CAMBRIDGE
Cambridge University Press & Assessment
978-1-316-51086-5 — Handbook of Constructive Mathematics
Edited by Douglas Bridges , Hajime Ishihara , Michael Rathjen , Helmut Schwichtenberg
Excerpt
More Information

PART I

INTRODUCTORY

Cambridge University Press & Assessment
978-1-316-51086-5 — Handbook of Constructive Mathematics
Edited by Douglas Bridges , Hajime Ishihara , Michael Rathjen , Helmut Schwichtenberg
Excerpt
More Information

1

An Introduction to Intuitionistic Logic

Michael Rathjen

1.1 Introduction

The constructive existence embodied in intuitionistic logic is very desirable in mathematics as it supports the computational view of mathematics. For a classically trained mathematician, though, it is often not that easy to switch to a mode of reasoning that maintains constructivity. For instance, arguing by making case distinctions is an almost automatic habit in classical mathematics, but one that is liable to introduce illegitimate employments of the *law of excluded middle*, LEM. Several examples will be discussed in Section 1.2.

The main aim of this chapter is to present an informal and intuitive approach to intuitionistic[1] logic – also known as constructive logic – for the working constructive mathematician.[2] The guiding idea is that this will be furnished via the *Brouwer–Heyting–Kolmogorov interpretation* (henceforth the BHK-interpretation) of the logical connectives and quantifiers. This is presented in Section 1.3.

Sometimes, however, uncertainties as to the constructive validity of an argument might still arise as the BHK-interpretation is based on an unexplained notion of constructive function. Moreover, it can also be cumbersome to ascertain constructivity of a mode of reasoning by means of the BHK-interpretation, or, venturing in the other direction, to demonstrate that an argument doesn't hold under an intuitionistic lens. In such situations, a more formal approach may be called for. For this reason, and for other equally important purposes, this chapter also features two formal proof systems in Sections 1.4 and 1.5: Gentzen's *natural deductions* and a *Hilbert-style calculus* for intuitionistic predicate logic. In the natural deduction style of reasoning there are no axioms, only rules of inference. This lack of axioms is compensated for by having permission to introduce any formula as a hypothesis at any time. To be able to get rid of such formulas at later stages in the proof, there are rules that allow

[1] No pun intended.
[2] There are several excellent introductions to intuitionistic logic; for example, [26, Chapter 2]. This one is for the reader's convenience, namely to have one to hand in the same volume.

one to discharge hypotheses. A Hilbert-style proof calculus, on the other hand, has a number of axioms but few inference rules. The latter are, moreover, of a local nature in that they do not involve the sometimes burdensome regime of tracking open and discharged hypotheses over the entire proof. Both formalizations have important roles to play. Natural deductions beautifully exhibit the connection between intuitionistic logic and computations known as the *Curry–Howard correspondence* (or *isomorphism*) and the *formulae-as-types interpretation*, whereas a Hilbert-style calculus is very useful in demonstrating that crucial concepts (e.g., realizability) are preserved under intuitionistic logic. The Curry–Howard correspondence will be briefly discussed in the final section, Section 1.7.

The penultimate section, Section 1.6 is devoted to Kleene's *1945-realizability* of intuitionistic number theory (or Heyting arithmetic), **HA**. The concept and technique of realizability is another nice illustration of the fact that intuitionistic proofs encapsulate numerical information, and it reveals how it is extractable from them. Furthermore, realizability bears out the fact that one option of instantiating the unexplained notion of constructive function of the BHK-interpretation consists in equating it with the notion of partial computable (= partial recursive = partial Turing machine computable) function.

1.2 Constructive Existence

Constructive mathematics is both old and new for the reason that, with few exceptions, mathematicians thought constructively until the 1870s, that is, before the set-theoretic shift initiated by Dedekind, Cantor, and others, while a substantial development of modern mathematics from a constructive base (largely thought to be impossible) had to await the work of Errett Bishop in the second half of the twentieth century.

