

# Constructivism and Proof Theory (draft)

A.S.Troelstra

September 29, 2003

## Contents

<b>1</b>	<b>Introduction</b>	<b>6</b>
1.1	Constructivism . . . . .	6
1.2	Proof Theory . . . . .	7
<b>2</b>	<b>Intuitionistic Logic</b>	<b>8</b>
2.1	The BHK-interpretation . . . . .	8
2.2	Natural deduction . . . . .	9
2.3	The systems <b>Hi</b> and <b>Hc</b> . . . . .	11
2.4	Metamathematics of <b>I</b> , and the relation to <b>C</b> . . . . .	12
<b>3</b>	<b>Semantics of intuitionistic logic</b>	<b>13</b>
3.1	I-completeness . . . . .	13
3.2	Kripke semantics . . . . .	13
3.3	Topological and algebraic semantics . . . . .	15
<b>4</b>	<b>Intuitionistic arithmetic</b>	<b>16</b>
4.1	Realizability . . . . .	17
4.2	Characterization of realizability . . . . .	18
<b>5</b>	<b>Constructive Mathematics</b>	<b>19</b>
5.1	Bishop's constructive mathematics BCM . . . . .	19
5.2	Constructive recursive mathematics CRM . . . . .	20
5.3	Intuitionism INT . . . . .	21
5.4	Lawless sequences and I-validity . . . . .	23
5.5	Comparison of BCM, CRM and INT . . . . .	24
<b>6</b>	<b>Proof Theory of first-order logic</b>	<b>24</b>
6.1	The Gentzen systems $G_c$ and $G_i$ . . . . .	25
6.2	Cut elimination . . . . .	26
6.3	Natural Deduction and normalization . . . . .	28
6.4	The Tait-calculus . . . . .	29
<b>7</b>	<b>Proof Theory of mathematical theories</b>	<b>30</b>
7.1	The language . . . . .	31
7.2	Order types . . . . .	31
7.3	Truth-complexity . . . . .	32
7.4	The proof-theoretic ordinal of a theory . . . . .	33
7.5	The method . . . . .	33

## Keywords

Algebraical semantics, axiom of open data, bar induction, Bishop's constructive mathematics, Brouwer-Heyting-Kolmogorov interpretation, choice sequence, Church's thesis, Church-Kleene ordinal, constructive recursive mathematics, constructivism, continuity axioms, contraction, cut elimination, disjunction property, elimination rule, explicit definability property, finitism, Gentzen system, Glivenko theorem, Gödel-Gentzen translation, Heyting arithmetic, Hilbert's programme, Hilbert-type system, I-completeness, introduction rule, intuitionism, intuitionistic arithmetic, intuitionistic logic, inversion lemma, Kreisel-Lacombe-Shoenfield-Tsejtin theorem, Kripke forcing, Kripke semantics, lawless sequence, Markov's principle, Markov's rule, natural deduction, negative formula, normalization, order type, ordinal notation, predicativism, proof theory, proof-theoretic ordinal, realizability, semi-formal system, singular cover, Specker sequence, Tait calculus, topological semantics, truth-complexity, typed lambda-calculus, weakening.

## Glossary

**algebraical semantics** Defined by valuations as classical semantics, but the values are not just 0,1, but elements of a Heyting algebra.

**almost negative formula** is an arithmetical formula not containing disjunction, and  $\exists$  only as part of a subformula of the form  $\exists x(t = s)$ .

**axiom of open data** An axiom in the theory of lawless sequences, stating that when a property  $A$  (not depending on choice parameters) holds for a lawless sequence  $\alpha$ , then there is an initial segment of  $\alpha$  such that all lawless sequences starting with the same initial segment have property  $A$ . There is a more general version concerns  $A$  with choice parameters.

**bar induction** An induction principle over countably branching trees closely related to transfinite induction.

**Bishop's constructive mathematics** A version of constructivism where it is required that all theorems have "numerical meaning"; in particular existence theorems  $\exists x \dots$  must be capable of being made explicit by exhibiting an object which can take the place of  $x$ .

**Brouwer-Heyting-Kolmogorov interpretation** An interpretation which explains the meaning of the logical operators by describing the proofs of logically compound statements in terms of the proofs of its immediate subformulas.

**choice sequence** An infinite sequence of natural numbers, not fixed in advance by a law or recipe, but created step by step by successive choices, possibly subject to certain restrictions. A key notion in intuitionistic analysis.

**Church's thesis** (relative to intuitionism): loosely, the statement that any definite (law-like) total number-theoretic function is partial recursive, or some closely related (more or less equivalent) statement.

**Church-Kleene ordinal** The least ordinal which cannot be represented by a recursive wellordering.

**classical arithmetic** is a first-order theory based on classical logic with equality and function symbols. The language contains 0,  $S$  (successor),  $=$  and constants for all primitive recursive functions. The axioms: 0 is not a successor, defining equations for all primitive recursive functions, the induction schema.

**constructive recursive mathematics** A version of constructivism, insisting that all mathematical objects and operations are representable as algorithms.

**constructivism** A normative demand for explicitness of the mathematical objects studied: concrete representability, or explicit definability, or presentable as mental constructions.

**continuity axioms** The typical continuity axiom in the theory of choice sequences states that if a natural number for *each* choice sequence  $\alpha$  is chosen such that a property  $A(\alpha, n)$

holds, then the  $n$  may be chosen continuously in  $\alpha$ , i.e., depending on an initial segment of  $\alpha$  only. Other continuity axioms are mostly weakenings or strengthenings of this principle.

**contraction** A rule in certain Gentzen systems, concluding  $\Gamma, A \Rightarrow \Delta$  from  $\Gamma, A, A \Rightarrow \Delta$ , and similarly with  $A$  on the right.

**cut elimination** A systematic procedure for removing the Cut rule from deductions in Gentzen systems.

**Cut rule** A rule in Gentzen systems with premises of the form  $\Gamma \Rightarrow \Delta, A$  and  $A, \Gamma' \Rightarrow \Delta'$  and conclusion  $\Gamma\Gamma' \Rightarrow \Delta\Delta'$ . There are variants.

**disjunction property** A rule which hold for most well-known formalisms based on intuitionistic logic: if  $A \vee B$  is derivable, then either  $A$  or  $B$  is derivable ( $A \vee B$  closed).

**elimination rule** for an operator  $X$ . In natural deduction systems a rule which is a counterpart to the introduction rules and tells us that the introduction rules give all possibilities for introducing the operator  $X$ .

**explicit definability property** A rule which holds for many well-known formalisms based on intuitionistic logic: if a sentence  $\exists x A(x)$  is derivable, there is a closed term  $t$  such that  $A(t)$  is derivable. In particular this usually holds if  $x$  ranges over the natural numbers; in other cases it only holds if the theory considered contains enough terms for naming elements in the range of  $x$ .

**Extended Church's thesis** is a generalization of Church's thesis to functions defined over subsets of the natural numbers defined by an almost negative formula.

**fan theorem** A principle in the intuitionistic theory of choice sequences stating, essentially, that the binary tree (the Cantor discontinuum) is compact. It is a special case of bar induction. Usually combined with a continuity axiom: if a natural number for *each* choice sequence  $\alpha$  in the binary tree is chosen such that a property  $A(\alpha, n)$  holds, then there is an  $m$  such that the  $n$  may be chosen so as to depend on an initial segment of  $\alpha$  of length  $m$  only.

**finitism** A normative view in the foundations of mathematics insisting on concrete representability of mathematical objects as finite combinatorial structures.

**Gentzen system** A type of system with axioms and rules, deducing a sequent expression from certain sequents which are its premises; an axiom is a rule with empty collection of premises. For each logical operator there left- and right rules.

**Glivenko theorem** A negated formula of propositional logic is provable in classical logic if and only if it is intuitionistically provable.

**Gödel-Gentzen translation** An embedding of classical arithmetic into intuitionistic arithmetic, or of classical predicate logic into intuitionistic predicate logic, which eliminates disjunction and existence.

**Heyting algebra.** A lattice with top and bottom with a binary operation  $\rightarrow$  on lattice elements satisfying  $a \rightarrow b \leq c$  iff  $a \leq b \wedge c$  (*conj* is the meet-operation of the lattice). In a complete Heyting algebra all joins and meets exist.

**Heyting arithmetic** Defined like classical arithmetic, but based on intuitionistic logic instead of classical logic.

**Hilbert's program** aims at obtaining a proof of consistency for formalized mathematical theories by finitistic means.

**Hilbert-type system** A system based on a number of axiom schemata and (for propositional logic) the rule of modus ponens: from  $A \rightarrow B, A$  deduce  $B$ , and for predicate logic in addition the rule of generalization: if  $A(x)$  has been deduced without assumptions, then we may deduce  $\forall x A(x)$ . There are variants with more rules, but none of these have a mechanism for discharging or closing assumptions as in natural deduction.

**I-completeness** This is the statement that if a formula of predicate logic holds in all intuitionistically meaningful structures (structure defined as in classical model theory), then it is intuitionistically derivable in intuitionistic predicate logic. Similarly for propositional logic.

**induction (axiom) schema** is the schema  $A(0) \wedge \forall x(A(x) \rightarrow A(Sx)) \rightarrow \forall x A(x)$ , where  $S$  is the successor function, and  $A$  an arbitrary formula of the theory under consideration.

**interpretational proof theory** studies syntactically defined interpretations from a theory into another theory.

**introduction rule** A rule in natural deduction systems which obtains a deduction of a formula  $A$  with principal operator  $X$  from deductions of the immediate subformulas of  $X$ .

**intuitionism** A foundational point of view, due to the Dutch mathematician L.E.J. Brouwer, insisting that mathematics is a mental construction throughout; a consequent development of this point of view leads to a mathematical practice which deviates from classical mathematics.

**intuitionistic arithmetic** Similar to classical first-order arithmetic but based on intuitionistic logic instead.

**intuitionistic logic** The logic which is the basic of most constructive formalisms; an axiomatization is obtained by dropping from one of the usual formalisms for classical logic the principle of the excluded third or the double negation law.

**inversion lemma**

**Kreisel-Lacombe-Shoenfield-Tsejtin theorem** in constructive recursive mathematics states that each map from a complete separable metric space to a separable metric space is continuous.

**Kripke semantics**

**lawless sequence** A special type of choice sequence, where at any time only a finite initial segment is known, without restrictions on future choices.

**Markov's principle** states that if it is impossible that an algorithm does not terminate, then it does terminate. In the language of arithmetic this may be expressed as  $\neg\neg\exists x(t(x) = 0) \rightarrow \exists x(t(x) = 0)$ .

**Markov's rule** is the rule which states that if the premise of an instance of Markov's principle is derivable, then so is the conclusion.

**natural deduction** A (class of) deduction system(s) for first order logic without axioms. For each logical operator  $*$  there are rules telling us how to deduce formulas with  $*$  as principal operator (the introduction rules for  $*$ ) and elimination rules for  $*$  which tell us how to draw conclusions from formulas with  $*$  as principal operator. Typically, the elimination rules tell us that the corresponding introduction rules exhaust the possibilities for introducing  $*$ . Characteristic for natural deduction systems is the possibility of making deductions from open assumptions which are later closed (discharged).

**negative formula** In arithmetic, a formula which does not contain disjunctions or existence quantifiers; for predicate logic, we ask in addition that prime formulas  $P$  only occur under a negation (i.e., for our choice of primitives, in a subformula of the form  $P \rightarrow \perp$ ).

**normalization** A procedure for removing detours in natural deductions; a detour occurs when a formula is introduced and at the next step eliminated (namely when it occurs as main premise of an elimination rule).

**ordinal notation** A collection of constants and operations from which notations for ordinals are generated. From the constants  $0, \omega$  and the operations of successor, ordinal sum, ordinal multiplication and exponentiation we can construct notations for all ordinals below  $\varepsilon$ .

**Peano arithmetic.** Theory based on classical predicate logic, with quantifiers for the natural numbers, and constants for for all primitive recursive functions and predicates, and arithmetical axioms stating that  $0$  is never a successor, that the successor function is bijective, defining axioms for the primitive recursive functions and predicates, and the induction axiom schema.

**predicativism** A foundational point of view which accepts the natural numbers but avoids impredicative definitions of sets, that is to say a set must not be defined by reference to the totality of all sets.

**proof theory** The study of formal systems as combinatorial objects.

**proof-interpretation** See Brouwer-Heyting-Kolmogorov interpretation.

**proof-theoretic ordinal** Roughly, a least upper bound for the ordinals associated with provable well-orderings in a theory.

**realizability** An interpretation of formulas of first-order arithmetic in the language of arithmetic which may be described as a recursive analogue of the Brouwer-Heyting-Kolmogorov interpretation, and which validates Extended Church's thesis.

**sequent** An expression of the form  $\Gamma \Rightarrow \Delta$  (" $\Gamma$  yields  $\Delta$ "), where  $\Gamma$  and  $\Delta$  are finite sets, multisets, or sequences, depending on the formalism considered. Intuitively, a sequent  $A_1, A_2, \dots, A_n \Rightarrow B_1, B_2, \dots, B_m$  means " $B_1 \vee B_2 \vee \dots \vee B_m$  follows from  $A_1 \wedge A_2 \wedge \dots \wedge A_n$ ".

**singular cover** is a notion from constructive recursive mathematics. A cover  $I_0, I_1, I_2, \dots$  of intervals of the closed interval  $[0, 1]$  is singular, if the lengths of the unions  $I_0 \cup I_1 \cup \dots \cup I_n$  is bounded by a number less than 1.

**Specker sequence:** a recursive, bounded, monotone sequence of rationals which does not have a (recursive) real as a limit. This is a notion from constructive recursive mathematics.

**structural proof theory** is that part of proof theory which studies formal proofs as combinatorial objects.

**Tait calculus** A Gentzen-type system for classical predicate logic with one-sided sequents  $\Rightarrow \Delta$  ( $\Delta$  a finite set) only (hence  $\Rightarrow$  may be dropped altogether). Negation is treated as defined, and a prime formula  $Rt_1 \dots t_n$  and its negation are both treated as atomic;  $\neg \neg Rt_1 \dots t_n$  is by definition identical with  $Rt_1 \dots t_n$ . There are infinitary variants for systems containing numerical quantifiers.

**topological semantics** Defined by valuations as classical semantics, but with values not from  $\{0, 1\}$ , but from the open sets of a topological space. A special case of algebraical semantics.

**transfinite induction** over an order  $\prec$  is the principle: if  $A(x)$  follows from  $\forall y \prec x A(y)$ , then  $A(x)$  for all  $x$  in the field of the relation  $\prec$ , for all formulas  $A$  in the theory under consideration.

**typed lambda-calculus** A language for terms built-up with application and functional abstraction, denoting functions on the type structure over given, basic sets.

**weakening** A rule in Gentzen systems of the form: from  $\Gamma \Rightarrow \Delta$  conclude  $\Gamma, A \Rightarrow \Delta$ , or  $\Gamma \Rightarrow A, \Delta$ .

## Summary

Introduction to the constructive point of view in the foundations of mathematics, in particular intuitionism due to L.E.J. Brouwer, constructive recursive mathematics due to A.A. Markov, and Bishop's constructive mathematics. The constructive interpretation and formalization of logic is described. For constructive (intuitionistic) arithmetic Kleene's realizability interpretation is given; this provides an example of the possibility of a constructive mathematical practice which diverges from classical mathematics. The crucial notion in intuitionistic analysis, choice sequence, is briefly described and some principles which are valid for choice sequences are discussed. The second half of the article deals with some aspects of proof theory, i.e. the study of formal proofs as combinatorial objects. Gentzen's fundamental contributions are outlined: his introduction of the so-called Gentzen systems which use sequents in stead of formulas and his result on first-order arithmetic showing that (suitably formalized) transfinite induction up to the ordinal  $\varepsilon$  cannot be proved in first-order arithmetic.

# 1 Introduction

## 1.1 Constructivism

Since the beginning of the twentieth century several positions w.r.t. the foundations of mathematics have been formulated which might be said to be versions of constructivism.

