Short Introduction to Topology

For Computer Science Grad Students

and other people who don't give a crap about topology

DRAFT version 0.6

Mark Jason Dominus (`mjd@plover.com`)

November 15, 2010

Abstract

This document is a **draft**, circulated only to attract comment and correction. It contains numerous errors of fact, vague overgeneralizations, and misleading implications. Many claims about general topological spaces in fact apply only to Hausdorff spaces. Much of the Applications section should be considered placeholders that bear only a vague resemblance to the correct, accurate explanations. Readers not already familiar with topology may come away with severe misapprehensions. Do not rely on it for anything.

Please send comments, suggestions, and corrections to the author at `mjd@plover.com`.

Please do not distribute this draft after **15 December, 2010**. The most recent version is available at `http://blog.plover.com/math/topology-doc.html`.

Topology is the branch of mathematics that tries to understand continuity and continuous functions. You probably don't need to know much topology for your CS degree, but people will bring it up from time to time. For example, you will hear in your category theory class that topological spaces form a category with continuous functions as the arrows. Or you will hear that certain topological spaces are natural models for intuitionistic logic. The goal of these notes is to be the minimal explanation of topology that will enable you to understand those two things.

On the real line, continuity is defined in terms of distance: a function from $\mathbb{R}$ to $\mathbb{R}$ is *continuous* if points that were sufficiently close together in the domain are mapped to points that are close together in the codomain. The use of "close together" implies a distance function, which in $\mathbb{R}$ is simply $|x - y|$. But what about spaces that are more complex, for which distance doesn't make sense, or where no distance function is apparent?

Topology reformulates this idea of "closeness" in terms of special sets called "open sets".

Open sets

The single fundamental concept of topology is the "open set".

You probably already know that in $\mathbb{R}$, an *open interval* is a set of the form $\{x : a < x < b\}$ for some real a and b. This interval is written (a, b). An *open set in* $\mathbb{R}$ is a union of open intervals. The union might be infinite.

For example, $\{x : x > 0\}$ is an open set because it is the union of the intervals $(0, n)$ for all positive n. $\mathbb{R}$ itself is open, and the empty set is also considered open, because it's an empty union of intervals.

This generalizes to $\mathbb{R}^n$ as follows: An *open ball* is the set of all points x whose distance from some center point p is less than some radius ϵ. In $\mathbb{R}$, an open ball is nothing but an open interval. (In $\mathbb{R}^3$, open balls are actually ball-shaped.) An open set in $\mathbb{R}^n$ is a union of open balls.

Open sets are topology's generalization of the analytic notion of "close together". In topology, two entities are "close" if every open set containing one intersects the other one. They are boundedly far apart if there are disjoint open sets around them.

Topologies

A *topological space* is a set X equipped with an open set structure. That is, we pick out some of the subsets of X and designate them as "open". The collection of "open" subsets of X that we designate is called the *topology* of the space. Since a topology is supposed to be a collection of open sets, its members are required to satisfy certain properties that are analogous to those satisfied by open sets in $\mathbb{R}^n$:

1. X itself must be open, and $\varnothing$ must be open
2. Any union of open sets must be open
3. The intersection of any two open sets must be open

Property 3 implies that any finite intersection of open sets must be open, by induction. But infinite intersections might *not* be open. This matches the behavior of $\mathbb{R}^n$, where finite intersections of open sets are open, but infinite intersections are sometimes not. For example, the intersection of the open sets $(-x, x)$ for all positive reals x is the set $\{0\}$, which is not open. In contrast, the union described in property 2 can be any union whatever, even a huge uncountable union.

This is obviously a very general definition, and the usual practice here would be to present a number of examples, some of which would be very weird. I am not going to do that. $\mathbb{R}^n$ is the only example you need to remember. But keep in mind that the definition *is* very general, and applies in all sorts of surprising places and spaces. The end of these notes will discuss a couple of examples of that.

With the definition in place we can define a number of topological properties, such as interiors, boundaries, limits, continuous functions, and so on, which are generalizations of those properties in $\mathbb{R}^n$.

I will present a few of these, and then talk about computer science and logic.

