A Topological Analysis of Classification Systems

Wilson Poulter

March 15, 2021

1 Introduction

Classification may be described as a process by which the abstract space known as the *Universe of Knowledge* (UK) is given a cartographic structure [1–3]. One might think of this as similar to the way that the Earth and its various localities have been mapped out, as seen in atlases and other formats. A purpose of this mapping is to provide the means by which one can cognitively traverse the UK with some degree of orientation or spatiality, placing documents/concepts with similar content near to one another.

In mathematics, a similar method of inducing spatiality on a structureless data-set is accomplished through the construction of what is called a topology. Though I will define it precisely in §2, a topology is at its essence a collection of special subsets of a data-set (called *open sets*) whose members are in some sense near to one another. In this way, the open sets of a topological space resemble the categories of a classification scheme, with the documents/concepts taking the place of the data point.

In this essay, I introduce various topological concepts to the audience of classification theorists, and propose that the analogy described above provides an opportunity for classification theorists to study the structures of classification systems using the expansive tools that the mathematical subject of topology provides. For purposes of demonstration, I analyze what I call *the shelving problem* using a topological analysis. Following this analysis, I describe a few other problems that I believe this analysis invites and potentially resolves.

2 Topological Prerequisites

In this section, the various topological concepts that will be used in the sections that follow are discussed, following [4]. The reader with a background in formal topology should feel empowered to proceed to the next section. For those who may require an introduction to the basic concepts of set theory as well, see Appendix A. Further reading on set-theoretic concepts here and beyond can be found in [5].

To begin, we formally define the primary concern of mathematical topology, the *topological space*:

Definition 2.1. Let X be a set, and $\mathcal{T}_X$ a subset of $\mathcal{P}(X)$. We say that $\mathcal{T}_X$ is a *topology on X* if it satisfies the three following conditions:

- the empty set and X are both members of $\mathcal{T}_X$,
- all arbitrary unions of members of $\mathcal{T}_X$ are themselves members of $\mathcal{T}_X$, and
- all finite intersections of members of $\mathcal{T}_X$ are themselves members of $\mathcal{T}_X$.

If $\mathcal{T}_X$ is a topology on X, then we say the pair $(X, \mathcal{T}_X)$ is a *topological space*.

Given a topological space $(X, \mathcal{T}_X)$, we say that the members of $\mathcal{T}_X$ are the *open sets of* X, often using the capital letters U, V, and W to denote them (although this is not necessary). When it is non-obstructive to discussion, we will simply denote topological spaces by their underlying set X without reference to their topology $\mathcal{T}_X$, as is common in practice. To provide some semantic meaning to the highly formal notion of an open set, if two members x and y of X belong to an open set, then there are in some sense near to one another. The more open sets that x and y both belong to, the nearer they are. In fact, if x and y occupy all of the same open sets, then x and y are spatially indistinguishable. This notion of nearness also allows for the formal discussion of continuity in terms of mappings and functions:

Definition 2.2. Let X and Y be topological spaces, and let $f : X \to Y$ be a function. We say that f is *continuous* if for all open subsets U of Y, the inverse image of U under f is an open subset of X. That is, for all $U \in \mathcal{T}_Y$, $f^{-1}(U) \in \mathcal{T}_X$ as well.

Essentially, this definition claims that a mapping between two spaces is continuous if points nearby one another in the domain are mapped to points nearby one another in the codomain. On the graph of a function represented in the Cartesian plane, this would mean that there are no breaks in the line that represents the graph. The following theorem shows that two continuous processes that follow one after the other combine to create one continuous process. Proof is deferred to Appendix B.

Theorem 2.3. Let X, Y, and Z all be topological spaces. Suppose that $f : X \to Y$ and $g : Y \to Z$ are both continuous functions, then $(g \circ f) : X \to Z$ is continuous as well.

The above property of continuous maps allow us to construct what is a more general construction in mathematics called a *category*, which is simply a collection of spaces and a class of maps between these spaces, with certain properties. An important notion in an category is the equivalence of objects, which is determined by special maps in the category. We define these special maps and this equivalence now in turn.

Definition 2.4. Let X and Y be topological spaces, and let $f : X \to Y$ be a function. We say that f is a *homeomorphism* if f has an inverse $f^{-1} : Y \to X$, and both f and f^{-1} are continuous.