Typically, a constructivist view demands of mathematics some form of explicitness of the objects studied, they must be concretely representable, or explicitly definable, or capable of being viewed as mental constructions. We distinguish five variants of constructivism in this article: finitism, predicativism, intuitionism (INT), constructive recursive mathematics (CRM), and Bishop's constructive mathematics (BCM). We will be short about finitism and predicativism, and concentrate on the other three instead.

*Finitism* insists on concrete representability of the objects of mathematics and avoids the higher abstractions. Thus particular functions from  $\mathbb{N}$  to  $\mathbb{N}$  are considered, but the notion of an arbitrary function from  $\mathbb{N}$  to  $\mathbb{N}$  is avoided, etc. This curtails the use of logic, in particular the use of quantifiers over infinite domains. Infinite domains are regarded as indefinitely extendable finite domains rather than as completed infinite totalities. Finitism is not only of interest as a version of constructivism, but also as a key ingredient in Hilbert's original program: Hilbert wanted to establish consistency of formal mathematical theories by 'finitistic' means, since he regarded these as evidently justified and uncontroversial (see also below under 1.2).

*Predicativism* concentrates on the explicitness and non-circular character of definitions. As a rule, in the predicativist approach the natural numbers are taken for granted; but sets of natural numbers have to be explicitly defined, and in defining a mathematical entity  $A$  say, the definition should not refer to the totality of objects of which  $A$  is an element.

### Intuitionism (INT)

Intuitionism, as it is understood here, is due to the Dutch mathematician L.E.J. Brouwer (1881–1966). The basic tenets of intuitionism may be summarily described as follows.

1. Mathematics is not formal; the objects of mathematics are mental constructions in the mind of the (ideal) mathematician. Only the thought constructions of the (idealized) mathematician are exact.
2. Mathematics is independent of experience in the outside world, and mathematics is in principle also independent of language. Communication by language may serve to suggest similar thought constructions to others, but there is no guarantee that these other constructions are the same. (This is a solipsistic element in Brouwer's philosophy.)
3. Mathematics does not depend on logic; on the contrary, logic is part of mathematics.

The first item not only leads to the rejection of certain theorems of classical logic, but also opens a possibility for admitting deviant objects, the "forever incomplete" choice sequences. Just as for CRM, the mathematical theories of INT are not simply sub-theories of their classical counterparts, but may actually be incompatible with the corresponding classical theory.

### Constructive Recursive Mathematics (CRM)

A.A. Markov (1903–1979) formulated in 1948–49 the basic ideas of constructive recursive mathematics (CRM for short). They are the following:

1. Objects of constructive mathematics are constructive objects, concretely: words in various alphabets.
2. The abstraction of potential existence (potential realizability) is admissible but the abstraction of actual infinity is not allowed. Potential realizability means e.g., that we may regard addition as a well-defined operation for all natural numbers, since we know how to complete it for arbitrarily large numbers. This admissibility is taken to include acceptance of ‘Markov’s Principle’ : if it is impossible that an algorithmic computation does not terminate, it does in fact terminate. The rejection of actual infinity is tantamount to the rejection of classical logic.
3. A precise notion of algorithm is taken as a basis (Markov chose for this his own notion of ‘Markov-algorithm’). Since Markov-algorithms are encoded by words in suitable alphabets, they are objects of CRM; conversely, each word in some definite alphabet may be interpreted as a Markov algorithm.
4. Logically compound statements have to be interpreted so as to take the preceding points into account.

Markov’s principle holds neither in INT nor in BCM.

### **Bishop’s Constructive Mathematics (BCM)**

Errett Bishop (1928–1983) formulated his version of constructive mathematics around 1967. There is a single “ideological” principle underlying BCM:

1. proofs of existential statements must provide a method of constructing the object satisfying the specifications,

and three more pragmatic guiding rules for the development of BCM:

2. avoid concepts defined in a negative way;
3. avoid defining irrelevant concepts — that is to say, among the many possible classically equivalent, but constructively distinct definitions of a concept, choose the one or two which are mathematically fruitful ones, and disregard the others;
4. avoid pseudo-generality, that is to say, do not hesitate to introduce an extra assumption if it facilitates the theory and the examples one is interested in satisfy the assumption.

Starting from the principles outlined above, three distinct versions of mathematics have been developed, which differ notably in their respective theories of the continuum, as will be explained further on.

## **1.2 Proof Theory**

Proof theory owes its origin to Hilbert’s Program, i.e., the project of establishing freedom of contradiction for formally codified (substantial parts of) mathematics, using elementary, “evident” reasoning (finitistic reasoning). As shown by Gödel, in its original form this program was bound to fail. However, a modification of the program has been successful; one then asks to establish consistency using “evident” means of proof, possibly stronger than the system whose consistency is to be established.

*Structural proof theory* studies formal mathematical (logical) proofs as combinatorial structures; various styles of formalization are compared.

*Hilbert-Schütte* style proof theory takes its starting point from Gentzen's consistency proof for arithmetic, and compares formal systems with respect to their proof-theoretic strength, by analyzing the structure of suitably devised deduction systems.

*Interpretational proof theory* compares formalisms via syntactic translations/interpretations. we shall encounter various examples below.

## 2 Intuitionistic Logic

Although Brouwer was positively averse to formalization of mathematics, he was nevertheless the first to formulate and establish some principles of intuitionistic logic. But it were Kolmogorov in 1925 and Heyting in 1930 who demonstrated that intuitionistic logic could be studied as a formalism. But formalizations of intuitionistic logic need an informal interpretation to justify them as codifications of *intuitionistic* logic. Heyting (1930, 1934) and Kolmogorov (1932) each developed such an interpretation; their interpretations were later to be seen to be essentially equivalent. Heyting in particular built on some of Brouwer's early papers.

### 2.1 The BHK-interpretation

The need for a different logic in the setting of INT, BCM and CRM becomes clear by considering some informal examples.

The following is not acceptable as a constructive definition of a natural number:

$$n = 2 \text{ if } R \text{ holds, } n = 3 \text{ if } \neg R \text{ holds.}$$

where  $R$  stands for some mathematically unsolved problem, e.g.,  $R =$  "The Riemann hypothesis holds". This is not a constructive definition because we cannot identify  $n$  with one of the explicitly given natural numbers  $0, 1, 2, 3, 4, \dots$ ; for such an identification to be possible, we have to decide whether  $R$  or  $\neg R$  holds, i.e., to decide the Riemann hypothesis. Note that the definition becomes acceptable as soon as problem  $R$  has been solved.

*Example of a non-constructive proof.* Consider the following statement: there exist irrational  $a, b$  such that  $a^b$  is rational. This statement has a very simple proof:  $\sqrt{2}^{\sqrt{2}}$  is either rational or irrational. In the first case, take  $a = b = \sqrt{2}$ . In the second case, take  $a = \sqrt{2}^{\sqrt{2}}, b = \sqrt{2}$ . The proof is obviously non-constructive, since it does not permit us to compute  $a$  with any desired degree of accuracy. A constructive proof of the statement is possible, for example, by an appeal to a non-trivial theorem of Gelfond: if  $a \notin \{0, 1\}$ ,  $a$  algebraic,  $b$  irrational algebraic, then  $a^b$  is irrational, even transcendental.

INT, BCM and CRM have the same logical basis, called intuitionistic logic or constructive logic, and which is a subsystem of classical predicate logic. The standard informal interpretation of logical operators in intuitionistic logic is the so-called *proof-interpretation* or *Brouwer-Heyting-Kolmogorov interpretation* (*BHK-interpretation* for short). The formalization of intuitionistic logic started before this interpretation was actually formulated, but it is preferable to discuss the BHK-interpretation first since it facilitates the understanding of the more technical results.

We use capitals  $A, B, C, \dots$  for arbitrary formulas. Our logical operators are  $\wedge, \vee, \rightarrow, \perp, \forall, \exists$ . We treat  $\neg A$  as an abbreviation of  $A \rightarrow \perp$ .  $\equiv$  denotes identity of strings,  $:\equiv$  is the definition symbol. If  $\mathcal{E}$  is a syntactic expression, we write  $\mathcal{E}[x/t]$  for the result of substituting the term  $t$  for the free variable  $x$  in  $\mathcal{E}$ ; it is tacitly assumed that  $t$  is free for  $x$  in  $\mathcal{E}$ , that is to say, no variable free in  $t$  becomes bound after substitution. We often use a more informal notation: if  $\mathcal{E}(x)$  has been introduced in the discourse as an expression  $\mathcal{E}$  with some free occurrences of the variable  $x$ , we write  $\mathcal{E}(t)$  for  $\mathcal{E}[x/t]$ .

On the BHK-interpretation, the meaning of a statement  $A$  is given by explaining what constitutes a proof of  $A$ , and *proof of  $A$*  for logically compound  $A$  is explained in terms of what it means to give a proof of its constituents. Thus:

1. A proof of  $A \wedge B$  is given as a pair of proofs  $\langle p, q \rangle$ ,  $p$  a proof of  $A$ ,  $q$  a proof of  $B$ .
2. A proof of  $A \vee B$  is of the form  $\langle 0, p \rangle$ ,  $p$  a proof of  $A$ , or  $\langle 1, q \rangle$ ,  $q$  a proof of  $B$ .
3. A proof of  $A \rightarrow B$  is a construction  $q$  which transforms any proof  $p$  of  $A$  into a proof  $q(p)$  of  $B$ .
4. Absurdity  $\perp$  ('the contradiction') has no proof; a proof of  $\neg A$  is a construction which transforms any supposed proof of  $A$  into a proof of  $\perp$ .

In the quantifier clauses, we assume the individual variables to range over a domain  $D$ ; the fact that  $d \in D$  for some  $d$  is not supposed to need further proof. (This is sometimes expressed by calling  $D$  a *basic* domain;  $\mathbb{N}$  is an example.)

5. A proof  $p$  of  $\forall x A(x)$  is a construction transforming any  $d \in D$  into a proof  $p(d)$  of  $A(d)$ .
6. A proof  $p$  of  $\exists x A(x)$  is a pair  $\langle d, q \rangle$  with  $d \in D$ ,  $q$  a proof of  $A(d)$ .

The concepts of *proof* and *construction* in these explanations are to be taken as primitive; "proof" is not to be identified with any notion of deduction in any formal system. Obviously, the constructions in the clauses for implication and the universal quantifier are (constructive) functions.

Let us write  $\lambda x.t(x)$  for  $t(x)$  as a function of  $x$  (so  $(\lambda x.t(x))(d) = t(d)$ ). As an example of the BHK-interpretation, let us argue that  $\neg\neg(A \vee \neg A)$  is valid on this interpretation.

- (i) If  $c$  proves  $A$ , then  $\langle 0, c \rangle$  proves  $A \vee \neg A$ ;
- (ii) if  $d$  proves  $\neg A$ , then  $\langle 1, d \rangle$  proves  $A \vee \neg A$ ;
- (iii) assume  $b$  proves  $\neg(A \vee \neg A)$ ;
- (iv) if  $c$  proves  $A$ , then  $b\langle 0, c \rangle$  proves  $\perp$ , hence  $\lambda c.b\langle 0, c \rangle$  proves  $\neg A$  (i, iii),
- (v) if  $d$  proves  $\neg A$ , then  $b\langle 1, d \rangle$  proves  $\perp$  (ii), hence  $b\langle 1, \lambda c.\langle b\langle 0, \rangle \rangle \rangle$  proves  $\perp$  (iv,v),
- (vii)  $\lambda b.b\langle 1, \lambda c.\langle 0, c \rangle \rangle$  proves  $\neg(A \vee \neg A) \rightarrow \perp$ , or  $\neg\neg(A \vee \neg A)$  (iii,vi).

## 2.2 Natural deduction

We now turn to a deduction system for constructive logic which is readily motivated by the BHK-interpretation, namely *Natural Deduction*. The following two rules for  $\rightarrow$  are obviously valid on the basis of the BHK-interpretation:

- (a) If, starting from a hypothetical (unspecified) proof  $u$  of  $A$ , we can find a proof  $t(u)$  of  $B$ , then we have in fact given a proof of  $A \rightarrow B$  (without the assumption that  $u$  proves  $A$ ). This proof may be denoted by  $\lambda u.t(u)$ .
- (b) Given a proof  $t$  of  $A \rightarrow B$ , and a proof  $s$  of  $A$ , we can apply  $t$  to  $s$  to obtain a proof of  $B$ . For this proof we may write  $\text{App}(t, s)$  or  $ts$  ( $t$  applied to  $s$ ).

These two rules yield a formal system for intuitionistic implication logic, the Natural Deduction system  $\rightarrow\text{Ni}$  ( $\text{Ni}$  restricted to implication), with two rules for constructing deductions:

$$\frac{[A]^u \quad \vdots \quad B}{A \rightarrow B} \rightarrow\text{I}, u \qquad \frac{A \rightarrow B \quad \vdots \quad B}{B} \rightarrow\text{E}$$

The first rule reads as follows:  $\dot{\cdot}$  represents an argument which from a hypothesis or open assumption  $A$  (labeled  $u$ ) obtains  $B$ . (The argument may contain other open assumptions, not shown, besides  $A$ .) Then the  $\rightarrow$ -introduction rule concludes  $A \rightarrow B$  and the assumption  $A$  is now *closed* (not any longer active), or the hypothesis  $A$  has been *discharged*. In other words, the conclusion  $A \rightarrow B$  does not depend on the assumption  $A$ . In the  $\rightarrow$ -elimination rule  $\rightarrow$ E, deductions of  $A \rightarrow B$  and of  $A$  are combined into a deduction of  $B$ . An example of a deduction is shown below on the left:

$$\frac{\frac{\frac{A \rightarrow (A \rightarrow B)^v \quad A^u}{A \rightarrow B} \quad A^u}{\frac{B}{A \rightarrow B} u} v}{(A \rightarrow (A \rightarrow B)) \rightarrow (A \rightarrow B)} \quad \frac{\frac{\frac{v: A \rightarrow (A \rightarrow B) \quad u: A}{vu: A \rightarrow B} \quad u: A}{(vu)u: B} \quad \lambda u.(vu)u: A \rightarrow B}{\lambda v \lambda u.(vu)u: (A \rightarrow (A \rightarrow B)) \rightarrow (A \rightarrow B)}$$

The tree on the right is the same as on the left, but annotated with notations for the constructions corresponding to it according to the BHK-interpretation (application for an application of rule  $\rightarrow$ E, abstraction for an application of  $\rightarrow$ I, deduction variables for the hypothetical constructions associated with assumptions).

We may also read the tree on the right as a tree of formulas with to each formula associated a term of a typed-lambda calculus; the formulas then are seen as types. It is to be noted that not only the term shown gets associated to a type (= formula), but that also each subterm of the term shown has a type (since the subterms always appear higher up in the tree).

If we regard the formula (Type) of each expression to be inseparable from the expression, and moreover we treat  $t: A$  and  $t^A$  as synonym, we may write the conclusion as of the tree in full as

$$\lambda v^{A \rightarrow (A \rightarrow B)}.(\lambda u^A.((v^{A \rightarrow (A \rightarrow B)} u^A)^{A \rightarrow B} u^A)^B)^{A \rightarrow B}: (A \rightarrow (A \rightarrow B)) \rightarrow (A \rightarrow B).$$

Here we have also indicated the types of the subterms. Clearly, the lambda-expression completely encodes the proof tree and may be regarded as an alternative notation for this tree (there is a good deal of redundancy in this).

A more formal, inductive, definition of deductions in the full system **Ni** for intuitionistic first-order logic (without equality) is as follows.

*Basis.* The single-node tree with label  $A$  (i.e., a single occurrence of  $A$ ) is a (natural) *deduction* from the open assumption  $A$ ; there are no closed assumptions.