The meaning of 'existence' in mathematics in the first phase was essentially what we equate with constructive existence nowadays. In general, the requirement for the latter is the demand that E be respected:

> (E) *The correctness of an existential claim* $(\exists x \in A)\varphi(x)$ *is to be guaranteed by warrants from which both an object* $x \in A$ *and a further warrant for* $\varphi(x)$ *are constructible.*

Or as Bishop ([2, p. 2]) put it:

> *When a man proves a positive integer to exist, he should show how to find it. If God has mathematics of his own that needs to be done, let him do it himself.*

The year 1888 saw Hilbert's proof of the basis theorem (Gordan's problem of invariants). Hilbert demonstrated the existence of a finite basis via a proof by contradiction. It is telling that he had to convince Cayley and Gordan that he had really proved the theorem, since they, like other mathematicians, expected a solution along the lines of E that exhibited a finite basis.[3]

The later part of the nineteenth century and the first part of the twentieth century was a period of great advances in mathematics, but also one of uncertainty and opposing views. A central role in the discussions about mathematical existence was played by Zermelo's proof that the reals can be well-ordered, presented at the International Congress of Mathematicians in 1905. While many mathematicians were apt to dismiss the paradoxes as peripheral to mathematics, contradictions of a somewhat philosophical nature, Zermelo's result concerned a core object of mathematics: $\mathbb{R}$. His proof notoriously used the axiom of choice, AC. While Zermelo argued that AC was self-evident, it was also criticized as an excessively non-constructive principle by some of the most distinguished analysts of the day.[4] Zermelo's proof furnishes absolutely no idea as to how a well-ordering of $\mathbb{R}$ can be defined (let alone be constructed). At the time it was natural to single out AC as the sole villain that engenders undefinable mathematical entities. With the advent and tools of modern mathematical logic, however, it emerged that the venerable logical principle (or law) of excluded middle, $\phi \vee \neg\phi$, suffices to produce such examples. For example, one can produce existential statements of the form $\exists x \subseteq \mathbb{R}^2 \varphi(x)$ that are provable in pure logic with the aid of excluded middle, however, $\mathbf{ZFC}$ (Zermelo–Fraenkel set theory with the axiom of choice), even when augmented by the generalized continuum hypothesis, GCH, cannot prove that there is a definable such set.[5] In a similar vein, there are number-theoretic statements such that first-order number theory $\mathbf{PA}$ (Peano arithmetic) proves $\exists x\, \theta(x)$ but for no term t does $\mathbf{PA}$ prove $\theta(t)$.

Brouwer made his famous criticism of the law of excluded middle, LEM, in his 1907 dissertation [4] and his 1908 article 'De onbetrouwbaarheid der logische principes'[6] [5]. He was not the first person, though, to raise doubts about its validity. The German mathematician Paul du Bois-Reymond in his book [9] *Die allgemeine Functionentheorie* from 1882 clearly separated actual infinities from potential infinities and argued that the logic governing potential but non-actual infinite sets would not countenance LEM.

Brouwer called his mathematics *intuitionistic mathematics*. The formal logic that drops LEM and related principles such as the double negation shift $\neg\neg A \to A$ is

[3] Hence Gordan's famous remark, 'this is not mathematics, this is theology', although later Gordan came to appreciate 'theology' in mathematics.

[4] At the end of a note sent to the *Mathematische Annalen* in December 1905, Borel writes about the axiom of choice: *'It seems to me that the objection against it is also valid for every reasoning where one assumes an arbitrary choice made an uncountable number of times, for such reasoning does not belong in mathematics.'* ([3, pp. 1251–1252]; translation by H. Jervell, cf. [16, p. 96]).

[5] This means that for no set-theoretic formulas $\psi(x)$ does one have $\mathbf{ZFC} + \mathsf{GCH} \vdash \exists! x[x \subseteq \mathbb{R}^2 \wedge \varphi(x) \wedge \psi(x)]$. The latter follows from a result of Feferman [11] obtained by forcing in 1963.

[6] The Unrealiability of the Logical Principles.

called *intuitionistic logic* and sometimes *constructive logic* or *Heyting's predicate calculus*. The first name is well ingrained, but Brouwer did not develop intuitionistic logic. The first explicit formulation of the laws of intuitionistic logic is due to the Russian logician Kolmogorov [18]. Kolmogorov accepted Brouwer's critique of LEM when applied to infinite domains. He then took Hilbert's formalization of classical logic [13] as the starting point for his investigation, deselecting those axioms that have validity only in the domain of the finitary. With the exception of the axiom $A \to (\neg A \to B)$ (which is not valid in minimal logic), Kolmogorov arrived at a complete formalization of intuitionistic logic. The main achievement of his paper, though, was to prove that classical logic is translatable into intuitionistic logic, thereby largely anticipating the independent discoveries of translations by Gentzen and Gödel in 1933. The full formalization of intuitionistic logic was obtained in 1930 by Heyting [12], who was unaware of Kolmogorov's work.

