Solutions: Induction

1. (a) Show it doesn't matter where you 'start the induction'; i.e. use induction to show that if, for some k, $\varphi(k)$ is true and also $(\forall n \geq k)(\varphi(n) \rightarrow \varphi(n+1))$, then $(\forall n \geq k)\varphi(n)$.

   (b) *IGT2* states the principle of course-of-values induction as follows: given (i) $\varphi(0)$ and (ii) $\forall n\{(\forall k \leq n)\varphi(k) \rightarrow \varphi(n+1)\}$ we can infer (iii) $\forall n\varphi(n)$. Give another version of the principle where the induction starts by making a *single* assumption instead of using both (i) and (ii).

   (c) Without looking back at the reading, show that the simple principle of arithmetical induction and the principle of course-of-values induction imply each other other.

   (a) Just put $\psi(n) =_{\text{def}} \varphi(n+k)$. Then we are given that $\psi(0)$ and $\forall n(\psi(n) \rightarrow \psi(n+1))$. So our standard induction principle entails that $\forall n\psi(n)$, which is in turn equivalent to $(\forall n \geq k)\varphi(n)$.

   (b) Just instead say: given $\forall n\{(\forall k < n)\varphi(k) \rightarrow \varphi(n)\}$, you can infer $\forall n\varphi(n)$. For note the given assumption implies in particular $(\forall k < 0)\varphi(k) \rightarrow \varphi(0)$, that is to say $\forall k(k < 0 \rightarrow \varphi(k)) \rightarrow \varphi(0)$. That conditional has a vacuously true antecedent, so the conditional implies (i) $\varphi(0)$. And the new conditional trivially implies the old conditional (ii).

   (c) Bookwork which won't be repeated here.

2. Show that, for any (natural number) n,

   (a) $1 + 3 + 5 + \ldots + (2n + 1) = (n + 1)^2$

   (b) $1^3 + 2^3 + 3^3 + \ldots + n^3 = [n(n + 1)/2]^2$

   (c) $n^3 - n$ is divisible by 6.

   Also

   (d) By guesswork or otherwise, find a formula for $3^0 + 3^1 + 3^2 + \ldots + 3^n$ and use induction to confirm it is correct.

   (a) Let's do it the laborious way first. Say $\varphi(n)$ holds just if equation (a) holds for n. Trivially we have the base case (i) $\varphi(0)$. Suppose next that we have $\varphi(n)$, i.e.

$$\begin{aligned} 1+3+5+\ldots+(2n+1) &= (n+1)^2 \\ \text{So}\quad 1+3+5+\ldots+(2n+1)+(2n+3) &= (n+1)^2+(2n+3) \\ \text{So}\quad 1+3+5+\ldots+(2n+3) &= n^2+4n+4 \\ &= ((n+1)+1)^2 \end{aligned}$$

   which demonstrates $\varphi(n+1)$. Since n in our supposition was arbitrary, we have the induction step (ii) $\forall n(\varphi(n) \rightarrow \varphi(n+1))$. Given (i) and (ii), it follows by induction that (iii) $\forall n\varphi(n)$.

That was a bit tedious but it works. But you should have spotted that there is much cuter proof which doesn't appeal to induction. For note that, summing vertically term by term,

$$\begin{array}{rl} & [\ 1 + 3 + 5 + \ldots + (2n-3) + (2n-1) + (2n+1)] \\ + & [(2n+1) + (2n-1) + (2n-3) + \ldots + 5 + 3 + 1\ ] \\ = & (2n+2) + (2n+2) + (2n+2) + \ldots + (2n+2) + (2n+2) + (2n+2) \end{array}$$

So *twice* the desired sum is $2(n+1)$ summed $(n+1)$ times, i.e. is $2(n+1)^2$.

(b) This time, say $\varphi(n)$ holds just if equation (b) holds for n. Trivially again (i) $\varphi(0)$. Suppose next that we have $\varphi(n)$, i.e.

$$\begin{array}{rcl} 1^3 + 2^3 + 3^3 + \ldots + n^3 & = & [n(n+1)/2]^2 \\ \text{So}\quad 1^3 + 2^3 + 3^3 + \ldots + n^3 + (n+1)^3 & = & [n(n+1)/2]^2 + (n+1)^3 \\ \text{and rearranging on the right gives} & = & [(n+1)(n+2)/2]^2 \end{array}$$

which demonstrates $\varphi(n+1)$. That gives us the induction step (ii) $\forall n(\varphi(n) \to \varphi(n+1))$. Hence again it follows by induction that (iii) $\forall n \varphi(n)$.

(c) We can again crank the handle and use induction. Suppose for some n, $n^3 - n$ is divisible by 6. Now consider $(n+1)^3 - (n+1)$, i.e. $n^3 + 3n^2 + 2n$, which is $n^3 - n$ plus $3n(n+1)$. Which is a number divisible by 6, by hypothesis, plus another number divisible by 6 (since the second term has 3 as a factor, and also one of n and $(n+1)$ is even). So that sum, and hence $(n+1)^3 - (n+1)$, is indeed divisible by six. Which gives us the induction step for a routine inductive argument that $n^3 - n$ is always divisible by 6.