*Inductive step.* Let  $\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3$  be given deductions. A (natural) *deduction*  $\mathcal{D}$  may be constructed according to one of the rules below. The classes  $[A]^u$ ,  $[B]^v$ ,  $[\neg A]^u$  below contain open assumptions of the deductions of the premises of the final inference, but are closed in the whole deduction. For  $\wedge, \vee, \rightarrow, \forall, \exists$  we have *introduction rules* (*I-rules*) and *elimination rules* (*E-rules*). For absurdity or falsehood  $\perp$  there is only a single rule,  $\perp_i$ .

$$\begin{array}{c} \mathcal{D}_1 \\ \frac{\perp}{A} \perp_i \\ \\ \mathcal{D}_1 \quad \mathcal{D}_2 \\ \frac{A \quad B}{A \wedge B} \wedge I \\ \\ \mathcal{D}_1 \\ \frac{A \wedge B}{A} \wedge E_R \\ \\ \mathcal{D}_1 \\ \frac{A \wedge B}{B} \wedge E_L \\ \\ [A]^u \\ \mathcal{D}_1 \\ \frac{B}{A \rightarrow B} \rightarrow I, u \\ \\ \mathcal{D}_1 \quad \mathcal{D}_2 \\ \frac{A \rightarrow B \quad A}{B} \rightarrow E \\ \\ \mathcal{D}_1 \quad \mathcal{D}_2 \quad \mathcal{D}_3 \\ \frac{A \vee B \quad C \quad C}{C} \vee E, u, v \\ \\ \mathcal{D}_1 \\ \frac{A}{A \vee B} \vee I_R \\ \\ \mathcal{D}_1 \\ \frac{B}{A \vee B} \vee I_L \end{array}$$

$$\begin{array}{c}
\mathcal{D}_1 \\
\frac{A[x/y]}{\forall x A} \forall I \\
\\
\mathcal{D}_1 \quad \mathcal{D}_2 \\
\frac{A[x/t]}{\exists x A} \exists I \quad \frac{[A[x/y]]^u}{\exists x A} \frac{C}{\exists E, u}
\end{array}
\quad
\begin{array}{c}
\mathcal{D}_1 \\
\frac{\forall x A}{A[x/t]} \forall E \\
\\
\mathcal{D}_1 \quad \mathcal{D}_2 \\
\frac{[A[x/y]]^u}{C} \frac{C}{\exists E, u}
\end{array}$$

In  $\exists E$ :  $y \equiv x$  or  $y \notin \text{FV}(A)$ , and  $y$  not free in  $C$  nor in any assumption open in  $\mathcal{D}_2$  except in  $[A[x/y]]^u$ . In  $\forall I$ :  $y \equiv x$  or  $y \notin \text{FV}(A)$ , and  $y$  not free in any assumption open in  $\mathcal{D}_1$ . To obtain the classical system **Nc** we add the more general *classical absurdity rule*  $\perp_c$

$$\begin{array}{c}
[\neg A]^u \\
\mathcal{D}_1 \\
\frac{\perp}{A} \perp_{c,u}
\end{array}$$

The informal argument of the correctness of the BHK-interpretation is now easily seen to correspond to the following **Ni**-proof

$$\begin{array}{c}
\frac{v: (A \vee \neg A) \rightarrow \perp \quad \frac{u: A}{\langle 0, u \rangle: A \vee \neg A}}{\langle 0, u \rangle v: \perp} \\
\frac{\lambda u. \langle 0, u \rangle v: A \rightarrow \perp}{\langle 1, \lambda u. \langle 0, u \rangle v \rangle: A \vee \neg A} \\
\frac{v: (A \vee \neg A) \rightarrow \perp \quad \langle 1, \lambda u. \langle 0, u \rangle v \rangle: A \vee \neg A}{\langle 1, \lambda u. \langle 0, u \rangle v \rangle v: \perp} \\
\frac{\langle 1, \lambda u. \langle 0, u \rangle v \rangle v: \perp}{\lambda v. \langle 1, \lambda u. \langle 0, u \rangle v \rangle v: \neg \neg (A \vee \neg A) \rightarrow \perp}
\end{array}$$

The expressions to the left of the  $:$  in this tree may be taken either as descriptions of BHK-constructions, or as terms in typed lambda calculus enriched with pairing and constants  $0, 1$ .

### 2.3 The systems **Hi** and **Hc**

For metamathematical investigations, where we often encounter proofs by induction on the length of deductions, natural deduction is not always optimal; Hilbert-type systems are often found to be more convenient. A *Hilbert-type system* is based on axioms (“primitive theorems”) and rules for deriving new theorems from theorems obtained before. If we also add a set of assumptions  $\Gamma$ , we obtain the notion of *A is derivable from  $\Gamma$* . However, assumptions are always open, there is no discharging or closing of assumptions.

The axioms for **Hi** are all formulas of one of the following forms:

$$\begin{array}{l}
A \rightarrow (B \rightarrow A), \quad (A \rightarrow (B \rightarrow C)) \rightarrow ((A \rightarrow B) \rightarrow (A \rightarrow C)); \\
A \rightarrow A \vee B, \quad B \rightarrow A \vee B; \\
(A \rightarrow C) \rightarrow ((B \rightarrow C) \rightarrow (A \vee B \rightarrow C)); \\
A \wedge B \rightarrow A, \quad A \wedge B \rightarrow B, \quad A \rightarrow (B \rightarrow (A \wedge B)); \\
\forall x A \rightarrow A[x/t], \quad A[x/t] \rightarrow \exists x A; \\
\forall x (B \rightarrow A) \rightarrow (B \rightarrow \forall y A[x/y]) \quad (x \notin \text{FV}(B), y \equiv x \text{ or } y \notin \text{FV}(A)); \\
\forall x (A \rightarrow B) \rightarrow (\exists y A[x/y] \rightarrow B) \quad (x \notin \text{FV}(B), y \equiv x \text{ or } y \notin \text{FV}(A)); \\
\perp \rightarrow A.
\end{array}$$

Rules for deductions from a *set* of assumptions  $\Gamma$  are:

- Ass      If  $A \in \Gamma$ , then  $\Gamma \vdash A$ .  
 $\rightarrow$ E     If  $\Gamma \vdash A \rightarrow B$ ,  $\Gamma \vdash A$ , then  $\Gamma \vdash B$ ;  
 $\forall$ I        If  $\Gamma \vdash A$ , then  $\Gamma \vdash \forall yA[x/y]$ , ( $x \notin \text{FV}(\Gamma)$ ,  $y \equiv x$  or  $y \notin \text{FV}(A)$ ).

Ass is known as the assumption rule;  $\rightarrow$ E is also known as *Modus Ponens* (MP), and  $\forall$ I as the rule of *Generalization* (G).

**Hc** is **Hi** plus an additional axiom schema  $\neg\neg A \rightarrow A$  (*law of double negation*). Instead of the law of double negation, one can also take the *law of the excluded middle*  $A \vee \neg A$ .

A deduction may be regarded as a finite tree, where the nodes are labeled with formulas; at the top nodes appear axioms or assumptions; the bottom node is the conclusion; the immediate predecessors of a node  $A$  are the premises of a correct rule application with conclusion  $A$ . We may furthermore assume that each node is also labeled either with the indication assumption, axiom, or the name of the rule applied to obtain it from its predecessors.

*Example of a deduction.*

$$\frac{(A \rightarrow B \vee A) \rightarrow ((B \rightarrow B \vee A) \rightarrow (A \vee B \rightarrow B \vee A)) \quad A \rightarrow B \vee A}{\frac{(B \rightarrow B \vee A) \rightarrow (A \vee B \rightarrow B \vee A)}{A \vee B \rightarrow B \vee A} \quad B \rightarrow B \vee A}$$

Equivalence of the Hilbert-type systems with Natural deduction holds in the following sense: if  $A$  is derivable from open assumptions  $\Gamma$  in the natural deduction system, then  $A$  is derivable from assumptions  $\Gamma$  in the Hilbert-type system, and vice versa. The crucial step is a proof of the so-called deduction theorem for the Hilbert-type systems:

$$\text{If } \Gamma, B \vdash A, \text{ then } \Gamma \vdash A \rightarrow B.$$

*Logic with equality.* The preceding formalisms **Hi** and **Hc** may be extended to incorporate equality by addition of ( $R$  relation symbol,  $f$  function symbol of the language)

$$\begin{aligned} &\forall x(x = x), \\ &\forall xyz(x = y \wedge z = y \rightarrow z = x), \\ &R\vec{s} \wedge \vec{s} = \vec{t} \rightarrow R\vec{t}, \\ &\vec{s} = \vec{t} \rightarrow f\vec{s} = f\vec{t}. \end{aligned}$$

## 2.4 Metamathematics of I, and the relation to C

Some striking metamathematical properties of many intuitionistic formalisms are the following properties of *closed* disjunctions and existential statements

- $\vdash A \vee B$  then  $\vdash A$  or  $\vdash B$  (the *disjunction property*);
- if  $\vdash \exists xA(x)$  then there is a term  $t$  such that  $\vdash A(t)$  (the *existence property*).

For intuitionistic first-order logic, these hold in general, not only for closed formulas.

*Relationship between C and I.* *Negative* formulas are formulas without  $\vee, \exists$ , where each atomic formula appears under a negation. The following translation  $^g$  of formulas into *negative* formulas is due to Gödel and Gentzen:

### Definition 1

$$\begin{aligned} P^g &::= \neg\neg P \text{ for atomic } P, \\ \perp^g &::= \perp, \\ (A \wedge B)^g &::= A^g \wedge B^g, \\ (A \rightarrow B)^g &::= A^g \rightarrow B^g, \\ (\forall xA)^g &::= \forall xA^g, \\ (A \vee B)^g &::= \neg(\neg A^g \wedge \neg B^g), \\ (\exists xA)^g &::= \neg\forall x\neg A^g. \end{aligned}$$

The translation may be made slightly more elegant by replacing afterwards all occurrences of  $\neg\neg\neg P$  by  $\neg P$  (for atomic  $P$ ), repeatedly, until no occurrences of  $\neg\neg\neg P$  are left. We can now readily show (lemma) that for all negative formulas  $\neg\neg A \rightarrow A$ , hence  $\neg\neg A \leftrightarrow A$ . Then we prove from this

$$\mathbf{C} \vdash A \leftrightarrow A^g; \text{ and } \mathbf{C} \vdash A \text{ iff } \mathbf{I} \vdash A^g.$$

As a consequence,  $\mathbf{I}$  and  $\mathbf{C}$  prove the same negative formulas. A variant is the translation  $^q$ ;  $A^q$  is obtained by inserting  $\neg\neg$  after each occurrence of  $\forall$ , and in front of the whole formula. The same properties hold; moreover, we obtain the Glivenko theorem for propositional logic: for all propositional formulas  $A$ ,  $\mathbf{I}p \vdash \neg A$  iff  $\mathbf{C}p \vdash \neg A$ .

### 3 Semantics of intuitionistic logic

We discuss three types of semantics, classical-style, modal-logic-style, and topological. The definition of each of these may be interpreted classically as well as intuitionistically. The classical reading of the classical-style semantics is uninteresting, since it coincides with classical semantics. The modal-logic-style (Kripke) semantics is a special case of topological semantics, which in turn is a special case of algebraic semantics; the latter is not treated in any detail here. Algebraic semantics is so general (i.e., there are so many models) that an intuitionistic completeness proof can be given for first-order logic. For the other types of semantics this only holds for intuitionistic *propositional* logic; there is a completeness proof for first-order logic relative to topological semantics, but it uses classical reasoning.

#### 3.1 I-completeness

An obvious question is, whether there is a completeness theorem for intuitionistic logic, analogous to classical completeness. At first sight the BHK-interpretation does not readily lend itself to formal semantical treatment, because of the mathematically untractable character of the basic notions (construction and constructive proof) used. But we can attempt to define validity parallel to the classical notion. Let  $F(R_1, \dots, R_n)$  be a formula of first-order predicate logic, containing predicate letters  $R_1, \dots, R_n$ . We shall say that  $F$  is *I-valid* if for *all* intuitionistically meaningful domains  $D$ , and *all* intuitionistically meaningful relations  $R_i^*$  of arity corresponding to  $R_i$ , we have that

$$F^*(R_1^*, \dots, R_n^*) \text{ is intuitionistically true}$$

where  $F^*$  is obtained by relativizing quantifiers  $\forall x, \exists y$ , to  $D$  as  $\forall x \in D, \exists y \in D$ , and substituting  $R_i^*$  for  $R_i$ . This helps, since in order to show I-completeness

$$F \text{ I-valid} \Rightarrow F \text{ provable,}$$

it is in certain cases possible to show  $F$  provable from  $F^*(R_1^*, \dots, R_n^*)$  intuitionistically true for a *limited* collection of domains and relations. (This is to be compared to the classical situation: there too, to obtain completeness, we do not need validity over *all* structures; in fact, structures with the natural numbers as domain, and  $\Delta_2^0$ -definable arithmetical predicates suffice.) This will be explained after we have introduced lawless sequences.

#### 3.2 Kripke semantics

Kripke semantics, first formulated for modal logics, also applies to intuitionistic logic.

In defining validity of sentences for various kinds of semantics for first-order logic, one either uses assignments of domain elements to individual variables, or one considers an

extended language in which names for all the domain elements are available. In the latter case, one can restrict attention to validity for sentences; in the first case, one has to define validity of formulas under assignments. We choose for the option with the extended language. As a rule we use elements of the domain as their own name.

**Definition 2** A Kripke model for intuitionistic logic is a quadruple  $\langle K, R, D, \Vdash \rangle$  where  $K$  is a set, the elements of  $K$  are called nodes,  $R$  is a partial order on  $K$ ,  $D$  is a function associating a domain set  $D(k)$  with each  $k \in K$ , and  $\Vdash$  is a binary relation between nodes of  $K$  and atomic formulas  $P$ , satisfying

$$kRk' \text{ implies } D(k) \subset D(k'), \quad k \Vdash P \text{ and } kRk' \text{ implies } k' \Vdash P.$$

We often write  $\leq$  for  $R$ , and  $k < k'$  if  $kRk'$ , but  $k \neq k'$ .  $k \Vdash P$  is pronounced "k forces P". For an easy formulation of the forcing relation, we prefer to deal with sentences in the language of predicate logic extended with constants naming the elements of the domains  $D(k)$ , for all  $k \in K$ . The forcing relation is extended to compound formulas by the following clauses:

1.  $k \Vdash A \wedge B$  iff  $k \Vdash A$  and  $k \Vdash B$ ;
2.  $k \Vdash A \vee B$  iff  $k \Vdash A$  or  $k \Vdash B$ ;
3.  $k \Vdash A \rightarrow B$  iff  $\forall k'$  (if  $kRk'$  and  $k' \Vdash A$  then  $k' \Vdash B$ );
4.  $k \not\Vdash \perp$ .
5.  $k \Vdash \forall x A(x)$  iff  $\forall k' \forall d' \in D(k')$  (If  $kRk'$  then  $k' \Vdash A(d')$ );
6.  $k \Vdash \exists x A(x)$  iff for some  $d \in D(k)$   $k \Vdash A(d)$ .

A sentence  $A$  is said to be valid in model  $\mathcal{K}$  at node  $k$  iff  $k \Vdash A$  in the model;  $A$  is valid in  $\mathcal{K}$  iff  $A$  is valid at all  $k$  in  $\mathcal{K}$ .  $A$  is Kripke-valid iff  $A$  is valid in all Kripke models  $\mathcal{K}$ . We write  $\Gamma \Rightarrow A$ , iff for all models  $\mathcal{K}$ , if all formulas of  $\Gamma$  are valid in  $\mathcal{K}$ , then  $A$  is valid in  $\mathcal{K}$ .

**Theorem 1** For Kripke models we have correctness: all formulas derivable in **I** are Kripke-valid, and completeness: Kripke-valid formulas are derivable in **I**. For predicate logic, completeness proofs unavoidably use classical reasoning.

*Example 1.* Consider a model where  $K = \{0, 1\}$ ,  $0 < 1$ , and  $1 \Vdash P$ ,  $\Vdash$  holds in no other case. Then it is readily verified that  $0 \Vdash P \vee \neg P$ . This example may be motivated by a weak counterexample: let  $P$  stand for a mathematically undecided question, 0 for our present state of knowledge. 1 represents a possible future state of knowledge where we have discovered that  $P$  holds (it may be that 1 is never reached). At 0 we have not yet learnt that  $P$  holds, but also not that  $\neg P$  holds, because that would exclude that later on we might discover that  $P$  holds after all.

