philosophies

Article

Analogy in Terms of Identity, Equivalence, Similarity, and Their Cryptomorphs

Marcin J. Schroeder

Department of International Liberal Arts, Akita International University, Akita 010-1211, Japan; mjs@aiu.ac.jp; Tel.: +81-18-886-5984

Received: 10 October 2018; Accepted: 5 June 2019; Published: 12 June 2019

Abstract: Analogy belongs to the class of concepts notorious for a variety of definitions generating continuing disputes about their preferred understanding. Analogy is typically defined by or at least associated with similarity, but as long as similarity remains undefined this association does not eliminate ambiguity. In this paper, analogy is considered synonymous with a slightly generalized mathematical concept of similarity which under the name of tolerance relation has been the subject of extensive studies over several decades. In this approach, analogy can be mathematically formalized in terms of the sequence of binary relations of increased generality, from the identity, equivalence, tolerance, to weak tolerance relations. Each of these relations has cryptomorphic presentations relevant to the study of analogy. The formalism requires only two assumptions which are satisfied in all of the earlier attempts to formulate adequate definitions which met expectations of the intuitive use of the word analogy in general contexts. The mathematical formalism presented here permits theoretical analysis of analogy in the contrasting comparison with abstraction, showing its higher level of complexity, providing a precise methodology for its study and informing philosophical reflection. Also, arguments are presented for the legitimate expectation that better understanding of analogy can help mathematics in establishing a unified and universal concept of a structure.

Keywords: analogy; identity; equality; equivalence; similarity; resemblance; tolerance relation; structure; cryptomorphism

---

1. Introduction

The concept of analogy belongs to the class of concepts which can be considered elusive, i.e., concepts which almost everyone claims to understand well, which are used in a majority of the domains of inquiry from philosophy to computer science as well as in the everyday discourse, but which under scrutiny escape any commonly accepted definition. The same applies to concepts signified by expressions which include the adjective "analog", such as analog information, analog computing, analog model, etc. As in cases of other elusive concepts such as identity, structure, information, computation, or mind, the omnipresence of the word "analogy" in everyday discourse where there is no expectation for the semantic clarification makes it very difficult to establish a common foundation for its analysis. Every attempt to establish a commonly accepted definition of this and other elusive concepts evokes strong resistance as the precision or lack thereof is usually perceived as a limitation or an incursion into someone's freedom of thinking. However, when we see how much of intellectual effort is put into fruitless discussions in which the only objective is to elevate one view of the meaning over others without much care for establishing any common conceptual framework allowing comparisons, then the need for setting a foundation for the study becomes clear. Of course, setting a foundation for the study of the meaning of a concept does not have to be and should not have to be considered a conclusion of its discussion, but rather an entry into another level of inquiry.

My claim that analogy belongs to elusive concepts may be objected by someone who refers to one of the many dictionary or encyclopedic definitions as a sufficient common ground. For instance, we can find in the respected Stanford Encyclopedia of Philosophy that "An *analogy* is a comparison between two objects, or systems of objects, that highlights respects in which they are thought to be similar. *Analogical reasoning* is any type of thinking that relies upon an analogy. An *analogical argument* is an explicit representation of a form of analogical reasoning that cites accepted similarities between two systems to support the conclusion that some further similarity exists" [1]. This definition, as with many other attempts, is guilty of the common sin of addressing the meaning of hard words which consists of sweeping the dirt under the carpet of another hard but undefined word in the *definiendum*. In this case, this word is "similarity". Not only is similarity not defined in the proposed definitions, but it is not clear whether the words "similar" and "analogous" are, or should be considered different in their meaning.

The main objective of this paper is to search for the common ground and to establish a firm foundation for the study of analogy in the experience of mathematics and in its conceptual framework. It is a surprisingly well-kept secret, that there is a quite extensive mathematical theory of general similarity relation initiated by Eric Christopher Zeeman more than fifty years ago, unfortunately under the misleading term of tolerance relation [2,3]. Zeeman introduced this term because it accurately represented the application of the concept of mathematical similarity to the context of the article, i.e., the mathematical modeling of the brain mechanisms involved in visual perception. For instance, his "[...] notion of 'tolerance' within which we allow an object to move before we notice any difference" [2] (p. 241) was first formulated using the concepts of geometry with the intention to provide a mathematical tool for the study of cognition, but with the following this specific case the general definition as a binary, reflexive, and symmetric relation intended to build the bridge to topological description. Zeeman probably did not expect that his tolerance relation would initiate extensive research reported in hundreds of papers. In mathematics, the choice of terminology is rarely a subject of deeper reflection. However, in this case, the decision to use the term "tolerance" had significant negative consequences for the dissemination of the results in mathematical studies of similarity outside of the relatively narrow group of experts. It is very unlikely that someone unfamiliar with general algebra looking for a theory of similarity would guess its mathematical name in such a disguise or realize its affinity to some fundamental mathematical theories. There are many other commonly-known, and others confined to more specialized literature, pieces of mathematical theory relevant to the study of analogy.

To avoid confusion, the present paper is not intended as a contribution to mathematics, but rather a contribution to philosophy of analogy through the use of the familiar rudimentary mathematical concept of similarity and its well-known and rich theory for the purpose of establishing the conceptual framework defining analogy and for relating it to other fundamental philosophical concepts such as identity, equality, equivalence or difference. Moreover, the objectives of this paper do not include the resolution of any particular philosophical controversies over analogy. The main purpose is to establish a common foundation and to present a tool-kit of mathematical methods to assist future studies of analogy and its philosophical interpretation. If more specific issues in the study of analogy are invoked here, it is only to show the effectiveness of these intellectual tools or to demonstrate the need for their use.

While the paper has as its main objective to contribute a mathematical formalism to the methodology for the study of analogy and similarity, it has an additional objective to present the possibility of using analogy for the purpose of establishing a clear meaning for the general concept of a structure.

2. Symmetry of Analogy

Another legitimate issue regarding the use of tolerance relation as a mathematical model for similarity is whether the concept of tolerance relation is sufficiently general. At first, this may seem obvious since tolerance is a binary relation on a set with only two simple defining conditions: The relation is reflexive (every element is similar to itself) and symmetric (if x is similar to y, then y is similar to x). This paper goes in the direction of a slightly increased generality when similarity in the most general sense does not have to be reflexive but has to satisfy the condition that whatever is not similar to itself is not similar to anything else. This weaker than reflexivity condition introduced in the present author's earlier publications is dictated by the instances of similarity in mathematical theories, but it is of some importance for the philosophical discourse on the relationship between identity, similarity, and difference [4,5].

We do not have to go beyond the generality of reflexive tolerance relations to wonder whether it is not excessive for the concept of similarity. However, the existence of the large variety of applications of this concept in virtually all domains of inquiry makes this concern unjustified. On the other hand, there is one domain, psychology where the symmetry of similarity was questioned and arguments were presented calling for a more general definition.

Probably the most representative for this way of thinking is the article by Amos Tversky in which he claims that symmetry is too much restrictive [6]. Tversky admits that "Similarity has been viewed by both philosophers and psychologists as a prime example of a symmetric relation. Indeed, the assumption of symmetry underlies essentially all theoretical treatments of similarity." [6] (p. 328). However, he claims that his paper provides "empirical evidence for asymmetric similarities". Tversky's paper has some minor factual errors (e.g., his metric space is not metric, but pseudo-metric and his metric introduced to measure similarity is not a distance or metric, but quasi-metric with little affinity to geometric concepts), but more problematic is the logical error leading to his conclusion of the asymmetry of similarity. He claims that "Similarity judgments can be regarded as extensions of similarity statements, that is statements of the form 'a is like b.' Such a statement is directional; it has a subject, a, and a referent, b, and it is not equivalent in general to the converse similarity statement 'b is like a.' [...] We say 'the son resembles the father' rather than 'the father resembles the son'."

This and other examples in the paper do not provide any empirical proof that similarity is not symmetric, but that the statements about similarity may be strongly context dependent. If you do not know anything about the additional structure serving as a context (in this case family relationship) there is no reason to claim any direction in similarity statements. "John is similar to Joe" carries exactly the same meaning as "Joe is similar to John", but when we learn that Joe is a son of John and that sons inherit features of fathers, then we may consider the latter statement more appropriate. Even this is not obvious, because we may talk to someone who knows Joe better than John. In this case, it would be more appropriate to use the former statement. In each of Tversky's examples, there is a hidden ordering "directing" the similarity statement. Of course, there is nothing wrong in considering relational structures consisting of tolerance relations and some other relational structures, for instance of the type of order. However, a general theory has to present all component relations, and here is Tversky's error. He does not identify any additional relations that are "smuggled" into his examples. Tolerance relations with additional structures have been studied from the very beginning of the theory, even in the paper of Zeeman that initiated their study.

