Induction

John Stalker[*]

March 4, 2015

Induction is often somewhat mysterious the first time one sees it. It's often presented in a way which looks at first sight circular. There is in fact nothing circular about it, and little that's mysterious. Proofs by induction are simple applications of one of the basic properties of the integers:

Theorem 1 *Every non-empty subset of the integers which has a lower bound has a least element.*

Different presentations of the integers differ on whether this is a theorem or an axiom. In any case I won't prove it here. A simple corollary is the Principle of Induction.

Theorem 2 *If i is an integer and φ is a predicate such that $\varphi(i)$ and, for each $k \geq i$, $\varphi(k)$ implies $\varphi(k+1)$ then $\varphi(n)$ for all integers $n \geq i$.*

This is usually stated either for $i = 0$ or $i = 1$, in which case it is a statement about the natural numbers. Unfortunately there are two contradictory conventions about what the natural numbers are. Rather than choosing a side I'll just keep i arbitrary. "Predicate" is a technical term from Mathematical Logic. Roughly speaking it means a reasonable statement about one or more variables. More precisely, it's a statement which can be used to define a set. Not every statement defines a set. To see why, consider the statement "x is not very large." This is not a predicate. If it were then we could prove by induction that there are no very large non-negative integers because 0 is not very large and if l is very large the surely subtracting 1 will not change this, so $l - 1$ is also very large. In other words if k is not very large then $k + 1$ is also not very large. But no one would be likely to try to define the set of very large integers or its complement, except as a joke. A more subtle example is the statement "x cannot be defined in 140 characters or fewer." This also is not a predicate. If it were then there would be a smallest non-negative integer which cannot be defined in 140 characters or fewer. But "the smallest non-negative integer which cannot be defined in 140 characters or fewer" is a definition with fewer than 140 characters. It's beyond the scope of these notes to define the term "predicate" precisely, but its definition is very broad. Any statement you are likely to want to write down is allowed. An alternate way of thinking about predicates, which you may prefer, is in terms of Boolean-valued functions. A predicate takes some arguments and returns a value of "true" or "false" depending on their values. This is a correct interpretation as long a we are clear that the word "function" is used in it's usual mathematical sense, which is more general than "computable function". The correct syntax is as you would expect from this interpretation: It is redundant to say "$\varphi(n)$ is true" or "$\varphi(n)$ holds." It's sufficient to say "$\varphi(n)$", as in the theorem above. Similarly the negation of this statement is "not $\varphi(n)$". Statements like like "$\varphi(n)$ is false" or "$\varphi(n)$ does not hold" are redundant.

Now that the statement of the theorem has been made precise, or rather now that it has been made as precise as I intend to make it, we can proceed to its proof. Until further notice

$$U = \{n \in \mathbf{Z} \colon n \geq i\}$$

and all other sets are subsets of U. **Proof:** Let

$$T_\varphi = \{n \in U \colon \varphi(n)\}, \quad F_\varphi = \{n \in U \colon \text{ not } \varphi(n)\}.$$

[*]School of Mathematics, TCD, Copyright 2015

In words, T_φ and F_φ are the sets where φ is true and false, respectively. If F_φ is non-empty then, by the previous theorem, it has a least element. Call this element l. $l \neq i$ because $\varphi(i)$ by hypothesis. Except for $l = j$, Every $l \in U$ is $k+1$ for some $k \in U$. So there is a non-negative integer k such that $l = k+1$. $k < l$ and l is the *least* element of F_φ, so $k \notin F_\varphi$. Equivalently $k \in T_\varphi$. So we have a $k \in U$ such that $\varphi(k)$, but not $\varphi(k+1)$, contradicting our hypotheses. So our assumption the F_φ is non-empty is untenable. There is no $n \in U$ such that not $\varphi(n)$. Equivalently, $\varphi(n)$ for all $n \in U$. ■

It is of course also true that sets of integers bounded above have greatest elements and there is a corresponding principle of downward induction. A more interesting extensions is The Principle of Generalised Induction:

Theorem 3 *Suppose φ is a predicate, $\varphi(i)$ for some integer i, and, for each $k \geq i$, the statements $\varphi(i)$, $\varphi(i+1)$, ..., $\varphi(k)$ together imply $\varphi(k+1)$. Then $\varphi(n)$ for all $n \geq i$.*

This is easier to apply than the original Principle of Induction because in proving $\varphi(k+1)$ we are allowed to assume $\varphi(j)$ for all $i \leq j \leq k$, not just for $j = k$.

Proof: We introduce a new predicate ψ, where $\psi(k)$ is the statement that $\varphi(j)$ for all integers j between i and k. Then $\psi(i)$ because $\varphi(i)$. By hypothesis, $\psi(k)$ implies $\varphi(k+1)$. Trivially, $\psi(k)$ implies $\psi(k)$. So $\psi(k)$ implies $\psi(k)$ and $\varphi(k+1)$. But $\psi(k)$ and $\varphi(k+1)$ is just $\psi(k+1)$. We therefore have $\psi(i)$ and the statement that $\psi(k)$ implies $\psi(k+1)$, which are exactly the two things we need in order to apply the preceding theorem. ■

Induction is normally used for statements about integers. At first sight there's no hope of using it for statements about real numbers, since there are non-empty subsets of the real numbers which have lower bounds but no least element. There is, however, a property of the real numbers which plays a role analogous to the role played for the integers by the first theorem of this section.