Note that if a function f is a homeomorphism, then it creates a one-to-one correspondence between the open sets of X and Y that respects union and intersection, meaning that X and Y are formally the same space. This leads to the definition below:

Definition 2.5. Let X and Y be topological spaces. We say that X and Y are *topologically equivalent* if there exists some function $f : X \to Y$ such that f is a homeomorphism.

We now take a moment to note that this definition of equivalence is truly an equivalence relation, in the sense that it is *reflexive*, *symmetric*, and *transitive.* Proof is deferred to Appendix B.

Theorem 2.6. The relationship of topological equivalence on topological spaces is an equivalence relation.

The following construction allows us to construct a topology from an selection of sets that may be of interest. This will prove invaluable in what follows, as we must construct topologies from data-sets that do not already have one.

Definition 2.7. Let X be a set, and let $\mathcal{S}$ be a subset of $\mathcal{P}(X)$. We may construct a topology on X determined by $\mathcal{S}$ as follows:

1. construct a collection $\mathcal{B}$ whose membership is defined by the subsets given by finite intersections of members of $\mathcal{S}$, then

2. construct a topology $\mathcal{T}_X$ on X whose membership is defined by the subsets given by arbitrary unions of members of $\mathcal{B}$.

$\mathcal{S}$ and $\mathcal{B}$ are called the subbase and base of $\mathcal{T}_X$, respectively. We say that $\mathcal{T}_X$ is generated from the subbase $\mathcal{S}$ and the base $\mathcal{B}$ respectively.

Lastly, these two definitions and theorem will be useful as every topology we construct will have the properties defined in Definition 2.9, as shown in Theorem 3.3.

Definition 2.8. Let X be a topological space. We say that a collection of open subsets $\mathcal{U}$ is a cover if for every $x \in X$, there is some open subset $U \in \mathcal{U}$ such that $x \in U$.

Definition 2.9. Let X be a topological space. We say that X is *compact* if for every open cover $\mathcal{U}$ of X, there is a finite subcollection $\mathcal{V}$ of $\mathcal{U}$ such that $\mathcal{V}$ is still an open cover.

Essentially, a compact topological space is one whose members are in some manner finitely close to one another. Put another way, compact spaces are small, while non-compact spaces are big. This last theorem shows that compactness is a property that is preserved by continuous functions. See Appendix B for proof.

Theorem 2.10. Let X and Y be topological spaces, let X be compact, and let $f : X \to Y$ be a continuous and surjective function, then Y is also compact.

3 Constructing the Classification Topology

A classification system organizes documents/concepts into a structure that through syntactic and semantic means allows users to navigate the UK. Without the intention of claiming that semantic aspects of classification systems are unimportant [6], we turn our attention to purely syntactical aspects here. By this, I mean that for a moment, we are going to forget the meaning that is endowed in a classification system, and look at the skeleton that is left behind when doing so. Hence, the framework that I introduce in this section is not a complete framework for the study of classification systems. That said, the skeleton is as much a part of the body as the flesh, and I suggest that study of this skeleton is illuminating not only for itself, but for the body as a whole, as well as the flesh on its own.

Mathematics is a tool that allows for the thorough analysis of syntactical structures. Moreover, the syntactical structure of a classification system resembles the syntactical structure of a topological space. Thus, in the spirit of [7], we define here a methodology by which every classification system may be endowed with a topological structure, and thus studied via mathematical analyses.

In the discussion that follows, we consider the UK as an infinite set denoted by $\mathbb{U}$, though we do not specify its cardinality. The definition below sets out the formal assumptions that we are making about our classification systems in relation to the UK.

Definition 3.1. A *practical classification system* C is a subset of $\mathcal{P}(\mathbb{U})$ with the following properties:

- C is finite,
- the members of C form a cover of $\mathbb{U}$, and
- the members of C are infinite.

We call the members of C the *classes* of C.

Less formally, this definition assumes that a practical classification system has a finite number of classes, that these classes are adequate to classify all knowledge, and that there is room in every class for an infinite number of documents/concepts. We can imagine familiar classification systems as practical classification systems as I have defined above. For example, the Dewey Decimal Classification (DDC) system's notations each represent a class, these classes are ostensibly exhaustive, and moreover, there is no reason to assume the classes could not contain an infinite number of documents/concepts.