*Example 2.* Let  $K \equiv \{k_0, k_1, k_2, \dots\}$  such that  $k_0 < k_1 < k_2 < \dots$ , and where  $D(i) \equiv \{0, 1, \dots, i\}$ , and  $k_i \Vdash Rj$  for all  $j < i$  and in no other case. Then it follows that

$$0 \Vdash \forall x \neg \neg Rx, \quad 0 \not\Vdash \neg \neg \forall x Rx.$$

The first follows because for each  $i \in D(k)$ , and each  $k \in \mathbb{N}$ , it holds that  $k + 1 \Vdash Ri$ ; and the second follows because for all  $k \in \mathbb{N}$  we have  $k \not\Vdash \forall x Rx$ , and this is so because  $k \not\Vdash Rk$ . As a result, we see that  $\forall x \neg \neg Rx \rightarrow \neg \neg \forall x Rx$  is not Kripke-valid. In fact, in the model  $\neg(\forall x \neg \neg Rx \rightarrow \neg \neg \forall x Rx)$  holds.

We can add function symbols and equality in several different ways. We describe only one. Equality is treated as a binary relation which at every node is interpreted by an

equivalence relation  $\sim_k$  (so  $k \Vdash d = d'$  iff  $d \sim_k d'$ , for  $d, d' \in D(k)$ , which in addition satisfies: if  $d \sim_k d' \wedge k' \geq k$  then  $d \sim_{k'} d'$ . the  $n$ -ary function symbols  $f$  are interpreted by  $n$ -ary functions  $f_k$  such that

$$\begin{aligned} &\text{If } \vec{d} \sim_k \vec{d}' \text{ then } f_k \vec{d} \sim_k f_k \vec{d}', \\ &\text{if } f_k \vec{d} \sim_k d' \text{ and } k' \geq k \text{ then } f_{k'} \vec{d} \sim_{k'} d'. \end{aligned}$$

### 3.3 Topological and algebraic semantics

**Definition 3** A topological model consists of a triple  $\mathcal{T} \equiv \langle T, D, \llbracket \cdot \rrbracket \rangle$ , where  $T$  is a topological space,  $D$  a fixed domain, and  $\llbracket \cdot \rrbracket$  a valuation function for prime sentences in  $\mathcal{L}(D)$  which assigns open sets in  $T$ .  $\llbracket \cdot \rrbracket$  is extended to arbitrary sentences by:

$$\begin{aligned} \llbracket A \wedge B \rrbracket &::= \llbracket A \rrbracket \cap \llbracket B \rrbracket, \\ \llbracket A \vee B \rrbracket &::= \llbracket A \rrbracket \cup \llbracket B \rrbracket, \\ \llbracket A \rightarrow B \rrbracket &::= \text{Int}\{x : x \in \llbracket A \rrbracket \rightarrow x \in \llbracket B \rrbracket\}, \\ \llbracket \perp \rrbracket &::= \emptyset, \\ \llbracket \forall y A \rrbracket &::= \text{Int}(\bigcap_{d \in D} \llbracket A[y/d] \rrbracket), \\ \llbracket \exists y A \rrbracket &::= \bigcup_{d \in D} \llbracket A[y/d] \rrbracket. \end{aligned}$$

A sentence  $A$  is valid in the model  $\mathcal{T}$  if  $\llbracket A \rrbracket = T$ .

Kripke semantics may be seen to be a special case of topological semantics: a Kripke model  $\langle K, \leq, D, \Vdash \rangle$  corresponds to a topological model  $\langle T, D', \llbracket \cdot \rrbracket \rangle$ , where  $T$  is the space with  $K$  as set of points and as opens the collection of upwards monotone sets;  $D' = \bigcup_{k \in K} D(k)$ . For the atomic sentences we put

$$\llbracket R(d_1, \dots, d_n) \rrbracket ::= \{k : d_1, \dots, d_n \in D(k) \wedge k \Vdash R(d_1, \dots, d_n)\}.$$

Then we readily check that

$$k \in \llbracket A \rrbracket \text{ iff } k \Vdash A.$$

Another important special case are the *Beth models*, which are, in essence, topological models over the Cantor Discontinuum  $2^{\mathbb{N}}$  (sequences of 0, 1). Validity in Beth models may also be formulated as a notion of forcing over the tree of 01-sequences, rather similar to Kripke forcing. The main differences are: there is a constant domain for the variables, and the clauses for disjunction and the existential quantifier are modified.

$k \Vdash A$  corresponds to  $V_k \subset \llbracket A \rrbracket$ , where

$$V_k ::= \{\alpha : k \text{ is an initial segment of } \alpha\}$$

Here  $\alpha$  ranges over infinite 01-sequences. Now  $k \Vdash A \vee B$  iff there are  $k_0, \dots, k_{p-1}$  such that for each  $k_i$ ,  $k_i \Vdash A$  or  $k_i \Vdash B$ , and the  $k_i$  are a *cover*, that is to say  $V_{k_0} \cup \dots \cup V_{k_{p-1}}$  is a cover of the whole space. A similar clause holds for disjunctions.

Topological semantics may be generalized still further, by considering assignments which map variables to the elements of a complete Heyting algebra. A complete Heyting algebra is a structure  $\langle H, \wedge, \vee, \bigvee, \top, \rightarrow \rangle$ , such that  $H$  is a complete Heyting algebra with  $\wedge$  as meet and  $\bigvee$  as arbitrary join,  $\top$  as maximal element, satisfying  $y \wedge \bigvee_{i \in I} x_i = \bigvee_{i \in I} (y \wedge x_i)$ , and  $x \rightarrow y \leq z$  iff  $x \leq y \wedge z$ .

The open sets of a topological space constitute a complete Heyting algebra, with intersection corresponding to  $\wedge$ ,  $\bigcup$  to  $\bigvee$ , and  $\text{Int}\{x : x \in X \rightarrow x \in Y\}$  to  $X \rightarrow Y$ .  $\text{Int}(\bigcap X)$  corresponds to  $\bigwedge X$ .

Completeness relative valuations in a suitable complete Heyting algebra can be proved easily, even *constructively*: the complete Heyting algebra required may be constructed by

applying a suitable completion operator to the Lindenbaum algebra of the equivalence classes of formulas in predicate logic.

For topological models, we can prove *classically* completeness for topological models over the Cantor-discontinuum, that is to say we have classical completeness for Beth-models.

## 4 Intuitionistic arithmetic

In intuitionistic arithmetic the variables range over the natural numbers, the logical basis is **I** with equality, there are symbols for all primitive recursive functions, the so-called defining equations for the primitive recursive functions,  $\forall x(Sx \neq 0)$ , and the induction schema. Hence among the axioms there is a symbol for the predecessor function, say *prd*, satisfying

$$\text{prd}(0) = 0, \text{prd}(Sx) = x.$$

Hence, if  $Sx = Sy$ , then by the equality axioms for functions,  $\text{prd}(Sx) = \text{prd}(Sy)$ , so  $x = y$ . hence  $Sx = Sy \rightarrow x = y$ .

One can now prove by induction  $\forall xy(x = y \vee x \neq y)$ , and hence  $\neg\neg t = s \leftrightarrow t = s$ . This has as a consequence that we can apply the Gödel-Gentzen translation in a simplified form:  $(t = s)^g \equiv (t = g)$ .

Some notable metamathematical properties of **HA** are

### Theorem 2

1. **PA** is conservative over **HA** for  $\Pi_2^0$ -sentences (i.e., all  $\Pi_2^0$ -sentences provable in **PA** are already provable in **HA**), and **HA** is closed under Markov's Rule (see below). (A  $\Pi_2^0$ -sentence is a sentence of the form  $\forall x\exists y(t = s)$ .)
2. The Disjunction Property *DP* and the Explicit Definability property *ED* hold for sentences.
3. **HA** is closed under the rule *IP* (Independence of Premise):

$$IP \quad \text{If } \neg A \rightarrow \exists xB \text{ then } \vdash \exists x(\neg A \rightarrow B) \text{ (} x \text{ not free in } A\text{)}.$$

The first property explains why intuitionistic arithmetic based on **HA** differs so little from classical arithmetic **PA** (which is just **HA** with classical logic): the bulk of the theorems has a logical complexity of at most  $\Pi_2^0$ .

*Markov's principle* (basic form) may be stated as:

$$MPR \quad \vdash \neg\neg A \rightarrow A \text{ (} A \text{ primitive recursive)}$$

or equivalently

$$\vdash \neg\neg\exists x(t(x) = 0) \rightarrow \exists x(t(x) = 0)$$

for an arbitrary term  $t$ . This principle is accepted in CRM, but not in BCM and INT. *Markov's rule* is the rule corresponding to this implication:

$$MR \quad \text{If } \vdash \neg\neg\exists x(t(x) = 0) \text{ then } \vdash \exists x(t(x) = 0)$$

and it may readily be seen that closure under this rule implies the conservativity result.

## 4.1 Realizability

Realizability by numbers was originally devised by Kleene as a semantics for intuitionistic arithmetic, by defining for arithmetical sentences  $A$  a notion “the number  $\mathbf{n}$  realizes  $A$ ”, intended to capture some essential aspects of the intuitionistic meaning of  $A$ . Here  $\mathbf{n}$  is not a term of the arithmetical formalism, but an element of the natural numbers  $\mathbb{N}$ . The definition is by induction on the complexity of  $A$ :

1.  $\mathbf{n}$  realizes  $t = s$  iff  $t = s$  holds;
2.  $\mathbf{n}$  realizes  $A \wedge B$  iff  $\mathbf{p}_0\mathbf{n}$  realizes  $A$  and  $\mathbf{p}_1\mathbf{n}$  realizes  $B$ ;
3.  $\mathbf{n}$  realizes  $A \vee B$  iff  $\mathbf{p}_0\mathbf{n} = 0$  and  $\mathbf{p}_1\mathbf{n}$  realizes  $A$  or  $\mathbf{p}_0\mathbf{n} > 0$  and  $\mathbf{p}_1\mathbf{n}$  realizes  $B$ ;
4.  $\mathbf{n}$  realizes  $A \rightarrow B$  iff for all  $\mathbf{m}$  realizing  $A$ ,  $\mathbf{n}\bullet\mathbf{m}$  is defined and realizes  $B$ ;
5.  $\mathbf{n}$  realizes  $\neg A$  if for no  $\mathbf{m}$ ,  $\mathbf{m}$  realizes  $A$ ;
6.  $\mathbf{n}$  realizes  $\exists y A$  iff  $\mathbf{p}_1\mathbf{n}$  realizes  $A[y/\overline{\mathbf{p}_0\mathbf{n}}]$ .
7.  $\mathbf{n}$  realizes  $\forall y A$  iff  $\mathbf{n}\bullet\mathbf{m}$  is defined and realizes  $A[y/\overline{\mathbf{m}}]$ , for all  $\mathbf{m}$ .

Here  $\mathbf{p}_1$  and  $\mathbf{p}_0$  are the inverses of some standard primitive recursive pairing function  $\mathbf{p}$  coding  $\mathbb{N}^2$  onto  $\mathbb{N}$ , and  $\overline{\mathbf{m}}$  is the standard term  $S^{\mathbf{m}}0$  (numeral) in the language of intuitionistic arithmetic corresponding to  $\mathbf{m}$ ;  $\bullet$  is partial recursive function application, i.e.,  $\mathbf{n}\bullet\mathbf{m}$  is the result of applying the function with code  $\mathbf{n}$  to  $\mathbf{m}$ . (Later on we also use  $\bar{m}, \bar{n}, \dots$  for numerals.)

The definition may be extended to formulas with free variables by stipulating that  $\mathbf{n}$  realizes  $A$  if  $\mathbf{n}$  realizes the universal closure of  $A$ .

Reading “there is a number realizing  $A$ ” as “ $A$  is constructively true”, we see that a realizing number provides witnesses for the constructive truth of existential quantifiers and disjunctions, and in implications carries this type of information from premise to conclusion by means of partial recursive operators. In short, realizing numbers “hereditarily” encode information about the realization of existential quantifiers and disjunctions.

Realizability, as an interpretation of “constructively true” may be regarded as a concrete implementation of the Brouwer-Heyting-Kolmogorov explanation (BHK for short) of the intuitionistic meaning of the logical connectives. Realizability corresponds to BHK if (a) we concentrate on (numerical) information concerning the realizations of existential quantifiers and the choices for disjunctions, and (b) the constructions considered for  $\forall, \rightarrow$  are encoded by (partial) recursive operations.

Realizability gives a classically meaningful definition of intuitionistic truth; the set of realizable statements is closed under deduction and must be consistent, since  $1=0$  cannot be realizable. It is to be noted that some principles are realizable which are classically false, for example

$$\neg\forall x[\exists yTxy \vee \forall y\neg Txy]$$

is easily seen to be realizable. ( $T$  is Kleene’s T-predicate, which is assumed to be available in our language;  $Txyz$  is primitive recursive in  $x, y, z$  and expresses that the algorithm with code  $x$  applied to argument  $y$  yields a computation with code  $z$ ;  $U$  is a primitive recursive function extracting from a computation code  $z$  the result  $Uz$ .) For  $\neg A$  is realizable iff no number realizes  $A$ , and realizability of  $\forall x[\exists yTxy \vee \forall y\neg Txy]$  requires a total recursive function deciding  $\exists yTxy$ , which does not exist (more about this below). In this way realizability shows how in constructive mathematics principles may be incorporated which cause it to diverge from the corresponding classical theory, instead of just being included in the classical theory.

In order to exploit realizability proof-theoretically, we have to formalize it. Let us first discuss its formalization in ordinary intuitionistic first-order arithmetic **HA**.

$x, y, z, \dots$  are numerical variables,  $S$  is successor. We use the notation  $\bar{n}$  for the term  $S^n 0$ ; such terms are called *numerals*.  $\mathbf{p}_0, \mathbf{p}_1$  bind stronger than infix binary operations, i.e.,  $\mathbf{p}_0 t + s$  is  $(\mathbf{p}_0 t) + s$ . For primitive recursive predicates  $R$ ,  $Rt_1 \dots t_n$  may be treated as a prime formula since the formalism contains a symbol for the characteristic function  $\chi_R$ .

Now we are ready for a formalized definition of “ $x$  realizes  $A$ ” in **HA**.

**Definition 4** *By recursion on the complexity of  $A$  we define  $x \underline{\mathbf{rn}} A$ ,  $x \notin \text{FV}(A)$ , “ $x$  numerically realizes  $A$ ” :*

$$\begin{aligned} x \underline{\mathbf{rn}} (t = s) &::= (t = s) \\ x \underline{\mathbf{rn}} (A \wedge B) &::= (\mathbf{p}_0 x \underline{\mathbf{rn}} A) \wedge (\mathbf{p}_1 x \underline{\mathbf{rn}} B), \\ x \underline{\mathbf{rn}} (A \vee B) &::= (\mathbf{p}_0 x = 0 \wedge \mathbf{p}_1 x \underline{\mathbf{rn}} A) \vee (\mathbf{p}_0 x \neq 0 \wedge \mathbf{p}_1 x \underline{\mathbf{rn}} B), \\ x \underline{\mathbf{rn}} (A \rightarrow B) &::= \forall y (y \underline{\mathbf{rn}} A \rightarrow \exists z (Txyz \wedge Uz \underline{\mathbf{rn}} B)), \\ x \underline{\mathbf{rn}} \forall y A &::= \forall y \exists z (Txyz \wedge Uz \underline{\mathbf{rn}} A), \\ x \underline{\mathbf{rn}} \exists y A &::= \mathbf{p}_1 x \underline{\mathbf{rn}} A[y/\mathbf{p}_0 x]. \end{aligned}$$

Note that  $\text{FV}(x \underline{\mathbf{rn}} A) \subset \{x\} \cup \text{FV}(A)$ .

**Remarks.** (i) We have omitted a clause for negation, since in arithmetic we can take  $\neg A ::= A \rightarrow 1 = 0$ . We can in fact also omit a clause for disjunction, using the fact that in arithmetic we may define disjunction by  $A \vee B ::= \exists x ((x = 0 \rightarrow A) \wedge (x \neq 0 \rightarrow B))$ . The resulting definition of realizability is not identical, but equivalent in the following sense: a variant,  $\underline{\mathbf{rn}}'$ -realizability say, is *equivalent* to  $\underline{\mathbf{rn}}$ -realizability if for each formula  $A$  there are two partial recursive functions  $\phi_A$  and  $\psi_A$  such that  $\vdash x \underline{\mathbf{rn}} A \rightarrow \phi_A(x) \underline{\mathbf{rn}}' A$  and  $\vdash x \underline{\mathbf{rn}}' A \rightarrow \psi_A(x) \underline{\mathbf{rn}} A$ . Another variant is obtained, for example, by changing the clause for prime formulas to  $x \underline{\mathbf{rn}}'(t = s) ::= (x = t \wedge t = s)$ .

