebraic Structures*, Springer, Berlin, Heidelberg, New York, 2005.

[2] M. Erné, J. Koslowski, A. Melton, and G, E. Strecker, “A primer on Galois connections”, *Proceedings of the 1991 Summer Conference on Topology and Its Applications in Honor of Mary Ellen Rudin and Her Work*, Annals of the New York Academy of Sciences, 1993, Vol. 704, pp. 103–125.

[3] J. W. Lloyd, *Foundations of Logic Programming*, Springer-Verlag, Berlin, Heidelberg, New York, 1987.

[4] E. H. Moore, *Introduction to a Form of General Analysis*, Yale University Press, New Haven, Connecticut, 1910.

[5] O. Prosorov, “Compositionality and contextuality as adjoint principles”, In: M. Werning, E. Machery, and G. Schirtz (Eds.), *The Compositionality of Meaning and Content, Vol. II, Applications to Linguistics, Psychology, and Neuroscience*, Frankfurt, Ontos-Verlag, 2005, pp. 149–174.

[6] O. Prosorov, “Sheaf-theoretic formal semantics”, *TRAMES: A Journal of the Humanities and Social Sciences*, 2006, Vol. 10, No. 1, pp. 57–80.

[7] O. Prosorov, “Semantic topologies as linguistic universals”, In: T. Stolz (Ed.), *Proceedings of the 39th Annual Meeting of the Societas Linguistica Europea SLE'2006, Relativism and Universalism in Linguistics*, Bremen, Germany, August 30 – September 2, 2006, pp. 109–110.

[8] O. Prosorov, *Topologies et faisceaux en sémantiques des textes. Pour une herméneutique formelle*, Ph.D. Dissertation, Université Paris X, Nanterre, France.

[9] O. Prosorov, “Topologies and sheaves in linguistics”, In: R. Bhatia (ed.), *Abstracts of the International Congress of Mathematicians ICM'2010*, Hyderabad, India, August 19–27, 2010, pp. 623–624.

[10] O. Prosorov, “Topologies and sheaves appeared as syntax and semantics of natural language”, In: O. Prosorov (ed.), *Proceedings of the International Interdisciplinary Conference Philosophy, Mathematics, Linguistics: Aspects of Interaction PhML'2012*, St. Petersburg, Russia, May 22–25, 2012, pp. 146–163.

[11] O. Prosorov, “Linguistic universals of topological nature”, *Proceedings of the International Interdisciplinary Conference Philosophy, Mathematics, Linguistics: Aspects of Interaction 2014 PhML'2014*, St. Petersburg, Russia, April 21–25, 2014, pp. 55–68.

[12] A. Tarski, “On the Concept of Logical Consequence”, In: J. Corcoran, ed., *Logic, Semantics, Metamathematics: Papers from 1923 to 1938 by Alfred Tarski*, Oxford University Press, 1983.

[13] S. Willard, *General Topology*, Addison-Wesley, Reading, Massachusetts, 1970.