For this paper, it is important that the claim of Tversky's paper that similarity requires more general theory without the requirement of symmetry is clearly false. The same objection applies to papers making similar claims of directionality. Probably the most relevant here is an example of the paper "Structure Mapping: A Theoretical Framework for Analogy" by Dedre Gentner [7].

Here too we have the claim of asymmetry of analogy expressed by the statements of the form "A T is (like) a B" and corresponding to this claim of directionality the formalization of analogy as a mapping from the (relational) structure B (base) to (relational) structure T (target). Gentner's article involves several undefined terms (e.g., "object nodes", "discarding of attributes", "large number", "few", etc.), some formally meaningless common sense "postulates" (e.g., the second postulate for mappings formulated as "Try to preserve relations between objects") and lastly, at least one fundamental logical error (in the definition "An *abstraction* is a comparison in which the base domain is an abstract relational structure" which without an explanation of the meaning of "abstract relational structure" independent from "abstraction" becomes circular). These defects make the text a caricature of mathematical formalization worth mentioning only because it reveals the intuitive symmetric character of analogy. Although Gentner explicitly writes about the directionality of analogy at the beginning, later analogy becomes a symmetric domain comparison as for instance in "In the structure-mapping framework, the interpretation rules for analogy can be distinguished from those for other kinds of domain comparisons." [7] (p. 159).

Thus, the only philosophical claim of the present paper is that analogy is a (weak) tolerance relation and as such its study may benefit from the consideration of the results of mathematical analysis of this concept. It is important to remember that the concept of tolerance relation belongs to the sequence of philosophically fundamental concepts of increasing generality starting from identity and going through equality, equivalence.

Thus, a brief review of concepts forming the context for analogy, such as identity, equality, equivalence, similarity, etc. will be followed by the exposition of corresponding mathematical concepts. All these concepts come from the same theoretical framework of general algebra, so mathematics gives us not only a conceptual framework but also an insight into the interdependence of relevant concepts. In this paper, all effort will be made to avoid technical detail and to make the presentation accessible to everyone with or without a deeper mathematical background. However, this article is about the contribution of mathematical theory to philosophy and the study of analogy, so the presentation of its theses in the complete absence of some exposition of mathematical concepts and relevant mathematical theory is impossible.

Contributions of mathematics to the study of analogy are not limited to the unification of concepts present in the discourse on similarity, but they permit a critical analysis of some claims or ideas within earlier studies. This paper is not intended as a comprehensive review of all applications of mathematical knowledge of tolerance relations to the study of analogy. Such applications are just examples supporting the view that we get a useful tool. For instance, in her classic work *Models and Analogies in Science* Mary Hesse introduced the distinction between positive, negative and neutral analogies [8]. The positive analogy between items consists of the properties or relations they share. The negative analogy in the ones they do not share and the neutral analogy comprises the properties of yet to be distinguished as positive or negative. This type of distinction, although justified in the restricted meaning of the particular socio-scientific context considered by Hesse, becomes meaningless in the general context-independent case where statements of the type "not known yet" (by whom? when? etc.) cannot be considered legitimate. Better examples of the power of mathematical conceptualization can be given in the critical analysis of Hesse's claim that "similarity must resolve into identity and difference" [8] (pp. 70–71). Using mathematical tools, this paper will demonstrate that this claim is false and at most replaceable, since it can be substituted by the "resolution" of every similarity into equivalence relation and a pure form of similarity (Wittgenstein's family resemblance) in which all objects, although similar, are distinctive.

Mathematical theory of similarity defined as a tolerance relation can be used to demonstrate the fallacy of the idea to measure similarity by counting the number of all common properties or by using for this purpose the proportion of common properties to all possessed by the objects, as for instance in the mentioned before paper by Amos Tversky [6].

The first step in this direction, but without a mathematical theory of similarity, was done by Nelson Goodman [9]. He objected the concept of similarity as an absolute relation independent from any specifications of properties with respect to which it is considered. He correctly insisted that similarity should always be considered with respect to some family of attributes. He overstated his claim asserting that objects can be similar in an unlimited number of ways. It is true that in some situations that they can be similar in the infinite number of ways, but strictly speaking the issue is not in the "unlimited" number of ways but in the indefiniteness of this number. Later we will see that the same tolerance relation can be defined by several different families of attributes and that the size of these families can be different. Thus, the strength of similarity of two objects with respect to one family of attributes measured by common attributes can be different when we use a different family of attributes. This contradicts the much later claim by Satosi Watanabe that "To persuade that two objects are similar, it is natural to enumerate the properties that are commonly owned by the objects. The more properties are shared, the more similar they are." [10]. This error could be avoided, if Goodman's objections, formulated as philosophical arguments, opened the discussion supported by mathematical proof from the theory of tolerance relations showing that there are no possible counterarguments.

There is another and, in some sense, independent and tangential theme of this paper. It is the role of analogy as an intellectual instrument in the study of structures. This is not restricted to the issue of the extension of analogy as similarity between objects understood as elements of some set S to similarity between ordered pairs as elements of a binary relation on S, i.e., a subset of the direct square of S. Relational structures, binary or of higher order, with one or with many relations are not the only structures within mathematics and its applications to more specific domains of inquiry. Analogy is sometimes understood as interdependence or similarity of structures (we could see an example of a not very successful attempt of this way of thinking in Gentner's paper criticized above), but this does not eliminate problems in its understanding. After all, the general question "What is a structure?" is not at all easier to answer than "What is analogy?" Both structure and analogy are elusive concepts.

Instead of providing the ultimate answers to both questions, this paper presents an outline of the mutual dependence between these two concepts. We claim only that analogy can be defined as a (weak) tolerance relation on some set, but this definition leaves open several questions regarding the pragmatic aspects of this concept, for instance, its role in logic. On the other hand, the question about the general concept of a structure is without any definite answer. We could already see in the examples presented above that a typical interpretation of the word "structure" is a relational structure, i.e., a set with some number of relations defined in it. However, we have many examples (for instance in mathematics) of structures which are not relational. Also, some relational structures are considered equivalent to structures which are not relational.

Questions about the general meaning of the concept of structure cannot be easily resolved by the existing tools of mathematics, such as the concept of a morphism (structure preserving mapping) which requires pre-existence of a conventional—and to some extent arbitrary—choice of implementation. The same structure (for instance a topological space) can be implemented in several different, but apparently equivalent ways (topological space is defined by the class of open subsets, closed subsets, closure operator, or a long sequence of other equivalent operators, base for open subsets, base for closed subsets, nets, filters, etc.) On the other hand, seemingly very different structures may be easily associated as essentially identical. For instance, even mathematicians are sometimes surprised that every topological space on a finite set corresponds in a unique way to a quasi-ordering and in the case of a topological space at the lowest level of separability (i.e., satisfying axiom T_0 which very rarely is not required) this is simply a partial ordering [11].

The equivalence of definitions for many commonly equated structures may seem in some cases obvious, but this impression may be deceptive because the transitions between defining concepts require in some cases additional ad hoc assumptions and there is no uniform description of the way they should be performed. Justification of the claim that the two conceptual frameworks produce essentially the same structure is done in the case by case manner, not by following a universal rule. Different levels of abstraction come with ramifications or mergers of the categories of implementation. Because of this lack of a universal method for comparing and relating mathematical structures, despite having the same conventional name they have quite frequently different and non-equivalent definitions.

Paradigmatic example can be found in the diverse menagerie of the high-level generalizations of topology. Why dropping of one axiom defining a topological space keeps it in the category of topological spaces, but dropping another does not? At which level (if any) of the axioms of separation does topology begin? If the answer is that it is just an accidental matter, then the very concept of mathematical structure becomes questionable.

In everyday mathematical practice, in the absence of isomorphic mappings which by the definition require presence of identical objects on both sides of the function so that their preservation can be assessed, we are expressing the fact of similarity within the class of structural implementations by referring to "*cryptomorphic presentations of a structure*" (or *cryptomorphs*). The term "*cryptomorphism*" was introduced by Garett Birkhoff already in 1967 and defined in a quite narrow context, but it was never clearly defined in full generality because in each instance this type of similarity relation requires an ad hoc procedure. Thus, there are many unanswered questions regarding this thus far intuitive concept. How is it possible to describe the identity of a cryptomorph or cryptomorphic class of the structure independently from the particular choice of its defining concepts and corresponding to the sets of axioms? What actually is "*cryptomorphism*"? How is the concept of cryptomorphism related to the general concept of a structure? All we can say is that it is expressing our intuitive sense of similarity. We are simply content to grasp analogy, or rather we are forced to be content.

