MATHEMATICAL SEMANTICS OF INTUITIONISTIC LOGIC

SERGEY A. MELIKHOV

ABSTRACT. This work is a mathematician's attempt to understand intuitionistic logic. It can be read in two ways: as a research paper interspersed with lengthy digressions into rethinking of standard material; or as an elementary (but highly unconventional) introduction to first-order intuitionistic logic. For the latter purpose, no training in formal logic is required, but a modest literacy in mathematics, such as topological spaces and posets, is assumed.

The main theme of this work is the search for a formal semantics adequate to Kolmogorov's informal interpretation of intuitionistic logic (whose simplest part is more or less the same as the so-called BHK interpretation). This search goes beyond the usual model theory, based on Tarski's notion of semantic consequence, and beyond the usual formalism of first-order logic, based on schemata. Thus we study formal semantics of a simplified version of Paulson's meta-logic, used in the `Isabelle` prover.

By interpreting the meta-logical connectives and quantifiers constructively, we get a generalized model theory, which covers, in particular, realizability-type interpretations of intuitionistic logic. On the other hand, by analyzing Kolmogorov's notion of semantic consequence (which is an alternative to Tarski's standard notion), we get an alternative model theory. By using an extension of the meta-logic, we further get a generalized alternative model theory, which suffices to formalize Kolmogorov's semantics.

On the other hand, we also formulate a modification of Kolmogorov's interpretation, which is compatible with the usual, Tarski-style model theory. Namely, it can be formalized by means of sheaf-valued models, which turn out to be a special case of Palmgren's categorical models; intuitionistic logic is complete with respect to this semantics.

CONTENTS

1. Introduction

1.1. Formal semantics

Much of this long treatise is devoted to understanding in rigorous mathematical terms A. Kolmogorov's 1932 short note "On the interpretation of intuitionistic logic" [108]. What the logical community has grasped of this text is contained in the so-called BHK interpretation, often considered to be "the intended interpretation" of intuitionistic logic and very extensively analyzed in the literature (see §§**3.7**–**3.8** below). However, as observed in the classic textbook by Troelstra and van Dalen [202], "the BHK-interpretation in itself has no 'explanatory power'" since "on a very 'classical' interpretation of [the informal notions of] construction and mapping," its six clauses "justify the principles of two-valued (classical) logic." But as discussed in §§**3.9**–**3.13** below, this arguably does not apply to Kolmogorov's original interpretation.

The problem with Kolmogorov's original interpretation is that it is incompatible with the usual model theory, based on Tarski's notion of semantic consequence; worse yet, it is an interpretation of something that does not exist in the usual formalism of first-order logic, based on schemata.

Instead of this usual formalism we use a simplified (and, eventually, slightly extended) version of L. Paulson's 1989 meta-logic [157], which is the basis of the `Isabelle` prover (see §§**4.1**–**4.6**). (In fact, it builds not on the standard modern formalism but rather on the early tradition of first-order logic, as in the textbooks by Hilbert–Bernays [90] and Hilbert–Ackermann [89].) The meta-logical connectives and quantifiers can be used to express, for instance, the horizontal bar and comma in inference rules (in fact, the entire deductive system of a first-order logic is a single meta-formula); derivability of a formula or a rule; implication between principles (in the usual sense, which has nothing to do with ordinary implication between formulas or schemata); variables fixed in a derivation (in the sense of Kleene), etc. In this approach, principles and rules, and even implications between them become finite entities within the formal language (without any meta-variables — even in situations where the standard formalism would have side conditions such as "where x is not free in F" or "where t is free for x in $F(x)$") and so one can speak of their interpretation by mathematical objects.

Formal semantics of the meta-logic of first-order logics seems to have never been developed, however; thus we have to undertake this task (see §§**4.7**–**4.8** and §7). In usual model theory, regarded as a particular semantics of the meta-logic, the meta-logical connectives and quantifiers are interpreted classically (and moreover in the two-valued way) — even though they are constructive. By interpreting them constructively, we get a generalized model theory (see §**7.1**). It suffices to formalize, in particular, realizability-type interpretations of intuitionistic logic, which do not fit in the usual model theory (see §*7.1.4*). However, to formalize Kolmogorov's interpretation of inference rules (which is a substantial part of his interpretation of intuitionistic logic, and seems to be uniquely possible in his approach), we need an alternative generalized model theory (see §**7.3**). It is nothing new for formulas and principles, but an entirely different interpretation of

rules (and more complex meta-logical constructs) already in the case where the meta-logical connectives and quantifiers have a classical, two-valued interpretation (see §7.2 concerning this case).

We also formulate an alternative interpretation of intuitionistic logic, called "modified BHK-interpretation", which is compatible with the usual, Tarski-style model theory. A specific class of usual models of intuitionistic logic that fully meets the expectations of the modified BHK-interpretation is described. These are sheaf-valued, or "proof-relevant topological" models (§6) of intuitionistic logic — not to be confused with the usual open set-valued "sheaf models" (see [13] and references there). As one could expect, sheaf-valued models turned out to be a special case of something well-known: the "categorical semantics" (cf. [121]), and more specifically Palmgren's models of intuitionistic logic in locally cartesian closed categories with finite sums [156]. But unexpectedly, this very simple and natural special case with obvious relevance for the informal semantics of intuitionistic logic does not seem to have been considered per se (apart from what amounts to the $\wedge, \rightarrow$ fragment [11]). We prove completeness of intuitionistic logic with respect to sheaf-valued models by showing that in a certain rather special case (sheaves over zero-dimensional separable metrizable spaces) they interpret the usual topological models over the same space. Normally (e.g. for sheaves over Euclidean spaces or Alexandroff spaces) this is far from being so. On the other hand, our sheaf-valued models also interpret Medvedev–Skvortsov models (§3.13).

1.2. Informal semantics

We adopt and develop Kolmogorov's understanding of intuitionistic logic as the logic of schemes of solutions of mathematical problems (understood as tasks). In this approach, intuitionistic logic is viewed as an extension package that upgrades classical logic without removing it — in contrast to the standard conception of Brouwer and Heyting, which regards intuitionistic logic as an alternative to classical logic that criminalizes some of its principles. The main purpose of the upgrade comes, for us, from Hilbert's idea of equivalence between proofs of a given theorem, and from the intuition of this equivalence relation as capable of being nontrivial. This idea of "proof relevance" amounts to a categorification; thus, in particular, our sheaf-valued semantics of intuitionistic logic can be viewed as a categorification of the familiar Leibniz–Euler–Venn subset-valued semantics of classical logic.

Traditional ways to understand intuitionistic logic (semantically) have been rooted either in philosophy — with emphasis on the development of knowledge (Brouwer, Heyting, Kripke) or in computer science — with emphasis on effective operations (Kleene, Markov, Curry–Howard). Kolmogorov's approach emphasizes the right order of quantifiers, and is rooted in the mathematical tradition, dating back to the Antiquity, of solving of problems (such as geometric construction problems) and in the mathematical idea of a method (of solution or proof). The approach of Kolmogorov has enjoyed some limited appreciation in Europe in the 1930s, and had a deeper and more lasting impact

on the Moscow school of logic[1] — but nowadays is widely seen as a mere historical curiosity. One important reason for that[2] could have been the lack of a sufficiently clear formulation and of any comprehensive exposition of intuitionistic logic based on this approach. The present treatise attempts to remedy this deficiency.

Our introduction to some basic themes of intuitionistic logic (§5) is in the language of its usual topological models, which were discovered independently by Stone, Tang and Tarski in the 1930s but largely went out of fashion since the 60s. (However, in the zero-order case these models include, in disguise, Kripke models, which have had far better luck with fashion.) We present a motivation of topological models through a model of a classical first-order theory extracted from the clauses of the BHK interpretation (§5.1).

Disclaimers

(i) This unconventional introduction to intuitionistic logic should suffice for the purposes of the author's papers [1], [2], [3], but as such it does not pretend to be complete or even finished. It is hoped that a future version will cover certain additional topics and will be accompanied by an unconventional introduction to homotopy type theory.

(ii) Most translations quoted in this paper have been edited by the present author in order to improve syntactic and semantic fidelity. When emphasis is present in quoted text, it is always original.

Acknowledgements

---

[1]In the introduction to her survey of logic in the USSR up to 1957, S. A. Yanovskaya emphasizes "The difference of the viewpoints of the 'Moscow' school of students and followers of P. S. Novikov and A. N. Kolmogorov and the 'Leningrad' school of students of A. A. Markov on the issue of the meaning of non-constructive objects and methods", where the former viewpoint "consists in allowing in one's work not only constructive, but also classical methods of mathematics and mathematical logic" [213]. In the West, the principles of the 'Leningrad' school are better known as "Russian constructivism" (cf. [202]).

[2]Besides the lack of an English translation of Kolmogorov's paper [108] until the 1990s, and the lack of a word in English to express the notion of a "mathematical task" (German: Aufgabe, Russian: задача) as distinguished from the notion of an "open problem" (German: Problem, Russian: проблема). Let us note, in particular, that Kolmogorov himself, when writing in French, took care to mention the German word and so spoke of "problèmes (Aufgaben)" [109].

it in a way more compatible with the mentality of a "working mathematician". I'm also grateful D. Botin and D. Cenzer, who taught me some bits of logic many years ago.

2. MOTIVATION: HOMOTOPIES BETWEEN PROOFS

2.1. Hilbert's 24th Problem

In 1900, Hilbert published (and partly presented in his lecture at the International Congress of Mathematicians in Paris) a list of 23 open problems in mathematics [85]. Shortly before the 100th anniversary of the list, the historian Rüdiger Thiele discovered a would-be 24th problem in Hilbert's notebooks [195] (for a photocopy see [106] and [196]). We don't know why Hilbert refrained from publicizing it, but one obvious possibility is that he simply did not see how to state it precisely enough.

> "The 24th problem in my Paris lecture was to be: Criteria of simplicity, or proof of the greatest simplicity of certain proofs. Develop a theory of the method of proof in mathematics in general. Under a given set of conditions there can be but one simplest proof. Quite generally, if there are two proofs for a theorem, you must keep going until you have derived each from the other, or until it becomes quite evident what variant conditions (and aids) have been used in the two proofs. Given two routes, it is not right to take either of these two or to look for a third; it is necessary to investigate the area lying between the two routes. Attempts at judging the simplicity of a proof are in my examination of syzygies and syzygies between syzygies (see [[84]], Lectures XXXII–XXXIX). The use or the knowledge of a syzygy simplifies in an essential way a proof that a certain identity is true. Because any process of addition [is] an application of the commutative law of addition etc. [and because] this always corresponds to geometric theorems or logical conclusions, one can count these [processes], and, for instance, in proving certain theorems of elementary geometry (the Pythagoras theorem, [theorems] on remarkable points of triangles), one can very well decide which of the proofs is the simplest."[3]

A rather narrow understanding of the problem, as merely asking for "a criterion of simplicity in mathematical proofs and the development of a proof theory with the power to prove that a given proof is the simplest possible"[4] has became widespread in the literature (see e.g. [20; p. 38], [77], [195]). However, Hilbert's original text above is unambiguously posing a totally different problem as well (in the two sentences starting with "Quite generally"): to develop a proof theory that can tell whether two given proofs of a theorem are "essentially same", or _homotopic_, in the sense that one can _derive each from the other_ (whatever that means). It is this aspect of the 24th problem that we shall be concerned with.

---

[3]Translation by Thiele, who also comments that "a part of the last sentence is not only barely legible in Hilbert's notebook but also grammatically incorrect", and points out that "corrections and insertions that Hilbert made in this entry show that he wrote down the problem in haste".

[4]Quoted from Wikipedia's "Hilbert's 24th problem" (English) as of 11/11/11

Moreover, Hilbert's reference to syzygies between syzygies is obviously[5][6] provoking one to think of a hypothetical space of proofs of a given theorem and of whether it should be acyclic or contractible (so that proofs can be simplified in a canonical way) under a given set of conditions.

Certainly, Hilbert was not alone in being concerned with (informal) questions of this kind. For instance, a sample of attempts by mathematicians other than proof theorists at judging "essential sameness" of certain proofs (of the irrationality of $\sqrt{2}$ and of Gauss' quadratic reciprocity law) can be found in T. Gowers' blog [72]. A more systematic study of the question of identity of proofs has taken place in intuitionistic proof theory, starting in the 70s (Prawitz [165], Kreisel [112], Martin-Löf [125]; see a brief survey in [42; §2]) and continued in the context of higher categories and higher λ-calculus [121], [178], [78], [91]. The more recent approach of homotopy type theory [92] is to a large extent a development of this proof-theoretic tradition. We should also point out some very different approaches to somewhat similar questions, which cannot be discussed here in detail: [120] (and references there), [174], [35], [20], [145; §8.2], [23].

So how to judge if two proofs of a theorem are "essentially same"? Hilbert's idea that there might be the *simplest* proof of a theorem Θ and his mention of the *use* of variant conditions (and aids) in proving Θ establish a context where it seems natural to identify a proof p of Θ with every its *simplification*, that is, a proof of Θ obtained from p by crossing out some redundant steps. Of course, the notion of a "redundant step" (or a step that is not "essentially used" in the proof) needs an accurate definition; we will return to this issue later.

On the other hand, Hilbert's words ("keep *going* until you have *derived* each from the other") may seem to hint at a formal theory of proofs containing a number of specified elementary transformations between proofs of the same assertion. In this case the relation of *homotopy* between proofs of the same assertion can be defined as the equivalence relation generated by elementary transformations.

Such formal theories endowed with elementary transformations between proofs are best known in the case of constructive proofs (in particular, the natural deduction system of Gentzen–Prawitz and the corresponding λ-calculus, see [64]). Before turning to constructive proofs, let us review the situation with usual, non-constructive ones.

---

[5]Let us also note the connection between (higher) syzygies and (higher) Martin-Löf identity types. In Voevodsky's model of Martin-Löf's type theory, types are interpreted as Kan complexes (see [92]), so one kind of a type is an abelian simplicial group, which by the Dold–Kan correspondence (see [67]) can be identified with a pointed chain complex C of abelian groups: $b \in C_{-1} \xleftarrow{\partial_0} C_0 \xleftarrow{\partial_1} C_1 \xleftarrow{\partial_2} C_2 \xleftarrow{\partial_3} \ldots$ Write $t : C$ to mean $t \in \partial_0^{-1}(b)$, and $t = u : C$ to mean that $t : C$ and $u : C$ are homologous relative to $\partial_0(t) = \partial_0(u)$, that is, $t - u \in \partial_1(C_1)$. Then the identity type $Id_C(t, u)$ will be the following pointed chain complex: $t - u \in C_0 \xleftarrow{\partial_1} C_1 \xleftarrow{\partial_2} C_2 \xleftarrow{\partial_3} \ldots$

[6]Perhaps we should also note the connection between (higher) identity types and (higher) "homotopical syzygies" (see [123]), which goes back to J. H. C. Whitehead's classical work on identities among relations (see [4]). The role of the Dold–Kan correspondence is played here by the not so well understood connection (see [146]) between extended group presentations of [30], where higher relators correspond to higher cells of a CW complex, and free simplicial groups, such as the Kan loop group (see [67]) of a simplicial set.

2.2. The collapse of non-constructive proofs

A standard theory of derivations based on classical logic is Gentzen's sequent calculus [65] (see [64]). In particular, the elementary inferences

(Cut) from $A \vee C$ and $C \rightarrow B$ infer $A \vee B$;

(C_R) from $A \vee A$ infer A;

(W_{R1}) from A infer $A \vee B$;

(W_{R2}) from A infer $B \vee A$;

(W_L) from B infer $C \rightarrow B$

can be interpreted as instances of the inference rules of sequent calculus.[7] A number of elementary transformations between derivations of the same proposition in sequent calculus can be found in the proof of Gentzen's Hauptsatz on cut elimination (see [64], [65]). These Γ-*transformations* include, in particular, the following (Γ_L) and (Γ_R).

(Γ_L) Suppose that q is a proof of $C \rightarrow B$. Then the following fragment of a proof:
- A implies $A \vee C$ by (W_{R1});
- $C \rightarrow B$ by q;
- hence $A \vee B$ by (Cut)

transforms into
- A implies $A \vee B$ by (W_{R1}).

We note that (Γ_L) can be seen as a simplifying transformation: it eliminates the proof q of the redundant lemma $C \rightarrow B$ at the cost of replacing one weakening (A implies $A \vee C$) by another (A implies $A \vee B$).

(Γ_R) Suppose that p is a proof of $A \vee C$. Then the following fragment of a proof:
- B implies $C \rightarrow B$ by (W_L);
- $A \vee C$ holds by p;
- hence $A \vee B$ by (Cut)

transforms into
- B implies $A \vee B$ by (W_{R2}).

The symmetry between (Γ_L) and (Γ_R) is highlighted in Gentzen's original notation [65; III.3.113.1-2]. We can see it explicitly by rewriting $C \rightarrow B$ and $A \vee C$ respectively as $B \vee \neg C$ and $\neg C \rightarrow A$, using tautologies of classical logic.

Let us also consider two additional Λ-*transformations*:

(Λ_1), (Λ_2): The following fragment of a proof can be eliminated:
- A implies $A \vee A$ by (W_{R1}) or by (W_{R2});
- $A \vee A$ implies A by (C_R).

---

[7]Namely, (Cut) can be interpreted as an instance of the cut rule, (C_R) as an instance of the right contraction rule, (W_{R1}) and (W_L) as instances of the right and left weakening rules, and (W_{R2}) as (W_{R1}) followed by an instance of the exchange rule.

Theorem 2.1 (Lafont [64; Appendix B.1]; see also [42; §7]). *All proofs of any given theorem are equivalent modulo Γ- and Λ-transformations.*

Lafont's own summary of his result is that "classical logic is inconsistent, not from the logical viewpoint ([falsity] is not provable) but from an algorithmic one."

Proof. Let p and q be proofs of a theorem Θ. They can be augmented to yield a proof $p + W_{R1}$ of $\Theta \vee \Phi$, where Φ is some fixed assertion, and a proof $q + W_L$ of the implication $\Phi \to \Theta$. Combining the latter two proofs, we obtain a proof $(p + W_{R1}) + (q + W_L) + Cut$ of the assertion $\Theta \vee \Theta$. Hence we get a third proof, $(p + W_{R1}) + (q + W_L) + Cut + C_R$, of Θ, which we will denote $p \oplus q$ for brevity.[8]

It remains to show that $p \oplus q$ can be reduced to p, as well as to q by Γ- and Λ-transformations. Indeed, Γ_L reduces $p \oplus q$ to the proof $p + W_{R1} + C_R$ of Θ (which augments p by saying that Θ implies $\Theta \vee \Theta$, which in turn implies Θ). The latter proof is reduced to p by Λ_1. Similarly, Γ_R reduces $p \oplus q$ to the proof $q + W_{R2} + C_R$ of Θ, which in turn is reduced to q by Λ_2. $\square$

It should be noted that, despite Lafont's theorem, some authors have been able to find inequivalent proofs in the framework of classical logic. See, in particular, Došen [42] (see details in [43]; see also [44] and references there), Kuznets [117] and references there; and Guglielmi [73], [74] (see also [75] and references there)

Remark 2.2. One issue with Γ-transformations is that (Γ_L) does not merely eliminate q, but only does so at the expense of changing a weakening. So it is not immediately clear that the lemma $C \to B$ is redundant in the initial proof of (Γ_L), in the strict sense of not being used at all in that proof.

To address this concern, let us revisit Lafont's construction.[9] Recall that we were free to choose any assertion as Φ. We can set Φ, for instance, to be the negation $\neg\Theta$ (any weaker assertion, such as $\Theta \vee \neg\Theta$, would also do). Then $\Theta \vee \Phi$ holds in classical logic regardless of validity of Θ. Therefore the step W_{R1} of the proof $p \oplus q$ replaces the proved theorem Θ with the tautology $\Theta \vee \Phi$. The effect of this step is that we forget all that we have learned from the proof p. Thus $p \oplus q$ does not essentially use p, so p must be redundant even in the stricter sense.

On the other hand, we could take Φ to be Θ itself (any stronger assertion, for instance $\Theta \wedge \neg\Theta$, would also do). Then the step W_L replaces the proved theorem Θ by the tautology $\Phi \to \Theta$. The effect of this step is that we forget all that we have learned from q, and we similarly conclude that q must be redundant. However, it appears that the meaning of the proof $p \oplus q$ should not vary depending on the strength of Φ, since for the purposes of this proof we did not need to know anything at all about Φ.

Our conclusion, that neither p nor q is "essentially used" in $p \oplus q$, shows that the notion of a proof "essentially using" its part cannot be sensibly formalized in the framework of

---

[8]A version of this $\oplus$ is also found in Artemov's Logic of Proofs (see [101]).

[9]What follows is only an informal discussion, perhaps raising more issues than it attempts to settle.

classical logic. It turns out, however, that this notion can be consistently defined once we abandon classical logic in favor of intuitionistic.

2.3. About constructive proofs

In contrast to classical logic, where one can sometimes prove that a certain object exists without giving its explicit construction, proofs in intuitionistic logic can be interpreted, at least very roughly, as algorithms (the Curry–Howard correspondence, see [64], [188]). For example, if S is a set, an intuitionistic proof of the assertion "S is nonempty" can be roughly thought of as a program (executed on a Turing machine, say) that returns a specific element of S. (These issues will be discussed in more detail in §**3.12** and §*7.1.4*.)

In writing a computer program, it is not enough to know that some subsidiary computation is possible in principle; for the program to run, every step must be actually implemented. In fact, the value of every particular variable can be chased through all dependent subroutines using a debugger. This is in sharp contrast with classical mathematics, where the particular way that a lemma is proved has absolutely no effect on the proof of the theorem modulo that lemma. In intuitionistic logic, there can be no dependence between statements other than a dependence between their specific proofs.

With the above in mind, although $\Theta \vee \Phi$ is intuitionistically a tautology as long Φ is (for instance, Φ can be taken to be $\Theta \to \Theta$), the intuitionistic step W_{R1}, replacing the proved theorem Θ with the tautology $\Theta \vee \Phi$ does not have the effect of forgetting how Θ was proved. On the contrary, it has the effect of specifying that $\Theta \vee \Phi$ is to be proved "via Θ" and not "via Φ", and extending a given proof of Θ to a specific proof of $\Theta \vee \Phi$. Since the latter proof does not depend on any proof of Φ, the intuitionistic step Cut promotes it to a proof of $\Theta \vee \Theta$ that does not depend on any proof assuming Φ as a premise.

It follows that if p and q are constructive, then $p \oplus q$, understood constructively, essentially uses p and does not essentially use q. In agreement with this, the intuitionistic version of the sequent calculus (see [64]) admits only asymmetric interpretations of $p \oplus q$, which can be reduced by the intuitionistic analogues of Γ- and Λ-transformations to p, but not to q.

3. What is intuitionistic logic?

3.1. Semantics of classical logic

In classical logic, the meaning of propositions is determined by answering the question: *When is a given proposition true?* A standard answer to this question is given by Tarski's definition of truth. The definition is by induction on the number of logical connectives and quantifiers:

- which primitive propositions are true is assumed to be known from context.

For instance, in the context of arithmetic, a proposition of the form $t = s$ (with t and s being arithmetical expressions) is true whenever t and s rewrite to the same numeral.

- $P \vee Q$ is true if P is true or Q is true;
- $\neg P$ is true if P is not true;
- $\exists x \, R(x)$ is true if $R(x_0)$ is true for some $x_0 \in \mathscr{D}$.

Here P, Q are propositions and R is a unary predicate, i.e., a proposition with one parameter. In first-order logic, all parameters are understood to run over a fixed set $\mathscr{D}$, which is called the *domain of discourse*. Thus $R(x_0)$ is a proposition for any $x_0 \in \mathscr{D}$

The other classical connectives and quantifiers $\top$, $\bot$ $\wedge$, $\rightarrow$, $\forall$ do not need to be mentioned since they are expressible in terms of $\vee$, $\neg$ and $\exists$. For the record, we get for them, as a consequence of the above:

- $\top$ is true;
- $\bot$ is false;
- $P \wedge Q$ is true if P is true and Q is true;
- $P \rightarrow Q$ is true if Q is true or P is not true;
- $\forall x \, R(x)$ is true if $R(x_0)$ is true for all $x_0 \in \mathscr{D}$.

We can also understand the latter six clauses the other way round, as an "explanation" of classical connectives and quantifiers in terms of truth of propositions.

Remark 3.1. In this chapter we always assume our meta-logic to be classical; that is, when discussing intuitionistic or any other logic, we always mean *this discussion* to take place in classical logic. It the next chapter, some of that discussion will be formalized (in fact, more of it than in any standard text on first-order logic), so we will speak of a formal meta-logic, which in fact will turn out to be not exactly classical (but for many purposes it can assumed to be classical); but then the meta-meta-logic, in which all discussions of that formal meta-logic take place, will still be classical.

Intuitionistic logic can be approached in several ways, of which we will discuss the two earliest ones: the Brouwer–Heyting approach (briefly), and Kolmogorov's approach (in more detail).

3.2. Platonism vs. Verificationism

The tradition initiated by Brouwer in 1907 and continued by his school in Amsterdam (Heyting, Troelstra, et al.) and elsewhere till the present day — perhaps more successfully in computer science than in mathematics — presupposes a deep rethink (along the lines of Logical Positivism and prior development of verificationist thinking; see Martin-Löf [131], [127; pp. 414–416]), of the conventional understanding of what mathematics is and what it is not. In the words of Troelstra and van Dalen [202; p. 4]:

"It does not make sense to think of truth or falsity of a mathematical statement independently of our knowledge concerning the statement. A statement is *true* if we have [a constructive] proof of it, and *false* if we can show that the assumption that there is a [constructive] proof for the statement leads to a contradiction. For an arbitrary statement we can therefore not assert that it is either true or false."

Thus, in the tradition of Brouwer and Heyting, propositions are considered to have a different — intuitionistic — meaning, which is determined by answering the question: *When do we know a constructive proof of a given proposition?* We will return to this question in §3.7, once it will have been restated in §3.4 so as to avoid the terminological clash with the classical meaning of propositions.

Whether the classical meaning of propositions "makes sense" or not is a philosophical question which we will try to avoid — so as to focus on answering mathematical questions. But at the very least, the author is determined to keep good relations with the 99% of mathematicians who are convinced that the classical meaning of propositions does make sense; and so considers the *terminology* of Brouwer and Heyting, which is in direct conflict with the convictions of the 99% of mathematicians, to be unfortunate. Indeed, a useful terminology would provoke mathematical, and not philosophical questions.

The depth of the rethink of mathematics promoted by the battlefield terminology of Brouwer and Heyting should not be underestimated. In the above quote, we are asked to abstain not only from using the principle of excluded middle (in the form asserting that every statement is either true or false), but also from understanding mathematical objects with the customary mental aid of Platonism (i.e. Plato's vision of mathematical objects as ideal entities existing independently of our knowledge about them). To be sure, Heyting has stated it boldly enough [80]:

> "Faith in transcendental existence, unsupported by concepts, must be rejected as a means of mathematical proof, [... which] is the reason for doubting the law of excluded middle."

The present author is no proponent of Platonism, but, on the contrary, feels that properly understood mathematics should not depend on any particular philosophy (for reasons related to Occam's Razor). Heyting certainly has a point in that when we mathematicians speak of "existence" of mathematical objects and "truth" of mathematical theorems, it is not at all clear what we mean by this; the implicit hypothesis that this talk "makes sense" at all can certainly be considered as a religious dogma. On the other hand, basic concepts of Intuitionism such as "construction", "proof" and "contradiction" are just as obscure as "existence" and "truth"; and the intuitionists' conviction that these concepts (as well as their "explanation" by clauses of the BHK interpretation, cf. §3.7 below) "make sense" is certainly no less of a religious dogma.

Thus the claims of greater meaningfulness and lesser dogmatism that Intuitionism pretends to have over classical mathematics appear to be highly controversial. Can we, nevertheless, *understand* what is intuionistic mathematics?

3.3. The issue of understanding

From the viewpoint of Intuitionism, it is trivial to understand classical mathematics: one just needs to explicitly mention the principle of excluded middle in the hypothesis of every classical theorem, thus transforming it into a conditional assertion.

Moreover, as shown and emphasized by Kolmogorov [107], the use of this principle alone cannot be the source of a contradiction (see §**5.6**); nor it can lead to a contradiction when combined with the axioms of first-order arithmetic (Gödel–Gentzen, cf. [129]). The same is true of the second-order arithmetic (Kreisel, cf. [34]) and, with some additional effort, also of the set theory ZF (without the axiom of choice) [58], [163] — albeit this requires impredicative intuitionism, which is something that not all constructivists accept [129], [34]. However, in the presence of the countable axiom of choice (which suffices for classical analysis) in its standard form, interpreted intuitionistically according to Kolmogorov, the addition of the principle of excluded middle generally increases consistency strength even with respect to impredicative intuitionism [21] (see also [153]). In fact, the issue of precise formulation of the axiom of choice is important here; see [34; 2.2.4] and [130].

This is entirely parallel to the situation within classical mathematics. Some principles of classical set theory, including the continuum hypothesis and the axiom of choice, are internally provable to be consistent with ZF; in particular, their addition to ZF does not increase its consistency strength — or in other words, their use alone cannot be the source of a contradiction. Other principles, like large cardinal axioms or the Axiom of Determinacy, are known or conjectured to increase consistency strength when added to ZF; yet there are various good reasons to believe that resulting systems are still consistent, perhaps not significantly different from the reasons for one's belief in the consistency of ZF. Thus classical mathematics with or without the axiom of choice should be, from the viewpoint of an intuitionist, on a par with areas of classical mathematics that depend on the said principles as seen by a classical mathematician who does not find these principles to be intuitively justified. Conversely, it would be beneficial to *understand* intuitionistic logic and its relations with classical logic in a way that is compatible with the plain Platonic intuition arguably shared by most "working mathematicians".

First attempts at meaning explanations of intuitionistic logic in terms of ordinary practices of classical mathematics can be found in a 1928 paper by Orlov [154; §§6,7] (see also [40]) and in the 1930 and 1931 papers by Heyting [79], [80]. Both Orlov and Heyting gave explanations (see [1; *6.2.1*] for a detailed review) — in terms of provability of propositions ("propositions" and "provability" being understood in a sense compatible with classical logic) which directly anticipate Gödel's 1933 provability translation — to be reviewed in §*5.7.3*. Gödel's translation certainly yields an accurate explanation in arguably classical terms (albeit not exactly in terms of classical logic) of the standard formalization of intuitionistic logic. But nevertheless it arguably misses the essence of intuitionistic logic, in that it does not explain how one proposition can have essentially different proofs (cf. §2).

3.4. Solutions of problems

The second approach to intuitionistic logic, initiated by Kolmogorov in 1932 and continued to some extent by his school in Moscow aimed at being less demanding of one's philosophical convictions. In the words of Kolmogorov [108]:

> "On a par with theoretical logic, which systematizes schemes of proofs of theoretical truths, one can systematize schemes of solutions of problems — for example, of geometric construction problems. For instance, similarly to the principle of syllogism we have the following principle here: *If we can reduce the task of solving b to that of a, and the task of solving c to that of b, then we can also reduce the task of solving of c to that of a.*
>
> Upon introducing appropriate notation, one can specify the rules of a formal calculus that yield a symbolic construction of a system of such problem solving schemes. Thus, in addition to theoretical logic, a certain new *calculus of problems* arises. In this setting there is no need for any special, e.g. intuitionistic, epistemological presuppositions.
>
> The following striking fact holds: *The calculus of problems coincides in form with Brouwer's intuitionistic logic, as recently formalized by Mr. Heyting.*
>
> In the second section we undertake a critical analysis of intuitionistic logic, accepting general intuitionistic presuppositions; and observe that intuitionistic logic should be replaced with the calculus of problems, since its objects are in reality not theoretical propositions but rather problems."

In this paper, we use the word *problem* solely in the sense of a request (or desire) to find a construction meeting specified criteria on output and permitted means (as in "chess problem", "geometric construction problem", "initial value problem"). With *problem* understood in this sense, Kolmogorov's interpretation is, in the words of Martin-Löf [126; p. 6],

> "very close to programming. '*a* is a method [of solving the problem (doing the task) *A*] can be read as '*a* is a program ...'. Since programming languages have a formal notation for the program *a*, but not for *A*, we complete the sentence with '... which meets the specification *A*'. In Kolmogorov's interpretation, the word *problem* refers to something to be done and the word *program*[10] to how to do it."

A key difference between problems and propositions is the lack of a direct analogue of the Platonic concept of truth in the case of problems. In the words of Kolmogorov [108]:

> "It is, however, essential that every proved proposition is *correct*; for problems one has no concept corresponding to this correctness."

As a matter of convention, we will always understand a *solution* of a problem positively, in that a proof that a problem Γ has no solutions is not considered to be a solution of Γ. Instead, we consider it to be a solution of a different problem, denoted $\neg\Gamma$; and the problem of either finding a solution of Γ or proving that one does not exist is denoted $\Gamma \vee \neg\Gamma$. We say that Γ is *solvable* if Γ has a solution, and *decidable* if $\Gamma \vee \neg\Gamma$ has a solution.

---

[10]Kolmogorov's original interpretation does not mention any "programs", so this is likely a misprint, and the phrase is to be read as "the word *solution*".

3.5. The Hilbert–Brouwer controversy

A problem of the form $\Gamma \vee \neg\Gamma$ may in general be very hard. There is nothing surprising here, of course, since the principle of excluded middle applies to propositions, whereas Γ is a problem. Nevertheless, the principle of decidability of every mathematical problem has been stressed forcefully by Hilbert in the preface to his famous list of problems [85]:

"Occasionally it happens that we seek the solution [of a mathematical problem] under insufficient hypotheses or in an incorrect sense, and for this reason do not succeed. The problem then arises: to show the impossibility of the solution under the given hypotheses, or in the sense contemplated. Such proofs of impossibility were effected by the ancients, for instance when they showed that the ratio of the hypotenuse to the side of an isosceles right triangle is irrational. In later mathematics, the question as to the impossibility of certain solutions plays a preeminent part, and we perceive in this way that old and difficult problems, such as the proof of the axiom of parallels, the squaring of the circle, or the solution of equations of the fifth degree by radicals have finally found fully satisfactory and rigorous solutions, although in another sense than that originally intended. It is probably this important fact along with other philosophical reasons that gives rise to the conviction (which every mathematician shares, but which no one has as yet supported by a proof) that every definite mathematical problem must necessarily be susceptible of an exact settlement, either in the form of an actual answer to the question asked, or by the proof of the impossibility of its solution and therewith the necessary failure of all attempts. Take any definite unsolved problem, such as the question of the irrationality of the Euler–Mascheroni constant C, or of the existence of an infinite number of prime numbers of the form $2n + 1$. However unapproachable these problems may seem to us and however helpless we stand before them, we have, nevertheless, the firm conviction that their solution must follow by a finite number of purely logical processes."

A counterexample to Hilbert's ascription of his conviction to all mathematicians was soon provided by Brouwer, who himself fiercely opposed it [26; Statement XXI], [27], [28]. Yet one can hardly deny that the task of *settlement*, in the sense of Hilbert, of mathematical problems (including, indeed, those in Hilbert's list), i.e., the task of solving problems of the form $\Gamma \vee \neg\Gamma$, is something that we mathematicians all constantly undertake, with great efforts — including those 99% of us who consider the principle of excluded middle to be a triviality, which can be used freely whenever the need arises.

We thus come to distinguish the *principle of excluded middle* (for propositions) and the *principle of decidability* (for problems). Yet the two have been systematically conflated by Heyting (see [80] and [201; p. 235]) and especially by Brouwer; in fact, their conflation was at the very root of Brouwer's opposition to the principle of excluded middle, whose first published record [27] reads:

Firstly, the *syllogism* infers from the embedding [inpassing][11] of a system [systeem] b into a system c, along with the embedding of a system a into the system b, a direct embedding of the system a into the system c. This is nothing more than a tautology.

Likewise the principle of *non-contradiction* is indisputable: The accomplishment of the embedding of a system a into a system b in a prescribed manner, and the obstruction showing the impossibility of such an embedding, exclude each other.

Now consider the principle of *excluded middle*: It claims that every hypothesis is either true or false; in mathematics this means that for every hypothetical embedding of one system in another, satisfying certain given conditions, we can either accomplish such an embedding by a construction, or we can construct the obstruction showing its impossibility. It follows that the question of the validity of the principle of excluded middle is equivalent to the question whether undecidable mathematical problems can exist. There is no indication of a justification for the occasionally pronounced (see [Hilbert's problem list [85] ...]) conviction that there exist no undecidable mathematical problems.

This all is remarkably similar to what one could say about the *problem* of "embedding of one system in another" (whatever that means), except for the alleged connection with the principle of excluded middle. Twenty years later, Brouwer has even complained that after his repeated declarations that the principle of excluded middle is the same as the principle of decidability, Hilbert still fails to recognize their alleged identity [28]. Unfortunately, neither this complaint nor the repeated declarations cited in it seem to bring in any arguments as to why the two principles should be equated (apart from the insistence that they obviously should). In fact, in the opinion of Kolmogorov [110],

> "The principle of excluded middle is, according to Brouwer, inapplicable only to judgements of a special kind, in which a theoretical proposition is inseparably linked with construction of the object of the proposition. Therefore, one may suppose that Brouwer's ideas are actually not at all in contradiction with the traditional logic, which has in fact never dealt with judgements of this kind."

3.6. Problems vs. conjectures

Obviously, Brouwer committed what most mathematicians still do today: he conflated a problem (understood as a request for finding, for instance, a proof of a certain assertion, according to a certain specification of what counts as a proof) with a conjecture (understood as a judgement that a certain assertion is true). Indeed, most of Hilbert's

---

[11]This is the first occurrence of this term in the text, but it was explained in Brouwer's dissertation [26; p. 77], where he apparently rejects the notion of a subset defined by a property of its elements in favor of an explicitly constructed subset: "Often it is quite easy to construct inside such a structure, independently of how it originated, new structures, whose elements are taken from the elements of the original structure or systems of these, arranged in a new way, but bearing in mind their original arrangement. The so-called 'properties' of a system express the possibility of constructing such new systems having a certain connection with the given system. And it is exactly this *embedding* of new systems *in a given system* that plays an important part in building up mathematics, often in the form of an inquiry into the possibility or impossibility of an embedding satisfying certain conditions, and in the case of possibility into the ways in which it is possible."

problems consist in proving a conjecture. Let us take, for instance, his first problem, which asks to prove the Continuum Hypothesis (CH): *Every infinite subset of $\mathbb{R}$ is of the same cardinality as $\mathbb{R}$ or $\mathbb{N}$.*

Is it the case that CH is either true or false? Yes: this is a theorem of ZFC — a direct consequence of the principle of excluded middle.

So, is CH true? Not really: CH cannot be proved in ZFC, as shown by P. Cohen.

So, is CH false? Not really: CH cannot be disproved in ZFC, as shown by Gödel.

This is why in modern classical logic, *truth* is considered to be a semantic notion, depending on the choice of a two-valued model of the theory (cf. §**3.1**; *two-valued* means that every formula without parameters is either true or false). For instance, CH is true in Gödel's constructible universe, and false in Cohen's models of ZFC.

In contrast, the following problem (!CH): *Prove CH in ZFC* — has no solutions (by Cohen); and the problem (!¬CH): *Disprove CH in ZFC* — also has no solutions (by Gödel). And we do not need to fix a model of ZFC for this to make sense. Thus we should not equate the problem !CH with any question pertaining to *truth*. Then, if we want to be precise, Hilbert's first problem should indeed be seen as a *problem* — arguably, best formalized as !CH — and not a yes/no *question*.[12] But had Hilbert used a *whether* question in his first problem, it could as well be formalized as the problem !CH ∨ !¬CH, i.e., *Prove or disprove CH in ZFC*. Note, incidentally, that insolubility of the latter problem (which follows from Gödel *and* Cohen) does not yield a counterexample to Hilbert's principle of decidability of all mathematical problems. Indeed, !CH is decidable (by Cohen), !¬CH is decidable (by Gödel), and even !CH ∨ !¬CH is decidable (by Gödel and Cohen). But the decidability of these problems is not, of course, merely a consequence of the principle of excluded middle.

In general, a problem — as understood in the present paper — should not be conflated with any yes/no question. In particular, the problem *Prove P*, where P is a proposition, is closely related to two distinct questions: *Is P true?* and *Does there exist a proof of P?* Truth should not be conflated with provability, since by Gödel's incompleteness theorem most theories are not complete with respect to any model. On the other hand, if G and H are groups, the question *Is G isomorphic to H?* is closely related to two distinct problems: *Prove that G is isomorphic to H* and *Find an isomorphism between G and H*. Proofs should not be conflated with constructions of isomorphisms: one proof might correspond to several isomorphisms or not correspond to any specific isomorphism.

---

[12]This is even reasonably consistent with Hilbert's original wording [85]: "The investigations of Cantor on such assemblages of points suggest a very plausible theorem, which nevertheless, in spite of the most strenuous efforts, no one has succeeded in proving. This is the theorem: Every system of infinitely many real numbers, i.e., every assemblage of numbers (or points), is either equivalent to the assemblage of natural integers, $1, 2, 3, \ldots$ or to the assemblage of all real numbers". Alongside the Continuum Hypothesis, Hilbert also poses the problem of well-ordering of the continuum: "It appears to me most desirable to obtain a direct proof of this remarkable statement of Cantor's, perhaps by actually giving an arrangement of numbers such that in every partial system a first number can be pointed out."

3.7. The BHK interpretation

The mathematical meaning of problems is determined by answering the question: *What does it mean to find a solution to a given problem?* A standard but informal answer to this question was given essentially in Kolmogorov's 1932 paper [108] and can also be seen as an interpretation of the intuitionistic connectives, and to some extent also quantifiers[13] in terms of solutions of problems. Independently[14] of Kolmogorov, Heyting gave rather similar interpretations for some intuitionistic connectives in terms of proofs of propositions (¬ and ∨ in [80] and → in a 1930 letter to Freudenthal [208], [205] and in print in a 1934 book [81]). Kolmogorov went somewhat further than this in that he described a systematic interpretation of zero-order formulas of intuitionistic logic. But much more importantly, Kolmogorov's paper also offers what in essence is an interpretation of the meta-logical connectives and quantifiers (as discussed in §**3.11** and §*7.2.1* below); to the best of the author's knowledge, this aspect of Kolmogorov's interpretation has been completely ignored in all of the previous literature dealing with the "BHK-interpretation".

A formal answer to the question *What does it mean to find a solution to a given problem?* (or *What does it mean to give a constructive proof to a given proposition?*) was given, in particular, by Gentzen's 1935 natural deduction calculus [65]. (An independent but essentially equivalent answer is given by the λ-calculus, see [64], [93], [119], [34], [188; §4], [202; §10.8].) This answer can of course be read syntactically, on a par with, say, the sequent calculus for classical logic. Yet the introduction rules of the natural deduction calculus can be also read as a semantic interpretation of the intuitionistic connectives and quantifiers, which turns out to be closely related to Kolmogorov's and Heyting's interpretations (see [54], [127; pp. 410–412]).

Initially, Heyting has switched to Kolmogorov's language; in particular, the introduction into intuitionistic logic found in his 1934 book [81] is based on Kolmogorov's understanding of the connectives and quantifiers (but not the meta-quantifiers, which Heyting ignores in [81]). More influential, however, was his later return to the ideologically loaded language of Brouwer, most notably in his 1956 book [82]. This in turn inspired the modern tradition of the so-called "BHK interpretation", named so by Troelstra after Brouwer, Heyting and Kreisel [197], with "K" later reassigned to Kolmogorov by Troelstra and van Dalen [202]. It is also known as the "standard interpretation", and even the "intended interpretation" of intuitionistic logic (see [199], [188], [190], [64], [206], [171], [211]). Compared to Kolmogorov's original interpretation, the BHK interpretation (i) omits Kolmogorov's interpretation of the meta-connectives and meta-quantifiers; (ii)

---

[13]An interpretation of ∃ is discussed in a slightly different language in §2 of Kolmogorov's paper.

[14]A note inserted at proof-reading in Kolmogorov's paper [108] acknowledged the connection with Heyting's [80]. A remark found in the body of [108]: "This explanation of the meaning of the sign ⊢ is very different from that by Heyting, although it leads to the same formal calculus" must refer to Heyting's syntactic definition of ⊢ and indicates that at the time of writing, Kolmogorov was unaware of Heyting's semantic interpretation of ⊢ in [80], which is rather similar to Kolmogorov's.

treats the quantifiers in a more systematic way; (iii) gives a more detailed interpretation of $\to$ due to Heyting (which is similar to Kolmogorov's interpretation of $\forall$).

Most presentations of the BHK interpretation in the literature are in Brouwer's language, but here is one in the language of Kolmogorov.

- What are solutions of primitive problems is assumed to be known from context.

For instance, Euclid's first three postulates are the following primitive problems:[15]

(1) draw a straight line segment from a given point to a given point;
(2) extend any given straight line segment continuously to a longer one;
(3) draw a circle with a given center and a given radius.

We may thus stipulate that each of the primitive problems (1) and (3) has a unique solution, and give some description of the possible solutions of (2).

- A solution of $\Gamma \wedge \Delta$ consists of a solution of Γ and a solution of Δ.
- A solution of $\Gamma \vee \Delta$ consists of an explicit choice between Γ and Δ along with a solution of the chosen problem.
- A solution of $\Gamma \to \Delta$ is a *reduction* of Δ to Γ; that is, a general method of solving Δ on the basis of any given solution of Γ.
- The *absurdity* $\times$ has no solutions; $\neg\Gamma$ is an abbreviation for $\Gamma \to \times$, and the *triviality* $\checkmark$ is an abbreviation for $\neg\times$.[16]
- A solution of $\exists x\,\Theta(x)$ is a solution of $\Theta(x_0)$ for some explicitly chosen $x_0 \in \mathscr{D}$.
- A solution of $\forall x\,\Theta(x)$ is a general method of solving $\Theta(x_0)$ for all $x_0 \in \mathscr{D}$.

Kolmogorov's notion of a "general method" roughly corresponds to the notion of a "construction" advocated by Brouwer and Heyting, but is perhaps less rhetorical in that it puts a central emphasis on the tangible issue of the right order of quantifiers. According to Kolmogorov, if $\Gamma(\mathscr{X})$ is a problem depending on the parameter $\mathscr{X}$ "of any sort", then "to present a general method of solving $\Gamma(\mathscr{X})$ for every particular value of $\mathscr{X}$" should be understood as "to be able to solve $\Gamma(\mathscr{X}_0)$ for every given specific value of $\mathscr{X}_0$ of the variable $\mathscr{X}$ by a finite sequence of steps, known in advance (i.e. before the choice of $\mathscr{X}_0$)". This does not pretend to be a fully unambiguous definition, but "hopefully cannot lead to confusion in specific areas of mathematics".

An important clarification to the BHK interpretation, emphasized by Kreisel (in somewhat different terms; see details and further discussion in [39] and [17; §5.1]), is that a solution of a problem must not only consist of a construction of a prescribed type, but also include a proof that the proposed construction does solve the problem. In fact, this stipulation is clearly expressed already in Proclus' ancient commentary on Euclid's

---

[15]See [3; §??, §??] for a thorough discussion of postulates vs. axioms.

[16]The notation $\times$, $\checkmark$ is non-standard; the usual one is $\bot$, $\top$ — same as in classical logic. We will eventually need to differentiate notationally between classical and intuitionistic logics, and to this end we use $\top$, $\bot$ and Roman letters in classical contexts; and $\checkmark$, $\times$ and Greek letters in intuitionistic contexts. This suffices to differentiate classical formulas from intuitionistic ones, and so we can afford to keep the other connectives, and both quantifiers, identical in appearance in classical and intuitionistic contexts.

Elements (see [3]), and so may be called the *Proclus–Kreisel principle.* If we parse the clauses of the BHK interpretation, assuming that solutions of primitive problems come with correctness proofs, we see that these easily yield correctness proofs for solutions of problems in the $\wedge$, $\vee$, and $\exists$ clauses; thus in the context of the BHK interpretation, the Proclus–Kreisel principle essentially amounts to the convention that every general method (such as those in the $\rightarrow$ and $\forall$ clauses) is supplied with a proof that it actually does what it is meant to do.[17]

Finally, two more precautions on interpreting the BHK interpretation:

- The domain $\mathscr{D}$ must be a "simple domain" such as the set $\mathbb{N}$ of natural numbers; a subset of $\mathbb{N}$ defined with aid of quantifiers would require a more elaborate version of the BHK interpretation (see [39]). In fact, Martin-Löf's type theory [126] is based on extending the BHK interpretation to more general domains.
- The status of all primitive problems (i.e. if they are considered to have a solution) is supposed to be fixed before the logical connectives can be explained, and cannot be updated as we determine the meaning of the connectives (Prawitz; see [173]).

3.8. Understanding the connectives

3.8.1. $\vee$ and $\exists$. The seemingly innocuous interpretations of $\vee$ and $\exists$ are already in sharp contrast with classical logic. Let us consider the following classical proof that x^y can be rational for irrational x and y: if $\sqrt{2}^{\sqrt{2}}$ is rational, then $x = y = \sqrt{2}$ will do; else let $x = \sqrt{2}^{\sqrt{2}}$ and $y = \sqrt{2}$. Although we have proved that the proposition $P(x,y) =$ "x^y *is rational*" holds either for $(x,y) = (\sqrt{2}, \sqrt{2})$ or for $(x,y) = (\sqrt{2}^{\sqrt{2}}, \sqrt{2})$, this method is not going to prove it for any specific pair. Thus if $!Q$ denotes the problem *Prove the proposition Q,* then we have solved the problem $!\big(P(\sqrt{2}, \sqrt{2}) \vee P(\sqrt{2}^{\sqrt{2}}, \sqrt{2})\big)$, and hence also the problem $!\exists x \exists y\, P(x,y)$ (with x and y ranging over all irrational numbers), but we have no clue about solving either $\exists x \exists y\, !P(x,y)$ or $!P(\sqrt{2}, \sqrt{2}) \vee !P(\sqrt{2}^{\sqrt{2}}, \sqrt{2})$.

On the other hand, with classical meta-logic we may wonder if for a given particular problem Γ there exists a solution of Γ. According to the BHK interpretation (as presented above), the set of solutions of a problem $\Gamma \vee \Delta$ is the disjoint union of those of Γ and Δ. The disjoint union of two sets is nonempty if and only if at least one of them is nonempty. Thus we conclude that $\Gamma \vee \Delta$ is soluble if and only if either Γ is soluble or Δ is soluble. In symbols, if $?\Gamma$ denotes the proposition *"The problem Γ is soluble"*, then $?(\Gamma \vee \Delta)$ is equivalent to $?\Gamma \vee ?\Delta$; for similar reasons, the proposition $?\exists x\, \Gamma(x)$ is equivalent to $\exists x\, ?\Gamma(x)$.

Let us note that the explicit choice between Γ and Δ that is present in any solution of $\Gamma \vee \Delta$ is forgotten (and cannot be recovered in general) when only the existence of

---

[17]A well-known alternative to the Proclus–Kreisel principle is to deal with bare constructions, and defer their correctness verifications until the solubility of the entire problem (built out of primitive problems with the logical connectives) is asserted. This approach necessitates a clear distinction of "propositions" (or problems) vs. "judgements" and is used in Martin-Löf's type theory (see [17; §5.1.2]). In fact, according to [71; p. 36], this alternative approach is implicit already in Frege's work.

such a solution is asserted. For instance, the proposition $?\bigl(!P(\sqrt{2},\sqrt{2}) \vee !P(\sqrt{2}^{\sqrt{2}},\sqrt{2})\bigr)$ might follow from a proof-theoretic result asserting that rationality of real numbers of a certain form can always be either proved or disproved (in ZFC, say); this would still give us no clue which of the two numbers is rational (unless the proof-theoretic result is proved in a constructive way).[18]

3.8.2. $\forall$ and $\wedge$. The significance of the Proclus–Kreisel principle is clear from the following *Schwichtenberg's paradox.* Let us consider the problem $!P(x,y,z,n)$ of proving the following proposition $P(x,y,z,n)$:

$$x^n + y^n = z^n \quad \to \quad n \leq 2.$$

Of course, it is trivial to devise a general method m to verify the inequality $x^n + y^n \neq z^n$ and thus to solve the problem $!P(x,y,z,n)$ for every particular choice of positive integers x,y,z and $n > 2$ for which it can be solved. What is hard is to prove that m actually succeeds on all inputs. Thus it is only due to Kreisel's clarification that the problem $\forall x,y,z,n\,!P(x,y,z,n)$ (with x,y,z,n ranging over all positive integers) is nontrivial. Yet this problem is easily seen to be equivalent to the problem $!\forall x,y,z,n\,P(x,y,z,n)$ of proving Fermat's last theorem. Indeed, any solution of the problem $\forall x\,!Q(x)$ is clearly a proof of the proposition $\forall x\,Q(x)$; conversely, any proof of the latter proposition yields by specialization a proof of every its instance $Q(x)$, and this specialization procedure is a general method with respect to the input x, whose success is guaranteed by the original proof of $\forall x\,Q(x)$. Let us note that, due to the existence of the general method m, the problem of proving Fermat's last theorem is also equivalent to the problem $!\neg?!\neg\forall x\,,y,z,n\,P(x,y,z,n)$ of proving that Fermat's last theorem cannot be disproved.[19]

If there exists a general method of solving a problem $\Gamma(x)$ for every x, then, trivially, $\Gamma(x)$ is soluble for each x; in symbols, $?\forall x\,\Gamma(x) \to \forall x\,?\Gamma(x)$. This implication cannot be reversed in general. For example, if $\Gamma(f)$ asks to find a complex root of an algebraic equation $f(x) = 0$ of degree > 0, then a solution of $\Gamma(f)$ exists by the fundamental theorem of algebra; yet a general method of solving all such problems $\Gamma(f)$ could be interpreted, for instance, as a formula providing a solution in radicals. Similarly, if $\Gamma(f)$ is an initial value problem of the form $y' = f(x,y)$, $y(x_0) = y_0$, $f \in C^1$, then a solution of $\Gamma(f)$ exists in some neighborhood of x_0 by the well-known theorem; yet a general method of solving all such problems could be interpreted, for instance, as a formula providing a solution in terms of elementary functions.

The intuitionistic conjunction per se does not behave very differently compared to the classical conjunction. In particular, one can see that the proposition $?(\Gamma \wedge \Delta)$ is

---

[18]In fact, irrationality of $\sqrt{2}^{\sqrt{2}}$ follows from the Gelfond–Schneider theorem (which answered a part of Hilbert's 7th problem): a^b is transcendental if a and b are algebraic, $a \neq 0, 1$, and b is not rational.

[19]A similar remark about Goldbach's conjecture is found in Kolmogorov's letter to Heyting (see [1; §6.2.2]), which indicates that he did implicitly assume general methods to include verification of their own correctness and seems to have been aware of the significance of this provision. In his paper [108], he also emphasized universal acceptability of the validity of solutions as an inherent property of logical and mathematical problems.

equivalent to $?\Gamma \wedge ?\Delta$ (similarly to the above discussion of $\vee$) and that the problem $!(P \wedge Q)$ is equivalent to $!P \wedge !Q$ (similarly to the above discussion of $\forall$).

3.8.3. $\rightarrow$. By the *modus ponens* rule

$$\frac{P,\ P \rightarrow Q}{Q}$$

any proof of the proposition $P \rightarrow Q$ yields a general method of converting proofs of P into proofs of Q, along with a proof that this method indeed effects such conversion. In other words, any solution of the problem $!(P \rightarrow Q)$ yields a solution of $!P \rightarrow !Q$. But not conversely. For instance, a general method of converting ZFC-proofs of the Continuum Hypothesis CH into ZFC-proofs of $0 = 1$ (where $0 = \varnothing$ and $1 = \{\varnothing\}$, as usual) is given by simply keeping the given ZFC-proof; and a verification of the correctness of this method is given by Cohen's proof that ZFC-proofs of CH do not exist. Yet we cannot possibly obtain from this a ZFC-proof that CH implies $0 = 1$, since no such ZFC-proofs exist, as proved by Gödel. Of course, what this argument really shows is that the problems $!\neg P$ and $\neg !P$ are inequivalent for any independent statement P.[20]

By the $\rightarrow$ clause of the BHK interpretation, or by the intuitionistic *modus ponens* rule

$$\frac{\Gamma,\ \Gamma \rightarrow \Delta}{\Delta}$$

if there exists a reduction of the problem Δ to the problem Γ, then the solubility of Γ implies the solubility of Δ; in symbols, $?(\Gamma \rightarrow \Delta) \rightarrow (?\Gamma \rightarrow ?\Delta)$. The converse implication does not hold in general. For example, let Γ be the problem $!?\Delta$ of proving the solubility of Δ. Clearly, if Γ is soluble, i.e., there exists a proof of the solubility of Δ, then Δ is soluble. On the other hand, a proof of the solubility of Δ does not necessarily yield an actual solution of Δ.[21]

---

[20] Some issues need to be clarified here. For the sake of definiteness, we may demand that the verifications of correctness of general methods take place in some theory T. If all atomic problems request to prove something in a certain theory, a natural choice for T would be that theory. It might seem that identifying $\times$ with $!0 = 1$ would then get us in trouble with the second incompleteness theorem, which says that T itself does not prove $\mathrm{Bew}(\#0 = 1) \rightarrow 0 = 1$, where Bew is a formula expressing Gödel's provability predicate for T (see [185]). However, we are dealing here with solutions rather solubility, but the incompleteness theorems arise only because of the existential quantifier. Namely, let us recall that $\mathrm{Bew}(\#P)$ is defined as $\exists n\, \mathrm{Prov}(n, \#P)$, where $\mathrm{Prov}(n, m)$ is a formula expressing the predicate that n is the Gödel number of a proof of the formula with Gödel number m. T proves $\mathrm{Prov}(n, \#P)$ for some numeral n if and only if T proves P; and T proves $\neg\,\mathrm{Prov}(n, \#P)$ for every numeral n if and only if T does not prove P (see [185; 3.2.4]). In particular, T proves that $\mathrm{Prov}(n, \#0 = 1) \rightarrow 0 = 1$. Thus from a proof of $0 = 1$ in T we can certainly get a contradiction. Also, returning to our example (where $T = $ ZFC), since ZFC does not prove either CH or $\neg$CH by Gödel and Cohen, ZFC proves $\neg\,\mathrm{Prov}(n, \mathrm{CH})$ and $\neg\,\mathrm{Prov}(n, \neg\mathrm{CH})$.

[21] According to a certain formal logic QHC of ?'s and !'s [1], for this to be possible, $?\Delta$ must be independent from the formal theory being considered [2; Remark 3.10]. QHC also places further limitations on potential examples of the above type. Although the implication $\neg ?\Gamma \rightarrow ?\neg\Gamma$ does not hold in general, the rule $\neg ?\Gamma\ /\ ?\neg\Gamma$ is valid; thus one cannot prove $\neg ?\Gamma$ without also proving $?\neg\Gamma$ [1; 3.12]. Also, there are the rule $?!P \rightarrow ?!Q\ /\ ?(!P \rightarrow !Q)$ and the implication $(!?\Gamma \rightarrow !?\Delta) \rightarrow !(?\Gamma \rightarrow ?\Delta)$ [1; 3.16].

3.8.4. Absurdity. Let us now discuss the meaning of the negation, $\neg\Gamma$. Using ! and ? as before, and writing $\neg$ also for the classical negation of propositions, we get the problem $!\neg?\Gamma$, which reads: *Prove that Γ has no solutions.* (As before, this refers to classical proofs, not constructive ones.) We will now argue that this problem is equivalent to $\neg\Gamma$ on the BHK interpretation (with our classical meta-logic), by describing solutions of the problems $\neg\Gamma \to !\neg?\Gamma$ and $!\neg?\Gamma \to \neg\Gamma$.

Indeed, suppose that we have a solution of $\neg\Gamma$, that is, a general method m of obtaining a solution of $\times$ based on any given solution of Γ, including a proof that this method works. But since $\times$ has no solutions, we get the following proof p_m that Γ has no solutions: "If s is a solution of Γ, then m applied to s yields a non-existent object, which is a contradiction; thus Γ has no solutions." Moreover, it is clear that the procedure associating p_m to m is a general method, which works regardless of the nature of m.

Conversely, let p be a proof that Γ has no solutions. If s is a solution of Γ, then p yields a proof that s is *itself* also a solution of $\times$; indeed, p proves that s does not exist, but everything is true of a non-existent object. Thus, given a p, we get a general method m_p of obtaining a solution of $\times$ on the basis of any given solution s of Γ; namely, m_p returns s itself and includes a verification that s is a solution of $\times$ (which is given by p). Thus m_p is a solution of $\neg\Gamma$. Moreover, it is clear that the procedure associating m_p to p is a general method, which works regardless of the nature of p.

Curiously, this analysis is at odds with Kolmogorov's remark [108]:

> "Let us note that $\neg a$ should not be understood as the problem 'prove the unsolvability of a'. In general, if one considers 'unsolvability of a' as a well-defined notion, then one only obtains the theorem that $\neg a$ implies the unsolvability of a, but not vice versa. If, for example, it were proved that the well-ordering of the continuum surpasses our capabilities, one could still not claim that the existence of such a well-ordering leads to a contradiction."