Definition 3.2. Let C be a practical classification system. We define the *classification topology* of C, denoted by $\mathcal{C}$, as the topology on $\mathbb{U}$ generated by the subbase C, as defined in Definition 2.7.

It should be noted that while the classification topology of a practical classification system seemingly adds more information than is necessary to the system, because the topology has been generated from the practical classification system's classes, this is merely a transformation of information. Indeed, it is a transformation that allows us to expand our vocabulary and analyses regarding classification systems, using those of mathematical topology.

As an example, I show that the classification topologies are topologically compact.

Lemma 3.3. Let $\mathcal{C}$ be a classification topology arising from a practical classification system C. Then $\mathcal{C}$ is topologically compact.

Proof. This result is trivial when one realizes that $\mathcal{C}$ is finite. Indeed, since C is finite, there is a number N such that $|C| = N$. Thus, there are 2^N possible intersections of classes in C, and thus 2^{2^N} possible unions of these intersections in turn. Thus, the order of $\mathcal{C}$ is bounded by 2^{2^N}, hence it is finite. Thus, if $\mathcal{U}$ is a cover of $\mathbb{U}$ in the topology $\mathcal{C}$, then $\mathcal{U}$ is finite, and thus its own finite subcover. It follows from definition that $\mathcal{C}$ is compact. □

4 The Shelving Problem

Consider a shelf in a library. Abstractly, it organizes documents/concepts through its *shelf order*, that is, by the order in which the documents/concepts appear on the shelf, going from left to right. Intuitively, documents/concepts on the shelf are more similar the closer they are to one another, or at least that is the ideal.

In reality, every classification theorist and cataloger knows that the shelf ordering is a lousy way to physically manifest a conceptual system like the DDC. In light of this, I in no way claim novelty to the results of this section. In what follows, I will define the topology that is inherent to the shelf ordering, and use this to discuss formally what everyone knows already about the problematic nature of the shelf order. I have chosen this problem purposely because I see it as well enough understood to clarify some of topological notions that have been discussed, and to provide a simple example of their application.

To begin, I will define the shelf ordering in a similar way to which I defined practical classification systems above.

Definition 4.1. A shelf-ordering is an unbounded dense ordering $<$ on $\mathbb{U}$. In other words, the relation $<$ on $\mathbb{U}$ has the following properties:

- transitivity,
- anti-reflexivity,
- anti-symmetry,
- linearity, i.e., for all $x, y \in \mathbb{U}$, either $x < y$ or $y < x$,
- density, i.e., for all $x, y \in \mathbb{U}$ such that $x < y$, there is some $z \in \mathbb{U}$ such that $x < z < y$, and

- unboundedness, i.e., for all $x \in \mathbb{U}$ there exist $y, z \in \mathbb{U}$ such that $y < x < z$.

We may think of the notation $x < y$ as informally saying that x sits to the left of y on the shelf. Thus, a shelf ordering is an arrangement of documents/concepts belonging to the UK based on horizontal positioning. The first three properties are easy enough to see: (transitivity) if a member x sits to the left of a member y that sits to the left of a member z, then x sits the left of z. (anti-reflexivity) No member of the UK sits to the left of itself, (anti-symmetry) no two members sit simultaneously to the left of one another, and (linearity) every pair of members can be oriented left to right. The property of density means that there is always another document/concept that sits between two, meaning that there is ideally always a document/concept that sits between two others (perhaps a document/concept that combines the two originals). Lastly, the property of unboundedess states similarly that there is no ideal first or last document/concept on the shelf.

We use this ordering to define a topology on $\mathbb{U}$:

Definition 4.2. Given a shelf ordering $<$, the *shelf topology* $\mathcal{S}_<$ is the topology generated by the subbase whose members are of the following two forms:

- $\{x : x < x_0 \text{ for } x \in \mathbb{U}\}$ for some $x_0 \in \mathbb{U}$, and
- $\{x : x_0 < x \text{ for } x \in \mathbb{U}\}$ for some $x_0 \in \mathbb{U}$.

Unlike the practical classification systems, a shelf topology is not compact.