Hands up, though, those who didn't immediate remark that we don't need to appeal to induction, because $n^3 - n = (n-1)n(n+1)$. And given three consecutive numbers, at least one must be even and divisible by 2, and exactly one will be divisible by 3, so their product must be divisible by 6.

(d) You could have used guesswork. Or you could put $3^0 + 3^1 + 3^2 + \ldots + 3^n = S(n)$, and then noted that $1 + 3(3^0 + 3^1 + 3^2 + \ldots + 3^n) = (3^0 + 3^1 + 3^2 + \ldots + 3^n) + 3^{(n+1)}$, that is to say $1 + 3S(n) = S(n) + 3^{(n+1)}$, so $S(n) = (3^{(n+1)} - 1)/2$.

If we want to confirm this by induction, note that trivially $S(0) = (3^{(0+1)} - 1)/2$. And suppose now that $S(n) = (3^{(n+1)} - 1)/2$. Then

$$\begin{array}{rcl} S(n+1) & = & (3^{(n+1)} - 1)/2 \ + \ 3^{n+1} \\ & = & (3.3^{(n+1)} - 1)/2 \\ & = & (3^{(n+2)} - 1)/2 \end{array}$$

So the equation for the sum holds for $n = 0$, and if it holds for n it holds for $n+1$. Hence by induction it always holds.

3. The Fibonacci numbers F_k are defined by: $F_0 = 0$, $F_1 = 1$, and if $k > 1$, $F_k = F_{k-1} + F_{k-2}$. Write down the first few Fibonacci numbers, and then show that, for any n,

   (a) $F_0 + F_1 + F_2 + \ldots + F_n = F_{n+2} - 1$

   (b) $F_0^2 + F_1^2 + F_2^2 + \ldots + F_n^2 = F_n F_{n+1}$

(c) $F_{n-1}F_{n+1} = F_n^2 + (-1)^n$

The first few Fibonacci numbers are, of course,

$$0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, 610, 987, \ldots$$

(a) The base case is trivial. So suppose the displayed equation holds for some n. Then,

$$\begin{aligned} F_0 + F_1 + F_2 + \ldots + F_n + F_{n+1} &= F_{n+2} - 1 + F_{n+1} \\ &= F_{n+3} - 1 \end{aligned}$$

So the supposition that the given equation (a) holds for n immediately implies that it holds for $n + 1$, so by induction it always holds.

Or for fun, though not so quickly, we can do this without induction. For brevity, put $S(n) = F_0 + F_1 + F_2 + \ldots + F_n$, then note

$$\begin{aligned} 2S(n) &= F_0 + (F_0 + F_1) + (F_1 + F_2) + \ldots + (F_{n-2} + F_{n-1}) + (F_{n-1} + F_n) + F_n \\ &= F_0 + F_2 + F_3 + \ldots + F_n + F_{n+1} + F_n \\ &= (S(n) - F_1) + (F_{n+1} + F_n) \\ &= S(n) - 1 + F_{n+2} \\ \text{So} \quad S(n) &= F_{n+2} - 1 \end{aligned}$$

(b) This example is similar. So the base case is trivial. Suppose the displayed equation holds for some n. Then

$$\begin{aligned} F_0^2 + F_1^2 + F_2^2 + \ldots + F_n^2 + F_{n+1}^2 &= F_nF_{n+1} + F_{n+1}^2 \\ &= F_{n+1}(F_n + F_{n+1}) \\ &= F_{n+1}F_{n+2} \end{aligned}$$

So the given equation (b) holds for $n+1$ too, and off we go again on another argument by induction.

This time too we can avoid arguing by induction, if we recall that $F_k = F_{k+1} - F_{k-1}$. So

$$\begin{aligned} F_0^2 + F_1^2 + \ldots + F_n^2 &= F_0^2 + F_1(F_2 - F_0) + F_2(F_3 - F_1) + \ldots + F_n(F_{n+1} - F_{n-1}) \\ &= F_nF_{n+1} \quad \text{by cancelling terms} \end{aligned}$$

(c) Once again the base case is trivial. So suppose the displayed equation holds for some n. Then

$$\begin{aligned} F_nF_{n+2} &= F_n(F_n + F_{n+1}) \\ &= F_n^2 + F_nF_{n+1} \\ &= (F_{n-1}F_{n+1} - (-1)^n) + F_nF_{n+1} \\ &= (F_{n-1} + F_n)F_{n+1} + (-1)^{n+1} \\ &= F_{n+1}^2 + (-1)^{n+1} \end{aligned}$$

So again, the supposition that the equation (c) holds for n implies that it holds for $n + 1$, so by induction it always holds.

4. (The Towers of Hanoi: see picture at http://bit.ly/1c4QSUf.) Suppose you have three posts and a stack of n different sized disks, initially placed on one post with the largest disk on the bottom and with each disk above it smaller than the disk below. You are to move the disks so they end up all on another post, again in decreasing order of size with the largest disk on the bottom. The only moves you are allowed involve taking the top disk from one post and moving it so that it becomes the top disk on another post, without being put on a smaller disk.