Here is an example of a non-constructive existence proof that one finds in almost every book and article concerned with constructive issues.[7]

Proposition 1.1 *There exist irrational numbers $\alpha, \beta \in \mathbb{R}$ such that α^β is rational.*

Proof We know that $\sqrt{2}$ is irrational, and ${\sqrt{2}}^{\sqrt{2}}$ is either rational or irrational. If it is rational, let $\alpha := \beta := \sqrt{2}$. If not, put $\alpha := {\sqrt{2}}^{\sqrt{2}}$ and $\beta := \sqrt{2}$. Thus in either case a solution exists. □

This proof provides two pairs of candidates for solving the equation $x^y = z$ with x and y irrational and z rational, without giving a means of determining which. From a non-trivial result of Gelfand and Schneider, it is known that ${\sqrt{2}}^{\sqrt{2}}$ is transcendental, and thus the second pair provides an explicit answer.

Similarly, classical proofs of disjunctions can be unsatisfactory. H. Friedman pointed out that classically either $e - \pi$ or $e + \pi$ is a irrational number since assuming that both $e - \pi$ and $e + \pi$ are rational entails the contradiction that e is rational. But to this day we don't know which of these numbers is irrational.

Another example is the standard proof of the Bolzano–Weierstraß Theorem.

Example 1.2 If S is an infinite subset of the closed interval $[a, b]$, then $[a, b]$ contains at least one point of accumulation of S.

Proof We construct an infinite nested sequence of intervals $[a_i, b_i]$ as follows.

Put $a_0 = a$, $b_0 = b$. For each i, consider two cases:

(i) if $\left[a_i, \frac{1}{2}(a_i + b_i)\right]$ contains infinitely many points of S, put $a_{i+1} = a_i$, $b_{i+1} = \frac{1}{2}(a_i + b_i)$;

[7] Dummett [10] writes that this example is due to Peter Rososinski and Roger Hindley.

(ii) if $\left[a_i, \frac{1}{2}(a_i + b_i)\right]$ contains only finitely many points of S, put $a_{i+1} = \frac{1}{2}(a_i + b_i)$, $b_{i+1} = b_i$.

With the help of LEM, it is plain that each interval $[a_i, b_i]$ contains infinitely many points of S. This being a sequence of nested intervals, $(a_i)_{i\in\mathbb{N}}$ converges to a point every neighbourhood of which contains infinitely many points of S. □

The foregoing proof specifies a 'method' which, in general, a constructivist cannot carry out.

1.2.1 Counterexamples from Analysis

Certain basic principles of classical mathematics, which are taboo for the constructive mathematician, were called *principles of omniscience* by Bishop. They can be stated in terms of binary sequences, where a binary sequence is a function $\alpha : \mathbb{N} \to \{0, 1\}$. Below, the quantifier $\forall\alpha$ is supposed to range over all binary sequences and the variables n, m range over natural numbers. Let $\alpha_n := \alpha(n)$.

Definition 1.3 *Limited Principle of Omniscience* (LPO):

$$\forall\alpha\,[\exists n\,\alpha_n = 1 \;\;\vee\;\; \forall n\,\alpha_n = 0].$$

Weak Limited Principle of Omniscience (WLPO):

$$\forall\alpha\,[\forall n\,\alpha_n = 0 \;\;\vee\;\; \neg\forall n\,\alpha_n = 0].$$

Lesser Limited Principle of Omniscience (LLPO):

$$\forall\alpha\,\big(\forall n, m[\alpha_n = \alpha_m = 1 \to n = m] \;\to\; [\forall n\,\alpha_{2n} = 0 \;\;\vee\;\; \forall n\,\alpha_{2n+1} = 0]\big).$$

Theorem 1.4 *The following implications hold constructively:*

$$\mathsf{LPO} \;\Rightarrow\; \mathsf{WLPO} \;\Rightarrow\; \mathsf{LLPO}. \tag{1.1}$$

Proof The first implication is obvious. For the second, assume $\forall n, m[\alpha_n = \alpha_m = 1 \to n = m]$. Applying WLPO to $\beta(n) := \alpha_{2n}$, we have $\forall n\,\beta_n = 0$ or $\neg\forall n\,\beta_n = 0$. Clearly, the first case yields $\forall n\,\alpha_{2n} = 0$. So assume $\neg\forall n\,\beta_n = 0$. From $\alpha_{2k+1} = 1$ one obtains $\beta_n = 0$ for all n, contradicting the latter assumption. Hence $\alpha_{2k+1} \neq 1$, whence $\alpha_{2k+1} = 0$ for all k since $\forall k\,[\alpha_{2k+1} = 0 \vee \alpha_{2k+1} = 0]$. □

Classically one has the principle

$$\forall x, y \in \mathbb{R}\,[x = y \vee x \neq y].$$

This principle entails WLPO and is thus not acceptable constructively. Many well-known theorems of classical analysis only require LPO or just WLPO. The story with LLPO, though, is much more subtle.[8] One of the best-known consequences of LLPO is $\forall x, y \in \mathbb{R}\,[x \leq y \lor y \leq x]$.