Could one make any sense out of the last sentence, mathematically? It could be referring, for instance, to the fact that the existence of a well-ordering of the continuum might be (i) not provable in some formal theory, say ZF, and yet (ii) consistent with this theory. (For example, this is how Kolmogorov's words are interpreted by Coquand [34].)

To interpret (i) as the unsolvability of some problem Γ_1, this Γ_1 could be the problem of deriving the existence of the well-ordering of the continuum (i.e., a special case of the axiom of choice) from the axioms of ZF; a solution of Γ_1 would be such a derivation, and the proof of the non-existence of such derivation would presumably be either in ZF or in some stronger theory. But in this case, $\neg\Gamma_1$ was solved by Cohen: given such a derivation, he does get a contradiction (by a method whose correctness is verified in ZF; cf. footnote 20 above).

And to interpret (ii) as the unsolvability of $\neg\Gamma_2$ for some problem Γ_2, we could understand it syntactically or semantically. On the syntactic reading of (ii), $\neg\Gamma_2$ could be the problem of deriving a contradiction from the axioms of ZF and the hypothesis of the well-orderability of the continuum; but this does not seem to be of the form $\neg\Gamma_2$ for any problem Γ_2. On the semantic reading of (ii), Γ_2 could be the problem of constructing

a model of ZF and a well-ordering of the continuum in this model; indeed, the unsolvability of this $\neg\Gamma_2$ amounts to the (non-constructive) existence of a model of ZF where the continuum can be well-ordered (compare [1; §5.4]). But in this case, Γ_2 was solved by Gödel, who constructed a certain model of ZF (Gödel's constructible universe), and a well-ordering of the continuum in that model.

Thus it is possible that Kolmogorov might have simply conflated Γ_1 and Γ_2, thinking of a single informal problem, "Find a well-ordering of the continuum".

On the other hand, Kolmogorov's wording ("$\neg a$ implies the *unsolvability* of a"; "the *existence* of such a well-ordering leads to a contradiction") certainly conflates implications between problems with implications between propositions; this makes us suspect that in writing of what must have been the former, he might have actually been thinking of the latter (in this particular remark about the well-ordering of the continuum). Indeed, if P is the proposition *The continuum is well-orderable*, then the *unsolvability* of the problem of proving it, or $\neg?!P$ in the above notation, certainly does not imply $\neg P$, the proposition that the *existence* of a well-ordering of the continuum leads to a contradiction. In this connection, let us also not forget Kolmogorov's initial rejection of the explosion principle $\times \to \alpha$ [107], [208].

It should be noted that Heyting, while using Kolmogorov's terminology (and so being aware of Kolmogorov's remark quoted above) has explicitly identified $\neg\Gamma$ with $!\neg?\Gamma$ [81]:

> "It is appropriate to interpret the notion of 'reduction' in such a way that the proof of the impossibility of solving a at the same time reduces the solution of any problem whatsoever to that of a."

3.9. Something is missing here

Troelstra and van Dalen note of the BHK interpretation [202; p. 9 and p. 33]:

> "This explanation is quite informal and rests itself on our understanding of the notion of construction and, implicitly, the notion of mapping; it is not hard to show that, on a very 'classical' interpretation of construction and mapping, H1–6 [the six clauses of the BHK interpretation] justify the principles of two-valued (classical) logic."
>
> "This exercise [[202; Exercise 1.3.4]] shows that the BHK-interpretation in itself has no 'explanatory power': the possibility of recognizing a classically valid logical schema as being constructively unacceptable depends entirely on our interpretation of 'construction', 'function', 'operation'."

Indeed (see also [175]), let $\{\!\!\{\Gamma\}\!\!\}$ denote the set of solutions of the problem Γ. Then the BHK interpretation guarantees that:

- $\{\!\!\{\Gamma \wedge \Delta\}\!\!\}$ is the product $\{\!\!\{\Gamma\}\!\!\} \times \{\!\!\{\Delta\}\!\!\}$;
- $\{\!\!\{\Gamma \vee \Delta\}\!\!\}$ is the disjoint union $\{\!\!\{\Gamma\}\!\!\} \sqcup \{\!\!\{\Delta\}\!\!\}$;
- there is a map $\mathscr{F}\colon \{\!\!\{\Gamma \to \Delta\}\!\!\} \to \mathrm{Hom}(\{\!\!\{\Gamma\}\!\!\}, \{\!\!\{\Delta\}\!\!\})$ into the set of all maps;
- $\{\!\!\{\times\}\!\!\} = \varnothing$;
- $\{\!\!\{\exists x\, \Gamma(x)\}\!\!\}$ is the disjoint union $\bigsqcup_{d \in \mathscr{D}} \{\!\!\{\Gamma(d)\}\!\!\}$;

- there is a map $\mathscr{G} \colon \{\!\}\forall x\,\Gamma(x)\{\!\} \to \prod_{d\in\mathscr{D}} \{\!\}\Gamma(d)\{\!\}$ into the product.

If we force $\mathscr{F}$ to be the identity map, we obtain what might be called the "classical BHK". Indeed, $\{\!\}\Gamma \vee \neg\Gamma\{\!\} = \{\!\}\Gamma\{\!\} \sqcup \mathrm{Hom}(\{\!\}\Gamma\{\!\}, \varnothing)$ is never empty; thus $\Gamma \vee \neg\Gamma$ has a solution for each problem Γ. Note that we get the same result as long as $\mathscr{F}$ is surjective.

Thus the six clauses of the BHK interpretation, as presented either above or in Troelstra–van Dalen [202], fail to capture the essence of intuitionistic logic. The way to deal with this issue in the intuitionistic tradition has been, unfortunately, just to sweep it under the carpet.[22] Thus, elsewhere in the same book by Troelstra and van Dalen [202; p. 10], we find the assertion

> "Even if the explanations H1-6 [the six clauses of the BHK interpretation] leave a lot of questions open, they suffice to show that certain logical principles should be generally acceptable from a constructive point of view, while some other principles from classical logic are not acceptable."

This is followed by an analysis of "the principle of excluded middle" (i.e. the principle of decidability in our terminology), which ends with a striking conclusion [202; p. 11]:

> "Thus we cannot accept PEM [the Principle of Excluded Middle] as a universally valid principle on the BHK-interpretation."

Don't the latter two quotations directly contradict the previous two quotations from the same book, in which the same BHK interpretation was found to "justify the principles of two-valued (classical) logic"?

The only sign of hope for understanding what is going on here is the clause *"as a universally valid principle"*, whose meaning is not really explained in Troelstra–van Dalen [202]. We are told, however, that

> "Constructively, accepting PEM [the Principle of Excluded Middle] as a general principle means that we have a universal method for obtaining, for any proposition A, [a proof of $A \vee \neg A$, that is,] either a proof of A or a proof of $\neg A$"

This is parallel to Heyting's original assertion [80; p. 59]:

> "a proof that the law of excluded middle is a general law must consist in giving a method by which, when given an arbitrary proposition, one could always prove either the proposition itself or it negation."

Yet these demands for a *"method"* appear out of the blue both in Heyting [80] and in Troelstra–van Dalen [202], without any explanations. Of course, what they could have said to settle the issue is that (i) it is precisely their demands for a "method" that are lacking to give "explanatory power" to the six clauses of the BHK-interpretation; (ii) thus these demands should really be included in an additional clause of the BHK-interpretation, in order to guarantee that it does explain intuitionistic rather than classical logic.

---

[22]And it is arguably this particular sweeping under the carpet that is largely responsible for the fact that most mathematicians are aware of intuitionistic logic, yet fail to understand it.

In fact, Kolmogorov did explicitly incorporate a similar demand (for a "method") in his original version of the BHK interpretation [108]. But before we are ready to formulate and discuss it in a clear language, we will need to understand what could have prevented Troelstra and van Dalen, as well as all other writers in the intuitionistic tradition (to the best of the author's knowledge) to add such a stipulation to the BHK interpretation.

One factor, which could be deciding in the case of Heyting's early work, was the attitude that, in view of Heyting's derivation system, intuitionistic logic needs to be explained philosophically, rather than mathematically. It is obviously due to this attitude that Heyting did not attempt a systematic contentual interpretation of intuitionistic logic prior to Kolmogorov's paper, even though he had clearly gathered the necessary bits and pieces. But this attitude also meant that one does not even aim at delineating exactly what is included in the interpretation.

Another, much more serious factor was the widespread acceptance of first-order logic with its usual model theory and a particular tradition of trying to express notions that are not expressible in it (including the "principle of excluded middle") as the default language in foundations of mathematics. Like any language (natural or artificial), it comes with a plethora of implicit assumptions that on the one hand facilitate communication and thought, but on the other hand subjugate them to invisible limitations. We will now proceed to review and overcome some of these limitations.

3.10. Classical logic revisited

A detailed treatment of first-order formal logic will be given in the next chapter (§4). Here we include a quick glimpse of it, just enough to proceed with clarification of the BHK interpretation in the next section (§**3.11**).

3.10.1. Syntax. Formal classical logic deals with formulas, such as

$$p \vee \neg q$$

and

$$q \vee \neg p.$$

Here p and q are *propositional variables* and $\vee$, $\neg$ are *(formal) connectives*; all of these are some of the ingredients of the formal language of classical logic. In classical logic, one cannot (formally) derive any of these two formulas from the other one (as we will discuss shortly). Similarly, none of the formulas

$$p \vee \neg p$$

and

$$q \vee \neg q$$

is derivable from each other without using the law of excluded middle.

In modern treatments of first order classical logic, the law of excluded middle is usually expressed by the *schema*

$$f \vee \neg f$$

in which $\vee$ and $\neg$ have the same meaning as before, but f is a *metavariable* or schematic variable (not to be confused with the propositional variable f), which is absent from the formal language of classical logic, but is understood to denote a formula. Formulas like $\mathsf{p} \vee \neg\mathsf{p}$, $\mathsf{q} \vee \neg\mathsf{q}$ and $(\mathsf{p} \vee \mathsf{q}) \vee \neg(\mathsf{p} \vee \mathsf{q})$, obtained by replacing each meta-variable with some formula are *instances* of this schema, just like e.g. $3^2 - 2^2 = (3 + 2)(3 - 2)$ is an instance of the identity $x^2 - y^2 = (x + y)(x - y)$.

The schema itself is sometimes identified with the set of its instances. But in any case it does not belong to the formal language — because infinite sets of formulas, like metavariables, exist only in meta-language and not within the formal language. (A finite set of formulas can often be identified with their conjunction, which is a single formula, but infinite conjunctions do not exist in traditional classical logic.) In particular, when the ingredients of the formal language are interpreted in some way, we get interpretations of formulas, but no interpretations of schemata.

Thus, even though the law of excluded middle is constantly used in classical logic, it is absent as an entity from both syntax and semantics in the standard modern formalism (like all other laws and inference rules that define classical logic). One could argue that "only specific instances of the law of excluded middle are really used", but this is only a part of the story, since to recognize some formula as an instance of the schema one does refer to the actual schema.

This recognition step can actually be bypassed. There is a one-to-one correspondence between formulas and schemata given by replacing predicate variables by metavariables and conversely (at least if they are drawn from the same alphabet, meaning either a finite human alphabet, for practical purposes, or a countably infinite mathematical alphabet, for theoretical ones). Then instead of instances of a schema one can speak of *substitution instances* of the corresponding formula (which are the same set of formulas).

Moreover, classical logic (and all other logics that we will consider) satisfies the substitution theorem: if some formula is derivable in classical logic, then also all its substitution instances are derivable. Because of this, it suffices for practical purposes to speak only of derivability of formulas, and so often there is no practical need for a separate notion of "derivability of schemata"; for example, all instances of the schema $\Big(((f \to g) \to f) \to f\Big)$ are derivable if and only if the formula $\Big(((\mathsf{f} \to \mathsf{g}) \to \mathsf{f}) \to \mathsf{f}\Big)$ is derivable. (They are indeed derivable in classical logic, but this does not concern us here.) But as one goes deeper into the modern formalism of first-order logic there appear more and more of such surrogates, i.e. writing one thing while actually thinking

about another one.[23] So it can be argued that to sacrifice the notion of "derivability of schemata" is to step on a slippery slope from a conceptual viewpoint.

Remark 3.2. Let us pause here to note that the substitution theorem does not say that each substitution instance F' of a formula F is derivable from F in classical logic (and it is generally not derivable!). It only says that the *judgement* of derivability of F in classical logic, denoted $\vdash F$, implies the judgement $\vdash F'$. The difference is that both these judgements belong to the meta-theory, and so the implication between them is "contentual" rather than formal (in other words, one is derived from the other not in formal classical logic, but in textbooks on formal classical logic).

It not hard to formally include a device that amounts to the use schemata into the formal language. Let us add a new symbol "·" to the formal language, and write it in front of a formula whenever we want to "read it as a schema"; such expressions of the extended formal language, consisting of a formula prefixed by "·", will be called *principles*. (Thus the notion of a formula is in effect extended to include principles; however, to avoid confusion, we will keep the old meaning for the word "formula", using it to refer to formulas of the original language, without "·".) That is, instead of the schema

$$f \vee \neg f$$

(which does not belong to the formal language) or rather the infinite conjunction of its instances, which could be written symbolically as

$$\forall f \ f \vee \neg f$$

(but anyway does not belong to the formal language) we will write the principle

$$\cdot f \vee \neg f$$

(which now belongs to the formal language). Of course, the principle

$$\cdot p \vee \neg p$$

---

[23]Suffice it to say that in intuitionistic logic, the formula $\varphi \vee \neg\varphi$ is not a consequence of the formula $\neg\neg\varphi \rightarrow \varphi$; but nevertheless people say and write, and for a good reason, that "the principle of excluded middle", which they express by the schema $\varphi \vee \neg\varphi$, is a consequence of "the principle of double negation", which express by the schema $\neg\neg\varphi \rightarrow \varphi$ (see Example 4.36 and Remark 4.37 below for the details). And people obviously do think in these terms of "principles", which are made illegitimate by the standard formalism. Worse yet, the modern tradition of dealing with schemata, according to which "principles" are expressed by schemata, is in itself a source of confusion. The relation between the meta-language and the language is akin to the relation between syntax and semantics. Just like the arithmetical formula $x + y = y + x$ denotes any of its interpretations $0 + 1 = 1 + 0$, $2 + 3 = 3 + 2$, etc. (see also §4.1.4 concerning the semantics of non-closed formulas), the schema $f \vee \neg f$ denotes any of its instances $p \vee \neg p$, $(p \vee q) \vee \neg(p \vee q)$, etc. But just like the formula $\forall x \forall y \ x + y = y + x$ denotes effectively the infinite conjunction $0+1 = 1+1 \wedge 2+3 = 3+2 \wedge \ldots$, it would be more accurate to say that the "principle of excluded middle" is expressed not by the schema $\varphi \vee \neg\varphi$, but rather by the infinite conjunction of all its instances, which might be written symbolically as $\forall \varphi \ \varphi \vee \neg\varphi$ (and similarly for the "principle of double negation").

will then serve the same purpose — but this is not a big deal, as the formula

$$\forall x\, r(x)$$

already serves the same purpose as[24]

$$\forall y\, r(y).$$

Here we encounter some further ingredients of the formal language of classical logic: the *quantifier* $\forall$, the unary *predicate variable* r and the *individual variables* x, y. It should be noted in this connection that instances of the schema $p \vee \neg p$ also include formulas like $\forall x\, r(x) \vee \neg \forall x\, r(x)$ and $r(x) \vee \neg r(x)$; and instantiation and substitution become a bit more complicated to define precisely in the presence of individual variables (this will be discussed in §4).

But there are more substantial questions to worry about:

(1) Can one derive arbitrary instances of a schema from the corresponding principle, without resorting to meta-language?
(2) Can one reasonably formalize derivability of principles?
(3) Does "·" have any natural semantics?

The answers are affirmative, and will be briefly discussed here and in more detail in §4.

Our way to answer the first two questions is along the lines of the well-known proof assistant Isabelle, whose approach is in many ways closer to the early tradition of first-order logic, as e.g. in the textbooks by Hilbert–Ackermann [89] and Hilbert–Bernays [90] than to the modern formalism.

For our present purposes, we can identify the judgement of derivability of a principle $\cdot F$, denoted $\vdash \cdot F$, with the judgement $\vdash F$. A more reasonable definition is in §4.

Now let us turn to the third question.

3.10.2. Formal semantics. As discussed in §**??** above, the elements of the formal language of classical logic (or any other formal language) are themselves devoid of any meaning but can be given meaning in an interpretation. For example, one can interpret

- each propositional variable by a subset of the set $\mathbb{R}$ of real numbers (for example, $|p|$ could consist of all integers, and $|q|$ of all positive reals);
- each individual variable by an element of the set $\mathbb{N}$ of natural numbers (for example, $|x| = 3$, $|y| = 5$);
- each k-ary predicate variable by an $\mathbb{N}^k$-indexed family of subsets of $\mathbb{R}$ (for example, $|r|(n) = (-\infty, n]$); and
- each connective or quantifier by an appropriate set-theoretic operation; in particular, $\vee$ by taking the union of the two given sets, $\neg$ by taking the complement of the given set in $\mathbb{R}$, and $\forall$ by taking the intersection of all members of the given $\mathbb{N}$-indexed family of sets.

---

[24]We will see in §4 that in both cases this is a matter of replacing an expression of λ-calculus with an α-equivalent expression, because "·" is really just an abbreviation for the "meta-quantifier closure" (whatever it means).

Then each formula also gets an interpretation, defined by a straightforward recursion on its subformulas. In our example, $|p \vee \neg q|$ consists of all integers and all nonnegative reals, whereas $|q \vee \neg p|$ consists of all positive reals and all non-integer reals; incidentally, these two sets would have to be equal if the formulas $p \vee \neg q$ and $q \vee \neg p$ were interderivable.

The interpretation of individual variables, which is called an *individual assignment*, is needed to interpret formulas with *free occurrences* of individual variables, such as $r(x) \vee \forall x \big(r(x) \vee r(y)\big)$. Here the first occurrence of x and the only occurrence of y are free, but the second occurrence of x is not free, because it is "bound" by the quantifier. But we obviously do not need an individual assignment to interpret *closed* formulas, that is, those in which no individual variable occurs freely. For example, $|\forall x\, r(x)| = \bigcap_{n \in \mathbb{N}}(-\infty, n] = (-\infty, 0]$.

It is also not hard to see that all instances of the law of excluded middle, including $p \vee \neg p$, $q \vee \neg q$, $\forall x\, r(x) \vee \neg \forall x\, r(x)$ and $r(x) \vee \neg r(x)$ have the same interpretation, $\mathbb{R}$. Moreover, this remains true regardless of exactly how propositional, predicate and individual variables are interpreted.

The latter fact suggests that the principle $\cdot p \vee \neg p$ could also be assigned a definite interpretation: "(semantic) truth", denoted $\mathbb{T}$. In general,

- every principle $\cdot F$ is assigned an interpretation $|\cdot F| \in \{\mathbb{T}, \mathbb{\bot}\}$, which is set to be $\mathbb{T}$ if for every interpretation of propositional and predicate variables that occur in F and individual variables that occur freely in F, the interpretation of F is the entire set $\mathbb{R}$; and otherwise, $\mathbb{\bot}$.

Thus $|\cdot p \vee \neg q| = \mathbb{\bot}$ (which is called "(semantic) falsity") since e.g. $|q \vee \neg p| \neq \mathbb{R}$.

The interpretation of propositional and predicate variables, called a *predicate valuation*, is needed to interpret formulas, but is not needed to interpret principles. By an *interpretation of the language* of classical logic we will mean only an interpretation of connectives and quantifiers (which includes the choices of $\mathbb{N}$ and $\mathbb{R}$). This suffices to interpret principles.

If $|\cdot F| = \mathbb{T}$, we also say that the principle $\cdot F$ is *valid* in our interpretation of the language of classical logic. The judgement that $\cdot F$ is valid in the interpretation being considered is denoted by $\vDash \cdot F$. It is well-known that $\vDash \cdot F$ is equivalent to $\vdash \cdot F$ in the present example (see §*4.8.2*, where this example is discussed in more detail). Validity of formulas with respect to a fixed predicate valuation and/or a fixed individual assignment will be discussed in §4, but it is not needed for our present purposes.

Remark 3.3. An interpretation of the language of a theory T over classical logic (such as Peano Arithmetic, Tarski's elementary geometry or Zermalo–Fraenkel set theory) extends an interpretation of the language of classical logic by interpreting additional ingredients in the language of T: function symbols (such as those for the binary operations of addition and multiplication and for numeric constants) and predicate constants (such as those for the binary predicates of equality and membership and the ternary predicate of betweenness for points on a line).

Connectives, quantifiers, function symbols and predicate constants are constants; whereas individual variables and predicate variables are variables (in the sense of λ-calculus, see §**4.2** below). This distinction is mathematically significant, as predicate constants and predicate variables do behave differently even in classical logic (see examples in §*4.8.1* and §*4.8.2*), and was clear in classic texts such as those of Hilbert–Ackermann [89], Hilbert–Bernays [90], Church [32] and Novikov [149], [151].

Nevertheless the modern tradition of first-order logic effectively conflates predicate variables and predicate constants under the guise of "predicate symbols" (or "predicate letters").[25] Consequently, the distinction between the language of a logic and the language of a theory is also blurred, and hence an interpretation of such language is assumed to include an interpretation of "predicate symbols" in modern treatments of first-order logic — even though modern treatments of propositional (=zero-order) logic usually clearly separate valuation of propositional variables from interpretation of connectives.

3.10.3. Informal semantics. Let us now look at another way of interpreting the ingredients of the language of classical logic:

- each propositional variable is interpreted by a contentful (e.g. mathematical) proposition;
- each individual variable by an element of a set $\mathscr{D}$, called the *domain of discourse*;
- each k-ary predicate variable by a k-ary contentful predicate on $\mathscr{D}$;
- each formal connective or quantifier is interpreted according to the usual truth tables (see §**3.1**), or in other words by the corresponding contentual[26] connective or quantifier ($\vee$ by "or", $\neg$ by "not", $\forall$ by "for all", etc.);
- a principle $\cdot F$ is interpreted by the *meta-proposition* (that is, a judgement about propositions, itself treated as a higher-level proposition) asserting that for every interpretation of propositional and predicate variables that occur in F and individual variables that occur freely in F, the contentful proposition that interprets F is true;[27]

---

[25]They are usually introduced on a par with function symbols, and are indeed treated as constants in the context of theories — in which case the substitution theorem does not apply (unless all axioms are given by schemata). But when dealing with a bare logic and its models, "predicate symbols" can often be substituted with formulas, which one would not normally do to constants; and one often assumes a countable list of "predicate symbols" of each arity, none of them standing for anything specific (as e.g. in the texts of Kleene [103] and Troelstra–van Dalen [202]), which is also a strange way of dealing with constants. But it should be noted that in the case of propositional logic, "propositional symbols" are clearly distinguished from the (syntactic) propositional constants $\top$ and $\bot$; and in logic with equality, the equality is clearly distinguished from "predicate symbols". The difference is that $\top$, $\bot$ and $=$ are not supposed to be substitutable by formulas.

[26]"Contentual" (or rather its more widely used German equivalent *inhaltlich*) is Hilbert's term, the opposite of "formal", which he used to explain the Formalist program.

[27]In more detail, let $\vec{x} = (x_1, \ldots, x_n)$ be the tuple of all individual variables that occur freely in F, let $\vec{p} = (p_1, \ldots, p_m)$ be the tuple of all propositional and predicate variables that occur in F, and let $\vec{r} = (r_1, \ldots, r_m)$ be the tuple of their arities. Then $\cdot F$ is interpreted by the meta-proposition $\forall \vec{P} \, \forall \vec{X} \, (\vec{p}, \vec{x} \mapsto \Phi)(\vec{P}, \vec{X})$ asserting the truth of all the contentful propositions $(\vec{p}, \vec{x} \mapsto F)(\vec{P}, \vec{X})$ obtained from F by substituting $\vec{x}$ by an n-tuple $\vec{X}$ of elements of $\mathscr{D}$ and $\vec{p}$ by an m-tuple $\vec{P}$ of contentful predicates

- $\vDash \cdot F$ asserts that the meta-proposition that interprets $\cdot F$ is true.

This is not really a definite interpretation but rather an informal sketch of what could possibly be made into a definite interpretation. To be sure, this can be done; for instance:

- each propositional variable is interpreted by a closed formula in the language of Peano Arithmetic (PA);
- each individual variable by a closed term of PA (such as $(1+1)*(1+1)+1$);
- each k-ary predicate variable by a formula $F(\mathsf{x}_1, \ldots, \mathsf{x}_k)$ in the language of PA such that $\forall \mathsf{x}_1 \ldots \forall \mathsf{x}_k \, F(\mathsf{x}_1, \ldots, \mathsf{x}_k)$ is a closed formula;
- each formal connective or quantifier by the namesake connective or quantifier in the language of PA, or rather of its underlying (classical) logic;
- a principle $\cdot F$ is interpreted by the formula of the language of PA which expresses, using the Gödel numbering, the assertion that for every interpretation of propositional and predicate variables that occur in F and individual variables that occur freely in F, the formula that interprets F is derivable in PA;
- $\vDash \cdot F$ asserts that the formula that interprets $\cdot F$ is true in the standard model $\mathbb{N}$ of PA.

In this example, if $|\mathsf{p}|$ is a formula expressing that there exist infinitely many twin primes and $|\mathsf{q}|$ is Gödel's formula, "asserting its own non-derivability", then $|\mathsf{p} \vee \neg \mathsf{p}|$ and $|\mathsf{q} \vee \neg \mathsf{q}|$ are different formulas (in contrast to the previous example), though both are derivable in PA. Incidentally, that neither $|\mathsf{q}|$ nor $|\neg \mathsf{q}|$ is derivable in PA is nothing special, because already in the previous example neither $|\mathsf{q}|$ nor $|\neg \mathsf{q}|$ was the entire $\mathbb{R}$. Also, $|\cdot \mathsf{p} \vee \neg \mathsf{p}|$ and $|\cdot \neg \neg \mathsf{p} \to \mathsf{p}|$ are different formulas (in contrast to the previous example), though both are true in $\mathbb{N}$ (and derivable in PA, though this is less obvious and does not matter for the purposes of our interpretation).

Admittedly, this example is silly enough in that it can be said to interpret formal classical logic by means of formal classical logic — except that the new ingredient "$\cdot$" is interpreted in a more meaningful way. Indeed, derivability of formulas in PA is no longer the business of PA and its underlying logic, but rather of the meta-theory of PA and its underlying (meta-)logic — which has been formalized by Gödel within PA and its underlying logic. Thus the interpretation of $\cdot$ takes us from "propositions" to "meta-propositions" (which are realized here as formalized judgements about formalized propositions). But this is not all. Validity of arithmetical formulas in the standard model of PA is also no longer the business of PA. Thus the the definition of validity takes us further from meta-propositions to judgements about meta-propositions (which are this time left at an informal level).

Returning to the informal sketch above, let us emphasize that an informal interpretation like above is in general not necessarily formalizable in a reasonable way; but if it happens to be formalizable, there ought to some flexibility in how it can be formalized. (In our example, we could look at a non-standard model of PA, or at Tarski's plane

---

of arities $r_1, \ldots, r_m$ on $\mathscr{D}$, and then applying the usual truth tables to interpret the connectives and quantifiers.

geometry along with its various models, or at the set theory ZFC along with its various models.) Also, the informal sketch above seems to be adding nonzero value to its particular formalization above, just like reading an informal sketch/idea of proof often turns out to be more helpful than going through the details of a carefully written, but unmotivated formal proof. Because of this, in what follows we will occasionally still resort to informal "contentual" interpretations on a par with the sketch above; but it should always be clear what is supposed to be rigorous and what is not.

Let us note that according to the informal sketch above, the principle of excluded middle,

$$\cdot p \vee \neg p,$$

is interpreted by the meta-proposition asserting that *for each contentful proposition p, the proposition $p \vee \neg p$ is true.* The judgement

$$\vDash \cdot p \vee \neg p$$

asserts that this meta-proposition is true. If we ignore the subtleties of different levels of reflection, which are anyway far from clear in such vague formulations, we can roughly express both the informal interpretation of the principle of excluded middle and the judgement of its validity as

$$\forall P \; P \vee \neg P,$$

where P is understood to run over contentful (e.g. mathematical) propositions.

3.11. Clarified BHK interpretation

Let us now try to do the same with intuitionistic logic and with the principle of decidability,

$$\cdot \gamma \vee \neg \gamma.$$

Here γ is a propositional variable in the formal language of intuitionistic logic, which we denote by a Greek letter in order to avoid confusion with the language of classical logic. In fact, in view of our previous discussions, it makes sense to call γ a *problem variable*, because intuitionistic logic is supposed to be about problems rather than propositions (at least, certainly not about propositions in the usual sense of classical mathematics).

At this point, we are completely ignorant about the laws and inference rules of intuitionistic logic, and know little about its semantics. But at least we could hope that an informal interpretation of the principle of decidability parallel to the above informal interpretation of the principle of excluded middle can be roughly expressed as

$$\forall \Gamma \; \Gamma \vee \neg \Gamma,$$

where Γ is understood to run over contentful (e.g. mathematical) problems. But what does this expression mean? If we read it by analogy with the $\forall$ clause in the BHK interpretation, then it expresses *the (meta-)problem of finding a general method of solving the problem $\Gamma \vee \neg \Gamma$ for every contentful problem Γ.* And this is precisely Kolmogorov's original interpretation of the principle of decidability [108], except that his use of the

symbol "⊢" where we use the symbol "·" raises some issues (which we will discuss in a moment).

If we now turn to the judgement

$$\vDash \cdot\, \gamma \vee \neg\gamma$$

that the principle of decidability is valid in our informal interpretation, the natural way to express it is that *there exists* a solution of the said meta-problem; that is, *there exists a general method of solving the problem* $\Gamma \vee \neg\Gamma$ *for every contentful problem* Γ. Using the notation of §**3.8**, this can be roughly written as

$$?\forall \Gamma\ \Gamma \vee \neg\Gamma.$$

And this is reminiscent of Brouwer's *"Third insight. The identification of the principle of excluded middle with the principle of decidability of every mathematical problem."*[28] [28; §1].

Remark 3.4. Let us now discuss Kolmogorov's original approach [108]. It is conceivable that he did not really mean his semantic definition of ⊢ (which coincides with our semantic interpretation of ·) to serve as an interpretation of Heyting's syntactic definition of ⊢ ("To indicate that a formula is included in the list of 'correct formulas' the sign ⊢ will be put in front of it" [83]). This reading is supported by the fact Kolmogorov separately lists "*rules* of our calculus of problems", which are themselves not problems. In this case our interpretation above can be seen simply as a slight extension of Kolmogorov's.

However, Kolmogorov's words "This explanation of the meaning of the sign ⊢ is quite different from that of Heyting, even though it leads to the same rules of the calculus" [108] suggest that he could have actually meant one to interpret the other. In this case our interpretation above is rather a "correction" of Kolmogorov's, since in our approach the judgement ⊢ ·γ ∨ ¬γ is naturally interpreted by the judgement ⊨ ·γ ∨ ¬γ, and not by the meta-problem that interprets ·γ ∨ ¬γ. It should be noted here that Heyting, in his review of Kolmogorov's problem interpretation [81], already chose to interpret a formula prefixed by the ⊢ symbol as a judgement and not a problem (though without any mention of a general method).

The "correction", if it indeed occurs, may have some philosophical implications. Kolmogorov has claimed that in his calculus of problems, "there is no need for any special, e.g. intuitionistic, epistemological presuppositions" (cf. his quote in §**3.4**). But he arguably did not quite achieve this stated goal, because he appears to implicitly refer to the

---

[28]Brouwer's four "insights" intended to summarize his work in intuitionism; the third is the shortest, and essentially reiterates after 20 years the identification discussed in §**3.6** — except that this "Third insight" taken alone can be seen as a *definition* of the principle of excluded middle. The "principle of decidability of every mathematical problem", originally formulated by Hilbert (see §**3.4**), hardly intended to assert a *general method* for settlement of all mathematical problems. However, it is likely that Brouwer, with his philosophy that all mathematical statements should be read constructively, did actually mean this principle to assert a *general method* — at least in the later paper [28]. See, for instance, footnote 6 in [28], which guarantees this sort of reading for another principle.

same epistemological presuppositions as Heyting.[29] It is also arguable that our choice of the judgement of existence instead of Kolmogorov's implicit or Heyting's explicit judgements of knowledge does provide independence from "epistemological presuppositions", even if gained by the price of what Heyting called "faith in transcendental existence".

On a more practical level, the switch from knowledge to existence corresponds to the use of classical, rather constructive meta-logic (or meta-meta-logic in the setting of §4).

Let us now extend the above to the following informal interpretation of the ingredients of the language of intuitionistic logic:

- each individual variable is interpreted by an element of a set $\mathscr{D}$;
- each k-ary problem variable is interpreted by a contentful (e.g. mathematical) problem with k parameters, each running over $\mathscr{D}$;
- each formal connective or quantifier is interpreted according to the usual BHK interpretation (see §**3.7**);
- a principle $\cdot\,\Phi$ is interpreted by the meta-problem of finding a general method of solving all the contentful problems that result from all possible interpretations of all problem variables that occur in Φ and all individual variables that occur freely in Φ;[30]
- $\vDash \cdot\,\Phi$ asserts that the meta-problem that interprets $\cdot\,\Phi$ has a solution.

To fix a name, we will call this the *clarified BHK* (or briefly cBHK) interpretation of intuitionistic logic.

---

[29]According to Kolmogorov [108], "the formula [$\vdash a \vee \neg a$] reads as follows: to give a general method that for every problem a either finds a solution of a or deduces a contradiction on the basis of such a solution! In particular, when the problem a consists in proving a proposition, one must possess a general method to either prove or reduce to a contradiction any proposition. If the reader does not consider himself to be omniscient, he will probably determine that [$\vdash a \vee \neg a$] cannot be on the list of problems solved by him." Let us note that this effectively appeals to an assumption regarding the reader's knowledge. Elsewhere in [108] Kolmogorov introduces a list of axioms of intuitinistic logic by stating that "we must assume that the solutions of some elementary problems are already known. We take as postulates that we already have solutions to the following groups A and B of problems. The subsequent presentation is directed only to the reader who has solved all these problems". Here Kolmogorov again clearly assumes something regarding the reader's knowledge. In fact, his assumption that solutions of certain problems beginning with $\vdash$ *are already known* is not so different from Heyting's epistemological definition of $\vdash$ [79]: "To satisfy the intuitionistic demands, the assertion must be the observation of an empirical fact, that is, of the realization of the expectation expressed by the proposition p. Here, then, is the *Brouwerian assertion* of p: *It is known how to prove p.* We will denote this by $\vdash p$. The words 'to prove' must be taken in the sense of 'to prove by construction'."

[30]In more detail, let $\vec{x} = (x_1, \ldots, x_n)$ be the tuple of all individual variables that occur freely in Φ, let $\vec{\gamma} = (\gamma_1, \ldots, \gamma_m)$ be the tuple of all problem variables that occur in Φ, and let $\vec{r} = (r_1, \ldots, r_m)$ be the tuple of their arities. Then $\cdot\,\Phi$ is interpreted by the meta-problem $\forall\vec{\Gamma}\,\forall\vec{X}\,(\vec{\gamma}, \vec{x} \mapsto \Phi)(\vec{\Gamma}, \vec{X})$ of finding a general method of solving all the contentful problems $(\vec{\gamma}, \vec{x} \mapsto \Phi)(\vec{\Gamma}, \vec{X})$ obtained from Φ by substituting $\vec{x}$ by an n-tuple $\vec{X}$ of elements of $\mathscr{D}$ and $\vec{\gamma}$ by an m-tuple $\vec{\Gamma} = (\Gamma_1, \ldots, \Gamma_m)$ of contentful problems, where each Γ_i has r_i parameters running over $\mathscr{D}$, and then applying the usual BHK interpretation to interpret the connectives and quantifiers.

3.12. The principle of decidability and its relatives

So is the principle of decidability valid on the clarified BHK interpretation? That is, does there exist a general method[31] of solving the problem $\Gamma \vee \neg\Gamma$ for every contentful problem Γ?

If G_p is a group given by a finite presentation p (i.e. a finite list of generators and relations), let Γ_p be the problem, *Find an isomorphism between G_p and the trivial group 1*. Then from any solution of Γ_p we can extract a proof that $G_p \simeq 1$ (by Kreisel's clarification), whereas from any solution of $\neg\Gamma_p$ we can extract a proof that $G_p \not\simeq 1$ (by our solution of $\neg\Gamma_p \to !\neg?\Gamma_p$). Hence any solution of $\Gamma_p \vee \neg\Gamma_p$ would tell us, in particular, whether G_p is isomorphic to 1 or not (since any solution of $\Gamma \vee \Delta$ involves, in the first place, an explicit choice between Γ and Δ).

Now if we have a general method of solving all problems of the form $\Gamma \vee \neg\Gamma$, then in particular we have a general method of solving the problem $\Gamma_p \vee \neg\Gamma_p$ for all values of p. If the latter general method is interpreted as an *algorithm* (in the sense of Turing machines) with input p (which is perfectly consistent with Kolmogorov's explanation of a general method), then such a general method would yield an algorithm deciding whether G_p is isomorphic to 1. But it is well-known that there exists no such algorithm.[32] This shows that the principle of decidability is not valid under the clarified BHK interpretation:

$$\not\vdash \cdot \gamma \vee \neg\gamma.$$

In the next section (§**3.13**) we will review a much more elementary example, which shows incidentally that it is not really necessary to interpret general methods constructively, as algorithms.

Let us pause to note that we have *not* established validity, on the clarified BHK interpretation, of the negated principle of decidability, $\cdot\neg\gamma \vee \neg\gamma$. For that we would need a solution of $\forall\Gamma\,!\neg?(\Gamma \vee \neg\Gamma)$, that is, a general proof (=a general method to prove) that each contentful problem of the form $\Gamma \vee \neg\Gamma$ has no solutions. But there exists no such general proof, since, in fact, any problem of the form $(\Delta \to \Delta) \vee \neg(\Delta \to \Delta)$ does have a solution. Thus we see, quite trivially, that

$$\not\vdash \cdot \neg(\gamma \vee \neg\gamma)$$

---

[31]One can also wonder about Hilbert's conviction, that $\Gamma \vee \neg\Gamma$ admits a (possibly *ad hoc*) solution for every contentful problem Γ. This judgement, which can be expressed as $\forall\Gamma\,?(\Gamma \vee \neg\Gamma)$ in the above notation, falls outside of the scope of intuitionistic logic on the clarified BHK interpretation,. A certain extension QHC of both intuitionistic and classical logics, which includes additional connectives ? and ! (and no further connectives) is studied in the author's papers [1], [2], [3]. Hilbert's conviction is naturally seen as an interpretation of the principle $\cdot?(\gamma \vee \neg\gamma)$ in the language of QHC, which turns out to be an independent principle with respect to QHC.

[32]The geometrically-minded reader might prefer the modification of this example based of S. P. Novikov's theorem (improving on an earlier result of A. A. Markov, Jr.), that for a certain sequence of finite simplicial complexes $K_1, K_2, \ldots$ there exists no algorithm to decide, for any positive integer input n, whether K_n is piecewise-linearly homeomorphic to the 5-dimensional sphere (see [31]).

on the clarified BHK interpretation. In fact, a stronger assertion turns out to hold:

$$\vDash \cdot \neg\neg(\gamma \vee \neg\gamma);$$

thus on the clarified BHK interpretation, there exists a general proof (we will describe it explicitly in §**3.14**, (32)) that for each problem Γ there exists no proof that $\Gamma \vee \neg\Gamma$ has no solutions. Yet on the other hand, we have just seen a proof that the problem $\forall p\,(\Gamma_p \vee \neg\Gamma_p)$ has no solutions. In particular, our former argument works also to establish

$$\nvDash \cdot \neg\neg\forall \mathsf{x}\left(\gamma(\mathsf{x}) \vee \neg\gamma(\mathsf{x})\right)$$

on the clarified BHK interpretation; this was originally observed by Brouwer, who presented "counterexamples to the freedom from contradiction of the Multiple Principle of Excluded Middle of the second kind" [28; §2], and now is better known as independence of the Double Negation Shift principle (see §*5.6.1*).

3.13. Medvedev–Skvortsov problems

Let us fix a set X, and consider the class of problems Γ_f of the form: *Solve the equation $f(x) = 0$, where $f\colon X \to \{0, 1\}$ is an arbitrary (set-theoretic) function.* Thus a solution of Γ_f is any $x \in X$ such that $f(x) = 0$. Such problems will be called *Medvedev–Skvortsov problems* with domain X. Of course, f is determined by the pair of sets $\left(X, f^{-1}(0)\right)$. Thus we may write a pair of sets, (X, Y), to encode the problem Γ_f, where $f\colon X \to \{0, 1\}$ is such that $f^{-1}(0) = Y$. It is this language of pairs that was used by Medvedev [136], [137], Skvortsov [183] and Läuchli [119] (see also [179]). If $\mathscr{D}$ is a fixed set (the domain of discourse) and $k \in \mathbb{N}$, we may also consider k-parameter families of Medvedev–Skvortsov problems with a fixed domain X. Every such family is encoded by a k-parameter family of pairs of sets $\left(X, Y(x_1, \ldots, x_k)\right)$, where $x_1, \ldots, x_k$ run over $\mathscr{D}$. Here Y is a function from $\mathscr{D}^k$ to the set 2^X of all subsets of X.

Let us consider the following interpretation of the ingredients of the language of intuitionistic logic.

- each nullary problem variable is interpreted by a pair of sets (X, Y);
- each individual variable is interpreted by an element of $\mathscr{D}$;
- each k-ary problem variable is interpreted by a k-parameter family of pairs sets of the form $\left(X, Y(t_1, \ldots, t_k)\right)$, where each t_i runs over $\mathscr{D}$;
- each formal connective or quantifier is interpreted as follows:
  - $(X, Y)\,|\wedge|\,(X', Y') = (X \times X',\ Y \times Y')$;
  - $(X, Y)\,|\vee|\,(X', Y') = (X \sqcup X',\ Y \sqcup Y')$;
  - $(X, Y)\,|\to|\,(X', Y') = \left(\operatorname{Hom}(X, X'), \{f\colon X \to X' \mid f(Y) \subset Y'\}\right)$;
  - $|\times| = (\{\varnothing\}, \varnothing)$;
  - $|\exists|t\left(X, Y(t)\right) = \left(\mathscr{D} \times X, \{(d, x) \in \mathscr{D} \times X \mid x \in Y(d)\}\right)$;
  - $|\forall|t\left(X, Y(t)\right) = \left(\operatorname{Hom}(\mathscr{D}, X), \{f\colon \mathscr{D} \to X \mid \forall d \in \mathscr{D}\ f(d) \in Y(d)\}\right)$.
- a principle $\cdot\Phi$ that contains occurrences of precisely m problem variables is interpreted by a function $|\cdot\Phi|$ that assigns to every m-tuple of sets $X_1, \ldots, X_m$

a pair of sets (S, S'), defined as follows. Let $x_1, \ldots, x_n$ be all individual variables that occur freely in Φ, let $\gamma_1, \ldots, \gamma_m$ be all problem variables that occur in Φ, and let $r_1, \ldots, r_m$ be their arities. Upon substituting each x_i by some $d_i \in \mathscr{D}$, and each γ_i by an r_i-parameter family of pairs sets of the form $\big(X_i, Y_i(t_1, \ldots, t_{r_i})\big)$, and applying the above interpretation of the connectives and quantifiers, we obtain an interpretation of Φ by a pair or sets (S, T). Here S depends only on $\vec{X} = (X_1, \ldots, X_m)$ and $T = T(\vec{d}, \vec{Y})$ is a subset of S that depends additionally on $\vec{d} = (d_1, \ldots, d_n)$ and $\vec{Y} = (Y_1, \ldots, Y_m)$. The desired pair $|\cdot \Phi|(\vec{X})$ is (S, S'), where S' is the intersection of the sets $T(\vec{d}, \vec{Y})$ over all $(\vec{d}, \vec{Y}) \in \mathscr{D}^n \times \mathrm{Hom}(\mathscr{D}^{r_1}, 2^{X_1}) \times \ldots \times \mathrm{Hom}(\mathscr{D}^{r_m}, 2^{X_m})$.

- $\vDash \cdot \Phi$ asserts that the function $|\cdot \Phi|$ assigns to every tuple of sets a pair (S, S') with nonempty S'.

This is very much in the spirit of the clarified BHK interpretation. Indeed, let us recall that a pair of sets (X, Y) is actually supposed to encode a problem whose set of solutions is Y. Thus let us write $\{(X, Y)\} = Y$. If we denote pairs of sets by Greek letters, then we obtain, just like in §3.9:

- $\{\Gamma \mid \wedge \mid \Delta\} = \{\Gamma\} \times \{\Delta\}$;
- $\{\Gamma \mid \vee \mid \Delta\} = \{\Gamma\} \sqcup \{\Delta\}$;
- there is a map $\mathscr{F} \colon \{\Gamma \mid \to \mid \Delta\} \to \mathrm{Hom}(\{\Gamma\}, \{\Delta\})$ (actually, a surjection);
- $\{\mid \times \mid\} = \varnothing$;
- $\{\mid \exists \mid t\, \Gamma(t)\} = \bigsqcup_{d \in \mathscr{D}} \{\Gamma(d)\}$;
- there is a map $\mathscr{G} \colon \{\mid \forall \mid t\, \Gamma(t)\} \to \prod_{d \in \mathscr{D}} \{\Gamma(d)\}$ (actually, a bijection).

Moreover, a "general method" of solving the problems encoded by $\big(S, T(\vec{d}, \vec{Y})\big)$ for all $(\vec{d}, \vec{Y})$ is naturally interpreted as a common solution to all these problems, i.e., an element in the intersection of all the sets $T(\vec{d}, \vec{Y})$. And then of course such a general method exists if and only if the intersection is nonempty. Thus the Medvedev–Skvortsov interpretation of intuitionistic logic is arguably a formalization of the clarified BHK interpretation.

Let us now consider the interpretation of the principle of decidability, $|\cdot \gamma \vee \neg \gamma|$. It is a function assigning to a set X the pair of sets $(S, \bigcap_{Y \subset X} T(Y))$, where

$$\big(S, T(Y)\big) = (X, Y) \mid \vee \mid \mid \neg \mid (X, Y).$$

We have $S = X \sqcup \mathrm{Hom}(X, \{\varnothing\})$ and $T(Y) = Y \sqcup \{f \colon X \to \{\varnothing\} \mid f(Y) \subset \varnothing\}$, that is,

$$T(Y) = \begin{cases} Y \sqcup \varnothing, & \text{if } Y \neq \varnothing; \\ \varnothing \sqcup \mathrm{Hom}(X, \{\varnothing\}), & \text{if } Y = \varnothing. \end{cases}$$

Thus $\bigcap_{Y \subset X} T(Y) = \varnothing$. So the principle of decidability is not valid under the Medvedev–Skvortsov interpretation.

3.14. Some intuitionistic validities

The point of the clarified BHK interpretation is that it provides a reasonable explanation of intuitionistic validities, without relying on any formal system of axioms and inference

rules. Although this explanation is highly informal, it works — and is often more helpful than any formal calculus if one needs to quickly verify whether a given formula is an intuitionistic validity.[33]

The order of precedence of connectives, quantifiers and the symbol $\cdot$ is (in groups of equal priority, starting with higher precedence/stronger binding): 1) $\neg$, $\exists$ and $\forall$; 2) $\wedge$ and $\vee$; 3) $\rightarrow$ and $\leftrightarrow$; 4) $\cdot$.

Another convention (perhaps obvious): we do not omit parameters of problem variables, e.g. in the formula $\forall x \, \big(\alpha \vee \beta(y) \vee \gamma(x, y)\big)$, α is a nullary problem variable, β is a unary one and γ is a binary one. (This agrees with the notation of §4.)

3.14.1. Basic validities.

(1) $\vDash \cdot \gamma \rightarrow \neg\neg\gamma$

Given a problem Γ and a solution s of Γ, we need to produce a solution of $\neg\neg\Gamma$ by a general method. Indeed, given a solution of $\neg\Gamma$, that is, a method turning solutions of Γ into solutions of $\times$, we simply apply this method to s and get a contradiction.

(2) (contrapositive) $\vDash \cdot (\gamma \rightarrow \delta) \; \longrightarrow \; (\neg\delta \rightarrow \neg\gamma)$

Indeed, from $\Gamma \rightarrow \Delta$ and $\Delta \rightarrow \times$ we infer $\Gamma \rightarrow \times$.

(3) $\vDash \cdot \neg\gamma \; \longleftrightarrow \; \neg\neg\neg\gamma$

This follows from (1): the "$\rightarrow$" implication by substitution, and the "$\leftarrow$" implication via (2).

(4) $\vDash \cdot \neg\neg\times \; \longleftrightarrow \; \times$

Here the "$\leftarrow$" implication is a special case of (1), and the "$\rightarrow$" implication can be rewritten as $\vDash \cdot \neg\neg\neg\times$, which by (3) is equivalent to $\vDash \cdot \neg\times$. The latter is a special case of $\vDash \cdot \gamma \rightarrow \gamma$ (where the problem $\Gamma \rightarrow \Gamma$ has an obvious solution).

(5) (explosion) $\vDash \cdot \times \rightarrow \gamma$

If s is a solution of $\times$, then s does not exist; in particular, s is itself also a solution of any given problem Γ.

(6) $\vDash \cdot \neg\delta \vee \gamma \; \longrightarrow \; (\delta \rightarrow \gamma)$

Given a solution of $\neg\Delta \vee \Gamma$ and a solution of Δ, we get a solution of $\times \vee \Gamma$, hence a solution of $\Gamma \vee \Gamma$. This yields a solution of Γ by considering two cases.

(7) $\vDash \cdot \beta \vee \gamma \; \longrightarrow \; (\neg\beta \rightarrow \gamma)$

This follows from (6) with δ substituted by $\neg\beta$, using (1).

(8) (decidability implies stability) $\vDash \cdot \gamma \vee \neg\gamma \; \longrightarrow \; (\neg\neg\gamma \rightarrow \gamma)$

---

[33]Why do these informal BHK-arguments work? One possible answer is that they do just because they are in fact sketches of proofs that the principles in question are satisfied in the sheaf-valued models of §6. Intuitionistic logic is shown to be complete with respect to this class of models in §6, so by formalizing these sketches one would indeed establish that the principles in question are intuitionistic validities. Alternatively, one can interpret the informal BHK-arguments as textual representations of λ-terms (see [93], [188; §4], [202; §10.8]) or, equivalently, as sketches of proofs that the principles in question are satisfied in Läuchli's models [119]. In fact, there seems to be no significant difference between the two alternatives (see [11]).

This follows from (6) with δ substituted by $\neg\neg\gamma$, using (3).

3.14.2. Quantifiers.

$$(9) \vDash \cdot \exists x \, \forall y \, \gamma(x, y) \;\longrightarrow\; \forall y \, \exists x \, \gamma(x, y)$$

Indeed, if we found an x_0 and a method to turn every y into a solution of $\Gamma(x_0, y)$, then we have a method to produce for every y an x and a solution of $\Gamma(x, y)$.

Here is a good place to note that, like in classical logic, $\exists$ and $\forall$ can be thought of as generalizations of $\vee$ and $\wedge$. Thus similarly to (9) we get

$$(10) \vDash \cdot \exists x \, \big(\gamma(x) \wedge \delta(x)\big) \;\longrightarrow\; \exists x \, \gamma(x) \wedge \exists x \, \delta(x)$$
$$(11) \vDash \cdot \forall y \, \gamma(y) \vee \forall y \, \delta(y) \;\longrightarrow\; \forall y \, \big(\gamma(y) \vee \delta(y)\big)$$

Intuitionistic logic features an additional connection: $\exists$ and $\forall$ behave as if they were generalizations of $\wedge$, and $\to$, respectively.[34] Thus similarly to (9) we also get

$$(12) \vDash \cdot \exists x \, \big(\theta \to \gamma(x)\big) \;\longrightarrow\; \big(\theta \to \exists x \, \gamma(x)\big)^{[35]}$$

Moreover, similarly to the obvious validities: $\vDash \cdot \forall x \, \forall y \, \gamma(x, y) \;\longleftrightarrow\; \forall y \, \forall x \, \gamma(x, y)$ and $\vDash \cdot \exists y \, \exists x \, \gamma(x, y) \;\longleftrightarrow\; \exists x \, \exists y \, \gamma(x, y)$ we get something otherwise not so obvious:

$$(13) \vDash \cdot \forall x \, \big(\theta \to \gamma(x)\big) \;\longleftrightarrow\; \big(\theta \to \forall x \, \gamma(x)\big)$$
$$(14) \vDash \cdot \theta \wedge \exists x \, \gamma(x) \;\longleftrightarrow\; \exists x \, \big(\theta \wedge \gamma(x)\big)$$

In addition, by specializing (11) we get

$$(15) \vDash \cdot \theta \vee \forall x \, \gamma(x) \;\longrightarrow\; \forall x \, \big(\theta \vee \gamma(x)\big)^{[36]}$$

There are two more validities of the above sort. Firstly,

$$(16) \vDash \cdot \forall x \, \big(\gamma(x) \to \theta\big) \;\longleftrightarrow\; \big(\exists x \, \gamma(x) \to \theta\big)$$

Indeed, a solution of $\forall x \, \big(\Gamma(x) \to \Theta\big)$ turns each x into a solution of $\Gamma(x) \to \Theta$. A solution of $\exists x \, \Gamma(x) \to \Theta$ turns any given x_0 and solution of $\Gamma(x_0)$ into a solution of Θ. These are obviously reducible to each other.

Secondly,

$$(17) \vDash \cdot \exists x \, \big(\gamma(x) \to \theta\big) \;\longrightarrow\; \big(\forall x \, \gamma(x) \to \theta\big)^{[37]}$$

Indeed, given an x_0 and a solution of $\Gamma(x_0) \to \Theta$, and assuming a method of solving $\Gamma(x)$ for every x, we apply this method to $x = x_0$ to get a solution of $\Gamma(x_0)$ and consequently a solution of Θ.

$$(18) \vDash \cdot \neg \exists x \, \gamma(x) \;\longleftrightarrow\; \forall x \, \neg\gamma(x)$$
$$(19) \vDash \cdot \exists x \, \neg\gamma(x) \;\longrightarrow\; \neg\forall x \, \gamma(x)^{[38]}$$

These are special cases of (16) and (17).

---

[34]This connection is made precise in dependent type theory, where no distinction is made between the domain of a variable, like $\mathscr{D}$ in the BHK clauses for $\exists$ and $\forall$, and the set of solutions of a problem, like that of Γ in the BHK clauses for $\wedge$ and $\to$.

[35]The converse implication is known as the Principle of Independence of Premise; see §*5.2.2*.

[36]The converse implication is known as the Constant Domain Principle; see §*5.2.1* below.

[37]The converse implication is an independent principle; see §*5.2.2*.

[38]The converse implication is known as the Generalized Markov Principle; see §**5.5**.

3.14.3. Distributivity.

(20) $\vDash \cdot\, \theta \wedge (\gamma \vee \delta) \;\longleftrightarrow\; (\theta \wedge \gamma) \vee (\theta \wedge \delta)$

(21) $\vDash \cdot\, \theta \vee (\gamma \wedge \delta) \;\longleftrightarrow\; (\theta \vee \gamma) \wedge (\theta \vee \delta)$

Here (20) and the "$\rightarrow$" implication in (21) follow similarly to (14) and (15).

The remaining implication says that if we have (i) either a solution of Θ, or a solution of Γ, and (ii) either a solution of Θ, or a solution of Δ, then we can produce, by a general method, (iii) either a solution of Θ, or solutions of Γ and Δ. Clearly, there is such a general method which proceeds by a finite analysis of cases; in fact, there are two such distinct methods m_i, $i = 1, 2$, which in the case that we have two solutions of Θ select the ith one.

(22) $\vDash \cdot\, (\theta \rightarrow \gamma) \wedge (\theta \rightarrow \delta) \;\longleftrightarrow\; \big(\theta \rightarrow (\gamma \wedge \delta)\big)$

(23) $\vDash \cdot\, (\theta \rightarrow \gamma) \vee (\theta \rightarrow \delta) \;\longrightarrow\; \big(\theta \rightarrow (\gamma \vee \delta)\big)$[39]

These follow similarly to (13) and (12).

(24) $\vDash \cdot\, (\gamma \rightarrow \theta) \wedge (\delta \rightarrow \theta) \;\longleftrightarrow\; \big((\gamma \vee \delta) \rightarrow \theta\big)$

(25) $\vDash \cdot\, (\gamma \rightarrow \theta) \vee (\delta \rightarrow \theta) \;\longrightarrow\; \big((\gamma \wedge \delta) \rightarrow \theta\big)$[40]

These follow similarly to (16) and (17).

(26) (exponential law) $\vDash \cdot\, (\alpha \wedge \beta \rightarrow \gamma) \;\longleftrightarrow\; \big(\alpha \rightarrow (\beta \rightarrow \gamma)\big)$

(27) $\vDash \cdot\, \alpha \wedge (\beta \rightarrow \gamma) \;\longrightarrow\; \big((\alpha \rightarrow \beta) \rightarrow \gamma\big)$

These are again similar to (16) and (17).

(28) (de Morgan law) $\vDash \cdot\, \neg\gamma \wedge \neg\delta \;\longleftrightarrow\; \neg(\gamma \vee \delta)$

(29) (de Morgan law) $\vDash \cdot\, \neg\gamma \vee \neg\delta \;\longrightarrow\; \neg(\gamma \wedge \delta)$[41]

These are special cases of (24) and (25).

(30) $\vDash \cdot\, \neg(\gamma \wedge \delta) \;\longleftrightarrow\; (\gamma \rightarrow \neg\delta)$

(31) $\vDash \cdot\, \gamma \wedge \neg\delta \;\longrightarrow\; \neg(\gamma \rightarrow \delta)$

These follow from (26) and (27).

3.14.4. Consequences and refinements.

(32) $\vDash \cdot\, \neg\neg(\gamma \vee \neg\gamma)$

Indeed, a solution of $\neg(\Gamma \vee \neg\Gamma)$ yields by (28) a solution of $\neg\Gamma \wedge \neg\neg\Gamma$, that is, a solution of $\neg\Gamma$ together with a method to turn such a solution into a contradiction.

(33) $\vDash \cdot\, (\gamma \rightarrow \neg\delta) \;\longleftrightarrow\; (\delta \rightarrow \neg\gamma)$

This follows from (30).

(34) $\vDash \cdot\, (\neg\neg\gamma \rightarrow \neg\neg\delta) \;\longleftrightarrow\; (\neg\delta \rightarrow \neg\gamma)$

This "idempotence of contrapositive" follows from (3) and (33).

(35) $\vDash \cdot\, (\neg\neg\gamma \rightarrow \neg\delta) \;\longleftrightarrow\; (\gamma \rightarrow \neg\delta)$

---

[39] The converse implication is equivalent to the Gödel–Dummett principle $\cdot\, (\gamma \rightarrow \delta) \vee (\delta \rightarrow \gamma)$; see §5.2.2.

[40] The converse implication, Skolem's principle, is equivalent to the Gödel–Dummett principle; see §5.2.2.

[41] The converse implication is equivalent to Jankov's principle $\neg\alpha \vee \neg\neg\alpha$, see Proposition 5.9 below.

This follows from (33) and (3), or alternatively from (1), (2) and (3).

$$(36) \quad \vDash \cdot \neg(\gamma \wedge \neg\neg\delta) \; \longleftrightarrow \; \neg(\gamma \wedge \delta)$$
$$(37) \quad \vDash \cdot \neg(\gamma \vee \neg\neg\delta) \; \longleftrightarrow \; \neg(\gamma \vee \delta)$$

These follow from (30) and (28) respectively, using (3).

Surprisingly, (31) is a consequence of a reversible implication:

$$(38) \quad \vDash \cdot \neg\neg\gamma \wedge \neg\delta \; \longleftrightarrow \; \neg(\gamma \rightarrow \delta)$$

Here the "←" implication follows from (28) and the contrapositive of (6).

Conversely, suppose we are given a solution of $\neg\neg\Gamma \wedge \neg\Delta$ and a solution of $\Gamma \rightarrow \Delta$. Then we have solutions of $\neg\Delta$, of $\neg\Gamma \rightarrow \times$, and of $\neg\Delta \rightarrow \neg\Gamma$ (the contrapositive). Applying the latter to the solution of $\neg\Delta$, we get a solution of $\neg\Gamma$, and hence a contradiction.

$$(39) \quad \vDash \cdot \neg(\neg\delta \vee \gamma) \; \longleftrightarrow \; \neg(\delta \rightarrow \gamma)$$

This improvement on the contrapositive of (6) follows from (38) and (28).

$$(40) \quad \vDash \cdot \neg(\beta \vee \gamma) \; \longleftrightarrow \; \neg(\neg\beta \rightarrow \gamma)$$

This improvement on the contrapositive of (7) follows from (39) and (37).

$$(41) \quad \vDash \cdot \neg(\neg\gamma \vee \neg\delta) \; \longleftrightarrow \; \neg\neg(\gamma \wedge \delta)$$

This improvement on the contrapositive of (29) follows from (39) and (30).

$$(42) \quad \vDash \cdot \neg(\gamma \wedge \neg\delta) \; \longleftrightarrow \; \neg\neg(\gamma \rightarrow \delta)$$

This improvement on the contrapositive of (31) follows from (38) and (36).

$$(43) \quad \vDash \cdot \neg\neg\gamma \wedge \neg\neg\delta \; \longleftrightarrow \; \neg\neg(\gamma \wedge \delta)$$

This follows from (41) and (28), or alternatively from (38) and (30).

$$(44) \quad \vDash \cdot (\neg\neg\gamma \rightarrow \neg\neg\delta) \; \longleftrightarrow \; \neg\neg(\gamma \rightarrow \delta)$$

This also follows from (38) and (30).

$$(45) \quad \vDash \cdot \neg\neg\gamma \vee \neg\neg\delta \; \longrightarrow \; \neg\neg(\gamma \vee \delta)^{[42]}$$

This follows from (29) using (28).

$$(46) \quad \vDash \cdot \neg\neg\forall x\, \gamma(x) \; \longrightarrow \; \forall x\, \neg\neg\gamma(x)^{[43]}$$
$$(47) \quad \vDash \cdot \exists x\, \neg\neg\gamma(x) \; \longrightarrow \; \neg\neg\exists x\, \gamma(x)^{[44]}$$

Each of these follows from (19) using (18).