Lemma 4.3. A shelf topology $\mathcal{S}_<$ on $\mathbb{U}$ is *not* compact.

Proof. To show this, we construct an open cover of $\mathbb{U}$ with no finite subcover. By the unboundedness property of $<$, it follows that there exists sequences $x_1, x_2, \ldots$ and $y_1, y_2, \ldots$ in $\mathbb{U}$ with the property that

$$\cdots < x_2 < x_1 < y_1 < y_2 < \ldots$$

and for all $z \in \mathbb{U}$, there is some number i such that $x_i < z < y_i$. By the definition of $\mathcal{S}_<$, we see that the sets $U_i := \{z : x_i < z < y_i \text{ for } z \in \mathbb{U}\}$ are each open, and moreover that the collection given by $\mathcal{U} = \{U_i : i < \infty\}$ is an open cover of $\mathbb{U}$. Now, let $\mathcal{W}$ be a finite subcollection of $\mathcal{U}$. By finiteness, there is some number N such that if $U_i \in \mathcal{W}$, then $i < N$. But, then $x_N \notin U_i$ for any $U_i \in \mathcal{W}$. Thus, if $\mathcal{W}$ is a finite collection of $\mathcal{U}$, it does not cover $\mathbb{U}$. $\square$

This shows that shelf topologies and the classification topologies arising from practical classification systems are not equivalent, but as stated above, this was expected. But, I take this observation a step further and demonstrate that the process of shelving is not even a continuous process. To do so, I abstractly associate shelving with a function β that assigns every document/concept to a unique place on the shelf. I show that when this is the case, the function β cannot be continuous, that is, there are discontinuities that arise from moving between a practical classification system, and the shelf.

Theorem 4.4. Let $\mathcal{C}$ be a classification topology arising from a practical classification system C, let $\mathcal{S}_<$ be a shelf topology arising from a shelf ordering $<$, and let $\beta : (\mathbb{U}, \mathcal{C}) \to (\mathbb{U}, \mathcal{S}_<)$ be an invertible function, then β is not continuous.

Proof. Suppose towards a contradiction that β is continuous. From Theorem 2.10, it follows that $(\mathbb{U}, \mathcal{S}_<)$ is compact, since $\mathcal{C}$ is a compact topology on $\mathbb{U}$ by Lemma 3.3. But, the statement of $(\mathbb{U}, \mathcal{S}_<)$'s compactness is in direct contradiction to Lemma 4.3. $\square$

Corollary 4.5. A shelf topology arising from a shelf ordering is not topologically equivalent to any classification topology arising from a practical classification system.

Proof. Topological equivalence requires the existence of a homeomorphism, which would in turn requires some invertible and continuous function $\beta : (\mathbb{U}, \mathcal{C}) \to (\mathbb{U}, \mathcal{S}_<)$, contrary to Theorem 4.4. $\square$

5 Questions for Further Research

Having demonstrated the conceptual usefulness of formally defining and resolving the shelving problem, I now consider some further areas of research that a topological framework lends itself towards. This section should be considered as a non-exhaustive list of this methodology's potential.

5.1 Dimensional Analysis

Although the notion of dimension is often thought about in terms of geometric systems that bare resemblance to Euclidean systems, there is nonetheless a generalized notion of dimension that can be applied to a topological space. We first define a refinement of a cover, followed by the definition of this dimension:

Definition 5.1. Let X be a topological space, and let $\mathcal{U}$ be an open cover. We say that $\mathcal{V}$ is a *refinement* of $\mathcal{U}$ if $\mathcal{V}$ is also an open cover, and moreover, for every member $V \in \mathcal{V}$ there is some $U \in \mathcal{U}$ such that $V \subseteq U$.

Definition 5.2. Let X be a topological space. We define the *Lebesgue covering dimension* a X to be the least number n (if such a number exists) such that every open cover $\mathcal{U}$ of X has a refinement $\mathcal{W}$ such that every member $x \in X$ is contained in no more than $n + 1$ members of $\mathcal{W}$. If no such number n exists, the dimension of X is said to be infinite.

Preliminary investigations on my part suggest that hierarchical classification systems like DDC or LCC have a dimension of zero. Meanwhile, faceted classification systems with $n+1$ facets have dimension n. For example, the colon classification (CC) would be four dimensional as it has five facet categories.