In terms of partial recursive function application  $\bullet$  and the definedness predicate  $\downarrow$  ( $t \downarrow$  means “ $t$  is defined”), we can write more succinctly:

$$\begin{aligned} x \underline{\mathbf{rn}} (A \rightarrow B) &::= \forall y (y \underline{\mathbf{rn}} A \rightarrow x \bullet y \downarrow \wedge x \bullet y \underline{\mathbf{rn}} B), \\ x \underline{\mathbf{rn}} \forall y A &::= \forall y (x \bullet y \downarrow \wedge x \bullet y \underline{\mathbf{rn}} B). \end{aligned}$$

Of course, the partial operation  $\bullet$  and the definedness predicate  $\downarrow$  are not part of the language, but expressions containing them may be treated as abbreviations, using the following equivalences:

$$\begin{aligned} t_1 = t_2 &::= \exists x (t_1 = x \wedge t_2 = x), \\ t_1 \bullet t_2 = x &::= \exists y z u (t_1 = y \wedge t_2 = z \wedge T y z u \wedge U u = x), \\ t \downarrow &::= \exists z (t = z). \end{aligned}$$

(here  $t_1, t_2$  terms containing  $\bullet$ , and  $x, y, z, u$  do not occur free in  $t_1, t_2$ ).

## 4.2 Characterization of realizability

From the preceding it will be clear that more formulas are being realized than just the formulas provable in **HA**. The aim of this subsection is to characterize the realizable formulas of **HA**, that is to say we want to find an axiomatization for the formulas which are **HA**-provable realizable.

Let us call an arithmetical formula *almost negative*, if it is built from formulas  $t = s$ ,  $\exists x(t = s)$  with the help of  $\forall, \wedge, \rightarrow$  alone. Let *Extended Church's Thesis* be the schema

$$\text{ECT}_0 \quad \forall x [Ax \rightarrow \exists y B(x, y)] \rightarrow \exists z \forall x [Ax \rightarrow z \bullet x \downarrow \wedge B(x, z \bullet x)].$$

where  $A$  is almost negative, or expressed without the use of  $\bullet, \downarrow$ :

$$\forall x[Ax \rightarrow \exists yB(x, y)] \rightarrow \exists z\forall x[Ax \rightarrow \exists u(Tz xu \wedge B(x, Uu))].$$

Then the following holds

**Theorem 3** (*Troelstra, 1971*)

$$\begin{aligned} \mathbf{HA} + \text{ECT}_0 \vdash A &\leftrightarrow \exists x(x \underline{\text{rn}} A), \\ \mathbf{HA} + \text{ECT}_0 \vdash A &\text{ iff } \mathbf{HA} \vdash \exists x(x \underline{\text{rn}} A). \end{aligned}$$

The mathematical practice of CRM is more or less captured axiomatically by the formalism  $\mathbf{HA} + \text{ECT}_0 + \text{M}$ , where  $\text{M}$  is a general form of *Markov's Schema*:

$$\text{M} \quad \forall x(Ax \vee \neg Ax) \wedge \neg\neg\exists xAx \rightarrow \exists xAx,$$

which in fact, in the presence of  $\text{ECT}_0$  is equivalent to Markov's schema  $\text{M}_{\text{PR}}$  we encountered before. In the presence of  $\text{M}_{\text{PR}}$  we can replace  $\text{ECT}_0$  by the schema

$$\text{ECT} \quad \forall x[\neg Ax \rightarrow \exists yB(x, y)] \rightarrow \exists z\forall x[\neg Ax \rightarrow z\bullet x\downarrow \wedge B(x, z\bullet x)]$$

without further restrictions on the  $A$ .

## 5 Constructive Mathematics

Below follows a brief sketch, describing similarities and differences in mathematical practice (more specifically, in analysis) between INT, BCM and CRM. There is also a short digression on the use of lawless sequences in obtaining completeness for I-validity (cf. 5.4).

### 5.1 Bishop's constructive mathematics BCM

Natural numbers are regarded as unproblematic, and most of elementary number theory can be treated on the axiomatic basis of  $\mathbf{HA}$ . Similarly, the rationals ( $\mathbb{Q}$ ), and the integers ( $\mathbb{Z}$ ) which can be encoded into  $\mathbb{N}$  in a straightforward way are unproblematic.

The obvious requirement for a constructive real is that we must be able to approximate it by rational numbers with any required degree of accuracy. This can be achieved by interpreting the standard method of introducing the real numbers via fundamental sequences (Cauchy-sequences) of rationals in a constructive way: and indeed, this is the usual method in BCM, in CRM as well as in INT.

A sequence  $\langle r_n \rangle$  is a fundamental sequence (f.s.) with modulus  $\alpha$ , if

$$\forall kmm'(|r_{\alpha k+m} - r_{\alpha k+m'}| < 2^{-k},$$

an alternative definition uses a fixed modulus, e.g.,

$$\forall nm(|r_n - r_m| < n^{-1} + m^{-1}).$$

Two fundamental sequences  $\langle r_n \rangle, \langle s_n \rangle$  are equivalent, notation  $\langle r_n \rangle \sim \langle s_n \rangle$  if

$$\forall k\exists n\forall m(|r_{n+m} - s_{n+m}| < 2^{-k}).$$

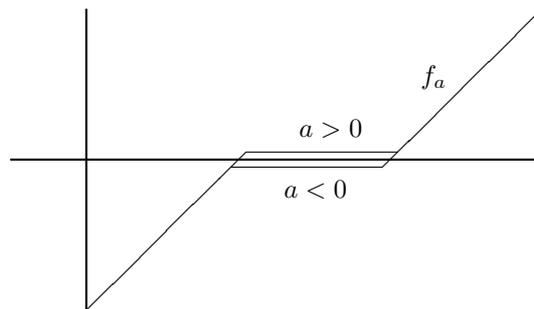
The equivalence classes of fundamental sequences modulo  $\sim$  are the Cauchy reals.

A function from the reals to the reals must enable us to compute us an arbitrarily close approximation of the value from a sufficiently close approximation of the argument, etc.

In analysis we soon meet with examples of theorems which need revision in a constructive setting. (In many such cases, the classical proof relies at one point on the Bolzano-Weierstrass theorem: an infinite set  $X$  in a bounded interval has a condensation point, i.e., a point  $y$  such that every neighborhood of  $y$  contains infinitely many points of  $X$ .) One of the first examples of this kind one encounters is the *intermediate value theorem*.

In its simplest form, the theorem is as follows: let  $f$  be a continuous function from the closed interval  $[b, c]$  to  $\mathbb{R}$ , with  $f(b) < 0$ ,  $f(c) > 0$ , then there is a real number  $x$ ,  $b < x < c$ , such that  $f(x) = 0$ . Constructively, this theorem does not hold. To see this, consider the function  $f_a$ , depending on a parameter  $a \in \mathbb{R}$ , defined as

$$f_a(x) = \begin{cases} x - 2 & \text{if } x - 2 \leq a, \\ a & \text{if } x - 4 \leq a \leq x - 2, \\ x - 4 & \text{if } a \leq x - 4. \end{cases}$$



As long as we do not know whether  $a \leq 0$ , or  $a \geq 0$ , we are unable to determine a zero of  $f_a$ . An arbitrarily small change of  $a < 0$  to a value  $> 0$  changes the zero from 4 to 2, hence we cannot determine the zero with any desired degree of accuracy.

For a constructive variant of this theorem we can follow two distinct strategies: weakening the conclusion, or strengthening the hypotheses. The following variant is constructively correct and obtained by weakening the conclusion:

Let  $f$  be a function from the closed interval  $[b, c]$  to  $\mathbb{R}$ , with  $f(b) < 0$ ,  $f(c) > 0$ , and let  $\varepsilon > 0$ , then there is a real number  $x$ ,  $b < x < c$ , such that  $|f(x)| < \varepsilon$ .

Alternatively, we may strengthen the hypotheses of the theorem as follows: Let  $f$  be a function from the closed interval  $[b, c]$  to  $\mathbb{R}$ , with  $f(b) < 0$ ,  $f(c) > 0$ , such that for all  $d, e$  with  $a < d < e < b$  there is a  $y \in [d, e]$  such that  $f(y) \# 0$ . (We write  $t \# s$  iff  $|t - s| > (n + 1)^{-1}$  for a suitable natural number  $n$ . I.e.,  $\#$  is a positive form of inequality.) Then there is a  $y \in [a, b]$  such that  $f(y) = 0$ .

This version of the theorem looks a bit clumsy, but it is precisely what we need to prove constructively that the reals are algebraically closed, that is to say, any polynomial of odd degree over the reals has a real root.

BCM has been developed at length: complex analysis, integration and measure theory, normed spaces, locally compact abelian groups, parts of algebra with an algorithmic character. All results of BCM also hold in CRM and INT, since in BCM no principles are accepted which do not also hold in CRM and INT.

## 5.2 Constructive recursive mathematics CRM

In CRM the fundamental sequences used to define real numbers are *recursive* modulo a standard enumeration of the rationals. The resulting mathematics deviates from classical mathematics. On the negative side, there are a number of results which in a strong sense falsify statements from classical mathematics:

1. There is a bounded monotone increasing sequence of rationals  $\langle r_n \rangle$  which positively differs from each real:

$$\forall x \exists km \forall n (|r_{m+n} - x| > 2^{-k}).$$

(this is called a *strong Specker sequence*; a *Specker sequence* is a bounded monotone sequence of rationals which cannot have a real as its limit.)

2. There exists a *singular cover* of open intervals  $\langle I_n \rangle_n$  of  $[0, 1]$ , that is to say

$$\forall x \in [0, 1] \exists n (x \in I_n), \quad \exists \varepsilon > 0 \forall n \sum_{i=0}^n |I_i| < \varepsilon.$$

3. There exists a continuous real-valued function on  $[0, 1]$  which is unbounded, hence not uniformly continuous.
4. There exists a uniformly continuous function on  $[0, 1]$  which is everywhere positive, but the infimum on  $[0, 1]$  is zero.
5. There is a function, uniformly continuous on  $[0, 1]$ , which cannot be integrated over  $[0, 1]$  in any reasonable sense.

At first sight it seems as if results such as (2) prevent a decent measure theory. However, it turns out that whenever a cover  $\langle I_n \rangle_n$  is singular, then the sequence  $\langle r_n \rangle_n$  given by  $r_n := |I_0 \cup I_1 \cup \dots \cup I_n|$  is a Specker sequence; and if we insist on a limit for  $\langle r_n \rangle_n$ , singular covers are impossible. The results (iii)–(iv) are easily derived from (ii).

On the other hand, there is also a beautiful positive result, the KLST-theorem, or the *Kreisel-Lacombe-Shoenfield-Tsejtin* theorem:

**Theorem 4** (Kreisel-Lacombe-Shoenfield-Tsejtin, 1957, 1959) *Every total function from a complete, separable, metric space into a separable metric space is continuous.*

The proof of this uses  $\text{ECT}_0 + \text{M}$ , that is, Extended Church's Thesis and Markov's principle. This theorem cannot be proved in  $\mathbf{HA} + \text{ECT}_0$ .

### 5.3 Intuitionism INT

In intuitionism, not only classical logic is rejected, which leads to the rejection of certain theorems of classical mathematics as invalid, but on the other hand objects are admitted which in a positive sense make intuitionistic mathematics deviate from classical mathematics, namely Brouwer's notion of choice sequence.

Choice sequences (of natural numbers) are generated (by the idealized mathematician) by successively choosing elements of the sequence, while possibly stipulating restrictions on future choices at each stage of the process. So there is no commitment in advance to a law or recipe for determining the values; the only commitment which must hold is that more and more elements will be determined. Choice sequences are unfinished objects; nevertheless we can make meaningful statements about the properties of the universe of all choice sequences, we can define operations which can be applied to arbitrary choice sequences, namely continuous operations as term-by-term addition of two choice sequences, etc.

For this informally described notion one may argue the plausibility of certain axiomatic principles; some of these principles do not hold for the classical universe of all number sequences. before continuing, some

**Notation** Until further notice,  $\alpha, \beta, \gamma$  range over (choice) sequences. We write  $\bar{\alpha}x$  for the (code of) the finite sequence  $\langle \alpha 0, \alpha 1, \dots, \alpha(x-1) \rangle$ .  $\beta \in \bar{\alpha}m$  means the same as  $\bar{\beta}m = \bar{\alpha}m$ . For (codes of) finite sequence  $n, m$  we write  $n \subseteq m$  if  $n$  is an initial segment of  $m$ , and  $n \subset m$  if  $n$  is a proper initial segment of  $m$ . We use  $n * m$  for the (code of) the concatenation of  $n$  with  $m$ .

Among the crucial principles which may be argued to hold for this notion, is the (*weak*) *continuity principle*: an operator  $\Gamma$  associating a natural number to every choice sequence  $\alpha$ , is necessarily continuous, that is to say  $\Gamma\alpha$  is determined on the basis of a finite initial segment of  $\alpha$ :  $\forall \alpha \exists n \forall \beta \in \bar{\alpha}n (\Gamma\alpha = \Gamma\beta)$ . If we combine this with the insight

$$\forall \alpha \exists n A(\alpha, n) \rightarrow \exists \Gamma \forall \alpha A(\alpha, \Gamma\alpha)$$

where  $\Gamma$  is a functional from functions to numbers, we obtain, without explicit reference to functionals from functions to numbers, the principle of *weak continuity for numbers*.

$$\text{WC-N} \quad \forall \alpha \exists n A(\alpha, n) \rightarrow \forall \alpha \exists mn \forall \beta \in \bar{\alpha}m A(\beta, n).$$

The motivation for this principle is, that if  $\Gamma$  is to compute a number from a choice sequence  $\alpha$ , there must be a stage  $n$  for which the available information suffices for computing  $\Gamma\alpha$ . Since the only type of information which at any given stage is guaranteed to exist for every choice sequence  $\alpha$ , is a finite initial segment of the sequence of values, say  $\bar{\alpha}n := \langle \alpha 0, \dots, \alpha(n-1) \rangle$ , it is assumed that  $\Gamma$  can always be computed on the basis of a finite initial segment of values.