Connected sets

A "connected" set is one that is in all one piece. For example, intervals in $\mathbb{R}$ are connected, and $\mathbb{R}$ itself is connected, but the set $(0, 1) \cup (2, 3)$ is not (it's in two pieces) and the set of integers is not (it's in many pieces).

To prove that a set is not connected, we find a "separation" of it: we split the set into two nonempty, disjoint pieces that have some space in between. Mere disjointness is not sufficient for separation. After all, we can split $\mathbb{R}$ itself into two pieces $\{x : x < 0\}$ and $\{x : x \geq 0\}$ which are disjoint. But the two pieces touch each other, and their union is $\mathbb{R}$, which is connected. We want to say that in a separation the two pieces are boundedly far apart.

The way we express this in topology is to say that the two pieces are separated if they are contained in disjoint open sets. We can't find disjoint open sets containing $\{x : x < 0\}$ and $\{x : x \geq 0\}$, because every open set that contains $\{x : x \geq 0\}$ must also contain some negative numbers. And one can show that $\mathbb{R}$ is connected, as it should be. In fact it follows directly from the definition of $\mathbb{R}$ that it is connected.[1]

So let's define a connected set: A *separation* of a set S is a partition of S into two nonempty disjoint parts, X and Y, such that there are disjoint open sets X' and Y' with $X \subset X'$ and $Y \subset Y'$. A set is *disconnected* if there is a separation of it, and *connected* if not.

Under this definition, is $\mathbb{R}\backslash\{0\}$ connected or disconnected? Disconnected, of course. The separation is obvious: X is the positive reals, Y is the negative reals, $X' = X$ is an open set containing X, and $Y' = Y$ is an open set containing Y that is disjoint from X'.

- **Exercise 1.** A finite subset of $\mathbb{R}$ is connected only if it has fewer than two elements.
- **Exercise 2.** $\mathbb{Q}$, the set of rationals, is not connected.

Limits

Let S be a set in a topological space. A point p is called a *limit point* of S if it is arbitrarily close to the rest of S. As usual, the notion of "arbitrarily close to" is formulated in terms of open sets. The formal definition is that a point p (whether in S or not) is a limit point of S if every open set that contains p also intersects some point of S other than p.

Let's think about limit points in $\mathbb{R}$. We have some set, say $I^\circ = (0, 1)$. When is p a limit point of I°? Well, every point x in I° is a limit point, because every open set containing x contains some open interval around x, which then intersects I° in some point close to but not equal to x.

But also, 0 and 1 are limit points of I°, because every open interval around 0 intersects I°, and 1 similarly.

[1] If you know the definition of $\mathbb{R}$ in terms of Dedekind cuts, you should be able to see this.

No other points of $\mathbb{R}$ are limit points of I°. Consider -1 for example. Not every open set containing -1 intersects I°. $(-1.5, -0.5)$ is a counterexample.

Let's consider another example, $S = \{1, \frac{1}{2}, \frac{1}{3}, \ldots\}$. None of the points of S are limit points, because for each point $p \in S$, it's possible to find a small open interval around p that does not intersect S anywhere else. But the point 0 is a limit point of S, because every open interval around 0 intersects some points of S. 0 is indeed the limit of the sequence $\langle 1, \frac{1}{2}, \frac{1}{3}, \ldots \rangle$, which explains the name "limit point".

- **Exercise 3.** A finite subset of $\mathbb{R}^n$ has no limit points.
- **Exercise 4.** If S is a sequence in $\mathbb{R}^n$ that converges to p, then p is a limit point of the range of S.
- **Exercise 5.** But the range of S might have limit points even if S does not converge.
- **Exercise 6.** Which points of $\mathbb{R}^2$ are limit points of an open disc? Which points are limit points of a closed disc?

As we saw with I°, limit points of a set might be in the set, or not. A set which contains all its limit points is called a *closed* set. The union of a set S with its set of limit points is called the *closure* of S, written $S^\bullet$. The name "closed" was chosen because, in $\mathbb{R}$, the closure of an open interval (a, b) is the the closed interval $[a, b]$, and in $\mathbb{R}^n$, the closure of an open ball is a closed ball.