4. What is a logic, formally?

This chapter provides a rather unconventional introduction to basic first-order logic.

On the one hand, it aims at being relatively readable by including just enough detail to make it possible for the interested reader to reconstruct the rest. (For instance, we use parentheses without discussing their official treatment in the formal language.)

---

[42]The converse implication is equivalent to Jankov's principle $\neg\alpha \vee \neg\neg\alpha$, see Proposition 5.9 below.

[43]The converse implication is known as the principle of Double Negation Shift; see §*5.6.1*.

[44]The converse implication is known as the Strong Markov Principle; see §**5.5**.

On the other hand, we are much more careful about meta-level reasoning than is usually done. Some reasons for doing this are discussed in the following Introduction.

4.1. Introduction

This Introduction is not formally used in the sequel; it is addressed to the reader who has some acquaintance with logic textbooks and is wondering what is the point of yet another treatment of basic first-order logic.

4.1.1. Two conceptions of a logic. We identify a logic with the collection of its derivable rules (including derivable principles, i.e. derivable rules with no premisses). This is equivalent to identifying a logic with its consequence relation. In the terminology of Wójcicki [212; §1.6] and Blok–Pigozzi [24] this is the "inferential" conception of logic, as opposed to the "formulaic" one, in which a logic is identified with the collection of its derivable principles (or, equivalently, derivable formulas); these authors as well as Avron [8] argue in favor of the inferential conception.

In particular, as noted by Wójcicki [212; §1.6.0], if the purpose of a logic is to elucidate schemes of reasoning, then logical schemes of inferences found in a mathematical practice that is considered to be in agreement with a logic L are captured by the derivable rules of L; in general, these cannot be reconstructed from the derivable principles of L alone, which capture only direct use of logical validities in the said mathematical practice.

The two conceptions of a logic ascribe somewhat different meanings to the notion of an inference rule, which is related to the difference between derivable and admissible rules. (A rule is called admissible if adding it as a new inference rule to the derivation system does not affect the collection of derivable principles. Thus, the notion of admissibility is "formulaic" in that it only depends on the collection of derivable principles.) In classical logic, the difference between derivable and admissible rules exists but is rather insignificant (see Examples 4.68 and 4.69, Remark 4.70 and Proposition 4.73 below). So the entire issue of two different conceptions of a logic only really manifests itself in intuitionistic and other non-classical logics.

4.1.2. What is a rule? Another variety of phenomena that one cannot fully appreciate by only looking at classical logic is independent principles, as well as consequences between them — and more generally, consequences between rules (see, however, Examples 4.75 and 4.82 and Proposition 4.74 concerning the situation with these in classical logic). Analyzing such consequences is going to be quite helpful in the next chapter to get some feel of how intuitionistic logic works.

Now, even to give a definition of derivable rules or consequences between principles, some precise definition of a rule — and, in particular, of a principle — is needed. Indeed, it is not so clear exactly what they are, since the formal language provides no syntax for their usual side conditions such as "provided that x is not free in α" or "provided that t is free for x in $\alpha(x)$", and *a fortiori* no formal rules for dealing with these side conditions. Often principles and rules are treated as arbitrary infinitary collections of formulas (or

of tuples of formulas). But this is clearly at odds with the fact that in essence they are finite objects (written with finitely many symbols in logic textbooks). The well-known workaround [162] (see also [47]): a rule is called structural if it is preserved by all substitutions without anonymous variables (in the sense of Kleene [103]) — is clearly still far from being a satisfactory solution.

The usual way to understand (and represent in writing) principles and rules as finite objects is by means of schemata, which belong to the meta-language. While it is not difficult to make this meta-language just as formal as the language itself (for example, by regarding the set of terms and the sets of α-equivalence classes of n-formulas of the language as semantic objects, and creating an obvious syntax to describe them), this approach does not seem to be technically advantageous (see, however, [176], [152], [61] concerning formal calculi of meta-variables, and [172; §3.7], [99] concerning schematic rules). And, as it turns out, there is a good reason for not taking this path: the extra level of linguistic abstraction (where syntax is regarded as an abstraction of semantics) is actually an overkill, because the desired effect can already be achieved by means of the usual λ-abstraction (where a function is regarded as an abstraction of the dependence of a prescribed expression on prescribed variables).

4.1.3. Meta-logic. To get rid of side conditions of the form "provided that x is not free in α", we will use a *meta-quantifier*. In the author's initial approach, the desired effect was achieved by means of a device that is like a usual quantifier in that it binds an (otherwise free) metavariable in a schema, but is invisible on the level of formulas, since the binding is only used to specify which formulas are considered to be instances of the schema. Thus it was natural to call this device a "meta-quantifier".

However, by googling for the word *meta-quantifier*, the author discovered that a similar, but much more convenient device has been used for the same purpose (of getting rid of the side conditions) in the `Isabelle` theorem prover [148; §5.9], [157]. It has also been called the "meta-quantifier" (e.g. in the index of the old manual [158]) but for a somewhat different reason [157; §2.1]: it is the (universal) quantifier of `Isabelle`'s meta-logic, which serves the purpose of a common ground for a number of different logics supported by `Isabelle`.

This view of `Isabelle` and its creator Paulson that the meta-quantifier is really a quantifier of the meta-logic turns out to be very beneficial. The side conditions do not disappear, they are only relegated from rules to meta-rules, where they can be seen to belong naturally. No schemata are needed whatsoever, because the meta-quantifier operates on formulas. In this respect Paulson's approach builds on the early tradition of first-order logic, found in the textbooks by Hilbert–Ackermann [89], Hilbert–Bernays [90], Church [32] (alongside the modern schematic approach) and P. S. Novikov [149], [151]. In this tradition, one does not speak of any schemata, but only of formulas; instead, the derivation system includes a substitution rule (see also Church [32] and Feferman [53] for a comparison and the history of the two approaches). However, inference rules were anyway stated in schematic form; and the substitution rule may be seen as problematic

because it is not structural. In Paulson's approach, inference rules involve no meta-variables, and the substitution rule effectively becomes a meta-rule.

Another principal ingredient of `Isabelle`'s metalogic is the meta-implication, which is closely related, in particular, to the entailment (the horizontal bar in rules) and to the consequence ($\vdash$) — not to be conflated with the implication ($\rightarrow$) — in object logics. These relations will be clarified below; the former is described implicitly in Paulson's example [157; §4.4, subsection "Hypothetical rules"], whereas of the latter Paulson only notes that $\vdash$ is not the same as the meta-implication [157; Remark at the end of §3].

To summarize, meta-logic enables one to "explain" many standard notions in logic that one just gets used to rather than understands in the traditional approach, as well as to avoid some traditional ambiguities, which can easily lead to errors through mechanical application of the official definitions.

4.1.4. Two conceptions of consequence. In mathematics, there are two ways to interpret formulas with parameters. The following example is given in Kleene's textbook [103; §32]: the arithmetical formula $(x + y)^2 = x^2 + 2xy + y^2$ begs to be understood as an identity (valid for all natural numbers x), whereas the arithmetical formula $x^2 + 2 = 3x$ begs to be understood as an equation (i.e., as a condition on x). There is no difference in syntax to reflect this obvious distinction in meaning. But the latter is not illusory: it is reflected in use. For, as noted by Avron [10], when "dealing with identities [...] the substitution rule is available, and one may infer $\sin x = 2 \sin \frac{x}{2} \cos \frac{x}{2}$ from the identity $\sin 2x = 2 \sin x \cos x$. In contrast, [...] substituting $\frac{x}{2}$ for x everywhere in an equation is an error".

In first-order logic, similar phenomena in the use of substitution are normally associated with the distinction between free and bound variables: it is only bound variables that can be harmlessly renamed.[45] However, in all of the above examples, x is intended (by Kleene and Avron) to signify a free variable (in the sense of first-order logic). Logicians often speak of "fresh free variables", which can be seen to behave like variables in identities rather than equations. While this idea suffices for practical applications (cf. [157; §7]), it does not seem to be appropriate for theoretical/foundational purposes.

Instead, the way that traditional first-order logic accounts for the semantic distinction between "identities" and "equations" is by distinguishing two notions of syntactic consequence and two corresponding notions of semantic consequence. Thus Kleene [103; §22, §32] speaks of variables being *varied* or *held constant* in a deduction (effectively distinguishing two variants of syntactic consequence); whereas Avron [8], [9], [10] speaks of *truth* and *validity* semantic consequence relations (both are mentioned also e.g. in the Kolmogorov–Dragalin textbook [111]). The same sort of distinction is also thought to exist in propositional modal logic [8], [9] (see also [76] and references there).

---

[45] In the words of Paulson, "The purpose of the proviso 'x not free in ...' is to ensure that the premise may not make assumptions about the value of x, and therefore holds for all x." [158; §1.2.2].

However, the view that first-order logic (classical or intuitionistic) must include two distinct consequence relations at the same time has never gained wide acceptance, presumably due to its technical inconvenience and philosophical awkwardness. Most authors appear to believe that there is a unique, "true" consequence relation (either syntactic or semantic), but then they disagree with each other on which one it is. Thus, the "truth"/"fixed variables" approach, corresponding to the equational interpretation of free variables, is chosen e.g. by Church [32], Enderton [48], Troelstra [200] and van Dalen [207]; whereas the "validity"/"varied variables" approach, corresponding to the identity interpretation of free variables, is chosen e.g. by Shoenfield [181] and, in the syntactic part, by Mendelson [141].

Each of the two approaches has its advantages. In particular, with the "validity"/"varied variables" approach, the unrestricted Hilbert-style generalization rule holds (and so the notion of a *rule* does not have to be very complicated). With the "truth"/"fixed variables" approach, the unrestricted deduction theorem holds (and so proof-theoretic formalisms such as sequent calculus and natural deduction apply in their usual form). The two approaches are discussed and compared — with the opposite conclusions reached — by Blok–Pigozzi [24] and Hakli–Negri [76].

The meta-quantifier offers a somewhat different outlook on these issues, which finally clarifies the picture. There is only one consequence relation (coming from the meta-logic), from which both traditional ones can be easily recovered. An appropriate form of the generalization rule holds without restrictions (and so the notion of a rule does not have to be complicated; in fact, as discussed above, it is simpler than in either of the traditional approaches). The deduction theorem holds without restrictions — or rather under the automatic restriction that the meta-quantifier cannot be used inside a formula (because logical connectives and quantifiers cannot operate on expressions that involve meta-logical ones).

4.2. Typed expressions

This section is mostly standard material (cf. [188], [64], [121]), apart from some deviations in notation and terminology, which opt for a more mathematical (rather than computer science) style, and some additional conventions in §*4.2.4*, §*4.2.5*.

4.2.1. Simply typed λ-calculus. One fixes a collection of *base types* (we will encounter only three base types: of terms, of formulas, and of meta-formulas). Arbitrary *types* (or in more detail *types of λ-expressions*)[46] are defined as follows:

(1) base types are types;
(2) if Γ and Δ are types, then so is $\Gamma \to \Delta$;
(3) if Γ and Δ are types, then so is $\Gamma \times \Delta$.

---

[46]These are also called "arities" following Martin-Löf (see [147; §3.6]), but we will use the word "arity" in its traditional sense.

For each type Δ, one may fix a collection of *constants* of type Δ. For each type Δ, one always has a countable ordered collection of *variables* of type Δ, denoted x_i^Δ, $i = 1, 2, \ldots$. (Thus, by saying "x is a variable of type Δ" we mean that $x = x_i^\Delta$ for some i.) Arbitrary *expressions*[47] *of type* Δ are defined as follows, where $T : \Delta$ denotes the relation "T is an expression of type Δ":

(1) a constant of type Δ is an expression of type Δ;
(2) a variable of type Δ is an expression of type Δ;
(3) (λ-abstraction) if $T : \Delta$ on the assumption that x is a variable of type Γ, then $x \mapsto T : \Gamma \to \Delta$;[48]
(4) (function application) if $F : \Delta \to \Gamma$, and $T : \Delta$, then $F(T) : \Gamma$;[49]
(5) (tuples) if $S : \Gamma$ and $T : \Delta$, then $(S, T) : \Gamma \times \Delta$.
(6) (projections) if Γ, Δ are types, then $p_1 : \Gamma \times \Delta \to \Gamma$ and $p_2 : \Gamma \times \Delta \to \Delta$.

We will abbreviate $F\big((S, T)\big)$ by $F(S, T)$, and usually also $\big(F(T)\big)(S)$ by $F(T)(S)$ and $p_i(T)$ by $p_i T$.

Example 4.1. Let $\mathbb{R}$ be a type with one constant for each real number. Then addition of real numbers $(x, y) \mapsto x + y$ is of type $\mathbb{R} \times \mathbb{R} \to \mathbb{R}$; integration of functions $(f, (a, b)) \mapsto \int_a^b f(x)\, dx$ is of type $(\mathbb{R} \to \mathbb{R}) \times (\mathbb{R} \times \mathbb{R}) \to \mathbb{R}$; and the differential $f \mapsto df$ is of type $(\mathbb{R} \to \mathbb{R}) \to (\mathbb{R} \to (\mathbb{R} \to \mathbb{R}))$. (This example ignores some complications such as non-differentiable functions, etc.)

Example 4.2. Given $F : \Theta \to \Gamma$ and $G : \Gamma \to \Delta$, we can define their *composition* $G \circ F : \Theta \to \Delta$ by $r \mapsto G(F(r))$. Thus $\circ : (\Theta \to \Gamma) \times (\Gamma \to \Delta) \to (\Theta \to \Delta)$ is the following expression: $q \mapsto \Big(r \mapsto p_1(q)\big(p_2(q)(r)\big) \Big)$.

Example 4.3. In first-order logic as treated e.g. in `Isabelle`, we have a type of terms, which will be denoted below by $\mathbb{0}$, and a type of formulas, which will be denoted below by $\mathbb{1}$. Then a *quantifier* is a constant of type $(\mathbb{0} \to \mathbb{1}) \to \mathbb{1}$. If $\mathfrak{q}$ is a quantifier, the expression $\mathfrak{q}(x \mapsto F)$ is customarily denoted by $\mathfrak{q}x\, F$. (So e.g. for the universal quantifier $\forall : (\mathbb{0} \to \mathbb{1}) \to \mathbb{1}$ this takes the usual form $\forall x\, F$.)

Remark 4.4. Throughout this section we will continue to denote variables by lowercase Roman letters, constants by lowercase Fraktur letters, unknown terms by uppercase Roman letters, compound terms that are specific (at least up to α-equivalence, which is defined below) by uppercase Fraktur letters; and unknown types by uppercase Greek letters. These letters all represent *metavariables* in that they do not themselves belong to λ-calculus, as opposed to the variables x_i^Δ, the constants p_i, specific base types and their specific constants (these will introduced in the next section). It should also be noted that the indices of the variables x_i^Δ are plain numbers (i.e., sequences of bars, such

---

[47]In λ-calculus, these are usually called "terms", but we reserve this word for expressions of one specific type, in accordance with the terminology of first-order logic.

[48]In λ-calculus, $x \mapsto T$ is usually written as $\lambda x.T$, but we will use the traditional mathematical notation.

[49]In λ-calculus, $F(T)$ is usually written as FT, but we will normally use the more customary mathematical notation. In some cases we do resort to FT, regarding it as an abbreviation for $F(T)$.

as $|||||$, abbreviated as usual by means of decimal notation) and not numeric expressions; thus $\varkappa^{\Delta}_{10^{10^{10}}}$ is not an expression of type Δ, but only an informal meta-expression that may be taken as a euphemism referring to an actual expression of type Δ (which is unprintable).

Abstraction *binds* variables; variables that are not bound by an abstraction are called *free*. More precisely, whether a variable x of some type *occurs freely* in an expression T of some type is determined by the following recursive definition (recursion over the number of symbols in T). Let y be a variable distinct from x (i.e. $x = \varkappa^{\Gamma}_i$ and $y = \varkappa^{\Delta}_j$ where either $\Gamma \neq \Delta$ or $i \neq j$), let $\mathfrak{c}$ be a constant of some type and F, R and S be expressions of some types. Then:

(1) x occurs freely in x;
(2) x does not occur freely in y;
(3) x does not occur freely in $\mathfrak{c}$;
(4) x occurs freely in (R, S) if it occurs freely in R or in S;
(5) x occurs freely in $F(R)$ if it occurs freely in F or in R;
(6) x does not occur freely in $x \mapsto R$;
(7) x occurs freely in $y \mapsto R$ if it occurs freely in R.

For example, x does not occur freely in $x \mapsto x$; but y occurs freely in $(x \mapsto x)(y)$. Also, f and g occur freely in $(f \mapsto f)\big(x \mapsto f\big(g(x)\big)\big)$.

An expression of some type is *closed* if no variable of any type occurs freely in it.

Example 4.5. A closed expression $\mathfrak{E} : \big(\Lambda \to (\Delta \to \Gamma)\big) \to \big((\Lambda \to \Delta) \to (\Lambda \to \Gamma)\big)$ can be defined by $f \mapsto \Big(g \mapsto \big(l \mapsto f(l)\big(g(l)\big)\big)\Big)$.

Example 4.6. A closed expression $\mathfrak{D} : \big((\Delta \to \Gamma) \to \Theta\big) \to \Big(\big(\Delta \to (\Lambda \to \Gamma)\big) \to (\Lambda \to \Theta)\Big)$ can be defined by $f \mapsto \Big(g \mapsto \big(l \mapsto f\big(t \mapsto g(t)(l)\big)\big)\Big)$.

4.2.2. Substitution. Let $T : \Delta$, and let x be a variable of type Δ. The *(capture-avoiding, strict) substitution* $S|_{x:=T}$, where $S : \Gamma$, is, when defined, an expression of type Γ, which is determined recursively as follows. Let y be a variable distinct from x, let $\mathfrak{c}$ be a constant of some type and F, R and S be expressions of some types. Then:

(1) $x|_{x:=T} = T$;
(2) $y|_{x:=T} = y$;
(3) $\mathfrak{c}|_{x:=T} = \mathfrak{c}$;
(4) $(R, S)|_{x:=T} = (R|_{x:=T}, S|_{x:=T})$;
(5) $F(R)|_{x:=T} = (F|_{x:=T})(R|_{x:=T})$;
(6) $x \mapsto R|_{x:=T} = x \mapsto R$;
(7) $y \mapsto R|_{x:=T} = \begin{cases} \text{undefined}, & \text{if } x \text{ occurs freely in } R \text{ and } y \text{ occurs freely in } T; \\ y \mapsto (R|_{x:=T}) & \text{otherwise.} \end{cases}$

Of course, the recursion halts once it reaches the "undefined" case, and in this event $S|_{x:=T}$ is set to be undefined.

Example 4.7. In the notation of Example 4.3, suppose that $>$ is a constant of type $\mathbb{0} \times \mathbb{0} \to \mathbb{1}$ and $+$ is a constant of type $\mathbb{0} \times \mathbb{0} \to \mathbb{0}$, and let us write $x > y$ to mean $>(x, y)$ and $x + y$ to mean $+(x, y)$. Then

- $\forall x \; x > y|_{y:=y+z} = \forall x \; x > y + z$;
- $\forall x \; x > y|_{y:=x+z}$ is undefined;
- $\forall x \; x > y|_{x:=y+z} = \forall x \; x > y$.

The following auxiliary definition will be needed on several occasions. A relation $\succ$ on expressions is called *compatible* if it satisfies the following conditions, where x is a variable of some type, and F, G, R, S, T are expressions of some types.

(1) if $S \succ T$, then $x \mapsto S \succ x \mapsto T$;
(2) if $S \succ T$, then $F(S) \succ F(T)$;
(3) if $F \succ G$, then $F(T) \succ G(T)$;
(4) if $S \succ T$, then $(R, S) \succ (R, T)$ and $(S, R) \succ (T, R)$;
(5) if $S \succ T$, then $\mathbb{p}_1 S \succ \mathbb{p}_1 T$ and $\mathbb{p}_2 S \succ \mathbb{p}_2 T$.

Two expressions are considered to be the same for all practical purposes if one can be obtained from another by renaming of bound variables. In more detail, the α-*equivalence* relation is the least compatible equivalence relation on expressions of arbitrary types such that whenever x, y are variables of some type and T is an expression of some type,

- if y does not occur freely in T and $T|_{x:=y}$ is defined, then $x \mapsto T \overset{\alpha}{=} y \mapsto (T|_{x:=y})$.

Clearly, α-equivalent expressions are always of the same type. Two α-equivalent expressions are also said to be obtained from one another by α-*conversion*.

Example 4.8. $x \mapsto \big(y \mapsto F(x, y)\big)$ is α-equivalent to $a \mapsto \big(b \mapsto F(a, b)\big)$ and to $y \mapsto \big(x \mapsto F(y, x)\big)$. However, $x \mapsto (y \mapsto x)$ is not α-equivalent to $y \mapsto (x \mapsto y)$ (because of the free variable).

Also, $x \mapsto (x \mapsto x)$ is α-equivalent to $x \mapsto (y \mapsto y)$ and to $y \mapsto (x \mapsto x)$ but is not α-equivalent to $y \mapsto (x \mapsto y)$. Informally speaking, since bound variables can be renamed, they are essentially only nameless placeholders. In particular, once a variable gets bound (like in $x \mapsto x$), it loses its name to an outside observer; this name may then be recycled (like in $x \mapsto (x \mapsto x)$), should one need to do so.

The point of α-equivalence that if a substitution $S|_{x:=T}$ is undefined, then we can replace S by an α-equivalent term S' so that $S'|_{x:=T}$ is defined. (For example, $x \mapsto f(y)|_{y:=g(x)}$ is undefined, but becomes defined upon renaming x in $x \mapsto f(y)$ into z.) Although S' can be chosen in many ways, the result will be well-defined up to α-equivalence.

Let us note that the α-equivalence class $S|_{x:=T}^{\alpha}$ includes the results (when they are defined) of all substitutions of the form $S'|_{x:=T'}$, where S' is α-equivalent to S and T' is α-equivalent to T. This is useful, for instance, since one can always make a double substitution of the form $(S|_{x:=T})|_{y:=U}$ defined by replacing S and T by α-equivalent terms (see [188]).

4.2.3. βη-equivalence. The relation of *β-reduction* is the least compatible relation such that

(1) every expression of the form $(x \mapsto S)(T)$ β-reduces to $S|_{x:=T}$;
(2) every expression of the form $\mathbb{p}_i(T_1, T_2)$ β-reduces to T_i for $i = 1, 2$;
(3) α-equivalent expressions β-reduce to each other.

For example, $(x \mapsto x)(T)$ β-reduces to T; and $(x \mapsto T)(x)$ β-reduces to T.

Example 4.9. In Example 4.5, $\mathfrak{E}(F)(G)(L)$ β-reduces $F(L)\big(G(L)\big)$ in three steps. In Example 4.6, $\mathfrak{D}(F)(G)(L)$ β-reduces to $F(t \mapsto G(t)(L)$ in three steps.

Example 4.10. The expression $\big(f \mapsto f(z)\big)\big(x \mapsto (z \mapsto x)\big)$ β-reduces to $f(z)|_{f:=x\mapsto(z\mapsto x)}$. This substitution is defined, with the result $\big(x \mapsto (z \mapsto x)\big)(z)$. But the latter β-reduces to $z \mapsto x|_{x:=z}$, which is not defined. Up to α-equivalence, this is the same as $y \mapsto x|_{x:=z}$, which evaluates to $y \mapsto z$.

The relation of *η-reduction* is the least compatible relation such that

(1) every expression of the form $x \mapsto F(x)$, where x does not occur freely in F, η-reduces to F;
(2) every expression of the form $\big(\mathbb{p}_1 T, \mathbb{p}_2 T\big)$ η-reduces to T;
(3) α-equivalent expressions η-reduce to each other.

Let us note that η-reduction converts between functions that return the same output on the same input (possibly by employing different algorithms).

The inverse of a β- or η-reduction is called a *β- or η-expansion.*

The *βη-equivalence* is the least compatible equivalence relation on expressions of arbitrary types generated by β-reduction and η-reduction. Clearly, $\beta\eta$-equivalent expressions are always of the same type. We will denote $\beta\eta$-equivalence by the symbol $=$.

Example 4.11. Let us construct closed expressions $\mathfrak{F} : (\Theta \times \Gamma \to \Delta) \to \big(\Theta \to (\Gamma \to \Delta)\big)$ and $\mathfrak{G} : \big(\Theta \to (\Gamma \to \Delta)\big) \to (\Theta \times \Gamma \to \Delta)$ such that $\mathfrak{F}\big(\mathfrak{G}(G)\big) = G$ and $\mathfrak{G}\big(\mathfrak{F}(F)\big) = F$.

We define $\mathfrak{F}$ by $f \mapsto \Big(r \mapsto \big(s \mapsto f(r,s)\big)\Big)$ and $\mathfrak{G}$ by $g \mapsto \Big(q \mapsto g\big(\mathbb{p}_1 q\big)\big(\mathbb{p}_2 q\big)\Big)$.

$$\mathfrak{F}\big(\mathfrak{G}(G)\big) \overset{\beta}{=} \mathfrak{F}\Big(q \mapsto G\big(\mathbb{p}_1 q\big)\big(\mathbb{p}_2 q\big)\Big) \overset{\beta}{=} r \mapsto \big(s \mapsto f(r,s)\big)\Big|_{f:=q\mapsto G(\mathbb{p}_1 q)(\mathbb{p}_2 q)}$$

$$= r \mapsto \left(s \mapsto \big(q \mapsto G\big(\mathbb{p}_1 q\big)\big(\mathbb{p}_2 q\big)\big)(r,s)\right) \overset{\beta}{=} r \mapsto \left(s \mapsto \Big(G\big(\mathbb{p}_1 q\big)\big(\mathbb{p}_2 q\big)\big|_{q:=(r,s)}\Big)\right)$$

$$= r \mapsto \Big(s \mapsto G\big(\mathbb{p}_1(r,s)\big)\big(\mathbb{p}_2(r,s)\big)\Big) \overset{\beta}{=} r \mapsto \Big(s \mapsto G(r)(s)\Big) \overset{\eta}{=} r \mapsto G(r) \overset{\eta}{=} G.$$

Also,

$$\mathfrak{G}\big(\mathfrak{F}(F)\big) \overset{\beta}{=} \mathfrak{G}\Big(r \mapsto \big(s \mapsto F(r,s)\big)\Big) \overset{\beta}{=} q \mapsto g\big(\mathbb{p}_1 q\big)\big(\mathbb{p}_2 q\big)\Big|_{g:=r\mapsto\big(s\mapsto F(r,s)\big)}$$

$$= q \mapsto \Big(r \mapsto \big(s \mapsto F(r,s)\big)\Big)\big(\mathbb{p}_1 q\big)\big(\mathbb{p}_2 q\big) \overset{\beta}{=} q \mapsto \big(s \mapsto F(r,s)\big|_{r:=\mathbb{p}_1 q}\big)\big(\mathbb{p}_2 q\big)$$

$$= q \mapsto \big(s \mapsto F(\mathbb{p}_1 q, s)\big)\big(\mathbb{p}_2 q\big) \overset{\beta}{=} q \mapsto \Big(F(\mathbb{p}_1 q, s)\big|_{s:=\mathbb{p}_2 q}\Big) = q \mapsto F\big(\mathbb{p}_1 q, \mathbb{p}_2 q\big)$$

$$\overset{\eta}{=} q \mapsto F(q) \overset{\eta}{=} F.$$

Remark 4.12. Here is an alternative argument, which uses (in addition to β- and η-reductions) also η-expansions (but is perhaps conceptually clearer).

We have $\mathfrak{F}(F) \overset{\beta}{=} r \mapsto \big(s \mapsto F(r,s)\big)$, and consequently $\mathfrak{F}(F)(R) \overset{\beta}{=} s \mapsto F(R,s)$, and hence also $\mathfrak{F}(F)(R)(S) \overset{\beta}{=} F(R,S)$.

Also, $\mathfrak{G}(G) \overset{\beta}{=} q \mapsto G\big(\mathbb{p}_1 q\big)\big(\mathbb{p}_2 q\big)$, and therefore $\mathfrak{G}(G)(Q) \overset{\beta}{=} G\big(\mathbb{p}_1 Q\big)\big(\mathbb{p}_2 Q\big)$, and hence also $\mathfrak{G}(G)(R,S) \overset{\beta}{=} G(R)(S)$.

Thus $\mathfrak{F}\big(\mathfrak{G}(G)\big)(R)(S) = \mathfrak{G}(G)(R,S) = G(R)(S)$.

Also, $\mathfrak{G}\big(\mathfrak{F}(F)\big)(R,S) = \mathfrak{F}(F)(R)(S) = F(R,S)$.

Finally,

$$\mathfrak{G}\big(\mathfrak{F}(F)\big) \overset{\eta}{=} q \mapsto \mathfrak{G}\big(\mathfrak{F}(F)\big)(q) \overset{\eta}{=} q \mapsto \mathfrak{G}\big(\mathfrak{F}(F)\big)(\mathbb{p}_1 q, \mathbb{p}_2 q) = q \mapsto F(\mathbb{p}_1 q, \mathbb{p}_2 q)$$

$$\overset{\eta}{=} q \mapsto F(q) \overset{\eta}{=} F.$$

Also,

$$\mathfrak{F}\big(\mathfrak{G}(G)\big) \overset{\eta}{=} r \mapsto \mathfrak{F}\big(\mathfrak{G}(G)\big)(r) \overset{\eta}{=} r \mapsto \Big(s \mapsto \mathfrak{F}\big(\mathfrak{G}(G)\big)(r)(s)\Big) = r \mapsto \Big(s \mapsto G(r)(s)\Big)$$

$$\overset{\eta}{=} r \mapsto G(r) \overset{\eta}{=} G.$$

Example 4.13. Let us construct closed expressions $\mathfrak{A} : \Gamma \times \Delta \to \Gamma \times (\Gamma \to \Delta)$ and $\mathfrak{B} : \Gamma \times (\Gamma \to \Delta) \to \Gamma \times \Delta$ such that $\mathfrak{B}\big(\mathfrak{A}(Q)\big) = Q$, but in general $\mathfrak{A}\big(\mathfrak{B}(Q)\big) \neq Q$.

We set $\mathfrak{A} = q \mapsto \big(\mathbb{p}_1 q, \ s \mapsto \mathbb{p}_2 q\big)$ and $\mathfrak{B} = q \mapsto \Big(\mathbb{p}_1 q, \ \mathbb{p}_2(q)\big(\mathbb{p}_1 q\big)\Big)$.

Then $\mathfrak{A}(S,T) \overset{\beta}{=} (S, s \mapsto T)$ and $\mathfrak{B}(S,F) \overset{\beta}{=} \big(S, F(S)\big)$.

Hence $\mathfrak{B}\big(\mathfrak{A}(S,T)\big) = \mathfrak{B}(S, s \mapsto T) \overset{\beta}{=} (S,T)$. It follows that $\mathfrak{B}\big(C(Q)\big) = Q$.

However, $\mathfrak{A}\big(\mathfrak{B}(S,F)\big) = \mathfrak{A}\big(S, F(S)\big) = \big(S, s \mapsto F(S)\big) \neq \big(S, s \mapsto F(s)\big) \overset{\eta}{=} (S,F)$.

4.2.4. Simultaneous substitution.

Convention 4.14. For $n \geq 2$ let us write $\Delta_1 \times \ldots \times \Delta_{n+1}$ to mean $(\Delta_1 \times \ldots \times \Delta_n) \times \Delta_{n+1}$ and $(T_1, \ldots, T_{n+1})$ to mean $\big((T_1, \ldots, T_n), T_{n+1}\big)$. We also write $\mathbb{p}_i : \Delta_1 \times \ldots \times \Delta_{n+1} \to \Delta_i$ to mean $\mathbb{p}_i \circ \mathbb{p}_1$ for $i \leq n$ and to mean $\mathbb{p}_2$ for $i = n+1$.

We abbreviate $F\big((T_1, \ldots, T_n)\big)$ by $F(T_1, \ldots, T_n)$. We also write Δ^n for the n-tuple product $\Delta \times \ldots \times \Delta$.

We will further identify $\Delta_1 \times \Delta_2 \times \Delta_3$ with $\Delta_1 \times (\Delta_2 \times \Delta_3)$ via the "function" of type $\Delta_1 \times \Delta_2 \times \Delta_3 \to \Delta_1 \times (\Delta_2 \times \Delta_3)$, defined by $q \mapsto \left(\mathbb{p}_1 q, \left(\mathbb{p}_2 q, \mathbb{p}_3 q\right)\right)$, and the "function" of type $\Delta_1 \times (\Delta_2 \times \Delta_3) \to \Delta_1 \times \Delta_2 \times \Delta_3$, defined by $q \mapsto \left(\mathbb{p}_1 q, \mathbb{p}_1\left(\mathbb{p}_2 q\right), \mathbb{p}_2\left(\mathbb{p}_2 q\right)\right)$, which are easily seen to be mutually inverse.

By an iterated use of this identification we also identify $\Delta_1 \times \ldots \times \Delta_n$ with any iterated product of $\Delta_1, \ldots, \Delta_n$.

For each $i = 1, \ldots, n$ let $T_i : \Delta_i$, and let x_i be a variable of type Δ_i. Let us write $\vec{x} := (x_1, \ldots, x_n)$ and $\vec{T} := (T_1, \ldots, T_n)$, and also $\vec{x}|\hat{k}$ and $\vec{T}|\hat{k}$ for the same tuples with x_k and T_k omitted. The *(capture-avoiding, strict) simultaneous substitution* $S|_{\vec{x}:=\vec{T}}$, where $S : \Gamma$, is, when defined, an expression of type Γ, which is determined recursively as follows. Let y be a variable distinct from each x_i, let $\mathfrak{c}$ be a constant of some type and F, R and S be expressions of some types. Then:

(1) $x_k|_{\vec{x}:=\vec{T}} = T_k$;

(2) $y|_{\vec{x}:=\vec{T}} = y$;

(3) $\mathfrak{c}|_{\vec{x}:=\vec{T}} = \mathfrak{c}$;

(4) $(R, S)|_{\vec{x}:=\vec{T}} = (R|_{\vec{x}:=\vec{T}}, S|_{\vec{x}:=\vec{T}})$;

(5) $F(R)|_{\vec{x}:=\vec{T}} = (F|_{\vec{x}:=\vec{T}})(R|_{\vec{x}:=\vec{T}})$;

(6) $x_k \mapsto R|_{\vec{x}:=\vec{T}} = x_k \mapsto (R|_{\vec{x}|\hat{k}:=\vec{T}|\hat{k}})$;

(7) $y \mapsto R|_{\vec{x}:=\vec{T}} = \begin{cases} \text{undefined}, & \text{if for some } i, \\ & x_i \text{ occurs freely in } R \text{ and } y \text{ occurs freely in } T_i \\ y \mapsto (R|_{\vec{x}:=\vec{T}}) & \text{otherwise.} \end{cases}$

Example 4.15. $\left((x, y)|_{x:=y}\right)\big|_{y:=x} = (y, y)$ and $\left((x, y)|_{y:=x}\right)\big|_{x:=y} = (x, x)$. However, $(x, y)|_{x:=y, y:=x} = (y, x)$.

Example 4.16. Somewhat unexpectedly, $\mathfrak{G}\left(x \mapsto (y \mapsto T)\right) = \mathfrak{G}\left(y \mapsto (x \mapsto T)\right)$. Indeed, $\mathfrak{G}\left(x \mapsto (y \mapsto T)\right)(R)(S) = (y \mapsto T|_{x:=R})(S)$, where the substitution $|_{x:=R}$ is undefined if y occurs freely in R. If z is a variable that does not occur freely in R and T, then up to α-equivalence we have

$$(y \mapsto T|_{x:=R})(S) = \left(z \mapsto (T|_{y:=z})|_{x:=R}\right)(S) = \left((T|_{y:=z})|_{x:=R}\right)\big|_{z:=S} = T|_{x:=R, y:=S}.$$

Similarly, $\mathfrak{G}\left(y \mapsto (x \mapsto T)\right)(R)(S) = T|_{x:=R, y:=S}$, and the assertion follows.

Convention 4.17. If $x_0, \ldots, x_n$ are variables of types $\Delta_0, \ldots, \Delta_n$, and $T : \Delta$, an expression

$$x_0, \ldots, x_n \mapsto T$$

of type $\Delta_0 \times \ldots \times \Delta_n \to \Delta$ is defined recursively as $\mathfrak{G}\left(x_0 \mapsto (x_1, \ldots, x_n \mapsto T)\right)$ (see Convention 4.14). Thus, for instance, $\left(x, y \mapsto (x, y)\right)(y, x) = (y, x)$.

Up to α-equivalence, $x_1, \ldots, x_n \mapsto T$ is the same as $q \mapsto T|_{x_1:=\mathbb{p}_1 q, \ldots, x_n:=\mathbb{p}_n q}$, where q is a variable of type $\Delta_1 \times \ldots \times \Delta_n$ that does not occur freely in T. The new expression is defined without recursion, but is well-defined only up to α-equivalence (because of the

variable q). Let us note that the substitutions $|_{x_i:=\mathbb{p}_i q}$ commute, since no x_j occurs in the expression $\mathbb{p}_i q$. So it does not matter if we make them simultaneously or in some order.

Remark 4.18. If $\vec{x}$ denotes an n-tuple of variables of some types with $n > 0$, one may use the notation $\vec{x} \mapsto T$ as if it were an expression of the appropriate type — and not just an abbreviation for such an expression. In particular, we have a version of the β-reduction: $(\vec{x} \mapsto S)(\vec{T}) = S|_{\vec{x}:=\vec{T}}$, and a version of the η-reduction: $\vec{x} \mapsto F(\vec{x}) = F$.

One may also use the notation $\vec{T}$ as if it were an expression of the appropriate type — and not just an abbreviation for such an expression. We have a version of the β-reduction: $\mathbb{p}_i(T_1, \ldots, T_n) = T_i$, and a version of the η-reduction: $(\mathbb{p}_1 T, \ldots, \mathbb{p}_n T) = T$. However, one must be careful not to assume that this all works also for $n = 0$, since it does not. Something much more involved does [36] (see also [38]).

Convention 4.19. We will regard $\Delta^0 \times \Gamma$, $\Gamma \times \Delta^0$ and $\Delta^0 \to \Gamma$ as abbreviations for Γ (without attempting to define Δ^0).

4.2.5. Free substitution. It is well-known (see e.g. [188], [121]) that for every expression T of any type Δ there exists an expression $T_0 : \Delta$ such that T $\beta\eta$-reduces to T_0, and if T $\beta\eta$-reduces to T_1, then T_1 $\beta\eta$-reduces to T_0; moreover, every T_0' satisfying the same property is α-equivalent to T_0. Such a T_0 is called a *normal form* of T; and if T is its own normal form, it is said to be *in normal form*. (For our purposes the normal form could be replaced in what follows by the β-normal form, which is defined similarly using β-reduction.)

If T_0 is a normal form of T, and some $\beta\eta$-reduction of T to T_0 involves no α-conversions, then T_0 will be called a *strict normal form* of T, and T will be said to *have a strict normal form*. It is not hard to see that when a strict normal form of a λ-expression T exists, it is unique (as a λ-expression, and not just up to α-equivalence). Clearly, T is its own strict normal form if and only if it is in normal form.

If S is in normal form, and the expression $(\vec{x} \mapsto S)(\vec{T})$ has a strict normal form, where $\vec{x} = (x_1, \ldots, x_n)$, $\vec{T} = (T_1, \ldots, T_n)$, each x_i is a variable of some type Δ_i, and each $T_i : \Delta_i$, then this strict normal form of $(\vec{x} \mapsto S)(\vec{T})$ will be denoted by $S[\vec{x}/\vec{T}]$, or in more detail by $S[x_1/T_1, \ldots, x_n/T_n]$, and $\vec{T}$ will be called *free for (substitution for) $\vec{x}$ in S*, or in more detail $T_1, \ldots, T_n$ will be called *free for (simultaneous substitution for) $x_1, \ldots, x_n$ in S*.

4.3. Languages and structures

4.3.1. Terms and n-terms. A *first-order language* $\mathscr{L}$ provides, in particular, formal means to speak about *individuals*, which are the elements of a certain nonempty set $\mathscr{D}$, called the *domain of discourse*. For instance, if $\mathscr{L}$ is the standard language of arithmetic, a good choice for $\mathscr{D}$ would be the set of natural numbers. Now $\mathscr{D}$ and its elements are regarded as "semantic" objects, whereas $\mathscr{L}$ itself is about "syntax" only. This means, in particular, that $\mathscr{L}$ does not refer on its own to the elements of any specific set $\mathscr{D}$, but

is capable of doing it in multiple ways. One particular way is specified by an $\mathscr{L}$-*structure* $\imath$, also known as an *interpretation* of $\mathscr{L}$, which includes a choice of a set $\mathscr{D}$ along with a number of other choices. The formal definitions of $\mathscr{L}$ and $\imath$ will be summarized at the end of §*4.3.2*, once we will have introduced all of their ingredients.

The "individual" fragment of $\mathscr{L}$ is can be formulated in terms of a base type, which we denote $\mathbb{0}$. The variables $\mathbb{x}_1^{\mathbb{0}}, \mathbb{x}_2^{\mathbb{0}}, \ldots$ of this type are the *individual variables* of the language $\mathscr{L}$. For reasons of readability we will also use the alternative spelling $\mathsf{a}, \mathsf{b}, \mathsf{c}, \ldots, \mathsf{x}, \mathsf{y}, \mathsf{z}$ for the first 26 ones, reserving an upright sans-serif font for this purpose. These are abbreviations, and not metavariables, as they are determined by the specific identifications $\mathsf{a} = \mathbb{x}_1^{\mathbb{0}}$, $\mathsf{b} = \mathbb{x}_2^{\mathbb{0}}$, etc. Metavariables are used, for instance, in a phrase like "let a and b be individual variables", which amounts to saying "let $a = \mathbb{x}_i^{\mathbb{0}}$ and $b = \mathbb{x}_j^{\mathbb{0}}$ for some unspecified i and j (possibly $i = j$)".

$\mathscr{L}$ may also contain *function symbols* (such as $+$ and $*$), each standing for an operation on $\mathscr{D}$ (such as addition and multiplication of natural numbers) that inputs an n-tuple of individuals (for some specified $n \geq 0$) and outputs one individual. Thus an n-ary function symbol of $\mathscr{L}$ is a constant $\mathfrak{f}$ of type $\mathbb{0}^n \to \mathbb{0}$, and its interpretation $\mathfrak{f}_\imath$ is an n-ary operation $\mathscr{D}^n \to \mathscr{D}$. Nullary function symbols, which are constants of type $\mathbb{0}$ (such as $\mathsf{0}$ and $\mathsf{1}$), are called *individual constants* of $\mathscr{L}$ and stand for specific individuals (such as the natural numbers 0 and 1).

Terms of $\mathscr{L}$ are defined inductively, as built out of individual variables using the function symbols. A sample term in the case of the language of arithmetic is $(\mathsf{x} + \mathsf{1}) * \mathsf{y} + \mathsf{x}$. Thus terms of $\mathscr{L}$ are those expressions of type $\mathbb{0}$ that involve only variables of type $\mathbb{0}$, tuples, function application, and a prescribed collection of constants of types $\mathbb{0}^n \to \mathbb{0}$. (In particular, terms of $\mathscr{L}$ do not involve λ-abstraction.)

If T is a *closed term* (i.e. a term that is a closed expression), its interpretation $|T|_\imath$ is a specific element of $\mathscr{D}$. To interpret an arbitrary term T as an element of $\mathscr{D}$, one additionally needs a *variable assignment* u, which assigns an individual $\mathit{u}(x)$ to each individual variable x. Indeed, u extends in the obvious way to a function $|\cdot|_\imath^{\mathit{u}} \colon \mathscr{T} \to \mathscr{D}$, where $\mathscr{T} = \mathscr{T}(\mathscr{L})$ is the set of terms of $\mathscr{L}$. Thus for every term T we have $|T|_\imath^{\mathit{u}} \in \mathscr{D}$.

It is often convenient to consider expressions of the form $x_1, \ldots, x_n \mapsto T$, where T is a term and the x_i are pairwise distinct individual variables (in other words, each $x_i = \mathbb{x}_{i_n}^{\mathbb{0}}$, where $i_1, \ldots, i_n$ are pairwise distinct). We will call these n-*terms*; of course, they are of type $\mathbb{0}^n \to \mathbb{0}$. Let us note that the arithmetical 1-terms $\mathsf{x} \mapsto \mathsf{x} + \mathsf{y}$, $\mathsf{y} \mapsto \mathsf{x} + \mathsf{y}$ and $\mathsf{z} \mapsto \mathsf{x} + \mathsf{y}$ should not be conflated, since application works differently for them: $(\mathsf{x} \mapsto \mathsf{x} + \mathsf{y})(\mathsf{t}) = \mathsf{t} + \mathsf{y}$, $(\mathsf{y} \mapsto \mathsf{x} + \mathsf{y})(\mathsf{t}) = \mathsf{x} + \mathsf{t}$ and $(\mathsf{z} \mapsto \mathsf{x} + \mathsf{y})(\mathsf{t}) = \mathsf{x} + \mathsf{y}$. One can define n-terms inductively, as built out of *atomic k-terms*, which are individual variables and expressions of the form $x \mapsto x$, where x is an individual variable, by using appropriate lifts of the function symbols. For instance, using the lifted sum, from the arithmetical 2-terms $\mathsf{x}, \mathsf{y} \mapsto \mathsf{x} * \mathsf{y}$ and $\mathsf{y}, \mathsf{z} \mapsto 3 * \mathsf{y}$ one obtains the arithmetical 3-term $\mathsf{x}, \mathsf{y}, \mathsf{z} \mapsto \mathsf{x} * \mathsf{y} + 3 * \mathsf{y}$. (However, one does not obtain anything in this fashion from the arithmetical 1-terms $\mathsf{x} \mapsto \mathsf{x} * \mathsf{y}$ and $\mathsf{y} \mapsto \mathsf{y}$.)

In more detail, every k-term can be lifted to a $(k + l)$-term by introducing l dummy variables via $\mathfrak{F}_l : (\mathbb{0}^k \to \mathbb{0}) \to (\mathbb{0}^k \times \mathbb{0}^l \to \mathbb{0})$, defined by $f \mapsto \left(q \mapsto f\big(\mathbb{p}_1 q\big) \right)$; and every function symbol, $\mathfrak{f} : \mathbb{0}^n \to \mathbb{0}$, can be lifted to $\mathfrak{f}_m : (\mathbb{0}^m \to \mathbb{0})^n \to (\mathbb{0}^m \to \mathbb{0})$ via $\mathfrak{G}_m :$ $(\mathbb{0}^n \to \mathbb{0}) \to \big((\mathbb{0}^m \to \mathbb{0})^n \to (\mathbb{0}^m \to \mathbb{0})\big)$, $c \mapsto \left(r \mapsto \Big(q \mapsto c\big(\mathbb{p}_1(r)(q), \ldots, \mathbb{p}_n(r)(q)\big) \Big) \right)$. Thus $\mathfrak{f}_m(F_1, \ldots, F_n) = \mathfrak{f} \circ (F_1 \times \ldots \times F_n)$, where $F_1 \times \ldots \times F_n : \mathbb{0}^m \to \mathbb{0}^n$ is defined by $q \mapsto \big(F_1(q), \ldots, F_n(q)\big)$. Let us note that an interpretation of a function symbol induces, via the analogue of $\mathfrak{G}_m$, an interpretation of its lifted version, which is by a function $\mathrm{Hom}(\mathscr{D}^m, \mathscr{D})^n \to \mathrm{Hom}(\mathscr{D}^m, \mathscr{D})$, where $\mathrm{Hom}(X, Y)$ denotes the set of all maps $X \to Y$.

Using the inductive construction of n-terms, every *closed n-term* F (i.e. an n-term that is a closed expression) is interpreted by a specific function $|F|_\imath : \mathscr{D}^n \to \mathscr{D}$. Also, given a variable assignment $\boldsymbol{u}$, any n-term F is interpreted by a function $|F|_\imath^u : \mathscr{D}^n \to \mathscr{D}$.

Since λ-abstraction is traditionally not included in first-order languages, n-terms are, strictly speaking, not included in the language for $n > 0$. This is not a serious issue, since an n-term uniquely corresponds to a pair consisting of a term and an n-tuple of pairwise distinct positive integers. However, there is no such simple connection between terms and α-equivalence classes of n-terms.

4.3.2. Formulas. In fact, the language $\mathscr{L}$ is not supposed to be a means to speak directly of individuals, but rather of functions of individuals such as propositions or problems about individuals or tuples of individuals. The case of 0-tuples amounts to propositions/problems without parameters, which are really the subject of zero-order logic — but are also included in first-order logic.

Instead of dealing with actual human-language sentences expressing propositions or problems, one fixes a set $\mathscr{O}$ of mathematical objects, which can be thought of as Platonic "ideas" (or rather Fregean "senses") of propositions or problems (thus an element of $\mathscr{O}$ is supposed to encode all propositions or problems without parameters that express the same "idea" in different words). A function $\mathscr{D}^m \to \mathscr{O}$ can then be thought of as encoding a proposition/problem with m parameters.

Of course, on assuming that a proposition has no "sense" beyond its "reference" (that is, beyond its truth value), one must have $\mathscr{O} = \{0, 1\}$. However, even with classical logic, there are more interesting options, discovered by Leibniz and Euler and reviewed in §4.8.2 below.

Now these all are "semantic" objects, whereas $\mathscr{L}$ itself is about "syntax" only. Thus, a choice of $\mathscr{O}$ is included not in the language $\mathscr{L}$, but in its interpretation $\imath$, which will still include a number of further choices.

Now, just like elements of $\mathscr{D}$ are represented by terms and other expressions of type $\mathbb{0}$, elements of $\mathscr{O}$ will be represented by expressions of a new base type, $\mathbb{1}$. $\mathscr{L}$ may contain constants of type $\mathbb{0}^n \to \mathbb{1}$, which are called *$n$-ary predicate constants*. Thus an n-ary predicate constant $\mathfrak{p}$ is interpreted by a predicate $|\mathfrak{p}|_\imath : \mathscr{D}^n \to \mathscr{O}$. Examples include the arithmetical binary predicates $=$ and $<$ and the geometric ternary predicate of betweenness (for points on a line).

The variables $\varkappa_1^{\mathbb{0}^n \to \mathbb{1}}, \varkappa_2^{\mathbb{0}^n \to \mathbb{1}}, \ldots$ of type $\mathbb{0}^n \to \mathbb{1}$ for each $n = 0, 1, 2 \ldots$ are called the *predicate variables* of $\mathscr{L}$. In dealing with classical logic (or modal logics based on classical logic), we will also call nullary predicate variables *propositional variables*, and use the nicknames $\mathsf{a}, \mathsf{b}, \mathsf{c}, \ldots, \mathsf{x}, \mathsf{y}, \mathsf{z}$ for the first 26 predicate variables of each arity, reserving a fancy (Euler) upright serif font for this purpose. To indicate the arity of a predicate variable f (here f is a metavaraible, in contrast to the abbreviations $\mathsf{a}, \mathsf{b}, \ldots, \mathsf{f}, \ldots$), we may write the appropriate η-expansion of f such as $\mathsf{x} \mapsto f(\mathsf{x})$ or $\mathsf{x}, \mathsf{y} \mapsto f(\mathsf{x}, \mathsf{y})$ instead of f. (See Convention 4.17.) In dealing with intuitionistic logic (and its non-classical extensions), the predicate variables can be called *problem variables*; in this case, we will also write $\alpha, \beta, \gamma, \ldots, \chi, \psi, \omega$ for the first 24 ones of each arity, reserving also the Greek part of the font.

The language $\mathscr{L}$ may further contain constants of type $\mathbb{1}^n \to \mathbb{1}$, which stand for logical operations, and are called *n-ary connectives*. Thus for every n-ary connective $\mathfrak{f}$, its interpretation $|\mathfrak{f}|_\iota$ is a logical operation $\mathcal{O}^n \to \mathcal{O}$.

It should be noted that nullary connectives are the same kind of λ-expressions as nullary predicate constants (all of them are constants of type $\mathbb{1}$). It is nevertheless convenient to distinguish them, since connectives belong to the language of a logic, whereas predicate constants only belong to the language of a theory over a logic.

$\mathscr{L}$ may also contain constants of type $(\mathbb{0} \to \mathbb{1}) \to \mathbb{1}$, which are called *quantifiers*. A quantifier $\mathfrak{q}$ is interpreted by a function $|\mathfrak{q}|_\iota \colon \mathrm{Hom}(\mathscr{D}, \mathcal{O}) \to \mathcal{O}$. If $\mathfrak{q}$ is a quantifier, F is an expression of type $\mathbb{1}$, and x is an individual variable, then $\mathfrak{q}x\, F$ abbreviates the expression $\mathfrak{q}(x \mapsto F)$. Here the individual variable x is bound by the λ-abstraction $\mapsto$; but traditionally it is also said to be *bound* by the quantifier. Also, F is called the *scope* of (the occurrence of) the quantifier. Moreover, $\mathfrak{q}x_0, \ldots, x_n\, F$ abbreviates $\mathfrak{q}x_0\,(\mathfrak{q}x_1, \ldots, x_n\, F)$.

Example 4.20. The language of intuitionistic logic contains a nullary connective $\times$ ("absurdity"), which stands for a problem that has no solutions, binary connectives $\wedge$ ("conjunction"), $\vee$ ("disjunction") and $\to$ ("implication"; not to be confused with the type constructor $\to$), and quantifiers $\exists$ ("existential") and $\forall$ ("universal"). There is no redundancy between these ingredients of the language (see §5.8 below).

It is also convenient to add, as is often done, some "syntactic sugar" to the language of intuitionistic logic: a logical constant $\checkmark$ ("triviality") is defined as $\times \to \times$; a unary connective $\neg$ ("negation") is defined as $\alpha \mapsto \alpha \to \times$ (thus $\checkmark = \neg\times$), and a binary connective $\leftrightarrow$ ("equivalence") is defined as $\alpha, \beta \mapsto (\alpha \to \beta) \wedge (\beta \to \alpha)$.

Example 4.21. The language of classical logic can be taken to be the same as that of intuitionistic logic, except for the logical constants, which we will write as $\top$ ("truth") and $\perp$ ("falsity"). (This complements our convention of denoting predicate variables by Greek letters in the intuitionistic case and by Roman letters in the classical case). There is additional redundancy here: the classical $\wedge$ can be defined as $p, q \mapsto \neg(p \to \neg q)$, the classical $\vee$ as $p, q \mapsto \neg p \to q$, and the classical $\exists$ as $f \mapsto \neg\forall x\, \neg f(x)$.

Example 4.22. The language of a theory over classical or intuitionistic logic (for instance, the "empty theory", which has no axioms) may additionally contain finitely many function symbols and/or predicate constants.

An *atomic formula* of $\mathscr{L}$ is an expression of type $\mathbb{1}$ obtained by applying either an n-ary predicate constant or an n-ary predicate variable to an n-tuple of terms. A *formula* of $\mathscr{L}$ is an expression built out of atomic formulas using the connectives and the quantifiers. In other words, formulas are expressions of type $\mathbb{1}$ involving only variables of types $\mathbb{0}$ and $\mathbb{0}^n \to \mathbb{1}$, tuples, function application, a prescribed collection of constants of types $\mathbb{0}^n \to \mathbb{0}$, $\mathbb{0}^n \to \mathbb{1}$ and $\mathbb{1}^n \to \mathbb{1}$, and a prescribed collection of constants of type $(\mathbb{0} \to \mathbb{1}) \to \mathbb{1}$ used in conjunction with certain specific λ-abstractions. (No other uses of λ-abstraction are allowed.)

A formula is called *closed* (or "closed as a formula") if individual variables do not occur freely in it, and *λ-closed* (or "closed as a λ-expression") if no variables occur freely in it. The arithmetical formula $\forall x, y\ \big(\alpha(x, y) \wedge \alpha(y, x) \to x = y\big)$ is closed, but not λ-closed. The arithmetical formula $\forall x, y, z\ \big(x + z = y + z \to x = y\big)$ is λ-closed.

This completes our definition of a first-order language $\mathscr{L}$. Namely, it consists of the following sets of typed λ-expressions (variables and constants only), where n ranges over $\mathbb{N} = \{0, 1, 2, \dots\}$:

(1_n) of n-ary function symbols (of type $\mathbb{0}^n \to \mathbb{0}$);
(2_n) of n-ary predicate constants (of type $\mathbb{0}^n \to \mathbb{1}$);
(3_n) of n-ary connectives (of type $\mathbb{1}^n \to \mathbb{1}$);
  (4) of quantifiers (of type $(\mathbb{0} \to \mathbb{1}) \to \mathbb{1}$);
  (5) of individual variables (of type $\mathbb{0}$);
(6_n) of n-ary predicate variables (of type $\mathbb{0}^n \to \mathbb{1}$);

Each of the sets (5), (6_n) is countably infinite, and each of the remaining sets is required to be finite. The sets (3_n)–(6_n) are called the *purely logical part* of the first-order language $\mathscr{L}$. The set of terms $\mathscr{T}(\mathscr{L})$ and the set of α-equivalence classes of formulas $\mathscr{F}_0(\mathscr{L})$ can be reconstructed from the sets (1)–(6), so they do not have to be included in $\mathscr{L}$.

The definition of an $\mathscr{L}$-structure $\imath$ is also complete: it consists of the sets $\mathscr{D}$ and $\mathscr{O}$ and of the following functions, where n ranges over $\mathbb{N}$:

($1_n'$) from the set of n-ary function symbols to $\mathrm{Hom}(\mathscr{D}^n, \mathscr{D})$;
($2_n'$) from the set of n-ary predicate constants to $\mathrm{Hom}(\mathscr{D}^n, \mathscr{O})$;
($3_n'$) from the set of n-ary connectives to $\mathrm{Hom}(\mathscr{O}^n, \mathscr{O})$;
  (4') from the set of quantifiers to $\mathrm{Hom}(\mathrm{Hom}(\mathscr{D}, \mathscr{O}), \mathscr{O})$.

We have not yet defined interpretation of formulas in an $\mathscr{L}$-structure; this will be done next, with the aid of n-formulas.

4.3.3. n-Formulas and $(n, \vec{r})$-formulas. By an *n-formula* we mean an expression of the form $x_1, \dots, x_n \mapsto F$, where the x_i are pairwise distinct individual variables and F is a formula. Thus n-formulas are of type $\mathbb{0}^n \to \mathbb{1}$; Kleene [103; §34] called them "name

forms". An n-formula is *closed* if individual variables do not occur freely in it, and λ-*closed* if no variables occur freely in it. Let us note that quantifiers do not operate on formulas, but rather transform a 1-formula into a formula; thus in some sense n-formulas are inherent in first-order languages — even though traditionally these do not include n-formulas.

It should be noted that an arbitrary formula can be viewed as a closed n-formula applied to an n-tuple of individual variables; thus closed n-formulas can be viewed as more fundamental than formulas. Moreover, an arbitrary k-formula can be viewed as a closed $(n + k)$-formula pre-composed with an expression of the form $x_1, \ldots, x_n \mapsto (x_1, \ldots, x_n, y_1, \ldots, y_k)$, where the x_i and the y_j are pairwise distinct individual variables. Thus closed n-formulas can be viewed as more fundamental even than arbitrary k-formulas.

Using k-formulas, one can give an alternative recursive description of formulas (and n-formulas), which is particularly suitable for closed formulas (and closed n-formulas). An *atomic (closed) n-formula* is an $(n + k)$-ary predicate variable or constant pre-composed with a (closed) expression of the form $x_1, \ldots, x_n \mapsto (T_1, \ldots, T_{n+k})$, where the x_i are pairwise distinct individual variables and each T_j is a term. In particular, each $x_1, \ldots, x_n \mapsto T_j$ is a (closed) n-term. A *(closed) n-formula* is an expression built out of atomic (closed) k-formulas using appropriate lifts of the connectives and the quantifiers. For instance, using the lifted conjunction, from closed 2-formulas $\mathsf{x}, \mathsf{y} \mapsto \varphi(\mathsf{x}, \mathsf{y})$ and $\mathsf{y}, \mathsf{z} \mapsto \psi(\mathsf{y}, \mathsf{z})$ one obtains the closed 3-formula $\mathsf{x}, \mathsf{y}, \mathsf{z} \mapsto \varphi(\mathsf{x}, \mathsf{y}) \wedge \psi(\mathsf{y}, \mathsf{z})$; and using the lifted existential quantifier, from the latter one obtains the closed 2-formula $\mathsf{x}, \mathsf{z} \mapsto \exists \mathsf{y} \, \varphi(\mathsf{x}, \mathsf{y}) \wedge \psi(\mathsf{y}, \mathsf{z})$.

