Partial orders, well-founded orders and the principle of mathematical induction

Written by Prakash Panangaden
McGill University

31st August 2023

These are supplementary notes which Prakash wrote. In them, he discusses partial orders, well-founded orders and *proves* the principle of mathematical induction. This extra bit of reading is for fun and is not "testable" material.

1 Partial orders

Just as equivalence relations are an abstraction of equality, there is another class of relations that abstract inequality. These are the relations that abstract the notion of "ordering." However, unlike the usual notion of ordering of real numbers or integers we will **not** insist that every pair of elements are related.

Definition 1. Let S be any set. A **partial order** on S is a binary relation, usually written $\leq$, satisfying

1. $\forall x \in S. x \leq x$ [reflexivity]
2. $\forall x, y \in S. (x \leq y) \wedge (y \leq x) \Rightarrow (x = y)$ [antisymmetry].
3. $\forall x, y, z. (x \leq y) \wedge (y \leq z) \Rightarrow (x \leq z)$ [transitivity].

We are **not** requiring

$$\forall x, yS,\ x \leq y \text{ or } y \leq x.$$

If a partial order does satisfy this additional condition we call it a **total order** or a **linear order**.

Typical examples are the set of integers ordered by the usual notion of less than or equal, the set of subsets of a given set ordered by inclusion, or the set of complex numbers ordered by their magnitude. Henceforth we shall say *poset* rather than "set together with a partial order". Of course a given set may have many partial orders defined on it. Note that this notion does not include the usual notion of "strictly less than". One can introduce the symbol $x < y$ as an abbreviation for $x \leq y$ and $x \neq y$. When I refer to this concept I will use the phrase "strictly smaller" or "strictly less than" or "strictly decreasing". When I just say "less than" I mean what one would call "less than or equal to" in ordinary language.

Given a subset $X \subset S$ of a poset S we say that the element x_0 is the *least* element of X if it is less than every other element of X. In symbols

$$\forall x \in X,\ x_0 \leq x.$$

A given set may or may not have a least element. For example if we look at all the negative integers there is no least element. If we look at all the positive fractions there is no least element. In both these examples we had a total order. If we have a partially ordered set we can have the following situation. We say that an element $m \in X$ is *minimal* if it is not strictly greater than anything else in X. In symbols

$$\forall x \in X,\ \neg(x < m).$$

The element m could be less than x *or unrelated to it.* Consider the non-empty subsets of the $\{a, b, c\}$ ordered by set inclusion. The singleton sets $\{a\}, \{b\}, \{c\}$ are all minimal but there is no least element in the set. If a set does have a least element then it is obviously also (the only) minimal element.

2 Well-founded orders and Induction

In this section I will prove the principle of induction for general well-founded orders. The note is a bit terse and is intended to be a supplement to the lecture. It is, however, complete and you can - in principle - learn everything that you need to know for this class and beyond, about induction, from this note.

I assume that you have studied the principle of mathematical induction and used it to prove identities like $\Sigma_{i=1}^{n} i^2 = \frac{1}{6}n(n+1)(2n+1)$.

For many students this principle appears as a magic wand. My experience is that many are unsure of the status of this principle; "is it an assumption?", "a definition?" or a "theorem?". The short answer is that it is a theorem. Its range of applicability is much wider than the exercises that you had to do in discrete mathematics classes may have suggested.

The theory of induction is deep and difficult as can be seen by glancing at a book like "Elementary Induction on Abstract Structures" by Y. Moschovakis. What we do in this note is a small exercise in chapter 1 of that book. It will, however, suffice for the applications that we have in mind. The main theorem that we prove rests on the crucial notion of well-founded order.

Definition 2. Let $(S, \leq)$ be a poset. The order relation $\leq$ is said to be **well founded** if every nonempty subset has a minimal element.

Note, I did not say "minimum", just "minimal". The set of subsets of a *finite set* form a well-founded order. The set of positive integers forms a well-founded order, but the set of all integers do not. The set of subsets of an infinite set, ordered by inclusion, does not form a well-founded subset. The set of positive rationals does not form a well-founded order. But almost every structure that arises in computer science does form a well-founded order.

Proposition 3. A poset S is a well-founded order iff there are no infinite descending sequences in S.

Proof Immediate from the definition. ∎

One can see why the negative integers (and hence all integers) do not form a well-founded order.

2.1 Induction

So who cares about well-founded orders anyway?

Definition 4. Suppose that $(S, \leq)$ is a poset, we say that the **principle of induction** holds for S if for any predicate P, we have

$$(\forall x \in S.((\forall y \in S.(y < x \Rightarrow P(y)) \Rightarrow P(x))) \Rightarrow (\forall x \in S.P(x)).$$

This generalizes the usual notion of induction. You may notice that the usual notion of "base case" appears to be missing. This is implicitly included in the above formula since, if u is a minimal element, the (empty) antecedent above implies that $P(u)$ is true.

The reason for studying well-founded orders is contained in the following theorem.

Theorem 5. For a poset $(S, \leq)$ the principle of induction holds if and only if the order relation $\leq$ is well-founded.