- **Exercise 7.** Some sets are both open and closed.
- **Exercise 8.** Some sets are neither open nor closed.
- **Exercise 9.** The closure of a set is a closed set.
- **Exercise 10.** If S is a closed set, then $S^\bullet = S$.
- **Exercise 11.** Every closed set S is the closure of some set S'.

If $A \subset B$, and $A^\bullet = B$, we say that A is *dense in* B: every point of B is "close to" many points of A. The canonical example of this is:

- **Exercise 12.** $\mathbb{Q}$ is dense in $\mathbb{R}$.

Probably the most important theorem about closed sets is that a set is closed if and only if its complement is open. This is true on $\mathbb{R}$ for the usual meanings of "closed" and "open". For example, the complement of the closed interval $[0, 1]$ is the open set $(-\infty, 0) \cup (1, \infty)$.

This theorem is so important that it's the only one I'm going to prove in this whole document. Let G be an open set and H be its complement. I want to show that H is closed. Let p be a limit point of H. I want to show that $p \in H$. Since p is a limit point of H, every open set containing p must intersect H. G is an open set not intersecting H, so $p \notin G$, so $p \in H$.

Now conversely, suppose H is a closed set and G is its complement. I want to show that G is open. For each $p \in G$, p is not a limit point of H, so there is some open set G_p that contains p and does not intersect H. Since each G_p is disjoint from H, each is a subset of G. Let $G' = \bigcup G_p$. $G' \subset G$, because G' is a union of subsets of G. $G \subset G'$ because $p \in G_p \subset G'$ for each $p \in G$. So $G = G'$. G' is a union of open sets, and thus is open, and therefore so is G.

- **Exercise 13.** The closure of a set S is the smallest closed set C with $S \subset C$.
- **Exercise 14.** The closure of a set S is the intersection of all closed sets C with $S \subset C$.

One reason that closed sets are important in analysis is the *Heine-Borel theorem:* A continuous function on a closed, bounded set has bounded values. This is not true of continuous functions on bounded sets in general. For example, the function $x \mapsto 1/x$ on the bounded set $(0, 1)$ is continuous but not bounded.

The topological version of the Heine-Borel theorem replaces "closed, bounded set" with "compact set". (In $\mathbb{R}^n$, these are equivalent.) Compactness is probably the most important topological concept that I'll omit from these notes.

Boundaries

We saw that sometimes some of the limit points of a set are in the set, and some are not. But the limit points that are not in the set are close to it. That is the source of the topological definition of a boundary: The *boundary* of a set S is the set of points that are limit points of both S and its complement.

- **Exercise 15.** The boundary of an open interval (a, b) is the set $\{a, b\}$.
- **Exercise 16.** The boundary of a closed interval $[a, b]$ is the set $\{a, b\}$.
- **Exercise 17.** The boundary of the open ball $\{(x, y, z) : x^2 + y^2 + z^2 < 1\}$ in $\mathbb{R}^3$ is the sphere $\{(x, y, z) : x^2 + y^2 + z^2 = 1\}$. So is the boundary of the closed ball.
- **Exercise 18.** The boundary of an open set S is $S^\bullet - S$.

Topology also defines the *interior* of a set, which is dual to the closure of the set. The interior of a set S, written S°, is the set of points in S that are *not* on the boundary.

- **Exercise 19.** S° is open.
- **Exercise 20.** If S is open, then $S = S^\circ$.
- **Exercise 21.** S° is the largest open set G with $G \subset S$.
- **Exercise 22.** The interior of a closed ball in $\mathbb{R}^n$ is an open ball.
- **Exercise 23.** Write S^C for the complement of set S. $S^{\circ C} = S^{C\bullet}$, and $S^{\bullet C} = S^{C\circ}$.