   (a) Show that for any n there must be a sequence of moves that does indeed end with all the disks on a post different from the original one in the desired configuration.

   (b) How many moves are required given an initial stack of n disks in the sequence of moves revealed by your answer to the previous question?

   (a) We proceed by induction (what else?). The base case, with with only one disk to play with, is trivial: move it to another post and you are done.

   Suppose then that it is possible to move n disks to another post, according to the rules. Then if you instead start with $n + 1$ disks, leave the bottom, largest, disk untouched and play with the top n disks. By assumption we can now move these top n disks to a different post (for the post with the largest disk kept on it is still always usable, because that any other disk can be placed on it). Then you move the so-far-untouched largest disk to the remaining empty post. Now keep that disk fixed, but off you go again, moving the n remaining disks so that they end up on top of the largest disk.

   We have shown, then, that we can legally move one disk, and that if we can legally move n disks we can legally move $n+1$ disks. Hence we can legally move any number of disks.

   (b) The described method requires $2^n - 1$ steps to move n disks. That's trivial for the case $n = 1$. Suppose it is true for some n that it takes $2^n - 1$ steps to move n disks. Then on our method, to move $n + 1$ disks, we will need $2^n - 1$ steps to move the top n disks to an empty post, 1 step to shift the largest disk, and then another $2^n - 1$ steps to move the top n disks again. That sums to $2^{n+1} - 1$ steps. So by induction, the claim that it always takes $2^n - 1$ steps follows.

   Bonus question: Is that the fastest solution?

Note now the answers given to the remaining questions are quite verbose, and no doubt more than you need to have given. But I've amplified a little where this seems interesting/useful.

5. Use some form of arithmetical induction to show that

   (a) Every wff of your favourite system for the propositional calculus is balanced, i.e. has the same number of left and right parentheses. (We'll assume that you aren't using a Polish bracket-free notation!)

   (b) No proper initial part of a wff (i.e. initial part shorter than the whole wff itself) is itself a wff. [Hint: Assume the desired result holds for wffs with up to n connectives, and then consider whether there could be counterexamples when dealing with a wff with $n + 1$ connectives.]

   Conclude that

(c) Any non-atomic wff can be decomposed into a main connective and one or two subformulae in exactly one way.

(a) Since we are being asked to use induction over numbers, we need some numbers to play with! Consider then *the number of connectives in an expression.*

So let $\varphi(n)$ say that any wff with n connectives is balanced, i.e. has the same number of left and right connectives.

Then trivially, (i) $\varphi(0)$ (since 0-connective wffs are atomic wffs which are bracket less).

Now suppose that for $k \leq n$, $\varphi(k)$ and consider now an arbitrary $n+1$-connective wff A. There are two cases to consider. Either (i) A is of the form $\neg A_1$ (or $(\neg A_1)$ if negation introduces brackets in your preferred syntax), where A_1 has n connectives. Or (ii) A is of the form $(A_1 \circ A_2)$, where $\circ$ is a binary connective.

In case (i), since we have $\varphi(n)$, we know that A_1 has balanced parentheses, and we have added none (or we have added a left and a right on the bracketed syntax for negation), it follows that $\neg A_1$ is still balanced.

In case (ii), both A_1 and A_2, having no more than n connectives, must both be balanced, by the assumption $(\forall k \leq n)\varphi(k)$. [NB here note that it is important that we using course-of-values induction, to cover all the possible cases for the number of connectives in A_1 and A_2.] Since A_1 and A_2 are both balanced, and we've added a left and a right bracket in forming $(A_1 \circ A_2)$, this too is balanced.

So either way, our $n+1$-connective wff A is balanced. Hence we have $\varphi(n+1)$.

So we've shown (ii) $\forall n((\forall k \leq n)\varphi(k) \rightarrow \varphi(n{+}1))$. Hence it follows by course-of-values induction that (iii) $\forall n \varphi(n)$.

(b) Again, we do an induction over the number of connectives in a wff. This time we'll write $\varphi(n)$ if and only if no proper initial part of a wff with n connectives is itself a wff.

Then vacuously, (i) $\varphi(0)$ (since 0-connective wffs are atomic wffs which have no proper initial parts).

Now suppose that for all $k \leq n$, $\varphi(k)$. Consider now an arbitrary $n+1$-connective wff A. There are various cases to consider.

Either (i) A has the form $\neg A_1$ where A_1 has n connectives (I'll leave it as a variant proof for you do the case if you like brackets with negation). Suppose for reductio that A has a proper initial part D which is a wff. Since it starts with a negation, D would have to be of the form $\neg E$ where E is a wff. But then E would be a wff which is an initial part of A_1 – but that's impossible by the induction hypothesis.