The principle is easily seen to conflict with classical logic, for example we can *refute* the following universally quantified form of the *Principle of the Excluded Middle* PEM:

$$\forall \alpha (\forall x (\alpha x = 0) \vee \neg \forall x (\alpha x = 0)).$$

A stronger expression of the same insight is obtained by using the notion of a neighborhood function:

$$\gamma \in K_0 := \forall nm (\gamma n > 0 \rightarrow \gamma n = \gamma(n * m)) \wedge \forall \alpha \exists x (\gamma(\bar{\alpha}x) > 0).$$

The idea is that  $\gamma$  encodes a functional  $\Gamma: \mathbb{N}^{\mathbb{N}} \rightarrow \mathbb{N}$ ; if  $n$  is too short to compute  $\Gamma\alpha$  for  $\alpha \in n$ , we put  $\gamma n = 0$ ; and if  $n$  suffices to compute  $\Gamma\alpha$  for  $\alpha \in n$ , we put  $\gamma n = \Gamma\alpha + 1$ . Now we can express the continuity principle in a stronger form:

$$\text{C-N} \quad \forall \alpha \exists n A(\alpha, n) \rightarrow \exists \gamma \in K_0 \forall m (\gamma m \neq 0 \rightarrow \forall \alpha \in m A(\alpha, \gamma m - 1)).$$

With the notation  $\gamma(\alpha) = x := \exists n (\gamma(\bar{\alpha}n) = x + 1)$  we can express this in a more compact way as  $\forall \alpha \exists n A(\alpha, n) \rightarrow \exists \gamma \in K_0 A(\alpha, \gamma(\alpha))$ . Another principle defended by Brouwer for the notion of choice sequences, which is also classically valid, is (in a formulation due to S.C. Kleene) the principle of (*decidable*) *Bar Induction*:

$$\text{BI}_D \quad \forall \alpha \exists x P(\bar{\alpha}x) \wedge \forall n (Pn \vee \neg Pn) \wedge \forall n (Pn \rightarrow Qn) \wedge \forall n (\forall y Q(n * \langle y \rangle) \rightarrow Qn) \rightarrow Q\langle \rangle.$$

or the stronger *monotone bar induction*

$$\text{BI}_M \quad \forall \alpha \exists x Q(\bar{\alpha}x) \wedge \forall n (\forall y Q(n * \langle y \rangle) \rightarrow Qn) \rightarrow Q\langle \rangle.$$

In the presence of C-N the versions  $\text{BI}_D$  and  $\text{BI}_M$  are equivalent. An important consequence of  $\text{BI}_M$  is the so-called *Fan Theorem*. Let  $T$  be a *fan*, that is to say a decidable finitely branching tree, such that all nodes have at least one successor:

$$\begin{aligned} & \forall n (n \in T \text{ or } n \notin T), \\ & \langle \rangle \in T, \\ & n \in T \wedge m \subset n \rightarrow m \in T, \\ & \forall n \in T \exists x (n * \langle x \rangle \in T), \\ & \forall n \in T \exists y \forall x (n * \langle x \rangle \in T \rightarrow x \leq y). \end{aligned}$$

Let  $\alpha_T$  range over infinite branches of  $T$ . Then the Fan Theorem may be stated as:

$$\text{FAN} \quad \forall \alpha_T \exists x A(\bar{\alpha}x) \rightarrow \exists z \forall \alpha_T \exists y \leq z A(\bar{\alpha}y).$$

One of the mathematical consequences of this principle is the compactness of the closed interval and of the Cantor discontinuum.

In the preceding paragraphs we sketched a simplified picture of “general” choice sequences: all restrictions on future choices are permitted at any stage of generating the sequence, provided the sequence can always be continued. However, if we do not admit arbitrary restrictions on future choices of numerical values, but only restrictions from a specific class  $\mathcal{C}$ , many distinct universes, with distinct principles, may be described.

An extreme example of this is the notion of *lawless sequence*; for a lawless sequence, no general restriction on future choices is made, so at any stage in the generation of such a sequence only an initial segment is known. It has been a point of debate whether the consideration of such subdomains of the universe of all choice sequences is meaningful; here we take the position that this is meaningful.

A characteristic axiom for lawless sequences is the so-called Axiom of Open Data:

$$A(\alpha) \rightarrow \exists n(\alpha \in n \wedge \forall \beta \in n A(\beta)).$$

This is equivalent to

$$\forall \alpha(A\alpha \rightarrow B\alpha) \leftrightarrow \forall n(\forall \alpha \in n A\alpha \rightarrow \forall \alpha \in n B\alpha).$$

The picture is still further complicated by the consideration of Brouwer’s controversial empirical sequences, sequences constructed by explicit reference to the activity of an idealized mathematician, which is supposed to be carried out in discrete stages  $0, 1, 2, \dots$  these considerations lead to Kripke’s schema, for arbitrary meaningful statements  $A$ :

$$\text{KS} \quad \exists \alpha[A \leftrightarrow \exists n(\alpha n \neq 0)].$$

It is debatable whether a conceptually coherent notion of choice sequences exists, which embraces (1) empirical sequences, (2) choice sequences as outlined above and in addition satisfying  $\text{BI}_M$ ,  $\text{KS}$ ,  $\text{C-N}$ .

## 5.4 Lawless sequences and I-validity

The lawless sequences enable us to connect I-validity with validity in a Beth model, or, what amounts to the same, a topological model over the Cantor Discontinuum  $\mathcal{C}$ .

Let  $V_n := \{\alpha : \alpha \in n\}$ , where  $n$  encodes a finite 01-sequence. Consider a topological model  $\langle \mathcal{C}, \mathbb{N}, \llbracket \cdot \rrbracket \rangle$ . With each  $n$ -ary relation  $R$  we can now associate a  $n$ -ary relation  $R_\alpha^*$  depending on a lawless parameter  $\alpha$ , as follows:

$$R_\alpha^*(n_1, \dots, n_k) := \exists n(\alpha \in n \wedge V_n \subset \llbracket R(n_1, \dots, n_k) \rrbracket).$$

The axioms of lawless sequences then permit us to show for all sentences  $A$ :

$$A_\alpha^* \leftrightarrow \exists n(\alpha \in n \wedge V_n \subset \llbracket A \rrbracket),$$

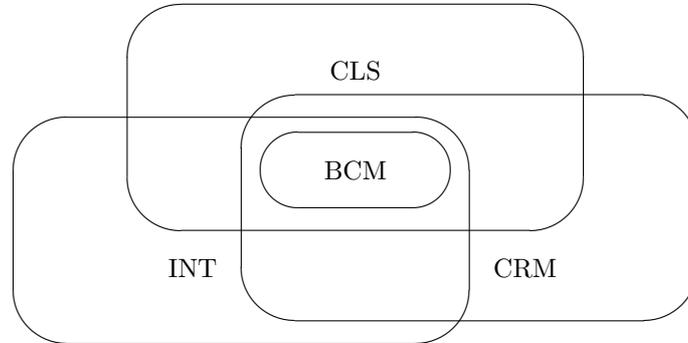
where  $A_\alpha^*$  is obtained from  $A$  by replacing occurrences of prime formulas  $P$  by  $P_\alpha^*$ . Conversely, if we start from the predicates  $R_\alpha^*$ , we can obtain a topological model putting

$$\alpha \in \llbracket R(n_1, \dots, n_k) \rrbracket := R_\alpha^*(n_1, \dots, n_k).$$

So  $\forall \alpha \in \mathcal{C} A_\alpha^* \leftrightarrow \llbracket A \rrbracket = T$ . These constructions are inverse to each other. For propositional logic we can give an intuitionistically acceptable completeness proof relative to topological models over  $\mathcal{C}$ , and also for predicate logic without  $\perp$ . Finally, this idea still works for all of predicate logic, provided we permit a non-standard interpretation of  $\perp$  as an arbitrary proposition such that  $\llbracket \perp \rrbracket \subset \llbracket P \rrbracket$  for all atomic sentences  $P$ . The correspondence outlined above then gives us completeness for I-validity.

## 5.5 Comparison of BCM, CRM and INT

The formal relations between intuitionistic mathematics INT, classical mathematics CLS, Bishop's constructive mathematics BCM and constructive recursive mathematics CRM in the area of mathematical analysis may be pictured as follows.



Examples showing the non-emptiness of the various areas shown in the picture are

- (i)  $(\text{INT} \cap \text{CRM}) \setminus \text{CLS}$ : functions on  $[0, 1]$  are continuous;
- (ii)  $(\text{CRM} \cap \text{CLS}) \setminus \text{INT}$ : for reals  $x, y$ , if  $x \neq y$ , then  $x \# y$ ;
- (iii)  $(\text{INT} \cap \text{CLS}) \setminus \text{CRM}$ : continuous functions on  $[0, 1]$  are uniformly continuous, a consequence of the compactness of  $[0, 1]$ ;
- (iv)  $\text{INT} \setminus (\text{CRM} \cup \text{CLS})$ : all functions on  $[0, 1]$  are uniformly continuous (a combination of (i) and (iii));
- (v)  $\text{CLS} \setminus (\text{CRM} \cup \text{INT})$ : decidability of equality between reals;
- (vi)  $\text{CRM} \setminus (\text{INT} \cup \text{CLS})$ : the existence of strong Specker sequences, entailing the non-compactness of  $[0, 1]$  and the existence of continuous, but not uniformly continuous, unbounded functions on  $[0, 1]$

A diagram as shown is possible, because we can express the mathematical results of BCM, CRM and INT in a common language. However, this picture ought to be interpreted with caution, because it indicates a formal relationship only. The real numbers have quite different interpretations in CLS, INT and CRM, and all the picture does is to say how the results differ if we give the mathematical objects their respective interpretations suited to each version of constructivism.

## 6 Proof Theory of first-order logic

Proof Theory made a big step forward with the work of Gerhard Gentzen. He devised a type of logical calculus which is called a *Gentzen system*, or a *sequent calculus*. Gentzen treated only classical and intuitionistic logic, but nowadays Gentzen systems have been described and used for a whole range of other logics, such as various systems of modal logic.

Sequent calculi deal with sequents, expressions of the form  $\Gamma \Rightarrow \Delta$ , where the  $\Gamma, \Delta$  are finite sets, sequences or multisets, depending on the precise version of Gentzen system one is discussing. (A multiset is a set with multiplicity, in other words, a sequence modulo the order of the formulas.) Intuitively, a sequent  $A_1, A_2, \dots, A_n \Rightarrow B_1, B_2, \dots, B_m$  means " $B_1 \vee B_2 \vee \dots \vee B_m$  follows from  $A_1 \wedge A_2 \wedge \dots \wedge A_n$ ".

In a sequent  $\Gamma \Rightarrow \Delta$ ,  $\Gamma$  is called the *antecedent*, and  $\Delta$  the *succedent* of the sequent. We present here first versions of such Gentzen systems which are quite close to the systems as originally described by Gentzen for classical and intuitionistic logic.

## 6.1 The Gentzen systems $\mathbf{Gc}$ and $\mathbf{Gi}$

The antecedents and succedents are multisets. Proofs or deductions are labelled finite trees with a single root, with axioms at the top nodes, and each node-label connected with the labels of the (immediate) successor nodes (if any) according to one of the rules. The rules are divided into *left-* ( $L$ -) and *right-* ( $R$ -) *rules*. For a logical operator  $\otimes$  say,  $L\otimes$ ,  $R\otimes$  indicate the rules where a formula with  $\otimes$  as main operator is introduced on the left and on the right respectively. The axioms and rules for  $\mathbf{Gc}$  are:

*Axioms*

$$\text{Ax } A \Rightarrow A$$

$$\text{L}\perp \perp \Rightarrow$$

*Rules for weakening (W) and contraction (C)*

$$\text{LW } \frac{\Gamma \Rightarrow \Delta}{A, \Gamma \Rightarrow \Delta}$$

$$\text{RW } \frac{\Gamma \Rightarrow \Delta}{\Gamma \Rightarrow \Delta, A}$$

$$\text{LC } \frac{A, A, \Gamma \Rightarrow \Delta}{A, \Gamma \Rightarrow \Delta}$$

$$\text{RC } \frac{\Gamma \Rightarrow \Delta, A, A}{\Gamma \Rightarrow \Delta, A}$$

*Rules for the logical operators*

$$\text{L}\wedge \frac{A_i, \Gamma \Rightarrow \Delta}{A_0 \wedge A_1, \Gamma \Rightarrow \Delta} \quad (i = 0, 1)$$

$$\text{R}\wedge \frac{\Gamma \Rightarrow \Delta, A \quad \Gamma \Rightarrow \Delta, B}{\Gamma \Rightarrow \Delta, A \wedge B}$$

$$\text{L}\vee \frac{A, \Gamma \Rightarrow \Delta \quad B, \Gamma \Rightarrow \Delta}{A \vee B, \Gamma \Rightarrow \Delta}$$

$$\text{R}\vee \frac{\Gamma \Rightarrow \Delta, A_i}{\Gamma \Rightarrow \Delta, A_0 \vee A_1} \quad (i = 0, 1)$$

$$\text{L}\rightarrow \frac{\Gamma \Rightarrow \Delta, A \quad B, \Gamma' \Rightarrow \Delta'}{A \rightarrow B, \Gamma, \Gamma' \Rightarrow \Delta, \Delta'}$$

$$\text{R}\rightarrow \frac{A, \Gamma \Rightarrow \Delta, B}{\Gamma \Rightarrow \Delta, A \rightarrow B}$$

$$\text{L}\forall \frac{A[x/t], \Gamma \Rightarrow \Delta}{\forall x A, \Gamma \Rightarrow \Delta}$$

$$\text{R}\forall \frac{\Gamma \Rightarrow \Delta, A[x/y]}{\Gamma \Rightarrow \Delta, \forall x A}$$

$$\text{L}\exists \frac{A[x/y], \Gamma \Rightarrow \Delta}{\exists x A, \Gamma \Rightarrow \Delta}$$

$$\text{R}\exists \frac{\Gamma \Rightarrow \Delta, A[x/t]}{\Gamma \Rightarrow \Delta, \exists x A}$$

where in  $L\exists$ ,  $R\forall$ ,  $y$  is not free in the conclusion.

The variable  $y$  in an application  $\alpha$  of  $R\forall$  or  $L\exists$  is called the *proper* variable of  $\alpha$ . The proper variable of  $\alpha$  occurs only above  $\alpha$ .

In the rules the  $\Gamma, \Delta$  are called the *side formulas* or the *context*. In the conclusion of each rule, the formula not in the context is called the *principal* or *main* formula. The formula(s) in the premise(s) from which the principal formula derives (i.e., the formulas not belonging to the context) are the *active* formulas. In the axiom Ax, both occurrences of  $A$  are principal; in  $L\perp$  the occurrence of  $\perp$  is principal.

The intuitionistic system  $\mathbf{Gi}$  is the subsystem of  $\mathbf{Gc}$  obtained by restricting all axioms and rules to sequents with *at most one* formula on the right.

To these systems one may add the *Cut rule*; an application of Cut has the form:

$$\text{Cut } \frac{\Gamma \Rightarrow \Delta, A \quad A, \Gamma' \Rightarrow \Delta'}{\Gamma, \Gamma' \Rightarrow \Delta, \Delta'}$$

$A$  is the cutformula of this application. Gentzen showed

**Theorem 5** *The Cut rule may be eliminated, that is, for every proof in the sequent calculus with Cut, there may be found by a systematic elimination procedure a proof not using Cut.*

Proofs without Cut have the subformula property: all formulas occurring in the deduction tree are subformulas of the formulas occurring in the conclusion. Using Cut permits short proofs, but these proofs may be “indirect”: a formula  $A$  may be cut out by an application of Cut which does not occur in the conclusion. It is not difficult to prove completeness of the system  $\mathbf{Gc} + \text{Cut}$ , for example by proving it to be equivalent to other formalizations of classical predicate logic. The (completeness of the) cutfree calculus provides us with a convenient route to several important results on predicate logic, for example the *interpolation theorem*.

Kleene was the first to devise a version of the Gentzen calculi without structural rules; the structural rules were “absorbed” in a manner of speaking in the logical rules. Weakening was absorbed by generalizing the axioms, and contraction by including the principal formula also in the premises. Thus for example, the classical Gentzen-Kleene calculus  $\mathbf{GKc}$  has the following axioms and rules for implication:

$$\begin{array}{ll} \text{Ax } P, \Gamma \Rightarrow \Delta, P \text{ (} P \text{ atomic)} & \text{L}\perp \perp, \Gamma \Rightarrow \Delta \\ \text{L}\rightarrow \frac{\Gamma, A \rightarrow B \Rightarrow \Delta, A \quad B, A \rightarrow B, \Gamma \Rightarrow \Delta}{A \rightarrow B, \Gamma \Rightarrow \Delta} & \text{R}\rightarrow \frac{A, \Gamma \Rightarrow \Delta, B, A \rightarrow B}{\Gamma \Rightarrow \Delta, A \rightarrow B} \end{array}$$

For the intuitionistic version  $\mathbf{GKi}$  these take the form:

$$\begin{array}{ll} \text{Ax } P, \Gamma \Rightarrow P \text{ (} P \text{ atomic)} & \text{L}\perp \perp, \Gamma \Rightarrow A \\ \text{L}\rightarrow \frac{A \rightarrow B, \Gamma \Rightarrow A \quad B, A \rightarrow B, \Gamma \Rightarrow C}{A \rightarrow B, \Gamma \Rightarrow C} & \text{R}\rightarrow \frac{A, \Gamma \Rightarrow B}{\Gamma \Rightarrow A \rightarrow B} \end{array}$$

Note that in going from the conclusions to one of the premises, in the classical system no formula ever disappears; in the intuitionistic system, this only holds for the left hand side. The absence of the rules of weakening and contraction in the Gentzen-Kleene systems is often very convenient.