The dimension of a topological space in some way conveys the complexity of its structure, with higher dimensional spaces being the more complex. It is known that Ranganathan thought of the UK as an infinite-dimensional space [2]. It would be interesting to consider the ways in which this complexity affects organization and the user's navigation of the UK.

5.2 Comparison of Structures

While classification theorists have distinguished the various classifications systems from one another based on syntactical considerations [1], one could take it upon themselves to show the varying degrees of difference or similarity between two classification systems using a formal topological analysis.

5.2.1 Topological Equivalence

We have already discussed the notion of topological equivalence. This gives a very strict notion of equivalence that is based fundamentally on the topology of a space. It would be interesting to investigate the topological equivalence of various classification systems, and perhaps define a broad generalization of different structures that give rise to the same topologies (n-faceted schemes, m-branching hierarchies, etc.). In practice, however, it is unlikely that topological equivalence will be found between differing classification systems. In fact, there likely will be differences between different editions of classification systems, as these would be sensitive to the addition of even one new class. Thus, it could be useful to investigate similarity and difference with a looser sense of equivalence.

5.2.2 Homotopy Equivalence

Homotopy equivalence is too complex to define here precisely (see [8] for details). Nonetheless, it is a topological equivalence that can be used to compare different topological spaces in interesting ways. It is preserved by topologically equivalent spaces, that is, if two spaces are topologically equivalent, then they are homotopically equivalent as well. It also lends itself well to computational and combinatorial methods, especially if one can compute the so called *nerve* of a space. This in turn requires the existence of a so called *good covering*; investigating the requirements for a practical classification system to have a good covering would be interesting, and would then motivate a computational program to study the homotopical aspects of classification system.

5.3 Speculative Classifications

Topology provides examples of many very interesting ways that data can be organized. For example, a circle (or 1-sphere) is topologically interesting because its entire surface can be traversed be travelling with respect to a single orientation, yet by doing so you end where you began. What would it mean for a classification system to have a circular structure like this? Can our semantics lend itself to a syntactical system such as this? We can apply this same line of questioning to any topological space. What we might find is that practical classification systems with only finitely many classes are not flexible enough to create topological phenomena such as those we speculate about. Perhaps fuzzier systems of classification could be used to allow for a new syntax that lends itself to potentially useful topological phenomena.

6 Conclusion

Classification systems, as we have come to see, resemble mathematical topological spaces in many ways. Using this analogy in our analyses of these systems allows us to clarify the syntactical claims that classification theorists make in a mathematically rigorous way. The analysis of the shelving problem here is only a brief glimpse of the potential that this framework could offer this field of study.

References

[1] M. P. Satija and D. Mart, "Mapping of the Universe of Knowledge in Different Classification Schemes," *International Journal of Knowledge Content Development & Technology*, p. 21, 2017.

[2] F. L. Miksa, "The concept of the universe of knowledge and the purpose of LIS classification," in *Classification Research for Knowledge Representation and Organization* (N. Williamson and M. Hudon, eds.), Elsevier, 1992.

[3] B. K. Sen, "Universe of knowledge from a new angle," *Annals of Library and Information Studies*, p. 6, 2009.

[4] J. R. Munkres, *Topology.* Upper Saddle River, NJ: Prentice Hall, Inc, 2nd ed ed., 2000.

[5] P. R. Halmos, *Naive Set Theory.* Undergraduate Texts in Mathematics, New York, NY: Springer New York, 1974.

[6] B. Hjørland, "Are relations in thesauri "context-free, definitional, and true in all possible worlds"?: Are Relations in Thesauri "Context-Free, Definitional, and True in All Possible Worlds"?," *Journal of the Association for Information Science and Technology*, vol. 66, pp. 1367–1373, July 2015.

[7] K. Kurakawa, Y. Sun, and S. Ando, "Application of a Novel Subject Classification Scheme for a Bibliographic Database Using a Data-Driven Correspondence," *Frontiers in Big Data*, vol. 2, p. 48, Jan. 2020.

[8] A. Hatcher, *Algebraic Topology.* Cambridge University Press, 2002.

