# Mathematical Induction, Well-Foundedness, and the Axiom of Choice

## 0 introduction

In an essential way, formal proofs differ from informal reasoning: what in an informal argument may seem evident may nevertheless require non-trivial properties in a formal rendering of the same argument. This is the main reason why I consider formal proofs important: they reveal structure that otherwise may remain hidden. Thus, formal proofs contribute to a better understanding of the structure of a mathematical argument. (Apart from this, formal proofs are used for verification, either mechanical or not, but that is a different story.)

Particularly for this sake of insight, it may be relevant to know, first, that several seemingly different notions are actually equivalent, and, second, what is needed to prove this. For example, by now it is well-known that Mathematical Induction and Well-Foundedness are equivalent concepts, which can be defined in (at least) three different ways. The proof of one of these equivalences entails no more than a change of representation, but I do not recall ever having met a formal proof of the other one; hence this note.

For the sake of distinction, we call the three versions "Mathematical Induction" , "Well-Foundedness" , and "Finite Decrease" . The three versions pertain to a property of a binary relation $<$ on its domain $\Omega$; in what follows, we consider a fixed set $\Omega$. We will need no particular properties of $<$, but in practical applications $<$ often is a partial order. We adopt the following typing conventions, which enable us to omit the (constant) type denotations from the formulae themselves:

| | | |
|---|---|---|
| $x, y$ | : | $\Omega$ |
| $V$ | : | subset of $\Omega$ |
| $b$ | : | $\Omega \rightarrow \mathsf{Bool}$ ("predicates on $\Omega$") |
| $s$ | : | $\mathsf{Nat} \rightarrow \Omega$ ("infinite lists over $\Omega$") |
| $i, j$ | : | $\mathsf{Nat}$ (used as indices in $s$ only) |
| $f$ | : | $V \rightarrow V$ (in the context of a fixed $V$) |

I am aware that the following story possibly can also be told, partly or wholly, in the point-free relational calculus, but I feel no urge to do so here (if only because not every reader will be familiar with this calculus).

## 1 definitions

The principle of Mathematical Induction states that to prove a proposition of the shape $(\forall y :: b{\cdot}y)$ , it suffices to prove the (formally weaker) proposition $(\forall y :: \; b{\cdot}y \Leftarrow (\forall x : x{<}y : b{\cdot}x)\,)$ , for any predicate $b$ on $\Omega$ . Thus, that relation $<$ admits Mathematical Induction, can be rendered formally as:

MI: $(\forall b :: \; (\forall y :: b{\cdot}y) \;\Leftarrow\; (\forall y :: \; b{\cdot}y \Leftarrow (\forall x : x{<}y : b{\cdot}x)\,)\,)$

A relation is called Well-Founded if every non-empty subset of its domain has at least one *minimal element*, which means that the subset contains no elements smaller than such minimal element. Formally, for our relation $<$ , Well-Founded means:

WF: $(\forall V :: \; V{\neq}\phi \;\Rightarrow\; (\exists y :: \; y{\in}V \wedge (\forall x : x{\in}V : \neg(x{<}y))\,)\,)$

The property of Finite Decrease is a different formulation of well-foundedness that, at first sight, bears no resemblance to the previous one; it states that every decreasing sequence over $\Omega$ is finite. For technical reasons, it is more convenient to rephrase this as "no infinite sequence over $\Omega$ is decreasing" ; thus the somewhat awkward question what a proposition like "every ... sequence is finite" should mean is avoided. So we have:

FD: $\neg\,(\exists s :: \; (\forall j :: \; s{\cdot}(j{+}1) < s{\cdot}j)\,)$

## 2 proofs of equivalence

By just looking at the formulae we see that MI and WF have very similar shapes, whereas FD is enterily different: the latter involves infinite sequences, whereas the other two involve boolean functions and subsets, which are related concepts −see below−.

Indeed, the proof of $\text{MI} \equiv \text{WF}$ is more or less trivial; for the sake of completeness we include it anyhow.