## 6.2 Cut elimination

Let us write  $\vdash_n \Gamma \Rightarrow \Delta$  for “the sequent  $\Gamma \Rightarrow \Delta$  has a deduction of depth at most  $n$ ”. The following three properties are easily provable for the system  $\mathbf{G3c}$ .

1. if  $\vdash_n \Gamma \Rightarrow \Delta$ , then  $\vdash_n \Gamma' \Gamma \Rightarrow \Delta \Delta'$  (*closure under weakening*);
2. if  $\vdash_n \Gamma, A, A \rightarrow \Delta$ , then  $\vdash_n \Gamma, A \Rightarrow \Delta$  and if  $\vdash_n \Gamma \rightarrow A, A, \Delta$  then  $\vdash_n \Gamma \Rightarrow A, \Delta$  (*closure under contraction*);
3. if  $\vdash_n \Gamma \Rightarrow A \rightarrow B, \Delta$  then  $\vdash_n \Gamma, A \Rightarrow B, \Delta$  (*inversion lemma for implication on the right*).

There are inversion properties for the other logical operators as well, e.g., if  $\vdash_n \Gamma, A \wedge B \Rightarrow \Delta$  then  $\vdash_n \Gamma, A, B \rightarrow \Delta$ . However, below only (iii) is used. Closure under contraction and weakening on the left, and inversion for implication on the right also hold for  $\mathbf{GKi}$ .

(i)–(iii) are proved by induction on the depth of proofs. For (i) and (ii) this is almost trivial. For (iii) we distinguish three cases. Case (a): if  $\mathcal{D}$  is an axiom, the result is easy. So let  $\mathcal{D}$  be a deduction of depth at most  $n + 1$  of  $\Gamma \Rightarrow A \rightarrow B, \Delta$ . Case (b):  $A \rightarrow B$  is not principal in the conclusion of  $\mathcal{D}$ . Then we apply the induction hypothesis to the deduction(s) of the premise(s) of the conclusion of  $\mathcal{D}$ , and then apply the last rule used in  $\mathcal{D}$  to the result(s). Case (c): if  $A \rightarrow B$  is principal in the conclusion of  $\mathcal{D}$ , then the premise of the conclusion is  $\Gamma, A \Rightarrow B, A \rightarrow B, \Delta$ . We apply the induction hypothesis

and find  $\vdash_n \Gamma, A, A \Rightarrow B, B, \Delta$ . We then apply closure under contraction twice and find  $\vdash_n \Gamma, A \Rightarrow B, \Delta$ .

*Sketch of the proof of the elimination of Cut.* We give the sketch for implication logic only. We have to show how to get rid of an application of Cut as above; let  $\mathcal{D}_0, \mathcal{D}_1$  be the cutfree deductions of the left and right premises, and  $d_0, d_1$  their respective heights. We apply a main induction on the complexity of  $D + 1$  (the *rank* of the application of Cut), with a subinduction on  $d_0 + d_1$ , the *level* of the application of Cut considered. There are three main cases:

1. Either  $\mathcal{D}_0$  or  $\mathcal{D}_1$  is an axiom.
2. The cutformula  $D$  is not principal in at least one of the premises.
3.  $D$  is principal in both premises.

Case (i): for example, let  $\mathcal{D}_0$  be an axiom with principal formula  $P$ . If  $P \equiv D$ , we have a deduction of the form  $(\Gamma \equiv \Gamma'', P)$

$$\frac{\frac{\mathcal{D}_1}{\vdash_n \Gamma', P \Rightarrow \Delta'} \quad \vdash_0 \Gamma'', P \Rightarrow P, \Delta, P}{\Gamma''\Gamma', P \Rightarrow \Delta, \Delta''}}$$

Now the conclusion can be obtained by weakening  $\mathcal{D}_1$ . Or  $P \not\equiv D$ , then the deduction has the form  $(\Delta'', P \equiv \Delta)$

$$\frac{\frac{\mathcal{D}_1}{\vdash_n \Gamma', D \Rightarrow \Delta'} \quad \vdash_0 \Gamma'', P \Rightarrow \Delta'', P, D}{\Gamma''\Gamma', P \Rightarrow \Delta, \Delta', P''}}$$

and in this case the conclusion is again an axiom.

Case (ii). In this case we permute applications of rules. Assume, for example, that  $D \equiv A \rightarrow B$  is not principal in the left premise. Then the deduction is of the form

$$\frac{\frac{\frac{\mathcal{D}_0}{\Gamma_0 \Rightarrow \Delta_0, D} \quad \frac{\mathcal{D}_1}{\Gamma_1 \Rightarrow \Delta_1, D}}{\Gamma \Rightarrow \Delta, D} \text{R} \quad \frac{\mathcal{D}_2}{D, \Gamma' \Rightarrow \Delta'}}{\Gamma, \Gamma' \Rightarrow \Delta, \Delta'}}$$

and is transformed into

$$\text{Cut} \frac{\frac{\frac{\mathcal{D}_0}{\Gamma_0 \Rightarrow \Delta_0, D} \quad \frac{\mathcal{D}'}{D, \Gamma' \Rightarrow \Delta'}}{\Gamma'\Gamma_0 \Rightarrow \Delta_0\Delta'} \quad \frac{\frac{\mathcal{D}_1}{\Gamma_1 \Rightarrow \Delta_1, D} \quad \frac{\mathcal{D}'}{D, \Gamma' \Rightarrow \Delta'}}{\Gamma'\Gamma_1 \Rightarrow \Delta_1\Delta_1'} \text{R}}{\Gamma'\Gamma \Rightarrow \Delta\Delta'} \text{Cut}$$

The two new Cuts have the same rank, but are of lower level than the original Cut; hence can be removed by the subinduction hypothesis.

Case (iii).  $D$  is principal in both premises. So if  $D \equiv A \rightarrow B$ , we have a deduction terminating in:

$$\frac{\frac{\frac{\mathcal{D}_{00}}{\Gamma, A \Rightarrow B, A \rightarrow B, \Delta} \quad \frac{\mathcal{D}_{10}}{\Gamma', A \rightarrow B \Rightarrow A, \Delta'}}{\Gamma \Rightarrow A \rightarrow B, \Delta} \quad \frac{\frac{\mathcal{D}_{11}}{\Gamma', A \rightarrow B, B \Rightarrow \Delta'}}{\Gamma', A \rightarrow B \Rightarrow \Delta'}}{\Gamma\Gamma' \Rightarrow \Delta\Delta'}}$$

becomes

$$\frac{\frac{\mathcal{D}_0^* \quad \Gamma A \Rightarrow B, \Delta}{\Gamma' \Gamma \Gamma \Rightarrow B \Delta \Delta \Delta'} \quad \frac{\mathcal{D}_0 \quad \Gamma \Rightarrow A \rightarrow B, \Delta \quad \mathcal{D}_{10} \quad \Gamma', A \rightarrow B \Rightarrow A \Delta'}{\Gamma' \Gamma \Rightarrow \Delta \Delta' A}}{\Gamma' \Gamma \Gamma \Gamma \Rightarrow \Delta \Delta \Delta \Delta' A} \quad \frac{\mathcal{D}_0 \quad \Gamma \Rightarrow A \rightarrow B, \Delta \quad \mathcal{D}_{11} \quad \Gamma', A \rightarrow B, B \Rightarrow \Delta'}{\Gamma' \Gamma B \Rightarrow \Delta \Delta'}}$$

where  $\mathcal{D}_0^*$  is the deduction obtained by application to  $\mathcal{D}_0$  of the inversion lemma followed by contraction. By repeated closure under contraction we find a deduction with cuts on  $A \rightarrow B$  of lower level, plus cuts of lower rank.

An important result for classical first-order logic is

**Theorem 6** (Herbrand, 1928) *A prenex formula, say*

$$B \equiv \forall x \exists x' \forall y \exists y' \dots A(x, x', y, y', \dots)$$

where  $A$  is quantifier-free, is derivable in classical first-order logic (no  $=$ ) iff there is a disjunction of the form

$$D \equiv \bigvee_{i=0}^n A(x_i, t_i, y_i, s_i, \dots)$$

such that  $D$  is provable propositionally, and  $B$  is derivable from  $D$ .

This theorem may be obtained as a corollary of cut elimination.

### 6.3 Natural Deduction and normalization

*Normalization* for natural deduction corresponds to cut elimination for Gentzen systems. We concentrate on the implication fragment of **Ni**. In an arbitrary deduction in this system local maxima of formula complexity can occur:

$$\frac{\frac{[A]^u \quad \mathcal{D}}{B} \quad \frac{A \rightarrow B \quad u}{A}}{B} \quad \mathcal{D}'$$

This happens when the conclusion of an introduction rule is at the same time the major premise of an elimination rule. The local maximum, the formula occurrence of  $A \rightarrow B$  may be removed by transforming the deduction above as follows:

$$\mathcal{D}' \\ [A] \\ \mathcal{D} \\ [B] \\ \mathcal{D}''$$

If we chose among the local maxima of highest complexity a topmost occurrence in the rightmost branch of the proof tree containing such local maxima, this transformation reduces the number of local maxima of highest complexity, since no new ones are introduced. Thus we can ultimately obtain a deduction without local maxima; a detailed analysis of such a *normal deduction* then shows that all formulas occurring are subformulas of either the conclusion or of one of the open assumptions. This idea can be extended to the full language of predicate logic. Summing up, we have just sketched a proof of

**Theorem 7** *Any deduction in the system of natural deduction for intuitionistic first-order logic can be transformed into a normal deduction with the same conclusion and the same open assumptions.*

There is also a normalization theorem for classical first-order logic, but this needs more care in formulating, and will not be discussed here.

The proof of the normalization theorem is quite simple, simpler than the proof of cut elimination. On the other hand, one needs a closer analysis of the structure of normal derivations before the subformula property can be established.

The normalization procedure gives added meaning to the formulas-as-types paradigm: a reduction step as just described corresponds precisely with a reduction step in the typed lambda calculus:

$$(\lambda x.t)t' \text{ reduces to } t[x/t'].$$

The subformula property is a desirable property of proofs, but there is a price to pay: the depth of the proof trees may increase hyperexponentially. That is to say, there is a sequence of non-normal proofs  $\langle \mathcal{D}_n \rangle_n$  with the depth of  $\mathcal{D}_m$  bounded by a linear function of  $m$ , such that their (unique) normal forms  $\mathcal{D}_m^n$  has a depth bounded by  $2m + 2$ , where  $2_{m+1} = 2^{2^m}$ ,  $2_0 = 1$ .

In the case of predicate logic an even stronger result holds: there is a sequence of proofs  $\langle \mathcal{D}_k \rangle_k$  with depth linear in  $k$ , conclusion  $C_k$ , such that *every normal proof* of  $C_k$ , ( $k > 1$ ) (not only the proofs obtained by normalizing  $\mathcal{D}_k$ ) has a size of at least  $2_k$ , hence depth at least  $2_{k-1}$ . Via the correspondence between systems of natural deduction and cut elimination, similar lower bounds apply to Gentzen systems.

## 6.4 The Tait-calculus

The symmetry present in classical logic permits the formulation of one-sided Gentzen systems; one may think of the sequents of such a calculus as obtained by replacing a two-sided sequent  $\Gamma \Rightarrow \Delta$  by a one-sided sequent  $\Rightarrow \neg\Gamma, \Delta$  (with intuitive interpretation the disjunction of the formulas in  $\neg\Gamma, \Delta$ ), and if we restrict attention to one-sided sequents throughout, the symbol  $\Rightarrow$  is redundant.

In order to achieve this, we need a different treatment of negation. We shall assume that formulas are constructed from *positive literals*  $P, P', P'', R(t_0, \dots, t_n), R(s_0, \dots, s_m)$  etc., as well as *negative literals*  $\neg P, \neg P', \neg P'', \neg R(t_0, \dots, t_n), \dots$  by means of  $\vee, \wedge, \forall, \exists$ . Both types of literals are treated as *primitives*.

Negation  $\neg$  satisfies  $\neg\neg P \equiv P$  for literals  $P$ , and is defined for compound formulas by De Morgan duality:

- (i)  $\neg(A \wedge B) \equiv (\neg A \vee \neg B)$ ;
- (ii)  $\neg(A \vee B) \equiv (\neg A \wedge \neg B)$ ;
- (iii)  $\neg\forall x A \equiv \exists x \neg A$ ;
- (iv)  $\neg\exists x A \equiv \forall x \neg A$ . ⊠

The one-sided calculus **GT** has the following rules and axioms:

$$\Gamma, P, \neg P \quad \frac{\Gamma, A_i}{\Gamma, A_0 \vee A_1} (i \in \{0, 1\}) \quad \frac{\Gamma, A \quad \Gamma, B}{\Gamma, A \wedge B}$$

$$\frac{\Gamma, A[x/y]}{\Gamma, \forall x A} \quad \frac{\Gamma, A[x/t]}{\Gamma, \exists x A}$$

under the obvious restrictions on  $y$  and  $t$ . The Cut rule takes the form

$$\text{Cut} \quad \frac{\Gamma, A \quad \Delta, \neg A}{\Gamma, \Delta}$$

This calculus is especially convenient in investigating classical logic. It easily generalizes to logic with infinite conjunctions and disjunctions, and a calculus of this type also plays an important role in the investigations of the proof theory of mathematical theories.

## 7 Proof Theory of mathematical theories

This is a very wide area. The oldest questions originated in Hilbert's program in combination with Gödel's second incompleteness theorem: exactly how much do we need to prove the consistency of certain axiomatic theories? Can we measure in some way the proof-theoretic strength of mathematical theories? (We call a theory  $\mathbf{T}$  stronger than a theory  $\mathbf{T}'$  if we can prove the suitably formalized consistency of  $\mathbf{T}'$  in  $\mathbf{T}$ .)

More recent topics are for example the characterization of the class of recursive functions which are provably total in a given theory, or the study of weak systems of arithmetic, i.e., systems based on weakened versions of induction.

Here the discussion will be limited to the question of proof-theoretic strength. The earliest substantial result in this area is due to Gentzen (1936, 1938). Gentzen's result for classical first-order arithmetic states, approximately, that the so-called principle of transfinite induction over an ordering is provable for wellorderings of an ordertype less than  $\varepsilon_0$ , but not up to  $\varepsilon_0$ , and that the consistency of arithmetic can be proved by transfinite induction up to  $\varepsilon_0$  for a particular primitive recursive well-ordering.

$\varepsilon_0$  is the least ordinal closed under exponentiation, i.e., the least  $\alpha$  such that  $\omega^\alpha = \alpha$ . More precisely, let us write  $\alpha'$  for the successor of  $\alpha$ , and let  $\lambda$  stand for a limit ordinal, then addition, multiplication and exponentiation may be extended from  $\mathbb{N}$  to the ordinals; they are characterized by the following equations:

$$\begin{aligned} \alpha + 0 &= \alpha, & \alpha + \beta' &= (\alpha + \beta)', & \alpha + \lambda &= \sup_{\xi < \lambda} (\alpha + \xi); \\ \alpha \cdot 0 &= 0, & \alpha \cdot \beta' &= \alpha \cdot \beta + \alpha, & \alpha \cdot \lambda &= \sup_{\xi < \lambda} (\alpha \cdot \xi); \\ \alpha^0 &= 1, & \alpha^{\beta'} &= \alpha^\beta \cdot \alpha, & \alpha^\lambda &= \sup_{\xi < \lambda} (\alpha^\xi). \end{aligned}$$

Let  $\prec$  be a definable order relation on  $\mathbb{N}$ , and let us write  $\text{field}_\prec$  for  $\{x : \exists y(x \prec y \vee y \prec x)\}$ .