A Naive Set Theory

Naively, a *set* is a simply a collection of objects. The objects collected within a set are called its members. For example, one could consider the set of all socks in the world, which I denote as S. If you were to go to your sock drawer and pick some sock denoted by s, then s would be a member of S. I denote this relationship of membership as $s \in S$. More generally, if X is a set, and x is a member of X, then I summarize this relationship as $x \in X$. We often define sets based on some property that its members possess. To do so, we use a builder notation, that takes the general form of $\{x : P(x)\}$, where $P(x)$ is short for saying "x has property P." For example, we could construct the set Y of all yellow coloured socks by stating that $Y = \{x : x \in S \text{ and } x \text{ is yellow}\}$.

A.1 Union and Intersection

Given two sets X and Y, we can create two new sets whose membership is given by (i) belonging either to X or Y (or both), and (ii) belonging to both X and Y. The former is referred to as the *union of X and Y* (denoted $X \cup Y$), and the latter is referred to as the *intersection of X and Y* (denoted $X \cap Y$). These are analogous to the Boolean operators used in catalogue search engines. As an example, consider the set Y of all yellow coloured socks and the set W of all woolen socks. Then the set $Y \cup W$ represents the set of socks that are either coloured yellow or made of wool, and the set $Y \cap W$ represents the set of socks that are both coloured yellow and made of wool.

Note that we may in fact find the union or the intersection of arbitrarily as many sets as we would like to. This could be a finite number, like 100, or it could even be some infinite value. In either case, union combine all of the selected sets into one single set, and intersections take the common part of all the sets.

A.2 Subsets and Power Set

Given two sets X and Y, if every member of X is also a member of Y, then we say that X is a *subset* of Y, and denote this relationship as $X \subseteq Y$. For example, if we take S as the set of all socks and W as the set of all woolen socks, then $W \subseteq S$ because every sock made of wool is also a sock. Given any set X, we may form a new set that contains all of its subsets. This set is called *the power set of* X, and it is denoted $\mathcal{P}(X)$. In this case, we see that the members of the power set are themselves sets. For example, if S still denotes the set of all socks, then the members of $\mathcal{P}(S)$ would consist of all possible collections of socks, such as W and Y above, the socks in your sock drawer, and even S itself. We must also not forget to include the *empty set* in $\mathcal{P}(S)$, which is the set that contains no members, denoted as $\emptyset$. Since $\emptyset$ contains no members, the empty set is counter-intuitively a subset of *every* set.

A.3 Functions and Inverses

Given two sets X and Y, we may associate every member of X to a member of Y. We often denote such an association by the letter f, and call this association a *function from* X *to* Y, denoted $f : X \to Y$. We say that X is the *domain of* f, and Y is the *codomain of* f. We denote the member of Y that a member x of X is associated to by $f(x)$. For example, consider the set L of left footed shoes and the set R of right footed shoes. We may define a function $r : L \to R$ such that every left-footed shoe is sent to its matching right-footed shoe. Here, the domain of r is the left-footed shoes, and its codomain is the right-footed shoes. Thus, if x represents the shoe I wear on my left foot, then $r(x)$ represents the shoe I wear on my right foot. Now, suppose that $U \subseteq X$ and $V \subseteq Y$, and let f be as above. We denote the *image of* U *under* f as $f(U)$. This is a subset of Y that is made of the members of Y that the members of U are associated to. That is, $y \in f(U)$ if and only if there is some $x \in U$ such that $f(x) = y$. Similarly, we define the *preimage of* V *under* f as $f^{-1}(V)$. This is a subset of X that is made of up the members of X whose image is contained in V, that is, $f^{-1}(V)$ is the largest subset of X such that $f(f^{-1}(V)) \subseteq V$. For example, if we consider the the set S of socks, the set C of colours, and the function $c : S \to C$ that sends every sock to its colour, the image of the wool socks W under c is in all likely hood all of C, i.e., $c(W) = C$. On the other hand, if Y' represents the shades of yellow, then the preimage of Y' under c would be the set of all socks whose colour is a shade of yellow Y, i.e., $c^{-1}(Y') = Y$.