At this point it is worth mentioning that LPO is still much weaker than LEM. Indeed, it is interesting to study semi-intuitionistic systems with LPO. Particularly noteworthy seems to be the fact that adding LPO to constructive Zermelo–Fraenkel set theory, $\mathbf{CZF}$, does not change the proof-theoretic strength whereas adding LEM to $\mathbf{CZF}$ yields classical $\mathbf{ZF}$ (for details see [21, Section 2.7] in this volume).

One way to refute all of these principles is via a recursive reading of the BHK-interpretation.

1.3 The Brouwer–Heyting–Kolmogorov Interpretation

The difference between the classical and the intuitionistic understanding of the logical connectives and quantifiers is partiually well illuminated by the BHK-interpretation, to which we turn next.

In a first approach, a mathematical assertion could be construed as a meaningful statement describing a state of affairs, which traditionally is something that is either true or false. In the case of mathematical statements involving quantifiers ranging over infinite domains, however, by adopting such a view one is compelled to postulate an objective transcendent realm of mathematical objects which determines their meaning and truth value. Most schools of constructive mathematics reject such an account as unconvincing. Kolmogorov observed that the laws of the constructive propositional calculus become evident upon conceiving propositional variables as ranging over problems or tasks. The constructivist's restatement of the meaning of the logical connectives is known as the *BHK-interpretation*. It is couched in terms of an informal notion of proof. It is instructive to view such proofs as pieces of *evidence* sometimes referred to as *proof objects*.

Definition 1.5

(i) p proves $\varphi \land \psi$ iff p is pair $\langle a, b\rangle$, where a is proof for φ and b is proof for ψ.
(ii) p proves $\varphi \lor \psi$ iff p is pair $\langle n, q\rangle$, where $n = 0$ and q proves φ, or $n = 1$ and q proves ψ.
(iii) p proves $\varphi \to \psi$ iff p is a function (or rule) which transforms any proof s of φ into a proof $p(s)$ of ψ.
(iv) p proves $\neg\varphi$ iff p proves $\varphi \to \bot$.
(v) p proves $(\exists x \in A)\varphi(x)$ iff p is a pair $\langle a, q\rangle$ where a is a member of the set A and q is a proof of $\varphi(a)$.

[8] LPO and LLPO can be separated at the level of full intuitionistic Zermelo–Fraenkel. For this and more references see [6, Section 9].

(vi) p proves $(\forall x \in A)\varphi(x)$ iff p is a function (rule) such that for each member a of the set A, $p(a)$ is a proof of $\varphi(a)$.

(vii) p proves $\perp$ is impossible, so there is no proof of $\perp$.

Many objections can be raised against the above definition. The explanations offered for implication and universal quantification are notoriously imprecise because the notion of function (or rule) is left unexplained. Another problem is that the notions of set and set membership are in need of clarification. But in practice these rules suffice to codify the arguments which mathematicians want to call constructive. Note also that the above interpretation (except for $\perp$) does not address the case of atomic formulas.

Definition 1.6 We say that a formula φ is *valid under the BHK-interpretation*, if a construction (or proof object) p can be exhibited that is a proof of φ in the sense of the BHK-interpretation.

Example 1.7 Here are some examples of the BHK-interpretation. We sometimes use λ-notation for functions.

(i) The identity map, $\lambda x.x$, is a proof of any proposition of the form $\varphi \to \varphi$ since $(\lambda x.x)(p) = p$.

(ii) A proof of $\varphi \wedge \psi \to \psi \wedge \varphi$ is provided by the function $f(\langle a, b\rangle) = \langle b, a\rangle$.

(iii) Perhaps a bit wondrous, but any function is a proof of $\perp \to \varphi$ as $\perp$ has no proof.