Let  $\text{TI}(\prec, A)$  be *transfinite induction* over an ordering  $\prec$  w.r.t.  $A$ :

$$\text{TI}(\prec, A) \quad \forall x \in \text{field}_\prec (\forall y \prec x Ay \rightarrow Ax) \rightarrow \forall x \in \text{field}_\prec Ax$$

and let  $\text{TI}(\prec)$  be  $\text{TI}(\prec, A)$  for all  $A$  of the language. We write  $\prec_y$  for  $\prec$  restricted to  $\{x : x \prec y\}$ . Gentzen's result may now be restated as:

**Theorem 8** *For a particular primitive recursive  $\prec$  of order type  $\varepsilon_0$ , the consistency of  $\mathbf{PA}$  (as formalized in Gödel's second incompleteness theorem) is provable by a transfinite induction over  $\prec$  w.r.t. a quantifier-free statement. Hence  $\text{TI}(\prec)$  cannot be proved in  $\mathbf{PA}$ ; however, for all  $y \in \text{field}_\prec$   $\text{TI}(\prec_y)$  is provable.*

What matters is not just the ordertype of the ordering, but the way the ordering is given to us. In fact, the result concerns not so much ordertypes, but rather ordinal notations, that is to say an encoding of a well-ordering, together with some extra structure on it. For the ordinals  $\varepsilon_0$  we can obtain notations (encodings in the natural numbers) using the fact that each ordinal  $\beta$  below  $\varepsilon_0$  has a unique Cantor normal form:

$$\beta = \omega^{\alpha_n} + \omega^{\alpha_{n-1}} + \dots + \omega^{\alpha_0},$$

where  $\alpha_0 \leq \alpha_1 \leq \dots \leq \alpha_n$ . When  $\beta < \varepsilon_0$ , then  $\alpha_n < \beta$ . Addition, multiplication, exponentiation may be defined in terms of this standard representation, and it is not hard to see that this can all be encoded in terms of the natural numbers.

Below we describe how to define the proof-theoretic ordinal for theories formulated in the language of second-order arithmetic.

## 7.1 The language

The Tait-style language  $\mathcal{L}_{\mathbb{N}}^2$  contains

- A constant 0 for zero, and constants for all primitive recursive functions and predicates.
- Variables  $x, y, z, \dots$  for numbers and variables  $X, Y, Z, \dots$  for unary predicates (sets). For “ $X$  applies to  $t$ ” we write  $t \in X$ , (with  $Xt$  as an abbreviation).
- Logical operators  $\vee, \wedge, \exists, \forall$ .
- The membership symbol  $\in$  and its negation  $\notin$  (both are primitive symbols!)

*Negation* is defined:

$$\neg(t \in X) := t \notin X, \quad \neg(t \notin X) := (t \in X),$$

$$\neg(Rt_1 \dots t_n) := \bar{R}t_1 \dots t_n \quad (\bar{R} \text{ stands for the complement of } R)$$

and for the logical operators as before for the Tait-calculus.

## 7.2 Order types

From now on a *tree* is a set of (codes of) finite sequences closed under initial segments. The elements of the tree are called *nodes*. A set of nodes is a *thread* if it is linearly ordered by the relation ‘initial segment of’; a maximal thread is a *path*. let  $\prec$  be a binary relation on the natural numbers; it is *well-founded* iff there are no infinite  $\prec$ -descending threads. We put:

$$\text{otyp}_{\prec}(n) := \begin{cases} \sup\{\text{otyp}_{\prec}(m) + 1 : m \prec n\} & \text{for } n \text{ in the field of } \prec, \\ \omega_1 & \text{otherwise.} \end{cases}$$

$$\text{otyp}(\prec) := \sup\{\text{otyp}_{\prec}(n) : n \in \text{field}_{\prec}\}$$

$$s \prec_T t := s \in T \wedge t \in T \wedge t \subset s$$

$$\text{otyp}_T(s) := \text{otyp}_{\prec_T}(s)$$

$$\text{otyp}(T) := \text{otyp}_T(\langle \rangle)$$

$$\omega_1^{\text{CK}} := \sup\{\text{otyp}(\prec) : \prec \text{ is a recursive well-ordering}\}$$

$\omega_1^{\text{CK}}$  is the least non-recursive ordinal and is usually called the *Church-Kleene ordinal*.

### 7.3 Truth-complexity

We define a relation  $\models^\alpha \Delta$ ,  $\alpha$  an ordinal,  $\Delta$  a set of first-order formulas by the following clauses ( $t^{\mathbb{N}}$  is the value of a closed term  $t$  in the structure  $\mathbb{N}$ ):

1. If  $\Delta$  contains a true atomic sentence, then  $\models^\alpha \Delta$  for all ordinals  $\alpha$ .
2. If  $t^{\mathbb{N}} = s^{\mathbb{N}}$  then  $\models^\alpha \Delta, s \notin X, t \in X$  for all ordinals  $\alpha$ .
3. If  $\models^{\alpha_i} \Delta, A_i, \alpha_i < \alpha$  for  $i = 1, 2$  then  $\models^\alpha \Delta, A_1 \wedge A_2$ ,
4. If  $\models^{\alpha_i} \Delta, A_i, \alpha_i < \alpha$  for  $i = 1$  or  $i = 2$ , then  $\models^\alpha \Delta, A_1 \vee A_2$ ,
5. If  $\models^{\alpha_i} \Delta, A(\bar{n})$ , and  $\alpha_i < \alpha$  for all  $i \in \mathbb{N}$ , then  $\models^\alpha \Delta, \forall x A(x)$ ,
6. If  $\models^{\alpha_i} \Delta, A(\bar{n})$ , and  $\alpha_i < \alpha$  for some  $i \in \mathbb{N}$ , then  $\models^\alpha \Delta, \exists x A(x)$

If we forget about the ordinal superscripts, the preceding clauses represent an inductive definition of truth in  $\mathbb{N}$  for the disjunction of sentences in  $\Delta$ , where  $\forall, \exists$  are treated as infinite conjunctions and disjunctions. (Thus the third clause states that  $\forall x A(x)$  is true if  $A(\bar{n})$  is true for all  $n$ .) Alternatively, we may think of the clauses as representing an infinitary version of the Tait-calculus. For example, the third clause may be stated as

$$\frac{A(0), \Delta \quad A(1), \Delta \quad \dots \quad A(\bar{n}), \Delta \quad \dots}{\forall x A(x), \Delta}$$

It will be clear that a true arithmetical formula will have a well-founded infinitary proof tree in such a system: each branch ends in an axiom. Adding the ordinal superscript is in fact assigning an ordinal to a well-founded infinitary proof tree.

A  $\Pi_1^1$ -sentence is a sentence  $\forall X A(X)$ ,  $A$  arithmetic (i.e., with numerical quantifiers only). The clauses of the truth definition above may also serve as a truth definition for  $\Pi_1^1$ -sentences:  $\forall X A(X)$  for first-order  $A$  is true if we can prove  $A(X)$  in the system.

More precisely, one can establish (by a non-trivial proof) an

**Theorem 9** ( *$\omega$ -completeness theorem*): Let  $\forall \vec{X} F(\vec{X})$  be a  $\Pi_1^1$ -sentence, then we have  $\mathbb{N} \models \forall \vec{X} F(\vec{X})$  iff there is an  $\alpha < \omega_1^{\text{CK}}$  such that  $\models^\alpha F(\vec{X})$ .

*Truth-complexity*  $\text{tc}$  for a  $\Pi_1^1$ -sentence  $G \equiv \forall \vec{X} F(\vec{X})$  or  $G \equiv F(\vec{X})$  is defined by:

$$\text{tc}(G) := \begin{cases} \omega_1 & \text{if } \mathbb{N} \not\models G, \\ \min\{\alpha : \models^\alpha F(\vec{X})\} & \text{otherwise} \end{cases}$$

Here  $\omega_1$  is the first uncountable ordinal. The  $\omega$ -completeness theorem is now equivalent to

$$\mathbb{N} \models F \text{ iff } \text{tc}(F) < \omega_1^{\text{CK}}.$$

We have the following

**Theorem 10** (*Boundedness theorem*): Let  $\prec$  be an arithmetical definable relation, then

$$\text{otyp}(\prec) \leq \text{tc}(\forall X \text{TI}(\prec, X)).$$

## 7.4 The proof-theoretic ordinal of a theory

We put

$$\text{spec}(\mathbb{N}) := \{\text{tc}(F) : F \text{ is a } \Pi_1^1\text{-sentence, } \mathbb{N} \models F\}$$

and for any enumerable theory  $\mathbf{T}$ :

$$\begin{aligned} \text{spec}(\mathbf{T}) &:= \{\text{tc}(\forall \vec{X} F(\vec{X})) : \mathbf{T} \models F(\vec{X})\} \\ \|\mathbf{T}\|_{\Pi_1^1} &:= \sup(\text{spec}(\mathbf{T})) \\ \|(T)\| &:= \sup\{\text{otyp}(\prec) : \prec \text{ is primitive recursive, } \mathbf{T} \vdash \text{TI}(\prec, X)\} \end{aligned}$$

By the boundedness theorem stated above for all theories  $\mathbf{T}$ :

$$\|\mathbf{T}\| \leq \|\mathbf{T}\|_{\Pi_1^1} < \omega_1^{\text{CK}}$$

If we can also show that for all  $\alpha < \|\mathbf{T}\|_{\Pi_1^1}$  there is a primitive recursive  $\prec$  such that  $\alpha \prec \text{otyp}(\prec)$  and  $(T) \vdash \text{TI}(\prec, X)$ , we may deduce

$$\|\mathbf{T}\| = \|\mathbf{T}\|_{\Pi_1^1}.$$

## 7.5 The method

Let us now outline the method by which the proof-theoretic ordinal for many theories may be found. Let  $\mathbf{T}$  be a recursively axiomatizable theory, such as first-order classical arithmetic  $\mathbf{PA}$ , in the language of  $\mathcal{L}_{\mathbb{N}}^2$ . If  $\mathbf{T} \vdash F$ , there is a cutfree deduction in first-order logic for

$$(a) \quad A_1, \dots, A_n \vdash F$$

where  $A_1, \dots, A_n$  are axioms of  $\mathbf{T}$ . We can determine ordinals  $\alpha, \alpha_i$  ( $1 \leq i \leq n$ ) such that

$$(b) \quad \models^\alpha \neg A_1, \dots, \neg A_n, F$$

and

$$(c) \quad \models^{\alpha_i} A_i \quad (1 \leq i \leq n)$$

From this we want to estimate an ordinal  $\beta$  such that  $\models^\beta F$ . This is done by extending the relation  $\models^\alpha$  to a semiformal system with deducibility relation  $\vdash_\rho^\alpha$ . The system is called *semiformal* because there is a rule with infinitely many premises for introducing  $\forall$ . The rules are similar to those for  $\models^\alpha$ , except that we have now also a Cut rule:

$$\text{If } \vdash_\rho^{\alpha_1} \Gamma, B \text{ and } \vdash_\rho^{\alpha_2} \Gamma', \neg B \text{ then } \vdash_\rho^\alpha \Gamma, \Gamma' \quad (\alpha_0, \alpha_1 < \alpha, \text{complexity}(B) < \rho)$$

$\rho$  is the *cutrank*, and is the maximum logical complexity of cutformulas + 1.  $\vdash_\rho^\alpha$  is not a truth definition, but we do have

$$\models^\alpha B \text{ iff } \vdash_0^\alpha B.$$

Using Cut on (b) and (c), we find indeed an ordinal  $\beta$  and a  $\rho$  such that  $\vdash_\rho^\beta F$ . The semiformal system permits cutelimination, that is to say we can find a specific function  $\tau$  on ordinals such that

$$(d) \quad \text{If } \vdash_\rho^\alpha B \text{ then } \vdash_0^{\tau(\alpha, \rho)} B.$$

In fact, in the case of arithmetic we can show that

$$\text{If } \vdash_{\rho+1}^\alpha \Delta \text{ then } \vdash_\rho^{2^\alpha} \Delta.$$

This easily leads to  $\models^{\varepsilon_0} F$  in the case of arithmetic.

## References

- [BB85] E. Bishop and D. Bridges. *Constructive Analysis*. Springer-Verlag, Berlin, 1985.
- [BFPS81] Wilfried Buchholz, Solomon Feferman, Wolfram Pohlers, and Wilfried Sieg. *Iterated Inductive Definitions and Subsystems of Analysis: Recent Proof-Theoretical Studies*. Lecture Notes in Mathematics 897. Springer-Verlag, Berlin, Heidelberg, New York, 1981.
- [BR87] D. Bridges and F. Richman. *Varieties of Constructive Mathematics*. Cambridge University Press, Cambridge, 1987.
- [Bro75] L. E. J. Brouwer. *Collected Works 1. Philosophy and Foundations of Mathematics*. North-Holland Publ. Co., Amsterdam, and American Elsevier Publ. Co., New York, 1975.
- [Bus98] S. R. Buss, editor. *Handbook of Proof Theory*. Elsevier, Amsterdam, 1998.
- [Dra87] A. G. Dragalin. *Mathematical Intuitionism. Introduction to Proof Theory*. Translations of Mathematical Monographs, volume 67. American Mathematical Society, Providence Rh.I., 1987. Translated from the Russian original, published by Nauka, Moscow 1979.
- [Dum77] M. A. E. Dummett. *Elements of Intuitionism*. Clarendon Press, Oxford, 1977. 2nd ed. 2000.
- [Fit69] M. D. Fitting. *Intuitionistic Logic, Model Theory and Forcing*. North-Holland Publ. Co., Amsterdam, 1969.
- [Gen69] G. Gentzen. *The Collected Papers of Gerhard Gentzen*. North-Holland Publ. Co., Amsterdam, 1969. English translation of Gentzen's papers, edited and introduced by M. E. Szabo.
- [Gir87] J.-Y. Girard. *Proof Theory and Logical Complexity, Volume 1*. Bibliopolis, Napoli, 1987.
- [HB34] D. Hilbert and P. Bernays. *Grundlagen der Mathematik, Bd. I*. Springer-Verlag, Berlin, Heidelberg, New York, 1934. 2nd edition 1968.
- [HB39] D. Hilbert and P. Bernays. *Grundlagen der Mathematik, Bd. II*. Springer-Verlag, Berlin, Heidelberg, New York, 1939. 2nd edition 1970.
- [Her68] J. Herbrand. *Écrits Logiques*. Presses Universitaires de France, Paris, 1968. Translated as *Logical Writings*, Harvard University Press, Cambridge, MA, 1971.
- [Hin97] J. R. Hindley. *Basic Simple Type Theory*. Cambridge University Press, Cambridge, UK., 1997.
- [Kle52] S. C. Kleene. *Introduction to Metamathematics*. North-Holland Publ. Co., Amsterdam, 1952.
- [Kus84] B. A. Kushner. *Lectures on Constructive Mathematical Analysis*. Translations of Mathematical Monographs, volume 60. American Mathematical Society, Providence Rh.I., 1984. Translated from the Russian original, published by Nauka, Moscow 1973.
- [KV65] S. C. Kleene and R. E. Vesley. *The Foundations of Intuitionistic Mathematics*. North-Holland Publ. Co., 1965.

- [Min92a] G. E. Mints. *Selected Papers in Proof Theory*. North-Holland Publ. Co., Amsterdam; Bibliopolis, Napoli, 1992.
- [Min92b] G. E. Mints. *Selected Papers in Proof Theory*. Bibliopolis, Napoli and North-Holland, Amsterdam, 1992.
- [Poh89] W. Pohlers. *Proof Theory. An Introduction*. Springer-Verlag, Berlin, 1989.
- [Pra65] D. Prawitz. *Natural Deduction. A Proof-Theoretical Study*. Almqvist and Wiksell, Stockholm, 1965.
- [Sch68] K. Schütte. *Vollständige Systeme modaler und intuitionistischer Logik*. Springer-Verlag, Berlin, Heidelberg, New York, 1968.
- [Sch77] K. Schütte. *Proof Theory*. Springer-Verlag, Berlin, Heidelberg, New York, 1977.
- [Tak87] G. Takeuti. *Proof Theory*. 2nd edition. North-Holland Publ. Co., Amsterdam, 1987.
- [Tro73] A. S. Troelstra. *Metamathematical Investigation of Intuitionistic Arithmetic and Analysis. Chapters 1–4*. Lecture Notes in Mathematics 344. Springer-Verlag, Berlin, Heidelberg, New York, 1973.
- [TS96] A. S. Troelstra and H. Schwichtenberg. *Basic Proof Theory*. Cambridge University Press, Cambridge UK., 1996. To appear.
- [TvD88] A. S. Troelstra and D. van Dalen. *Constructivism in Mathematics*. Studies in Logic and the Foundations of Mathematics. North-Holland Publ. Co., Amsterdam, 1988. Two vols.