Or (ii) A has the form $(A_1 \circ A_2)$, where both A_1 and $A2$ have no more than n connectives. Suppose for reductio that A has a proper initial part D which is a wff. Since it has an initial bracket, D must be of the form $(B_1 * B_2)$ for some binary connective *, and where the wffs B_1 and B_2 again have no more than n connectives. By the induction hypothesis, B_1 can't be a proper initial segment of A_1 or vice versa. So they will have to be the same. So that implies that the remainder $*B_2)$ is a proper initial segment of $\circ A_2)$, which means the connectives must be the same, and then $B_2)$ is a proper initial segment of $A_2)$, which entails that the wff B_2 is a proper initial segment of the wff A_2 – but that's impossible by the induction hypothesis.

Hence either way, it follows from our supposition $k \leq n$, $\varphi(k)$ that if A has $n+1$ connectives, then it too has no wffs as proper initial parts. That is to say $\varphi(n+1)$ holds.

So that gives us the induction principle (ii) $\forall n((\forall k \leq n)\varphi(k) \to \varphi(n+1))$. Hence it again follows by course-of-values induction that (iii) $\forall n \varphi(n)$.

(c) If A is a wff of the form $\neg A_1$, then we know that this initial negation is the main connective. No problem!

What if A has the form $(A_1 \circ A_2)$, for binary connective $\circ$ and wffs A_1 and A_2? Could it also have the form $(B_1 * B_2)$, for a distinct binary connective $*$ and other wffs B_1 and B_2? No. Being wffs neither of A_1 and B_1 can be proper initial parts of the other. So the strings $(A_1$ and $(B_1$ must match exactly. Hence so must $\circ A_2)$ and $*B_2)$, which shows that the binary connective is fixed uniquely, and likewise that $A_2 = B_2$.

And then

(d) Explain how arithmetical induction can be used to give a proof that your favourite deductive system for the classical propositional calculus is sound in the sense that any theorem is a tautology.

(d) This sort of proof is particularly easy if we are dealing with an axiomatic proof-system. Indeed that's why books still often given an axiomatic system as their preferred style of formal logic: such systems may be horribly unnatural if we want to construct formalized deductions *inside* them, but it can be relatively easy to prove metalogical results about such systems from the *outside*.

(A) Suppose then that you have an axiomatic presentation of the propositional calculus (as in Mendelson's book or other texts). A logical proof is a linear sequence of wffs where each wff is either a logical axiom or follows from earlier wffs in the sequence by modus ponens, and a logical theorem is the last wff in such a proof sequence.

To prove that every theorem is a tautology is a simple proof by course-of-values induction on the length of the proof.

Define $\varphi(n)$ to be true if and only if any theorem produced by a proof up to n wffs long is a tautology.

Then we have, vacuously, $\varphi(0)$. (Or if you prefer to start the induction at $n = 1$, note that a one-step proof has to be an axiom, and we just check that any axiom – i.e. any instance of an axiom-schema – is a tautology).

Now suppose that $(\forall k \leq n)\varphi(k)$, and consider proofs n-wffs long. The final wff C in such a proof Π must

1. be an axiom, and hence a tautology, *or*
2. follow from two earlier wffs A and $(A \to C)$ by modus ponens. In this case, the earlier wffs are the last wffs in proofs consisting of two proper initial segments of Π. Each has its own proof nor more than n wffs long, and hence by our supposition each is a tautology. But if A and $(A \to C)$ are tautologies, then C must be a tautology.

So either way C is a tautology. Which shows that $\varphi(n+1)$. We can now use course-of-values induction to derive the required conclusion $\forall n\varphi(n)$.

(B) What if you are using a natural deduction system where you have temporary assumptions which are later discharged, and a theorem is at the root of a tree which has no remaining undischarged assumptions at the top of branches? Then the induction will have to be over e.g. the number of wffs in the tree which is a proof, and the relevant property to consider will be $\varphi'(n)$ which holds when, for any properly constructed tree of n wffs, the root wff is true on all valuations which make all the undischarged assumptions true. Suppose this holds for all proof-trees with up to n wffs, then we can show that it holds for any tree with with $n+1$ wffs by considering the final various rules that could have been applied to derive the final wff.

For example, we can paste together two proofs which end respectively with A and B to get a proof of $(A \wedge B)$ on their combined assumptions. And if the smaller proofs are such that, respectively, A and B are true on all valuations which make all the various undischarged assumptions true, then $(A \wedge B)$ will also be true on those valuations. For another example, suppose we have a proof with n wffs from undischarged assumption A and other assumptions B_i to conclusion C, then by cancelling through assumption A and adding one line at the foot of the tree we can form the proof with $n+1$ wffs from just the assumptions B_i to conclusion $(A \to C)$. And again if the former proof is such that the root wff C is true on all valuations which make A and all the B_i true, then evidently the new longer proof is such that the root wff $(A \to C)$ is true on all valuations which make all the B_i true.

Going through all the natural deduction rules for proof building like this, we can show in the end that if $(\forall k \leq n)\varphi'(k)$ then $\varphi(n+1)$. Course-of-values induction then yields the required conclusion $\forall n\varphi'(n)$, and then applying this to cases where there are no undischarged assumptions gives us our target result.