Continuity

In analysis, a function is continuous if you can make the image in the codomain as small as you like, by choosing a small enough part of the domain. That is, a function f is continuous if, given a subset C of the codomain that is contained in some small open set C', we can find a small open set D' in the domain such that $p \in D'$ implies that $f(p) \in C$. This motivates the following topological definition of continuity:

Suppose A and B are topological spaces. A function $f : A \to B$ is *continuous* if $f^{-1}(R)$ is open whenever R is open. ($f^{-1}(R)$ is the set of all points of the domain of f that map to points of R. That is, $f^{-1}(R) = \{x : f(x) \in R\}$.)

- **Exercise 24.** The identity function is continuous.
- **Exercise 25.** Constant functions are continuous.
- **Exercise 26.** Compositions of continuous functions are continuous.
- **Exercise 27.** Continuous functions of connected sets have connected images.
- **Exercise 28.** Continuous functions of open sets are *not* necessarily open.
- **Exercise 29.** The analytic definition of continuity on $\mathbb{R}^n$ coincides with the topological definition.

Homeomorphic spaces

If there's a bijective mapping between two sets A and B that is continuous in both directions (that is, both f and f^{-1} are continuous) then there is a bijection between the open sets of A and B, and so they have the same open set structure. Since all topological properties are formulated in terms of open sets, A and B have exactly the same topological properties, and are topologically equivalent. The topology jargon for this is that they are *homeomorphic*, and the bijective mapping is a *homeomorphism*.

Topologists do not distinguish between different homeomorphic spaces.

- **Exercise 30.** Even though $(0, 1)$ and $\mathbb{R}$ are geometrically quite different, they are topologically equivalent.

Applications

The category of topological spaces

One basic example of a category is the category **Top** of topological spaces. The objects are topological spaces, that is, each object is a set X equipped with a topology T, which is a collection of open sets satisfying the open-sets properties I listed back on page 2. The arrows are continuous functions between these spaces.

When we have an arrow $f : X \to Y$ and an arrow $g : Y \to X$ that are inverses, then f and g are bijections, and the spaces are homeomorphic. In the categorial view, the objects are isomorphic. Indeed, this is one of the main examples of isomorphic objects. They may be different sets, but a topologist doesn't care.

This category was an important motivator for category theory in the first place. Suppose you want to show that two topological spaces, say a torus and a sphere, are not homeomorphic. One way is to find a topological property that one has that the other does not.

An important topological property is the "fundamental group" of a space. Without getting into too many details, one chooses a "base point" x in the space, and considers the set of loops that start and end at x. Loops are considered equivalent if one can be smoothly transformed into the other. It turns out that the choice of base point doesn't matter, at least for connected sets. One can then put a group structure on the equivalence classes of loops in a straightforward way. The group operation is essentially concatenation of loops. This group is called the fundamental group of the space.

On a sphere, all the loops fall into a single equivalence class, because any loop can be contracted smoothly down to a single point. So the resulting group has only one element.[2]

But the fundamental group of a torus is quite different. Consider a very small loop a on the side of the torus. Such a loop can be shrunk to a point. But now consider a larger loop b that starts on the outside edge, goes around through the hole, and comes back to the outside. Such a loop cannot get any smaller, so a and b are different. This shows that the group is nontrivial. In fact the fundamental group of the torus is $\mathbb{Z}^2$. This proves that the torus is not a sphere.

I bring this up because this fundamental group construction is precisely a functor, one which maps the category of topological spaces to the category of groups. It is one of the very first functors ever considered as such. The study of such functors from **Top** to various categories of algebraic structures such as **Grp** is called *algebraic topology*.

Models of intuitionistic logic

Classical logic can be modeled with boolean algebras. The typical approach is to take the trivial boolean algebra, which has only two elements, $\top$ (true) and $\bot$ (false). A *valuation* is an assignment of one of these values to each propositional variable. This then extends to a truth value for every propositional formula:

- $v(f \vee g) = \bot$, if $v(f) = \bot$ and $v(g) = \bot$; otherwise $\top$
- $v(f \wedge g) = \top$, if $v(f) = \top$ and $v(g) = \top$; otherwise $\bot$
- $v(f \to g) = \bot$, if $v(f) = \top$ and $v(g) = \bot$; otherwise $\top$

[2] A space whose fundamental group is trivial is called "simply connected".