3. Analogy as a Universal Intellectual Tool

Analogy, although understood in many different ways throughout the ages of its use in philosophical and scientific inquiries was always a tool for eliminating, reducing or handling complexity. For instance, Aristotle writing in *Metaphysics* $(1048^a 25\text{-}^b 17)$ about the antithesis of potentiality and actuality escaped the trouble to explain the complexity of the concepts involved in it by invoking analogy "[...] we must not seek a definition of everything but be content to grasp the analogy" [12] (p. 82). In this case, the escape from complexity was achieved by building analogy between the relationship of metaphysical concepts at a very high level of abstraction (potential existence and actual existence) and the relationship between particulars coming from our everyday experience.

However, if analogy was simply a replacement of the abstract, general by that which is particular (as in the case of so-called ostensive definition), then it would have been just an illustration, possibly confusing and misleading. So, what is analogy and why does it have such an important role as an intellectual tool? The etymology of its name refers to the Greek word for proportion derived from geometric analysis of figures and therefore apparently related to the quantitative, metrical analysis of objects in human perception. But in fact, analogy belongs to fundamental concepts of the structural, i.e., qualitative methodology. Even in this original, literal meaning of the Greek word "*analogia*" as a proportion of geometric measures the equality of mutual relationships of the components within a whole, not their numerical values is important. It is no wonder that already in the philosophy of Greek antiquity analogy acquired much more general meaning of the equality or similarity in structural relations expressed frequently in terms of non-numerical, qualitative, intuitive proportions.

At this point, it is worth to reflect on the abuse of numerical proportions which can be misleading in the search for analogy. Not all magnitudes describing even most familiar systems of everyday experience have meaningful proportions. This is why statistics makes the distinction between the so-called interval and ratio levels of measurement to avoid the abuse of proportions. For instance, data consisting of the measurements of temperature expressed in Celsius or Fahrenheit scales belong to the former type. A twofold increase of temperature in either of the two scales is meaningless, as it can be easily observed that the same increase will produce different ratio dependent on the scale. Thus, saying that the twofold increase of the temperature measured in Celsius degrees indicates that it is twice warmer, as the majority of people would be inclined to say, is meaningless. On the other hand, the proportion of temperatures expressed in Kelvin scale is meaningful and reflects structural characteristics of energetic processes.

The example of the common-sense abuse of proportions in the interpretation of temperature or other quite frequent mistakes in interpretation of numerical values of other magnitudes should not be considered evidence against the role of intuition in analogy judgment. They may serve as a warning against fallacies in oversimplification of analogy. In fact, there is sufficient evidence for the surprisingly high human perceptual skills in the detection of similarities between structural characteristics expressed by proportions. For instance, from the observation of the interdependence between the proportions of the lengths of string in monochords and musical harmony, Pythagoreans derived their general concept of harmony which guided scientific inquiry in the centuries to come. Obviously, musical harmony was already a well-established concept when Pythagoreans associated it with proportions of numbers.

The ability to detect even small deviation from the musical harmony demonstrates the more general capacity of human perception in the recognition of structural characteristics without any direct analytical tools. We can associate it with the special role of analogy in the human capacity to identify structural resemblances which cannot be easily described or formalized. However, this astonishing capacity cannot be overextended to the domains of inquiry based on analytical methods. We should not be "content to grasp analogy" as advised by Aristotle, if we leave the judgment of the validity of the analogy in the analytical philosophical discourse exclusively to our intuition.

Thus, even if we fully appreciate the role of intuitive detection of analogy between structures, there is a legitimate question about the function of analogy in the study of structural characteristics. It is quite clear that analogy works through similarity or even equality (as in the case of proportions understood literally). However, even if frequently, but mistakenly, analogy is reduced to a binary relation on some set S of the type of identity (in logical sense), equivalence (binary relation which is reflexive, symmetric and transitive) or its generalization similarity (in mathematics tolerance relation which is reflexive and symmetric, but not necessarily transitive [13]), it actually can describe correspondence between structures built over the set of all subsets of the set S (power set of S), or objects of even higher set theoretical rank. Of course, there is nothing wrong in calling the similarity relation on a given set of objects to be an analogy, but we should consider the special role of analogy in the study of structural characteristics which may be lost in the reduction to tolerance relations at the lowest level of elements devoid of their own internal structural characteristics.

If we have predefined structures of a particular type (e.g., algebraic structures, partially ordered sets, topological spaces, etc.), then we could consider the description of analogy in terms of functions (homomorphisms, isomorphisms, etc.) between structures which preserve structural characteristics (algebraic operations, order, topology). In this approach structures are primary concepts and analogy is introduced as a secondary concept defined by selected functions determined by the condition of these structures' preservation. However, this approach trivializes analogy.

First, we lose the role of analogy as a tool for the inquiry of the structure, or for the determination of structural characteristics. If the structure is already defined and fully characterized, there is not much use for analogy. Moreover, these specific types of mathematical structures mentioned above are just examples of only apparently special importance. There are many other examples of at least equal philosophical, theoretical and practical significance.

A much more fundamental problem is that even for specific types of structures (for instance algebraic structures) for which the concept of morphism is well defined, it generates a relation which is transitive (composition of morphisms is a morphism), while transitive relations of similarity between structures (equivalence relations) form a very restricted sub-category.

4. Identity, Equality, and Equivalence Relation

It is surprising that in the literature on analogy there are a lot of references to the concept of identity and similarity and their mutual relationship, but there is very little interest in relations of equality and equivalence [8,14]. Identity was a subject of philosophical disputes since the earliest time of European philosophical tradition. One of the sources of controversies arising in the study of identity was in its role in both epistemological and ontological contexts. Since these issues are of secondary importance for the subject of the present paper and the views of the present author were presented elsewhere [15,16], identity will be considered here mainly in its relation to the other concepts of equality, equivalence, and similarity. Therefore, there will be no much interest in the ontological aspect of identity as a condition for existence.

The four concepts can be considered as expressions of relations of decreasing level of restriction or alternatively of increasing generality in the order as listed above with identity preceding equality.

Every two identical objects are equal, every two equal objects are equivalent, and so on. Neither of the links can be reversed. For instance, the distinction between identity and equality becomes crucial for the distinction between logical relationship and the relationship within theories with equality [17]. A simple example of the difference between logical identity and equality can be given by expressions "1 + 1" and "2". They are equal within arithmetic, but not identical in logic. Additionally, the first three relations are transitive (e.g., if a is equal to b and b is equal to c, then a is equal to c). The last one, similarity, obviously does not have to be transitive. For instance, when similarity in natural numbers is defined by a selection of the properties of divisibility, then the pair 7 and 9 can be considered similar as these numbers are odd (not divisible by 2) and the pair 2 and 7 similar, as they are prime (divisible only by themselves and 1). However, 2 and 9 are not similar with respect to either of the two properties.

4.1. Identity and Equality

The classical criterion for identity formulated by Leibniz in his *Discourse on Metaphysics*, Section 9 [18] (p. 308), but already considered by Thomas Aquinas in *Summa Theologiae* (*ST, I, Q40, A1, O3*) [19], and in some sense by Nicholas of Cusa in *De Li Non Aliud* (NA, 5, 18:9) [20], called *principium identitatis indiscernibilium,* states that two individuals are identical, if for any intrinsic, non-relational property it can be asserted that one has it if and only if the other has it. Today this principle is better known under its English name "principle of the identity of indiscernibles". It is hardly surprising that this criterion was many times challenged by David Hume, Immanuel Kant, and others, for instance regarding the meaning of intrinsic property or problematic reference to all properties.

The latter condition of identity with respect to all properties is in clear contradiction with the long-standing Aristotelian tradition in which there is a fundamental distinction between essential and accidental properties exactly for the reason of saving the concept of identity, in particular in the diachronic perspective. Saying that a man who shaved his beard is not identical with himself before shaving seems to be absurd. Thus, we should make a distinction between the essential properties defining identity. However, all efforts to provide objective criteria for these properties failed.

As a concern about the requirement of intrinsic properties to be considered, Kant in the *Critique of Pure Reason* restricts validity of Leibniz's principle to the phenomenal level and provides an example of two drops of water with all intrinsic quantitative and qualitative properties the same, but in different locations to be two different entities with their own *numerica identitas* [21] (p. 192).

There are many ways to escape traps of the traditional concept of identity, but the logical distinction between identity and equality has the most solid formal foundation and is consistent with our intuition. The man, before shaving his beard and after, meets the criterion of equality, but not the criterion of identity. After all, the former had a beard and the latter does not, so they are not identical but only equal. The distinction may be considered a hidden import of essentiality. Why is the man after shaving equal to the man before shaving, but his cut beard is not? The answer is that equality requires consideration of the structural identity at the level where we consider objects as equipped with the internal structure [16].