**proof of** $\text{MI} \equiv \text{WF}$ **:** Via the connection $x{\in}V \equiv \neg(b{\cdot}x)$ we derive:

$$
\begin{array}{ll}
 & (\forall y :: b{\cdot}y) \;\Leftarrow\; (\forall y :: \; b{\cdot}y \Leftarrow (\forall x : x{<}y : b{\cdot}x)\,) \\
\equiv & \quad \{ \text{ contraposition, and De Morgan } \} \\
 & (\exists y :: \neg(b{\cdot}y)\,) \;\Rightarrow\; (\exists y :: \; \neg(b{\cdot}y) \wedge (\forall x : x{<}y : b{\cdot}x)\,) \\
\equiv & \quad \{ \text{ range-term trading } \}
\end{array}
$$

$$
\begin{array}{ll}
 & (\exists y :: \neg(b{\cdot}y)) \;\Rightarrow\; (\exists y :: \neg(b{\cdot}y) \wedge (\forall x : \neg(b{\cdot}x) : \neg(x<y))) \\
\equiv & \quad \{\ V - b \text{ connection } \} \\
 & (\exists y :: y \in V) \;\Rightarrow\; (\exists y :: y \in V \wedge (\forall x : x \in V : \neg(x<y))) \quad .
\end{array}
$$

For the sake of brevity, we have left implicit the universal quantifications over $b$ and $V$, and we have hidden the required dummy transformation (from $b$ to $V$) in the hint "$V - b$ connection"; the justification of this transformation is that the boolean functions on $\Omega$ and the subsets of $\Omega$ are isomorphic, that is, the "$V - b$ connection" is a bijection.

Actually, we have proved a stronger theorem now, namely that every instance of MI is equivalent with an instance of WF, so MI and WF are term-wise equivalent.

□

To establish the equivalence of all three, we are left with the obligation to prove either $\text{MI} \equiv \text{FD}$ or $\text{WF} \equiv \text{FD}$ ; in view of the similarity of MI and WF, the one proof should not be more difficult than the other. (In fact, we can now use MI and WF interchangeably, whatever seems most convenient.)

The only way I am able to prove $\text{WF} \equiv \text{FD}$ , is by mutual implication.

**proof of** $\text{WF} \Rightarrow \text{FD}$ **:** By means of The Morgan's rule, property FD can be rephrased as:

$$\text{FD}: \quad (\forall s :: (\exists j :: \neg(s{\cdot}(j{+}1) < s{\cdot}j))) \quad ,$$

and now we prove $\text{WF} \Rightarrow \text{FD}$ term-wise too. We do so for any infinite sequence $s$, by defining an associated set $V$, obviously nonempty, by:

$$V \;=\; (\mathsf{set}\, j :: s{\cdot}j) \quad ,$$

and for this $s$ and $V$ we derive:

$$
\begin{array}{ll}
 & (\exists y :: y \in V \wedge (\forall x : x \in V : \neg(x<y))) \\
\equiv & \quad \{\ \text{dummy transformation, using the definition of } V\ \} \\
 & (\exists j :: (\forall i :: \neg(s{\cdot}i < s{\cdot}j))) \\
\Rightarrow & \quad \{\ \text{instantiation } i := j{+}1\ \} \\
 & (\exists j :: \neg(s{\cdot}(j{+}1) < s{\cdot}j)) \quad ,
\end{array}
$$

which is the term of FD .

□

**proof of** WF ⇐ FD **:** Via ¬WF ⇒ ¬FD ; assuming $V$ to be a nonempty set without minimal elements, we must prove the existence of a decreasing infinite sequence; to this end, we observe:

$$\begin{array}{ll} & \neg\,(\exists y : y\in V :\ (\forall x : x\in V : \neg(x<y)\,)\,) \\ \equiv & \quad \{ \text{ De Morgan } \} \\ & (\forall y : y\in V :\ (\exists x : x\in V : x<y)\,) \\ \Rightarrow & \quad \{ \bullet \text{ see below } \} \\ & (\exists f : f\in V\to V :\ (\forall y : y\in V : f\cdot y<y)\,) \quad . \end{array}$$

Now for any $y$ in $V$ −which exists because $V$ is nonempty− and for any $f$ as specified here, sequence $s$ defined recursively by:

$$\begin{array}{lcl} s\cdot 0 & = & y \\ s\cdot(j{+}1) & = & f\cdot(s\cdot j) \quad , \text{ for all natural } j \quad , \end{array}$$

is decreasing, because by its very construction we have $s\cdot(j{+}1) < s\cdot j$ , for all $j$ .