6. The Least Number Principle says, informally, that if some number has a given numerical property, then there is a least number with that property.
   (a) Show that the simple induction principle implies and is implied by the Least Number Principle.
   (b) How does the Least Number Principle relate to the principle that any non-empty set of natural numbers has a smallest element?

(a) The Least Number Principle says that, whatever property φ expresses, if $\exists n\varphi(n)$ then $\exists n\{\varphi(n) \wedge (\forall k < n)\neg\varphi(n)\}$.

Contrapose, and we get: if $\neg\exists n\{\varphi(n) \wedge (\forall k < n)\neg\varphi(n)\}$ then $\neg\exists n\varphi(n)$. Which is equivalent to if $\forall n\{(\forall k < n)\neg\varphi(n) \to \neg\varphi(n)\}$ then $\forall n\neg\varphi(n)$. But that's just a version of course-of-values induction. Therefore the Least Number Principle (applied to the particular wff φ) is just course-of-values induction (applied to the corresponding $\neg\varphi$). So the two general principles are equivalent. But we already know that the course-of-values induction principle implies and is implied by the simple induction. So the Least Number Principle is also equivalent to simple induction.

(b) If we assume that (i) for every numerical property there is a corresponding set (the set of natural numbers which have that property, i.e. the extension of the property) and (ii) for every set of natural numbers there is a corresponding numerical property (the property of being a number in that set), then the Least Number Principle as we first stated it without mentioning sets is equivalent to a version which does mention sets, i.e. the principle that any non-empty set of natural numbers has a smallest element.

Now – assuming we believe in sets at all! – (i) is uncontentious. But we might pause over (ii), the idea that any arbitrary set of numbers, however wild and impossible to specify in finite terms, corresponds to a genuine numerical property: see *IGT2*, p. 207 for more.

7. Say that m is fess than n if and only if either (i) m is even and n is odd or (ii) m and n have the same parity (both are even or both are odd).
   (a) How far along the sequence of numbers ordered by fess-than does the number 5 come?
   (b) Formulate and prove a Feast Number Principle.
   (c) How would we formulate a corresponding induction principle? Explain informally why the principle is sound. [Hint: Think of course-of-values induction.]

(a) Here are the natural numbers, re-sequenced by the fess relation. We get the evens in their usual order before the odds in *their* usual order:

$$0, 2, 4, 6, \ldots, 1, 3, 5, 7, \ldots.$$

If we use '$\sqsubset$' to symbolize the order-relation here, then $m \sqsubset n$ just in case m is even and n is odd or else m and n have the same parity and $m < n$.

If we march through the naturals in their new $\sqsubset$-ordering, checking off the first one, the second one, the third one, etc., where does the number 5 come in the order? Plainly, we cannot reach it in any finite number of steps: it comes, in a word, *transfinitely* far along the $\sqsubset$-sequence.

So if we want a position-counting number (officially, an *ordinal* number) to indicate how far along the sequence the number 5 is located, we will need a transfinite ordinal. We will have to say something like this: We need to march through all the even numbers, which here occupy positions arranged exactly like all the natural numbers in their natural order (why?); and then we have to go on another 3 steps. Let's use 'ω' to indicate the length of the sequence of natural numbers in their natural order, and we'll call a sequence with that length an ω-sequence. To indicate how far along the re-sequenced numbers we find the number 5, it is then tempting to say that it occurs at '$\omega + 3$'-th place.

And what about the whole sequence of re-ordered numbers? How long is it? How might we count off the steps along it, starting 'first, second, third, ...'? We must again march along as many steps as there are natural numbers in order to treck through the evens, and then – pausing only to draw breath – we have to march on through the odds, again going through positions arranged like all the natural numbers in their natural ordering. So, we have two ω-sequences, put end to end. It is tempting to say that the positions in the whole sequence are tallied by a transfinite ordinal we can indicate by '$\omega + \omega$'.

Cantor's theory of transfinite ordinals shows that these initial tempting thoughts can be made good, and we can indeed develop a lovely theory of orderings and the ordinals which measure their length.

(b) If some number has a given numerical property P, then there is a feast number with that property – i.e. an earliest number in the $\sqsubset$-sequence. Why? Consider cases. Suppose some number with property P is even. Then by the Least Number Principle, there will be a least number which is even-and-P, and this number must come earlier in the $\sqsubset$-sequence than any other number which is P. Suppose no number with property P is even, but some odd number is P. Then by the Least Number Principle, there will be a least number which is odd-and-P, and *this* number must come earlier in the $\sqsubset$-sequence than any other number which is P. Either way, there is a feast number which is P.

(c) Here's one version of the principle of course-of-values induction again:

1. from $\forall n((\forall m < n)Fm \to Fn)$, we can infer $\forall nFn$,

given that the quantifiers run over natural numbers, $<$ is the standard ordering over the numbers, and F stands in for some numerical property. Now, similarly, we can state a new principle of induction:

2. from $\forall n((\forall m \sqsubset n)Fm \to Fn)$, we can infer $\forall nFn$,

Here, the quantifiers run over the naturals again and $\sqsubset$ is the evens-before-odds 'fess' ordering.