Since the issue is of special interest in this paper, it needs to be addressed in a more elaborate way. In our current understanding of fundamental philosophical concepts, whether we are aware of this or not, we carry a lot of luggage from the philosophical tradition. In the long run, there were two main positions already established in Mediterranean Antiquity in the discussion of ontological status of universals. "Universal" as a predicate of one variable referred to an actually existing object (universalist position) or was just a generic name for the collection of actually existing objects possessing the universal as an attribute of the secondary ontological status (nominalistic position). There was nothing in between.

We could see the reflection of the dispute in the discussion of the problem of the identity of indiscernibles. In the universalistic approach, universals were entities with their own structures (for instance Platonic forms) expressed in essential properties. In the approach of moderate realism, universals did not have an independent existence and they were comprehended as structures abstracted from the objects which they represented. However, without any developed tools for structural analysis in both approaches, the universals were defined and analyzed in terms of external relations, such as their genus, species, differentia.

The dispute of universals was revived by the philosophical issues arising from the development of axiomatic set theory. One of the critical points was the problem of the (unrestricted) axiom schema of comprehension (sometimes called axiom schema of specification) which states that the collection determined by $\{x: \varphi(x)\}$ with $\varphi(x)$ standing for any predicate having x as the only free variable is a set. Bertrand Russell's Paradox in which $\varphi(x)$ is chosen to be "$x \notin x$" shows that the schema has to be revised, for instance by the relativization to the values of the variable x that are elements of some set S, i.e., the schema changes to the following: $\{x: x \in S \ \& \ \varphi(x)\}$.

This restriction seemed to provide an argument for the nominalistic position. Properties did not have their own independent existence or even meaning without the reference to a set within which variables of their predicates take value. The consequence of this priority was that properties (universals) can be simply understood as subsets of some sets, which of course leads to the perversion and trivialization of the very idea of property. Every set can be identified with a property expressed by the predicate $\varphi(x)$ = "$x \in A$". We are not concerned here at all on how the property is generated in x nor what type of structure is responsible. The structure imposed on all properties relativized to set S becomes simply a Boolean algebra isomorphic to that of all subsets of a given set S.

The luggage which we inherited in our current view of these matters is that we have insight only to the right side of the statement "$x \in A$". The left side, set x, which is an element of set A (in usual forms of set theory all objects of inquiry are sets, but some of these sets can be elements of others), is opaque for our analysis. The object x has a property shared with all other elements of A, and only with them, but its internal structure is invisible for us. This makes the distinction between "$x = y$" understood as identity and "$x = y$" understood as equality difficult to comprehend. It is significant that in mathematical symbolic convention these two different relations have the same symbol "=". Fortunately, in mathematical logic, there is a very clearly defined distinction of theories with and without equality.

4.2. Equivalence

The equivalence relation is probably the most important tool of mathematics as it serves as a ladder to the higher levels of abstraction. Its origins are in the very remote past and they are difficult to trace. Some insight is possible through ethno-scientific research on folk categorization and folk taxonomies in cultures not influenced by educational systems propagating cognitive methods developed in European tradition and influenced by the Aristotelian analysis of abstraction.

Aristotelian logical genus-species relationship linking different levels of abstraction was abducted and transformed by Tournefort, Linnaeus and other biologists into a more rigid taxonomic system in which genus and species are only particular two lowest levels of taxonomic hierarchy of life. However, biologists most likely just reclaimed that which originally belonged to the human exploration of the diverse structure of living forms. The biological hierarchic structure of life may be associated with folk taxonomic systems of a large variety of cultures, some completely detached from the influence of Europe.

The early studies of cognition in pre-industrialized societies carried out by Alexander Luria in the 1930s in Uzbekistan suggested that the functions of objects are the most important factors organizing perception of reality among people not influenced by modern education [22]. This is in clear contrast with categorization based on inherent characteristics dominant in industrialized societies. However, this conclusion was questioned later as being likely a result of a misunderstanding of the purpose of tasks used in tests. Uzbeki peasants were asked to remove one of the four pictures which in the least degree fits the other. If three pictures presented trees and one of them an axe, they never removed the axe, and protested if it was suggested, as without axe you cannot do much with the trees. This was interpreted that in their perception function of the axe was binding their perception stronger than the morphological similarities of trees. However, they had a general concept of a tree excluding any relationship with tools, so this interpretation was not justified.

In the 1960s, the discipline of ethnoscience emerged initiated by the earlier works of J. H. Greenberg [23], H. C. Conklin [24], W. H. Goodenough [25], and F. G. Lounsbury [26]. Its methodology was based on linguistic analysis of the ways how concepts are structured in different cultures following the following argument of Sturtevant: "The main evidence for the existence of a category is the fact that it is named" [27]. The program's goal was described as a study of "classifications as reflected by native terminology; discerning how people construe their world of experience from the way they talk about it" [27].

It turned out that categorization in non-European cultures may be finer than that of systematic botany. C. O. Frake reports that the Hanoonoo tropical forest agriculturalists of the central Philippines partition their plant world into more than 1600 categories, whereas systematic botanists classify the same flora into less than 1200 species [28].

The lesson which can be learned from ethnoscience is that folk taxonomies are much closer to the biological systematics than it was expected. In particular, there are close similarities in the general patterns of using morphological regularities in constructing taxa, the formation of sequential forms in naming, i.e., going to lower levels of taxa by adding words restricting the application of the name. We can see that structural analysis of human experience is not conditioned by analytical methods developed in particular cultural formation. Moreover, the foundation for this universal methodology is in the partition of the objects into disjoint classes organized into a hierarchic system. Since these classes acquired names, we can consider each level of this hierarchic structure an equivalence relation.

Another example of the fundamental role of equivalence relation across different cultural formations and in the wide chronological span is the use of numbers. We can see here not only an expression of the universal character of equivalence as a cognitive tool, but also the universal human ability to reach the highest level of abstraction at which objects are considered devoid of any individual properties. In this case, the equivalence is a relation not between objects, but between sets of objects based on arbitrarily chosen one-to-one correspondence of their elements. The fact that in some modern languages spoken in large, highly-industrialized language communities (for instance in Japanese) there are exceptions in the form of partial level of abstraction employing so-called counters which pre-group counted items according to some properties (e.g., counters representing flat objects, long objects, etc.) shows that the attaining of the high level of abstraction required some form of cognitive evolution and was not trivial.

There are too many instances of the use of equivalence relations in modern science to address all of them. For instance, equivalence relation appears in a fundamental role associated with numbers in modern quantitative methods of science, for instance in probability and statistics. Probability theory can be formulated in terms of the set of outcomes Ω equipped with appropriately defined probability measure P on events understood as subsets of Ω. There is an alternative approach in which emphasis is not on the set of outcomes, but on random variables, functions defined on Ω with their values in the set of real numbers. The axioms for the probability measure on Ω can be translated into axioms for the probability distribution of the random variable and sometimes the set of outcomes does not appear at all. This convenient procedure involves an equivalence relation and the transition from the set Ω on which the random variable is defined to equivalence classes of this relation. When we define a random variable X, we define partition of Ω into subsets with the same value of X. Thus, each value of the random variable X represents an event, i.e., a subset of Ω. Different random variables correspond to different partitions of Ω, i.e., different equivalence relations. In a similar way majority of quantitative methods of science involve equivalence relations and their equivalence classes. The only difference between the qualitative and quantitative methods is that the former use predicates expressing properties, while the latter replace equivalence classes with numbers on which algebraic operations are defined.

Finally, we can observe that the difference between equality and equivalence relations is a matter of the difference between the levels of abstraction. Since every equivalence relation on a given set S is uniquely associated (through a cryptomorphism!) with a partition of the set S (a separation elements of S into a covering of S by mutually disjoint subsets), each subset in this covering can be considered an object or element of the higher rank set of all subsets of S (called a power set of S). In this transition, that which was the equivalence relation on S becomes equality in the power set of S.

4.3. From Equivalence to Similarity

While the concept of equivalence relation belongs to the most elementary tools of mathematics and appears in almost all of its applications, similarity, with its cryptic name of tolerance relation, may be unfamiliar even for many mathematicians, unless in the finite case it is re-identified as the subject of graph theory [13]. Of course, similarity was sporadically invoked in some mathematical texts of more remote past, for instance by Henri Poincaré [29], but its more systematic theory started to develop only in the 1960s. Most likely this delay was due to the lack of expectations for non-trivial results with highly-restricted defining conditions. Also, the concept of similarity in geometry understood as invariance with respect to a uniform re-scaling (e.g., similar triangles familiar from school geometry) which actually is an equivalence relation, might have contributed to confusion and lack of interest.