- $v(\neg f) = \bot$, if $v(f) = \top$; otherwise $\bot$

A formula is a *tautology* if it has a value of $\top$ regardless of how you assign values to the variables. That is, tautologies are those formulas to which every valuation assigns a value of $\top$. The tautologies are precisely the theorems of classical logic.

This is nothing more than a formalization of the method of truth tables. Each line in the truth table represents one valuation. You calculate the value of your formula for each valuation, and if it is $\top$ on every line, the formula is a tautology.

One can generalize this. Choose a set X. We will assign valuations which are subsets of X. Extend the valuation to formulas by saying that $v(f \vee g) = v(f) \cup v(g)$ and $v(f \wedge g) = v(f) \cap v(g)$. Say that tautologies are those which are assigned the value X by all such valuations.

But nobody bothers to do this, because the tautologies are the same regardless of which nonempty set X you use. So there is no reason to bother with any X bigger than one element, with $\top = X$ and $\bot = \{\}$.

Subsets of X form a boolean algebra, in which the boolean operations are set union and set intersection. But classical logic is faithfully modeled by the simple boolean algebra with only two elements.

In intuitionistic logic, the situation is somewhat different. Many formulas are theorems of classical logic but not of intuitionistic logic. The most well-known example is probably $\neg p \vee p$. This gets the value $\top$ for all boolean algebra valuations, but it is not a theorem of intuitionistic logic. So boolean algebras are not faithful models of intuitionistic logic. Other important examples of formulas that are classical tautologies but that are not theorems of intuitionistic logic are $\neg\neg p \to p$ and $((p \to q) \to p) \to p$.

Intuitionistic logic cannot be modeled by boolean algebras. Instead, one can model intuitionistic logic with a generalization of boolean algebras called *Heyting algebras*. But no finite Heyting algebra suffices: to model formulas with up to n variables requires a Heyting algebra of size at least 2^{2^n}.

But there is a simple example of an infinite Heyting algebra that does successfully model intuitionistic logic. One takes the values of propositional variables to be subsets of $\mathbb{R}$, then extends this to a valuation on all formulas as follows:

- $v(f \vee g) = v(f) \cup v(g)$
- $v(f \wedge g) = v(f) \cap v(g)$
- $v(f \to g) = (v(f)^C \cup v(g))^\circ$
- $v(\neg f) = v(f)^{C\circ}$

The notation F^C here means the complement of the set F, and F° means the topological interior of the set F. The value given to $f \to g$ is the *topological interior* of the set of points that are either in $v(g)$ or not in $v(f)$.

With this definition of valuation, theorems of intuitionistic logic are precisely the formulas which are assigned a value of $\mathbb{R}$ by every valuation.

For example, one can show that every valuation of $\neg(p \wedge \neg p)$ is $\mathbb{R}$. Suppose $v(p)$ is some set $P \subset \mathbb{R}$. Then $v(\neg(p \wedge \neg p)) = (P \cap P^{C\circ})^{C\circ}$. But $P^{C\circ} \subset P^C$, so $P \cap P^{C\circ} \subset P \cap P^C = \varnothing$. Then $v(\neg(p \wedge \neg p)) = \varnothing^{C\circ} = \mathbb{R}^\circ = \mathbb{R}$, regardless of what set P was assigned as the value of of p. Thus $\neg(p \wedge \neg p)$ is a tautology, and it is indeed a theorem of intuitionistic logic.

On the other hand, $\neg p \vee p$, which is not a theorem, is not a tautology. Say that $v(p)$ is $(0, \infty)$, the positive reals. Then $v(\neg p \vee p) = v(\neg p) \cup v(p) = (0, \infty)^{C\circ} \cup (0, \infty) = (-\infty, 0]^\circ \cup (0, \infty) = (-\infty, 0) \cup (0, \infty) \neq \mathbb{R}$.

This is the intuitionistic analogue of the truth-table method. For more complete details about this, see Sørensen and Urzyczyn, *Lectures on the Curry-Howard Isomorphism*, sections 2.3–4.