In more detail, every k-formula can be made into a $(k + l)$-formula by introducing l dummy variables via $\mathfrak{F}_l : (\mathbb{O}^k \to \mathbb{1}) \to (\mathbb{O}^k \times \mathbb{O}^l \to \mathbb{1})$, defined by $f \mapsto \left(q \mapsto f(\mathbb{p}_1 q) \right)$; and every connective on formulas, $\mathfrak{c} : \mathbb{1}^n \to \mathbb{1}$, can be made into a connective on m-formulas, $\mathfrak{c}_m : (\mathbb{O}^m \to \mathbb{1})^n \to (\mathbb{O}^m \to \mathbb{1})$, via $\mathfrak{G}_m : (\mathbb{1}^n \to \mathbb{1}) \to \left((\mathbb{O}^m \to \mathbb{1})^n \to (\mathbb{O}^m \to \mathbb{1}) \right)$, defined by $c \mapsto \left(r \mapsto \left(q \mapsto c(\mathbb{p}_1(r)(q), \ldots, \mathbb{p}_n(r)(q)) \right) \right)$. Thus $\mathfrak{c}_m(F_1, \ldots, F_n) = \mathfrak{c} \circ (F_1 \times \ldots \times F_n)$, where $F_1 \times \ldots \times F_n : \mathbb{O}^m \to \mathbb{1}^n$ is defined by $q \mapsto \left(F_1(q), \ldots, F_n(q) \right)$. Also, every predicate variable or predicate constant, $\varphi : \mathbb{O}^n \to \mathbb{1}$, can be lifted to an expression $\varphi_m : (\mathbb{O}^m \to \mathbb{O})^n \to (\mathbb{O}^m \to \mathbb{1})$, so as to produce m-formulas out of m-terms, via $\mathfrak{H}_m : (\mathbb{O}^n \to \mathbb{1}) \to \left((\mathbb{O}^m \to \mathbb{O})^n \to (\mathbb{O}^m \to \mathbb{1}) \right)$, which is defined similarly to $\mathfrak{G}_m$. Finally, every formula quantifier, $\mathfrak{q} : (\mathbb{O} \to \mathbb{1}) \to \mathbb{1}$, can be made into an m-formula quantifier $\mathfrak{q}_m : (\mathbb{O}^{m+1} \to \mathbb{1}) \to (\mathbb{O}^m \to \mathbb{1})$ via $\mathfrak{K}_m : \left((\mathbb{O} \to \mathbb{1}) \to \mathbb{1} \right) \to \left((\mathbb{O}^m \times \mathbb{O} \to \mathbb{1}) \to (\mathbb{O}^m \to \mathbb{1}) \right)$, defined by $q \mapsto \left(f \mapsto \left(r \mapsto q(\mathfrak{F}(f)(r)) \right) \right)$, where $\mathfrak{F}$ is as in Example 4.11. Thus $\mathfrak{q}_m(y_1, \ldots, y_m, y_{m+1} \mapsto F) = y_1, \ldots, y_m \mapsto \mathfrak{q}(y_{m+1} \mapsto F)$. Let us note that interpretations of formula connectives and quantifiers induce, by using the analogues of $\mathfrak{G}_m$, $\mathfrak{H}_m$ and $\mathfrak{K}_m$, interpretations of m-formula connectives, predicate constants/variables and quantifiers, which are, respectively, by functions $\mathrm{Hom}(\mathscr{D}^m, \mathscr{O})^n \to \mathrm{Hom}(\mathscr{D}^m, \mathscr{O})$, $\mathrm{Hom}(\mathscr{D}^m, \mathscr{D})^n \to \mathrm{Hom}(\mathscr{D}^m, \mathscr{O})$ and $\mathrm{Hom}(\mathscr{D}^{m+1}, \mathscr{O}) \to \mathrm{Hom}(\mathscr{D}^m, \mathscr{O})$.

A *predicate valuation* v assigns a function $v(\varphi)\colon \mathscr{D}^m \to \mathcal{O}$ to each m-ary predicate variable φ, for $m = 0, 1, 2, \ldots$. (This will also be called simply "valuation", especially in the context of intuitionistic logic, where the "predicate variables" are really problem variables.) Traditionally valuations are found only in zero-order logic, since modern treatments of first-order logic do not include predicate variables in the language.

Given a predicate valuation v, for every closed formula F we get a specific element $|F|_i^v \in \mathcal{O}$; and, more generally, for every closed n-formula F a specific function $|F|_i^v\colon \mathscr{D}^n \to \mathcal{O}$. This function is defined straightforwardly using the previous recursive description of closed n-formulas, and it is clear from this definition that it depends only on the α-equivalence class of F. Let us note that if F is λ-closed, then $|F|_i := |F|_i^v$ does not depend on v.

Given a variable assignment u, every predicate valuation v also yields, for each $n = 0, 1, \ldots$, a function $|\cdot|_i^{uv}\colon \mathscr{F}_n \to \mathrm{Hom}(\mathscr{D}^n, \mathcal{O})$, where $\mathscr{F}_n = \mathscr{F}_n(\mathscr{L})$ is the set of α-equivalence classes of arbitrary n-formulas of $\mathscr{L}$. Thus for every formula F we get a specific element $|F|_i^{uv} \in \mathcal{O}$, well-defined up to α-equivalence of F.

It is often convenient to consider expressions of the form $\varphi_1, \ldots, \varphi_l \mapsto F$, where F is an n-formula and the φ_i are pairwise distinct predicate variables of (not necessarily distinct) arities $r_1, \ldots, r_m$. We will call such an expression an $(n; r_1, \ldots, r_m)$-*formula*, or an $(n, \vec{r})$-*formula*, where $\vec{r} = (r_1, \ldots, r_m)$; it is of type $\prod_{i=1}^m (\mathbb{0}^{r_i} \to \mathbb{1}) \to (\mathbb{0}^n \to \mathbb{1})$. An $(n, \vec{r})$-formula is called *closed* if no individual variables occur freely in it, and λ-*closed* if no variables occur freely in it. Similarly to the above, one can define lifted connectives, predicate variables/constants and quantifiers for $(n, \vec{r})$-formulas and use them to give an inductive construction of $(n, \vec{r})$-formulas. Given a predicate valuation v and a variable assignment u, an arbitrary $(n; r_1, \ldots, r_m)$-formula G is interpreted, using this inductive construction, by a function $|G|_i^{uv}\colon \prod_{i=1}^m \mathrm{Hom}(\mathscr{D}^{r_i}, \mathcal{O}) \to \mathrm{Hom}(\mathscr{D}^n, \mathcal{O})$, which depends only on the α-equivalence class of G. If G is closed, then $|G|_i^v := |G|_i^{uv}$ does not depend on u; and if G is λ-closed, then $|G|_i := |G|_i^v$ does not depend on v either.

Remark 4.23. Let us note that λ-closed $(n, \vec{r})$-formulas are similar to traditional schemata. Indeed, term variables and r_i-ary formula variables that occur in a schema can be considered as "blanks" that may be filled in by terms and by r_i-formulas; to be precise, each blank must be highlighted by a marker so that blanks of the same color are to be filled in by the same term or r_i-formula. Let us note, however, that these "blanks" themselves do not correspond to any *variables* in the sense of λ-calculus, but rather to expressions of the form $v \mapsto v$, where v is an individual variable or a predicate variable.

4.3.4. Meta-formulas. We will now define the extension of the language $\mathscr{L}$ by the *language of meta-logic* — not to be confused with a meta-language of $\mathscr{L}$ (which we do not need to treat formally since what is usually expressed by meta-variables in schemata will be achieved by means of the language of meta-logic). The additional ingredients of the extension of $\mathscr{L}$ are no longer specific to one's choice of a logic (and for this reason we need no separate notation for the extension); instead, they will be used to express meta-problems (or rather their Platonic "ideas"/Fregean "senses"). Just like propositions and

problems without parameters are interpreted by elements of the set $\mathcal{O}$, meta-problems without parameters will be interpreted by elements of another set $\mathcal{Q}$. We will further assume that there is a "reflection" function $! : \mathcal{O} \to \mathcal{Q}$ converting a proposition into a meta-problem asking one to prove it (or a problem into a meta-problem asking one to solve this problem); and a "truth value" (or "meta-reflection") function $? : \mathcal{Q} \to \{\top, \bot\!\!\bot\}$ converting a meta-problem into a judgement (=meta-meta-proposition) asserting the solubility of this meta-problem. The choice of $\mathcal{Q}$, $!$ and $?$ is included, along with a choice of an $\mathscr{L}$-structure i and some further data to be discussed in a moment, in a *meta-$\mathscr{L}$-structure* j, also called a *meta-interpretation* of $\mathscr{L}$.

We will see in §*4.7.1* that the usual, Tarski-style model theory corresponds to the straightforward *two-valued meta-interpretation* $j = i_+$, which sets $\mathcal{Q} = \{\top, \bot\!\!\bot\}$ and $? = \mathrm{id}$ no matter what $\mathcal{O}$ is (i.e., meta-problems are postulated to have no observable "sense" beyond their "reference"). However, we will also see in §7 that, in fact, a more interesting option is contained in Kolmogorov's implicit problem interpretation of the *meta-logic* of intuitionistic logic [108]. Thus we leave $\mathcal{Q}$ unspecified at this point.

Now this all is semantics (or rather meta-semantics); to represent it syntactically, we need yet another base type, to be denoted μ, along with a constant $! : \mathbb{1} \to \mu$. Thus $!$ is supposed to be an interpretation of $!$. Now any formula can be regarded, via $!$, as an expression of type μ. In fact, this will be essentially our only source of "elementary" expressions of type μ, since we are not going to use any variables of type μ. Thus, we can already define an *atomic meta-formula* as an expression of the form $!(F)$, where F is an arbitrary formula.

Arbitrary *meta-formulas* are built out of atomic meta-formulas using only the following constants: two *meta-connectives* and two varieties of a *meta-quantifier* (the universal one). The meta-connectives are the *meta-conjunction* $\& : \mu \times \mu \to \mu$, and the *meta-implication* $\Rightarrow : \mu \times \mu \to \mu$.[50] Their interpretations $|\&|_j : \mathcal{Q} \times \mathcal{Q} \to \mathcal{Q}$ and $|\Rightarrow|_j : \mathcal{Q} \times \mathcal{Q} \to \mathcal{Q}$ are supposed to express conjunction and implication between judgements. Under the two-valued meta-interpretation these are given by the usual truth tables, i.e., $|\&|_{i_+}(q, r) = \top$ if and only if $q = r = \top$, and $|\Rightarrow|_{i_+}(q, r) = \bot\!\!\bot$ if and only if $q = \top$ and $r = \bot\!\!\bot$. Then there is the *first-order meta-quantifier* $\mathfrak{q} : (\mathbb{O} \to \mu) \to \mu$ and the *n-ary second-order meta-quantifier* $\mathfrak{q}_n : ((\mathbb{O}^n \to \mathbb{1}) \to \mu) \to \mu$ for each $n = 0, 1, 2, \ldots$. The first-order meta-quantifier $\mathfrak{q}$ is interpreted by a function $|\mathfrak{q}|_j : \mathrm{Hom}(\mathscr{D}, \mathcal{Q}) \to \mathcal{Q}$; and the n-ary second-order meta-quantifier $\mathfrak{q}_n$ by a function $|\mathfrak{q}_n|_j : \mathrm{Hom}(\mathrm{Hom}(\mathscr{D}^n, \mathcal{O}), \mathcal{Q}) \to \mathcal{Q}$. Under the two-valued meta-interpretation these behave like the classical universal quantifier, i.e., $|\mathfrak{q}|_{i_+}(f) = \top$ if and only $f(x) = \top$ for all $x \in \mathscr{D}$, and $|\mathfrak{q}_n|_{i_+}(f) = \top$ if and only $f(\varphi) = \top$ for all $\varphi \in \mathrm{Hom}(\mathscr{D}^n, \mathcal{O})$.

---

[50]The meta-absurdity, the meta-disjunction and the existential meta-quantifier are not really needed, but will eventually be discussed as a side remark (see §*4.6.3*). All of these, as well as the meta-conjunction, would be definable (see §*4.6.3*) in terms of quantification over variables of type μ (which we do not include in our meta-logic, in contrast to Paulson).

In practice, meta-quantifiers will be written like the old-style (early 20th century) universal quantifiers,[51] but with fancy parentheses $(\cdot)$ so as to avoid visual confusion with the ordinary parentheses $(\cdot)$. Namely, if $\mathfrak{q} : (\Delta \to \upmu) \to \upmu$ is a meta-quantifier (either of them), $\mathcal{F}$ is an expression of type $\upmu$, and x is a variable of type Δ, then $(x)\, \mathcal{F}$ abbreviates the expression $\mathfrak{q}(x \mapsto \mathcal{F})$. Here x is said to be *bound* by the meta-quantifier, and $\mathcal{F}$ is called the *scope* of (the occurrence of) the meta-quantifier. Moreover, $(x_0, \ldots, x_n)\, \mathcal{F}$ abbreviates $(x_0)\, ((x_1, \ldots, x_n)\, \mathcal{F})$, where the x_i may be of different types.

This completes our definition of the meta-logical extension of $\mathscr{L}$: it consists of the constants $\mathbb{!}$, $\&$, $\Rightarrow$, $\mathfrak{q}$ and $\mathfrak{q}_0, \mathfrak{q}_1, \ldots$. Since these are fixed forever (in the present paper), the set of meta-formulas can be reconstructed from $\mathscr{L}$ alone.

The definition of a meta-$\mathscr{L}$-structure $\jmath$ is also complete: it consists of an $\mathscr{L}$-structure $\imath$, the set $\mathcal{Q}$, the maps $!$ and $?$, and two elements $|\&|_{\jmath}$, $|{\Rightarrow}|_{\jmath}$ of $\mathrm{Hom}(\mathcal{Q} \times \mathcal{Q}, \mathcal{Q})$, an element $|\mathfrak{q}|_{\jmath}$ of $\mathrm{Hom}(\mathrm{Hom}(\mathscr{D}, \mathcal{Q}), \mathcal{Q})$ and an element $|\mathfrak{q}_n|_{\jmath}$ of $\mathrm{Hom}(\mathrm{Hom}(\mathrm{Hom}(\mathscr{D}^n, \mathcal{O}), \mathcal{Q}), \mathcal{Q})$ for each $n \in \mathbb{N}$.

When no confusion arises, we will omit the symbol $\mathbb{!}$ in writing out expressions of type $\upmu$. Accordingly, formulas will be regarded as special (namely, atomic) meta-formulas. However, one should keep in mind that meta-connectives and meta-quantifiers cannot be used *inside* of a formula.

As usual, $\mathcal{F} \Leftrightarrow \mathcal{G}$ abbreviates $(\mathcal{F} \Rightarrow \mathcal{G})\, \&\, (\mathcal{G} \Rightarrow \mathcal{F})$; "$\Leftrightarrow$" is called *meta-equivalence*. We will stick to the following order of precedence of logical and meta-logical symbols (in groups of equal priority, starting with higher precedence/stronger binding):

(1) $\neg$, $\exists$ and $\forall$;
(2) $\wedge$ and $\vee$;
(3) $\to$ and $\leftrightarrow$;
(4) $(\cdot)$;
(5) $\&$;
(6) $\Rightarrow$ and $\Leftrightarrow$.

By an *n-meta-formula* we mean an expression of the form $x_1, \ldots, x_n \mapsto \mathcal{F}$, where the x_i are pairwise distinct individual variables and $\mathcal{F}$ is a meta-formula. By an $(n;\, r_1, \ldots, r_m)$-*meta-formula*, or an $(n, \vec{r})$-*meta-formula*, where $\vec{r} = (r_1, \ldots, r_m)$, we mean an expression of the form $\varphi_1, \ldots, \varphi_m \mapsto \mathcal{G}$, where the φ_i are pairwise distinct predicate variables of (not necessarily distinct) arities $n_1, \ldots, n_m$, and $\mathcal{G}$ is an n-meta-formula. Thus an $(n, \vec{0})$-meta-formula (where $\vec{0}$ is the unique 0-tuple) is an n-meta-formula, and a 0-meta-formula is a meta-formula. Let us call an $(n, \vec{r})$-meta-formula *closed* if individual variables do not occur freely in it, and λ-*closed* in no variables occur freely in it.

Like before, one can lift the meta-connectives and the varieties of the meta-quantifier to the level of $(n, \vec{r})$-meta-formulas, and use the lifts to give an inductive construction of $(n, \vec{r})$-meta-formulas. (Let us note that for a meta-formula $\mathcal{F}$, such a recursive description will involve $(n;\, r_1, \ldots, r_m)$-meta-formulas with $n > 0$ if $\mathcal{F}$ contains either a

---

[51]Peano's notation for universal quantification, superseded by Gentzen's analogue $\forall$ of Peano's $\exists$.

meta-level quantifier or a first-order meta-quantifier, and with $m > 0$ if $\mathcal{F}$ contains a second-order meta-quantifier.)

Using this inductive construction, every $(n; r_1, \ldots, r_m)$-meta-formula $\mathcal{F}$ along with a variable assignment u and a predicate valuation v are easily seen to yield a function $|\mathcal{F}|_{\mathcal{J}}^{uv} \colon \prod_{i=1}^m \mathrm{Hom}(\mathscr{D}^{r_i}, \mathcal{Q}) \to \mathrm{Hom}(\mathscr{D}^n, \mathcal{Q})$. If $\mathcal{F}$ is closed, then this function, written $|\mathcal{F}|_{\mathcal{J}}^{v}$, does not depend on u; and if $\mathcal{F}$ is λ-closed, then the function, written $|\mathcal{F}|_{\mathcal{J}}$, also does not depend on v.

In particular, a λ-closed meta-formula $\mathcal{F}$ is interpreted by a specific element $|\mathcal{F}|_{\mathcal{J}} \in \mathcal{Q}$. A closed meta-formula $\mathcal{F}$ along with a predicate valuation v yield an element $|\mathcal{F}|_{\mathcal{J}}^{v} \in \mathcal{Q}$. Given a predicate valuation v and a variable assignment u, an arbitrary meta-formula $\mathcal{F}$ is interpreted by an element $|\mathcal{F}|_{\mathcal{J}}^{uv} \in \mathcal{Q}$.

Example 4.24. Let us parse the following λ-closed meta-formula:

$$(\gamma) \ \gamma \vee \neg\gamma.$$

We have omitted the reflection function, and here it is shown:

$$(\gamma) \ \lfloor\!\lfloor(\gamma \vee \neg\gamma).$$

Let us also expand the meta-quantifier, so as to make the λ-abstraction explicit:

$$\mathbb{q}_0\big(\gamma \mapsto \lfloor\!\lfloor(\gamma \vee \neg\gamma)\big).$$

Now let us assume that some meta-interpretation $\mathcal{J}$ is given. Then the λ-closed $(0;0)$-formula $\gamma \mapsto \gamma \vee \neg\gamma \colon \mathbb{1} \to \mathbb{1}$ is interpreted by a function $\mathcal{O} \to \mathcal{O}$. Therefore $\gamma \mapsto \lfloor\!\lfloor(\gamma \vee \neg\gamma) \colon \mathbb{1} \to \mathbb{\mu}$ is interpreted by a function $\mathcal{O} \to \mathcal{Q}$. Hence $\mathbb{q}_0\big(\gamma \mapsto \lfloor\!\lfloor(\gamma \vee \neg\gamma)\big) \colon \mathbb{\mu}$ is interpreted by a specific element of $\mathcal{Q}$.

Example 4.25. Let us parse the following λ-closed meta-formula:

$$(\gamma) \ \big((\mathbb{x}) \ \gamma(\mathbb{x}) \Rightarrow \forall\mathbb{x}\,\gamma(\mathbb{x})\big).$$

We have omitted the reflection functions, and here they are shown:

$$(\gamma) \ \Big((\mathbb{x}) \ \lfloor\!\lfloor\big(\gamma(\mathbb{x})\big) \Rightarrow \lfloor\!\lfloor\big(\forall\mathbb{x}\,\gamma(\mathbb{x})\big)\Big).$$

Let us also expand the meta-quantifiers, so as to make the λ-abstractions explicit:

$$\mathbb{q}_1\left(\gamma \mapsto \left(\mathbb{q}\Big(\mathbb{x} \mapsto \lfloor\!\lfloor\big(\gamma(\mathbb{x})\big)\Big) \Rightarrow \lfloor\!\lfloor\big(\forall\mathbb{x}\,\gamma(\mathbb{x})\big)\right)\right).$$

Now let us assume that some meta-interpretation $\mathcal{J}$ is given. Then the λ-closed $(0;1)$-formula $\gamma \mapsto \forall\mathbb{x}\,\gamma(\mathbb{x}) \colon (\mathbb{0} \to \mathbb{1}) \to \mathbb{1}$ is interpreted by a function $\mathrm{Hom}(\mathscr{D}, \mathcal{O}) \to \mathcal{O}$ (namely, the interpretation of $\forall$), and the λ-closed $(1;1)$-formula $\gamma \mapsto \big(\mathbb{x} \mapsto \gamma(\mathbb{x})\big) \colon (\mathbb{0} \to \mathbb{1}) \to (\mathbb{0} \to \mathbb{1})$ is interpreted by a function $\mathrm{Hom}(\mathscr{D}, \mathcal{O}) \to \mathrm{Hom}(\mathscr{D}, \mathcal{O})$ (namely, the identity function). Hence $\gamma \mapsto \lfloor\!\lfloor\big(\forall\mathbb{x}\,\gamma(\mathbb{x})\big) \colon (\mathbb{0} \to \mathbb{1}) \to \mathbb{\mu}$ is interpreted by a function $\mathrm{Hom}(\mathscr{D}, \mathcal{O}) \to \mathcal{Q}$, and $\gamma \mapsto \big(\mathbb{x} \mapsto \lfloor\!\lfloor\big(\gamma(\mathbb{x})\big)\big) \colon (\mathbb{0} \to \mathbb{1}) \to (\mathbb{0} \to \mathbb{\mu})$ is interpreted by a function $\mathrm{Hom}(\mathscr{D}, \mathcal{O}) \to \mathrm{Hom}(\mathscr{D}, \mathcal{Q})$. Next, $\gamma \mapsto \mathbb{q}\big(\mathbb{x} \mapsto \lfloor\!\lfloor\big(\gamma(\mathbb{x})\big)\big) \colon (\mathbb{0} \to \mathbb{1}) \to \mathbb{\mu}$ is

interpreted by a function $\text{Hom}(\mathscr{D}, \mathscr{O}) \to \mathcal{Q}$, and consequently

$$\gamma \mapsto \left(\mathfrak{q}\Big(\mathsf{x} \mapsto \text{!`}\big(\gamma(\mathsf{x})\big) \Big) \Rightarrow \text{!`}\big(\forall \mathsf{x}\, \gamma(\mathsf{x})\big) \right) : (\mathbb{0} \to \mathbb{1}) \to \mathbb{\mu}$$

is also interpreted by a function $\text{Hom}(\mathscr{D}, \mathscr{O}) \to \mathcal{Q}$. Finally, the meta-formula in question is interpreted by the interpretation of $\mathfrak{q}_1$ applied to the latter function; this is a specific element of $\mathcal{Q}$.

By a *principle* we mean a λ-closed meta-formula involving no meta-connectives. In other words, a principle is a formula meta-quantified over all its free individual variables and over all its predicate variables. A meta-formula will be called *purely logical* if it involves no function symbols and no predicate constants.

Example 4.26. Fermat's theorem is a formula, like most other mathematical theorems that are expressible in a first-order language, whereas the principle of excluded middle is a purely logical principle, like most other logical principles that are expressible in a first-order language. There are important first-order theories (such as Peano arithmetic, Tarski's elementary geometry and Zermelo–Fraenkel set theory) where most axioms and theorems are formulas, but some are principles, though none are purely logical (e.g. the principle of mathematical induction, Tarski's continuity principle and the set-theoretic principles of separation and replacement).

A meta-formula $\mathcal{F}$ is said to be *valid* in a meta-$\mathscr{L}$-structure $\jmath$ if $\text{?}(|\mathcal{F}|_{\jmath}^{uv}) = \mathbb{T}$ for arbitrary variable assignment u and predicate valuation v. When $\mathcal{F}$ is closed, this is equivalent to saying that $\text{?}(|\mathcal{F}|_{\jmath}^{v}) = \mathbb{T}$ for an arbitrary predicate valuation v; and when $\mathcal{F}$ is λ-closed, this is equivalent to saying that $\text{?}(|\mathcal{F}|_{\jmath}) = \mathbb{T}$.

A meta-formula $\mathcal{F}$ is said to be *valid* in an $\mathscr{L}$-structure $\imath$ if it is valid in the two-valued meta-interpretation $\imath_+$. We also call $\mathcal{F}$ *valid* in the pair $(\imath, v)$ if $\text{?}(|\mathcal{F}|_{\imath_+}^{uv}) = \mathbb{T}$ for arbitrary variable assignment u. Dually, $\mathcal{F}$ *satisfiable* in the pair $(\imath, v)$ if $\text{?}(|\mathcal{F}|_{\imath_+}^{uv}) = \mathbb{T}$ for some variable assignment u.

We denote by $\vDash \mathcal{F}$, or in more detail $\vDash_{\jmath} \mathcal{F}$, the judgement that the meta-formula $\mathcal{F}$ is valid in $\jmath$. The judgement $\vDash_{\jmath} (\mathcal{F}_1 \,\&\, \cdots \,\&\, \mathcal{F}_m \Rightarrow \mathcal{G})$ is also abbreviated by

$$\mathcal{F}_1, \ldots, \mathcal{F}_m \vDash_{\jmath} \mathcal{G}.$$

When this judgement is true, we also say that $\mathcal{G}$ is a *semantic consequence* of the $\mathcal{F}_i$ in the meta-$\mathscr{L}$-structure $\jmath$. When $\jmath = \imath_+$, we abbreviate $\vDash_{\imath_+}$ by $\vDash_{\imath}$, and speak of *semantic consequence* in the $\mathscr{L}$-structure $\imath$.[52]

4.4. Meta-logic

Our meta-logic is in essence a second-order fragment of `Isabelle`'s meta-logic [157], which is in turn essentially a fragment of higher-order intuitionistic logic.[53]

---

[52]When the $\mathcal{F}_i$ and $\mathcal{G}$ are formulas, this is Tarski's standard definition of semantic consequence [193].
[53]Our meta-logic only deals with meta-formulas, whereas `Isabelle`'s with arbitrary expressions of types generated from the base types $\mathbb{0}$, $\mathbb{1}$, $\mathbb{\mu}$ using the type constructor $\to$. `Isabelle`'s meta-logic additionally

One need not panic about the fact that the meta-logic is intuitionistic and not classical (just like with λ-calculus, which is closely related to a fragment of intuitionistic logic). Firstly, when arguing about meta-logic, we have to do it in some meta-meta-logic (along with some meta-meta-theory over it), and this one will be classical (and informal) in our approach. Secondly, the semantics of our meta-logic according to the standard (Tarski's) setup of model theory turns out to be classical. On the other hand, it is not unexpected that our meta-logic is intuitionistic: this simply reflects the fact that derivation of formulas is a constructive activity; thus the primitive notion is that of the *problem* (or task) of deriving a given formula, and the more familiar notion of derivability of the formula is secondary: it merely asserts the existence of a solution of the said problem. One can discuss semantics of the metalogic without degenerating into classical logic (we will do it in §7), but one does not have to do it before discussing semantics of an object logic, be it intuitionistic, classical or modal. Thus by defining a logic in terms of meta-logic we will not complicate our task of understanding intuitionistic logic; on the contrary, this will simplify matters, as it is always easier to understand something that has a formal definition.

4.4.1. Meta-rules of inference. There are just two meta-rules of inference for each meta-connective: the introduction rule and the elimination rule.[54] In the meta-rules below, $\mathcal{F}$ and $\mathcal{G}$ are meta-formulas, x is an individual variable, T is a term, γ is an n-ary predicate variable, and Σ is an n-formula, for some $n = 0, 1, 2, \ldots$.[55] The meta-rules are stated in the natural deduction style, which means that they operate on deductions from assumptions. In particular, the $\Rightarrow$-introduction meta-rule below is understood thus: starting from a deduction of $\mathcal{G}$ from a list of assumptions that includes $\mathcal{F}$, we pass to a deduction of $\mathcal{F} \Rightarrow \mathcal{G}$ from the same list of assumptions but with $\mathcal{F}$ removed. (If some other entries of the list are identical with $\mathcal{F}$, they are kept.) See [64], [164] for further details on natural deduction (but it should be clear how it works from the examples below).

---

includes quantification over variables of arbitrary types (of which only two are needed for dealing with first-order logics, so we treat them separately), equality, λ-abstraction and function application (we prefer to leave these at the informal meta-meta-level, since their explicit treatment does not seem to add much conceptual clarity in our case). Due to our omission of `Isabelle`'s meta-rules for equality, our meta-logic is not capable of dealing with first-order logics with equality — although some theories with equality can be handled. (The author is unconvinced that this is a deficiency since modern type theories provide an arguably more adequate framework for dealing with equality than first-order logic.) Due to our omission of unrestricted quantification, Russell's definitions of connectives (see §*4.6.3*) no longer apply, and we find it convenient to add the meta-conjunction to our meta-logic (though it could have been eliminated in principle by using the exponential law).

[54]These meta-rules themselves belong to meta-meta-logic, which we acknowledge by drawing bold horizontal lines.

[55]It should be noted that this sentence sneaks in meta-meta-quantification over $\mathcal{F}$, $\mathcal{G}$, x, T, γ and Σ. One does not need to worry about it too much, because all substantial issues appear to be already addressed on the previous two levels. In particular, the meta-rules for ($\cdot$) take care of the entire business of capture-avoiding substitutions. This provides an opportunity to focus on more "substantial" issues at the object level (in a traditional sense that is ignorant of the interpretation of substitution by pullback).

$$(\alpha\text{-conversion}) \quad \begin{array}{c} \vdots \\ \mathcal{F} \\ \hline \mathcal{G} \end{array}, \text{ if } \mathcal{F} \text{ is } \alpha\text{-equivalent to } \mathcal{G}$$

$$(\&\text{-introduction}) \quad \begin{array}{cc} \vdots & \vdots \\ \mathcal{F} & \mathcal{G} \\ \hline \multicolumn{2}{c}{\mathcal{F} \,\&\, \mathcal{G}} \end{array}$$

$$(\&\text{-elimination}) \quad \begin{array}{c} \vdots \\ \mathcal{F} \,\&\, \mathcal{G} \\ \hline \mathcal{F} \end{array} \quad \text{and} \quad \begin{array}{c} \vdots \\ \mathcal{F} \,\&\, \mathcal{G} \\ \hline \mathcal{G} \end{array}$$

$$(\Rightarrow\text{-elimination}) \quad \begin{array}{cc} \vdots & \vdots \\ \mathcal{F} & \mathcal{F} \Rightarrow \mathcal{G} \\ \hline \multicolumn{2}{c}{\mathcal{G}} \end{array}$$

$$(\Rightarrow\text{-introduction}) \quad \begin{array}{c} [\mathcal{F}] \\ \vdots \\ \mathcal{G} \\ \hline \mathcal{F} \Rightarrow \mathcal{G} \end{array}$$

(meta-generalization)

$$\begin{array}{c} \vdots \\ \mathcal{F} \\ \hline (x)\,\mathcal{F} \end{array}, \text{ provided that } x \text{ does not occur freely in any of the assumptions;}$$

$$\begin{array}{c} \vdots \\ \mathcal{F} \\ \hline (\gamma)\,\mathcal{F} \end{array}, \text{ provided that } \gamma \text{ does not occur freely in any of the assumptions.}$$

$$(\text{meta-specialization}) \quad \begin{array}{c} \vdots \\ (x)\,\mathcal{F} \\ \hline \mathcal{F}[x/T] \end{array}, \text{ provided that } T \text{ is free for } x \text{ in } \mathcal{F};$$

$$\begin{array}{c} \vdots \\ (\gamma)\,\mathcal{F} \\ \hline \mathcal{F}[\gamma/\Sigma] \end{array}, \text{ provided that } \Sigma \text{ is free for } \gamma \text{ in } \mathcal{F}.$$

Let us note that $\mathcal{F}[x/T]$ has the same meaning as $\mathcal{F}|_{x:=T}$ in the first-order meta-specialization rule, and if $n = 0$, then also $\mathcal{F}[\gamma/\Sigma]$ has the the same meaning as $\mathcal{F}|_{\gamma:=\Sigma}$ in the second-order meta-specialization rule; but not so for $n > 0$, since in that case the λ-expression $\mathcal{F}|_{\gamma:=\Sigma}$ is not itself a meta-formula, and only β-reduces to the meta-formula $\mathcal{F}[\gamma/\Sigma]$.

One could in fact drop the "free for" hypotheses and understand the meta-specialization rules as enabling one to pass from a given deduction of $(x)\,\mathcal{F}$ or $(\gamma)\,\mathcal{F}$ to the deduction of any meta-formula in the α-equivalence class to which $(x \mapsto \mathcal{F})(T)$ or $(\gamma \mapsto \mathcal{F})(\Sigma)$ β-reduces. This seems reasonable, since we anyway have the α-conversion meta-rule. Nevertheless, the "free for" conditions are mentioned above in order to facilitate comparison with traditional treatments of first-order logics.

The "free for" conditions are defined in §*4.2.5*. Let us write them out explicitly in our situation to see that they are similar to the usual ones used in first-order logic (cf. e.g. Kleene [103; §18, §34], Church [32; §35]): only the clause (3) below is unusual, but it is analogous to (0).

Namely, that T is *free for x in $\mathcal{F}$* means that

(0) no free occurrence of x in $\mathcal{F}$ is in the scope of a quantifier or a meta-quantifier over any individual variable y that occurs in T.

And that the n-formula $\Sigma = x_1, \ldots, x_n \mapsto \Theta$ [56] is *free for γ in $\mathcal{F}$* means that

(1) every free occurrence of γ in $\mathcal{F}$ is in an atomic formula of the form $\gamma(T_1, \ldots, T_n)$ such that each T_i is free for x_i in Θ;

(2) no free occurrence of γ in $\mathcal{F}$ is in the scope of a quantifier or a meta-quantifier over any individual variable y that occurs freely in Σ;

(3) no free occurrence of γ in $\mathcal{F}$ is in the scope of a meta-quantifier over any predicate variable δ that occurs in Σ.

From the viewpoint of λ-calculus, conditions (2) and (3) are checked upon the substitution of Σ for γ; whereas (1) is checked upon the evaluation of the resulting expression, which involves an auxiliary substitution.

The point of the α-conversion rule is that object-level inference rules and laws such as $(\alpha, t)\ \alpha(t) \to \exists x\, \alpha(x)$ are now meta-formulas (and not just some schemata involving meta-variables) and so do not automatically yield their copies with different names of bound variables. Variables bound by meta-quantifiers can be easily renamed using the specialization and generalization meta-rules; but no such procedure exists for variables bound by usual quantifiers.

It is not hard to see that it suffices to assume the α-conversion rule only for atomic meta-formulas (i.e. the ♮-images of formulas). Indeed, an arbitrary meta-formula is built out of atomic ones using meta-connectives and meta-quantifiers; using the meta-rules of inference one can get rid of the meta-connectives and the meta-quantifiers, at the cost of

---

[56]Let us note that we cannot just write $\Sigma(x_1, \ldots, x_n)$ instead of Θ, because that does not exclude the unwanted situation that some x_i occurs freely in Σ.

multiplying the number of deductions, adding new assumptions, and remembering that some variables do not occur freely in the assumptions.

4.4.2. Examples of deductions. The only deductions that are provided from start are the *trivial* deductions, in which a meta-formula is deduced from itself. A meta-formula $\mathcal{F}$ is called *deducible* if using the meta-rules one can obtain (from the trivial deductions) a deduction of $\mathcal{F}$ from no assumptions. We will also abbreviate the phrase "to obtain (from the trivial deductions, using the meta-rules) a deduction of $\mathcal{F}$ from the assumption $\mathcal{G}$" by "to deduce $\mathcal{F}$ from $\mathcal{G}$".

Example 4.27. Starting from the trivial deduction of $\mathcal{F}$ from itself, we can apply the $\Rightarrow$-introduction meta-rule to obtain a deduction of $\mathcal{F} \Rightarrow \mathcal{F}$ from no assumptions. Thus the meta-formula $\mathcal{F} \Rightarrow \mathcal{F}$ is deducible.

Example 4.28. Starting from the trivial deduction of $\mathcal{G}$ from itself, we may view it as a deduction of $\mathcal{G}$ from the assumptions $\mathcal{F}$ and $\mathcal{G}$. Then the $\Rightarrow$-introduction yields a deduction of $\mathcal{F} \Rightarrow \mathcal{G}$ from the assumption $\mathcal{G}$. Applying the $\Rightarrow$-elimination once again, we get that $\mathcal{G} \Rightarrow (\mathcal{F} \Rightarrow \mathcal{G})$ is deducible. Also, in the case where $\mathcal{G}$ is of the form $\mathcal{H} \Rightarrow \mathcal{H}$, we can combine the deduction of $\mathcal{F} \Rightarrow \mathcal{G}$ from the assumption $\mathcal{G}$ with Example 4.27 to get that $\mathcal{F} \Rightarrow (\mathcal{H} \Rightarrow \mathcal{H})$ is deducible.

Example 4.29. Let us show that $(\mathcal{H} \Rightarrow \mathcal{F}) \,\&\, (\mathcal{H} \Rightarrow \mathcal{G}) \implies (\mathcal{H} \Rightarrow \mathcal{F} \,\&\, \mathcal{G})$ is deducible. For this, it suffices to deduce $\mathcal{H} \Rightarrow \mathcal{F} \,\&\, \mathcal{G}$ from $(\mathcal{H} \Rightarrow \mathcal{F}) \,\&\, (\mathcal{H} \Rightarrow \mathcal{G})$. For this, it in turn suffices to deduce $\mathcal{F} \,\&\, \mathcal{G}$ from the hypotheses $\mathcal{H}$ and $(\mathcal{H} \Rightarrow \mathcal{F}) \,\&\, (\mathcal{H} \Rightarrow \mathcal{G})$. This is done as follows: we first deduce $\mathcal{H} \Rightarrow \mathcal{F}$ and (separately) $\mathcal{H} \Rightarrow \mathcal{G}$ from $(\mathcal{H} \Rightarrow \mathcal{F}) \,\&\, (\mathcal{H} \Rightarrow \mathcal{G})$; and use the $\Rightarrow$-elimination meta-rule to deduce $\mathcal{F}$ from $\mathcal{H}$ and $\mathcal{H} \Rightarrow \mathcal{F}$, and $\mathcal{G}$ from $\mathcal{H}$ and $\mathcal{H} \Rightarrow \mathcal{G}$. This gives us deductions of $\mathcal{F}$ and $\mathcal{G}$ from the hypotheses $\mathcal{H}$ and $(\mathcal{H} \Rightarrow \mathcal{F}) \,\&\, (\mathcal{H} \Rightarrow \mathcal{G})$; thus we can use the $\&$-introduction meta-rule to obtain the desired deduction.

Example 4.30. Similarly to Example 4.29, $(x)\big(\mathcal{H} \Rightarrow \mathcal{F}\big) \implies \big(\mathcal{H} \Rightarrow (x)\,\mathcal{F}\big)$ is deducible, where x is an individual or predicate variable that does not occur freely in $\mathcal{H}$.

Example 4.31. Let us deduce $(\mathcal{F} \,\&\, \mathcal{G} \Rightarrow \mathcal{H}) \iff \big(\mathcal{F} \Rightarrow (\mathcal{G} \Rightarrow \mathcal{H})\big)$, the exponential meta-law. By the $\&$-elimination, it suffices to deduce both implications: "$\Rightarrow$" and "$\Leftarrow$". For this it suffices to deduce $\mathcal{F} \Rightarrow (\mathcal{G} \Rightarrow \mathcal{H})$ from $\mathcal{F} \,\&\, \mathcal{G} \Rightarrow \mathcal{H}$ and vice versa.

In one direction, it suffices, by the $\Rightarrow$-introduction, to deduce $\mathcal{G} \Rightarrow \mathcal{H}$ from the hypotheses $\mathcal{F}$ and $\mathcal{F} \,\&\, \mathcal{G} \Rightarrow \mathcal{H}$. For this it suffices, by the $\Rightarrow$-introduction, to deduce $\mathcal{H}$ from $\mathcal{G}$, $\mathcal{G}$ and $\mathcal{F} \,\&\, \mathcal{G} \Rightarrow \mathcal{H}$. Now from the hypotheses $\mathcal{F}$ and $\mathcal{G}$ we get, by the $\&$-introduction, $\mathcal{F} \,\&\, \mathcal{G}$. From the latter and $\mathcal{F} \,\&\, \mathcal{G} \Rightarrow \mathcal{H}$ we get $\mathcal{H}$ by the $\Rightarrow$-elimination.

Conversely, it suffices, by the $\Rightarrow$-introduction, to deduce $\mathcal{H}$ from the assumptions $\mathcal{F} \,\&\, \mathcal{G}$ and $\mathcal{F} \Rightarrow (\mathcal{G} \Rightarrow \mathcal{H})$. Now $\mathcal{F} \,\&\, \mathcal{G}$ yields, by the $\&$-elimination, $\mathcal{F}$ and $\mathcal{G}$. From $\mathcal{F}$ and $\mathcal{F} \Rightarrow (\mathcal{G} \Rightarrow \mathcal{H})$ we get $\mathcal{G} \Rightarrow \mathcal{H}$ by the $\Rightarrow$-elimination. Similarly, from $\mathcal{G}$ and $\mathcal{G} \Rightarrow \mathcal{H}$ we get $\mathcal{H}$, as desired.

Example 4.32. It is also easy to check that $(x)\,\mathcal{F}\,\&\,(x)\,\mathcal{G} \iff (x)\,(\mathcal{F}\,\&\,\mathcal{G})$ is deducible and $(x)\,(\mathcal{F} \Rightarrow \mathcal{G}) \implies ((x)\,\mathcal{F} \Rightarrow (x)\,\mathcal{G})$ is deducible, where x is either an individual variable or a predicate variable.

Example 4.33. Given a deduction of $\mathcal{H}$ from $\mathcal{G}$ and a deduction of $\mathcal{G}$ from $\mathcal{F}$, by combining them we obtain a deduction of $\mathcal{H}$ from $\mathcal{F}$. Using this observation, it is easy to show that $(\mathcal{F} \Rightarrow \mathcal{G})\,\&\,(\mathcal{G} \Rightarrow \mathcal{H}) \Rightarrow (\mathcal{F} \Rightarrow \mathcal{H})$ is deducible.

Remark 4.34. If one can deduce $\mathcal{G}$ from $\mathcal{F}$, then by the $\Rightarrow$-introduction, $\mathcal{F} \Rightarrow \mathcal{G}$ is deducible, and hence by the $\Rightarrow$-elimination, the deducibility of $\mathcal{F}$ implies that of $\mathcal{G}$.

However, the converse is not true: if the deducibility of $\mathcal{F}$ implies that of $\mathcal{G}$, it does not follow that $\mathcal{G}$ can be deduced from $\mathcal{F}$. Indeed, it can be shown that the meta-absurdity $(\gamma)\,\gamma$ and the Peirce meta-principle $(\alpha, \beta)\left(((\alpha \Rightarrow \beta) \Rightarrow \alpha) \Rightarrow \alpha\right)$ are both not deducible; but still the former cannot be deduced from the latter.

4.4.3. Hilbert-style formulation. For purposes of models (see §*4.7.1*) it is convenient to record a "Hilbert-style" formulation of the meta-logic, consisting mostly of *meta-laws*, i.e. inference meta-rules with no hypotheses. To simplify notation, we omit the horizontal bar in meta-laws and the vertical dots illustrating the deduction. Like before, $\mathcal{F}$, $\mathcal{G}$ and $\mathcal{H}$ stand for arbitrary meta-formulas, x is an individual variable, T is a term, γ is an n-ary predicate variable, and Φ is an n-formula, for some $n = 0, 1, 2, \ldots$.

(1) $\mathcal{F} \Rightarrow \mathcal{G}$, if $\mathcal{F}$ is α-equivalent to $\mathcal{G}$;

(2) $\mathcal{F} \Rightarrow \mathcal{F}$;

(3) $(\mathcal{F} \Rightarrow \mathcal{G})\,\&\,(\mathcal{G} \Rightarrow \mathcal{H}) \implies (\mathcal{F} \Rightarrow \mathcal{H})$;

(4) $((\mathcal{F}\,\&\,\mathcal{G}) \Rightarrow \mathcal{H}) \iff (\mathcal{F} \Rightarrow (\mathcal{G} \Rightarrow \mathcal{H}))$;

(5) $\mathcal{F}\,\&\,\mathcal{G} \Rightarrow \mathcal{F}$ and $\mathcal{F}\,\&\,\mathcal{G} \Rightarrow \mathcal{G}$;

(6) $(\mathcal{H} \Rightarrow \mathcal{F})\,\&\,(\mathcal{H} \Rightarrow \mathcal{G}) \implies (\mathcal{H} \Rightarrow \mathcal{F}\,\&\,\mathcal{G})$;

(7) $(x)\,\mathcal{F} \Rightarrow \mathcal{F}[x/T]$, provided that T is free for x in $\mathcal{F}$;

(8) $(\gamma)\,\mathcal{F} \Rightarrow \mathcal{F}[\gamma/\Phi]$, provided that Φ is free for γ in $\mathcal{F}$;

(9) $(x)\,(\mathcal{H} \Rightarrow \mathcal{F}) \implies (\mathcal{H} \Rightarrow (x)\,\mathcal{F})$, provided that x does not occur freely in $\mathcal{H}$;

(10) $(\gamma)\,(\mathcal{H} \Rightarrow \mathcal{F}) \implies (\mathcal{H} \Rightarrow (\gamma)\,\mathcal{F})$, provided that γ does not occur freely in $\mathcal{H}$;

(11) $\dfrac{\mathcal{F}}{(x)\,\mathcal{F}}$, provided that x does not occur freely in the assumptions;

(12) $\dfrac{\mathcal{F}}{(\gamma)\,\mathcal{F}}$, provided that γ does not occur freely in the assumptions;

(13) $\dfrac{\mathcal{F},\ \mathcal{F} \Rightarrow \mathcal{G}}{\mathcal{G}}$.

The three meta-rules with hypotheses are same as before, whereas the meta-laws have already been shown to be deducible above, except for (1), (5), (7) and (8), which follow immediately from the corresponding meta-rules using the $\Rightarrow$-introduction.

Conversely, we immediately get the α-conversion, &-elimination and specialization meta-rules from the corresponding meta-laws, using the $\Rightarrow$-elimination meta-rule. Also, from the meta-laws (2) and (4) we get $\mathcal{F} \Rightarrow (\mathcal{G} \Rightarrow \mathcal{F}\&\mathcal{G})$, which yields the &-introduction meta-rule by a double application of $\Rightarrow$-elimination.

It remains to recover the $\Rightarrow$-introduction meta-rule. (This is similar to the usual proof of the deduction theorem.) We induct on the length of the given deduction of $\mathcal{G}$ from $\mathcal{F}$. If this deduction is trivial, i.e. $\mathcal{G} = \mathcal{F}$, then the desired conclusion $\mathcal{G} \Rightarrow \mathcal{F}$ is yielded by the meta-law (2). Now given a deduction of $\mathcal{G}$ from $\mathcal{F}$, there are two possibilities for the last meta-rule used in this deduction.

Case 1. The last meta-rule is the $\Rightarrow$-elimination with hypotheses of the form $\mathcal{H}$ and $\mathcal{H} \Rightarrow \mathcal{F}$. We may assume that $\mathcal{G} \Rightarrow \mathcal{H}$ and $\mathcal{G} \Rightarrow (\mathcal{H} \Rightarrow \mathcal{F})$ are deducible. Then by a double application of the meta-law (4), $\mathcal{H} \Rightarrow (\mathcal{G} \Rightarrow \mathcal{F})$ is also deducible. Also, the meta-laws (3) and (4) yield that $(\mathcal{G} \Rightarrow \mathcal{H}) \Rightarrow \Big((\mathcal{H} \Rightarrow (\mathcal{G} \Rightarrow \mathcal{F})) \Rightarrow (\mathcal{G} \Rightarrow \mathcal{F})\Big)$ is deducible. Applying $\Rightarrow$-elimination twice, we conclude that $\mathcal{G} \Rightarrow \mathcal{F}$ is also deducible.

Case 2. The last meta-rule is the generalization with hypothesis of the form $\mathcal{F}'$, where $\mathcal{F} = (x)\,\mathcal{F}'$ and x is either an individual or a predicate variable that does not occur freely in $\mathcal{G}$. We may assume that $\mathcal{G} \Rightarrow \mathcal{F}'$ is deducible. Then by the generalization meta-rule, $(x)\,(\mathcal{G} \Rightarrow \mathcal{F}')$ is deducible. Hence by (9) or (10), $\mathcal{G} \Rightarrow \mathcal{F}$ is also deducible.

4.4.4. Enderton-style formulation. For purposes of models (see §*4.7.1*) we also need an "Enderton-style" formulation of the meta-logic, whose only meta-rule is $\Rightarrow$-elimination.[57] The notation is the same as in the Hilbert-style formulation, except that we now additionally need a tuple $\vec{z}$ of individual variables and a tuple $\vec{\beta}$ of predicate variables.

(1') $(\vec{z})\,(\vec{\beta})\,\big(\mathcal{F} \Rightarrow \mathcal{G}\big)$, if $\mathcal{F}$ is α-equivalent to $\mathcal{G}$;

(2') $(\vec{z})\,(\vec{\beta})\,\big(\mathcal{F} \Rightarrow \mathcal{F}\big)$;

(3') $(\vec{z})\,(\vec{\beta})\,\big((\mathcal{F} \Rightarrow \mathcal{G})\,\&\,(\mathcal{G} \Rightarrow \mathcal{H}) \implies (\mathcal{F} \Rightarrow \mathcal{H})\big)$;

(4') $(\vec{z})\,(\vec{\beta})\,\Big(\big((\mathcal{F}\,\&\,\mathcal{G}) \Rightarrow \mathcal{H}\big) \iff \big(\mathcal{F} \Rightarrow (\mathcal{G} \Rightarrow \mathcal{H})\big)\Big)$;

(5') $(\vec{z})\,(\vec{\beta})\,\big(\mathcal{F}\,\&\,\mathcal{G} \Rightarrow \mathcal{F}\big)$ and $(\vec{z})\,(\vec{\beta})\,\big(\mathcal{F}\,\&\,\mathcal{G} \Rightarrow \mathcal{G}\big)$;

(6') $(\vec{z})\,(\vec{\beta})\,\big((\mathcal{H} \Rightarrow \mathcal{F})\,\&\,(\mathcal{H} \Rightarrow \mathcal{G}) \implies (\mathcal{H} \Rightarrow \mathcal{F}\,\&\,\mathcal{G})\big)$;

(7') $(\vec{z})\,(\vec{\beta})\,\big((x)\,\mathcal{F} \Rightarrow \mathcal{F}[x/T]\big)$, provided that T is free for x in $\mathcal{F}$;

(8') $(\vec{z})\,(\vec{\beta})\,\big((\gamma)\,\mathcal{F} \Rightarrow \mathcal{F}[\gamma/\Phi]\big)$, provided that Φ is free for γ in $\mathcal{F}$;

---

[57]Hilbert-style deductive systems for first-order logics usually include some form of the generalization rule. In particular, so do the original systems of Hilbert–Bernays [90] and Hilbert–Ackermann [89]. Enderton's deductive system [48] includes only one inference rule, the *modus ponens*. The generalization rule has the status of a meta-theorem in his system, just like the Deduction Theorem.

$(9')$ $(\vec{z})\,(\vec{\beta})\,\big((x)\,(\mathcal{F} \Rightarrow \mathcal{G}) \implies ((x)\,\mathcal{F} \Rightarrow (x)\,\mathcal{G})\big)$;

$(10')$ $(\vec{z})\,(\vec{\beta})\,\big((\gamma)\,(\mathcal{F} \Rightarrow \mathcal{G}) \implies ((\gamma)\,\mathcal{F} \Rightarrow (\gamma)\,\mathcal{G})\big)$;

$(11')$ $(\vec{z})\,(\vec{\beta})\,\big(\mathcal{F} \Rightarrow (x)\,\mathcal{F}\big)$, provided that x does not occur freely in $\mathcal{F}$;

$(12')$ $(\vec{z})\,(\vec{\beta})\,\big(\mathcal{F} \Rightarrow (\gamma)\,\mathcal{F}\big)$, provided that γ does not occur freely in $\mathcal{F}$;

$(13')$ $\dfrac{\mathcal{F},\, \mathcal{F} \Rightarrow \mathcal{G}}{\mathcal{G}}$.

All the meta-laws of the Enderton-style system are easily deducible in the Hilbert-style system. Also, all the meta-laws of the Hilbert-style system are easily deducible in the Enderton-style system (in fact, all of them except for (9) and (10) are already present in the Enderton-style system, by considering empty $\vec{z}$ and $\vec{\beta}$).

It remains to recover the generalization meta-rules in the Enderton-style system. Let us first observe that if $\mathcal{F}$ and $\mathcal{G}$ are meta-formulas and v is either individual or predicate variable, $(v)\,\mathcal{G}$ is deducible from $(v)\,\mathcal{F}$ and $(v)\,(\mathcal{F} \Rightarrow \mathcal{G})$ in the Enderton-style system. Indeed, $(v)\,(\mathcal{F} \Rightarrow \mathcal{G}) \Rightarrow ((v)\,\mathcal{F} \Rightarrow (v)\,\mathcal{G})$ is an instance of $(9')$ or $(10')$, and the assertion follows by a double application of the $\Rightarrow$-elimination. Also, if $\mathcal{F}$ is a meta-formula and v is either individual or predicate variable that does not occur freely in $\mathcal{F}$, then $(v)\,\mathcal{F}$ is easily deducible from $\mathcal{F}$ in the Enderton-style system.

Now suppose that we are given a deduction from assumptions in the system combining the Enderton-style system and the two generalization meta-rules. Let us consider any application of these meta-rules, i.e. the meta-generalization of a meta-formula $\mathcal{F}$ with respect to an individual or predicate variable, call it v, that is not free in any of the assumptions. The only way that v can first arise as a free variable in the deduction of $\mathcal{F}$ from the assumptions is by occurring as a free variable in some instance $\mathcal{G}$ of some meta-law used in this deduction. Arguing by induction, we may assume that generalization meta-rules are not used within the deduction of $\mathcal{F}$ from the assumptions (regarded as a subtree of the deduction tree under consideration). But then we amend this deduction as follows: (i) $\mathcal{G}$ is replaced by another instance $(v)\,\mathcal{G}$ of the same meta-law; (ii) every application of the $\Rightarrow$-elimination, inferring $\mathcal{G}'$ from $\mathcal{F}'$ and $\mathcal{F}' \Rightarrow \mathcal{G}'$, is replaced as follows: if v occurs freely neither in $\mathcal{F}'$ nor in $\mathcal{G}'$, nothing is changed; if v occurs freely in $\mathcal{F}'$, the $\Rightarrow$-implication is replaced by the deduction (just mentioned above) of $(v)\,\mathcal{G}'$ from $(v)\,\mathcal{F}'$ and $(v)\,\mathcal{F}' \Rightarrow \mathcal{G}'$; and if v occurs freely in $\mathcal{G}'$ but not in $\mathcal{F}'$, then we first deduce $(v)\,\mathcal{F}'$ from $\mathcal{F}'$ and then $(v)\,\mathcal{G}'$ from $(v)\,\mathcal{F}'$ and $(v)\,\mathcal{F}' \Rightarrow \mathcal{G}'$; (iii) the application of meta-generalization to $\mathcal{F}$ is dropped. In the end of the inductive procedure, all applications of the generalization meta-rules will have been eliminated.

4.4.5. Deducing from assumptions. The following examples deal with specific formulas (regarded as atomic meta-formulas).

Example 4.35. Let us deduce β from the assumptions α, $\alpha \rightarrow \beta$ and the following *modus ponens* (object-level) rule:

(i) $(\varphi, \psi)\ \big(\varphi \to \psi\ \Longrightarrow\ (\varphi \Rightarrow \psi)\big)$.

Using the second-order meta-specialization, from (i) we infer $\alpha \to \beta\ \Longrightarrow\ (\alpha \Rightarrow \beta)$. From this and our assumption $\alpha \to \beta$ we infer, using the $\Rightarrow$-elimination, $\alpha \Rightarrow \beta$. Applying it again, we infer β.

Example 4.36. Assuming (i) and the following principle (which holds in intuitionistic logic, cf. §**3.14**, (32)):

(ii) $(\gamma)\ \neg\neg(\gamma \vee \neg\gamma)$,

let us deduce the following implication between principles:

$$(\gamma)\ \neg\neg\gamma \to \gamma\ \Longrightarrow\ (\gamma)\ \gamma \vee \neg\gamma.$$

Taking into account the $\Rightarrow$-introduction, it suffices to deduce $(\gamma)\ \gamma \vee \neg\gamma$ from (i), (ii) and $(\gamma)\ \neg\neg\gamma \to \gamma$. Now from (ii) we get $\neg\neg(\beta \vee \neg\beta)$ by the second-order meta-specialization; and from $(\gamma)\ \neg\neg\gamma \to \gamma$ we similarly get $\neg\neg(\beta \vee \neg\beta) \to (\beta \vee \neg\beta)$. Finally, similarly to Example 4.35, these yield $\beta \vee \neg\beta$; and then the second-order meta-generalization yields $(\beta)\ \beta \vee \neg\beta$. This is the same as $(\gamma)\ \gamma \vee \neg\gamma$ up to γ-equivalence.

Remark 4.37. By contrast, the rule

$$(\gamma)\ (\neg\neg\gamma \to \gamma \Rightarrow \gamma \vee \neg\gamma)$$

cannot be deduced from (i) and (ii), because it meta-specializes (by setting $\gamma = \neg\beta$) to $\neg\neg\neg\beta \to \neg\beta \Rightarrow \neg\beta \vee \neg\neg\beta$, where the premise holds in intuitionistic logic (see §**3.14**, (3)), but the conclusion does not (see §**5.4** below). (In other words, the said rule is not even admissible for intuitionistic logic.)

Example 4.38. Let us deduce $\exists x\,\big(\gamma(x) \to \gamma(y)\big)$ from (i) and the following assumptions (which will be among the laws of intuitionistic logic):

(iii) $(\gamma)\ \gamma \to \gamma$;
(iv) $(\beta, t)\ \beta(t) \to \exists x\,\beta(x)$.

Using the first-order meta-specialization, from (iii) we infer $\gamma(y) \to \gamma(y)$, and from (iv) we infer $(\beta)\ \beta(y) \to \exists x\,\beta(x)$. From the latter we infer, using the second-order meta-specialization, $\big(\gamma(y) \to \gamma(y)\big) \to \exists x\,\big(\gamma(x) \to \gamma(y)\big)$. (Here we use that $x \mapsto \gamma(x) \to \gamma(y)$ is free for β in $\beta(y) \to \exists x\,\beta(x)$. For this, it is essential that y is free in $\beta(y) \to \exists x\,\beta(x)$. Thus we would not be able to make this step directly from (iv).) The remainder of the argument is similar to that of Example 4.35.

Example 4.39. Let us deduce $\forall x\,\big(\gamma(x) \to \gamma(y)\big)$ from (i), (iii) and the following rule

(v) $(\beta, t)\ \big(\beta(t) \Rightarrow \forall x\,\beta(x)\big)$.

Using the first-order meta-specialization, from (v) we can still infer $(\beta)\ \big(\beta(y) \Rightarrow \forall x\,\beta(x)\big)$. From the latter we can still infer, using the second-order meta-specialization, the metaformula $\big(\gamma(y) \to \gamma(y)\big) \Rightarrow \forall x\,\big(\gamma(x) \to \gamma(y)\big)$. Since we still have $\gamma(y) \to \gamma(y)$ from (iii), by the $\Rightarrow$-elimination we conclude $\forall x\,\big(\gamma(x) \to \gamma(y)\big)$.

Remark 4.40. Since the formula deduced in Example 4.39 is contradictory (either classically or intuitionistically), we conclude that (v) is an incorrect interpretation of the generalization (object-level) rule. Moreover, the trouble is clearly in that $x \mapsto \gamma(x) \to \gamma(y)$ is free for β in $\beta(y) \Rightarrow \forall x\, \beta(x)$ — but it shouldn't be, if this were to be the generalization rule. There is no such trouble with the true generalization (object-level) rule:

$$(v')\ (\beta)\ \big((t)\ \beta(t) \Rightarrow \forall x\, \beta(x)\big).$$

This does not meta-specialize by substituting $x \mapsto \gamma(x) \to \gamma(y)$ for β, because it is not free for β in $(t)\ \beta(t) \Rightarrow \forall x\, \beta(x)$. The first-order meta-specialization does not apply either, since the scope of the meta-quantifier is not the entire meta-formula in (v').