Outside of mathematics, similarity was not faring better, at best as a poor cousin of equivalence relegated to the range of informal, intuitive or artistic skills. It is possible that the early interest in similarity in the time of Romanticism was stimulated by its appeal to those who avoided formality and certainty. It took quite a long time before similarity acquired its recognition as a competitor of the equivalence relation in its role in linguistics.

The most influential critique of the role of equivalence and support for replacing it by similarity is in the work of Ludwig Wittgenstein, in particular in his 1953 *Philosophical Investigations*, where he elaborated on the intellectual tool of the family resemblance (*Familienähnlichkeit*) for the analysis of language, meaning, and comprehension [30].

Although the metaphor of family resemblance in the context of philosophical reflection on classification and categorization currently firmly associated with Wittgenstein was traced back by Rodney Needham [31] already in the 1860s in Grimm's Dictionary and in the work of Friedrich Nietzsche [32], there is no doubt that *Philosophical Investigations* started the new era in philosophy and scientific methodology. To be sure, Wittgenstein was interested in the specific type of non-transitive similarity (similarity which is not equivalence) for which he used the metaphor of family resemblance (*Familienähnlichkeit*) and distinguished it from a more general similarity (*Ähnlichkeit*) which includes a transitive case of equivalence, although how much he was aware of the role and meaning of transitivity is not clear. For our purpose it is important to observe that in works analyzing family resemblance, as well as the variety of taxonomic methodologies (e.g., *polythetic classification*), we can observe some significant shift in thinking.

Wittgenstein and those interested in alternative forms of taxonomy (in biological sciences or anthropology) returned to the view that there is no reason to the claim that any selection of objects from a given set defines a property. Wittgenstein's view of the role of family resemblance in our comprehension was of course of the greatest importance, since it was going beyond just pragmatic consideration for taxonomy.

In the models of family resemblance, the objects (typically called items) were entered independently of properties (attributes) and the focus was on the construction of their interdependence. One model of family resemblance was, for instance, described by listing selections of properties (indicated by capital letters A-E) characterizing five items: {A, B, C, D}, {A, B, C, E}, {A, B, D, E}, {A, C, D, E}, {B, C, D, E} [31,33]. Of course, we could have described this family resemblance by merely listing selections of objects (marked here by small letters a-e) with given five properties: A ~ {a, b, c, d}, B ~ {a, b, c, e}, C ~ {a, b, d, e}, D ~ {a, c, d, e}, E ~ {b, c, d, e}. Wittgenstein's insistence on the specific form of the relationship between objects and properties (family resemblance) generated interest in the structural analysis of the relationship.

We can observe that the family resemblance is not an equivalence relation, as the classes of objects grouped according to the five properties cover the entire set of objects S = {a, b, c, d, e}, but they are not disjoint, and they do not allow for finer, disjoint partition. The relation $\boldsymbol{R}$ defined by x$\boldsymbol{R}$y if both x and y have at least one common property is symmetric (If x$\boldsymbol{R}$y, then y$\boldsymbol{R}$x), reflexive (x$\boldsymbol{R}$x), but not transitive (it is not true that, if x$\boldsymbol{R}$y and y$\boldsymbol{R}$z, then x$\boldsymbol{R}$z). Reflexivity and symmetry are conditions for a tolerance relation which is identified in mathematics with similarity.

Someone can ask whether these two very simple conditions can produce a theory of any interest. While this is a normative question addressing interests and preferences and it is difficult to answer it in an objective way, we can respond to it with another rhetorical question: Is graph theory of any interest? After all, tolerance relations on a finite set S and graphs on S are cryptoisomorphic implementations of the same structure.

5. Mathematical Theory of Similarity

This section presents the mathematical expression of the issues addressed earlier. Mathematical concepts give us tools for the analysis of analogy. In order to make this paper self-sufficient, the exposition is elementary and includes all necessary definitions and a few relevant propositions in order to provide a brief overview of the theory. However, all propositions will be presented without proofs which can be found elsewhere, together with an elaborate exposition of technical aspects [11,13].

5.1. Binary Relations and Their Algebra as Tools

Similarity can be formalized within an algebra of binary relations. A binary relation on a set S is a subset of the direct product S×S. If we have any predicate for two variables R(x,y) with variables assuming values in the set S, we can associate it with the relation $\boldsymbol{R}$ = {(x,y): ((x,y) ∈ S × S & R(x,y)}. As the set $\mathscr{R}$(S) of all binary relations on S is a set of subsets of S × S, and therefore a set, it can be partially ordered by inclusion. This partial order can be defined in $\mathscr{R}$(S): $\boldsymbol{R} \leq \boldsymbol{T}$ *iff* x$\boldsymbol{R}$y ⇒ x$\boldsymbol{T}$y. We can consider a Boolean algebra structure on $\mathscr{R}$(S) by importing set theoretical operations from S×S. Boolean operations distinguish the *empty relation* **Ø** and the *full or universal relation* S×S. We can also define a *complementary relation* $\boldsymbol{R}^c$ for relation $\boldsymbol{R}$ in $\mathscr{R}$(S) by:∀x,y ∈ S: x$\boldsymbol{R}^c$y *iff* not x$\boldsymbol{R}$y, or in other words: ∀x,y ∈ S: x$\boldsymbol{R}^c$y *iff* (x,y) ∉ $\boldsymbol{R}$.

The only nontrivial operations giving $\mathscr{R}$(S) its rich structure going beyond Boolean algebra are composition and converse operations. The *composition operation* is defined for any ordered pair of relations $\boldsymbol{R}$, $\boldsymbol{T}$ by: ∀x,y ∈ S: x$\boldsymbol{RT}$y *iff* ∃z ∈ S: x$\boldsymbol{R}$z and z$\boldsymbol{T}$y. The *equality relation* $\boldsymbol{E}$ = {(x,y): x = y} is compatible with the order and gives $\mathscr{R}$(S) the structure of a *partially ordered monoid*. The other specific relation algebraic unary operation on $\mathscr{R}$(S) is *converse* $\boldsymbol{R} \to \boldsymbol{R}^*$ defined by ∀x,y ∈ S: x$\boldsymbol{R}^*$y *iff* y$\boldsymbol{R}$x.

Binary relations are defined on the set S, but they generate binary relations on $\boldsymbol{2}^S$, the *power set of S* (set of all subsets of S): $\boldsymbol{2}^S$ = {A: A ⊆ S}. For instance, we can consider relations $\boldsymbol{R}^a$ and $\boldsymbol{R}^e$ on $\boldsymbol{2}^S$ defined by: ∀A ⊆ S: $\boldsymbol{R}^e$(A) = {y ∈ S: ∀x ∈ A: x$\boldsymbol{R}$y}, ∀A ⊆ S: $\boldsymbol{R}^e$(A) = {y ∈ S: ∃y ∈ A: x$\boldsymbol{R}$y}.

The definitions can be expressed in words that the subset Ra(A) of S consists of all elements in relation $\boldsymbol{R}$ with all elements of A, while the subset $\boldsymbol{R}^e$(A) of S consists of all elements in relation $\boldsymbol{R}$ with at least one of elements of A (this explains letters "a" and "e" in symbols $\boldsymbol{R}^a$(A) and $\boldsymbol{R}^e$(A), since "a" stands for "all", "e" for "exists").

For one-element sets the two corresponding sets coincide, so we can simplify our notation for single element subsets: $\boldsymbol{R}$(x) = $\boldsymbol{R}^a$({x}) = $\boldsymbol{R}^e$({x}).

Obviously: $\boldsymbol{R}^a$(A) = ∩{$\boldsymbol{R}$(x): x ∈ A} and $\boldsymbol{R}^e$(A) = ∪{$\boldsymbol{R}$(x): x ∈ A}.