Theorem 4 *Every non-empty subset of the non-negative real numbers has a greatest lower bound.*

Depending on where you learned Real Analysis this also might be either a theorem or an axiom. Also the greatest lower bound may be called an infimum. Again, this statement won't be proved here. What theorem plays the role of the Principle of Generalised Induction?

Theorem 5 *Suppose that s is real and φ is a predicate with a single real argument. Suppose that*

1. *for all real $x > s$, if $\varphi(t)$ for all t satisfying $s < t < x$ then there is a $y > x$ such that $\varphi(t)$ for all t satisfying $s < t < y$, and*
2. *there is a real $x > s$ such that $\varphi(t)$ for all t satisfying $s < t < x$.*

Then $\varphi(t)$ for all $t > s$.

Proof: Let

$$U = \{t \in \mathbf{R} : t > s\},$$

$$T_\varphi = \{t \in U : \varphi(t)\},$$

$$F_\varphi = \{t \in U : \text{ not } \varphi(t)\}.$$

$x > s$ is a lower bound for F_φ if and only if there is no $t \in F_\varphi$ with $t < x$, *i.e.* if and only if $\varphi(t)$ for all t satifying $s < t < x$. Similarly, $y > s$ is a lower bound if and only if $\varphi(t)$ for all t satifying $s < t < y$. So our first hypothesis is the statement that for any lower bound $x > s$ for F_φ there is another lower bound y with $y > x$. In other words, no $x > s$ can be a greatest lower bound for F_φ. But $F_\varphi \subset \mathbf{R}$ has a lower bound, namely s, and the second hypothesis prevents s from being a greatest lower bound. So, by the preceding theorem, F_φ must be empty. In view of the definition of F_φ we can conclude that $\varphi(t)$ for all $t > s$. ■

Comparing this to the earlier theorems, is a closer analogue to the Principle of Generalised Induction than to the usual Principle of Induction. The first hypothesis above corresponds roughly to the inductive step and the second case corresponds to the base case.

An corollary of the theorem is the following useful fact.

Theorem 6 *Suppose that s is real and f is a continuous real valued function defined for $t \geq s$. Suppose that*

$$f(s) < L$$

and, for each $x > s$ there is a $K < L$ such that if

$$f(t) < L$$

for all t in the interval $s < t < x$ then

$$f(t) < K$$

for all x in that interval. Then

$$f(t) < L$$

for all $t > s$.

Note that in the preceding theorem the "inductive step" was to extend a statement valid in one interval to the same statement for a larger interval. Here the inductive step is somewhat different. We keep the interval the same, but prove a sharper inequality from a weaker one.

Proof: Let $\varphi(x)$ be the statement "$\varphi(t)$ whenever $s < t < x$" and apply the preceding theorem. Because f is continuous at s there is a $\delta > 0$ such that $f(t) < L$ for all $t < s + \delta$. In other words $\varphi(t)$ for all $t < s + \delta$. This is the second condition from the preceding theorem, with $x = s + \delta$. To get the first condition, suppose $\varphi(t)$ for all $s < t < x$, *i.e.* that $f(t) < L$ whenever $s < t < x$. Then, by the hypotheses of this theorem there is a $K < L$ such that $f(t) < K$ in the same interval. The continuity of f at x then implies, first, that $f(x) \leq K$ and hence $f(x) < L$ and, second, that there is a $\delta > 0$ such that $f(t) < L$ whenever $|t - x| < \delta$. But then $f(t) < L$ for all $st < y$ where $y = x + \delta > x$. In other words, $\varphi(t)$ for all $t < y$. This completes the proof of the inductive step, and hence of the theorem. ■

As with proofs by induction, proofs using the preceding theorem may appear circular at first. The inductive step does not, however, assume what it purports to prove. It assumes a similar, but weaker, version of the same inequality. The theorem, or some variant thereof, is sometimes referred to as the "Bootstrap Lemma", a reference to Baron Münchhausen, a fictional character who, in one episode, pulls himself up by his own bootstraps.

It's sometimes useful to have a version of the preceding theorem which applies to finite intervals rather than semi-infinite intervals.

Theorem 7 *Suppose that s, z are real and g is a continuous real valued function defined for $s \leq t < z$. Suppose that*

$$g(s) < L$$

and, for each $x > s$ there is a $K < L$ such that if

$$g(t) < L$$

for all t in the interval $s < t < x$ then

$$g(t) < K$$

for all x in that interval. Then

$$g(t) < L$$

for all t in the interval $s < t < z$.

Proof: Let $f(t) = g\left(\frac{z-s}{z-t}s\right)$ and apply the preceding theorem. ■