Now, (ii) is said to be a principle of *transfinite induction* – the induction is over objects ordered in a sequence that gets tallied by a transfinite ordinal. Yet, for all that, we can give an entirely plodding informal justification for (ii):

> Assume $\forall n((\forall m \sqsubset n)Fm \to Fn)$. Then we have $(\forall m \sqsubset 0)Fm \to F0$ and, since nothing $\sqsubset$-precedes 0 (so the conditional's antecedent is vacuously true), we therefore have $F0$. We next have $(\forall m \sqsubset 2)Fm \to F2$, and since 0 (the only $\sqsubset$-predecessor of 2) satisfies F, it follows that $F2$. Next $(\forall m \sqsubset 4)Fm \to F4$, and since 0 and 2 (the only $\sqsubset$-predecessors of 4) satisfy F, it follows that $F4$. And so on, through all the evens. But then we have $(\forall m \sqsubset 1)Fm \to F1$ and since every $\sqsubset$-predecessor of 1, i.e. every even number, indeed satisfies F, it follows that $F1$. Which shows that $(\forall m \sqsubset 3)Fm$, and off we go again, now through all the odds. Hence any number, even or odd, satisfies F.

Note that this informal justification for (ii), a principle of transfinite induction along an ordering of length $\omega + \omega$, just involves a pair of ordinary inductions, one after the other – the first to show that $\forall mF(m)$ for even m, the second to show that the same holds for the odds. So there is evidently nothing scary going on here.

Reality check: We've shown we can do course-of-values induction over the fess-than ordering. Why can't we do an analogue of ordinary simple induction?

8. Start with two new definitions:

1. Let's say that the relation R defined over the objects X is *well-founded* just in case for any non-empty collection of objects $D \subseteq X$, then there is at least one R-minimal object a among D, i.e. an object a such that $\forall x(x \in D \to \neg Rxa)$.
2. Let's say that you can do *w-induction* on the relation R defined over the objects X just in case you can infer $\forall z\varphi(z)$ from $\forall x(\forall y(Ryx \to \varphi(y)) \to \varphi(x))$, where the

quantifications again are over the objects X. [That is, you can infer everything is φ from the premiss that if x's R-predecessors are all φ then so is x.]

Five problems:

(a) Give non-trivial examples of well-founded relations over three different kinds of objects.

(a) Here's a small selection of cases, showing some options ...

1. We've just shown that 'fess' relation on the natural numbers is well-founded.
2. Or take the $\prec$ order on the naturals where $m \prec n$ when $m \neq n$ and m is a factor of n. [Note in the set of numbers $D = \{2, 3, 4, 9, 12\}$ both 2 and 3 are $\prec$-minimal in D. Being well-founded doesn't require there to be a unique minimal element in each relevant set.]
3. Take ordered *pairs* of numbers $\langle m, n\rangle$ for $n > 0$. Then if we define $\langle m, n\rangle \prec \langle m', n'\rangle$ iff $m/n < m'/n'$, $\prec$ is not a well-founded relation (why not?). But if we put $\langle m, n\rangle \sqsubset \langle m', n'\rangle$ is either $m < m'$ or $m = m' \wedge n < n'$, then $\sqsubset$ is a well-founded relation over the pairs (why?).
4. Take the set of wffs of the propositional calculus, and consider the relation R such that Rw_1w_2 holds when the wff w_1 is a subformula of w_2.
5. Assuming you accept the axiom of foundation, take a collection of sets X and consider the relation $\in$.

(b) How does w-induction compare with course-of-values induction over the natural numbers $\mathbb{N}$?

(c) How does the claim that $<$ is well-founded over $\mathbb{N}$ compare with the Least Number Principle?

(d) Show that if the relation R defined over X is well-founded then we can do w-induction on the relation R. [Hint: prove the contrapositive.]

(e) Show that if we can do w-induction on the relation R defined over the objects X, then R must be well-founded. [Hint: obviously you need to use induction – do the induction on the property an object x has if every $D \subseteq X$ which contains x has an R-mininal element.]

Because we have (d) and (e), what we have temporarily called w-induction is usually called *well-founded induction*

(b) Ordinary course-of values induction is of course just an application of well-founded induction over $\mathbb{N}$ with respect to the order $<$.

(c) The Last Number Principle adds something to the claim that $<$ is well-founded over $\mathbb{N}$ – for the former says not just that, given any numbers $D \subset \mathbb{N}$, at least one has no predecessors in D, but also that all the other members of D are greater than it.

Suppose the relation R defined over the objects X is not only well-founded but *linear*, in the sense that for any $x, y \in X$, then $Rxy \vee x = y \vee Ryx$. Then if a and b are both R-minimal in D, we can't have Rab or Rba; so $a = b$; moreover, for any other object c in D, Rac. It is because $<$ is linear in addition to being well-founded that the full Least Number Principle holds.