Given three sets X, Y, and Z, and two functions $f : X \to Y$ and $g : Y \to Z$, we may combine f and g to get a new function. This function is called the *composition of* f *and* g, and is denoted $(g \circ f) : X \to Z$. Note that the domain of $(g \circ f)$ is X, while its codomain is Z. The function $(g \circ f)$ is formed by first applying f to x then directly afterwards applying g to $f(x)$, i.e., $(g \circ f)(x) = g(f(x))$. For example, consider the set P of people, the set S' of pairs of shoes, and the set R of right-footed shoes. We define the function $f : P \to S'$ that associates a person to their favourite pair of shoes, and a function $r : S' \to R$ that assigns a pair of shoes to its right-footed shoe. Thus, the function $r \circ f : P \to R$ associates every person to their favourite right-footed shoe.

Given sets X and Y and a function $f : X \to Y$, we say that that f is *injective* if for every member of y, the preimage $f^{-1}(\{y\})$ contains at most one member. We also say that f is *surjective* if the image $f(X) = Y$. In the case that f is both injective and surjective, we say that f is *invertible.* The reason for this is because if f is invertible, then there exists a function $f^{-1} : Y \to X$ such that $(f^{-1} \circ f)(x) = x$ for all $x \in X$ and $(f \circ f^{-1})(y) = y$ for all $y \in Y$. In such a case, f^{-1} is called the *inverse of* f. For example, if we again consider the set L of left-footed shoes, and the set R of right-footed shoes, then the function $r : L \to R$ that takes a left-footed shoe to its matching right-footed shoe, and the function $l : R \to L$ that takes a right-footed shoe to its matching left-footed shoe, then we see that $(r \circ l)$ and $(l \circ r)$ respectively take every left-footed shoe to itself and every right-footed shoe to itself.

A.4 Binary Relations

Let X be a set. A binary relation on X is an association between members of X. Some common notations for binary relations are $\cong$ or $<$, though they are not limited to these. For example, on the set S of socks, we might impose the binary relation $\cong$, that represents "being the same type of sock." Indeed, if I had two socks x and y that had the same *SpongeBob SquarePants* design, and were both size medium, then I would say that they are of the same type, i.e., $x \cong y$.

We can classify binary relations by several different properties. We say that a binary relation $\cong$ on X is *reflexive* if for every $x \in X$, $x \cong x$. In our example above, being the same type of sock is reflexive because every sock is the same type as itself. We say that a binary relation $\cong$ on X is *symmetric* if for every $x, y \in X$ such that $x \cong y$, then $y \cong x$ as well. In our example above, x "being the same type of sock" as y also means that y is the same type of sock as x. We say that a binary relation $\cong$ on X is *transitive* if for every $x, y, z \in X$, if $x \cong y$ and $y \cong z$, then $x \cong z$. In our example above, the sock x being the same type as the sock y, and the sock y being the same type as the sock z implies that the sock x is the same type as the sock z. Binary relations on X that are reflexive, symmetric, and transitive are called *equivalence relations on* X. Our example of a binary relation on S above is also an example of an equivalence relation on S, as we have shown.

Lastly, by placing the prefix *anti-* in front of any of the above properties, this means that the property is never exemplified. For example, if a binary relation $<$ on X is *anti-reflexive*, this means that there are no $x \in X$ such that $x < x$. Likewise, if a binary relation $<$ on X is *anti-symmetric*, this means that there are no distinct pairs of elements $x, y \in X$ such that both $x < y$ and $y < x$. As an example of both of these, let S' be the set of shoes, and let $<$ be the relation "is the matching left-footed shoe of". Thus, if I take a shoe, it is never the case that it is its own matching left-footed shoe. Moreover, if a shoe x is the matching left-footed shoe of a right-footed shoe y, then y is not the matching left-footed shoe of the left-footed shoe x (otherwise you would wear your shoes backwards).

A.5 Cardinality

Let X be a set. We call the value that corresponds to the number of members of X its *cardinality*, and denote it as $|X|$. For example, if D denotes the set of socks in my sock drawer, then $|D|$ tells me how many socks are in my drawer. The cardinal values of sets can either be finite or infinite. Though we will not explain the reasons here, there are actually different infinite cardinal values. Put colloquially, some infinities are more infinite than others. This was shown by Georg Cantor using what is now referred to as the method of *diagonalization*. Cantor demonstrated using this method that while there are just as many whole numbers as there are fractions, there is actually more real numbers (decimal numbers that may or may not be written as fractions, e.g., π, e) than there are whole numbers. It is for this reason that my set-theoretic definition of the Universe of Knowledge in §3 does not specify its cardinality, although I am assuming it is infinite, I am not sure which infinity captures its size.