Now we can distinguish the following classes of binary relations of special interest for us defined by conditions:

- $\boldsymbol{R}$ is *symmetric* if $\boldsymbol{R} = \boldsymbol{R}^*$,

- $\boldsymbol{R}$ is *reflexive* if $\boldsymbol{E} \leq \boldsymbol{R}$,
- $\boldsymbol{R}$ is *transitive* if $\mathbf{R^2} = \boldsymbol{RR} \leq \boldsymbol{R}$,
- $\boldsymbol{R}$ is *antisymmetric* if $\boldsymbol{R} \wedge \boldsymbol{R}^* \leq \boldsymbol{E}$,
- $\boldsymbol{R}$ is *weakly reflexive* if $\forall x \in S$: $(x\boldsymbol{R}^c x \Rightarrow \forall y \in S$: $x\boldsymbol{R}^c y)$,
- $\boldsymbol{R}$ is a *function* if $\forall x \in S\ \exists y \in S$: $x\boldsymbol{R}y$ & $\forall x \in S \forall y_1, y_2 \in S$: $\{y_1, y_2\} \subseteq \boldsymbol{R}(x) \Rightarrow y_1 = y_2$,
- $\boldsymbol{R}$ is a *surjective function* if it is a function and $\boldsymbol{Re}(S) = S$,
- $\boldsymbol{R}$ is an *injective function* if it is a function and $\forall y \in S \forall x_1, x_2 \in S$: $\{x_1, x_2\} \subseteq \boldsymbol{R}^*(y) \Rightarrow x_1 = x_2$,
- $\boldsymbol{R}$ is a *bijective function* if it is a surjective and injective function.

In the following part of the paper we will refer to relations not only on a given set S, but also to relations on its power set $\mathbf{2^S} = \{A\colon A \subseteq S\}$. Since we consider both the sets of objects associated with elements of S, as well as the sets of properties characterizing objects associated with subsets of S, this interest in the interdependence of relations at the two levels of set-theoretical hierarchy is natural.

One of the types of structures defined on the power set $\mathbf{2^S}$ of S of special interest for us called closure space is defined in one of the crypto-morphically equivalent ways by a family of subsets $\mathscr{M}$ satisfying conditions: Entire S is in $\mathscr{M}$, and together with every subfamily of $\mathscr{M}$, its intersection belongs to $\mathscr{M}$, i.e., $\mathscr{M}$ is a *Moore family*. It is easy to see that we can define this structure by a *closure operator* defined on S (i.e., a function f on the power set $\mathbf{2^S}$ of a set S such that:

(1) For every subset A of S, $A \subseteq f(A)$;
(2) For all subsets A, B of S, $A \subseteq B \Rightarrow f(A) \subseteq f(B)$;
(3) For every subset A of S, $f(f(A)) = f(A)$).

The Moore family $\mathscr{M}$ of subsets is simply the family f-Cl of all closed subsets, i.e., subsets A of S such that $A= f(A)$. The family of closed subsets $\mathscr{M}$ = f-Cl is equipped with the structure of a complete lattice L_f by the set-theoretical inclusion. The mutual relationship between the two cryptoisomorphic implementations of the closure space leads back from the Moore family to the closure operator f by the following: For every subset A of S: $f(A)= \cap\{B \in \mathscr{M}\colon A \subseteq B\}$.

The Moore family $\mathscr{M}$ or alternatively corresponding closure operator f, with some additional axioms describing properties of closure operator f (or alternatively additional conditions for the family $\mathscr{M}$), can represent a very large variety of structures of a particular type (e.g., geometric, topological, algebraic, logical, etc.) defined on the subsets of S. Terminology of the theory of closure spaces was adopted from some of the paradigmatic instances of topological spaces and vector spaces in which closed subspaces are simply vector subspaces.

In addition to the family of closed subsets f-Cl we can distinguish some other families of special importance:

- f-Ind = $\{A \subseteq S\colon \forall x \in A\colon x \notin f(A \setminus \{x\})\}$—the family of independent subsets of S,
- f-Gen = $\{A \subseteq S\colon f(A)=S\}$—the family of generating subsets of S,
- f-Base = f-Ind $\cap$ f-Gen—the family of bases.

Unlike what is well known in the special case of vector space closure space, the last family may be empty in general, i.e., some closure spaces do not have bases. The absence of bases may be a significant drawback, as bases are minimal generating subsets.

The same way as we considered above a closure space structure on S as a relation on its power set $\mathbf{2^S}$ we can and we will consider a closure space on $\mathbf{2^S}$ as a relation on its own power set.

5.2. Mathematical Formalism for Equality, Equivalence, and Similarity

Now we can focus our attention on the relations that are subject of this study. We already have distinguished our equality relation $\boldsymbol{E} = \{(x,y)\colon x = y\}$. *Equivalence relations* are defined as those which are reflexive, symmetric and transitive, conditions which combined can be written: $\boldsymbol{E} \leq \boldsymbol{R}^* = \boldsymbol{R} = \boldsymbol{R}^2$.

Of course, $\mathbf{E} \leq \boldsymbol{E}^* = \boldsymbol{E} = \boldsymbol{E}^2$, so equality is a special case of equivalence relation (the least equivalence relation on S).

It is a very elementary fact that equivalence relations correspond in a bijective manner to partitions of the set S on which they are defined. Subsets belonging to such partition $\mathfrak{C} \subseteq \mathbf{2^S}$ (i.e., family $\mathfrak{C}$ which satisfies the conditions $\cup\mathfrak{C} = S$ and $\forall A,B \in \mathfrak{C}$: $A \cap B = \varnothing$) are called classes of equivalence for the corresponding relation. If we start from a partition $\mathfrak{C}$, its corresponding equivalence relation is defined by the condition that the elements x and y are related, i.e., $x\boldsymbol{R}y$ if they both belong to one of the subsets of the partition ($x\boldsymbol{R}y$ *iff* $\exists A \in \mathfrak{C}$: $\{x,y\} \subseteq A$). If we start from the equivalence $\boldsymbol{R}$, the partition is uniquely determined by the condition $A \in \mathfrak{C}$ iff $A = \boldsymbol{R}^a(A)$.

Equivalence relations are very simple but of extremely high importance in mathematics, as they are involved in the process of abstraction understood as a transition from elements of a set S to the elements of the partition associated with this equivalence.

As it was extensively discussed before, *tolerance relations* are more general, because they do not have to be transitive, i.e., they are defined by $\boldsymbol{E} \leq \boldsymbol{T}^* = \boldsymbol{T}$. For the reason which soon will become clear it is worth considering one small step in generalization to *weak tolerance relations* which are simply symmetric ($\boldsymbol{T}^* = \boldsymbol{T}$) and which are *weakly reflexive* ($\forall x \in S$: ($x\boldsymbol{T}^c x \Rightarrow \forall y \in S$: $x\boldsymbol{T}^c y$)). Originally, the latter condition of weak reflexivity appeared in this theory because it is important in the study of more general mathematical structures which are not reflexive, but which satisfy this weaker form [13]. However, the consequences of the generalization have clear importance for our study, as will be shown later.

It turns out that an arbitrary covering of the set S (family of subsets $\mathscr{H} \subseteq \mathbf{2^S}$ which satisfies the condition $\cup\mathscr{H} = S$) defines a tolerance relation on S the same way as partitions defined equivalence relations, i.e., by: $x\boldsymbol{T}y$ *iff* $\exists A \in \mathscr{H}$: $\{x,y\} \subseteq A$. However, we do not have bijective correspondence as before. Different coverings can define the same tolerance relation and the relation between coverings and tolerance relations is highly nontrivial in comparison to the special case of equivalence relations.

Suppose we have a tolerance relation $\boldsymbol{T}$ on S. We can define a family of subsets $\mathscr{H}_T = \{A \subseteq S$: $\forall x,y \in S$: $\{x,y\} \subseteq A \Rightarrow x\boldsymbol{T}y\}$. This class will be called the family of all *pre-classes of tolerance* $\boldsymbol{T}$. Of course, $\forall x,y \in S$: $x\boldsymbol{T}y$ *iff* $\exists A \in \mathscr{H}_T$: $\{x,y\} \subseteq A$, but it is clear that this family is redundant. If $\boldsymbol{T}$ is an equivalence, then $\mathscr{H}_T$ in addition to all members of the family of equivalence classes $\mathfrak{C}$ includes all their subsets. Therefore, we want to reduce $\mathscr{H}_T$ as much as possible. Using Zorn's lemma, we can infer that in $\mathscr{H}_T$ every pre-class A can be extended to a maximal pre-class, which we will call a *class of tolerance relation*. The subfamily $\mathfrak{C}$
textsubscriptT of all classes of tolerance is sufficient for the reconstruction of tolerance $\boldsymbol{T}$: $\forall x,y \in S$: $x\boldsymbol{T}y$ *iff* $\exists A \in \mathfrak{C}_T$: $\{x,y\} \subseteq A$. So, we have an efficient way to represent given tolerance relation by its family of tolerance classes.