Proof Suppose that $\leq$ is indeed a well-founded order. Let P be any predicate. We need to show that the principle of induction holds; since this has the form of an implication we need to show that the consequent holds whenever we assume the antecedent. Accordingly we assume the antecedent. Suppose that $\forall x \in S.(\forall y \in S.y < x \Rightarrow P(y)) \Rightarrow P(x)$. We must show that $\forall x \in S.P(x)$. Consider the set $U = \{u \in S | \neg P(u)\}$. What we need to show is that U is empty. Suppose that U is not empty then, because U is a subset of S, which we have assumed is a well-founded order, there must be a minimal element u_0. Thus $\forall y \in S.y < u_0 \Rightarrow y \notin U$, in other words if y is less than u_0 it must satisfy P. But according to our assumption, whenever everything less than u_0 satisfies P, u_0 itself must satisfy P. But now we have a contradiction because $u_0 \in U$ which by definition consists of elements that *do not* satisfy P. Thus our original assumption that U was not empty must be false, which is to say that $\forall x \in S.P(x)$.

Now for the reverse direction. Suppose that the principle of induction holds for $(S, \leq)$. Consider the predicate $F(x)$ defined as "there are no infinitely decreasing chains of elements of S that are all strictly less than x". Now suppose that x is some fixed element of S and that somehow we know that for all elements y that are strictly less than x, $F(y)$ holds. It must be the case that $F(x)$ holds because suppose that there is a decreasing sequence $x_1 > x_2 > x_3 \ldots$ all below x. Then x_2 is strictly less than x and there is an infinite, strictly decreasing sequence below it; in other words $\neg F(x_2)$. But we assumed that for all y strictly less than x – and x_2 is certainly in this collection – that $F(y)$ does hold. Thus such a chain cannot exist. Thus we have shown that

$$\forall x \in S.(\forall y \in S.y < x \Rightarrow F(y)) \Rightarrow F(x)$$

which means we can apply the principle of induction to $F(x)$ to conclude that $\forall x.F(x)$. By our proposition above, this means that the poset is a well-founded order.

Here is another, slicker, argument for the reverse direction. Suppose that there is a set $U \subset S$ which has no minimal element. This means that $\forall u \in U\ \exists v, v < u$ and $v \in U$. Now consider the predicate $P(x) \stackrel{\text{def}}{=} x \notin U$. Suppose that for any $x \in S$ we know that $\forall y < x\ P(y)$. I claim that $P(x)$ must hold. If $P(x)$ does not hold then $x \in U$ but our assumption about x is that every element $y < x$ is not in U (this is what $P(y)$ means) so, in other words, x is a minimal element of U. But this contradicts the assumption about U. Thus we have proved

$$\forall x \in S\ [\forall y < x\ P(y)] \Rightarrow P(x).$$

We are assuming that the principle of induction holds and the formula above is exactly the premise of the principle of induction. Thus, we have proved $\forall x\ P(x)$. But what is P? We have just proved $\forall x\ x \notin U$, i.e. U is the empty set. In short every non-empty set must have a minimal element. ∎

Now we can check that the so called principle of mathematical induction is a special case of this. The structure of interest here is the natural numbers i.e. the set $N = \{0, 1, 2, 3, \ldots\}$. This is clearly a well-founded order so the principle of induction is true. It only remains to put this in the familiar form. Let P be any predicate. There is a unique minimum element namely 0. Thus if we choose x to be 0 in the statement of the principle of induction above we have to show that $(\forall y \in N.y < 0 \Rightarrow P(y)) \Rightarrow P(0)$. The antecedent is vacuously true thus we must show that $P(0)$ is true. Now if we choose x to be $n+1$ we get

$$(\forall y \in N.y < (n+1) \Rightarrow P(y)) \Rightarrow P(n+1)$$

or rewriting this less clumsily we can say $\forall y \leq n.P(y) \Rightarrow P(n+1)$. Putting the pieces together we get the usual statement i.e.

$$(P(0) \wedge (\forall n.(\forall y \leq n.P(y) \Rightarrow P(n+1)))) \Rightarrow \forall n \in N.P(n).$$

The other principle that interests us in the principle of structural induction. Let S be any inductively defined set. Associated with the elements of S is the stage at which they enter the set S. We can define a well-founded order on S by saying that $x \leq y$ if x enters S at the same stage or before y in the inductive definition. Thus we can apply induction to this collection. An example will illustrate the idea.

Suppose that we define the set of *binary trees* as follows. The empty tree, written [] is a tree. If t_1 and t_2 are trees then $maketree[t_1, t_2]$ is also a tree. This is a typical inductive definition. Now the principle of induction applied to these tree structures says the following. Let P be any predicate on trees. If you can prove $P([])$ and if you can prove that for any trees t_1 and t_2 that whenever $P(t_1)$ and $P(t_2)$ holds then $P(maketree(t_1, t_2))$ also holds you can invoke the principle of structural induction to conclude that for all trees, t, $P(t)$ must hold. Thus we can do induction on trees and a multitude of other structures as well.

2.2 Zermelo's theorem

Please do not read this section. It will torment you needlessly.[1]

Definition 6. A **well-order** is a well-founded total order.

The usual order on the natural numbers is a well-order. Given a well-order every element, except the largest one if such an element exists, has a unique next element. Many examples of well-ordered sets can be given. On the other hand try to find a well-order on the real numbers. After trying for a while one tends to think that this is not possible.

However, using the Axiom of Choice, Zermelo proved

Theorem 7. Every set can be well ordered.

This created a sensation. The axiom of choice seems like an "obvious" principle but Zermelo's theorem seems unbelievable. As of yet no one can write down an explicit example of a well-order on the reals. This leads many logicians to reject the axiom of choice. However, the axiom of choice is equivalent to Zermelo's theorem and also to Zorn's lemma. The latter is essential to prove all kinds of theorems in mathematics: for example, every vector space has a basis. Thus most mathematicians are loath to give up the axiom of choice.

[1] This is Cesare and I agree.