Example 4.41. Assuming (i), let us deduce $(x)\ \varphi(x) \to \psi(x) \Rightarrow \big((x)\ \varphi(x) \Rightarrow (x)\ \psi(x)\big)$.

By the $\Rightarrow$-introduction, it suffices to deduce $(x)\ \varphi(x) \Rightarrow (x)\ \psi(x)$, assuming (i) and $(x)\ \varphi(x) \to \psi(x)$. For this it suffices, by the same token, to deduce $(x)\ \psi(x)$, assuming (i), $(x)\ \varphi(x) \to \psi(x)$ and $(x)\ \varphi(x)$. The latter two assumptions yield $\varphi(x) \to \psi(x)$ and $\varphi(x)$ by the meta-specialization. From these we get $\psi(x)$ similarly to Example 4.35. Finally, the meta-generalization yields $(x)\ \psi(x)$.

Let us note that the last step is only possible since x is not free in our assumptions.

Example 4.42. The meta-formula in the conclusion of Example 4.38 can be restated as $\forall y\, \exists x\, \big(\gamma(x) \to \gamma(y)\big)$, if in addition to the original hypotheses (i), (iii), (iv) we assume the generalization rule (v').

Indeed, since y is not free in any of the assumptions, using meta-generalization we get $(y)\ \exists x\, \big(\gamma(x) \to \gamma(y)\big)$ from the conclusion of 4.38. Now the generalization rule (v') can be applied, using the second-order meta-specialization and the $\Rightarrow$-elimination, to obtain the desired closed formula $\forall y\, \exists x\, \big(\gamma(x) \to \gamma(y)\big)$.

Similarly, the conclusion of Example 4.39 follows in the form $\forall y\, \forall x\, \big(\gamma(x) \to \gamma(y)\big)$ from the assumptions (i), (iii), (v') and the faulty assumption (v).

4.4.6. Specialization. A *specialization* of a meta-formula $\mathcal{F}$ is any meta-formula that is a result of β-reduction (possibly involving α-conversions) of a λ-expression of the form $\big(\vec{\varphi} \mapsto (\vec{x} \mapsto \mathcal{F})\big)(\vec{\Sigma})(\vec{T})$, where $\vec{x}$ is an n-tuple of pairwise distinct individual variables, $\vec{\varphi}$ is an m-tuple of pairwise distinct predicate variables of arities r_i, $\vec{T}$ is an n-tuple of terms, and $\vec{\Sigma}$ is a tuple of r_i-formulas.

In other words, a meta-formula $\mathcal{F}'$ is a specialization of a meta-formula $\mathcal{F}$ if and only if $\mathcal{F}' = \mathcal{F}[\vec{x}/\vec{T},\ \vec{\varphi}/\vec{\Sigma}]$, where $\vec{T}$ is free for $\vec{x}$ in $\mathcal{F}$ and $\vec{\Sigma}$ is free for $\vec{\varphi}$ in $\mathcal{F}$.

Clearly, specialization can always be obtained by first using the generalization meta-rule, then the specialization meta-rule, and finally the α-conversion meta-rule.

Thus, if a meta-formula $\mathcal{F}$ is deducible from assumptions that contain no free variables, then any its specialization is deducible from the same assumptions.

Example 4.43. Let us consider the formula $\mathcal{F} = \varphi(\mathsf{x})$ (regarded as an atomic meta-formula), the term $T = \mathsf{z}$ and the 1-formula $F = \mathsf{y} \mapsto \exists \mathsf{z}\, \psi(\mathsf{y})$. Then

$$\big(\varphi \mapsto (\mathsf{x} \mapsto \mathcal{F})\big)(F)(T) = \Big(\varphi \mapsto \big(\mathsf{x} \mapsto \varphi(\mathsf{x})\big)\Big)\big(\mathsf{y} \mapsto \exists \mathsf{z}\, \psi(\mathsf{y})\big)(\mathsf{z})$$

$$\overset{\beta}{=} \Big(\mathsf{x} \mapsto \big(\mathsf{y} \mapsto \exists \mathsf{z}\, \psi(\mathsf{y})\big)(\mathsf{x})\Big)(\mathsf{z}) \overset{\beta}{=} \big(\mathsf{y} \mapsto \exists \mathsf{z}\, \psi(\mathsf{y})\big)(\mathsf{z}) \overset{\beta}{=} \mathsf{y} \mapsto \exists \mathsf{z}\, \psi(\mathsf{y})|_{\mathsf{y}:=\mathsf{z}} \overset{\alpha}{=} \exists \mathsf{x}\, \psi(\mathsf{z}).$$

Of course, the final α-conversion can be avoided by the cost of replacing F with the α-equivalent 1-formula $F' = \mathsf{y} \mapsto \exists \mathsf{x}\, \psi(\mathsf{y})$. The result depends on the choice of F', but any other choice will yield a formula that differs only in the name of the bound variable.

Example 4.44. The formula $\alpha \wedge \alpha \to \alpha$ is a specialization of $\alpha \wedge \beta \to \alpha$.

Similarly, the meta-formula $\big(\alpha(\mathsf{x}) \to \alpha(\mathsf{x})\big) \implies \big(\alpha(\mathsf{x}) \to \forall \mathsf{x}\, \alpha(\mathsf{x})\big)$ is a specialization of $\big(\beta \to \alpha(\mathsf{x})\big) \implies \big(\beta \to \forall \mathsf{x}\, \alpha(\mathsf{x})\big)$.

On the other hand, the meta-formula $(\mathsf{x})\big(\alpha(\mathsf{x}) \to \alpha(\mathsf{x})\big) \implies \big(\alpha(\mathsf{x}) \to \forall \mathsf{x}\, \alpha(\mathsf{x})\big)$ is not a specialization of $(\mathsf{x})\big(\beta \to \alpha(\mathsf{x})\big) \implies \big(\beta \to \forall \mathsf{x}\, \alpha(\mathsf{x})\big)$.

Example 4.45. The formula $\big(\beta(\mathsf{y}) \to \beta(\mathsf{y})\big) \longrightarrow \exists \mathsf{x}\big(\beta(\mathsf{x}) \to \beta(\mathsf{y})\big)$ is a specialization of $\alpha(\mathsf{y}) \to \exists \mathsf{x}\, \alpha(\mathsf{x})$ (see Example 4.38).

Similarly, the meta-formula $\big(\beta(\mathsf{y}) \to \beta(\mathsf{y})\big) \implies \forall \mathsf{x}\big(\beta(\mathsf{x}) \to \beta(\mathsf{y})\big)$ is a specialization of $\alpha(\mathsf{y}) \Rightarrow \forall \mathsf{x}\, \alpha(\mathsf{x})$ (see Example 4.39).

On the other hand, the meta-formula $(\mathsf{y})\big(\beta(\mathsf{y}) \to \beta(\mathsf{y})\big) \implies \forall \mathsf{x}\big(\beta(\mathsf{x}) \to \beta(\mathsf{y})\big)$ is not a specialization of $(\mathsf{y})\ \alpha(\mathsf{y}) \Rightarrow \forall \mathsf{x}\, \alpha(\mathsf{x})$ (see Remark 4.40).

4.5. Logics

The purpose of this section is to introduce and discuss some "syntactic meta-sugar".

4.5.1. Rules. The *first-order meta-closure* $(1)\mathcal{F}$ of the meta-formula $\mathcal{F}$ is $(\vec{x})\,\mathcal{F}$, where $\vec{x}$ is the tuple of all individual variables occurring freely in $\mathcal{F}$. The *second-order meta-closure* $(2)\mathcal{F}$ is $(\vec{\gamma})\,\mathcal{F}$, where $\vec{\gamma}$ is the tuple of all predicate variables occurring freely in $\mathcal{F}$. Their combination $(2)(1)\mathcal{F}$ will be called the *full meta-closure* of $\mathcal{F}$. If Δ is a formula (and only in this case) we abbreviate $(2)(1)\Delta$ by $\cdot\Delta$.

Thus every principle can be written in the form $\cdot\Delta$, and every meta-formula of the form $\cdot\Delta$ is a principle, where by our convention Δ must be a mere formula.

A (structural) *rule*, written $\Gamma_1, \ldots, \Gamma_m/\Delta$, or, in more detail,

$$\frac{\Gamma_1, \ldots, \Gamma_m}{\Delta},$$

where $\Gamma_1, \ldots, \Gamma_m$ and Δ are formulas, is an abbreviation for the meta-formula

$$(2)\Big((1)\Gamma_1 \ \&\ \cdots\ \&\ (1)\Gamma_m \implies (1)\Delta\Big).$$

The formulas $\Gamma_1, \ldots, \Gamma_m$ are called the *premisses* of the rule, and Δ its *conclusion.*

What happens when $m = 0$, i.e. when the list of premisses is empty? If the empty meta-conjunction is defined as an abbreviation of some specific deducible meta-formula

$\mathcal{T}$ (for example, $(\gamma)\,\gamma \Rightarrow (\gamma)\,\gamma$), then a rule with no premisses, $/\,\Delta$, denotes the meta-formula $(2)\big(\mathcal{T} \Rightarrow (1)\,\Delta\big)$. But this is the same as the meta-formula $(2)\,(1)\,\Delta$, since it is easy to deduce $\big(\mathcal{T} \Rightarrow \mathcal{F}\big) \iff \mathcal{F}$. Thus rules with no premisses can be identified with principles.

Let us spell out some informal justification for the definition of a rule. The point of the second-order meta-quantification ("(2)") is clear at once from Example 4.36 and Remark 4.37. The point of the first-order meta-quantification ("(1)") is that we want to use principles and rules as basic assumptions in our deductions; but having free variables in the assumptions would ruin many correct deductions, such as the ones in Examples 4.30, 4.41 and 4.42. Finally, the reason why the first-order meta-closure is applied separately to each premiss of the rule (in contrast to the second meta-closure) is clear from Example 4.39 and Remark 4.40. Let us note that $(1)\,\Gamma_1 \,\&\, \cdots \,\&\, (1)\,\Gamma_m \iff (1)\,(\Gamma_1 \,\&\, \cdots \,\&\, \Gamma_m)$ is deducible (see Example 4.32).

A byproduct of the fact that principles and rules are λ-closed is that they have no non-trivial specializations. Let us recall that a specialization of a meta-formula is obtained by first applying meta-generalization, then meta-specialization, and finally α-equivalence. By a *special case* of a rule we mean any rule that can be obtained from it by *first* applying meta-specialization, *then* meta-generalization, and finally α-equivalence.

This turns out to work out differently for principles and for rules with premisses. Clearly, a principle $\cdot\,\Delta'$ is a special case of a principle $\cdot\,\Delta$ if and only if the formula Δ' is a specialization of the formula Δ. (That is, if and only if $\Delta' = \Delta[\vec{x}/\vec{T},\ \vec{\varphi}/\vec{\Sigma}]$, where $\vec{T}$ is free for $\vec{x}$ in Δ and $\vec{\Sigma}$ is free for $\vec{\varphi}$ in Δ.) This is not equivalent to saying that $(1)\,\Delta'$ is a specialization of $(1)\,\Delta$. For instance, $(1)\,\alpha(x)$ is not a specialization of $(1)\,\beta$, but $\alpha(x)$ is a specialization of β.

But for rules with premisses, we cannot use first-order meta-specialization. Thus a rule $\Gamma'_1, \ldots, \Gamma'_m \,/\, \Delta'$ is a special case of a rule $\Gamma_1, \ldots, \Gamma_m \,/\, \Delta$ if and only if the meta-formula $(1)\,\Gamma'_1 \,\&\, \cdots \,\&\, (1)\,\Gamma'_m \Rightarrow (1)\,\Delta'$ is a specialization of $(1)\,\Gamma_1 \,\&\, \cdots \,\&\, (1)\,\Gamma_m \Rightarrow (1)\,\Delta$. (That is, if and only if for some tuple $\vec{\Sigma}$ of r_i-formulas that is free for $\vec{\varphi}$ in each Γ_j and in Δ, each $\Gamma'_j = \Gamma_j(\vec{\varphi}/\vec{\Sigma})$ and $\Delta' = \Delta(\vec{\varphi}/\vec{\Sigma})$.)

Example 4.46. (Cf. Example 4.44.) The principle $\cdot\,\alpha \wedge \alpha \to \alpha$ is a special case of the principle $\cdot\,\alpha \wedge \beta \to \alpha$.

Similarly, the principle $\cdot\,\big(\alpha(x) \to \alpha(x)\big) \;\longrightarrow\; \big(\alpha(x) \to \forall x\,\alpha(x)\big)$ is a special case of $\cdot\,\big(\beta \to \alpha(x)\big) \;\longrightarrow\; \big(\beta \to \forall x\,\alpha(x)\big)$.

However, the rule $\dfrac{\alpha(x) \to \alpha(x)}{\alpha(x) \to \forall x\,\alpha(x)}$ is not a special case of $\dfrac{\beta \to \alpha(x)}{\beta \to \forall x\,\alpha(x)}$.

Example 4.47. (Cf. 4.45.) The principle $\cdot\,\big(\beta(y) \to \beta(y)\big) \;\longrightarrow\; \exists x\,\big(\beta(x) \to \beta(y)\big)$ is a special case of $\cdot\,\alpha(y) \to \exists x\,\alpha(x)$.

Similarly, the principle $\cdot\,\big(\beta(y) \to \beta(y)\big) \;\longrightarrow\; \forall x\,\big(\beta(x) \to \beta(y)\big)$ is a special case of $\cdot\,\alpha(y) \to \forall x\,\alpha(x)$.

However, the rule $\dfrac{\beta(y) \to \beta(y)}{\forall x\,\big(\beta(x) \to \beta(y)\big)}$ is not a special case of $\dfrac{\alpha(y)}{\forall x\,\alpha(x)}$.

4.5.2. Logics and theories. A *derivation system*[58] $\mathcal{D}$ in a first-order language $\mathscr{L}$ is a meta-formula of the form

$$\mathcal{H}_1 \,\&\, \cdots \,\&\, \mathcal{H}_k,$$

where each $\mathcal{H}_i$ is a purely logical rule (possibly with no premises) in $\mathscr{L}$. In other words, $\mathcal{D}$ is the meta-conjunction of finitely many purely logical principles and other purely logical rules (with at least one premise), which are called, respectively, the *laws*[59] and the *inference rules* of $\mathcal{D}$.

A *logic* is a meta-equivalence class of derivation systems in some first-order language $\mathscr{L}$. In other words, derivation systems $\mathcal{D}$ and $\mathcal{D}'$ are said to *determine the same logic* if the meta-formula $\mathcal{D} \Leftrightarrow \mathcal{D}'$ is deducible. The purely logical part of $\mathscr{L}$ is called the *language of the logic L* and is denoted $\mathscr{L}(L)$. A logic L^+ is called an *extension* of a logic L if there exist derivation systems $\mathcal{D}$ and $\mathcal{D}'$ such that $\mathcal{D}$ determines L and $\mathcal{D} \,\&\, \mathcal{D}'$ determines L^+.

A meta-formula $\mathcal{F}$ is called *derivable* in the logic L determined by a derivation system $\mathcal{D}$ if the meta-formula $\mathcal{D} \Rightarrow \mathcal{G}$ is deducible (in the meta-logic) — or equivalently, if $\mathcal{F}$ is deducible from the hypothesis $\mathcal{D}$ (in the meta-logic). Clearly, adding a derivable principle or rule to a derivation system $\mathcal{D}$ does not affect derivability of principles and rules in the logic determined by $\mathcal{D}$.

A principle $\cdot\Delta$ is derivable in the logic if and only if either $\cdot\Delta$ is a special case of a law, or there exist derivable principles $\cdot\Gamma_1, \ldots, \cdot\Gamma_m$ such that the rule $\Gamma_1, \ldots, \Gamma_m / \Delta$ is a special case of an inference rule. More generally, a rule $\Gamma_1, \ldots, \Gamma_m / \Delta$ is derivable in the logic if and only if either the principle $\cdot\Delta$ is one of the $\cdot\Gamma_i$,[60] or it is a special case of a law, or there exist finitely many formulas $\Delta_1, \ldots, \Delta_k$ such that each rule $\Gamma_1, \ldots, \Gamma_m / \Delta_i$ is derivable, $i = 1, \ldots, k$, and $\Delta_1, \ldots, \Delta_k / \Delta$ is a special case of an inference rule. (The "if" assertion is trivial, and the converse is not hard.)

It is also not hard to see that every special case of a derivable rule is derivable, and that every special case of a derivable principle is a derivable principle. In fact, every specialization of a derivable meta-formula is a derivable meta-formula; in particular, every specialization of a derivable formula is a derivable formula.

A first-order *theory* $\mathcal{T}$ over a logic L is a finite meta-conjunction of principles in a first-order language $\mathscr{L}$ whose purely logical part is the language of L.[61] The language $\mathscr{L}$ is called *the language of the theory $\mathcal{T}$* and is denoted $\mathscr{L}(\mathcal{T})$, and these principles are

---

[58]Also called a "deductive system" in the literature. For our purposes it is convenient to distinguish *deductions* in meta-logic from *derivations* within a specific object logic.

[59]These correspond to what is usually called "axiom schemata". Apart from the conflation of schemata and principles, this standard terminology is objectionable in the case of intuitionistic logic, as it fits well with the Brouwer–Heyting approach, but not with Kolmogorov's approach. On Kolmogorov's approach, the laws of intuitionistic logic are not propositions accepted without proof, but rather a toolkit of elementary constructions that can be used as black boxes.

[60]But not just a special case of one of the $\cdot\Gamma_i$. Thus the formulas Γ_i are *not* treated just as if the principles $\cdot\Gamma_i$ were laws.

[61]This corresponds to the traditional notion of a theory given by finitely many axioms and finitely many axiom schemata.

called the *axioms*[62] of $\mathcal{T}$. In contrast to the laws of a derivation system, the axioms of a theory may involve function symbols and predicate constants. Because of this, a typical theory (such as the theory of groups or the theory of an order relation) would involve no predicate variables, and hence also no meta-quantifiers (as long as all its axioms are closed formulas). Meta-quantifiers would be needed, however, to express principles that are not purely logical (such as the principle of mathematical induction). An *empty theory* over L has no axioms (but its language can be different from the language of L).

A formula $\mathcal{G}$ of $\mathscr{L}$ is a *theorem*[63] of a theory $\mathcal{T}$ over a logic L if the meta-formula $\mathcal{T} \Rightarrow \mathcal{G}$ is derivable in L, in other words, if $\mathcal{G}$ is deducible from $\mathcal{D} \,\&\, \mathcal{T}$ (in the meta-logic), where $\mathcal{D}$ is a derivation system for L. This is equivalent to saying that either the principle $\cdot\,\mathcal{G}$ is a special case of a law of $\mathcal{D}$ or of an axiom of $\mathcal{T}$, or there exist theorems $\mathcal{F}_1, \ldots, \mathcal{F}_m$ such that the rule $\mathcal{F}_1, \ldots, \mathcal{F}_m / \mathcal{G}$ is a special case of an inference rule of $\mathcal{D}$. (Let us note that an axiom that involves no meta-quantifiers has no special cases other than itself.) A theory $\mathcal{T}$ is called *consistent* if its theorems do not include all formulas of $\mathscr{L}$. It is easy to see that $\mathcal{T}$ is inconsistent (i.e. not consistent) if and only if $\mathcal{T} \Rightarrow (\gamma)\,\gamma$ is derivable in L.

Of course, assertions called "theorems", "lemmas" or "propositions" in the present text are not supposed to be theorems of any first-order theory. They are meta-meta-theorems. But some assertions of the present text (such as those starting with "the following formula is derivable in such and such logic") do signify theorems of the empty theories over classical, intuitionistic or QS4 logics. Also, if logics are regarded as meta-theories over the meta-logic, then for instance §*4.4.2* is devoted to some meta-theorems of the empty meta-theory over the meta-logic, and §*4.4.5* to those of some non-empty meta-theories.

4.5.3. Syntactic consequence. If L is the logic determined by a derivation system $\mathcal{D}$, we denote by $\vdash \mathcal{F}$, or in more detail $\vdash_L \mathcal{F}$, the judgement that the meta-formula $\mathcal{F}$ is derivable in the logic. The judgement $\vdash (\mathcal{F}_1 \,\&\, \cdots \,\&\, \mathcal{F}_m \Rightarrow \mathcal{G})$ is also abbreviated by

$$\mathcal{F}_1, \ldots, \mathcal{F}_m \vdash \mathcal{G}.$$

When this judgement is true, we also say that $\mathcal{G}$ is a *(syntactic) consequence* of the $\mathcal{F}_i$.[64] Here the $\mathcal{F}_i$ and $\mathcal{G}$ can be arbitrary meta-formulas; but a few cases are of particular interest. Namely, $\mathcal{F}_1, \ldots, \mathcal{F}_m \vdash \mathcal{G}$ might be of the following form:

(1) $\vdash$ (2)(1) Δ, where Δ is a formula. This is saying that $\cdot\,\Delta$ is a derivable principle.

---

[62] In the context of intuitionistic logic (and its non-classical extensions) these may be called *postulates* following Euclid (and his ancient commentators Geminus and Proclus, who have actually made the distinction between axioms and postulate quite clear, see [3]). It should be noted, however, that the modern meaning of the word *postulate* is not very different from that of *axiom*, as opposed to the ancient Greek word which meant *request, demand* (see [3]).

[63] Or a *soluble problem* in intuitionistic logic (and its non-classical extensions).

[64] Thus consequence belongs to meta-meta-logic, and not meta-logic. It is clear from Example 4.50 below (see also Remark 4.34) that it would not be a good idea to explain the judgement $\vdash \mathcal{F}$ by identifying it with the meta-formula $\mathcal{D} \Rightarrow \mathcal{F}$ itself.

(1') $(2)(1)\Gamma_1, \ldots, (2)(1)\Gamma_m \vdash (2)(1)\Delta$, where the Γ_i and Δ are formulas. This is saying that the principle $\cdot\,\Delta$ is a consequence of the principles $\cdot\,\Gamma_i$.

(1'') $(2)\Big((1)\Theta_{11}\,\&\cdots\&\,(1)\Theta_{1n_1} \Rightarrow (1)\Gamma_1\Big), \ldots, (2)\Big((1)\Theta_{m1}\,\&\cdots\&\,(1)\Theta_{mn_m} \Rightarrow (1)\Gamma_m\Big) \vdash$
$(2)\Big((1)\Xi_1\,\&\cdots\&(1)\Xi_n \Rightarrow (1)\Delta\Big)$, where the Θ_{ij}, Ξ_j, Γ_i and Δ are formulas. This is saying that the rule $\Xi_1, \ldots, \Xi_n\,/\,\Delta$ is a consequence of the rules $\Theta_{i1}, \ldots, \Theta_{in_i}\,/\,\Gamma_i$.

(2) $\vdash (1)\Delta$, where Δ is a formula. This is equivalent to $\vdash (2)(1)\Delta$ (since there are no free predicate variables in $\mathcal{D}$).

(2') $(1)\Gamma_1, \ldots, (1)\Gamma_m \vdash (1)\Delta$, where the Γ_i and Δ are formulas. Similarly, this is the case if and only if the rule $\Gamma_1, \ldots, \Gamma_m\,/\,\Delta$ is derivable.[65]

(3) $\vdash \Delta$, where Δ is a formula. This is saying that the formula Δ is derivable, and is the case if and only if $\vdash (1)\Delta$ (since there are no free individual variables in $\mathcal{D}$).

(3') $\Gamma_1, \ldots, \Gamma_m \vdash \Delta$, where the Γ_i and Δ are formulas. This is saying that the formula Δ is a syntactic consequence of the formulas Γ_i.[66] As long as the Γ_i contain no free individual variables, we similarly have that $\Gamma_1, \ldots, \Gamma_m \vdash \Delta$ if and only if $(1)\Gamma_1, \ldots, (1)\Gamma_m \vdash (1)\Delta$. In general, for logics in which the deduction theorem and the exponential law hold, $\Gamma_1, \ldots, \Gamma_m \vdash \Delta$ if and only if $\vdash \Gamma_1 \wedge \cdots \wedge \Gamma_m \to \Delta$.

What is this *deduction theorem*? In logics where implication is present, a usual rule of inference is the *modus ponens* rule:

$$\frac{\varphi,\ \varphi \to \psi}{\psi}.$$

Using this rule, we get that $\vdash \Gamma \to \Delta$ implies $\Gamma \vdash \Delta$ (see Examples 4.35 and 4.31); and more generally, that $\Gamma_1, \ldots, \Gamma_n \vdash \Gamma \to \Delta$ implies $\Gamma_1, \ldots, \Gamma_n, \Gamma \vdash \Delta$ (where Γ and Δ are formulas). In some logics the converse also holds, in which case it is called the *deduction theorem*. In particular, the deduction theorem holds for classical and intuitionistic logics, where it is proved similarly to the argument in §*4.4.3* recovering the $\Rightarrow$-introduction meta-rule from the Hilbert-style formulation of the meta-logic. By combining the deduction theorem with the exponential law $\cdot\,\big((\alpha \wedge \beta) \to \gamma\big) \ \longleftrightarrow\ \big(\alpha \to (\beta \to \gamma)\big)$, we easily get the last assertion in (3') above.

The deduction theorem is not as banal as it might seem. Indeed, the *modus ponens* rule yields, in fact, slightly more: the rule $\Gamma\,/\,\Delta$ is a consequence of the principle $\cdot\,\Gamma \to \Delta$. (If the principle is purely logical, this is equivalent to saying that if we add $\cdot\,\Gamma \to \Delta$ as a new law to the derivation system, then $\Gamma\,/\,\Delta$ becomes derivable.) However, the deduction theorem, whenever it holds, does *not* imply that the principle $\cdot\,\Gamma \to \Delta$ is a consequence of the rule $\Gamma\,/\,\Delta$. (Which in the purely logical case is equivalent to saying that $\cdot\,\Gamma \to \Delta$ becomes derivable once $\Gamma\,/\,\Delta$ is set as a new inference rule.) This is simply not true in general — already in classical logic (see Examples 4.79 and 4.83). It follows,

---

[65]Thus $(1)\Gamma_1, \ldots, (1)\Gamma_m \vdash (1)\Delta$ is the traditional "varied variables" version of syntactic consequence for formulas, as in the textbooks by Schoenfield and Mendelson.

[66]Thus $\Gamma_1, \ldots, \Gamma_m \vdash \Delta$ is the more usual "fixed variables" version of syntactic consequence for formulas, as in the textbooks by Church, Enderton, Troelstra and van Dalen.

incidentally, that for closed formulas Γ and Δ, the meta-formula $(\Gamma \Rightarrow \Delta) \implies (\Gamma \to \Delta)$ need not be derivable, already in classical logic (see also Examples 4.78 and 4.81).

Consequence between principles and rules will often be referred to using the words such as "implies" and "follows", when this cannot lead to confusion; thus "$\cdot\,\Gamma$ implies $\cdot\,\Delta$" signifies a meta-formula different form that signified by "the principle that Γ implies Δ" (see Example 4.36 and Remark 4.37). A set of rules $\mathcal{R}_1, \ldots, \mathcal{R}_k$ is said to be *interderivable* (or simply *equivalent*, when this cannot lead to confusion) with a set of rules $\mathcal{S}_1, \ldots, \mathcal{S}_l$ if each $\mathcal{R}_i$ is a consequence of $\mathcal{S}_1, \ldots, \mathcal{S}_l$, and each $\mathcal{S}_i$ is a consequence of $\mathcal{R}_1, \ldots, \mathcal{R}_k$. The notion of consequence between principles and rules is going to be indispensable for our purposes; its explicit form seems to occur only rarely in the literature (see [203], [33], [172; 1.4.8], [102]), but indirectly it is used more often (in particular, through the equivalent notions of the order relation on extensions of a given logic and of the basis for a set of rules, such as the set of all admissible rules).

Of the judgements listed above, $(3')$ implies $(2')$, which in turn implies $(1')$ (see Example 4.32), and these implications are not reversible (see Example 4.39 and Remark 4.40 and, respectively, Example 4.36 and Remark 4.37). To these judgements we may add also:

$(0') \vdash \Gamma_1, \ldots, \vdash \Gamma_m$ imply $\vdash \Delta$;

$(0'') \vdash \Theta_{11}, \ldots, \Theta_{1n_1} / \Gamma_1, \ldots, \vdash \Theta_{m1}, \ldots, \Theta_{mn_m} / \Gamma_m$ imply $\vdash \mathcal{H}_1, \ldots, \mathcal{H}_n / \Delta$.

Then $(1'')$ implies $(0'')$ by the $\Rightarrow$-elimination meta-rule, and in particular $(1')$ implies $(0')$. But not conversely, since non-derivable principles need not be consequences of each other (even in classical logic, see Example 4.75 below; also, in zero-order intuitionistic logic, $\cdot\,\neg\neg\gamma \vee \neg\gamma$ does not imply $\cdot\,\gamma \vee \neg\gamma$, which in turn does not imply $\cdot\,\times$, see §5.4 below). In fact, it is easy to see that principles $\cdot\,\Gamma_1, \ldots, \cdot\,\Gamma_m$ imply $\cdot\,\Delta$ if and only if the judgements $\mathcal{G} \vdash \Gamma_1, \ldots, \mathcal{G} \vdash \Gamma_m$ imply $\mathcal{G} \vdash \Delta$ for each λ-closed meta-formula $\mathcal{G}$ (by considering $\mathcal{G} = \cdot\,\Gamma_1 \& \cdots \& \cdot\,\Gamma_m$). The same argument also establishes the same assertion with a different choice of $\mathcal{G}$'s, namely: $\cdot\,\Gamma_1, \ldots, \cdot\,\Gamma_m$ imply $\cdot\,\Delta$ over a logic L if and only if $\vdash_{L'} \Gamma_1, \ldots, \vdash_{L'} \Gamma_m$ imply $\vdash_{L'} \Delta$ for any extension L' of L by additional laws (or by additional laws and inference rules). Similarly, $(1'')$ over a logic L is equivalent to $(0'')$ over all extensions L' of L by additional laws and inference rules.

Remark 4.48. Kleene's sequent "$\Gamma_1, \ldots, \Gamma_m \vdash^{x_1, \ldots, x_k} \Delta$", where $\vec{x} = (x_1, \ldots, x_k)$ is the tuple of individual variables that "may be varied" in the deduction, can be identified with our judgement $(\vec{x})\,\Gamma_1, \ldots, (\vec{x})\,\Gamma_m \vdash (\vec{x})\,\Delta$. In fact, according to Kleene [103; §24], his sequent notation "is not fully explicit, as it does not show for which of the assumption formulas a given superscript variable may be varied. [...] When there is occasion to be more explicit, the facts may be stated verbally, e.g. as in Lemma 8a below." Our notation includes a formalization of such verbal comments, by simply meta-quantifying the assumption formulas over distinct collections of individual variables. (The meta-quantifiers in the conclusion are redundant, as long as they repeat those that are common to all of the assumptions.)

4.5.4. Admissible rules. A rule $\mathcal{R}$ is called *admissible* for the logic determined by a derivation system $\mathcal{D}$ if every derivable principle of the logic determined by the extended derivation system $\mathcal{D} \,\&\, \mathcal{R}$ is also derivable in the original logic. (That is, if $\vdash \mathcal{R} \Rightarrow \Phi$ implies $\vdash \Phi$ for each formula Φ.) Thus adding an admissible rule to a derivation system does not change the set of derivable principles, but unless this rule is derivable, it changes the logic (i.e. the syntactic consequence relation). This can happen even with classical logic, as we will see in §*4.7.1*. Clearly, a purely logical rule $\Gamma_1, \ldots, \Gamma_m \,/\, \Delta$ is admissible if and only if for every its special case $\Gamma_1[\vec{\varphi}/\vec{\Phi}], \ldots, \Gamma_m[\vec{\varphi}/\vec{\Phi}] \,/\, \Delta[\vec{\varphi}/\vec{\Phi}]$ (where $\vec{\Phi}$ is free for $\vec{\varphi}$ in each Γ_i and in Δ) such that each $\Gamma_i[\vec{\varphi}/\vec{\Phi}]$ is a derivable formula, $\Delta[\vec{\varphi}/\vec{\Phi}]$ is also a derivable formula.

Remark 4.49. For this reason, admissible rules of the form $\Gamma_1, \ldots, \Gamma_m \,/\, \Delta$ are sometimes stated in the literature in the form

$$\frac{\vdash \Gamma_1, \ldots, \vdash \Gamma_m}{\vdash \Delta}$$

or verbally, "if $\vdash \Gamma_1, \ldots, \vdash \Gamma_m$, then $\vdash \Delta$", where the predicate variables $\varphi_1, \ldots, \varphi_k$ occurring in the Γ_i and in Δ are tacitly understood to range over all r_j-formulas $\Phi_1, \ldots, \Phi_k$ that are free for $\varphi_1, \ldots, \varphi_k$ in each Γ_i and in Δ (that is, with an implicit meta-meta-quantifier $\forall \vec{\varphi}$ meant at the beginning of the sentence). The verbal version is particularly confusing, since its two alternative readings: (i) without assuming any implicit meta-meta-quantification as in the judgement $(0')$ above, (ii) assuming another implicit meta-meta-quntification: "if $\forall \vec{\varphi} \vdash \Gamma_1, \ldots, \forall \vec{\varphi} \vdash \Gamma_m$, then $\forall \vec{\varphi} \vdash \Delta$" — are entirely different judgements, which are equivalent to each other, but not to the admissibility of the rule. In older literature, however, one often finds *inference rules* stated in this ambiguous language (which is even more confusing, and certainly indicates the author's acceptance of the "formulaic" conception of logic).

Clearly, every derivable rule is admissible. Moreover, a purely logical rule $\Gamma_1, \ldots, \Gamma_m \,/\, \Delta$ is derivable if and only if for every its special case $\Gamma_1[\vec{\varphi}/\vec{\Phi}], \ldots, \Gamma_m[\vec{\varphi}/\vec{\Phi}] \,/\, \Delta[\vec{\varphi}/\vec{\Phi}]$ (where $\vec{\Phi}$ is free for $\vec{\varphi}$ in each Γ_i and in Δ) and for every formulas $\Psi_1, \ldots, \Psi_l$ such that $(\mathbf{1}) \Psi_1, \ldots, (\mathbf{1}) \Psi_l \vdash \Gamma_i[\vec{\varphi}/\vec{\Phi}]$ for each i, we also have $(\mathbf{1}) \Psi_1, \ldots, (\mathbf{1}) \Psi_l \vdash \Delta[\vec{\varphi}/\vec{\Phi}]$. (This follows by considering the case $l = k$, $\Psi_i = \Gamma_i[\vec{\varphi}/\vec{\Phi}]$.)

Let us note that if a new law or inference rule is added to a derivation system, previously derivable rules do not cease to be derivable, but previously admissible rules may cease to be admissible.

Example 4.50. One might be suspecting that a rule $\Gamma \,/\, \Delta$ is admissible if and only if the meta-formula $\mathcal{H} := (\mathbf{2})\Big(\big(\mathcal{D} \Rightarrow (\mathbf{1}) \Gamma \big) \implies \big(\mathcal{D} \Rightarrow (\mathbf{1}) \Delta \big) \Big)$ is derivable. This is not so, since the derivability of $\mathcal{H}$ is in fact equivalent to the derivability, and not admissibility of the rule $\Gamma \,/\, \Delta$.

Indeed, the derivability of $\mathcal{H}$ amounts to the derivability of $\mathcal{D} \Rightarrow (\mathbf{1}) \Delta$ from the assumption $\mathcal{D} \Rightarrow (\mathbf{1}) \Gamma$. This in turn amounts to the derivability of $(\mathbf{1}) \Delta$ from the assumptions $\mathcal{D}$ and $\mathcal{D} \Rightarrow (\mathbf{1}) \Gamma$. But derivability from these assumptions is equivalent to

derivability from the assumptions $\mathcal{D}$ and $(1)\,\Gamma$. In turn, the derivability of $(1)\,\Delta$ from the assumptions $\mathcal{D}$ and $(1)\,\Gamma$ amounts to the derivability of $(1)\,\Gamma \Rightarrow (1)\,\Delta$ from the assumption $\mathcal{D}$. This is the same as the derivability of the rule $\Gamma \,/\, \Delta$.

Clearly, derivability and admissibility of a rule are preserved when the rule is replaced by one that is interderivable with it. In particular, an admissible rule that is interderivable with a principle must be derivable. Jeřábek [102] proved that in a number of logics, including zero-order intuitionistic logic and S4, every rule is interderivable with a principle plus a finite set of admissible rules.

Lorentzen's book, where the term "admissible rule" was first used, also dealt with the following notion: a rule $\mathcal{R}$ is *admissible relative to* a set of rules $\mathcal{S}_1, \ldots, \mathcal{S}_k$ if $\mathcal{R}$ becomes an admissible rule upon adjoining $\mathcal{S}_1, \ldots, \mathcal{S}_k$ as new inference rules to the logic [124; pp. 24, 40] (see also [177]). When $\mathcal{R}$ is a principle, this coincides with the notion of consequence; and when $k = 0$, the two notions (i.e., consequence and relative admissibility) reduce, respectively, to derivability of a rule and admissibility of a rule.

We will call a rule $\mathcal{R}$ *stably admissible* over a given logic if it is admissible over any extension of this logic, obtained by adjoining any number of new laws and inference rules. Thus derivable rules are stably admissible; and in contrast to admissible rules, if a new law or inference rule is added to the derivation system, previously stably admissible rules do not cease to be stably admissible. (However, a stably admissible rule can cease to be stably admissible upon a mere extension of the language of the logic, see Examples 4.71 and 4.72.)

Clearly, a purely logical rule $\Gamma_1, \ldots, \Gamma_m \,/\, \Delta$ is stably admissible if and only if for every its special case $\Gamma_1[\vec{\varphi}/\vec{\Phi}], \ldots, \Gamma_m[\vec{\varphi}/\vec{\Phi}] \,/\, \Delta[\vec{\varphi}/\vec{\Phi}]$ (where $\vec{\Phi}$ is free for $\vec{\varphi}$ in each Γ_i and in Δ) the principle $\cdot\Delta[\vec{\varphi}/\vec{\Phi}]$ is a consequence of the principles $\cdot\Gamma_1[\vec{\varphi}/\vec{\Phi}], \ldots, \cdot\Gamma_m[\vec{\varphi}/\vec{\Phi}]$. This uses the characterization of consequence between principles given in §4.5.3. Since that characterization involved some freedom, we obtain as a byproduct that a purely logical rule $\mathcal{R}$ is stably admissible over a logic if and only if it is admissible over any extension of this logic, obtained by adjoining any number of new laws (but not inference rules). By using the same freedom, we also get that $\mathcal{R}$ is stably admissible if and only if $\mathcal{G} \vdash \mathcal{R} \Rightarrow \Phi$ implies $\mathcal{G} \vdash \Phi$ for every formula Φ and every λ-closed meta-formula $\mathcal{G}$. This is equivalent to saying that $\vdash (\mathcal{R} \Rightarrow \Phi) \Rightarrow \Phi$ for every formula Φ (by considering $\mathcal{G} = \mathcal{R} \Rightarrow \Phi$).

On the other hand, it is easy to see that a rule $\mathcal{R}$ is derivable if and only if $\vdash \mathcal{R} \Rightarrow \mathcal{F}$ implies $\vdash \mathcal{F}$ for every rule $\mathcal{F}$ (by considering $\mathcal{F} = \mathcal{R}$). Similarly, a rule $\mathcal{R}$ is derivable if and only if $\mathcal{G} \vdash \mathcal{R} \Rightarrow \mathcal{F}$ implies $\mathcal{G} \vdash \mathcal{F}$ for every meta-formula $\mathcal{F}$ and every λ-closed meta-formula $\mathcal{G}$. Thus, similarly to the above, a rule $\mathcal{R}$ is derivable if and only if $\vdash (\mathcal{R} \Rightarrow \mathcal{F}) \Rightarrow \mathcal{F}$ for every rule $\mathcal{F}$ (or for every λ-closed meta-formula $\mathcal{F}$).

Admissibility, stable admissibility and derivability are all distinct notions, already in classical logic (see Examples 4.69, 4.71 and 4.72).

4.6. Intuitionistic logic (syntax)

4.6.1. Derivation system. A formalization of intuitionistic logic QH in terms of a derivation system was found by Heyting [83] (1928 and 1931), extending partial formalizations by Kolmogorov [107] (1925; the $\to, \neg, \exists, \forall$ fragment[67] with the omission of the explosion principle $\cdot \times \to \alpha$ and with inadvertent omission of the universal specialization principle $\cdot \forall x\, \alpha(x) \to \alpha(y)$; see [159]) and Glivenko (1928 and 1929; the zero-order fragment).[68] Here is an equivalent formulation due to Spector (as in [200]; see also [199]), where the usual side conditions are eliminated as explained below.

(I) (modus ponens rule) $\dfrac{\alpha,\ \alpha \to \beta}{\beta}$

(II) $\cdot \alpha \to \alpha$

(III) $\dfrac{\alpha \to \beta,\ \beta \to \gamma}{\alpha \to \gamma}$

(IV) (exponential law) $\dfrac{(\alpha \wedge \beta) \to \gamma}{\alpha \to (\beta \to \gamma)}$ and $\dfrac{\alpha \to (\beta \to \gamma)}{(\alpha \wedge \beta) \to \gamma}$

(V) (explosion principle) $\cdot \times \to \alpha$

(VI) $\cdot \alpha \wedge \beta \to \alpha$ and $\cdot \alpha \wedge \beta \to \beta$

(VII) $\cdot \alpha \to \alpha \vee \beta$ and $\cdot \beta \to \alpha \vee \beta$

(VIII) $\dfrac{\gamma \to \alpha,\ \gamma \to \beta}{\gamma \to \alpha \wedge \beta}$

(IX) $\dfrac{\alpha \to \gamma,\ \beta \to \gamma}{\alpha \vee \beta \to \gamma}$

(X) $\cdot \forall x\, \alpha(x) \to \alpha(t)$

(XI) $\cdot \alpha(t) \to \exists x\, \alpha(x)$

(XII) $\dfrac{\gamma \to \alpha(x)}{\gamma \to \forall x\, \alpha(x)}$

(XIII) $\dfrac{\alpha(x) \to \gamma}{\exists x\, \alpha(x) \to \gamma}$

Classical predicate logic QC is obtained by adding just one more law:

(XIV) (law of excluded middle) $\cdot p \vee \neg p$

Let us note that the traditional schematic versions of (X) and (XI) with side conditions follow immediately from (X) and (XI) respectively by applying the two specialization meta-rules (see §*4.4.1*), in either order. We have a notational choice of specializing the

---

[67]Kolmogorov's point of departure was Hilbert's derivation system for classical logic, which did not use other connectives. Presumably Kolmogorov was also aware of Russell's definition of $\wedge, \vee$ and $\exists$ in terms of $\to$ and the second-order $\forall$, which is valid intuitionistically, see [5] and [164; p. 67].

[68]It should be mentioned that an algebraic formulation of the zero-order fragment is found already in Skolem's 1919 paper (see [210]).

unary predicate variable α either to a 1-formula of the form $x \mapsto \Theta$, in which case we get

$$\vdash \forall x\, \Theta \to \Theta[x/T], \text{ provided that } T \text{ is free for } x \text{ in } \Theta;$$
$$\vdash \Theta[x/T] \to \exists x\, \Theta, \text{ provided that } T \text{ is free for } x \text{ in } \Theta,$$

or to an arbitrary 1-formula Σ, in which case we get

$$\vdash \forall x\, \Sigma(x) \to \Sigma(T), \text{ provided that } T \text{ is free for } x \text{ in } \Sigma(x) \text{ and } x \text{ does not occur freely in } \Sigma;$$
$$\vdash \Sigma(T) \to \exists x\, \Sigma(x), \text{ provided that } T \text{ is free for } x \text{ in } \Sigma(x) \text{ and } x \text{ does not occur freely in } \Sigma,$$

where the λ-expressions $\Sigma(x)$, $\Sigma(T)$ stand for their normal forms (which are formulas).

Let us note that the second side condition, "provided that x does not occur freely in Σ", cannot be dropped. For instance, if Σ is the arithmetical 1-formula $y \mapsto x = y$, and T is the arithmetical term $x + 1$, then $\forall x\, \Sigma(x) \to \Sigma(T)$ is not a valid formula.

Similarly, the traditional schematic versions of (XII) and (XIII) with side conditions follow immediately from (XII) and (XIII) respectively by applying the second-order meta-specialization and (for one of the traditions) the first-order meta-generalization (see §4.4.1). Namely, by specializing γ to a formula Γ, and α to a 1-formula of the form $x \mapsto \Theta$, we get

$$(x)\, \Gamma \to \Theta \;\vdash\; \Gamma \to \forall x\, \Theta, \text{ provided that } x \text{ does not occur freely in } \Gamma;$$
$$(x)\, \Theta \to \Gamma \;\vdash\; \exists x\, \Theta \to \Gamma, \text{ provided that } x \text{ does not occur freely in } \Gamma,$$

which corresponds to the tradition of "varied variables" (in Kleene's terminology), found in the textbooks by Schoenfield and Mendelson. Now by combining these with the meta-generalization with respect to x, we get

$$\Gamma \to \Theta \;\vdash\; \Gamma \to \forall x\, \Theta, \text{ provided that } x \text{ does not occur freely in } \Gamma \text{ and in the assumptions;}$$
$$\Theta \to \Gamma \;\vdash\; \exists x\, \Theta \to \Gamma, \text{ provided that } x \text{ does not occur freely in } \Gamma \text{ and in the assumptions,}$$

which corresponds to the tradition of "fixed variables", found in the textbooks by Church, Troelstra and van Dalen. Like before, we can replace Θ with $\Sigma(x)$, at the cost of the additional side condition "provided that x does not occur freely in Σ".

In particular, the special case of (XII) with $\gamma = \checkmark$, when rewritten as

$(\text{XII}_\checkmark)$ (generalization rule) $\dfrac{\alpha(x)}{\forall x\, \alpha(x)}$

(see Example 4.51 below), corresponds to the following traditional schematic forms of the generalization rule:

$$(x)\,\Sigma(x)\ \vdash\ \forall x\,\Sigma(x), \text{ provided that } x \text{ does not occur freely in } \Sigma;$$

$$\Sigma(x)\ \vdash\ \forall x\,\Sigma(x), \text{ provided that } x \text{ does not occur freely in } \Sigma \text{ and in the assumptions,}$$

where the λ-expression $\Sigma(x)$ again stands for its normal form (which is a formula). The side conditions cannot be dropped here as well (see Example 4.39 and Remark 4.40).

Example 4.51. Let us show that formulas Φ and $\checkmark \to \Phi$ are interderivable in intuitionistic logic. Indeed, by (VI) $\vdash \Phi \wedge \checkmark \to \Phi$, hence by the exponential law, $\vdash \Phi \to (\checkmark \to \Phi)$, and thus by *modus ponens*, $\Phi \vdash \checkmark \to \Phi$. Conversely, since $\checkmark = \times \to \times$, by (II) $\vdash \checkmark$, and therefore by *modus ponens*, $\checkmark \to \Phi \vdash \Phi$.

4.6.2. Modifications and dualizations. Since intuitionistic logic satisfies the deduction theorem (see [200], where it is proved in Spector's derivation system) and the exponential law, $\Phi, \Psi \vdash \Xi$ is equivalent to $\vdash \Phi \wedge \Psi \to \Xi$ (see §*4.5.3*).

Thus, should we wish to have fewer inference rules, we may well replace the inference rules in (III), (IV), (VIII) and (IX) by the corresponding laws (in which the horizontal bar is replaced by $\to$, and the comma by $\wedge$).

As for the inference rules (XII) and (XIII), for many purposes it is convenient to replace them by the generalization rule (XII$_\checkmark$) along with the following ones:

$$(\text{XII}')\quad \frac{\forall x\,(\gamma \to \alpha(x))}{\gamma \to \forall x\,\alpha(x)}$$

$$(\text{XIII}')\quad \frac{\forall x\,(\alpha(x) \to \gamma)}{\exists x\,\alpha(x) \to \gamma}$$

Indeed, using the generalization rule (XII$_\checkmark$) and its converse, which follows from (X), we can clearly exchange (XII$'$) for (XII) and (XIII$'$) for (XII).

Now by the deduction theorem, the inference rules (XII$'$) and (XIII$'$) can be replaced by the corresponding laws (in which the horizontal bar is replaced by $\to$).

The generalization rule is, however, strictly weaker than the corresponding principle, $\cdot\,\alpha(x) \to \forall x\,\alpha(x)$ (see §*4.7.1*). Informally, this can be explained as follows. If, for instance, $\neg x = x + 1$ has been proved, we should be able to conclude $\forall x\ \neg x = x + 1$; yet from $\exists x\ x = 1$ we cannot conclude $\exists x\,\forall x\ x = 1$ (which is equivalent to $\exists y\,\forall x\ x = 1$); also, from $\neg\forall x\ x = 1$ we cannot conclude $\neg x = 1$ (which, if proved, would imply $\forall x\ \neg x = 1$).

Let us note dual forms of the rules (XII$'$) and (XIII$'$):

$$(\text{XII}'^*)\quad \frac{\exists x\,(\gamma \to \alpha(x))}{\gamma \to \exists x\,\alpha(x)}$$

$$(\text{XIII}'^*)\quad \frac{\exists x\,(\alpha(x) \to \gamma)}{\forall x\,\alpha(x) \to \gamma}$$

These, or rather their corresponding principles, follow from the laws (XI) and (X) respectively, using the rule (XIII) and the generalization rule. Similarly, there are also dual forms of the rules (VIII) and (IX).

The laws (X) and (XI) also imply the converses of the rules (XII) and (XIII), and hence (using the generalization rule) also the converses of (XII′) and (XIII′). However, the converses of the rules (XII′*) and (XIII′*) are not derivable in intuitionistic logic (see 5.2.2), and therefore not even admissible in it, since they are easily seen to be interderivable with the principles $\cdot \exists t \left(\exists x\, \alpha(x) \to \alpha(t) \right)$ and $\cdot \exists t \left(\alpha(t) \to \forall x\, \alpha(x) \right)$, respectively. In classical logic, these principles are easily seen to be derivable.

Dual to the generalization rule is the following instantiation rule

$$(\text{XII}^*_{\vee})\ \ \frac{\exists x\, \gamma}{\gamma},$$

whose corresponding principle follows from (XIII) and (II). The more general rule $\exists x\, \alpha(x)\, /\, \alpha(x)$, which is the "contrapositive" of the generalization rule, is not even admissible in classical logic (by considering any of the two classically derivable formulas just mentioned above, $\exists x \left(\exists y\, \beta(y) \to \beta(x) \right)$ or $\exists x \left(\beta(x) \to \forall y\, \beta(y) \right)$).

The rule $\exists x\, \alpha(x)\, /\, \alpha(x)$ is also not admissible in intuitionistic logic, by considering the intuitionistically derivable formula $\exists x \left(\beta(x) \to \beta(y) \right)$ (see Example 4.38). But the second variable is essential here: as long as the language contains no function symbols, and Φ is a closed 1-formula, $\vdash \exists x\, \Phi(x)$ implies $\vdash \Phi(x)$ by the "existence property" of intuitionistic logic (Gentzen; see [104]). The related "disjunction property" will be discussed in §5.3 below.

4.6.3. Additional meta-connectives. In studying models, it seems worthwhile to be armed with what was not needed previously: the *meta-absurdity* ✳, the *meta-disjunction* $\mathcal{F} \parallel \mathcal{G}$ and the *existential first-order meta-quantifier* $\langle x \rangle\, \mathcal{F}$. These are assumed to satisfy the following meta-rules, where $\mathcal{F}$, $\mathcal{G}$ and $\mathcal{H}$ are meta-formulas, x is an individual variable and T is a term.

$$(\text{✳-elimination}) \quad \frac{\displaystyle \genfrac{}{}{0pt}{}{\vdots}{\text{✳}}}{\mathcal{F}}$$

$$(\parallel\text{-introduction}) \quad \frac{\genfrac{}{}{0pt}{}{\vdots}{\mathcal{F}}}{\mathcal{F} \parallel \mathcal{G}} \quad \text{and} \quad \frac{\genfrac{}{}{0pt}{}{\vdots}{\mathcal{G}}}{\mathcal{F} \parallel \mathcal{G}}$$

$$(\parallel\text{-elimination}) \quad \frac{\mathcal{F} \parallel \mathcal{G} \qquad \overset{[\mathcal{F}]}{\underset{\mathcal{H}}{\vdots}} \qquad \overset{[\mathcal{G}]}{\underset{\mathcal{H}}{\vdots}}}{\mathcal{H}}$$

$$(\langle\cdot\rangle\text{-introduction}) \quad \frac{\begin{array}{c}\vdots\\\mathcal{F}[x/T]\end{array}}{\langle x\rangle\,\mathcal{F}}, \text{ provided that } T \text{ is free for } x \text{ in } \mathcal{F}$$

$$(\langle\cdot\rangle\text{-elimination}) \quad \frac{\langle x\rangle\,\mathcal{F} \qquad \overset{[\mathcal{F}]}{\underset{\vdots}{\mathcal{H}}}}{\mathcal{H}}, \text{ provided that } x \text{ does not occur freely in } \mathcal{H} \text{ and}$$

in the assumptions.

On the other hand, Russell's well-known second-order definitions of $\times$, $\vee$, $\exists$ and $\wedge$ (see [5] and [64; §11.3.2]) suggest the following versions of $\divideontimes$, $\|$, $\langle\cdot\rangle$ and $\&$:

- $\tilde{\divideontimes} := (\alpha)\,\alpha$;
- $\mathcal{F} \,\tilde{\|}\, \mathcal{G} := (\gamma)\left((\mathcal{F} \Rightarrow \gamma)\,\&\,(\mathcal{G} \Rightarrow \gamma) \Rightarrow \gamma\right)$;[69]
- $\widetilde{\langle x\rangle}\mathcal{F} := (\gamma)\left((x)\,(\mathcal{F} \Rightarrow \gamma) \Rightarrow \gamma\right)$;
- $\mathcal{F} \,\tilde{\&}\, \mathcal{G} := (\gamma)\left(\left(\mathcal{F} \Rightarrow (\mathcal{G} \Rightarrow \gamma)\right) \Rightarrow \gamma\right)$,

where γ is a nullary predicate variable that does not occur freely in the meta-formulas $\mathcal{F}$, $\mathcal{G}$.

When $\mathcal{F}$ and $\mathcal{G}$ are formulas (regarded as atomic meta-formulas), we can identify $\&$ with $\tilde{\&}$; $\|$ with $\tilde{\|}$; and $\langle\cdot\rangle$ with $\widetilde{\langle\cdot\rangle}$. We only include full details for the case of meta-conjunction:

Proposition 4.52. *(a)* $\Phi \,\tilde{\&}\, \Psi \iff \Phi\,\&\,\Psi$ *is deducible for all formulas* Φ, Ψ;

(b) each of Φ, Ψ *can be deduced from* $\Phi \,\tilde{\&}\, \Psi$, *and* $\Phi \,\tilde{\&}\, \Psi$ *can be deduced from* Φ *and* Ψ *without using the* $\&$*-introduction and the* $\&$*-elimination meta-rules;*

(c) Let $\mathcal{H}$ *be a meta-formula where all occurrences of* $\&$ *are applied to formulas. Then* $\mathcal{H}$ *is deducible if and only if the meta-formula obtained from* $\mathcal{H}$ *by replacing every occurrence of* $\&$ *with a* $\tilde{\&}$ *is deducible without using the* $\&$*-introduction and the* $\&$*-elimination meta-rules.*

Proof. (b). To deduce Φ from $\Phi \,\tilde{\&}\, \Psi$ we use the meta-specialization with $\gamma = \Phi$ along with the fact that $\Phi \Rightarrow (\Psi \Rightarrow \Phi)$ is deducible (see Example 4.28). To deduce Ψ from $\Phi \,\tilde{\&}\, \Psi$ we use the meta-specialization with $\gamma = \Psi$ along with the fact that $\Phi \Rightarrow (\Psi \Rightarrow \Psi)$ is deducible (see Example 4.28).

To deduce $\Phi \,\tilde{\&}\, \Psi$ from Φ and Ψ, let us first note that γ can be deduced from Φ, Ψ and $\Phi \Rightarrow (\Psi \Rightarrow \gamma)$ by a double application of the $\Rightarrow$-elimination. Now by the $\Rightarrow$-introduction, we get a deduction of $\left(\Phi \Rightarrow (\Psi \Rightarrow \gamma)\right) \Rightarrow \gamma$ from Φ and Ψ, and it remains to apply the meta-generalization. $\square$

(a). The $\&$-elimination meta-rule along with the $\tilde{\&}$-introducing deduction provided by (b) yield a deduction of $\Phi \,\tilde{\&}\, \Psi$ from $\Phi\,\&\,\Psi$. Similarly one gets a deduction of $\Phi\,\&\,\Psi$ from $\Phi \,\tilde{\&}\, \Psi$. The assertion now follows by the $\Rightarrow$-introduction. $\square$

---

[69]Let us note that $\&$ can be eliminated here by a mere application of the exponential meta-law.

(c). The "if" assertion follows from (a), and the "only if" assertion from (b). $\square$

Let us also discuss the special case of the $\langle\cdot\rangle$-rules where x does not occur freely in Φ.

Example 4.53. Let us show that if Φ is a formula, $\Phi \iff (\gamma)\big((\Phi \Rightarrow \gamma) \Rightarrow \gamma\big)$ is deducible as long as γ does not occur freely in Φ.

The $\Leftarrow$ implication follows by meta-specializing γ to Φ.

To get the converse implication, by the meta-generalization, it suffices to deduce $(\Phi \Rightarrow \gamma) \Rightarrow \gamma$ from Φ. But $\Phi \Rightarrow \big((\Phi \Rightarrow \gamma) \Rightarrow \gamma\big)$ is easily deducible by a double application of the exponential meta-law.

Proposition 4.54. *In intuitionistic logic,*
   (a) $\vdash \varphi \iff (\gamma)(\varphi \to \gamma) \to \gamma$;
   (b) $\vdash \alpha \wedge \beta \iff (\gamma)\big(\alpha \to (\beta \to \gamma)\big) \to \gamma$;
   (c) $\vdash \alpha \vee \beta \iff (\gamma)(\alpha \to \gamma) \to \big((\beta \to \gamma) \to \gamma\big)$;
   (d) $\vdash \exists x\, \alpha(x) \iff (\gamma)\forall x\big(\alpha(x) \to \gamma\big) \to \gamma$;
   (e) $\vdash \times \iff (\gamma)\gamma$.

Parts (b), (c), (d),(e) yield meta-logical "definitions", up to meta-equivalence, of $\times$; of $\wedge$ and $\vee$ in terms of $\to$; and of $\exists$ in terms of $\to$ and $\forall$. Of course, these "definitions" also apply in classical logic. Let us note that part (e) says that $\vdash \times \Leftrightarrow \tilde{\divideontimes}$.

Proof. (a). This is similar to Example 4.53. The $\Leftarrow$ follows by meta-specializing γ to φ.

To get the converse, by the meta-generalization, it suffices to derive $(\varphi \to \gamma) \to \gamma$ from φ. By the modus ponens rule, it suffices to derive $\varphi \to \big((\varphi \to \gamma) \to \gamma\big)$. But it is easily derivable by a double application of the exponential law. $\square$

(b),(c),(d),(e). These follow from (a) by appropriately meta-specializing φ. Namely, (b) follows by the exponential law; (c) by (24) and the exponential law; (d) by (16); and (e) is straightforward. $\square$

Proposition 4.55. *In intuitionistic logic,*
   (a) $\vdash \Phi \wedge \Psi \Leftrightarrow \Phi \,\&\, \Psi$ *for any formulas* Φ, Ψ;
   (b) $\vdash \forall x\, \Phi \Leftrightarrow (x)\,\Phi$ *for any formula* Φ;
   (c) $\vdash \divideontimes \Rightarrow \times$;
   (d) $\vdash \Phi \parallel \Psi \Rightarrow \Phi \vee \Psi$ *for any formulas* Φ, Ψ;
   (e) $\vdash \langle x \rangle\, \Phi \Rightarrow \exists x\, \Phi$ *for any formula* Φ;
   (f) $\vdash \Phi \to \Psi \Rightarrow (\Phi \Rightarrow \Psi)$ *for any formulas* Φ, Ψ.

This clearly transfers to any extension of intuitionistic logic, including classical logic.

The converses to parts (d), (e), (f) do not hold already in zero-order classical logic (and hence also in zero-order intuitionistic logic), see Examples 4.78 and 4.76.

Proof. (a). From (VIII) with $\gamma = \checkmark$ and the $\&$-elimination meta-rule we get $\vdash \Phi \,\&\, \Psi \Rightarrow \Phi \wedge \Psi$. Conversely, from (VI) and the *modus ponens* rule we have $\vdash \Phi \wedge \Psi \Rightarrow \Phi$ and $\vdash \Phi \wedge \Psi \Rightarrow \Psi$. Hence by the $\&$-introduction, $\vdash \Phi \wedge \Psi \Rightarrow \Phi \,\&\, \Psi$. $\square$

(b). Here $\Leftarrow$ follows from the generalization rule, and $\Rightarrow$ from (X) and the *modus ponens* rule, along with the generalization meta-rule. □

(c). This follows from the ✳-elimination rule. □

(d). From (VII) and the *modus ponens* rule we have $\vdash \Phi \Rightarrow \Phi \vee \Psi$ and $\vdash \Psi \Rightarrow \Phi \vee \Psi$. Hence by the $\parallel$-elimination, $\vdash \Phi \parallel \Psi \Rightarrow \Phi \vee \Psi$. □

(e). This is similar to (d). □

(f). This follows from the *modus ponens* rule. □

Remark 4.56. If $\mathcal{F}$ is a meta-formula with no occurrences of ✳, then $\tilde{✳} \Rightarrow \mathcal{F}$ is deducible. Indeed, from $\tilde{✳}$ we can deduce any formula (regarded as an atomic meta-formula) by the meta-specialization. The assertion now follows by induction, using &- $\Rightarrow$-, $\parallel$-, $\langle \cdot \rangle$-introduction and the meta-generalization.