The compactness principle

An important theorem of model theory is the so-called "compactness principle". Suppose you have an infinite set of axioms, which are sentences of first-order logic. And suppose there is a model for any finite subset of these axioms. The compactness principle says there must also be a model for the entire set of axioms, and so the axioms must be consistent.

Since I didn't go into detail about compactness, I can't discuss this in detail. But the models can be taken to be boolean algebras, which are themselves examples of compact topological spaces. (The topology is that *every* subset of the space is open; such spaces are called "discrete".)

If one has models for some subsets of the axioms, then one can exhibit a model for the union of these subsets: it is the product of the topological spaces that are models for the subsets. But this only works if one can interpret this product space as a boolean algebra. This requires that the product of compact topological spaces be compact, which is called the Tychonoff theorem.

Here is an important application of the compactness principle, which also illustrates why it might be surprising. Consider the usual axioms of the real numbers, such as $\forall a, b : a + b = b + a$; $\forall x, y : \exists z : x \cdot z > y$, and so on. Add to these axioms the following infinite set of axioms concerning the constant ϵ: $\epsilon < 1$, $\epsilon < \frac{1}{2}$, $\epsilon < \frac{1}{3}$, and so on. Still no problem. Now add $\epsilon > 0$.

Clearly there is a model for any finite subset of these. Because any finite subset of these axioms is required only to satisfy some of the usual properties of $\mathbb{R}$, plus the requirement that $\epsilon < \frac{1}{n}$ for some integer n, plus possibly that ϵ is positive. So take the model to be $\mathbb{R}$, and $\epsilon = \frac{1}{n+1}$.

But by the compactness principle, there must be a model for all of the axioms at once, and here $\mathbb{R}$ will not do, because there is no such ϵ in $\mathbb{R}$.

So the compactness principle tells us that there must be a model that satisfies all the usual properties of $\mathbb{R}$, but which also has an infinitesimal element ϵ which is still positive, but smaller than $\frac{1}{n}$ for every integer n.

And since there is a model for this set of axioms, they must be consistent. The compactness principle proves immediately that the existence of an infinitesimal number ϵ is not inconsistent with the other properties of the reals.

What we get in this case is called the *hyperreals* or *nonstandard reals*, the subject of the branch of mathematics called *nonstandard analysis*.

Monads are closure operators

The topological closure operator satisfies the following properties:

1. $S \subset S^{\bullet}$
2. $S^{\bullet\bullet} \subset S^{\bullet}$
3. $(S \subset T) \rightarrow (S^{\bullet} \subset T^{\bullet})$

Analogously, monads in Haskell are required to support the following functions:

1. `return :: s -> M s`
2. `join :: M M s -> M s`
3. `fmap :: (s -> t) -> (M s -> M t)`

(They are also required to support a 'bind' operation called `>>=`, but `a >>= b` = `join (fmap b a)` and `join x` = `x >>= id`, so the two formulations are equivalent.)

Categorially, these are instances of the same structure, called a *Kleisli triple* or often just a *monad*. More formally, a monad is a functor (in the category-theory sense) equipped with two natural transformations. In the programming language category, the endofunctor is a type constructor; in the topological space category it is a closure operator. The natural transformations manifest themselves in the programming language category as the `join` and `return` functions, and in the topological space category as the required closure properties.

Categorial monads were first studied because of their connections with topological closure operators. Their significance in programming languages only became clear later on. Closure operators were earlier shown by Kuratowski to be fundamental to topological spaces: there is an equivalent formulation of topological spaces in terms of closure operators instead of in terms of open sets.

There's also a close connection between topological closure operators and modal operators in modal logic, but I don't know yet what it is.

Further reading

A standard and commonly recommended text is *Topology*, by James Munkres. I think it's one of those books that is confusing and obscure even when you already know what it's going to say, but a lot of people do seem to like it.

My own favorite is *General Topology*, by John L. Kelley. It's extremely terse, but always clear and pithy. The appendix contains a brilliantly clear exposition of axiomatic set theory, if you're interested in that.

As usual, the *Schaum's Outline* series volume is clear, direct, and unpretentious. It's declassé, but so are these notes.