We still do not know how to recognize coverings which are families of pre-classes for some tolerance relation. For this purpose, we have to introduce the structure of closure space on $\mathbf{2^S}$, i.e., a relation on the power set of $\mathbf{2^S}$ or on the power set of the power set. We define the following closure operator: $\forall \mathscr{B} \subseteq \mathbf{2^S}$: $f(\mathscr{B}) = \{A \subseteq S$: $\forall x,y \in A \exists B \in \mathscr{B}$: $\{x,y\} \subseteq B\}$. Now we can characterize coverings which form the family of all pre-classes for some tolerance relation $\boldsymbol{T}$, as coverings which are closed with respect to this closure operator (covering $\mathscr{B}$ of S is a family $\mathscr{H}_T$ of all pre-classes of tolerance $\boldsymbol{T}$ on S *iff* $f(\mathscr{B}) = \mathscr{B}$). Moreover, we have a bijective correspondence between tolerance relations and f-closed coverings. Since this is only an overview of the theory of tolerance relations, we will not enter the issue of optimization, i.e., finding minimal families of subsets of S uniquely representing tolerance relation $\boldsymbol{T}$, which can be achieved using f-bases introduced in the preliminaries. However, such bases do not always exist, and only in the case of finite set S we can always minimize the family of subsets representing tolerance. Moreover, the minimal families may have different cardinality (size).

This is the result anticipated in the introduction to this paper that the families of sets representing a tolerance relation (which of course correspond to predicates describing attributes) may be different and of different size. Any measure of similarity derived from the number of shared sets from the family or in other words shared attributes may depend on the arbitrary decision of the choice of family, and therefore it is meaningless.

Another topic is the analysis of tolerance relation from the point of view of deviation from equivalence relation. For this purpose, we can consider the nucleus N_T of T defined as an equivalence relation: $\forall x,y \in S$: x N_T y *iff* T(x) = T(y). Of course, if T itself is an equivalence relation, then its nucleus is identical with itself, i.e., N_T = T. Otherwise the nucleus partitions S into subsets in which all elements are in relation T with each other. If all equivalence classes of nucleus N_T consist of only one element, i.e., N_T = E (equality relation), the tolerance is *non-nuclear*. This type of tolerance corresponds to Wittgenstein's concept of pure family resemblance. Every tolerance relation on set S can be mapped on the non-nuclear tolerance relation defined on the set of class of abstractions of nucleus N_T. Moreover, every tolerance relation T can be constructed from some equivalence relation R (playing the role of the nucleus N_T) and some non-nuclear tolerance relation $T^{\#}$. In this sense, we can say that similarity can be "resolved" into equivalence and family resemblance. This is the result which, as it was stated in the introduction, in clear contrast to the "resolution into identity and difference" proposed by Hesse.

Now we can explain what compels us to consider yet another level of generalization to weak tolerance relations. It turns out that a very similar theory can be developed for weak tolerances, but in this generalization, we do not need to restrict the families of sets to coverings (their union can be a proper subset of the set S). We get much simpler and clear correspondence between weak tolerances on S and closed subfamilies of its power set with respect to the closure operator f defined before as expressed by the proposition [13]:
There is a bijective correspondence between weak tolerance relations on set S which generalize equivalence relations extending them to a general concept of similarity and closed subsets of the closure operator f on the power set of S, i.e., closure space $<2^S,f>$ defined by:$\forall \mathscr{B} \subseteq 2^S$: $f(\mathscr{B}) = \{B \subseteq S: \forall x,y \in B\ \exists A \in \mathscr{B}: \{x,y\} \subseteq A\}$.

We can use the theory of tolerance and weak tolerance relations to relate more extensive class of relations with those two by a process of "symmetrization". For every binary relation $\boldsymbol{R}$ on a given set S, we can define a relation $\boldsymbol{T_R}$ as follows: $\boldsymbol{T_R} = \boldsymbol{RR^*}$. Then we have [13]:

(i) $\boldsymbol{T_R}$ is a symmetric relation on S.
(ii) $\forall x \in S \forall y \in S$: x $\boldsymbol{T_R}$ y *iff* $R(x) \cap R(y) \neq \emptyset$.
(iii) $\boldsymbol{T_R}$ is a tolerance *iff* $\boldsymbol{R}$ is defined everywhere (equivalent to $\boldsymbol{E} \leq \boldsymbol{RR^*}$).
(iv) $\boldsymbol{T_R}$ is a weak tolerance *iff* $\boldsymbol{R}$ is *weakly reflexive*, i.e., $\forall x \in S$: $(x\boldsymbol{R^c}x \Rightarrow \forall y \in S: x\boldsymbol{R^c}y)$.
(v) $\boldsymbol{R}$ is a function $\Rightarrow$ $\boldsymbol{T_R}$ is an equivalence relation, but the reverse implication is not necessarily true.
(vi) $\boldsymbol{T}$ is an equivalence relation *iff* there exists a relation $\boldsymbol{R}$ which is a function and $\boldsymbol{T} = \boldsymbol{T_R}$.
(vii) Let $\boldsymbol{E} \leq \boldsymbol{T}$. Then T is an equivalence relation *iff* $\boldsymbol{T} = \boldsymbol{T_T}$.

This proposition links together the four types of relations: Equality, equivalence, tolerance, and weak tolerance with each other and with the very general class of weakly reflexive relations. We can observe that the class of equivalence relations is here associated with functions, which in turn are the most typical instruments of the mathematical formalization of theories across all disciplines.

Thus far, we considered the process generating the relations describing different levels of similarity from the weakly reflexive relations on a given set S. Now we will consider induction of these types of relations on the power set $\mathbf{2^S}$ of S (the set of all subsets of S) by the relations on S introduced already by Zeeman.

Let $\boldsymbol{R}$ be a relation on S. Then we define a relation $\boldsymbol{R^S}$ on $\mathbf{2^S}$ as follows:

$$\forall A \subseteq S \forall B \subseteq S: A\ \boldsymbol{R^S}\ B \textit{ iff } B \subseteq \boldsymbol{R^e}(A) \text{ and } A \subseteq \boldsymbol{R^e}(B).$$

Then, if $\boldsymbol{T}$ is a tolerance relation on S, then $\boldsymbol{T^S}$ is a tolerance relation on $\mathbf{2^S}$ [2]. It is also easy to show that a weak tolerance relation $\boldsymbol{T}$ induces a weak tolerance relation $\boldsymbol{T^S}$ on $\mathbf{2^S}$ and an equivalence induces equivalence. Thus, the similarity relation defined on a given set induces similarity of sets of objects. We can consider the reversed process of rather trivial "downward induction" from the power set of a set S to S when we consider the definition of $\boldsymbol{R^S}$ on $\mathbf{2^S}$ restricted to one-element sets. Obviously, if we start from the induction and proceed to the downward induction, we return to the original tolerance relation.

6. Interpretation and Consequences of the Mathematical Formalism

The short overview of the mathematical formalism for similarity demonstrates once more the curious feature of mathematical theories whereby more general objects of study defined by simplified or reduced conditions acquire more complex description. A relatively simple idea of the unique family of equivalence classes ramifies into several different families with extreme instances of pre-classes and classes of tolerance together with many intermediate types. In case of a similarity relation which is not an equivalence relation (associated with equivalence classes of items which have them or in other words which are described by them), we have many different complexes of properties defining them instead of the unique system of properties.

Mathematical formalism gives us analytical tools to distinguish nontrivial types of similarity in the strongest contrast to the familiar type of equivalence (described by non-nuclear tolerance relations). More generally, we can assess the degree of separation of a given tolerance (similarity) relation from an equivalence type through the examination of its nucleus.

Someone could object the identification of similarity understood as a tolerance relation with analogy. After all, analogy is being understood in so many different ways. However, no matter what additional conditions are imposed on analogy in the literature (with the few exceptions which were critically reviewed in the introduction as having hidden additional assumptions interpreted as asymmetry of similarity), it is always conceptualized as a symmetric relation between analogs and, at least directly, this relation is rarely denied reflexivity outside of mathematics (each item is trivially an analog of itself). To accommodate the mathematical understanding of similarity, we can slightly generalize reflexivity to weak reflexivity and tolerance relations to weak tolerance relations. Thus, no matter what the preferred definition may be, for instance, additional conditions supplementing those which define weak tolerance, or tolerance relation, it can be described and analyzed within this type of relations.

At this point, it is appropriate to explain the meaning and importance of further generalizations of similarity to weak tolerance relation. At first sight, the condition of weak reflexivity may seem absurd. How can something be not similar to itself? If we think about similarity described by properties, a closer look shows that the case is not so strange. We have to consider the possibility that the ascription of a property to some objects can be not true or not false, but meaningless. Does number 5 have a smell? Does it make sense to say that number 5 smells like itself? Of course not! Whatever the other elements of the set S whose element is number 5 may include, when we consider similarity with respect to smell, we should exclude number 5 from pre-classes of the weak tolerance describing similarity. In this way, there is nothing strange in this generalization, which offers us a wonderful completion of the formalism in the proposition ending preceding section. We have a full description of all similarities as closed subfamilies with respect to a well-defined closure operation.