(d) Suppose w-induction is not reliable for relation R. That is to say, for some φ, we can have the premiss $\forall x(\forall y(Ryx \to \varphi(y)) \to \varphi(x))$ even though the corresponding $\forall z\varphi(z)$ is not true. So there's an object, call it a which such that $\neg\varphi(a)$. Given the general premiss, it follows, by elementary logic, that $\exists y(Rya \wedge \neg\varphi(y))$. So we have another object, call it b which such that $\neg\varphi(b)$, and Rba. Given the general premiss, it follows, by elementary logic, that $\exists y(Ryb \wedge \neg\varphi(y))$. So we have another object, call it c which such that $\neg\varphi(c)$, and Rcb. And so it goes ... Put $D = \{a, b, c, \ldots\}$. No object among D is R-minimal – the downward R-chain goes on for ever. So R is not well-founded.

Contraposing, if R *is* well-founded over the objects X, then we can do w-induction on R.

(e) Suppose we can do w-induction on the relation R.

Now let's say that $x \in X$ is *good* if at least every $D \subseteq X$ which contains x has an R-minimal element. – the idea, then, is is that the mere presence of a good x among D rather magically ensures that there is an R-minimal object among D. Can there be good objects in this sense? We'll use w-induction to show that *all* objects are good!

So, suppose all the R-predecessors of x, if any, are good. Then take any collection of objects D containing x. Either x is already an R-minimal object among D; or it has an R-predecessor in D, and since that predecessor is good, that ensures that there is an R-minimal object among D. So we have $\forall x(\forall y(Ryx \to \mathit{good}(y)) \to \mathit{good}(x))$. Hence be induction, $\forall x\, \mathit{good}(x)$.

But now take any non-empty collection of objects $D \subseteq X$. It must contain at least one object x. But x will be good. So D will have an an R-minimal element. So R is well-founded on X, as was to be proved.

9. Two more definitions.

1. An *inductive datatype* (a.k.a. a recursive datatype) is defined in terms of some objects G, the *generators*, and some functions or operators O, the *constructors*, which map an object or objects as input to an object as output. The inductive datatype is then the smallest set (by set-inclusion of course) containing G and closed under the constructors – i.e. if it contains some given objects, then it also contains the results of applying a constructor in O to those objects.
2. We can do *s-induction* over a set D with respect to generators G (where $G \subseteq D$) and constructors O if and only if, from the premisses (i) that any object in G has property P and (ii) that the constructors preserve property P (i.e applied to things with P, a constructor yields something which also has P), we can infer (iii) that everything in D has P.

Five problems:

(a) Show how (1) the natural numbers, (2) the wffs of the propositional calculus, (3) proofs in an axiomatic system of propositional logic, can be regarded as inductive datatypes.

(a)
1. The (set of) natural numbers: generator 0; constructor *the successor function.*
2. The wffs of the propositional calculus: generators *the atomic wffs*; constructors, *prefixing a negation sign* and *putting a binary connective between two wffs and surrounding with parentheses.*

3. The proofs in an axiomatic system of propositional logic: generator *the empty proof* (the proof zero wffs long!); constructors *appending an axiom* and *appending a wff which follows by modus ponens from wffs already in the proof.*

(b) Show that if D is an inductive datatype, then we can indeed do s-induction over D with respect to its generators and constructors.

(c) Show conversely, show that if can do s-induction over D with respect to certain generators and constructors, then D is an inductive datatype.

(d) Use s-induction rather than arithmetical induction to show that every wff of your favourite system of propositional calculus is balanced, i.e. has the same number of left and right parentheses.

(e) Outline a proof using s-induction rather than arithmetical induction to show that every theorem of your favourite system of propositional calculus is a tautology.

Note, what we have temporarily called s-induction is usually called *structural induction.*

(b) Consider the set $D_P \subseteq D$ of elements of the datatype which have property P. Suppose, as the premisses for an s-induction, that (i) D_P contains the generators, and that (ii) D_P is closed under the constructors. Then, by definition of D, it is included in this set, i.e. $D \subseteq D_P$. So D_P is in fact the whole of D – hence every object in D has property P.

(c) Suppose we can do s-induction over D with respect to generators G and constructors O. Take the property P to be the property of belonging to every every set which contains the generators and is closed under the constructors.

Then obviously (i) every generator has property P.

Now, suppose f is an n-place constructor belong to O, and $o_1, o_2, \ldots, o_n$ belong to D and have property P. Then if Σ is a set which contains the generators and is closed under the constructors, $fo_1o_2, \ldots o_n$ belongs to Σ. But Σ is arbitrary, so $fo_1o_2, \ldots o_n$ belongs to any set which contains the generators and is closed under the constructors. That is to say $fo_1o_2, \ldots o_n$ has property P – so the constructor f preserves property P. But f was arbitrary, so (ii) all the constructors preserve property P.

Given (i) and (ii), then s-induction tells us that everything in D has property P, which makes D an inductive datatype with generators G and constructors O.