4.7. Models

4.7.1. Meta-models. Let us call a meta-rule *absolute* if it involves no conditions on the assumptions of the deduction. In particular, the $\Rightarrow$-elimination meta-rule is absolute; and so is every meta-law. Thus the Enderton-style system of §*4.4.4* consists entirely of absolute meta-rules.

Let us fix a language $\mathscr{L}$. An absolute meta-rule is said to be *satisfied* in a meta-$\mathscr{L}$-structure $\jmath$ if for every its instance

$$\frac{\mathcal{F}_1, \ldots, \mathcal{F}_n}{\mathcal{G}}$$

and every individual assignment u and every predicate valuation v, if $?(|\mathcal{F}_i|_{\jmath}^{uv}) = \mathbb{T}$ for each i, then $?(|\mathcal{G}|_{\jmath}^{uv}) = \mathbb{T}$.

In particular, a meta-law is satisfied in $\jmath$ if and only $?(|\mathcal{G}|_{\jmath}^{uv}) = \mathbb{T}$ for every its instance $\mathcal{G}$ and every individual assignment u and every predicate valuation v. Let us note also that the α-conversion rule is trivially satisfied in every meta-$\mathscr{L}$-structure.

A meta-$\mathscr{L}$-structure $\jmath$ will be called a *model of the meta-logic* if it satisfies all meta-rules of the Enderton-style formulation of the meta-logic in §*4.4.4*.

It is easy to see that if an absolute meta-rule is derivable in the Enderton-style system, then it is satisfied in any model of the meta-logic.

This applies, in particular, to the &-introduction and &-elimination meta-rules. Thus we obtain

Proposition 4.57. *Let $\jmath$ be a model of the meta-logic, $\mathcal{F}$ and $\mathcal{G}$ meta-formulas, u a variable assignment and v a predicate valuation.*
Then $|\mathcal{F} \,\&\, \mathcal{G}|_{\jmath}^{uv} = \mathbb{T}$ if and only if $|\mathcal{F}|_{\jmath}^{uv} = \mathbb{T}$ and $|\mathcal{G}|_{\jmath}^{uv} = \mathbb{T}$.

The same goes for the specialization meta-rules. However, the generalization meta-rules are not absolute.

Let us call a meta-$\mathscr{L}$-structure j a *semi-standard model of the meta-logic* if it satisfies all absolute meta-rules of the Hilbert-style formulation of the meta-logic in §4.4.4 and also the following two conditions (corresponding to the generalization meta-rules):

(A) If $?(|\mathcal{F}|_{j}^{u'v}) = \top$ for every variable assignment u' that agrees with u on all individual variables other than x, then $?(|(x)\,\mathcal{F}|_{j}^{uv}) = \top$;

(B) if $?(|\mathcal{F}|_{j}^{uv'}) = \top$ for every predicate valuation v' that agrees with v on all predicate variables other than γ, then $?(|(\gamma)\,\mathcal{F}|_{j}^{uv}) = \top$.

It is easy to see, following the straightforward derivation of the Enderton-style system from the Hilbert-style system, that every semi-standard model of the meta-logic is a model of the meta-logic.

Proposition 4.58. *Let j be a semi-standard model of the meta-logic, $\mathcal{F}$ and $\mathcal{G}$ meta-formulas, u a variable assignment and v a predicate valuation.*

(a) $?(|(x)\,\mathcal{F}|_{j}^{uv}) = \top$ if and only if $?(|\mathcal{F}|_{j}^{u'v}) = \top$ for every variable assignment u' that agrees with u on all individual variables other than x.

(b) $?(|(\gamma)\,\mathcal{F}|_{j}^{uv}) = \top$ if and only if $?(|\mathcal{F}|_{j}^{uv'}) = \top$ for every predicate valuation v' that agrees with v on all predicate variables other than γ.

Proof. Let us prove (a). The "if" assertion is precisely the condition corresponding to the generalization meta-rule. So it remains to prove the "only if" assertion. The condition corresponding to the specialization meta-rule is: for any $\mathcal{F}$, u and v, if $?(|(x)\,\mathcal{F}|_{j}^{uv}) = \top$, then $?(|\mathcal{F}[x/T]|_{j}^{uv}) = \top$, as long as T is free for x in $\mathcal{F}$.

Let y be a variable that does not occur freely in $\mathcal{F}$. Given an assignment u' as above, let u'' be the assignment with $u''(y) = u'(x)$ and $u''(z) = u(z)$ for every variable z other than y. Since y does not occur freely in $\mathcal{F}$, we have $|(x)\,\mathcal{F}|_{j}^{uv} = |(x)\,\mathcal{F}|_{j}^{u''v}$. By the condition corresponding to the meta-specialization rule, if $?(|(x)\,\mathcal{F}|_{j}^{u''v}) = \top$, then $?(|\mathcal{F}[x/y]|_{j}^{u''v}) = \top$. Since y does not occur freely in $\mathcal{F}$, we also have $|\mathcal{F}[x/y]|_{j}^{u''v} = |\mathcal{F}|_{j}^{u'v}$. Thus if $?(|(x)\,\mathcal{F}|_{j}^{uv}) = \top$, then $?(|\mathcal{F}|_{j}^{u'v}) = \top$.

The proof of (b) is similar. $\square$

Let us note that while the $\Rightarrow$-elimination meta-rule is absolute, the $\Rightarrow$-introduction meta-rule is not.

Let us call a meta-$\mathscr{L}$-structure j a *standard model of the meta-logic* if it satisfies all absolute meta-rules of the original natural deduction formulation of the meta-logic in §4.4.1, the conditions (A) and (B) above (corresponding to the generalization meta-rules), and also the following condition (corresponding to the $\Rightarrow$-introduction meta-rule):

(C) If either $?(|\mathcal{F}|_{j}^{uv}) = \bot$ or $?(|\mathcal{G}|_{j}^{uv}) = \top$, then $?(|\mathcal{F} \Rightarrow \mathcal{G}|_{j}^{uv}) = \top$.

It is easy to see that any standard model of the meta-logic is a semi-standard model of the meta-logic.

Condition (C) is converse to the condition of satisfying the $\Rightarrow$-elimination meta-rule. Thus we obtain

Proposition 4.59. *Let j be a standard model of the meta-logic, $\mathcal{F}$ and $\mathcal{G}$ meta-formulas, u a variable assignment and v a predicate valuation.*

$?(|\mathcal{F} \Rightarrow \mathcal{G}|_j^{uv}) = \mathbb{T}$ *if and only if* $?(|\mathcal{F}|_j^{uv}) = \mathbb{T}$ *implies* $?(|\mathcal{G}|_j^{uv}) = \mathbb{T}$.

It is easy to see that a two-valued meta-interpretation $j = i_+$ is always a standard model of the meta-logic.

In fact, the converse also holds: any standard model of the meta-logic reduces to a two-valued one.

Theorem 4.60. *Let j be a standard model of the meta-logic. Then there exists a two-valued model i_+ of the meta-logic such that any given meta-formula $\mathcal{F}$ is valid in j if and only if it is valid in i_+.*

Proof. The given model j includes an $\mathscr{L}$-structure i, which in turn determines a two-valued model i_+ of the meta-logic, with $!_{i_+} \colon \mathcal{O} \to \{\mathbb{T}, \bot\}$ defined as the composition $\mathcal{O} \xrightarrow{!_j} \mathcal{Q} \xrightarrow{?_j} \{\mathbb{T}, \bot\}$. For any atomic meta-formula $\mathcal{F}$ and any variable assignment u and any predicate valuation v we have $?_j(|\mathcal{F}|_j^{uv}) = \mathbb{T}$ if and only if $|\mathcal{F}|_{i_+}^{uv} = \mathbb{T}$. Now from 4.57, 4.58 and 4.59 we get that the same holds for an arbitrary meta-formula $\mathcal{F}$. In particular, a meta-formula $\mathcal{F}$ is valid in j if and only if it is valid in i_+. $\square$

4.7.2. Two-valued meta-models. Let us now record the conditions corresponding to all meta-rules (also for the additional meta-connectives) in two-valued models of the meta-logic.

- $|\mathcal{F} \mathbin{\&} \mathcal{G}|_{i_+}^{uv} = \mathbb{T}$ if and only if $|\mathcal{F}|_{i_+}^{uv} = \mathbb{T}$ and $|\mathcal{G}|_{i_+}^{uv} = \mathbb{T}$.

- $|\mathcal{F} \mathbin{\|} \mathcal{G}|_{i_+}^{uv} = \mathbb{T}$ if and only if $|\mathcal{F}|_{i_+}^{uv} = \mathbb{T}$ or $|\mathcal{G}|_{i_+}^{uv} = \mathbb{T}$.

- $|\mathcal{F} \Rightarrow \mathcal{G}|_{i_+}^{uv} = \mathbb{T}$ if and only if $|\mathcal{F}|_{i_+}^{uv} = \mathbb{T}$ implies $|\mathcal{G}|_{i_+}^{uv} = \mathbb{T}$.

- $|(x)\,\mathcal{F}|_{i_+}^{uv} = \mathbb{T}$ if and only if $|\mathcal{F}|_{i_+}^{u'v} = \mathbb{T}$ for every variable assignment u' that agrees with u on all individual variables other than x.

- $|(\gamma)\,\mathcal{F}|_{i_+}^{uv} = \mathbb{T}$ if and only if $|\mathcal{F}|_{i_+}^{uv'} = \mathbb{T}$ for every predicate valuation v' that agrees with v on all predicate variables other than γ.

- $|\langle x \rangle\,\mathcal{F}|_{i_+}^{uv} = \mathbb{T}$ if and only if $|\mathcal{F}|_{i_+}^{u'v} = \mathbb{T}$ for some variable assignment u' that agrees with u on all individual variables other than x.

- $|\divideontimes|_{i_+} = \mathbb{T}$ if and only if $!(\mathcal{O}) = \{\mathbb{T}\}$.

Here are some special cases:

Proposition 4.61. *Let i be an $\mathscr{L}$-structure and $\mathcal{F}$ a meta-formula.*

(a) Let v be a predicate valuation. Then $|(1)\,\mathcal{F}|_{i_+}^{v} = \mathbb{T}$ if and only if $|\mathcal{F}|_{i_+}^{uv} = \mathbb{T}$ for every variable assignment u.

(b) Suppose that $\mathcal{F}$ is closed. Then $|(2)\,\mathcal{F}|_{i_+} = \mathbb{T}$ if and only if $|\mathcal{F}|_{i_+}^{v} = \mathbb{T}$ for every predicate valuation v;

(c) Let v be a predicate valuation. Then $|\langle 1 \rangle \mathcal{F}|^{v}_{\iota_+} = \top$ if and only if $|\mathcal{F}|^{uv}_{\iota_+} = \top$ for some variable assignment u.

Let us note that the proposition can be rephrased as follows:

(a) a meta-formula $\mathcal{F}$ is valid in (ι, v) if and only if $(1)\mathcal{F}$ is valid in (ι, v);

(b) a closed meta-formula $\mathcal{F}$ is valid in ι if and only if $(2)\mathcal{F}$ is valid in ι;

(c) a meta-formula $\mathcal{F}$ is satisfiable in (ι, v) if and only if $\langle 1 \rangle \mathcal{F}$ is valid in (ι, v).

Corollary 4.62. *Let ι be an $\mathscr{L}$-structure.*

(a) A principle $\cdot\,\Phi$ is valid in ι if and only if the formula Φ is valid in ι (that is, $!(|\Phi|^{uv}_{\iota}) = \top$ for every variable assignment u and every predicate valuation v).

(b) A rule $\Phi_1, \ldots, \Phi_n \,/\, \Psi$ is valid in ι if and only if for every predicate valuation v, if the formulas $\Phi_1, \ldots, \Phi_n$ are valid in (ι, v), then Ψ is also valid in (ι, v) (in other words, if $!(|\Phi_j|^{uv}_{\iota}) = \top$ for each j and every variable assignment u, then $!(|\Psi|^{uv}_{\iota}) = \top$ for every variable assignment u).

4.7.3. Models. Let us fix a logic L based on a language $\mathscr{L}$ and given by a derivation system $\mathcal{D}$.

An $\mathscr{L}$-structure M is called a *model* of L if the meta-formula $\mathcal{D}$ is valid in M (in other words, if all the laws and inference rules of $\mathcal{D}$ are valid in M_+).[70] If an $\mathscr{L}$-structure M is a model of L, then clearly all derivable principles of L are valid in M, and more generally, it is easy to see that

$$\mathcal{F}_1, \ldots, \mathcal{F}_k \vdash_L \mathcal{G} \quad \text{implies} \quad \mathcal{F}_1, \ldots, \mathcal{F}_k \vDash_M \mathcal{G}$$

for arbitrary meta-formulas $\mathcal{F}_i$ and $\mathcal{G}$. In particular,

if a rule $\Phi_1, \ldots, \Phi_k \,/\, \Psi$ is derivable in L, then it is valid in every model of L.

For many logics one also has the converse implication (see e.g. [172; 1.4.11] for the zero-order case). In fact, a logic L is said to be *complete* with respect to a family of its models M_t (possibly consisting of just one model), if every purely logical principle that is valid in each M_t is derivable in L, and *strongly complete* if every purely logical rule that is valid in each M_t is derivable in L. Thus the family M_t is not strongly complete if and only if there exists a proper extension L^+ of L such that each M_t is a model of L^+.

A *model* of a theory $\mathcal{T}$ over L is an $\mathscr{L}(\mathcal{T})$-structure M that is a model of L such that the meta-formula $\mathcal{T}$ is valid in M (in other words, such that all axioms of $\mathcal{T}$ are valid in M). Clearly, every theorem of $\mathcal{T}$ is valid in all models of $\mathcal{T}$. The theory $\mathcal{T}$ is called *complete* with respect to a family M_t of its models if every formula that is valid in each M_t is a theorem of $\mathcal{T}$. Thus $\mathcal{T}$ is complete with respect to the family M_t if and only if every principle (possibly not purely logical) that is valid in each M_t is derivable from $\mathcal{T}$ in L. We may also call $\mathcal{T}$ *strongly complete* with respect to the family M_t if every rule (possibly not purely logical) that is valid in each M_t is derivable from $\mathcal{T}$ in L.

Given a first-order language $\mathscr{L}$, let $\mathscr{L}_+$ denote the language obtained by extending $\mathscr{L}$ by countably many predicate constants of each arity. Thus if $\mathscr{L}$ is the language of a logic

---

[70]Let us note that models as defined here include neither a predicate valuation nor a variable assignment.

L, we have an arity-preserving bijection φ_L between the predicate variables of $\mathscr{L}$ and the predicate constants of $\mathscr{L}_+$. It extends to a bijection φ_L between the formulas of $\mathscr{L}$ and those formulas of $\mathscr{L}_+$ that involve no predicate variables. If M is an $\mathscr{L}^+$-structure, let M^- denote the $\mathscr{L}$-structure obtained from M by discarding the interpretation of the predicate constants that are not in $\mathscr{L}$.

Proposition 4.63. *If every consistent theory $\mathcal{T}$ over a logic L such that $\mathscr{L}(\mathcal{T}) = \mathscr{L}(L)_+$ and $\mathcal{T}$ involves no predicate variables is complete with respect to some family $M_{\mathcal{T},t}$ of its models, then L is strongly complete with respect to the two-parameter family $M^-_{\mathcal{T},t}$ of its models.*

The proof also works if L is replaced by an empty theory over L.

Proof. If a rule $\Phi_1, \ldots, \Phi_k \,/\, \Psi$ of the language $\mathscr{L}(L)$ is not derivable in L, let $\mathcal{T} = $ $(\mathbf{1})\,\varphi_L(\Phi_1)\, \&\, \cdots\, \&\, (\mathbf{1})\,\varphi_L(\Phi_k)$. Then $\varphi_L(\Psi)$ is not a theorem of $\mathcal{T}$. In particular, $\mathcal{T}$ is consistent, and hence $\varphi_L(\Psi)$ is not valid in $M_{\mathcal{T},t}$ for some t. But each $(\mathbf{1})\,\varphi_L(\Phi_i)$ is valid in $M_{\mathcal{T},t}$. Thus $(\mathbf{1})\,\varphi_L(\Phi_1)\, \&\, \cdots\, \&\, (\mathbf{1})\,\varphi_L(\Phi_k) \Rightarrow (\mathbf{1})\,\varphi_L(\Psi)$ is not valid in $M_{\mathcal{T},t}$, and hence the rule $\Phi_1, \ldots, \Phi_k \,/\, \Psi$ is not valid in $M_{\mathcal{T},t}$ or, equivalently, in $M^-_{\mathcal{T},t}$. □

Remark 4.64. Strong completeness is closely related to Dragalin's composite theories [46]. We call a pair $\mathcal{T} = (\mathcal{T}_+, \mathcal{T}_-)$ a *composite theory* if $\mathcal{T}_+$ is a theory, i.e., a finite meta-conjunction of principles, called the *axioms* of $\mathcal{T}$, and $\mathcal{T}_-$ is a finite meta-disjunction (see *4.6.3*) of principles, which might be called the *taboos* of $\mathcal{T}$. The empty meta-disjunction will be identified with $\tilde{\divideontimes} = (\gamma)\,\gamma$. A composite theory $\mathcal{T}$ is called *inconsistent* if $\mathcal{T}_+ \vdash \mathcal{T}_-$.

A *model* of a composite theory $\mathcal{T}$ over a logic L is a model M of L such that $\mathcal{T}_+$ is valid in M and $\mathcal{T}_-$ is not valid in M. Since meta-disjunction, like all other meta-connectives, is treated classically in the two-valued meta-interpretation, to say that $\mathcal{T}_-$ is not valid in M is equivalent to saying that none of the taboos of $\mathcal{T}$ is valid in M — as long as there is at least one taboo. Else we have $\mathcal{T}_- = (\gamma)\,\gamma$, and to say $\mathcal{T}_-$ is not valid in M is equivalent to saying that not all formulas are valid in M. This is equivalent to the usual requirement that $\mathscr{D}$ be nonempty (which is thus "explained" by composite theories).

If every consistent composite theory $\mathcal{T}$ over L such that $\mathscr{L}(\mathcal{T}) = \mathscr{L}(L)_+$ and $\mathcal{T}$ involves no predicate variables has a model $M_{\mathcal{T}}$, then L is strongly complete with respect to the family $M^-_{\mathcal{T}}$. Indeed, if a rule $\Phi_1, \ldots, \Phi_n \,/\, \Psi$ is not derivable, then the composite theory $\mathcal{T}$ with axioms $(\mathbf{1})\,\varphi_L(\Phi_1), \ldots, (\mathbf{1})\,\varphi_L(\Phi_n)$ and the only taboo $(\mathbf{1})\,\varphi_L(\Psi)$ is consistent. Then $(\mathbf{1})\,\varphi_L(\Phi_1), \ldots, (\mathbf{1})\,\varphi_L(\Phi_n)$ are valid in $M_{\mathcal{T}}$, whereas $(\mathbf{1})\,\varphi_L(\Psi)$ is not. Hence $(\mathbf{1})\,\varphi_L(\Phi_1)\, \&\, \cdots\, \&\, (\mathbf{1})\,\varphi_L(\Phi_n) \Rightarrow (\mathbf{1})\,\varphi_L(\Psi)$ is not valid in $M_{\mathcal{T}}$. Therefore the rule $\Phi_1, \ldots, \Phi_n \,/\, \Psi$ is not valid in $M_{\mathcal{T}}$ or, equivalently, in $M^-_{\mathcal{T}}$.

4.7.4. Valuation fields. Given an $\mathscr{L}$-structure $\imath$ given by $\mathscr{D}$, $\mathscr{O}$ and interpretations of function symbols, predicate constants, connectives and quantifiers, we call a family $\mathcal{H}$ of subsets $\mathcal{H}_n \subset \mathrm{Hom}(\mathscr{D}^n, \mathscr{O})$, $n = 0, 1, 2, \ldots$, a *valuation field* with respect to $\imath$ if it is closed with respect to the interpretations of function symbols, predicate constants, connectives and quantifiers provided by $\imath$. In other words, $\mathcal{H} = (\mathcal{H}_n)$ is a valuation field if and only if each valuation υ such that $\upsilon(\varphi) \in \mathcal{H}_n$ for every n-ary predicate

variable φ, $n = 0, 1, 2, \ldots$, also satisfies $|F|_i^v \in \mathcal{H}_n$ for each n-formula F of the language, $n = 0, 1, 2, \ldots$.

A valuation field $\mathcal{H}$ *contains* a valuation v if $v(\varphi) \in \mathcal{H}_n$ for every n-ary predicate variable φ, $n = 0, 1, 2, \ldots$. The valuation field *generated by* a valuation v, denoted $\langle v \rangle$, is the smallest valuation field that contains v. Clearly, $\langle v \rangle = (\mathcal{H}_n)$ where each $\mathcal{H}_n$ consists of $|F|_i^v$ where F runs over all n-formulas of the language.

Given a valuation field $\mathcal{H} = (\mathcal{H}_n)$ with respect to an $\mathscr{L}$-structure i, a *meta-L-structure* j *restricted over* $\mathcal{H}$ is defined like a usual meta-$\mathscr{L}$-structure, except that the second-order meta-quantifiers are now interpreted by functions $|\mathfrak{q}_n|_j \colon \mathrm{Hom}(\mathcal{H}_n, \mathcal{Q}) \to \mathcal{Q}$. In particular, the *two-valued meta-L-structure restricted over* $\mathcal{H}$ is the modification $i_+(\mathcal{H})$ of the usual two-valued meta-$\mathscr{L}$-structure i_+ obtained by re-interpreting the second-order meta-quantifiers by functions $|\mathfrak{q}_n|_{i_+(\mathcal{H})} \colon \mathrm{Hom}(\mathcal{H}_n, \mathcal{Q}) \to \mathcal{Q}$ defined in the same way as before, i.e. by $|\mathfrak{q}_n|_{i_+(\mathcal{H})}(f) = \top$ if and only $f(\varphi) = \top$ for all $\varphi \in \mathcal{H}_n$.

It is not hard to see that for any valuation field $\mathcal{H}$, $i_+(\mathcal{H})$ is a standard model of the meta-logic in the sense of §*4.7.1*, except that all valuations mentioned there are now required to be contained in $\mathcal{H}$. (It is here that we use that $\mathcal{H}$ is a valuation field and not just an arbitrary family of subsets of the function sets $\mathrm{Hom}(\mathscr{D}^n, \mathscr{O})$.)

Given a valuation field $\mathcal{H}$, with respect to an $\mathscr{L}$-structure M, the latter is called a *model* of $\mathscr{L}$ *with respect to* $\mathcal{H}$ if the meta-formula $\mathcal{D}$ is valid in $M_+(\mathcal{H})$.

Remark 4.65. Valuation fields are parallel to (and, in a sense, generalize) algebras of admissible sets, which distinguish the so-called "general frames" from usual frames.

4.8. Classical logic (semantics)

4.8.1. Two-valued models. In a *two-valued* model of classical logic, the set $\mathscr{O}$ is taken to be $\{\top, \bot\}$, with $! = \mathrm{id}$, and the connectives and the quantifiers are interpreted according to the usual truth tables (see §**3.1**). (Thus two-valued models of classical logic differ only in the choice of the domain $\mathscr{D}$.) It is not hard to check that, indeed, the laws and the inference rules of classical logic (see §**4.6**) hold under this interpretation.

A model of a theory over classical logic is *two-valued* if it extends a two-valued model of classical logic (by additionally interpreting the predicate constants and the function symbols in the language of the theory). By a strong form of Gödel's completeness theorem [168; VIII.5.2], every consistent theory $\mathcal{T}$ over classical logic is complete with respect to the family of its two-valued models with domain $\mathbb{N}$.[71] In fact, the proof of this assertion in [168] actually shows that $\mathcal{T}$ is complete with respect to a family $M_{\mathcal{T},t}$ of its two-valued models with domain $\mathbb{N}$ that differ only in the interpretation of the predicate constants, as they all have the same interpretation of the function symbols. (Namely, if $\mathfrak{f}$ is an n-ary function symbol, $|\mathfrak{f}| \colon \mathbb{N}^n \to \mathbb{N}$ is defined by $|\mathfrak{f}|(\overline{t_1}, \ldots, \overline{t_n}) = \overline{\mathfrak{f}(t_1, \ldots, t_n)}$,

---

[71]Let us recall that our theories correspond to the traditional notion of a theory given by finitely many axioms and finitely many axiom schemata (which due to our finiteness restrictions on the language consist of countably many axioms). This is included in the countably axiomatized traditional theory of [168; VIII.5.2] (in fact, the finitely axiomatized case would already suffice for our present purposes).

where $\bar{\cdot} \colon \mathscr{T}(\mathscr{L}) \to \mathbb{N}$ is some bijection between all terms of the language and natural numbers.) Thus from 4.63 we get

Theorem 4.66. *Classical logic is strongly complete with respect to its unique two-valued model with domain $\mathbb{N}$.*

Every empty theory over classical logic is strongly complete with respect to a family of its two-valued models with domain $\mathbb{N}$. This family consists of just one model if there are no predicate constants in the language.

The condition that there be no predicate constants in the language cannot be dropped by considering a propositional (=nullary predicate) constant $\mathfrak{c}$. In any two-valued model M, either $|\mathfrak{c}|_M = \top$ or $|\mathfrak{c}|_M = \bot$; hence one of the non-derivable principles $\cdot\,\mathfrak{c}$, $\cdot\,\neg\mathfrak{c}$ is valid in M.

We will next use two-valued models to differentiate derivable, admissible and stably admissible rules, as well as principles.

Example 4.67. Although the generalization rule $\mathsf{p}(\mathsf{x}) \,/\, \forall\mathsf{x}\,\mathsf{p}(\mathsf{x})$ is derivable in classical logic (see (XII$_\checkmark$)), the corresponding formula $\mathsf{p}(\mathsf{t}) \to \forall\mathsf{x}\,\mathsf{p}(\mathsf{x})$ is not.

Indeed, let us consider a two-valued model M of classical logic with $\mathscr{D}_M = \{0,1\}$ and a predicate valuation ν such that $|\mathsf{p}|_M^\nu(0) = \bot$ and $|\mathsf{p}|_M^\nu(1) = \top$. Then $|\mathsf{p}(\mathsf{x}) \to \forall\mathsf{x}\,\mathsf{p}(\mathsf{x})|_M^{u\nu}$ is $\top$ if $u(\mathsf{x}) = 0$ and $\bot$ if $u(\mathsf{x}) = 1$. Hence $\mathsf{p}(\mathsf{x}) \to \forall\mathsf{x}\,\mathsf{p}(\mathsf{x})$ is not valid in (M,ν), and therefore also in M. Thus it is not derivable in classical logic.

Let us note that the meta-formula $(\mathsf{x})\,\mathsf{p}(\mathsf{x}) \Rightarrow \forall\mathsf{x}\,\mathsf{p}(\mathsf{x})$ must be valid in (M,ν), since it is an immediate consequence of the generalization rule. This is easy to check directly: it is interpreted by the implication "if $\mathsf{p}(\mathsf{x})$ is valid in (M,ν), then $\forall\mathsf{x}\,\mathsf{p}(\mathsf{x})$ is valid in (M,ν)", in which both the hypothesis and the conclusion are false.

Example 4.68. In zero-order classical logic with a propositional constant $\mathfrak{c}$ added to the language, the rule $\mathfrak{c} \,/\, \bot$ is admissible but not derivable (and not stably admissible).

Indeed, a two-valued model M such that $|\mathfrak{c}|_M = \top$ shows that the rule $\mathfrak{c} \,/\, \bot$ is not derivable. Moreover it shows that the principle $\cdot\,\mathfrak{c}$ does not imply the principle $\cdot\,\bot$, and hence that the rule $\mathfrak{c} \,/\, \bot$ is not stably admissible. On the other hand, a two-valued model M such that $|\mathfrak{c}|_M = \bot$ shows that the formula $\mathfrak{c}$ is not derivable. Hence the rule $\mathfrak{c} \,/\, \bot$, which has no special cases apart from itself, is admissible.

Example 4.69. In classical logic, the rule

$$\frac{\exists\mathsf{x}\,\mathsf{p}(\mathsf{x}),\ \exists\mathsf{x}\,\neg\mathsf{p}(\mathsf{x})}{\bot}$$

is admissible but not derivable.

To see the latter, we can use the same M and ν as in the previous example (4.67). Then $\exists\mathsf{x}\,\mathsf{p}(\mathsf{x})$ and $\exists\mathsf{x}\,\neg\mathsf{p}(\mathsf{x})$ are both valid in (M,ν), yet $\bot$ is not. Hence the rule in question is not valid in M, and therefore it is not derivable in classical logic.

On the other hand, suppose that for some 1-formula P we have both $\vdash \exists\mathsf{x}\,P(\mathsf{x})$ and $\vdash \exists\mathsf{x}\,\neg P(\mathsf{x})$. Let N be a two-valued model of classical logic with domain consisting of a

single element, $\mathscr{D}_N = \{*\}$. Then $\exists x\, P(x)$ and $\exists x\, \neg P(x)$ must both be valid in N, that is, $|\exists x\, P(x)|_N^{uv} = |\exists x\, \neg P(x)|_N^{uv} = \top$ for every individual assignment u and every predicate valuation v. Since $*$ is the only element in $\mathscr{D}_N$, this means that $|P|_N^{uv}(*) = |\neg P|_N^{uv}(*) = \top$, which is a contradiction. Hence for any special case of the rule in question, the derivability of the premises is impossible, and in particular implies the derivability of the conclusion. Thus the rule is admissible.

Remark 4.70. In fact, these examples are characteristic of non-derivable admissible rules in classical logic. They are all such that any specialization of the conjunction of their premises is non-derivable, and even not satisfiable in any two-valued model with one-element domain [47].

Example 4.71. The rule of the previous example is, in fact, stably admissible, not only for classical but, more generally, for intuitionistic logic (and also for empty theories over them, with no predicate constants of arity > 0 in the language). Let us note that this rule is not derivable in intuitionistic logic, since it is not derivable in classical logic.

To prove the stable admissibility, we need to show that for any 1-formula Σ, the principles $\cdot\, \exists x\, \Sigma(x)$ and $\cdot\, \exists x\, \neg\Sigma(x)$ imply $\cdot\, \times$. Indeed, let $\vec{\gamma}$ be the tuple of all predicate variables that occur in Σ, and let $\vec{r} = (r_1, \ldots, r_k)$ be the tuple of their arities. Next let $\times_{\vec{r}} = (\times_{r_1}, \ldots, \times_{r_n})$, where $\times_r$ denotes the r-formula $\varkappa_1, \ldots, \varkappa_r \mapsto \times$. (Alternatively, we could have used $\checkmark$ in place of $\times$, or both of them.) Then the two principles in question meta-specialize to the formulas $\exists x\, \Sigma(x)[\vec{\gamma}/\times_{\vec{r}}]$ and $\exists x\, \neg\Sigma(x)[\vec{\gamma}/\times_{\vec{r}}]$ respectively. Since the language contains no predicate constants of arity > 0, the formula $\Sigma(x)[\vec{\gamma}/\times_{\vec{r}}]$ contains no occurrences of individual variables, including x. But then our formulas $\exists x\, \Sigma(x)[\vec{\gamma}/\times_{\vec{r}}]$ and $\exists x\, \neg\Sigma(x)[\vec{\gamma}/\times_{\vec{r}}]$ are equivalent (over intuitionistic logic) to $\Sigma(x)[\vec{\gamma}/\times_{\vec{r}}]$ and $\neg\Sigma(x)[\vec{\gamma}/\times_{\vec{r}}]$ respectively. Thus they indeed have $\times$ as a joint consequence.

Example 4.72. In classical logic with language extended by one binary predicate constant (denoted $=$), the rule $\exists x\, p(x)$, $\exists x\, \neg p(x)\, /\, \bot$ is not stably admissible.

Indeed, let us show that the principles $\cdot\, \exists x\, x = y$ and $\cdot\, \exists x\, \neg x = y$ do not imply $\cdot\, \bot$; or in other words that if they are added as laws to classical logic with equality, the resulting logic is still consistent. Indeed, $(y)\, \exists x\, x = y$ or, equivalently, $\forall y\, \exists x\, x = y$ is derivable from $\forall y\, y = y$, and therefore is valid in any model of classical logic with equality; whereas $(y)\, \exists x\, \neg x = y$ or, equivalently, $\forall y\, \exists x\, \neg x = y$ is valid in any model of classical logic with equality with non-singleton domain. (See e.g. [141] concerning classical logic with equality.) Thus classical logic with $=$ added to the language (and no additional laws or inference rules) has two-valued models where both $\exists x\, x = y$ and $\exists x\, \neg x = y$ are valid; whereas $\bot$ is not valid in any model.

Proposition 4.73. *In zero-order classical logic (with no propositional constants added to the language) admissible rules are derivable.*

Proof. If a rule $P_1, \ldots, P_k\, /\, Q$ is not derivable, then by the strong completeness theorem 4.66, this rule is not satisfied in the two-valued model M of zero-order classical

logic. Then there is a special case $P_1', \ldots, P_k' \,/\, Q'$ of the rule in question and a valuation v such that each $|P_i'|_M^v = \mathbb{T}$ and $|Q'|_M^v = \bot\!\!\!\bot$ (see e.g. Proposition 4.62(b)). Let $\vec{p} = (p_1, \ldots, p_n)$ be the tuple of all propositional variables that occur freely in any of the formulas $P_1', \ldots, P_k', Q'$, and let $\vec{\mathfrak{c}} = (\mathfrak{c}_1, \ldots, \mathfrak{c}_n)$, where $\mathfrak{c}_i = \top$ if $v(p_i) = \mathbb{T}$ and $\mathfrak{c}_i = \bot$ if $v(p_i) = \bot\!\!\!\bot$. Then $(P_1', \ldots, P_k' \,/\, Q')[\vec{p}/\vec{\mathfrak{c}}]$ is a new rule, which can be written $P_1'', \ldots, P_k'' \,/\, Q''$, and which is another special case of the original rule $P_1, \ldots, P_k \,/\, Q$. By design, each $|P_i''|_M^v = \mathbb{T}$ and $|Q''|_M^v = \bot\!\!\!\bot$. But since $P_1'', \ldots, P_k''$ and Q'' contain no propositional variables, their interpretations do not depend on v, and we may also write $|P_i''|_M = \mathbb{T}$ and $|Q''|_M = \bot\!\!\!\bot$, and consider P_i'' and Q'' as principles. But zero-order classical logic is complete with respect to M (see 4.66 again), so each P_i'' is derivable, and Q'' is not derivable (in fact, $\neg Q''$ is derivable due to $|\neg Q''|_M = \mathbb{T}$.) Thus $P_1, \ldots, P_k \,/\, Q$ is not admissible. $\qquad\square$

Two-valued models can also to be used to compare implication between principles or rules with that between the judgements of their derivability.

Proposition 4.74. *In zero-order classical logic (with no propositional constants added to the language) rules $\mathcal{R}_1, \ldots, \mathcal{R}_k$ imply a rule $\mathcal{S}$ if and only if the judgements $\vdash \mathcal{R}_1, \ldots, \vdash \mathcal{R}_k$ imply $\vdash \mathcal{S}$.*

If propositional constants are added to the language, the assertion is no longer true (see Example 4.82 below).

Proof. The "only if" assertion holds in any logic. Conversely, let us assume the implication between the judgements of derivability, and let us consider two cases. If $\mathcal{R}_1, \ldots, \mathcal{R}_k$ all happen to be derivable, then so must be $\mathcal{S}$, and so it is trivially a consequence of $\mathcal{R}_1, \ldots, \mathcal{R}_k$. If some $\mathcal{R}_i$ is not derivable, then by the proof of 4.73 it has a special case $P_1, \ldots, P_n \,/\, Q$ (involving only $\top$, $\bot$ and other connectives, but no propositional variables) such that each P_j is derivable, and $Q \to \bot$ is also derivable. Thus by adjoining $\mathcal{R}_i$ to the logic, we will be able to derive $\bot$. Then everything becomes derivable, including $\mathcal{S}$. Thus $\mathcal{S}$ is a consequence of $\mathcal{R}_i$ alone; in particular, it is a joint consequence of $\mathcal{R}_1, \ldots, \mathcal{R}_k$. $\qquad\square$

Proposition 4.74 does not generalize to first-order classical logic:

Example 4.75. In classical logic, the principle $\cdot\, \exists x\, p(x) \to \forall x\, p(x)$ does not imply $\cdot\, \bot$, even though the judgement $\vdash \exists x\, p(x) \to \forall x\, p(x)$ implies $\vdash \bot$.

Indeed, $\exists x\, p(x) \to \forall x\, p(x)$ is valid in all models with a singleton domain, whereas $\bot$ is not valid in any model; thus there is no implication between the principles. On the other hand, $\exists x\, p(x) \to \forall x\, p(x)$ is not valid in some models; thus the judgement $\vdash \exists x\, p(x) \to \forall x\, p(x)$ is false, and so implies any other judgement.

4.8.2. Leibniz–Euler models. It was known already to Leibniz and became well-known from the work of Euler that classical propositional logic can be interpreted by set-theoretic operations (of course, at a level of knowledge that is informal by modern

standards; see [16]). The predicate case (but of course not the modern formulations) essentially appeared in the work of Pierce and Schröder in the 19th century.

To describe a *Leibniz–Euler* model $V = V(X, \mathscr{D})$ of classical logic,[72] one starts with arbitrary sets X and $\mathscr{D}$. The set $\mathcal{O}$ will be the set of arbitrary subsets of X, and the function $!: \mathcal{O} \to \{\top, \bot\}$ sends the entire X to $\top$ and every proper subset $S \subsetneq X$ to $\bot$.

Thus, to specify a valuation, each predicate variable of arity n must be interpreted by a $\mathscr{D}^n$-indexed family of subsets of X. Upon a choice of a valuation $\boldsymbol{v}$, a closed formula will be interpreted by a subset of X, and more generally a closed n-formula F will be interpreted by a $\mathscr{D}^n$-indexed family $|F|_V^v: \mathscr{D}^n \to \mathcal{O}$ of subsets of X. Upon a further choice of a variable assignment $\boldsymbol{u}$, an arbitrary formula will be interpreted by a subset of X, and more generally an arbitrary n-formula F will be interpreted by a $\mathscr{D}^n$-indexed family $|F|_V^{uv}: \mathscr{D}^n \to \mathcal{O}$ of subsets of X. Validity of an n-formula F in the model, $\vDash F$, means that $|F|_V^{uv}(t_1, \ldots, t_n) = X$ for each tuple $(t_1, \ldots, t_n) \in \mathscr{D}^n$ and for an arbitrary valuation $\boldsymbol{v}$ and assignment $\boldsymbol{u}$.

Finally, disjunction of formulas is interpreted by union of sets, negation by complement (within X), and universal quantification by $\mathscr{D}$-indexed union. In symbols, if F and G are formulas and H is a 1-formula,

- $|F \vee G|_V^{uv} = |F|_V^{uv} \cup |G|_V^{uv}$;
- $|\neg F|_V^{uv} = X \setminus |F|_V^{uv}$;
- $|\exists \mathsf{x}\, H(\mathsf{x})|_V^{uv} = \bigcup_{d \in D} |H|_V^{uv}(d)$.

Consequently, conjunction is interpreted by intersection, and existential quantification by $\mathscr{D}$-indexed intersection. The interpretation of implication is also determined by those disjunction and negation. The interpretation of $\bot$, which is classically equivalent to $\mathsf{p} \wedge \neg\mathsf{p}$, is determined by those of conjunction and negation; and the interpretation of $\top$ is determined by those of $\bot$ and negation. Thus $\bot$ is interpreted by the empty union (that is, $\varnothing$) and $\top$ by the empty intersection (that is, X).

We call this interpretation a *Leibniz–Euler model* of classical logic; indeed, it is not hard to check that the laws and the inference rules of classical logic (see §**4.6**) hold under this interpretation (cf. [168]). For example, let us consider the generalization rule, $\mathsf{q}(\mathsf{x}) \,/\, \forall \mathsf{x}\, \mathsf{q}(\mathsf{x})$. To check that it is valid in a Leibniz–Euler model $V = V(X, \mathscr{D})$, we pick some predicate valuation $\boldsymbol{v}$ (cf. 4.62). Then $|\mathsf{q}|_V^v$ is some function from $\mathscr{D}$ to the set of all subsets of X; $|\forall \mathsf{x}\, \mathsf{q}(\mathsf{x})|_V^v$ is the intersection of all its values; and $|\mathsf{q}(\mathsf{x})|_V^{uv}$ is its value at $\boldsymbol{u}(x)$, where the individual assignment $\boldsymbol{u}$ associates to every individual variable, including x, an arbitrary element of $\mathscr{D}$. Thus if $|\mathsf{q}(\mathsf{x})|_V^{uv} = X$ for every individual assignment $\boldsymbol{u}$, then $|\forall \mathsf{x}\, \mathsf{q}(\mathsf{x})|_V^v = X$, as desired.

In the case where X is a singleton, $\{*\}$, Leibniz–Euler models reduce to the two-valued models discussed previously. As long as no predicate constants are added to the language, two-valued models suffice to discern classical truth (see Theorem 4.66).

---

[72]In the literature, these seem to be best known as a special case of Boolean-valued models. However, for the purposes of the present treatise there is no benefit in dealing with general Boolean algebras (and consequently, also general Heyting algebras) so there is no point in using that terminology here, as it would only obscure matters.

Moreover, with an arbitrary language, a formula F that is not valid in some Leibniz–Euler model $V = V(X, \mathscr{D})$ is also not valid in some two-valued model M with the same domain $\mathscr{D}$. Namely, we can get M from V by keeping the same interpretations of the function symbols, picking an $x \in X$ such that $x \notin |F|_V^{uv}$ for some individual assignment u and predicate valuation v, and setting the interpretation $|\mathfrak{p}|_M$ of a predicate constant $\mathfrak{p}$ to be the function sending a tuple $\vec{d} \in \mathscr{D}^n$ to $\{*\}$ if $x \in |\mathfrak{p}|_V(\vec{d})$ and to $\varnothing$ otherwise.

On the other hand, classical logic (or the empty theory over it, with any language) is complete with respect to some Leibniz–Euler model with domain $\mathbb{N}$ and with a fixed predicate valuation [168; VIII.3.3, II.8.1], but of course not with respect to any particular two-valued model with a fixed valuation — not even in the propositional case. Also, a λ-closed formula F is not valid in a two-valued model M if and only if $\neg F$ is valid in M; of course, this strange property does not hold in Euler models.

Nevertheless, an Euler model such that $X = \mathscr{D}^k$ for some k can be derived from a two-valued model, by introducing k implicit parameters and changing the arities of predicate constants. Thus we fix a domain $\mathscr{D}$, and to interpret an n-ary predicate constant $\mathfrak{p}$ in an Euler model V with $X = \mathscr{D}^k$ we look at the interpretation of an $(n+k)$-ary predicate constant $\mathfrak{p}'$ in a two-valued model M with the same domain $\mathscr{D}$, and set

$$|\mathfrak{p}|_V(x_1, \ldots, x_n) = \{(x_{n+1}, \ldots, x_{n+k}) \in \mathscr{D}^k \mid |\mathfrak{p}'|_M(x_1, \ldots, x_{n+k}) = \top\}.$$

Example 4.76. In classical logic, $F \parallel G \vdash F \vee G$ for any formulas F, G by 4.55; but the converse, $F \vee G \vdash F \parallel G$, does not hold already in the zero-order case.

Indeed, let V be a Leibniz–Euler model of classical logic with $X = \{0, 1\}$, and let v be a valuation such that $v(\mathsf{p}) = \{0\}$ and $v(\mathsf{q}) = \{1\}$. Then $|\mathsf{p} \vee \mathsf{q}|_V^v = \{0, 1\}$, so $|\mathsf{p} \vee \mathsf{q}|_{V_+}^v = \top$; whereas $|\mathsf{p}|_V^v = \{0\}$ and $|\mathsf{q}|_V^v = \{1\}$, so $|\mathsf{p}|_{V_+}^v = |\mathsf{q}|_{V_+}^v = \bot$, and consequently $|\mathsf{p} \parallel \mathsf{q}|_{V_+}^v = \bot$. Thus $\mathsf{p} \vee \mathsf{q} \Rightarrow \mathsf{p} \parallel \mathsf{q}$ is not valid in V, and consequently is not derivable in classical logic.

Let us note that $F \vee G \Rightarrow F \parallel G$ is clearly valid for any formulas F, G in every two-valued model of classical logic.

Example 4.77. The previous example (4.76) can be improved: (2)$F \vee G \vdash$ (2)$(F \parallel G)$ does not hold already in zero-order classical logic.

Indeed, let V be a Leibniz–Euler model of classical logic with $X = \{0, 1\}$, and let v be a valuation such that $v(\mathsf{p}) = \{0\}$. Then $|\neg\mathsf{p}|_V^v = \{1\}$, so $|\mathsf{p}|_{V_+}^v = |\neg\mathsf{p}|_{V_+}^v = \bot$, and consequently $|\mathsf{p} \parallel \neg\mathsf{p}|_{V_+}^v = \bot$. Thus (2)$(\mathsf{p} \parallel \neg\mathsf{p})$ is not valid in V, and consequently is not derivable in classical logic. On the other hand, (2)$\mathsf{p} \vee \neg\mathsf{p}$ is the principle of excluded middle, which is certainly derivable in classical logic.

Example 4.78. In classical logic, $F \to G \vdash F \Rightarrow G$ for any formulas F, G by the *modus ponens* rule; but the converse, $F \Rightarrow G \vdash F \to G$, does not hold already in the zero-order case.

Indeed, let V be a Leibniz–Euler model of classical logic with $X = \{0, 1\}$, and let v be a valuation such that $v(\mathsf{p}) = \{0\}$. Then $|\mathsf{p} \to \bot|_V^v = \{1\}$, so $|\mathsf{p} \to \bot|_{V_+}^v = \bot$; whereas $|\mathsf{p}|_V^v = \{0\}$ and $|\bot|_V^v = \{\varnothing\}$, so $|\mathsf{p}|_{V_+}^v = |\bot|_{V_+}^v = \bot$, and consequently $|\mathsf{p} \Rightarrow \bot|_{V_+}^v = \top$.

Thus $(\mathsf{p} \Rightarrow \bot) \Rightarrow \mathsf{p} \to \bot$ is not valid in V, and consequently is not derivable in classical logic.

Let us note that $(F \Rightarrow G) \Rightarrow F \to G$ is clearly valid for any formulas F, G in every two-valued model of classical logic.

Example 4.79. A rule $F \,/\, G$ need not imply the principle $\cdot\, F \to G$ in zero-order classical logic with language extended by a propositional constant $\mathfrak{c}$.

Indeed, similarly to the previous example (4.78), $\vDash_V \mathfrak{c} \Rightarrow \bot$ does not imply $\vDash_V \mathfrak{c} \to \bot$ for a Leibniz–Euler model V with $X_V = \{0, 1\}$ and $|\mathfrak{c}|_V = \{0\}$.

Remark 4.80. A rule $F \,/\, G$ implies the principle $\cdot\, F \to G$ in zero-order classical logic (with no propositional constants added to the language). Indeed, $\vdash F \,/\, G$ implies $\vdash \cdot\, F \to G$ by the deduction theorem, and therefore $(F \,/\, G)$ implies $\cdot\, F \to G$ by 4.74.

Example 4.81. $\vdash (2)\,(F \Rightarrow G) \Rightarrow (2)\, F \to G$ need not imply $\vdash (2)\big((F \Rightarrow G) \Rightarrow F \to G\big)$ already in zero-order classical logic (with no propositional constants added to the language).

Indeed, let $F = \mathsf{p}$ and $G = \bot$. Then the former meta-formula is derivable by Remark 4.79; whereas the latter one is not derivable by Example 4.78.

Example 4.82. If $\mathcal{R}$ and $\mathcal{S}$ are rules, and the judgement $\vdash \mathcal{R}$ implies $\vdash \mathcal{S}$, then the rule $\mathcal{R}$ need not imply $\mathcal{S}$ in zero-order classical logic with language extended by a propositional constant $\mathfrak{c}$.

Indeed, let $\mathcal{R} = \mathfrak{c} \,/\, \bot$ and $\mathcal{S} = \cdot\, \neg\mathfrak{c}$ (or, equivalently, $\mathcal{S} = \top \,/\, \neg\mathfrak{c}$). Then $\mathcal{R}$ is not derivable (see Example 4.68), and consequently $\vdash \mathcal{R}$ implies everything, including $\vdash \mathcal{S}$. On the other hand, $\mathcal{R}$ does not imply $\mathcal{S}$ (see Example 4.79).

Example 4.83. A rule $F \,/\, G$ need not imply the principle $\cdot\, F \to G$ in classical logic, even if the formulas F and G are closed.

Indeed, let $F = \exists\mathsf{x}\,\mathsf{p}(\mathsf{x}) \wedge \exists\mathsf{x}\,\neg\mathsf{p}(\mathsf{x})$ and $G = \bot$. We will construct an Euler model by combining the two-valued models of Examples 4.67 and 4.69.

Thus let M and N be the two-valued models of classical logic with $\mathscr{D}_M = \{0, 1\}$ and $\mathscr{D}_N = \{*\}$. Let V be an Euler model with $X = \{\mu, \nu\}$ and $\mathscr{D} = \{0, 1\}$. Given a predicate valuation υ_M for M and a predicate valuation υ_N for N, let $\upsilon = \upsilon(\upsilon_M, \upsilon_N)$ be the predicate valuation for V defined on each r-ary predicate variable q by $\mu \in \upsilon(q)(x_1, \ldots, x_r)$ if and only if $\upsilon_M(q)(x_1, \ldots, x_r) = \top$ and by $\nu \in \upsilon(q)(x_1, \ldots, x_r)$ if and only if $\upsilon_N(q)(*, \ldots, *) = \top$. It follows that for every closed formula P, $\mu \in |P|_V^\upsilon$ if and only if P is valid in (M, υ_M), and $\nu \in |P|_V^\upsilon$ if and only if P is valid in (N, υ_N).

It is easy to see that valuations of the form $\upsilon(\upsilon_M, \upsilon_N)$ form a valuation field K. Let us note that these are precisely those valuations υ for V that satisfy for each $r = 1, 2, \ldots$ and each r-ary predicate variable q the following alternative: either $\nu \in \upsilon(q)(x_1, \ldots, x_r)$ for all $x_1, \ldots, x_r \in \{0, 1\}$ or $\nu \notin \upsilon(q)(x_1, \ldots, x_r)$ for all $x_1, \ldots, x_r \in \{0, 1\}$.

Let υ_M be a predicate valuation for M such that $\upsilon_M(\mathsf{p})(0) = \bot$ and $\upsilon_M(\mathsf{p})(1) = \top$, and let υ_N be some predicate valuation for N. Let $\upsilon = \upsilon(\upsilon_M, \upsilon_N)$.

Clearly, F is valid in (M, v_M), and consequently $|F|_V^v$ contains μ. Hence $|F \to \bot|_V^v$ does not contain μ, and therefore $|F \to \bot|_{V_+}^v = \bot\!\!\bot$. Thus $\cdot\, F \to \bot$ is not valid in V with respect to the valuation field K.

On the other hand, F is clearly not valid in (N, v_N) (regardless of the choice of v_N), and consequently $|F|_V^v$ does not contain ν. Hence $|F \Rightarrow \bot|_{V_+}^v = \top\!\!\top$. This works regardless of the choice of v_M, so we get that $F \,/\, \bot$ is valid in V with respect to the valuation field K.

4.8.3. Platonist interpretation I (naive). We now turn to contentual interpretations of the meta-logical notions introduced above, in the case of classical object logic.

As discussed in §**3.10**, a natural understanding of classical logic, closely related to the Platonist philosophy, is that closed formulas signify propositions — which are either true or false. This view is compatible with the usual Tarski-style model theory, as presented in §*4.3.4* and §*4.7.1*, and extends to give an interpretation of the meta-logic of classical logic.

Namely, a domain $\mathscr{D}$ is fixed, and $\mathscr{O}$ is taken to be a class of propositions, consisting of a prescribed class of contentful (e.g. mathematical) primitive propositions and of composite propositions, which are obtained inductively from primitive ones by using contentual classical connectives and quantifiers over arbitrary $\mathscr{D}$-indexed families of propositions. For instance, primitive propositions could be of the form "$n = m$", where n and m are numerals, e.g. "$0 = 1$" and "$1 = 1$". In this case composite propositions would include "$0 = 1$ or not $0 = 1$" and, if $\mathscr{D}$ is the set $\mathbb{N}$ of numerals, also "there exists an n such that not $n = 0$". Formal classical connectives and quantifiers ($\vee : \mathbb{1}^2 \to \mathbb{1}$, $\exists : (\mathbb{0} \to \mathbb{1}) \to \mathbb{1}$, etc.) are interpreted straightforwardly by the contentual ones ("or": $\mathscr{O}^2 \to \mathscr{O}$, "exists": $\mathrm{Hom}(\mathscr{D}, \mathscr{O}) \to \mathscr{O}$, etc.).

The truth function $!\colon \mathscr{O} \to \mathscr{Q} = \{\top\!\!\top, \bot\!\!\bot\}$ is defined explicitly on primitive propositions, and is extended inductively to composite ones by the usual truth tables (see §**3.1**). This determines a contentual interpretation of classical logic and its meta-logic, which we will call the *Platonist interpretation.*

Since we are assuming that each proposition is either true or false, the Platonist interpretation is incompatible with the specific interpretation of §*3.10.3*, although it is compatible with the preceding sketch there (if "provability" in that sketch is taken to be in a complete theory, such as Tarski's elementary geometry).

Example 4.84. Let us record the Platonist interpretations of the judgements of admissibility, stable admissibility and derivability of a purely logical rule $F \,/\, G$. Let $\vec{\varphi} = (\varphi_1, \ldots, \varphi_m)$ be the tuple of all predicate variables that occur freely in F or in G, and let $\vec{r} = (r_1, \ldots, r_m)$ be the tuple of their arities. Below $\vec{\Phi}$ is understood to run over the set $\mathscr{F}_{\vec{r}} := \mathscr{F}_{r_1} \times \ldots \times \mathscr{F}_{r_m}$ of all tuples $(\Phi_1, \ldots, \Phi_m)$ such that each Φ_i is an r_i-formula.

- Derivability: $\forall \mathscr{D}, \mathscr{O}\ \forall v\ (\forall u\ |F|_u^v \to \forall u\ |G|_u^v)$;
- Stable admissibility: $\forall \vec{\Phi}\, \forall \mathscr{D}, \mathscr{O}\ (\forall v, u\ |F[\vec{\varphi}/\vec{\Phi}]|_u^v \longrightarrow \forall v, u\ |F[\vec{\varphi}/\vec{\Phi}]|_u^v)$;
- Admissibility: $\forall \vec{\Phi}\ (\forall \mathscr{D}, \mathscr{O}\ \forall v, u\ |F[\vec{\varphi}/\vec{\Phi}]|_u^v \longrightarrow \forall \mathscr{D}, \mathscr{O}\ \forall v, u\ |F[\vec{\varphi}/\vec{\Phi}]|_u^v)$.

If we identify with each other all true propositions and, separately, all false propositions, we will get a two-valued model of classical logic (see §4.8.1), which clearly determines the validity of formulas and meta-formulas in the Platonist interpretation. Thus the Platonist interpretation can be seen as a mere rewording of two-valued models. In particular, it may be argued to be an adequate understanding of classical logic on the grounds that classical logic is strongly complete with respect to a two-valued model with countable domain (see Theorem 4.66).

However, every two-valued model M of classical logic obviously has the property that if some rule $F \,/\, G$ is valid in M, where F and G are closed formulas, then also the principle $\cdot F \to G$ is valid in M. This property is not satisfied by Leibniz–Euler models of classical logic (see Examples 4.79 and 4.83). Thus $F \,/\, G$ generally does not imply $\cdot F \to G$ in classical logic, but this is not detected by the Platonist interpretation!

4.8.4. Modified Platonist interpretation. It is not hard to amend the Platonist interpretation so as to overcome the said deficiency.

In addition to a domain of discourse $\mathscr{D}$, we fix a "hidden parameter" domain $\mathscr{E}$. Now $\mathscr{O}$ is taken to be a class of predicates on $\mathscr{E}$, consisting of a prescribed class of contentful (e.g. mathematical) primitive predicates and composite predicates, which are obtained inductively from the primitive ones by using contentual classical connectives and quantifiers over arbitrary $\mathscr{D}$-indexed families of predicates. For instance, if $\mathscr{E}$ is the rational plane $\mathbb{Q} \times \mathbb{Q}$, primitive predicates could be of the form "$nq = mr$", where n and m are numerals, e.g. "$3q = 2r$". In this case composite predicates would include "$q = 2r$ or $2q = r$" and, if $\mathscr{D}$ is the set $\mathbb{N}$ of numerals, also "there exists an n such that $nq = r$". Formal classical connectives and quantifiers are interpreted straightforwardly by the contentual ones.

The value of $!\colon \mathscr{O} \to \mathcal{Q} = \{\top, \bot\}$ on a predicate $P \in \mathscr{O}$ will be $\top$ if and only if $P(e)$ is true for all $e \in \mathscr{E}$, where "true" is defined explicitly for primitive $P \in \mathscr{O}$ and all $e \in \mathscr{E}$ and is extended inductively to composite $P \in \mathscr{O}$ and all $e \in E$ by the usual truth tables, separately for each $e \in \mathscr{E}$. This determines a contentual interpretation of classical logic and its meta-logic, which we will call the *modified Platonist interpretation*.

If we again identify with each other all true propositions and, separately, all false propositions, we will get a Leibniz–Euler model of classical logic, which determines the validity of formulas and meta-formulas in the modified Platonist interpretation. Thus the modified Platonist interpretation can be seen as a mere rewording of the Leibniz–Euler models. In particular, it does detect that $F \,/\, G$ generally does not imply $\cdot F \to G$ in classical logic.

Example 4.85. In the notation of Example 4.84, the modified Platonist interpretations of the same judgements are as follows, where e is understood to run over $\mathscr{E}$:
- Derivability: $\forall \mathscr{D}, \mathscr{E}, \mathscr{O} \; \forall v \, \big(\forall u \, \forall e \, |F|_u^v(e) \to \forall u \, \forall e \, |G|_u^v(e)\big)$;
- Stable admissibility:
  $$\forall \vec{\Phi} \, \forall \mathscr{D}, \mathscr{E}, \mathscr{O} \; \big(\forall v, u \, \forall e \, |F[\vec{\varphi}/\vec{\Phi}]|_u^v(e) \; \longrightarrow \; \forall v, u \, \forall e \, |F[\vec{\varphi}/\vec{\Phi}]|_u^v(e)\big);$$
- Admissibility:

$$\forall \vec{\Phi} \left(\forall \mathscr{D}, \mathscr{E}, \mathscr{O} \ \forall v, u \ \forall e \, |F[\vec{\varphi}/\vec{\Phi}]|_u^v(e) \ \longrightarrow \ \forall \mathscr{D}, \mathscr{E}, \mathscr{O} \ \forall v, u \ \forall e \, |F[\vec{\varphi}/\vec{\Phi}]|_u^v(e) \right).$$

5. Topology, translations and principles

Our next goal is to naturally arrive at Tarski's topological models of intuitionistic logic by means of a symbolic analysis of the BHK interpretation.

5.1. Deriving a model from BHK

Given a problem Γ, let $?\Gamma$ denote, as in §**3.8**, the proposition *A solution of Γ exists*. In the notation of §**3.9**, $?\Gamma$ can be also expressed as $|\Gamma| \neq \varnothing$, where $|\Gamma|$ denotes the set of solutions of Γ. Validity of propositions of the form $?\Gamma$ is understood to be determined, at least to some extent, by the BHK interpretation (in particular, it should be known from context for primitive Γ). Then the BHK entails, in particular, the following judgements, which involve both intuitionistic and classical connectives:

(1) $?(\Gamma \wedge \Delta)$ if and only if $?\Gamma \wedge ?\Delta$;
(2) $?(\Gamma \vee \Delta)$ if and only if $?\Gamma \vee ?\Delta$;
(3) $?(\Gamma \rightarrow \Delta)$ implies $?\Gamma \rightarrow ?\Delta$;
(4) $\neg ?\times$;
(5) $?\exists x \, \Gamma(x)$ if and only if $\exists x \, ?\Gamma(x)$;
(6) $?\forall x \, \Gamma(x)$ implies $\forall x ?\Gamma(x)$.
(7) $?\checkmark$.