(iv) $(*)$ $(\varphi \to \psi) \to [(\psi \to \theta) \to (\varphi \to \theta)]$ is valid under the BHK-interpretation. Note that the latter entails as a special case the law of contraposition,

$$(\varphi \to \psi) \to (\neg\psi \to \neg\varphi)$$

as $\neg\vartheta$ is $\vartheta \to \perp$. To find a BHK-proof of $(*)$, assume that f proves $\varphi \to \psi$, g proves $\psi \to \theta$, and p proves φ. Then $f(p)$ proves ψ, and hence $g(f(p))$ proves θ. Consequently, $\lambda x.g(f(x))$ proves $\varphi \to \theta$, and therefore $\lambda g.\lambda x.g(f(x))$ proves $(\psi \to \theta) \to (\varphi \to \theta)$. Thus, $\lambda f.\lambda g.\lambda x.g(f(x))$ is a proof of $(*)$.

(v) The law of excluded middle is not valid under any reasonable reading of the BHK-interpretation. Given a sentence θ, we might not be able to find a proof of θ nor a proof of $\neg\theta$. Wondrously, the double negation of that principle is valid under the BHK-interpretation. This may be seen as follows. Suppose g proves $\neg(\psi \vee \neg\psi)$. One easily constructs functions $\mathfrak{f}_0$ and $\mathfrak{f}_1$ such that $\mathfrak{f}_0$ transforms a proof of ψ into a proof of $\psi \vee \neg\psi$ and $\mathfrak{f}_1$ transforms a proof of $\neg\psi$ into a proof of $\psi \vee \neg\psi$, respectively. Thus, $\lambda a.g(\mathfrak{f}_0(a))$ is a proof of $\neg\psi$ while $\lambda b.g(\mathfrak{f}_1(b))$ is a proof of $\neg\psi \to \perp$. Consequently, $g(\mathfrak{f}_1(\lambda a.g(\mathfrak{f}_0(a))))$ is a proof of $\perp$. As a result, $\lambda g.g(\mathfrak{f}_1(\lambda a.g(\mathfrak{f}_0(a))))$ proves $\neg\neg(\psi \vee \neg\psi)$ for any formula ψ.

1.4 Natural Deductions

Constructive mathematics, just as classical mathematics, is mostly carried out informally by humans. Going back to the BHK-interpretation provides a good tool for testing whether a piece of mathematical reasoning holds under the constructive lens. Still, it may be desirable and convenient to have a set of formal logical rules available, should questions about the constructive validity of a proof be raised. Even with the BHK-interpretation at one's disposal, doubts can arise, due to BHK being based on an unexplained notion of function.

This section presents two formal systems of axioms and rules for intuitionistic logic, the natural deduction calculus invented by Gentzen and the intuitionistic Hilbert-style calculus.

Definition 1.8 In the following it is assumed that we are given a language $\mathcal{L}$ of predicate logic (also called first-order logic) with equality $=$. The logical primitives are $\wedge, \vee, \rightarrow, \bot, \forall, \exists$, where $\bot$ stands for absurdity and the negation $\neg\psi$ of a formula ψ is defined by $\psi \rightarrow \bot$. Such a language is further specified by its constant, function and relation symbols, together with their arities in the latter two cases. It is convenient to use different symbols for free $a, b, c, a_0, a_1, a_2, \ldots$ and *bound* $x, y, z, x_0, x_1, x_2, x_3, \ldots$ variables.[9]

Terms are generated from constants and free variables via function symbols. Bound variables aren't terms. We use the convention that metavariables $s, t, s_0, s_1, \ldots$ range over terms.

Formulas are then mostly defined as usual, the exception being the quantifiers. It is convenient to use notations such as $\phi(), \psi(), \theta(), \ldots$ as metavariables ranging over finite strings made up of symbols from $\mathcal{L}$ and a place-holder symbol $\star$, where $\star$ is assumed not to belong to $\mathcal{L}$. They were called *nominal forms* by Schütte [23]. The purpose of these nominal forms is to describe substitutions succinctly. If $\mathfrak{s}$ is any string of symbols, $\phi(\mathfrak{s})$ is obtained from $\phi()$ by replacing every occurrence of $\star$ by $\mathfrak{s}$.

The formation rule for formulas commencing with a quantifier is the following. If $\phi(a)$ is a formula with free variable a and x is a bound variable that does not occur in $\phi(a)$, then $\forall x\phi(x)$ and $\exists x\phi(x)$ are formulas.

Note that in a formula, a bound variable x can only occur within the scope of a quantifier $\forall x$ or $\exists x$.

We say that the variable a is *fully indicated* in $\phi(a)$ if a does not occur in $\phi()$.

A *closed formula* is one without free variables.

[9] Using different symbols for free and bound variables is not absolutely essential but it is extremely useful and simplifies arguments a great deal. Terms can be freely substituted for both kinds of variables since variables occurring in them are always free and thus cannot be captured by quantifiers.