An equivalence relation is the basic tool of abstraction; a transition from the objects of the lower level of abstraction to the higher one implemented by the transition from the elements of set S to classes of equivalence, sometimes called classes of abstraction. In a similar way, we can consider tolerance relation (or relation of similarity) as the basic tool of analogy. Their common mathematical formalism permits their mutual comparative analysis, which in turn gives us an insight into the comparison between abstraction and analogy.

One of the most unexpected consequences of this comparison is that analogy is much more complex than abstraction. Now we can fully appreciate and share Wittgenstein's fascination with family resemblances. But this complexity comes without the luggage of limited methods of analysis.

Analogy does not belong anymore to vague concepts suitable for non-scientific or at most folk-scientific discourse guided by common sense.

Tolerance relations and weak tolerance relations have extensive, but rather esoteric mathematical literature with the possible exception of more accessible mathematical linguistics. However, in the finite case (when the fundamental set S is finite) there is a well-established link between their theory and graph theory [13]. This connection opens a vast resource to the study of analogy.

The study of tolerance (or weak tolerance) relations gives as tools to analyze similarity at two levels. The process of induction and downward induction described in the preceding section shows the close correspondence between the similarity of the objects and similarity of their sets. This shows how we can make a transition between similarity of objects and similarity of predicates applied to these objects. Typically, the similarity is defined and analyzed exclusively on one or the other level of abstraction.

Finally, we can ask whether mathematics can benefit from the study of analogy. My own expectation is that a better understanding of analogy can help us to overcome the impasse in the study of the general concept of a structure, thereby going beyond familiar relational structures (sets equipped with one or multiple relations of arbitrary finite n-arity). The Bourbaki-style approach is very ineffective, being practically obsolete, and newer approaches do not resolve the issue which doomed the original effort of the French group of mathematicians. We usually have many apparently "equivalent" ways to define structures, which frequently turn out to be not entirely equivalent (especially if we do not make some hidden assumptions). These different definitions provide cryptomorphic versions of the structure, which have all features of analogy rather than equivalence. Resolving the problem analogy would pay the debt to mathematics for its similarity formalism.

Funding: This research received no external funding.

Acknowledgments: I would like to express my gratitude to my university colleague Leigh Bennett for multiple suggestions how to improve the language and the style of presentation in my paper. Also, I would like to thank anonymous reviewers of the early version of my paper for their critical remarks and especially one of them who wrote many useful and valuable comments and suggestions for the improvement of the language and content of the manuscript. Of course, I am responsible for all remaining deficiencies of the text.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Bartha, P. Analogy and Analogical Reasoning. *The Stanford Encyclopedia of Philosophy*. Zalta, E.N., Ed.; Winter 2016 ed. Available online: https://plato.stanford.edu/archives/win2016/entries/reasoning-analogy/ (accessed on 3 October 2018).
2. Zeeman, E.C. The topology of the brain and visual perception. In *Topology of 3-Manifols and Related Topics, Proceedings of The University of Georgia Institute 1961*; Fort, M.K., Jr., Ed.; Prentice-Hall: Englewood Cliffs, NJ, USA, 1962; pp. 240–256.
3. Zeeman, E.C.; Buneman, O.P. Tolerance spaces and the brain. In *Towards a Theoretical Biology*; Waddington, C.H., Ed.; Edinburgh University Press: Edinburgh, UK, 1968; pp. 140–151.
4. Heidegger, M. *Identity and Difference*; Stambaugh, J.J., Ed.; Harper & Row: New York, NY, USA, 1969.
5. Schroeder, M.J. Theoretical Study of the Concepts of Information Defined as Difference and as Identification of a Variety. In *Philosophy and Methodology of Information—The Study of Information in a Transdisciplinary Perspective*; Dodig-Crnkovic, G., Burgin, M., Eds.; World Scientific Series in Information Studies; World Scientific: Singapore, 2019; Volume 10, pp. 289–314.
6. Tversky, A. Features of Similarity. *Psychol. Rev.* **1977**, *84*, 327–352. [CrossRef]
7. Gentner, D. Structure-Mapping: A Theoretical Framework for Analogy. *Cogn. Sci.* **1983**, *7*, 155–170. [CrossRef]
8. Hesse, M. *Models and Analogies in Science*; Sheed and Ward: London, UK, 1963.
9. Goodman, N. Seven Strictures on Similarity. In *Problems and Projects*; Goodman, N., Ed.; Bobs-Merril: Indianapolis, IN, USA, 1972; pp. 22–32.
10. Watanabe, S. *Pattern Recognition: Human and Mechanical*; Wiley: New York, NY, USA, 1985.

11. Birkhoff, G. *Lattice Theory*, 3rd ed.; American Mathematical Society Colloquium Publications: Providence, RI, USA, 1967; Volume XXV.
12. Ross, W.D. (Ed.) *Aristotle: Selections*; Charles Scribner Sons: New York, NY, USA, 1955.
13. Schroeder, M.J.; Wright, M.H. Tolerance and weak tolerance relations. *J. Combin. Math. Combin. Comput.* **1992**, *11*, 123–160.
14. Douven, I.; Decock, L. Identity and similarity. *Philos. Stud.* **2010**, *151*, 59–78. [CrossRef]
15. Schroeder, M.J. Ontological Study of Information: Identity and State. *Kybernetos* **2014**, *43*, 882–894. [CrossRef]
16. Schroeder, M.J. Invariance as a Tool for Ontology of Information. *Information* **2016**, *7*, 11. [CrossRef]
17. Tarski, A. *Introduction to Logic and to the Methodology of Deductive Sciences*; Helmer, O., Ed.; Dover: Mineola, NY, USA, 1995.
18. Loemker, L., Translator; *G. W. Leibniz: Philosophical Papers and Letters*, 2nd ed.; D. Reidel: Dordrecht, The Netherlands, 1969.
19. St. Thomas Aquinas THE SUMMA THEOLOGICA Translated by Fathers of the English Dominican Province: L.40, C.4. Question 40, Article 1, Objection 3. Available online: www.documenta-catholica.eu/d_1225-1274-%20Thomas%20Aquinas%20-%20Summa%20Theologiae%20-%20Prima%20Pars%20-%20EN.pdf (accessed on 19 May 2019).
20. Hopkins, J. *Nicholas of Cusa On God as Not-Other: A Translation and Appraisal of De Li Non Aliud*, 3rd ed.; The Arthur J. Banning Press: Minneapolis, MN, USA, 1987.
21. Kant, I. *Critique of Pure Reason*; Meiklejohn, J.M.D., Ed.; Henry G. Bohn: London, UK, 1855.
22. Luria, A.R. Towards the Problem of the Historical Nature of Psychological Process. *Int. J. Psychol.* **1971**, *6*, 259–272. [CrossRef]
23. Greenberg, J.H. The nature and uses of linguistic typologies. *Int. J. Am. Linguist.* **1957**, *23*, 68–77. [CrossRef]
24. Conklin, H.C. Hanuno's color categories. *Southwest. J. Anthropol.* **1955**, *11*, 339–344. [CrossRef]
25. Goodenough, W.H. Componential analysis and the study of meaning. *Language* **1956**, *32*, 195–216. [CrossRef]
26. Lounsbury, F.G. A semantic analysis of the Pawnee kinship usage. *Language* **1956**, *32*, 158–194. [CrossRef]
27. Sturtevant, W.C. Studies in ethnoscience. *Am. Anthropol. (Spec. Publ.)* **1964**, *66*, 99–131. [CrossRef]
28. Frake, C.O. Cultural Ecology and Ethnography. *Am. Anthropol.* **1962**, *64*, 53–59. [CrossRef]
29. Poincaré, M.H. Sur certains surfaces algébriques. Troisième complément à l' <<Analysis sitûs>>. *Bulletin de La S. M. F.* **1902**, *30*, 49–70.
30. Wittgenstein, L. *Philosophical Investigations*; Anscombe, G.E.M., Ed.; Basil Blackwell: London, UK, 1986.
31. Needham, R. Polythetic classification: Convergence and consequences. *Man* **1975**, *10*, 349–369. [CrossRef]
32. Nietzsche, F. *Jenseits von Gut und Bose: Vorspiel einer Philosophie der Zukunz*; Neumann: Leipzig, Germany, 1886.
33. Bambrough, R. Universals and family resemblances. *Proc. Aristot. Soc.* **1961**, *60*, 207–222. [CrossRef]

© 2019 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).