Judgements (1)–(6) are perhaps best understood as immediate consequences of the six assertions in §**3.9**, which were in turn obtained directly from the BHK interpretation. For instance, from $|\Gamma \vee \Delta| = |\Gamma| \sqcup |\Delta|$ we immediately obtain (2), since $S \sqcup T \neq \varnothing$ if and only if $S \neq \varnothing$ or $T \neq \varnothing$. On the other hand, since $\checkmark = \times \rightarrow \times$ by definition, judgement (7) can be considered as a special case of either $?(\Gamma \rightarrow \Gamma)$ or $?(\times \rightarrow \Gamma)$. Both are clearly true on the BHK interpretation: (i) given a solution of Γ, by doing nothing to it we get a solution of Γ, which is a general method — hence a solution of $\Gamma \rightarrow \Gamma$; (ii) $\times \rightarrow \Gamma$ was already discussed in §**3.14**, (5).

Let us formulate a theory T_{BHK} over classical logic attempting to capture the meaning of the above informal judgements. This will be only a countable theory in the traditional sense, i.e. a countable (and recursively enumerable) set of axioms, not a finite meta-conjunction of formulas and principles.

For each n and each n-formula Φ in the language of intuitionistic logic, let us introduce an n-ary predicate constant, denoted $?\Phi$. The following are the axioms of T_{BHK}:

(1) $?(\Phi \wedge \Psi) \ \longleftrightarrow \ ?\Phi \wedge ?\Psi$;
(2) $?(\Phi \vee \Psi) \ \longleftrightarrow \ ?\Phi \vee ?\Psi$;
(3) $?(\Phi \rightarrow \Psi) \ \longrightarrow \ (?\Phi \rightarrow ?\Psi)$;
(4) $\neg ?\times$;
(5) $?\exists x \, \Phi(x) \ \longleftrightarrow \ \exists x \, ?\Phi(x)$;
(6) $?\forall x \, \Phi(x) \ \longrightarrow \ \forall x ?\Phi(x)$;

(7) $?\checkmark$.

Suppose that we have a Leibniz–Euler model of T_{BHK} in subsets of a set X with some domain $\mathscr{D}$. Let us call a subset of X *intuitionistic* if it of the form $|?\Phi|_{u}$ for some formula Φ in the language of intuitionistic logic, and some assignment u in $\mathscr{D}$. Thus $\varnothing$ is intuitionistic by (4), and X is intuitionistic by (7). The intersection of two intuitionistic sets is a set of the form $|?\Phi|_{u} \cap |?\Psi|_{u'}$. Here there is no loss of generality in considering only such pairs of formulas Φ, Ψ that the free variables of Φ have no overlap with the free variables of Ψ; consequently, we may additionally assume that $u' = u$; we are also free to keep the latter assumption but drop the former one. Now $|?\Phi|_{u} \cap |?\Psi|_{u} = |?\Phi \wedge ?\Psi|_{u}$, which by (1) is the same as $|?(\Phi \wedge \Psi)|_{u}$, an intuitionistic set. Similarly by (2), the union of two intuitionistic sets in intuitionistic. Also (5) implies that unions of some $\mathscr{D}$-indexed families of intuitionistic sets are intuitionistic.

Thus intuitionistic subsets of X are very similar to open subsets in some topology on X, except that they are known to be closed only under some infinite unions.

Then let us try again. Suppose that we have a Leibniz–Euler model of T_{BHK} in subsets of a set X with some domain $\mathscr{D}$. Let us call a subset of X *open* if it a union of subsets of the form $|?\Phi|_{u}$, where Φ is a formula of intuitionistic logic, and u is an assignment in $\mathscr{D}$. Thus the union of any family of open sets is open. Also, similarly to the above, the intersection of two open sets is open, and $\varnothing$ and X are open. Thus we get a topology on X.

Another apparent deficiency here is that the axiom lists (3) and (6), which provide a minimal formal extract from the intrinsically vague BHK clauses for $\to$ and $\forall$, seem to be too weak to yield any interesting conclusions as they provide only one-sided constraints on the intuitinistic $\to$ and $\forall$.

One way to strengthen (3) and (6) to something bidirectional is as follows:

(3′) $\big(?\Xi \to ?(\Phi \to \Psi)\big) \iff \big(?\Xi \to (?\Phi \to ?\Psi)\big)$;

(6′) $\big(?\Xi \to ?\forall\mathsf{x}\,\Phi(\mathsf{x})\big) \iff \big(?\Xi \to \forall\mathsf{x}\,?\Phi(\mathsf{x})\big)$.

Indeed, (3′) implies (3) by considering $\Xi = \Phi \to \Psi$, and (6′) implies (6) by considering $\Xi = \forall\mathsf{x}\,\Phi(\mathsf{x})$. Conversely, (3) implies the $\Rightarrow$ meta-implication in (3′), and (6) implies the $\Rightarrow$ meta-implication in (6′), by using the modus ponens rule of classical logic. As for the $\Leftarrow$ meta-implications, they specialize by considering $\Xi = \checkmark$ and using (7) to the following partial converses to (3) and (6):

(3″) $?\Phi \to ?\Psi \implies ?(\Phi \to \Psi)$;

(6″) $\forall\mathsf{x}\,?\Phi(\mathsf{x}) \implies ?\forall\mathsf{x}\,\Phi(\mathsf{x})$.

Remark 5.1. Let us emphasize that there is nothing in the BHK interpretation that would suggest strengthening (3) and (6) in this particular way — although *some* strengthening of (3) and (6) is certainly present in the BHK interpretation. In the QHC calculus of [1], which pretends to capture some formal aspects of the BHK interpretation along the lines of the present section, the $\Leftarrow$ meta-implications in (3) and (6) correspond to the rules $?\gamma \to (?\alpha \to ?\beta) \,/\, ?\gamma \to ?(\alpha \to \beta)$ and $?\gamma \to \forall\mathsf{x}\,?\alpha(\mathsf{x}) \,/\, ?\gamma \to ?\forall\mathsf{x}\,\alpha(\mathsf{x})$, which are shown to be non-derivable, and even non-admissible in QHC [2].

Let us denote the extension of T_{BHK} by the meta-axiom lists (3′) and (6′) by T'_{BHK}. (Here (3) and (6) could of course be dropped from T'_{BHK}.) Then (3″) and (6″) yield:

(3‴) $T'_{BHK} \vdash ?\Phi \to ?\Psi$ implies $T'_{BHK} \vdash ?(\Phi \to \Psi)$;
(6‴) $T'_{BHK} \vdash \forall\mathsf{x}\, ?\Phi(\mathsf{x})$ implies $T'_{BHK} \vdash ?\forall\mathsf{x}\, \Phi(\mathsf{x})$.

Now let us assume that our Leibniz–Euler model of T_{BHK} happens to satisfy also the meta-axioms of T'_{BHK}. (One way to achieve this is to consider the theory T''_{BHK} consisting of all formulas that are consequences of T'_{BHK}, and consider a Leibniz–Euler model of T''_{BHK} along with a valuation with respect to which T''_{BHK} is complete, cf. [168; VIII.3.3, II.8.1]. This model with valuation will necessarily satisfy all admissible, and not only derivable, rules of T'_{BHK}.) Then (3′) guarantees that $|?(\Phi \to \Psi)|_u$ is the largest open set contained in $|?\Phi \to ?\Psi|_u$, and (6′) guarantees that $|?\forall\mathsf{x}\, \Phi(\mathsf{x})|_u$ is the largest open set contained in $|\forall\mathsf{x}\, ?\Phi(\mathsf{x})|_u$. Symbolically, $|?(\Phi \to \Psi)|_u = \mathrm{Int}\big((X \setminus |?\Phi|_u) \cup |?\Psi|_u\big)$ and $|?\forall\mathsf{x}\, \Phi(\mathsf{x})|_u = \mathrm{Int}\bigcap_{d \in D} |?\Phi(d)|_u$, where Int denotes topological interior.

Thus we can also go the other way round: given any topological space X and a domain $\mathscr{D}$, we can interpret every predicate constant $?\alpha$, where α is a problem variable, by an arbitrary open subset of X (for any given assignment). Then we can interpret all our predicate constants $?\Phi$ using inductively the above interpretations of $?(\Phi \wedge \Psi)$, $?(\Phi \vee \Psi)$, $?(\Phi \to \Psi)$, $?\times$, $?\exists\mathsf{x}\, \Xi(\mathsf{x})$, $?\forall\mathsf{x}\, \Xi(\mathsf{x})$, $?\checkmark$ in terms of given interpretations of $?\Phi$, $?\Psi$, $?\Xi(\mathsf{x})$. (In particular, each $|?\Phi|_u$ will be an open subset of X.) This yields a model of T'_{BHK}.

Now, how does T'_{BHK} fit with the official laws and inference rules of intuitionistic logic? (3) implies $T'_{BHK} \vdash ?\Phi \,\&\, ?(\Phi \to \Psi) \Rightarrow ?\Psi$, and from (3‴) we immediately obtain $T'_{BHK} \vdash ?(\Phi \to \Phi)$ and $T'_{BHK} \vdash ?(\times \to \Phi)$. If, in addition to (3‴), we take into account (1) and (2), we also get $T'_{BHK} \vdash ?(\Phi_1 \wedge \Phi_2 \to \Phi_i)$ and $T'_{BHK} \vdash ?(\Phi_i \to \Phi_1 \vee \Phi_2)$ for each $i = 1, 2$. Using (5) and (6), we similarly obtain $T'_{BHK} \vdash ?\big(\Phi(\mathsf{t}) \to \exists\mathsf{x}\, \Phi(\mathsf{x})\big)$ and $T'_{BHK} \vdash ?\big(\forall\mathsf{x}\, \Phi(\mathsf{x}) \to \Phi(\mathsf{t})\big)$.

To verify that $T'_{BHK} \vdash ?\big(\Xi \to \Phi(\mathsf{x})\big) \Rightarrow ?\big(\Xi \to \forall\mathsf{x}\, \Phi(\mathsf{x})\big)$, by (3″) and (6′) it suffices to check that $T'_{BHK} \vdash ?\big(\Xi \to \Phi(\mathsf{x})\big) \to \big(?\Xi \to \forall\mathsf{x}\, ?\Phi(\mathsf{x})\big)$, which indeed follows from (3). This takes care of (XII′), and similarly one can deal with (XIII′), (VIII) and (IX). The same approach also works for (III) and (IV).

Thus if we regard each open set $|?\Phi|_u$ as an interpretation of the intuitionistic formula Φ, and not just of the predicate constant $?\Phi$, we get a model of intuitionistic logic.

5.2. Tarski models

Let us summarize in closed terms the models of intuitionistic logic that we have just obtained. In the zero-order case, they were discovered independently by Stone [189], Tang [192] (see also §5.7.3) and Tarski [194] in the 1930s, with Tarski having also established completeness of zero-order intuitionistic logic with respect to this class of models. The models were extended to the first-order case by Mostowski, with completeness established by Rasiowa and Sikorski (see §5.3).

We will consider only models of intuitionistic logic, and not of empty theories over it. Thus there will be no problem constants and function symbols in the language.

We fix an arbitrary topological space X and a domain (i.e., an arbitrary set) $\mathscr{D}$. Then the set $\mathscr{O}$ consists of all open subsets of X, and the function $! \colon \mathscr{O} \to \{\top, \bot\!\!\!\bot\}$ sends the entire space X to $\top$, and each proper open subset of X to $\bot\!\!\!\bot$.

Thus, to specify a valuation, each problem variable α of arity n must be interpreted by a $\mathscr{D}^n$-indexed family of open subsets of X (i.e., by a function $|\alpha| \colon \mathscr{D}^n \to \mathscr{O}$). Upon a choice of a valuation v, a closed formula will be interpreted by an open subset of X, and more generally a closed n-formula Φ will be interpreted by a $\mathscr{D}^n$-indexed family of open subsets of X, to be denoted $|\Phi|$ or in more detail $|\Phi|^v$, or in still more detail $|\Phi|^v_X$. (The valuation will be suppressed in the notation when it is clear from context.) Upon a further choice of a variable assignment u, an arbitrary formula will be interpreted by an open subset of X, and more generally an arbitrary n-formula Φ will be interpreted by a $\mathscr{D}^n$-indexed family of open subsets of X, to be denoted $|\Phi|$ or in more detail $|\Phi|_u$ or in still more detail $|\Phi|^v_u$. (The assignment will be suppressed in the notation when it is clear from context.) Validity of an n-formula Φ in the model, $\vDash \Phi$, means that $|\Phi|^v_u(t_1, \ldots, t_n) = X$ for each tuple $(t_1, \ldots, t_n) \in \mathscr{D}^n$ and for an arbitrary valuation v and assignment u.

Finally, intuitionistic connectives and quantifiers are interpreted as follows:

- $|\Phi \vee \Psi| = |\Phi| \cup |\Psi|$;
- $|\Phi \wedge \Psi| = |\Phi| \cap |\Psi|$;
- $|\Phi \to \Psi| = \mathrm{Int}\big((X \setminus |\Phi|) \cup |\Psi|\big)$;
- $|\times| = \varnothing$;
- $|\exists \mathsf{x}\, \Xi(\mathsf{x})| = \bigcup_{d \in D} |\Xi|(d)$;
- $|\forall \mathsf{x}\, \Xi(\mathsf{x})| = \mathrm{Int} \bigcap_{d \in D} |\Xi|(d)$,

where Φ and Ψ are formulas and Ξ is a 1-formula, and u and v are fixed. Let us note that $|\Phi \to \Psi|$ is the union of all open sets U such that $U \cap |\Phi| \subset U \cap |\Psi|$.

In particular, we have $|\neg\Phi| = \mathrm{Int}(X \setminus |\Phi|) = X \setminus \mathrm{Cl}\,|\Phi|$ and $|\neg\neg\Phi| = \mathrm{Int}(\mathrm{Cl}\,|\Phi|)$. Thus *decidable* formulas (i.e. Φ such that $\vdash \Phi \vee \neg\Phi$) are represented by clopen (=closed and open) sets; and *stable* formulas (i.e. Φ such that $\vdash \neg\neg\Phi \to \Phi$) are represented by regular open sets (i.e. sets equal to the interior of their closure). Stable formulas coincide (by (3) and (1) in §**3.14**) with those that are equivalent to a negated formula. Decidable formulas are stable by (8). On the other hand, if $\Phi \vee \neg\Phi$ is stable, then Φ is decidable by (32).

Tarski models yield simple and intuitive proofs that many classically valid principles are not derivable in intuitionistic logic. Here are a few examples.

5.2.1. Constant Domain Principle. The converse to (15), known as the *Constant Domain Principle* (for reasons related to Kripke models):

$$\cdot \forall \mathsf{x}\, \big(\alpha \vee \pi(\mathsf{x})\big) \;\longrightarrow\; \alpha \vee \forall \mathsf{x}\, \pi(\mathsf{x})$$

is not derivable in intuitionistic logic, as witnessed by the Tarski model with $X = \mathbb{R}$ and $\mathscr{D} = \mathbb{N}$ and a valuation with $|\pi|(n) = (-\frac{1}{n}, \frac{1}{n})$ and $|\alpha| = \mathbb{R} \setminus \{0\}$. A simple modification of this model and valuation, which we purposely leave to the reader, shows non-derivability

of the following Negative Constant Domain Principle:

$$\cdot\, \forall \mathsf{x} \left(\neg \alpha \vee \pi(\mathsf{x}) \right) \;\longrightarrow\; \neg \alpha \vee \forall \mathsf{x}\, \pi(\mathsf{x}).$$

In fact, the contrapositive of the Constant Domain Principle is derivable in intuitionistic logic:

$$\vdash \neg \forall \mathsf{x} \left(\alpha \vee \pi(\mathsf{x}) \right) \;\longleftrightarrow\; \neg\left(\alpha \vee \forall \mathsf{x}\, \pi(\mathsf{x}) \right).$$

Indeed, the right hand side is equivalent, by (40) and (13), to $\neg \forall \mathsf{x} \left(\neg \alpha \to \pi(\mathsf{x}) \right)$, and the "$\leftarrow$" implication follows from (7). (The "$\to$" implication is the contrapositive of (15).)

However, the following two-variable generalization of the Constant Domain Principle:

$$\cdot\, \forall \mathsf{y}\, \forall \mathsf{x} \left(\rho(\mathsf{y}) \vee \pi(\mathsf{x}, \mathsf{y}) \right) \;\longrightarrow\; \forall \mathsf{y} \left(\rho(\mathsf{y}) \vee \forall \mathsf{x}\, \pi(\mathsf{x}, \mathsf{y}) \right)$$

does not turn into an intuitionistically derivable principle upon taking the contrapositive:

$$\cdot\, \neg \forall \mathsf{y} \left(\rho(\mathsf{y}) \vee \forall \mathsf{x}\, \pi(\mathsf{x}, \mathsf{y}) \right) \;\longrightarrow\; \neg \forall \mathsf{y}\, \forall \mathsf{x} \left(\rho(\mathsf{y}) \vee \pi(\mathsf{x}, \mathsf{y}) \right).$$

Indeed, the latter principle is not valid in the Tarski model with $X = \mathscr{D} = \mathbb{R}$, by considering a valuation with $|\pi|(x, y) = (y - e^x, y + e^x)$ and $|\rho|(y) = \mathbb{R} \setminus \{y\}$. We will call the latter principle (i.e., the contrapositive) the *Parametric Distributivity Principle*. It may be regarded as "dual" to the Constant Domain Principle in that they have a common generalization but no common special cases.

Let us note that the Parametric Distributivity Principle can be equivalently reformulated as

$$\cdot\, \neg\neg \left(\forall \mathsf{y}\, \forall \mathsf{x} \left(\rho(\mathsf{y}) \vee \pi(\mathsf{x}, \mathsf{y}) \right) \;\longrightarrow\; \forall \mathsf{y} \left(\rho(\mathsf{y}) \vee \forall \mathsf{x}\, \pi(\mathsf{x}, \mathsf{y}) \right) \right)$$

using the intuitionistically derivable principles (34) and (44). Hence it is a special case of a variant of Kleene's Principle:

$$\cdot\, \neg\neg \forall \mathsf{y} \left(\forall \mathsf{x} \left(\rho(\mathsf{y}) \vee \pi(\mathsf{x}, \mathsf{y}) \right) \;\longrightarrow\; \left(\rho(\mathsf{y}) \vee \forall \mathsf{x}\, \pi(\mathsf{x}, \mathsf{y}) \right) \right)$$

Kleene's principle [103; §80, Theorem 58(b)(iii)] has $\pi(\mathsf{x})$ instead of $\pi(\mathsf{x}, \mathsf{y})$.

5.2.2. Independence of Premise. The converse to (12) is known as the principle of *Independence of Premise*:

$$\cdot\, \left(\delta \to \exists \mathsf{x}\, \pi(\mathsf{x}) \right) \;\longrightarrow\; \exists \mathsf{x} \left(\delta \to \pi(\mathsf{x}) \right).$$

Its contrapositive

$$\cdot\, \neg\left(\delta \to \exists \mathsf{x}\, \pi(\mathsf{x}) \right) \;\longleftrightarrow\; \neg \exists \mathsf{x} \left(\delta \to \pi(\mathsf{x}) \right)$$

is derivable in intuitionistic logic, which follows from (38) and (18). By contrast, even if δ is specialized to $\neg \alpha$, and $\exists$ is "specialized" to $\vee$ in the principle of Independence of Premise, the resulting formula, known as the *Kreisel–Putnam Principle*:

$$\cdot\, (\neg \alpha \to \beta \vee \gamma) \;\longrightarrow\; (\neg \alpha \to \beta) \vee (\neg \alpha \to \gamma)$$

is not derivable in intuitionistic logic. Indeed, let $X = \mathbb{R}^2$,

$$|\beta| = \{(x, y) \mid x > 0\},$$
$$|\gamma| = \{(x, y) \mid y > 0\}$$

and $|\alpha| = |\neg(\beta \vee \gamma)|$, so that $|\neg\alpha| = \{(x, y) \mid x > 0 \vee y > 0\}$. Then $|\neg\alpha \to \beta \vee \gamma| = \mathbb{R}^2$,

whereas

$$|\neg\alpha \to \beta| = \{(x, y) \mid x > 0 \vee y < 0\};$$
$$|\neg\alpha \to \gamma| = \{(x, y) \mid x < 0 \vee y > 0\}.$$

Hence $|\neg\alpha \to \beta| \cup |\neg\alpha \to \gamma| = \mathbb{R}^2 \setminus \{(0, 0)\}$.

This Tarski model shows also that *Harrop's Rule*:

$$\frac{\neg\alpha \to \beta \vee \gamma}{(\neg\alpha \to \beta) \vee (\neg\alpha \to \gamma)}$$

is not derivable in intuitionistic logic. (Alternatively, if it were derivable, then the Kreisel–Putnam principle would also be derivable by the deduction theorem.) However, it is admissible, see Theorem 5.4 below (hence, in particular, this rule does not imply the Kreisel–Putnam principle). The stronger rule:

$$\frac{\delta \to \beta \vee \gamma}{(\delta \to \beta) \vee (\delta \to \gamma)}$$

is not even admissible, by considering its special case with δ substituted by $\beta \vee \gamma$. In this case the meta-quantified premiss is derivable in intuitionistic logic, but the meta-quantified conclusion, $\cdot(\beta \vee \gamma \to \beta) \vee (\beta \vee \gamma \to \gamma)$, is clearly equivalent to the following *Gödel–Dummett principle*:

$$\cdot(\beta \to \gamma) \vee (\gamma \to \beta),$$

which is certainly not derivable in intuitionistic logic (for instance, the previous model in $\mathbb{R}^2$ works to show this). Let us note that, in fact, our specialization of δ to $\beta \vee \gamma$ led to no loss of generality, since $\cdot(\beta \vee \gamma \to \beta) \vee (\beta \vee \gamma \to \gamma)$ clearly implies (using the exponential law) the quantifier-free independence of premise,

$$\cdot(\delta \to \beta \vee \gamma) \to (\delta \to \beta) \vee (\delta \to \gamma).$$

Thus the latter principle as well as its rule version considered above are each equivalent to the Gödel–Dummett principle. These equivalent principles are strictly stronger than the Kreisel–Putnam principle (see Example 5.6 below).

A similar argument shows that the following stable Harrop's rule:

$$\frac{\neg\neg(\beta \vee \gamma) \to (\beta \vee \gamma), \ \neg\alpha \to \beta \vee \gamma}{(\neg\alpha \to \beta) \vee (\neg\alpha \to \gamma)}$$

is equivalent to the following stable Gödel–Dummett principle:

$$\frac{\neg\neg(\beta \vee \gamma) \to (\beta \vee \gamma)}{(\beta \to \gamma) \vee (\gamma \to \beta)}$$

and to the following stable Kreisel–Putnam principle:

$$\frac{\neg\neg(\beta \vee \gamma) \to (\beta \vee \gamma)}{(\neg\alpha \to \beta \vee \gamma) \ \longrightarrow \ (\neg\alpha \to \beta) \vee (\neg\alpha \to \gamma)};$$

and, moreover, that $\neg\alpha$ can be replaced by δ in these stable versions. In fact, these all are equivalent forms of the original Harrop rule. Indeed, assuming $\neg\alpha \to \beta \vee \gamma$ and writing $\Phi = \neg\alpha \wedge \beta$ and $\Psi = \neg\alpha \wedge \gamma$, we have $\neg\alpha \leftrightarrow \Phi \vee \Psi$. Then also $\neg\neg(\Phi \vee \Psi) \to (\beta \vee \gamma)$, and hence the stable Harrop rule implies $(\neg\alpha \to \Phi) \vee (\neg\alpha \to \Psi)$, which in turn implies $(\neg\alpha \to \beta) \vee (\neg\alpha \to \gamma)$.

By another argument of the same type, the rule

$$\frac{\beta \wedge \gamma \to \delta}{(\beta \to \delta) \vee (\gamma \to \delta)}$$

implies the Gödel–Dummett principle $\cdot\,(\beta \to \gamma) \vee (\gamma \to \beta)$, which in turn implies Skolem's principle (see [210]), the converse of (25):

$$\cdot\,(\beta \wedge \gamma \to \delta) \to (\beta \to \delta) \vee (\gamma \to \delta);$$

thus the three are equivalent. Let us also note that by "generalizing" $\wedge$ and $\vee$ to $\forall$ and $\exists$, Skolem's form of the Gödel–Dummett principle transforms into the converse of (17):

$$\cdot\,(\forall x\,\gamma(x) \to \delta) \;\longrightarrow\; \exists x\,(\gamma(x) \to \delta),$$

which is therefore also an independent principle (by considering a two-element domain).

Remark 5.2. We will see in §5.4 that the principle

$$\cdot\,(\beta \wedge \gamma \to \neg\alpha) \to (\beta \to \neg\alpha) \vee (\gamma \to \neg\alpha)$$

is strictly stronger than the Kreisel–Putnam principle. This might seem surprising, since the two principles are clearly equivalent to the rules

$$\frac{\neg\neg\delta \to \delta}{(\beta \wedge \gamma \to \delta) \to (\beta \to \delta) \vee (\gamma \to \delta)} \quad \text{and} \quad \frac{\neg\neg\delta \to \delta}{(\delta \to \beta \vee \gamma) \to (\delta \to \beta) \vee (\delta \to \gamma)}$$

respectively — even though these rules have the same premiss and equivalent meta-closures of the conclusions. As we have seen, it is enough to meta-quantify over δ:

$$\vdash (\delta)\,(\beta \wedge \gamma \to \delta) \to (\beta \to \delta) \vee (\gamma \to \delta) \iff (\delta)\,(\delta \to \beta \vee \gamma) \to (\delta \to \beta) \vee (\delta \to \gamma).$$

5.3. Topological completeness

Rasiowa and Sikorski proved that every consistent first-order theory $\mathcal{T}$ over intuitionistic logic is complete with respect to the Tarski model in a certain subspace $X = X_{\mathcal{T}}$ of the Baire space[73] with $\mathscr{D} = \mathbb{N}$ [168; X.3.2].[74] In particular, intuitionistic logic is complete with respect to one such Tarski model (in fact, even with respect to a fixed valuation), and strongly complete with respect to the class of all such Tarski models (see 4.63). The latter result was improved by Dragalin, who showed that intuitionisitic logic is strongly complete with respect to the Tarski model with $\mathscr{D} = \mathbb{N}$ and with X the Baire space itself

---

[73]That is, $\mathbb{N}^{\mathbb{N}}$ with the product topology. This space is well-known to be homeomorphic to the set of irrational reals.

[74]Thus if Ψ is not a theorem of $\mathcal{T}$, then Ψ is not valid in $X_{\mathcal{T}}$. It is additionally shown in [168; X.3.4 and remark after X.3.5] that if $\neg\Phi$ is not a theorem of $\mathcal{T}$, then Φ is satisfiable in the Tarski model in a certain subspace $X = X'_{\mathcal{T}}$ of the Baire space with $\mathscr{D} = \mathbb{N}$.

[46][75] (and, in fact, complete even with respect to a fixed valuation in such a model). It is easy to see that intuitionistic logic is not strongly complete with respect to any Tarski model with connected X, including Euclidean spaces (see §5.5.1).

Originally, Tarski [194] (see also [135]; for a sketch, see [188; 2.4.1], and for simplified proofs see [184], [113], [114] and references there) proved that if X is a separable metrizable space with no isolated points (for instance, $\mathbb{R}^n$ for any $n > 0$), then zero-order intuitionistic logic is complete with respect to the Tarski model in X. For instance, the Kreisel–Putnam Principle (and, in fact, even Harrop's Rule) is also disproved by the Tarski model with $X = \mathbb{R}$, by considering a valuation with $|\neg\alpha| = (0, 1)$, $|\beta| = \bigcup_{i=1}^{\infty}(\frac{1}{2i+2}, \frac{1}{2i})$ and $|\gamma| = \bigcup_{i=1}^{\infty}(\frac{1}{2i+1}, \frac{1}{2i-1})$. In general, it should be noted that open subsets of $\mathbb{R}$ can be quite complicated (see [149; §III.6.9]).

Tarski's result was only recently extended to first-order intuitionistic logic by Kremer [116], who showed it to be complete with respect to the Tarski model with $\mathscr{D} = \mathbb{N}$ in any *zero-dimensional* separable metrizable space X with no isolated points; and strongly complete with respect to Tarski models with $\mathscr{D} = \mathbb{N}$ in open subspaces of X.[76] These include the Cantor set, the Baire space, and the set of rational numbers. On the other hand, intuitionistic logic is not complete with respect to the Tarski model with any $\mathscr{D}$ in any locally connected space, including open subsets of Euclidean spaces [116].

5.3.1. Disjunction property. Interestingly, if Φ is a closed formula such that $|\Phi \vee \neg\Phi|^v = X$ for some valuation v for the Tarski model in a connected space X, then either $|\Phi|^v = X$ or $|\neg\Phi|^v = X$ (since the only clopen subsets of X are $\varnothing$ and X). In other words, for closed formulas Φ the judgement $\Phi \vee \neg\Phi \vDash \Phi \parallel \neg\Phi$ is valid in Tarski models with connected X (for instance, for $X = \mathbb{R}$). It is clearly not valid for any disconnected X; in particular, the judgement $\Phi \vee \neg\Phi \vdash \Phi \parallel \neg\Phi$ does not hold in zero-order intuitionistic logic.

However, the corresponding external judgement holds: in zero-order intuitionistic logic, if $\vdash \Phi \vee \neg\Phi$, then either $\vdash \Phi$ or $\vdash \neg\Phi$. This is a special case of the following remarkable fact, originally announced by Gödel and proved by Gentzen (see also [103; §80, Theorem 57(a)], [168; X.8.1], [135; Theorem 4.4]):

Theorem 5.3 (Disjunction Property). *In intuitionistic logic, if $\vdash \Phi \vee \Psi$ then either $\vdash \Phi$ or $\vdash \Psi$, for any closed formulas Φ and Ψ.*

---

[75]Dragalin's theorem [46; 3.5.2] says that every composite theory (see Remark 4.64) has a Tarski model with $\mathscr{D} = \mathbb{N}$ and with X the Baire space; this trivially implies the assertion (see Remark 4.64).

[76]Kremer's theorem says, in particular, that if Γ and Δ are closed formulas such that $\Gamma \to \Delta$ is not a theorem of intuitionistic logic, then $|\Gamma|^v \setminus |\Delta|^v \neq \varnothing$ for some valuation v in the Tarski model in X with $\mathscr{D} = \mathbb{N}$. If Ψ is not a theorem of intuitionistic logic, then its universal closure Δ is also not a theorem, and hence $|\checkmark|^v \setminus |\Delta|^v \neq \varnothing$ for some v, i.e., $|\Delta|^v \neq X$, and so Ψ is not valid in X. If a rule $\Phi_1, \ldots, \Phi_n / \Psi$ is not derivable in intuitionistic logic, then $\Gamma \to \Delta$ is not a theorem of intuitionistic logic, where Γ is the universal closure of $\Phi_1 \wedge \cdots \wedge \Phi_n$ and Δ is the universal closure of Ψ. Hence $|\Gamma|^v \setminus |\Delta|^v \neq \varnothing$ for some valuation v in the Tarski model in X with $\mathscr{D} = \mathbb{N}$. If we let $Y = |\Gamma|^v$ and $|\Xi|_u^{v_Y} = |\Xi|_u^v \cap Y$ for every formula Ξ and every assignment u in $\mathscr{D} = \mathbb{N}$, then it is easily checked by induction that v_Y is a valuation in the Tarski model in Y with $\mathscr{D} = \mathbb{N}$; clearly, $|\Gamma|^{v_Y} = Y$ and $|\Delta|^{v_Y} \neq Y$. Hence $\Phi_1, \ldots, \Phi_n$ are valid in (Y, v_Y), and Ψ is not valid in (Y, v_Y). Thus $\Phi_1, \ldots, \Phi_n / \Psi$ is not valid in Y.

It is easy to deduce this from the completeness of Tarski models.

The proof employs the following construction, which will also be used in sequel: the Alexandroff dual cone C^*W over the space W is the union $W \cup \{\hat{0}\}$ (as a set) endowed with the topology whose open sets are as follows: (i) open subsets of W; (ii) $W \cup \{\hat{0}\}$.

Proof. Suppose that $\vdash \Phi \vee \Psi$, but $\not\vdash \Phi$ and $\not\vdash \Psi$. Then there are valuations v_X, v_Y for the Tarski models in topological spaces X, Y with domains $\mathscr{D}_X$ and $\mathscr{D}_Y$ such that $|\Phi|^{v_X} \neq X$ and $|\Psi|^{v_Y} \neq Y$. Let $\mathscr{D}$ be a set that admits surjections d_X onto $\mathscr{D}_X$ and d_Y onto $\mathscr{D}_Y$. Let us define valuations v'_X, v'_Y for the Tarski models in X and Y with domain $\mathscr{D}$ by $|\alpha|^{v'_X}_u = |\alpha|^{v_X}_{d_X \circ u}$ and $|\alpha|^{v'_Y}_u = |\alpha|^{v_Y}_{d_Y \circ u}$ for every problem variable α and every assignment u in $\mathscr{D}$. Then it is easy to see that $|\Xi|^{v'_X} = |\Xi|^{v_X}$ and $|\Xi|^{v'_Y} = |\Xi|^{v_Y}$ for every closed formula Ξ. In particular, $|\Phi|^{v'_X} \neq X$ and $|\Psi|^{v'_Y} \neq Y$.

Let $Z := C^*(X \sqcup Y)$ be the Alexandroff dual cone over the disjoint union $X \sqcup Y$. Let us define a valuation v_Z for the Tarski model in Z with domain $\mathscr{D}$ by $|\alpha|^{v_Z}_u = |\alpha|^{v'_X}_u \cup |\alpha|^{v'_Y}_u$ for every problem variable α and every assignment u in $\mathscr{D}$. Then it is easy to check by induction (using that X and Y are open in Z) that $|\Xi|^{v_X}_u = |\Xi|^{v_Z}_u \cap X$ and $|\Xi|^{v_Y}_u = |\Xi|^{v_Z}_u \cap Y$ for any formula Ξ and any assignment u in $\mathscr{D}$. In particular, $|\Phi|^{v_Z} \not\supset X$ and $|\Psi|^{v_Z} \not\supset Y$.

By our hypothesis, $|\Phi \vee \Psi|^{v_Z} = Z$. In particular, either $\hat{0} \in |\Phi|^{v_Z}$ or $\hat{0} \in |\Psi|^{v_Z}$. However, the minimal open set containing $\hat{0}$ is the entire space Z. Hence either $|\Phi|^{v_Z} = Z$ or $|\Psi|^{v_Z} = Z$, which is a contradiction. $\qquad \square$

5.3.2. Harrop's rule. The Disjunction Property can be regarded as a multiple-conclusion admissible rule in zero-order intuitionistic logic. A rather similar argument establishes the admissibility of one genuine (single-conclusion) rule in first-order intuitionisitc logic:

Theorem 5.4. *Harrop's rule*

$$\frac{\neg\gamma \to \varphi \vee \psi}{(\neg\gamma \to \varphi) \vee (\neg\gamma \to \psi)}$$

is an admissible rule for intuitionistic logic.

See [98] for a direct syntactic proof not appealing to the completeness theorem.

It should be mentioned that Harrop's rule is, in fact, stably admissible [166].

The admissibility of Harrop's rule implies the following generalization of the Disjunction Property:

Corollary 5.5. *In intuitionistic logic, if $\neg\Gamma \vdash \Phi \vee \Psi$ then either $\neg\Gamma \vdash \Phi$ or $\neg\Gamma \vdash \Psi$, for any closed formulas Γ, Φ and Ψ.*

Proof. The hypothesis along with the Deduction Theorem imply $\vdash \neg\Gamma \to \Phi \vee \Psi$. Since Harrop's rule is admissible, we have $\vdash (\neg\Gamma \to \Phi) \vee (\neg\Gamma \to \Psi)$. By the Disjunction Property, it follows that either $\vdash \neg\Gamma \to \Phi$ or $\vdash \neg\Gamma \to \Psi$. This along with the *modus ponens* implies the assertion. $\qquad \square$

Proof of Theorem 5.4. We will prove the admissibility of the following rule, which as shown in §5.2.2 is equivalent to Harrop's rule:

$$\frac{\neg\neg(\gamma \vee \delta) \to (\gamma \vee \delta)}{(\gamma \to \delta) \vee (\delta \to \gamma)}.$$

Suppose that $\vdash \neg\neg(\Gamma \vee \Delta) \to (\Gamma \vee \Delta)$ and $\nvdash (\Gamma \to \Delta) \vee (\Delta \to \Gamma)$ for some formulas Γ and Δ. Then by Tarski's completeness theorem, there exist a space X and a domain $\mathscr{D}$ such that $|(\Gamma \to \Delta) \vee (\Delta \to \Gamma)|^{v_X}_{u_0} \neq X$ for some assignment u_0 in $\mathscr{D}$ and some valuation v_X in the Tarski model in X with domain $\mathscr{D}$. Then, in particular, $|\Gamma|^{v_X}_{u_0} \not\subset |\Delta|^{v_X}_{u_0}$ and $|\Delta|^{v_X}_{u_0} \not\subset |\Gamma|^{v_X}_{u_0}$.

Let us define a valuation v_Y in the Tarski model in $Y := |\Gamma \vee \Delta|^{v_X}_{u_0}$ with domain $\mathscr{D}$ by $|\alpha|^{v_Y}_u = |\alpha|^{v_X}_u \cap Y$ for every problem variable α and every assignment u in $\mathscr{D}$. Then $|\Xi|^{v_Y}_u = |\Xi|^{v_X}_u \cap Y$ for every formula Ξ and every assignment u (using that Y is open in X). In particular, $|\Gamma \vee \Delta|^{v_Y}_{u_0} = Y$, and neither of $|\Gamma|^{v_Y}_{u_0}$, $|\Delta|^{v_Y}_{u_0}$ lies in the other one.

Let Z be the Alexandroff dual cone $C^*Y = Y \cup \{\hat{0}\}$. Let us define a valuation v_Z in the Tarski model in Z with domain $\mathscr{D}$ by $|\alpha|^{v_Z}_u = |\alpha|^{v_Y}_u$ for every problem variable α and every assignment u in $\mathscr{D}$. Then $|\Xi|^{v_Y}_u = |\Xi|^{v_Z}_u \cap Z$ for every formula Ξ and every assignment u (using that Y is open in Z). In particular, neither of $|\Gamma|^{v_Z}_{u_0}$, $|\Delta|^{v_Z}_{u_0}$ lies in the other one. Also, $|\Gamma \vee \Delta|^{v_Z}_{u_0}$ contains Y, and consequently $|\neg\neg(\Gamma \vee \Delta)|^{v_Z}_{u_0} = Z$ (since $\text{Cl}\, Y = Z$). But then our hypothesis $\vdash \neg\neg(\Gamma \vee \Delta) \to (\Gamma \vee \Delta)$ implies that $|\Gamma \vee \Delta|^{v_Z}_{u_0} = Z$.

In particular, either $\hat{0} \in |\Gamma|^{v_Z}_{u_0}$ or $\hat{0} \in |\Delta|^{v_Z}_{u_0}$. However, the minimal open set containing $\hat{0}$ is the entire space Z. Thus either $|\Gamma|^{v_Z}_{u_0} = Z$ or $|\Delta|^{v_Z}_{u_0} = Z$, which is a contradiction. □

5.3.3. Alexandroff spaces. McKinsey and Tarski also proved that zero-order intuitionistic logic is complete with respect to the class of Tarski models with finite X. Finite topological spaces are included in the class of *Alexandroff spaces*, where the intersection of any (possibly infinite) family of open subsets of X is open. The relation $x \in \text{Cl}\, \{y\}$ on the points x, y of an Alexandroff space is a preorder, i.e. it is reflexive and transitive. Conversely, any preordered set P can be endowed with its Alexandroff topology, where a subset of P is defined to be closed if it is an order ideal. This gives a one-to-one correspondence between Alexandroff spaces and preordered sets, under which T_0 spaces[77] correspond precisely to posets.

In the zero-order case, Tarski models in Alexandroff spaces with fixed valuations are also known (in somewhat different terms) as Kripke models. In general, Tarski models in Alexandroff spaces with fixed valuations are the same as "Kripke models with constant domain". (See **??** concerning the case of "non-constant domain"). First-order intuitionistic logic is not complete with respect to the class of Tarski models in Alexandroff spaces: clearly, the Constant Domain Principle holds in all such models. In fact, this is the only obstacle to completeness: by a result of S. Görneman, the extension of intuitionistic logic by the Constant Domain Principle is complete with respect to Tarski models in Alexandroff spaces (see [59; §3.3, Corollary 8]).

---

[77]That is, spaces where at least one out of any two distinct points is contained in an open set not containing the other point.

Let us note that if the Alexandroff topology on a poset is T_1,[78] then it is discrete. Let us note also that every Alexandroff space X admits a continuous open surjection onto the T_0 Alexandroff space obtained as the quotient of X by the equivalence relation whose equivalence classes are the minimal nonempty open sets of X.

If K is a simplicial complex (or more generally a cell complex or a cone complex, see [140]), viewed as a topological space (with the metric topology, see [139]), let $X(K)$ be the face poset of K, viewed as a T_0 Alexandroff space. By sending a point of K to the minimal simplex (or cell or cone) that contains it, we obtain a continuous open surjection $K \to X(K)$. Conversely, given a T_0 Alexandroff space X, viewed as a poset, we have its order complex, whose simplices can be arranged into cones of a cone complex $K(X)$ such that $X(K(X)) = X$ (see [140]). Thus we have a continuous open surjection $K(X) \to X$. Thus a Tarski model in an Alexandroff space X gives rise to a Tarski model in the polyhedron $K(X)$, and all completeness results can be transferred accordingly.

Example 5.6. The Gödel–Dummett principle $\cdot (\beta \to \gamma) \vee (\gamma \to \beta)$ is obviously valid in all Alexandroff spaces associated to totally ordered posets. It is not valid in every Alexandroff space X corresponding to a poset that contains a pair of non-comparable elements q, r with a lower bound p. Indeed, the smallest open set U_p containing p also contains q and r, and consequently p lies neither in $\mathrm{Int}\left(U_q \cup (X \setminus U_r)\right)$ (which does not contain r) nor in $\mathrm{Int}\left(U_r \cup (X \setminus U_q)\right)$ (which does not contain q).

Meanwhile, the Kreisel–Putnam principle $\cdot (\neg\alpha \to \beta \vee \gamma) \to (\neg\alpha \to \beta) \vee (\neg\alpha \to \gamma)$ is valid, for instance, in the three-element poset P consisting of a pair of incomparable elements q, r and their lower bound p. Indeed, $\neg\alpha$ can only be interpreted by a regular open set, and the only regular open sets in P are $\varnothing$, P, $\{q\}$ and $\{r\}$. It is easy to see that the Kreisel–Putnam principle is satisfied under any valuation such that $|\neg\alpha| = |\times|$ or $|\neg\alpha| = |\checkmark|$ or $|\gamma| \subset |\beta|$ or $|\beta| \subset |\gamma|$. This leaves only the following options: $|\beta| = \{q\}$ and $|\gamma| = \{r\}$ (or vice versa), with $\neg\alpha$ interpreted by either $\{q\}$ or $\{r\}$. In each of these cases either $|\neg\alpha| = |\beta|$ or $|\neg\alpha| = |\gamma|$, and the Kreisel–Putnam principle is again satisfied.

Thus the Kreisel–Putnam principle does not imply the Gödel–Dummett principle.

Remark 5.7. In fact, the Kreisel–Putnam principle holds in every poset P satisfying the following *van Benthem condition* (cf. [94; p. 886]):

$$\forall x \in P \; \forall y, z \geq x \; \exists u \geq x \; \left(u \leq y, z \text{ and } \forall v \geq u \; \exists w \geq v \; (w \geq y \text{ or } w \geq z)\right).$$

Indeed, suppose that P satisfies this condition, but the Tarski model in P does not satisfy the Kreisel–Putnam principle. Then

$$\mathrm{Int}(\mathrm{Cl}(A) \cup B \cup C) \not\subset \mathrm{Int}(\mathrm{Cl}(A) \cup B) \cup \mathrm{Int}(\mathrm{Cl}(A) \cup C)$$

for some open subsets A, B, C of P. So there exists an $x \in \mathrm{Int}(\mathrm{Cl}(A) \cup B \cup C)$ such that $x \notin \mathrm{Int}(\mathrm{Cl}(A) \cup B)$ and $x \notin \mathrm{Int}(\mathrm{Cl}(A) \cup C)$. Then the smallest open set U_x containing

---

[78]A topological space is called T_1 if for every pair of distinct points, each is contained in an open set not containing the other one; or equivalently if all singleton subsets are closed.

x contains some $y \notin \mathrm{Cl}(A) \cup B$ and some $z \notin \mathrm{Cl}(A) \cup C$. Thus $y, z \geq x$ and $y, z \notin \mathrm{Cl}\, A$; also, $y \notin B$ and $z \notin C$. Let u be given by the van Benthem condition. Then $u \notin B$, $u \notin C$, so $u \notin B \cup C$. Also, $u \in \mathrm{Cl}(A) \cup B \cup C$ since $u \geq x$. Hence $u \in \mathrm{Cl}\, A$. Then there exists a $v \geq u$ such that $v \in A$. If w is given by the van Benthem condition, then $w \in A$ and either $w \geq y$ or $w \geq z$. Then either $y \in \mathrm{Cl}\, A$ or $z \in \mathrm{Cl}\, A$, which is a contradiction.

Remark 5.8. Iemhoff [95; 3.14] characterized admissible rules of zero-order intuitionistic logic as those rules that are valid in the Tarski model in every T_0 Alexandroff space whose corresponding poset has a smallest element and a "tight" lower bound for every finite subset Y, that is, a lower bound z such that every $x > z$ satisfies $x \geq y$ for some $y \in Y$. (It should be noted here that the admissibility of a rule Φ/Ψ means not just that $\vdash \Phi$ implies $\vdash \Psi$, but that $\vdash \Phi'$ implies $\vdash \Psi'$ for every special case Φ'/Ψ' of Ψ/Ψ; whereas the validity of Φ/Ψ in a class of models means just that $\vDash_v \Phi$ implies $\vDash_v \Psi$ for all valuations v in all those models. But, of course, the validity of Φ/Ψ in a class of models implies the validity of every special case Φ'/Ψ' of Ψ/Ψ in the same class of models, due to the quantification over all valuations. It is essential here that valuations are not included in our definition of a "model".)

Iemhoff also gave another characterization of admissible rules of zero-order intuitionistic logic. Let us define several operations on Tarski models-with-valuations (X, v): (O_1) "restrict" a given (X, v) by reducing X to the smallest open neighborhood of some point; (O_2) extend a given (X, v) in any way over the Alexandroff dual cone C^*X; (O_3) combine a given collection of (X_i, v_i) by taking the disjoint union of X_i. (All these operations are used in the proofs of 5.3, 5.4 above.) Then a rule is admissible for zero-order intuitionistic logic if and only if it is valid in every Tarski model-with-valuation obtained by applying O_3 to any family F of Tarski models-with-valuation in finite T_0 spaces such that F is in the image of O_2, and F is closed under O_1 and $O_2 \circ O_3$ [95; 3.15].

5.4. Jankov's principle

A striking omission in our list of intuitionistic validities in §**3.14** is the converse to one the de Morgan laws (29):

$$\cdot \neg(\gamma \wedge \delta) \longrightarrow \neg\gamma \vee \neg\delta. \tag{JP}$$

To see that this *Jankov's principle* is not derivable in intuitionistic logic, it suffices to substitute a stable but not decidable formula Φ (e.g. $\Phi = \neg\alpha$) for γ and $\neg\Phi$ for δ. Jankov's principle (also known as "de Morgan's principle" or the "weak principle of excluded middle") may have various appearances:

Proposition 5.9. *The following are equivalent to (JP):*

(1) $\cdot \neg\neg(\alpha \vee \beta) \longleftrightarrow \neg\neg\alpha \vee \neg\neg\beta$;
(2) $\cdot \neg\alpha \vee \neg\neg\alpha$;

(3) $\dfrac{\neg\neg\beta \to \beta}{\beta \vee \neg\beta}$.

Let us note that the admissibility of the rule (3) amounts to the assertion that stable formulas are decidable (which of course does not hold in intuitionisitc logic).

Proof. To obtain (1) from (JP), we substitute $\neg\alpha$ for γ and $\neg\beta$ for δ, and use the contrapositive of the other de Morgan tautology (28), $\vdash \neg(\alpha \vee \beta) \leftrightarrow \neg\alpha \wedge \neg\beta$.

To obtain (2) from (1), we apply the latter with $\neg\alpha$ susbstituted for β, and use (3) and (32).

Next, from (30) and (35) we get $\vdash \neg(\gamma \wedge \delta) \longrightarrow (\neg\neg\gamma \to \neg\delta)$, and it follows that $\vdash \neg(\gamma \wedge \delta) \longrightarrow (\neg\gamma \vee \neg\neg\gamma \to \neg\gamma \vee \neg\delta)$. Upon a double application of the exponential law (26), we get $\vdash \neg\gamma \vee \neg\neg\gamma \to \big(\neg(\gamma \wedge \delta) \to \neg\gamma \vee \neg\delta\big)$, so that (2) implies (JP).

Finally, applying the rule (3) with $\neg\alpha$ substituted for β, we get (2). Conversely, we trivially have $\vdash (\neg\alpha \leftrightarrow \beta) \longrightarrow (\neg\alpha \vee \neg\neg\alpha \to \beta \vee \neg\beta)$. The exponential law transforms this validity into $\vdash \neg\alpha \vee \neg\neg\alpha \longrightarrow \big((\neg\alpha \leftrightarrow \beta) \to \beta \vee \neg\beta\big)$. Applying the latter with $\neg\beta$ substituted for α, we get that (2) implies (3). □

Jankov's principle is satisfied in all *extremally disconnected* spaces, that is, spaces where the closure of every open set is open (or equivalently, every regular open set is clopen). All extremally disconnected metrizable spaces are discrete (see [49; Exer. 6.2.G(a)]), and in a certain model of ZF, all extremally disconnected compact Hausdorff spaces are finite [144]. However, in ZFC, a well-known example of an extremally disconnected compact Hausdorff space is the Stone–Čech compactification of any infinite discrete space [49].

On the other hand, the Alexandroff space corresponding to a poset with a greatest element is always extremally disconnected. In general, the Alexandroff space corresponding to a poset P is extremally disconnected if and only if "principal filters of P are directed sets", that is, every pair of elements of P that has a lower bound also has an upper bound. Indeed, if U is an open subset of P whose closure is not open, then some $z \in \mathrm{Cl}\,U$ is bounded above by some $y \notin \mathrm{Cl}\,U$ (since $\mathrm{Cl}\,U$ is not open). Also $z \notin U$ (since U is open), but it is bounded above also by some $x \in U$ (since $z \in \mathrm{Cl}\,U$). However, if u is an upper bound of x and y, then $u \in U$ (since $x \in U$) and therefore $y \in \mathrm{Cl}\,U$, which is a contradiction. Conversely, if $x, y \in P$ have a lower bound z but no upper bound, then the minimal open set U_x containing x has a non-open closure (since $\mathrm{Cl}\,U_x$ contains z but not y).

The Gödel–Dummett principle implies Jankov's principle; more precisely, the latter is a special case of the principle

$$\cdot (\gamma \wedge \delta \to \neg\alpha) \longrightarrow (\gamma \to \neg\alpha) \vee (\delta \to \neg\alpha), \tag{JP$'$}$$

which is in turn a special case of Skolem's form of the Gödel–Dummett principle (see §5.2.2). Jankov's principle does not imply the Gödel–Dummett principle by considering, for instance, the four-element poset consisting of a pair of incomparable elements along with their upper and lower bounds.

On the other hand, Jankov's principle implies the Kreisel–Putnam principle,

$$\cdot (\neg\alpha \to \gamma \vee \delta) \longrightarrow (\neg\alpha \to \gamma) \vee (\neg\alpha \to \delta).$$

Indeed, we clearly have $\vdash (\neg\alpha \to \gamma \vee \delta) \longrightarrow (\neg\alpha \vee \neg\neg\alpha \to \gamma \vee \delta \vee \neg\neg\alpha)$. Hence by the exponential law, $\cdot\neg\alpha \vee \neg\neg\alpha$ implies $\cdot(\neg\alpha \to \gamma \vee \delta) \longrightarrow \gamma \vee \delta \vee \neg\neg\alpha$, and it remains to observe that $\vdash \gamma \vee \delta \vee \neg\neg\alpha \longrightarrow (\gamma \vee \neg\neg\alpha) \vee (\delta \vee \neg\neg\alpha)$ and by (6), $\vdash (\gamma \vee \neg\neg\alpha) \vee (\delta \vee \neg\neg\alpha) \longrightarrow (\neg\alpha \to \gamma) \vee (\neg\alpha \to \delta)$. The Kreisel–Putnam principle does not imply Jankov's principle by considering, for instance, the three-element poset of Example 5.6.

By contrast, (JP$'$) is in fact equivalent to Jankov's principle. Indeed, by the exponential law (JP$'$) is equivalent to

$$\cdot\big(\alpha \to \neg(\gamma \wedge \delta)\big) \longrightarrow (\alpha \to \neg\gamma) \vee (\alpha \to \neg\delta),$$

where substituting $\neg\beta$ for α leads to no loss of generality due to (35). Now Jankov's principle implies $\cdot\big(\neg\beta \to \neg(\gamma \wedge \delta)\big) \longrightarrow (\neg\beta \to \neg\gamma \vee \neg\delta)$, whereas the Kreisel–Putnam principle implies $\cdot(\neg\beta \to \neg\gamma \vee \neg\delta) \longrightarrow (\neg\beta \to \neg\gamma) \vee (\neg\beta \to \neg\delta)$.

To summarize, the following strict entailments of principles have been discussed:

Gödel–Dummett pr. $\vdash$ Jankov's principle $\vdash$ Kreisel–Putnam pr. $\vdash$ Harrop's rule

5.4.1. Mints-Citkin rule and Rose's formula. Harrop's rule is a consequence of the following Mints–Citkin rule [33]

$$\frac{(\lambda \to \mu) \to \varphi \vee \psi}{\big((\lambda \to \mu) \to \varphi\big) \vee \big((\lambda \to \mu) \to \psi\big) \vee \big((\lambda \to \mu) \to \lambda\big)}.$$

Indeed, the Mints–Citkin rule with $\mu = \times$ implies Harrop's rule, since $\vdash (\neg\lambda \to \lambda) \to (\neg\lambda \to \varphi)$ (due to $\vdash \neg\lambda \wedge (\neg\lambda \to \lambda) \wedge \neg\lambda \to \times$ and $\vdash \times \to \varphi$).

Since Harrop's rule is not derivable (see §*5.2.2*), the Mints–Citkin rule is also not derivable. But it is known to be admissible (see [95; proof of Theorem 3.13]). In fact, a slight generalization of this rule is the first in an infinite list of admissible rules (the so-called Visser rules) that imply every admissible rule of zero-order intuitionistic logic, by a celebrated result of P. Rozière [170] and R. Iemhoff [95].

But in contrast to Harrop's rule, the Mints–Citkin rule is not stably admissible [96; Theorem 24]. Indeed, if it were stably admissible, then the Kreisel–Putnam principle

$$\cdot(\neg\alpha \to \beta \vee \gamma) \longrightarrow (\neg\alpha \to \beta) \vee (\neg\alpha \to \gamma),$$

whose formula happens to be a special case of the premise of the Mints–Citkin rule, would imply the principle

$$\cdot\big((\neg\alpha \to \beta \vee \gamma) \to (\neg\alpha \to \beta)\big) \vee \big((\neg\alpha \to \beta \vee \gamma) \to (\neg\alpha \to \gamma)\big) \vee \big((\neg\alpha \to \beta \vee \gamma) \to \neg\alpha\big).$$

By a well-known theorem of Kreisel and Putnam, the extension of zero-order intuitionistic logic by the Kreisel–Putnam principle has the Disjunction Property (see [37; p. 3] for a short proof). Thus the Kreisel–Putnam principle would have to imply either $\cdot(\neg\alpha \to \beta \vee \gamma) \to (\neg\alpha \to \beta)$ or $\cdot(\neg\alpha \to \beta \vee \gamma) \to (\neg\alpha \to \gamma)$ or $\cdot(\neg\alpha \to \beta \vee \gamma) \to \neg\alpha$. However, none of these three principles are derivable even in classical logic, in which the Kreisel–Putnam principle is certainly derivable.

It follows from Proposition 5.9, (3)$\Rightarrow$(2), that the principle $\cdot(\neg\neg\alpha \to \alpha) \to \alpha \vee \neg\alpha$ implies $\cdot\neg\alpha \vee \neg\neg\alpha$. In contrast, the Lemmon–Scott rule

$$\frac{(\neg\neg\alpha \to \alpha) \to \alpha \vee \neg\alpha}{\neg\alpha \vee \neg\neg\alpha}$$

is not derivable. Indeed, let us consider the Tarski model with $X = \mathbb{R} \times (\mathbb{R}_+ \cup \mathbb{R}_-) \cup \{(0,0)\}$, where $\mathbb{R}_+ = \{x \in \mathbb{R} \mid x > 0\}$ and $\mathbb{R}_- = \{x \in \mathbb{R} \mid x < 0\}$ and a valuation with $|\alpha| = (\mathbb{R}_+ \cup \mathbb{R}_-) \times \mathbb{R}_+$. Then $|\neg\alpha| = \mathbb{R} \times \mathbb{R}_-$ and $|\neg\neg\alpha| = \mathbb{R} \times \mathbb{R}_+$, so that $|\neg\alpha \vee \neg\neg\alpha| = \mathbb{R} \times (\mathbb{R}_+ \cup \mathbb{R}_-) = X \setminus \{(0,0)\}$. On the other hand, $|\alpha \vee \neg\alpha| = \mathbb{R} \times \mathbb{R}_- \cup (\mathbb{R}_+ \times \mathbb{R}_-) \times \mathbb{R}_+ = |\neg\neg\alpha \to \alpha|$, so that $|(\neg\neg\alpha \to \alpha) \to \alpha \vee \neg\alpha| = X$. (See [172; 3.5.21] for a model in an Alexandroff space.)

Remark 5.10. Our example above works to show non-derivability of the special case of the Lemmon–Scott rule with α substituted by $\neg\beta \vee \neg\gamma$, by considering a valuation with $|\beta| = \mathbb{R}_+ \times \mathbb{R}_+ \cup \mathbb{R} \times \mathbb{R}_-$ and $|\gamma| = \mathbb{R}_- \times \mathbb{R}_+ \cup \mathbb{R} \times \mathbb{R}_-$. On the other hand, the Lemmon–Scott rule is clearly equivalent to the rule

$$\frac{(\neg\neg\alpha \to \alpha) \to \neg\alpha \vee \neg\neg\alpha}{\neg\alpha \vee \neg\neg\alpha}.$$

Thus our model shows non-derivability of Rose's formula $\big((\neg\neg\Phi \to \Phi) \to \neg\Phi \vee \neg\neg\Phi\big) \to \neg\Phi \vee \neg\neg\Phi$, where $\Phi = \neg\beta \vee \neg\gamma$. (See [149; §III.6.9] for a model in the real line.)

Remark 5.11. It is not essential in the previous examples that X is non-locally compact. The same effect can be achieved in a Tarski model in $Y = \{(x,y) \in \mathbb{R}^2 \mid |y| \geq |x|\}$, by considering a valuation with $|\alpha|, |\beta|, |\gamma|$ as above but intersected with Y.

However, the Lemmon-Scott rule is admissible; in fact, as observed in [99], it follows easily from a special case of the Mints–Citkin rule:

$$\frac{(\neg\neg\alpha \to \alpha) \to \alpha \vee \neg\alpha}{\big((\neg\neg\alpha \to \alpha) \to \alpha\big) \vee \big((\neg\neg\alpha \to \alpha) \to \neg\alpha\big) \vee \big((\neg\neg\alpha \to \alpha) \to \neg\neg\alpha\big)}.$$

Indeed, it is not hard to check that the second disjunct of the conclusion implies $\neg\alpha$, whereas the first and the third disjuncts each imply $\neg\neg\alpha$.

Let us now sketch an alternative argument that the Mints–Citkin rule is not stably admissible (a variation of [97; proof of Proposition 23]). If the Mints–Citkin rule were stably admissible, the Scott principle

$$\cdot\big((\neg\neg\alpha \to \alpha) \to \alpha \vee \neg\alpha\big) \longrightarrow \neg\alpha \vee \neg\neg\alpha,$$

whose formula happens to be a special case of the premise of the Mints–Citkin rule, would imply the principle

$$\cdot\big(\hat\alpha \to \neg\alpha\big) \vee \big(\hat\alpha \to \neg\neg\alpha\big) \vee \big(\hat\alpha \to (\neg\neg\alpha \to \alpha)\big),$$

where $\hat\alpha = (\neg\neg\alpha \to \alpha) \to \alpha \vee \neg\alpha$. By a well-known theorem of Scott, the extension of zero-order intuitionistic logic by the Scott principle has the Disjunction Property (see [37; pp. 4–5]). Thus the Scott principle would have to imply either $\cdot\big((\neg\neg\alpha \to \alpha) \to \alpha \vee \neg\alpha\big) \to \neg\alpha$ or $\cdot\big((\neg\neg\alpha \to \alpha) \to \alpha \vee \neg\alpha\big) \to \neg\neg\alpha$ or $\cdot\big((\neg\neg\alpha \to \alpha) \to \alpha \vee \neg\alpha\big) \to$

$(\neg\neg\alpha \to \alpha)$. However, the first two of the latter tree principles are not derivable even in classical logic, in which the Scott principle is certainly derivable; and the third implies $\cdot\,\beta \vee \neg\beta$ by substituting $\beta \vee \neg\beta$ for α. However, the Scott principle does not imply $\cdot\,\beta \vee \neg\beta$, since Jankov's principle $\cdot\,\neg\alpha \vee \neg\neg\alpha$ implies the Scott principle but does not imply $\cdot\,\beta \vee \neg\beta$ (see §**5.4**).