More details on all that I have written about in this section can be found in [5].

B Proof of Theorems in §2

Theorem B.1. Let X, Y, and Z all be topological spaces. Suppose that $f : X \to Y$ and $g : Y \to Z$ are both continuous functions, then $g \circ f : X \to Z$ is continuous as well.

Proof. To prove the continuity of $g \circ f$, we must show that for every open set U of Z, there is an open set V of X such that $V = (g \circ f)^{-1}(U)$. We can simply the term $(g \circ f)^{-1}$ to $(f^{-1} \circ g^{-1})$, hence $(g \circ f)^{-1}(U) = f^{-1}(g^{-1}(U))$. By continuity of g, there is some open set W of Y such that $W = g^{-1}(U)$, and by continuity of f, there is some open set V of X such that $V = f^{-1}(W) = f^{-1}(g^{-1}(U)) = (g \circ f)^{-1}(U)$. □

Theorem B.2. The relationship of topological equivalence on topological spaces is an equivalence relation.

Proof. To show that topological equivalence is an equivalence relation, we must show that the relation is (i) reflexive, (ii) symmetric, and (iii) transitive.

(i) Suppose that X is a topological space. Then the identity function $i : X \to X$ defined by $x \mapsto x$ is a homeomorphism. Indeed, the identity function is invertible as $i^{-1} = i$. Moreover, i is continuous, as for every open set $U \in \mathcal{T}_X$, it is clear that $i^{-1}(U) = i(U) = U \in \mathcal{T}_X$. Thus, X is topologically equivalent to X.

(ii) Let X and Y be topologically equivalent. By definition, there is some $f : X \to Y$ that is a homeomorphism. By definition of homeomorphism, f has and inverse which we call g. Moreover, g is continuous. Lastly, $g^{-1} = f$, and f is by definition continuous. It follows that $g : Y \to X$ is a homeomorphism. Thus, if X is topologically equivalent to Y, then Y is topologically equivalent to X.

(iii) Let X, Y, and Z be topological spaces, and suppose that X is topologically equivalent to Y, and Y is topologically equivalent to Z. By definition, there exist homeomorphisms $f : X \to Y$ and $g : Y \to Z$. By Theorem 2.3, we see that both $g \circ f : X \to Z$ and $f^{-1} \circ g^{-1} : Z \to X$ are continuous, and that $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$, so $g \circ f$ is a homeomorphism. Thus, if X is topologically equivalent to Y, and Y is topologically equivalent to Z, then X is topologically equivalent to Z. $\square$

Theorem B.3. Let X and Y be topological spaces, let X be compact, and let $f : X \to Y$ be a continuous and surjective function, then Y is also compact.

Proof. Let $\mathcal{U}_Y$ be an open cover of Y. Define $\mathcal{U}_X$ as the collection $\{V : V = f^{-1}(U) \text{ for } U \in \mathcal{U}_Y\}$. By continuity of f, it follows that the members of $\mathcal{U}_X$ are open. Moreover, $\mathcal{U}_X$ covers X. Indeed, let $x \in X$, then there is some $U \in \mathcal{U}_Y$ such that $f(x) \in U$. Let $V = f^{-1}(U)$, then $x \in V$ and $V \in \mathcal{U}_X$. By compactness of X, there exists a finite subcover $\mathcal{V}_X$ of $\mathcal{U}_X$. Define $\mathcal{V}_Y$ as the collection $\{U : U = f(V) \text{ for } V \in \mathcal{V}_X\}$. This collection is finite, as $\mathcal{V}_X$ is finite. Moreover, it is a subcollection of $\mathcal{U}_Y$, as $f(f^{-1}(U)) = U$ for all sets U. Lastly, $\mathcal{V}_Y$ covers Y. Indeed, for all $y \in Y$, there is some $x \in X$ such that $f(x) = y$, and since $\mathcal{V}_X$ covers X, there is some $V \in \mathcal{V}_X$ such that $x \in V$, so $y \in f(V) \in \mathcal{V}_Y$ by definition. $\square$