□

The property we have used in step • is an instance of what is known as (one of the renderings of) the Axiom of Choice. That it is needed in our proof is not that surprising, because the Axiom of Choice is about the existence of functions, and existence of a function is what we need.

**Axiom of Choice:** For any sets $V, W$ and any predicate $P$ on $V\times W$ :

$$(\forall y : y\in W :\ (\exists x : x\in V : P\cdot x\cdot y)\,) \quad \Rightarrow$$
$$(\exists f : f\in W\to V :\ (\forall y : y\in W : P\cdot(f\cdot y)\cdot y)\,) \quad .$$

□

In informal renderings of the above proof, the appeal to the Axiom of Choice tends to be overlooked, quite understandably so, because the construction of a decreasing sequence seems so “evidently correct”: after all, if $V$ has no minimal elements we can go on and select smaller and smaller elements indefinitely, can't we? Selecting elements indefinitely is, however, not a formal reasoning step; therefore, in the formal proof the step labelled with $\bullet$ cannot possibly be taken without an appeal to the Axiom of Choice; moreover, a function $f$ with $(\forall y : y{\in}V : f{\cdot}y{<}y)$ is a crucial ingredient for the definition of a decreasing sequence.

## 3 an example

The motive to write this note was the following lemma, the need for which I encountered during the formal development of an algorithm known as "Distributed Summation" . This lemma is crucial to the proof that that algorithm is deadlock-free.

**lemma:** In a finite, acyclic, directed graph, the predecessor relation admits Mathematical Induction.
□

With $\Omega$ for the set of nodes of the graph, and with $<$ for the predecessor relation − so, $x{<}y$ means "the graph has an arrow from $x$ to $y$" −, the lemma can be formulated as:

$$\mathit{finite}{\cdot}\Omega \;\wedge\; \mathit{acyclic}{\cdot}(<) \quad\Rightarrow\quad \mathrm{MI} \quad ,$$

but from the point of view of proof construction this version turns out to be more tractable:

$$\neg\mathrm{FD} \;\wedge\; \mathit{acyclic}{\cdot}(<) \quad\Rightarrow\quad \mathit{infinite}{\cdot}\Omega \quad .$$

In words: an acyclic directed graph containing an infinite path, is infinite.

## 4 about finite and infinite

A proposition like "every decreasing sequence is finite" is awkward because it is unclear what it means. Apparently, it presupposes the existence of some set of so-called "sequences", some but not all of which have the property of being "finite"; so, we are considering a set containing both finite and infinite sequences. (If all sequences would be finite to start with, the above proposition would be void.) This is technically awkward.

To stay within the realm of finite sequences altogether, we could rephrase the proposition thus: "every decreasing (finite) sequence can be extended into a decreasing sequence only finitely often", but this is even more awkward: to formalize it we probably need names both for the length of a sequence and for the number of times it can be extended; in addition, a new concept – extension – is introduced.

A recursive definition of finite sequences probably does away with (the need to name) their lengths, but brings about the unavoidable case distinction between empty and nonempty sequences, which is not attractive either.

For these reasons, I prefer to formulate property FD in terms of infinite sequences: it yields the simplest possible formula.

* * *

In the same vein, it seems almost impossible to define finite sets without discussing their (finite) sizes as well, which is awkward if we are not interested in these sizes (but only in their existence). In such cases, rephrasing a theorem involving finite sets into one involving infinite sets, may simplify its proof. In fact, this is what I did in the lemma about finite directed graphs, where the term "(in)finite" occurs *twice*: this makes it very worthwhile to ponder about the particular phrasing to be used.

What, however, is an infinite set? To all practical intents and purposes, a set is infinite if it "contains" the natural numbers; more precisely, if there is an *injective* function from the natural numbers into that set:

$$\mathit{infinite}\cdot V \;\equiv\; (\exists s : s \in \mathsf{Nat} \rightarrow V : \; (\forall i,j : i \neq j : s\cdot i \neq s\cdot j)) \quad .$$

In words, we may also read this as: a set is infinite if it contains an infinite sequence all whose elements are different. For example, in the lemma about directed graphs, the notion of infinite sequence now occurs twice and this, indeed, is the key to its proof (together with a properly chosen definition of acyclicity, namely by means of the transitive closure of $<$).

Eindhoven, 30 april 2000

Rob R. Hoogerwoord
department of mathematics and computing science
Eindhoven University of Technology
postbus 513
5600 MB Eindhoven