(d) Think of the wffs as forming an inductive datatype with generators *the atomic wffs*; constructors, *prefixing a negation sign* and *putting a binary connective between two wffs and surrounding with parentheses.* Then the generators have the property of being balanced wffs, and the constructors preserve balance. So s-induction immediately gives the desired result.

(e) Take the case of axiomatic proofs. Think of the proofs as forming an inductive datatype with generator the empty proof and constructors *appending an axiom* and *appending a wff which follows by modus ponens from wffs already in the proof.* Then the generators vacuously has the property of having a final wff which is a tautology, and the constructors preserve that property as we in effect noted in an earlier answer. So s-induction immediately gives the desired result.

If you want to think in terms of natural deduction proofs, then the constructors of datatype you need to consider are rather more complicated. The last part of the answer to Question 5 above indicates the sort of constructors you'll need to consider (e.g. pasting two proofs and adding a final conjunctive line, discharging an assumption and adding a final conditional line, etc.).

10. Given a relation R, suppose that R^*ab holds just when either Rab or $\exists x_1(Rax_1 \wedge Rx_1b)$ or $\exists x_1\exists x_2(Rax_1 \wedge Rx_1x_2 \wedge Rx_2b)$ or $\exists x_1\exists x_2\exists x_3(Rax_1 \wedge Rx_1x_2 \wedge Rx_2x_3 \wedge Rx_3b)$ or .... In other words, R^*ab if there is a (finite) chain of R-related objects linking a to b.

    Suppose, for example, that Rab says that *a is a parent of b*. Then the corresponding R^*ab holds when *a is an ancestor of b*. For this reason, R^* is said to be the *ancestral* of R.

    The *R-posterity* of a is the set of objects x such that R^*ax.

    A property P is *R-hereditary* if, an object x has P and Rxy, then y has P.

    The *Fregean ancestral* of R is the relation R^+ defined as follows: R^+ab if and only if b has every R-hereditary property had by every x such that Rax.

    (a) Express the definition of the Fregean ancestral of R more formally.
    (b) Show that R^*ab if and only if R^+ab.
    (c) Reality check: why isn't the Fregean definition more simply that R^+ab if and only if b has every R-hereditary property had by a?
    (d) Frege and Russell/Whitehead in effect define the natural numbers as 0 plus its S-posterity (meaning posterity with respect to the successor relation S such that Sxy iff y is the immediate successor of x). Show it follows from this definition that arithmetical induction is a sound principle.
    (e) How might we generalize this idea of induction over the posterity of a relation? Compare with our definitions of w-induction and s-induction.

(a) We need to generalize over properties. We can do this directly, using second-order quantification, to get

$$\forall F\{(\mathrm{Her}_R(F) \wedge \forall x(Rax \to Fx)) \to Fb\}$$

where $\mathrm{Her}_R(F)$, expressing that F is R-hereditary, abbreviates $\forall x\forall y((Fx \wedge Rxy) \to Fy)$. This, notation apart, was Frege's definition.

Or, if for one reason or another you are unhappy generalizing over properties by using second-order quantification into predicate position, we could use first-order quantification over sets (i.e. over property-extensions):

$$\forall X\{(\mathrm{Her}_R(F) \wedge \forall x(Rax \to x \in X)) \to b \in X\}$$

where now, in terms of sets, $\mathrm{Her}_R(X)$ abbreviates $\forall x\forall y((x \in X \wedge Rxy) \to y \in X)$.

(b) Suppose R^*ab. Then for some n there is a chain of objects $x_1, x_2, \ldots, x_n$ such that Rax_1, Rx_1x_2, Rx_2x_3, ..., Rx_nb. Take an arbitrary R-hereditary property had by x_1. It is passed down the R-chain to b. Which implies R^+ab.

Now consider the property P that y has if there is a chain of objects $x_1, x_2, \ldots, x_n$ such that Rax_1, Rx_1x_2, Rx_2x_3, ..., Rx_ny. Plainly this property is R-hereditary (if Py and Ryz, then we just need to append one element to the R-chain of objects witnessing that y has property P to get a chain witnessing that z has P). And even more trivially, this property is had by any x such that Rax. Hence if R^+ab, then b has property P. Which is just to say R^*ab.

(c) On the simpler definition, it is trivial that R^+aa. But a isn't one of its own ancestors in the ordinary sense, so this simpler Fregean definition (applied to the parent relation) won't capture the idea of an ancestor – though it would capture the notion that 'is either identical to or is an ancestor of'. We could work with a corresponding tweaked general notion of the ancestral if we wanted to.

(d) An object n is in 0's S-posterity just if S^*0n which implies S^+0n, which is to say that n has every S-hereditary property had by any object j such that Sj, i.e. n has every S-hereditary property had by 1. That is to say, for any n, if $P1$ and $\forall m(Pm \to P(m+1))$, then Pm. Which is just ordinary induction (starting at 1). So for Frege and Russell/Whitehead, the induction principle for natural numbers is true by definition of the very notion of the numbers.

(e) Left for you to explore!