Remark 5.12. The rule
$$\frac{(\neg\neg\alpha \to \alpha) \to \alpha \vee \neg\alpha}{\alpha \vee \neg\alpha}$$
and its corresponding principle $\cdot\,\big((\neg\neg\alpha \to \alpha) \to \alpha \vee \neg\alpha\big) \to \alpha \vee \neg\alpha$ are equivalent to $\cdot\,\alpha \vee \neg\alpha$. Indeed, the latter principle clearly implies the former principle, whereas the rule is clearly equivalent to
$$\frac{(\neg\neg\alpha \to \alpha) \to \alpha \vee \neg\alpha}{\neg\neg\alpha \to \alpha},$$
which implies $\cdot\,\alpha \vee \neg\alpha$ by substituting $\beta \vee \neg\beta$ for α.

Moreover, the formula $\big((\neg\neg\alpha \to \alpha) \to \alpha \vee \neg\alpha\big) \to \alpha \vee \neg\alpha$ is equivalent to $\neg\neg\alpha \to \alpha$. Indeed, it is clearly equivalent to $\big((\neg\neg\alpha \to \alpha) \to \alpha \vee \neg\alpha\big) \to (\neg\neg\alpha \to \alpha)$, and it suffices to show that
$$\vdash \Big(\big((\neg\neg\alpha \to \alpha) \to \alpha \vee \neg\alpha\big) \to (\neg\neg\alpha \to \alpha)\Big) \;\longrightarrow\; (\neg\neg\alpha \to \alpha).$$
Indeed, we can rewrite this judgement, by the exponential law, as
$$\vdash \neg\neg\alpha \wedge \Big(\big((\neg\neg\alpha \to \alpha) \to \alpha \vee \neg\alpha\big) \to (\neg\neg\alpha \to \alpha)\Big) \;\longrightarrow\; \alpha. \qquad (*)$$
The left hand side of $(*)$ implies, using (27),
$$\Big(\neg\neg\alpha \to \big((\neg\neg\alpha \to \alpha) \to \alpha \vee \neg\alpha\big)\Big) \to (\neg\neg\alpha \to \alpha),$$
which by the exponential law rewrites as
$$\big((\neg\neg\alpha \to \alpha) \to (\neg\neg\alpha \to \alpha \vee \neg\alpha)\big) \to (\neg\neg\alpha \to \alpha),$$
which in turn is equivalent to $\neg\neg\alpha \to \alpha$. On the other hand, the left hand side of $(*)$ also implies $\neg\neg\alpha$; thus it implies $\neg\neg\alpha \wedge (\neg\neg\alpha \to \alpha)$, which in turn implies the right hand side of $(*)$.

Remark 5.13. See [19] concerning stable admissibility of rules in arithmetic.

5.5. Markov's principle

Similarly missing among intuitionistic laws is the following converse to (19), which is sometimes referred to as the *Generalized Markov Principle*:
$$\cdot\,\neg\forall x\,\pi(x) \;\longrightarrow\; \exists x\,\neg\pi(x). \tag{GMP}$$
This is of course the "generalization" of Jankov's principle above with $\exists$ and $\forall$ in place of $\vee$ and $\wedge$. On the other hand, (GMP) can also be "specialized" by replacing $\exists$ and $\forall$ with

$\wedge$ and $\rightarrow$, and the resulting principle $\cdot \vdash \neg(\chi \rightarrow \pi) \longrightarrow (\chi \wedge \neg\pi)$ is also non-derivable. Indeed, it specializes to $\cdot \neg\neg\chi \rightarrow \chi$ by substituting π with $\times$.

The stable case of (GMP),

$$\frac{\neg\neg\pi(x) \rightarrow \pi(x)}{\neg\forall x\, \pi(x) \longrightarrow \exists x \neg\pi(x)},$$

is clearly equivalent to the principle $\cdot \neg\forall x \neg\rho(x) \longrightarrow \exists x \neg\neg\rho(x)$, which is in turn equivalent, via (18), to the following *Strong Markov Principle*

$$\cdot \neg\neg\exists x\, \rho(x) \longrightarrow \exists x \neg\neg\rho(x) \tag{SMP}$$

whose converse is the intuitionistic law (47). Using (18) and (3), (SMP) can be equivalently rewritten as

$$\cdot \neg\neg\exists x \neg\neg\rho(x) \longrightarrow \exists x \neg\neg\rho(x),$$

which is in turn clearly equivalent to the rule

$$\frac{\neg\neg\rho(x) \rightarrow \rho(x)}{\neg\neg\exists x\, \rho(x) \rightarrow \exists x\, \rho(x)},$$

whose admissibility amounts to the assertion that if $\rho(x)$ is stable, then so is $\exists x\, \rho(x)$. Let us note that $\forall x\, \rho(x)$ is always stable if $\rho(x)$ is, due to the intuitionistic law (46).

It is easy to see that (SMP) holds in all Alexandroff spaces whose corresponding preorder is a directed set (i.e., every pair of elements has an upper bound). Nevertheless, Jankov's principle does not imply (SMP) by considering the Stone–Čech compactification of $\mathbb{N}$ (where the union of the regular open sets $\{n\}$, $n \in \mathbb{N}$, is not regular open). It is not hard to construct a simpler model of this kind that does not depend on the uncountable axiom of choice (see Proposition 5.24 below).

If $\exists$ is "specialized" to $\wedge$ in (SMP), the resulting principle $\cdot \neg\neg(\chi \wedge \rho) \longrightarrow (\chi \wedge \neg\neg\rho)$ is still non-derivable (indeed, it specializes to $\cdot \neg\neg\chi \rightarrow \chi$ by substituting ρ with $\checkmark$), but its stable case $\cdot \neg\neg(\neg\psi \wedge \neg\pi) \longrightarrow (\neg\psi \wedge \neg\pi)$ is derivable by (41) and (28). If $\exists$ is "specialized" to $\vee$ in (SMP), the resulting principle $\cdot \neg\neg(\rho_1 \vee \rho_2) \longrightarrow \neg\neg\rho_1 \vee \neg\neg\rho_2$ is still non-derivable, and so is its stable case (since the union of two regular open sets may well fail to be regular open). However, the decidable case

$$\frac{\rho_1 \vee \neg\rho_1, \ \ \rho_2 \vee \neg\rho_2}{\neg\neg(\rho_1 \vee \rho_2) \longrightarrow \neg\neg\rho_1 \vee \neg\neg\rho_2}$$

is derivable, since the premise implies $(\rho_1 \vee \rho_2) \vee \neg(\rho_1 \vee \rho_2)$ using (28), which in turn implies $\neg\neg(\rho_1 \vee \rho_2) \longrightarrow \rho_1 \vee \rho_2$ by (8).

The following *Strong Markov Rule* asserts that (SMP) holds for all decidable $\rho(x)$:

$$\frac{\rho(x) \vee \neg\rho(x)}{\neg\neg\exists x\, \rho(x) \longrightarrow \exists x\, \rho(x)}; \tag{SMR}$$

or equivalently that (GMP) holds for all decidable $\pi(x)$:

$$\frac{\pi(x) \vee \neg\pi(x)}{\neg\forall x\, \pi(x) \longrightarrow \exists x \neg\pi(x)}. \tag{SMR$'$}$$

To see that (SMR) is not derivable, let us consider the Tarski model with $\mathscr{D} = \mathbb{N}$ and $X = \mathbb{N}_+$, the one-point compactification of the countable discrete space $\mathbb{N}$, and a valuation with $|\rho|(n) = \{n\}$. Then each $|\rho|(n)$ is clopen, but their union is not regular open. Another useful Tarski model where (SMR) fails has $\mathscr{D} = \mathbb{N}$ and $X = 2^{\mathbb{N}}$, the Cantor set of all functions $\mathbb{N} \to \{0, 1\}$; each subset $|\rho|(n)$ of $2^{\mathbb{N}}$ consisting of all functions f such that $f(n) = 1$ is clopen, but the union of these subsets, $2^{\mathbb{N}} \setminus \{(0, 0, \dots)\}$, is not regular open.

These examples actually show that even the following *Markov Rule* is not derivable:

$$\frac{\rho(\mathsf{x}) \vee \neg\rho(\mathsf{x}), \quad \neg\neg\exists\mathsf{x}\,\rho(\mathsf{x})}{\exists\mathsf{x}\,\rho(\mathsf{x})}.$$

But the latter is strictly weaker than the Strong Markov Rule. Indeed, let us consider a Tarski model with $\mathscr{D} = \mathbb{N}$ and $X = \mathbb{N}_+ \times [0, 1] \cup \{\infty\} \times (-1, 0]$, and a valuation with $|\rho|(n) = \{n\} \times [0, 1]$. Again each $|\rho|(n)$ is clopen, and their union is not regular open. On the other hand, if $\mathrm{Cl}(\bigcup_{i=1}^{\infty} U_i) = X$, where each U_i is clopen, then at least one U_i contains a neighborhood of $\{\infty\} \times [-1, 1]$, and it follows that in fact $\bigcup U_i = X$.

Theorem 5.14. *The Markov rule is admissible in intuitionistic logic.*

Proof. Suppose that for some 1-formula Φ we have $\vdash \Phi(\mathsf{x}) \vee \neg\Phi(\mathsf{x})$ and $\vdash \neg\neg\exists\mathsf{x}\,\Phi(\mathsf{x})$, but $\nvdash \exists\mathsf{x}\,\Phi(\mathsf{x})$. Then by the completeness theorem, there exists a Tarski model in some space X with some domain $\mathscr{D}$ such that some point $p_X \in X$ is not contained in $|\exists\mathsf{x}\,\Phi(\mathsf{x})|_{u_0}^{v_X}$ for some valuation v_X and some assignment u_0. (Of course, the latter is needed only if $\exists\mathsf{x}\,\Phi(\mathsf{x})$ is not a closed formula.)

Let Y be the Alexandroff space corresponding to the three element poset consisting of a pair of incomparable elements p_Y, q and of their lower bound r. Let Z be the gluing $X \cup_{p_X = p_Y} Y$, that is, the quotient space of the disjoint union $X \sqcup Y$ by the equivalence relation whose only non-singleton class, to be denoted by p, consists of p_X and p_Y. We will identify X and Y with their images in Z, and in particular from now on we will write p in place of both p_X and p_Y. Let v_Z be the valuation for the Tarski model in Z with domain $\mathscr{D}$ defined by $|\alpha|_u^{v_Z} = |\alpha|_u^{v_X}$ for every problem variable α and every assignment u. Since X is open in Z, it is easy to check by induction that $|\Xi|_u^{v_X} = |\Xi|_u^{v_Z} \cap X$ for every formula Ξ and every assignment u. In particular, $|\Phi|_{u_0}^{v_X}(d) = |\Phi|_{u_0}^{v_Z}(d) \cap X$ for each $d \in \mathscr{D}$.

Then $p \notin |\Phi|_{u_0}^{v_Z}(d)$ for each $d \in \mathscr{D}$. Since Y is connected and each $|\Phi|_{u_0}^{v_Z}(d)$ is clopen (due to our hypothesis that $\Phi(\mathsf{x})$ is decidable), we also have $q \notin |\Phi|_{u_0}^{v_Z}(d)$ for each $d \in \mathscr{D}$. Hence $q \notin |\exists\mathsf{x}\,\Phi(\mathsf{x})|_{u_0}^{v_Z}$. Since $Z \setminus \{q\}$ is closed, q also does not belong to $|\neg\neg\exists\mathsf{x}\,\Phi(\mathsf{x})|_{u_0}^{v_Z} = \mathrm{Int}\,\mathrm{Cl}\,|\exists\mathsf{x}\,\Phi(\mathsf{x})|_{u_0}^{v_Z}$. But this contradicts our hypothesis that $\vdash \neg\neg\exists\mathsf{x}\,\Phi(\mathsf{x})$. $\square$

The following is sometimes called the *Markov Principle*:

$$\cdot\forall\mathsf{x}\,\big(\rho(\mathsf{x}) \vee \neg\rho(\mathsf{x})\big) \;\longrightarrow\; \big(\neg\neg\exists\mathsf{x}\,\rho(\mathsf{x}) \to \exists\mathsf{x}\,\rho(\mathsf{x})\big).$$

It is strictly stronger than the Markov Rule due to the admissibility of the latter; moreover, it is strictly stronger than the Strong Markov Rule. Indeed, let us consider a

Tarski model with $\mathscr{D} = \mathbb{N}$ and $X = \mathbb{N}_+ \times [0,1] \cup [0,\infty] \times \{0\}$ and a valuation with $|\rho|(n) = U_n := \{n\} \times (0,1]$. Then $\bigcup_{i=1}^{\infty} U_i = \mathbb{N} \times (0,1]$ and $\mathrm{Int}\,\mathrm{Cl}(\bigcup_{i=1}^{\infty} U_i) = \mathbb{N}_+ \times (0,1]$, and consequently $|\neg\neg\exists x\,\rho(x) \to \exists x\,\rho(x)| = X \setminus (\{\infty\} \times [0,1])$. On the other hand, $U_i \cup \mathrm{Int}(X \setminus U_i) = X \setminus \{(n,0)\}$, and therefore $|\forall x\,(\rho(x) \vee \neg\rho(x))| = X \setminus (\mathbb{N}_+ \times \{0\})$, which is not a subset of $X \setminus (\{\infty\} \times [0,1])$. Nevertheless, the model satisfies the Strong Markov Rule, since the only clopen subsets of X are $\varnothing$ and X.

Remark 5.15. Markov's original formulation of his principle was concerned with the case where x ranges over $\mathbb{N}$.[79] A decidable problem $\rho(n)$ is equivalent to the problem of verifying that $\rho_n = 1$, where each $\rho_n \in \{0,1\}$ is defined by $\rho_n = 1$ if $\rho(n)$ is soluble, and $\rho_n = 0$ if $\neg\rho(n)$ is soluble. Under the BHK interpretation (assuming that ρ is decidable), the problem $\exists n\,\rho(n)$ asks essentially to find a one in the sequence (ρ_i), and the problem $\neg\neg\exists n\,\rho(n)$, or equivalently $\neg\forall n\,\neg\rho(n)$, asks essentially to prove that there exists no proof that each entry in the sequence is zero. Markov's Principle is then asserting that there exists a general method that, given a sequence of zeroes and ones, and a proof that this sequence cannot be proved to be all 0's, produces a 1 somewhere in this sequence. Of course, we are not really told how to find this 1 by constructive means, so this principle is not justified by the clarified BHK interpretation.

However, on a very computational reading of the BHK interpretation, Markov's principle is validated. Indeed, suppose we have a Turing machine that reads a sequence of zeroes and ones until it finds a one. If this machine continues forever on some sequence, this can (perhaps) be equated with a proof that the sequence is all zeroes. Now, given a proof that the machine does not continue forever on some sequence, by actually running it on this sequence we get a general method of finding a one in the sequence. In this connection Markov's constructivist school considered Markov's Principle to be constructively acceptable (cf. [202]).

5.5.1. Principle of Omniscience. The following *Rule of Omniscience*:[80]

$$\frac{\pi(x) \vee \neg\pi(x)}{\exists x\,\neg\pi(x) \ \vee \ \forall x\,\pi(x)} \tag{RO}$$

implies (SMR$'$) by (7). Using (8) and (18), one can equivalently reformulate (RO) as

$$\frac{\rho(x) \vee \neg\rho(x)}{\exists x\,\rho(x) \ \vee \ \neg\exists x\,\rho(x)}.$$

(To obtain this from (RO), substitute $\neg\rho(x)$ for $\pi(x)$; and for the converse, substitute $\neg\pi(x)$ for $\rho(x)$.)

From the latter form of (RO) it is clear that its admissibility amounts to the assertion that if $\rho(x)$ is decidable, then so is $\exists x\,\rho(x)$. Also, (RO) is valid in those spaces where arbitrary union of clopen sets is clopen. These include all Alexandroff spaces and all

---

[79]In some form, Markov's Principle was discussed long before A. A. Markov, Jr. (who himself called it the "Leningrad Principle"), in particular by P. S. Novikov [150] and, according to Hilbert [88; p. 268], already by Kronecker.

[80]This is a form of Bishop's "Principle of Omniscience" [22].

connected spaces. In particular, Tarski models in Euclidean spaces satisfy (RO), and so intuitionistic logic is not strongly complete with respect to this class of models. On the other hand, (RO) is not valid in a Tarski model in the subset $[-1, 0] \cup \{\frac{1}{n} \mid n \in \mathbb{N}\}$ of the real line, where (SMR) is valid; indeed, in this space, a union of clopen sets is regular open, but generally not clopen.

The symmetric assertion, that $\forall x\, \rho(x)$ is decidable whenever $\rho(x)$ is, amounts to the admissibility of what can be called the *Weak Rule of Omniscience*:

$$\frac{\pi(x) \vee \neg\pi(x)}{\forall x\, \pi(x) \ \vee \ \neg\forall x\, \pi(x)}, \tag{WRO}$$

because it follows from (RO) by (19). The converse implication fails by considering the Tarski model in $\mathbb{N}_+$, where (RO) is not valid (indeed, already (SMR) is not valid), but (WRO) is valid, since the only open set of $\mathbb{N}_+$ that is not clopen is $\mathbb{N}$, but it is neither the interior of any larger set nor the intersection of any family of clopen sets. On the other hand, (WRO) together with (SMR) obviously entail (RO). Thus in fact (RO) is equivalent to the meta-conjunction of (WRO) and (SMR):

$$\vdash \quad \text{(RO)} \quad \Longleftrightarrow \quad \text{(WRO)} \ \& \ \text{(SMR)}.$$

Remark 5.16. When x ranges over $\mathbb{N}$, the three principles can be recast under the BHK interpretation in terms of the Cantor set $2^{\mathbb{N}}$ of all functions $\mathbb{N} \to \{0, 1\}$. If ∞ denotes the constant function $n \mapsto 1$, and p is understood to range over $2^{\mathbb{N}}$, then the three principles assert the existence of general methods for solving the following problems:

- Weak Rule of Omniscience: $p = \infty \vee \neg(p = \infty)$;
- Strong Markov Rule: $\neg(p = \infty) \to p \neq \infty$;
- Rule of Omniscience: $p = \infty \vee p \neq \infty$,

where $p = q$ denotes the problem, *Prove that $p(n) = q(n)$ for all $n \in \mathbb{N}$*, and $p \neq q$ denotes the problem, *Find an $n \in \mathbb{N}$ such that $p(n) \neq q(n)$*. As observed by Escardo [52], there is no loss of generality in permitting p to range just over the subset $\mathbb{N}_+ \subset 2^{\mathbb{N}}$ consisting of all non-increasing functions. (Note that $\mathbb{N}_+ = N \cup \{\infty\}$, where N is a copy of $\mathbb{N}$.) This follows from the existence of a constructive retraction r of $2^{\mathbb{N}}$ onto $\mathbb{N}_+$, defined by $r(p)(n) = \min\{p(k) \mid k \leq n\}$.

5.6. ¬¬-Translation

A translation of classical predicate logic QC into intuitionistic predicate logic QH was discovered essentially by Kolmogorov (even though he only had his formalization of a fragment of QH available at the moment) [107] (see also [34], [204], [159]).

Theorem 5.17. *Given a formula F in the language of classical logic, insert the double negation $\neg\neg$ in front of every subformula, and replace $\top$ and $\bot$ by $\checkmark$ and $\times$, respectively. Then the resulting formula $F_{\neg\neg}$ is derivable in intuitionistic logic if and only if F is derivable in classical logic.*

Moreover, a rule $F_1, \ldots, F_k \,/\, G$ is derivable in classical logic if and only if the rule $(F_1)_{\neg\neg}, \ldots, (F_k)_{\neg\neg} \,/\, G_{\neg\neg}$ is derivable in intuitionistic logic.

Kolmogorov's translation was later rediscovered by Gödel and Gentzen, who gave proofs in the setup of QH and also observed that prefixing $\wedge$, $\forall$ and (in the case of Gentzen) $\to$ by $\neg\neg$ is superfluous:

Proposition 5.18. *The following holds in intuitionistic logic:*
(a) $\vdash \neg\neg\alpha \wedge \neg\neg\beta \;\longleftrightarrow\; \neg\neg(\neg\neg\alpha \wedge \neg\neg\beta)$*;*
(b) $\vdash (\neg\neg\alpha \to \neg\neg\beta) \;\longleftrightarrow\; \neg\neg(\neg\neg\alpha \to \neg\neg\beta)$*;*
(c) $\vdash \forall x\,\neg\neg\alpha(x) \;\longleftrightarrow\; \neg\neg\forall x\,\neg\neg\alpha(x)$*.*

Proof. By (28) and (30), if γ and δ are stable, then so are $\gamma \wedge \delta$ and $\gamma \to \delta$. Also, by (18), if $\gamma(x)$ is stable, then so is $\forall x\,\gamma(x)$. But by (3), $\vdash \neg\neg\epsilon \leftrightarrow \epsilon$ holds for stable ϵ. $\qquad\square$

Proof of Theorem 5.17. The "only if" assertion is trivial: any derivation of $F_{\neg\neg}$ in intuitionistic logic is also valid classically, but clearly $F_{\neg\neg}$ is classically equivalent to F. Similarly for rules.

For the converse, we may consider instead of Kolmogorov's interpretation $F \mapsto F_{\neg\neg}$ its (Gentzen's) simplified version described above, to be denoted $F \mapsto F'$. Then it suffices to check that if $\cdot\, F$ is a law of classical logic, then $\cdot\, F'$ is derivable in intuitionistic logic; and if $F_1, \ldots, F_k \,/\, G$ is an inference rule of classical logic, then $F_1', \ldots, F_k' \,/\, G'$ is derivable in intuitionistic logic.

For principles and rules that do not involve $\vee$ or $\exists$, this is trivial, by 5.18. From the law $\cdot\, \alpha(x) \to \exists x\,\alpha(x)$ we get that $\cdot\, \neg\neg\alpha(x) \to \neg\neg\exists x\,\neg\neg\alpha(x)$ is derivable, using that $\vdash \beta \to \neg\neg\beta$. From the inference rule $\alpha(x) \to \gamma \,/\, \exists x\,\alpha(x) \to \gamma$ we get that $\neg\neg\alpha(x) \to \neg\neg\gamma \,/\, \neg\neg\exists x\,\neg\neg\alpha(x) \to \neg\neg\gamma$ is derivable, using that $\vdash \neg\neg(\neg\neg\gamma) \to \neg\neg\gamma$. The analogous law and inference rule involving $\vee$ are treated similarly. Finally, $\cdot\, \neg\neg(\neg\neg\alpha \vee \neg(\neg\neg\alpha))$ is derivable, using $\vdash \neg\neg(\gamma \vee \neg\gamma)$. $\qquad\square$

As for $\vee$ and $\exists$, let us note that one can eliminate them altogether using the intuitionistic de Morgan laws (28), (18):

$$\vdash \neg\neg(\alpha \vee \beta) \;\longleftrightarrow\; \neg(\neg\alpha \wedge \neg\beta),$$

$$\vdash \neg\neg\exists x\,\alpha(x) \;\longleftrightarrow\; \neg\forall x\,\neg\alpha(x).$$

To summarize, the $\neg\neg$-translation of classical logic into intuitionistic can be presented as follows:

- predicate variables are interpreted as problem variables prefixed by $\neg\neg$;
- $\top$ and $\bot$ are replaced by $\checkmark$ and $\times$, respectively;
- classical $\wedge$, $\forall$ and $\to$ (and hence $\neg$) are interpreted as the intuitionistic ones;
- classical $\vee$ and $\exists$ are interpreted as the $\neg$-conjugates of intuitionistic $\wedge$ and $\forall$.

Topologically, we have interpreted classical logic as the logic of regular open sets, where $\wedge$, $\forall$ and $\to$ (and hence $\neg$) are interpreted as in Tarski models, whereas $\vee$ and $\exists$ are interpreted via the Tarski interpretations of $\wedge$ and $\forall$.

In embedding classical logic into intuitionistic in a type theoretic context it may be essential to avoid the prefixing of atomic subformulas (see [29]). This is achieved by the Kuroda embedding, which places $\neg\neg$ before the entire formula and also *after* each universal quantifier. It is still equivalent to Kolmogorov's embedding:[81]

Proposition 5.19. *The following holds in intuitionistic logic:*

$(a) \vdash \neg\neg(\alpha \wedge \beta) \longleftrightarrow \neg\neg(\neg\neg\alpha \wedge \neg\neg\beta);$

$(b) \vdash \neg\neg(\alpha \vee \beta) \longleftrightarrow \neg\neg(\neg\neg\alpha \vee \neg\neg\beta);$

$(c) \vdash \neg\neg(\alpha \rightarrow \beta) \longleftrightarrow \neg\neg(\neg\neg\alpha \rightarrow \neg\neg\beta);$

$(d) \vdash \neg\neg\exists x\, \alpha(x) \longleftrightarrow \neg\neg\exists x\, \neg\neg\alpha(x).$

Proof. Using respectively (41), (28), (38) and (18), the left hand sides can be equivalently rewritten in a form where each atomic subformula is prefixed by a $\neg$. In this form, each atomic subformula can be replaced by its double negation, due to (3). $\square$

Of course, instead of placing $\neg\neg$ in front of the entire formula it suffices (in view of 5.18) to put it before every $\exists$ and $\vee$. This yields an "essentially local" version of the translation (similar to Dowek's version [45]):

- prefix every classical $\exists$ and $\vee$ by a $\neg\neg$;
- postfix every classical $\forall$ and $\wedge$ by a $\neg\neg$;
- if the entire formula is a single atom, prefix it by a $\neg\neg$;
- replace $\top$, $\bot$ by $\checkmark$, $\times$.

(Here a classical $\wedge$ really needs to be postfixed by a $\neg\neg$ only when there are no other types of connectives "around".)

5.6.1. $\neg\neg$-*Shift principle*. Prefixing only the entire formula with $\neg\neg$ (and replacing $\top$, $\bot$ by $\checkmark$, $\times$), which is known as Glivenko's transformation, does not work in the presence of universal quantifiers. Indeed, let F denote the formula $\forall x\, \pi(x) \vee \exists x\, \neg\pi(x)$, which we encountered in the Rule of Omniscience and which is derivable in classical logic. If Glivenko's transformation were taking classically derivable formulas to intuitionistically derivable formulas, then $\cdot\neg\neg F$, that is,

$$\cdot\neg\neg\big(\forall x\, \pi(x) \vee \exists x\, \neg\pi(x)\big),$$

would be an intuitionistically derivable principle. The latter principle is equivalent, via (28) and (30), to

$$\cdot\neg\exists x\, \neg\pi(x) \longrightarrow \neg\neg\forall x\, \pi(x),$$

which is nothing but the contrapositive of (GMP), Generalized Markov's Principle. By (18), the latter principle is in turn equivalent to the following $\neg\neg$-*Shift Principle*:

$$\cdot\forall x\, \neg\neg\pi(x) \longrightarrow \neg\neg\forall x\, \pi(x). \tag{DNS}$$

Note that the converse to the $\neg\neg$-Shift Principle is nothing but the intuitionistically derivable principle (46). If $\forall$ is "specialized" to $\wedge$, the $\neg\neg$-Shift Principle itself becomes

---

[81]Other variants of the $\neg\neg$-translation achieve still more economy in the use of $\neg$ (see [55]).

intuitionistically derivable (43); and if $\forall$ is "specialized" to $\rightarrow$, the resulting principle $\cdot (\chi \rightarrow \neg\neg\pi) \longrightarrow \neg\neg(\chi \rightarrow \pi)$ is again intuitionistically derivable (by (44) and (35)).

To see that the $\neg\neg$-Shift Principle is not derivable, let us consider the Tarski model with $X = \mathbb{R}$ and $\mathscr{D} = \mathbb{Q}$, and a valuation with $|\pi|(q) = \mathbb{R} \setminus \{q\}$. Then $|\forall x \, \neg\neg\pi(x)| = \mathbb{R}$, whereas $|\neg\neg\forall x \, \pi(x)| = \varnothing$ (indeed, $|\forall x \, \pi(x)| = \varnothing$ since $\mathbb{R} \setminus \mathbb{Q}$ has empty interior in $\mathbb{R}$). Another Tarski model where the $\neg\neg$-Shift Principle is not valid can be obtained by observing that the poset of integers (ordered in the usual way) with its Alexandroff topology contains only countably many open sets, whose intersection is empty, and no regular open sets other than $\varnothing$ and the whole space.

By 5.18, the $\neg\neg$-Shift Principle can be equivalently rewritten as

$$\cdot \neg\neg\forall x \, \neg\neg\pi(x) \longrightarrow \neg\neg\forall x \, \pi(x). \tag{DNS$'$}$$

It follows that the $\neg\neg$-Shift Principle holds for stable $\pi(x)$. In other words, the contrapositive of (SMP), the Strong Markov Principle, is derivable in intuitionistic logic.

Yet another equivalent formulation of the $\neg\neg$-Shift Principle was discussed in §**3.12**:

$$\cdot \neg\neg\forall x \, (\pi(x) \vee \neg\pi(x)). \tag{DNS$''$}$$

To see that (DNS) implies (DNS$''$) it suffices to note that by (32), $\vdash \forall x \, \neg\neg(\pi(x) \vee \neg\pi(x))$. Conversely, (DNS$''$) implies (DNS$'$) using that $\vdash \neg\neg\alpha \wedge (\alpha \vee \neg\alpha) \longrightarrow \alpha$.

Let us note that the $\neg\neg$-Shift Principle is valid, for instance, in the Tarski model in the one-point compactification $\mathbb{N}_+$ of the countable discrete space $\mathbb{N}$ with any $\mathscr{D}$. Indeed, $|\alpha \vee \neg\alpha|$ must be dense in $\mathbb{N}_+$ due to $\vdash \neg\neg(\alpha \vee \neg\alpha)$ (see (32)). Hence $|\alpha \vee \neg\alpha|$ contains $\mathbb{N}$. Therefore $|\forall x \, (\pi(x) \vee \neg\pi(x))|$ also contains $\mathbb{N}$, and thus $|\neg\neg\forall x \, (\pi(x) \vee \neg\pi(x))| = \mathbb{N}_+$.

5.7. Provability translation

5.7.1. Modal logic QS4. Suppose that classical connectives and quantifiers are interpreted by set-theoretic operations on subsets of a set X (as in §*4.8.2*), and that X also happens to be a topological space. As long as the Leibniz–Euler model is closed under the topological interior operator Int, this operator can be regarded as interpreting some additional unary connective $\square$. In a Leibniz–Euler model, equivalent formulas are always interpreted by the same subset; thus such a $\square$ must certainly be well-defined on equivalence classes of formulas, in the sense that if $F \leftrightarrow G$ is derivable in classical logic, then $\square F \leftrightarrow \square G$ is to be postulated. This condition along with the usual axioms of a topological space X in terms of interior,

- Int $P \subset P$;
- Int $P \subset \text{Int}(\text{Int } P)$;
- $X \subset \text{Int } X$;
- Int $P \cap \text{Int } Q = \text{Int}(P \cap Q)$,

correspond to the following properties of $\Box$, which suggest the intended reading of $\Box F$ as "there exists a proof of F".[82]

- (reflection) $\cdot \Box p \to p$;
- (proof checking) $\cdot \Box p \to \Box\Box p$;
- (modus ponens) $\cdot \Box p \land \Box(p \to q) \longrightarrow \Box q$.

- (necessitation) $\dfrac{p}{\Box p}$;

Indeed, the modus ponens principle follows from:

(i) $\cdot \Box p \land \Box q \longleftrightarrow \Box(p \land q)$;

(ii) $\dfrac{p \leftrightarrow q}{\Box p \leftrightarrow \Box q}$.

using that in classical logic, $\vdash p \land (p \to q) \longleftrightarrow p \land q$. Conversely, the modus ponens principle implies

(iii) $\cdot \Box(p \to q) \longrightarrow (\Box p \to \Box q)$

by the exponential law (regarded as a law of classical logic). Now (iii) and necessitation imply (ii) and (i). In more detail, the $\leftarrow$ implication in (i) is proved using that in classical logic, $\vdash p \land q \to p$ and $\vdash p \land q \to q$; and the $\to$ implication in (i), transformed by the exponential law, is proved using that in classical logic, $\vdash p \to (q \to (p \land q))$.

The four properties of $\Box$ (reflection, proof checking, necessitation and modus ponens), regarded as three laws and one inference rule, are supplemented by the usual laws and inference rules of classical logic (see §**4.6**), with understanding that they now apply to all formulas, possibly containing the new connective $\Box$. The resulting logic is known as QS4, and its zero-order version as S4. The connective $\Box$ is called a modality and is assigned the same level of precedence as $\neg$.

By design, a Leibniz–Euler model of classical logic in subsets of a topological space extends to a *topological model* of QS4 by interpreting $\Box$ by the interior operator. It is well-known that QS4 is complete with respect to its topological models [168] with countable domain (see also [115]).

Let us note that the laws of QS4 can be equivalently restated in terms of $\Diamond := \neg\Box\neg$:

(1) $\cdot p \to \Diamond p$;
(2) $\cdot \Diamond\Diamond p \to \Diamond p$;
(3) $\cdot (\Diamond p \to \Diamond q) \to \Diamond(p \to q)$.

However, the rule $\Diamond p \,/\, p$, which is the "contrapositive" of the necessitation rule $p \,/\, \Box p$, is not even an admissible rule of QS4. Indeed, $q \to \Box q$ is not derivable in QS4, but $\Diamond(q \to \Box q)$ is (see §*5.7.2* below). Alternatively, since $\alpha \lor \neg\alpha$ is not derivable in intuitionistic

---

[82]It is remarkable that Orlov, who first introduced these axioms in 1928, did so in order to give a provability explanation of propositional intuitionistic logic [154; §§6,7] (see [1; §*6.2.1*]). His work remained virtually unknown, while the same axioms were rediscovered in a few years by Lewis, Becker and Gödel (see [69], [135; Theorem 2.1]).

logic, whereas $\cdot\neg\neg(\alpha \vee \neg\alpha)$ is, it follows (see §5.7.3 below) that $\Box q \vee \Box\neg\Box q$ is not derivable in QS4, whereas $\neg\Box\neg(\Box q \vee \Box\neg\Box q)$ is.

Example 5.20. An example of an admissible non-derivable rule of QS4 is

$$\frac{\neg\Box p \wedge \neg\Box\neg p}{\bot}.$$

Indeed, in a topological model, $|\neg\Box p \wedge \neg\Box\neg p| = \mathrm{Cl}\,|p| \setminus \mathrm{Int}\,|p|$, that is, the topological frontier $\mathrm{Fr}\,|p|$. Since $\mathrm{Fr}\,|p|$ may well be the entire space (for instance, the frontier of $\mathbb{Q}$ as a subset of $\mathbb{R}$ is $\mathbb{R}$), the rule is not derivable.

On the other hand, if F is a formula, we claim that $\nvdash \neg\Box F \wedge \neg\Box\neg F$. Indeed, in a topological model with a valuation and an assignment such that all predicate variables are interpreted by the entire space, $|F|$ will be either empty or the entire space; in either case, $|\neg\Box F \wedge \neg\Box\neg F| = \mathrm{Fr}(|F|) = \varnothing$.

5.7.2. QS4 and formal provability. In what sense does the modality $\Box$ of S4 and QS4 correpond to provability? Clearly, provability in the sense of the "classical BHK" (see §3.9) satisfies the laws and inference rules of QS4 — albeit in a trivial way, with $\Box F \leftrightarrow F$ being valid.

However, $\Box$ fails to directly represent formal provability in any theory containing Peano Arithmetic. Indeed, reflection and necessitation imply $\vdash \Box(\Box\bot \to \bot)$; if $\Box$ is to be interpreted as formal provability, then this would be saying that the consistency of the theory can be proved within the theory, contradicting Gödel's second incompleteness theorem (see [185]). Thus Gödel spoke of derivability "understood not in a particular system, but in the absolute sense (that is, one can make it evident)" and suggested an intended reading of $\Box$ as "is provable in the absolute sense" [68]. This, of course, raises the question of what, if anything, $\Box$ ever has to do with actual proof theory.

In fact, it is more natural to compare $\Box$ not with Gödel's predicate of formal provability $\mathrm{Bew}(\#F) := \exists m\,\mathrm{Prov}(m, \#F)$, where $\mathrm{Prov}(m, n)$ is a formula expressing the predicate that the mth proof is a proof of the nth formula in the Gödel numbering, but rather with its strengthened variants such as Mostowski's

$$M(\#F) := \exists m\big(\mathrm{Prov}(m, \#F) \wedge \neg\,\mathrm{Prov}(m, \#\,0{=}1)\big),$$

which is equivalent to $\mathrm{Bew}(\#F)$ but the equivalence is not a theorem of Peano Arithmetic. Indeed, the consistency formula for Mostowski's predicate, $\neg M(\#\,0 = 1)$, is trivially a theorem of Peano Arithmetic — just like $\neg\Box\bot$ is trivially a theorem of QS4.

However, neither $M(\#F)$ nor any other arithmetic provability predicate $B(\#F)$ expressible in Peano Arithmetic can interpret the modality $\Box$ of QS4. Indeed, by Löb's theorem (see [185; 4.1.1], [101]), any arithmetic provability predicate of the form $\boxdot F = B(\#F)$ satisfying the (schematic) principles $\boxdot(F \to G) \to (\boxdot F \to \boxdot G)$ and $\boxdot F \to \boxdot\boxdot F$ and the (schematic) rule $F\,/\,\boxdot F$ also satisfies Löb's principle: $\boxdot(\boxdot F \to F) \to \boxdot F$. On the other hand, the contrapositive of Löb's principle for $F = \bot$: $\neg\boxdot\bot \to \neg\boxdot(\neg\boxdot\bot)$ is a modal version of Gödel's second incompleteness theorem, which is inconsistent with the theorem $\neg\Box\bot$ of QS4. (It should be noted that Löb's theorem implies also the

original second incompleteness theorem, since $\boxdot F = \mathrm{Bew}(\#F)$ does satisfy the three principles in the hypothesis of Löb's theorem.)

The next natural step in the quest for a proof-theoretic interpretation of the modality $\square$ of QS4 would be to get rid of the constraint imposed by working within a fixed first-order theory (such as the Peano Arithmetic). If we extend a theory T that admits a Gödel numbering by the reflection principle $R_T := \mathrm{Bew}_T(\#F) \to F$, where Bew_T denotes Gödel's provability predicate for T, then the formula expressing the consistency of T, $\mathrm{Consis}(T) := \mathrm{Bew}_T(\#\,0\!=\!1) \to 0\!=\!1$ follows, and so does $\mathrm{Consis}\big(T+\mathrm{Consis}(T)\big) = \mathrm{Bew}_T\big(\#\,\mathrm{Bew}_T(\#\,0\!=\!1)\big) \to 0\!=\!1$; and so on. If we now write $\boxdot F = \mathrm{Bew}_T(\#F)$, then the principles $\boxdot F \to F$, $\boxdot F \to \boxdot\,\boxdot F$ and $\boxdot(F \to G) \to (\boxdot F \to \boxdot G)$ are derivable in $T + R_T$. However, the rule $F\,/\,\boxdot F$ is not even admissible for $T + R_T$: while $\neg\,\boxdot\,\bot$ is provable (due to R_T), $\boxdot(\neg\,\boxdot\,\bot)$ is not (since $\neg\,\boxdot\,(\neg\,\boxdot\,\bot)$ is provable, due to Löb's theorem for T).

An analysis of the latter failure suggests that maybe instead of finding a single proof-theoretic interpretation of $\square$, different occurrences of $\square$ could be interpreted differently, according to their "depth" in a formula. This effectively means that we are lifting a formula of QS4 to a formula of a *polymodal* logic, which has infinitely many modalities: $[0], [1], [2], \ldots$, by labelling every occurrence of $\square$ with some number; and then seeking an arithmetical interpretation of this polymodal logic, with each $[n]$ interpreted by formal provability in a theory T_n, where each T_{i+1} contains at least T_i and its reflection principle R_{T_i}. Because of Löb's theorem, we do not want to label an occurrence of $\square$ as $[n]$ if its scope contains an occurrence of $\square$ already labelled as $[k]$ for some $k \geq n$.

But it turns out that such a lift is impossible. Indeed, let us consider the formula $\Diamond(\mathsf{p} \to \square\mathsf{p})$, which can be equivalently presented as $\neg\square(\mathsf{p} \wedge \neg\square\mathsf{p})$. To see that it is derivable in QS4, observe that $\square(\mathsf{p} \wedge \neg\square\mathsf{p})$ implies, firstly, $\square\mathsf{p}$ and, secondly, $\square\neg\square\mathsf{p}$, which in turns implies $\neg\square\mathsf{p}$. Thus $\square(\mathsf{p} \wedge \neg\square\mathsf{p})$ implies $\bot$, as desired. As observed in [118], this argument has "a flavor of self-referentiality"; indeed, if we try to repeat it with labels, we get that $[1](\mathsf{p} \wedge \neg[0]\mathsf{p})$ implies, firstly, $[1]\mathsf{p}$ and, secondly, $[1]\neg[0]\mathsf{p}$, which in turn implies $\neg[0]\mathsf{p}$. Yet there is nothing contradictory about this, as $[0]\mathsf{p}$ implies $[1]\mathsf{p}$ but not conversely.

A different approach was suggested informally by Gödel, but remained unpublished until 1995; a closely related, but somewhat different approach was suggested and formalized independently by Artëmov (see [101; §10]). Their idea is to lift a formula of S4 to a polymodal logic whose modalities are indexed by "proof terms", which are to be interpreted by actual proofs in the arithmetical model. Artëmov's approach takes advantage of the fact that Peano Arithmetic proves $\mathrm{Prov}(n, \#F)$ for some numeral n if and only if it proves F (see [185; 3.2.4]). Thus it can be said that on Artëmov's interpretation, the modality $\square$ of S4 represents "existence of proofs" in Peano Arithmetic (and not just in its extensions by reflection principles), where "proofs" have the usual meaning of formal proofs (except that Artëmov needs one "proof" to be able to prove several formulas), but "existence" is understood in an explicit sense, not expressible internally in Peano Arithmetic.

Another well-known approach is that $\Box F$ can be represented in the form $F \wedge \boxdot F$, where it is $\boxdot$ that should be interpreted as an arithmetic provability predicate. (Topologically, $\neg \boxdot \neg$ corresponds to Cantor's derivative operation; see [50] for further details.) In particular, by well-known results of (i) Solovay and (ii) Kuznetsov–Muravitsky, Goldblatt and Boolos (see [198], [50], [101]), $\Box F := F \wedge \mathrm{Bew}(\#F)$ satisfies the laws and inference rules of QS4, and besides those only one additional quantifier-free principle, namely the following principle Grz: $\Box(\Box(F \to \Box F) \to F) \to F$. This does not contradict Löb's theorem because $\mathrm{Bew}(\#F) \wedge F$ is not expressible in the form $B(\#F)$.

5.7.3. Provability translation. Tarski models motivate the following embedding of intuitionistic logic into QS4: interpret every problem variable as a predicate variable prefixed by a box (because each problem variable corresponds to an open set in a Tarski model with chosen valuation and assignment), replace $\times$ and $\checkmark$ by $\bot$ and $\top$ respectively, and insert a box in front of every implication and in front of every universal quantifier (because of their special treatment in Tarski models). This is a version of Gödel's translation (see below) due to McKinsey–Tarski, as extended to the first-order case by Rasiowa–Sikorski, Maehara and Prawitz–Malmnäs (see [198], [149], [56]).

Thus every formula Φ in the language of intuitionistic logic corresponds to a formula Φ' in the language of QS4 (namely, the formula obtained from Φ by the modification just described) so that the interpretation of Φ in a Tarski model of intuitionistic logic is the same as the interpretation of Φ' in the corresponding topological model of QS4.

It is straightforward to check directly that if $\cdot\,\Phi$ is a law of intuitionistic logic, then $\cdot\,\Phi'$ is derivable in QS4; and if $\Phi_1, \ldots, \Phi_k \,/\, \Psi$ is an inference rule of intuitionistic logic, then $\Phi'_1, \ldots, \Phi'_k \,/\, \Psi'$ is derivable in QS4. Alternatively, this follows from the strong completeness of QS4 with respect to topological models.

The converse implications follow from the strong completeness of intuitionistic logic with respect to Tarski models. (A syntactic verification is also possible, but it relies on Gentzen's cut elimination theorem; see e.g. [143; p. 489].)

Of course, one could as well insert a box in front of *every* subformula of Φ (in addition to replacing $\times$ and $\checkmark$ by $\bot$ and $\top$); the resulting formula $\Phi_\Box$ will be equivalent to Φ' in QS4:

Proposition 5.21. *The following holds in QS4:*
   (a) $\vdash (\Box p \wedge \Box q) \longleftrightarrow \Box(\Box p \wedge \Box q)$;
   (b) $\vdash (\Box p \vee \Box q) \longleftrightarrow \Box(\Box p \vee \Box q)$;
   (c) $\vdash \exists x\, \Box p(x) \longleftrightarrow \Box \exists x\, \Box p(x)$.

The topological counterpart of this proposition is saying essentially that finite intersections and arbitrary unions of open sets are open. Any proof of this fact from the axioms of a topological space in terms of Int should translate into a proof of 5.21.

By combining the above, we obtain

Theorem 5.22. *Given a formula Φ in the language of intuitionistic logic, insert a box in front of every subformula, and replace $\times$, $\checkmark$ by $\bot$, $\top$ respectively. Then the resulting formula $\Phi_\square$ is derivable in QS4 if and only if Φ is derivable in intuitionistic logic.*

Moreover, a rule $\Phi_1, \ldots, \Phi_k \,/\, \Psi$ is derivable in intuitionistic logic if and only if the rule $(F_1)_\square, \ldots, (F_k)_\square \,/\, G_\square$ is derivable in QS4.

Gödel's original translation, as extended to the first-order case (cf. [198]) is as follows: given a formula Φ, insert a box *after* each instance of $\vee$, $\rightarrow$ and $\exists$. Then the resulting formula Φ'' is derivable in QS4 if and only if Φ is derivable in intuitionistic logic. Indeed, Φ'' is derivable in QS4 if and only if $\square\Phi''$ is. On the other hand, $\Phi_\square \leftrightarrow \square\Phi''$ is derivable in QS4:

Proposition 5.23. *The following holds in QS4:*
(a) $\vdash \square(\mathsf{p} \wedge \mathsf{q}) \longleftrightarrow \square(\square\mathsf{p} \wedge \square\mathsf{q})$;
(b) $\vdash \square\forall\mathsf{x}\,\mathsf{p}(\mathsf{x}) \longleftrightarrow \square\forall\mathsf{x}\,\square\mathsf{p}(\mathsf{x})$.

Here each of the two judgements is proved similarly to their topological counterpart, $\operatorname{Int}\bigcap_i S_i = \operatorname{Int}\bigcap_i \operatorname{Int} S_i$. Indeed, here $\supset$ follows from $S_i \supset \operatorname{Int} S_i$. Conversely, we have $\bigcap_i S_i \subset S_n$ for each n, whence $\operatorname{Int}\bigcap_i S_i \subset \operatorname{Int} S_n$. Since this holds for each n, we get $\operatorname{Int}\bigcap_i S_i \subset \bigcap_i \operatorname{Int} S_i$. Applying Int to both sides completes the proof.

Here is a sample application of the $\square$-translation:

Proposition 5.24. *Jankov's principle does not imply the Strong Markov Principle.*

Proof. Let us consider the Tarski model in the one-point compactification $\mathbb{N}_+$ of the countable discrete space $\mathbb{N}$ with $\mathscr{D} = \mathbb{N}$. Let us consider the valuation v such that $|\mathsf{q}|^v(n) = \{n\}$ and $|\mathsf{r}|^v = \{\infty\}$, where q is unary and r is nullary, and any other problem variable is also interpreted by $\{\infty\}$ on every input. Since the union of the clopen sets $\{n\}$ is not regular open, (SMR) and in particular (SMP) are not valid with respect to the valuation field $\langle v \rangle$ (since their formulas are not valid with respect to v). To show that, nevertheless, Jankov's principle is valid with respect to $\langle v \rangle$, it is convenient to use the $\square$-translation. Let us note that every formula of QS4 is v-equivalent to (i.e., has the same interpretation under v as) a formula whose only problem variables are q, r.

If F is a quantifier-free closed m-formula of QS4 containing no $\square$'es, then either ∞ is contained in $|F|^v(\vec{n})$ for all $\vec{n}$, or ∞ is contained in $|\neg F|^v(\vec{n})$ for all $\vec{n}$; also, either $|F|^v(\vec{n})$ contains a punctured neighborhood of ∞ for all $\vec{n}$, or $|\neg F|^v(\vec{n})$ contains a punctured neighborhood of ∞ for all $\vec{n}$. Indeed, both assertions are true if F is either one of q, $\neg\mathsf{q}$, r, $\neg\mathsf{r}$, and remain true under conjunction and disjunction; thus they follow by considering the disjunctive normal form of F.

If $|F|^v(\vec{n})$ contains ∞ and $|\neg F|^v(\vec{n})$ contains a punctured neighborhood of ∞ for all $\vec{n}$, then $\square F$ is v-equivalent to $F \wedge \neg\mathsf{r}$; and in the remaining cases either $|F|^v(\vec{n})$ contains a punctured neighborhood of ∞ for all $\vec{n}$, or $|\neg F|^v(\vec{n})$ contains a neighborhood of ∞ for all $\vec{n}$, whence $\square F$ is v-equivalent to F. It follows by induction that every quantifier-free closed m-formula is v-equivalent to a quantifier-free closed m-formula without $\square$'es. In

other words, every quantifier-free formula is v-equivalent to a quantifier-free formula without $\square$'es.

If F is a quantifier-free 1-formula without $\square$'es, by considering its conjunctive normal form we get that $F(\mathsf{x})$ is v-equivalent to a formula of the form $\big(\mathsf{q}(\mathsf{x}) \vee G\big) \wedge \big(\neg\mathsf{q}(\mathsf{x}_1) \vee H\big)$, where G and H contain no occurrences of x. Consequently $\forall\mathsf{x}\, F(\mathsf{x})$ is v-equivalent to $\big(\forall\mathsf{x}\,\mathsf{q}(\mathsf{x}) \vee G\big) \wedge \big(\forall\mathsf{x}_1\,\neg\mathsf{q}(\mathsf{x}) \vee H\big)$, which is v-equivalent to $G \wedge \big(\mathsf{r} \vee H\big)$. By similarly using the disjunctive normal form, we get that $\exists\mathsf{x}\, F(\mathsf{x})$ is v-equivalent to a formula that contains no quantification over x. It follows by induction that every formula is v-equivalent to a quantifier-free formula without $\square$'es.

Finally, by combining the above results, we get that for each m-formula F, every set $|F|^v(n_1, \ldots, n_m)$ is either clopen or differs from a clopen set only in the point ∞. (So, for example, it cannot be the set of all even integers.) Then for each $v' \in \langle v \rangle$ and for each m-ary predicate variable p, every set $|p|^{v'}(n_1, \ldots, n_m)$ is either clopen or differs from a clopen set only in the point ∞. In either case, it is easy to see that $|\square\neg\square p|^{v'}(n_1, \ldots, n_m)$ is clopen. Thus the $\square$-translation of $\neg\alpha$, where α is nullary, is interpreted by a clopen set under v'. Hence the $\square$-translation of $\cdot\neg\alpha \vee \neg\neg\alpha$, a form of Jankov's principle (see 5.9), is valid with respect to $\langle v \rangle$. $\square$

5.8. Independence of connectives and quantifiers

Theorem 5.25. *Neither of the intuitionistic connectives and quantifiers $\times$, $\wedge$, $\vee$, $\rightarrow$, $\exists$, $\forall$ can be expressed in terms of the other ones.*

Our argument attempts to focus on properties that distinguish the connectives, but occasionally resorts to specific models. For a proof-theoretic approach see [164; p. 59].

Proof. It is easy to see that $\times$ is not expressible in terms of the other connectives and quantifiers by considering a Tarski model in a nonempty space X with the valuation that associates X to every problem variable and every tuple of elements of the domain. Then every formula not involving $\times$ is interpreted by X, but in every Tarski model, $\times$ is interpreted by $\varnothing$.

In intuitionistic logic, decidability is preserved under $\vee$ (using (28)), $\wedge$ (using (29)) and $\rightarrow$ (using (6) and (31)). In the model showing that the Weak Rule of Omniscience does not imply the Rule of Omniscience (see §5.5.1), decidability is also preserved under $\forall$, but not under $\exists$. Thus $\exists$ is not expressible in terms of the other connectives and quantifiers.

Due to the intuitionistically derivable principles (43), (44), (45), (46) and (47), and to the independence of the $\neg\neg$-Shift Principle (see §5.6.1), $\forall$ is the only connective or quantifier through which $\neg\neg$ cannot be "pushed inside". Thus $\forall$ is not expressible in terms of the other connectives and quantifiers.

Due to the same intuitionistically derivable principles and to the independence of Jankov's principle (see §5.5), $\vee$ and $\exists$ are the only connectives or quantifiers through which $\neg\neg$ cannot be "pushed outside". In models with singleton domain, however, $\neg\neg$

can be pushed outside through $\exists$ (but generally not through $\vee$). Thus $\vee$ cannot be expressed in terms of other connectives or quantifiers.

(Alternatively, one can combine the use of the singleton domain with the observation that $\vee$ is the only connective that does not preserve stability; or with the observation that $\vee$ is the only connective that is not preserved under the $\neg\neg$-translation, but if it were expressible in terms of other connectives and quantifiers in intuitionistic logic, then the same expression would also work in classical logic.)

If $\to$ is expressible in terms of $\wedge$, $\vee$, $\forall$, $\exists$ and $\times$, then so is $\neg$, and if so, such an expression would also work in classical logic (with $\times$ replaced by $\bot$), and consequently also in QS4. Now $\wedge$, $\vee$ and $\exists$ are preserved under the $\square$-translation (in the McKinsey–Tarski form, see 5.21), $\times$ turns into $\bot$, and $\forall$ is preserved from the viewpoint of models with singleton domain (Tarksi models of intuitionisitc logic and topological models of QS4). But $\neg$ is not preserved: $\neg\square\mathsf{p} \to \square\neg\square\mathsf{p}$ is not derivable in QS4, since in topological models, the complement of an open set does not have to be open.

Finally, given a collection $\Gamma_1, \ldots, \Gamma_n$ of closed formulas such that $\vdash \neg\neg\Gamma_i$ for each i, any closed formula Ξ obtained from $\Gamma_1, \ldots, \Gamma_n$ using $\vee$, $\to$ and $\times$ satisfies either $\vdash \Gamma_i \to \Xi$ for some i or $\vdash \Xi \to \times$. Indeed, arguing by induction, we may assume that Ξ is either $\Phi \vee \Psi$ or $\Phi \to \Psi$ or $\times$, where either $\vdash \Gamma_i \to \Phi$ or $\vdash \Phi \to \times$, and either $\vdash \Gamma_j \to \Psi$ or $\vdash \Psi \to \times$. If $\Xi = \times$, then $\vdash \Xi \to \times$. In the case $\Xi = \Phi \vee \Psi$, if $\vdash \Gamma_i \to \Phi$, then $\vdash \Gamma_i \to \Xi$; if $\vdash \Gamma_j \to \Psi$, then $\vdash \Gamma_j \to \Xi$; and if $\vdash \Phi \to \times$ and $\vdash \Psi \to \times$, then $\vdash (\Phi \vee \Psi) \to \times$. It remains to consider the case $\Xi = \Phi \to \Psi$. If either $\vdash \Gamma_j \to \Psi$ or $\vdash \Phi \to \times$, then $\vdash \Gamma_j \to \Xi$. So we may assume that $\vdash \Gamma_i \to \Phi$ and $\vdash \Psi \to \times$. Then $\vdash \Xi \to \neg\Phi$ and $\vdash \neg\neg\Gamma_i \to \neg\neg\Phi$. By the hypothesis $\vdash \neg\neg\Gamma_i$, so $\vdash \neg\neg\Phi$. Hence $\vdash \Xi \to \times$.

Now as far as models with singleton domain are concerned, the use of $\exists$ and $\forall$ does not yield anything new — but the use of $\wedge$ does, by considering the Tarski model in $\mathbb{R}$ (with singleton domain) and a valuation with $|\alpha| = \mathbb{R} \setminus \{0\}$ and $|\beta| = \mathbb{R} \setminus \{1\}$, which suffices to detect that $\Xi = \alpha \wedge \beta$ satisfies neither $\vdash \alpha \to \Xi$ nor $\vdash \beta \to \Xi$ nor $\vdash \Xi \to \times$. $\qquad\square$

6. Sheaf-valued models

The first two sections below contain, in particular, a review of some basic sheaf theory. This goes only slightly beyond an exposition of a standard material (for sheaves of sets), which can be found in Bredon's [25] and/or Godement's [66] books (mostly for sheaves of modules).

6.1. Sheaves

6.1.1. Sheaves and presheaves. A *sheaf* (of sets) on a (topological) space B is a map $\mathcal{F} \colon E \to B$ that is a local homeomorphism (i.e. each $e \in E$ has a neighborhood O in E such that $\mathcal{F}(O)$ is a neighborhood of $f(e)$ and $\mathcal{F}|_O \colon O \to \mathcal{F}(O)$ is a homeomorphism). Clearly, such an $\mathcal{F}$ is an open map; in particular, $\operatorname{Supp} \mathcal{F} := \mathcal{F}(E)$ is always open in B. If $\mathcal{F} \colon E \to B$ is a sheaf and $b \in B$, the point-inverse $\mathcal{F}_b := \mathcal{F}^{-1}(b)$ is called the *stalk* of $\mathcal{F}$ at b.

A *section* of a sheaf or more generally of a continuous map $f\colon E \to B$ over an open subset $U \subset B$ is a continuous map $s\colon U \to E$ such that $fs = \mathrm{id}_U$. Clearly, every section is a homeomorphism onto its image. A section $s\colon U \to E$ is said to *extend* a section $t\colon V \to E$ if $V \subset U$ and $t = s|_V$. If $\mathcal{F}\colon E \to B$ is a sheaf, it is easy to see that a base of topology on E is given by the images of all sections of $\mathcal{F}$ (over all open subsets of B).

A *morphism* $\mathcal{F} \to \mathcal{F}'$ between sheaves $\mathcal{F}\colon E \to B$ and $\mathcal{F}'\colon E' \to B$ is a continuous map $f\colon E \to E'$ such that $\mathcal{F}'f = \mathcal{F}$. Clearly, such a map is itself a local homeomorphism. A *constant* sheaf is the projection $B \times \Xi \to B$, where Ξ is a discrete space. A sheaf is *locally constant* if it is a covering map, that is, its restriction over a sufficiently small neighborhood of every point of B is isomorphic to a constant sheaf.

There are two basic reasons why a sheaf can fail to be locally constant:

- Even locally, sections need not extend, as is the case for the *characteristic sheaf* $\chi_U\colon U \to B$, where U is an open subset of B and χ_U is the inclusion map.
- Even locally, sections may extend non-uniquely, as is the case for the *amalgamated union* $\chi_U \cup_{\chi_W} \chi_V\colon U \cup_W V \to B$, where U and V are open subsets of B and W is an open subset of $U \cap V$. Here $U \cup_W V$ is the quotient space of the disjoint union $U \sqcup V = U \times \{0\} \cup V \times \{1\} \subset B \times \{0, 1\}$ by the equivalence relation $(b, 0) \sim (b, 1)$ if $b \in W$, and $\chi_U \cup_{\chi_W} \chi_V$ sends the class of (b, i) to b.

Note that $U \cup_W V$ is non-Hausdorff, as long as some $b \in (U \cap V) \setminus W$ lies in the closure of W — in this case $(b, 0)$ and $(b, 1)$ have no disjoint neighborhoods in $U \cup_W V$. In general, if $s\colon U \to E$ and $t\colon V \to E$ are sections of a sheaf $\mathcal{F}\colon E \to B$ over open sets, then $\{b \in U \cap V \mid s(b) = t(b)\}$ is open. If E happens to be Hausdorff, then this set must also be closed in $U \cap V$.

A *presheaf* (of sets) on a (topological) space B is a contravariant functor from the category of open subsets of B and their inclusions into the category of sets. A *morphism* $\varphi\colon F \to G$ of presheaves on B is a natural transformation of functors, that is, a collection of maps $\varphi_U\colon F(U) \to G(U)$ for all open $U \subset B$ that commute with the restriction maps $F(j)\colon F(U) \to F(V)$ and $G(j)\colon G(U) \to G(V)$ for all inclusions $j\colon V \hookrightarrow U$ between open subsets of B.

Given a sheaf, or more generally a continuous map $f\colon X \to B$, its *presheaf of sections* σf assigns to an open subset $U \subset B$ the set $(\sigma f)(U)$ of all continuous sections of f over U, and to an inclusion $j\colon V \hookrightarrow U$ of open subsets of B the map $(\sigma f)(U) \to (\sigma f)(V)$ given by restriction of sections. In general, for an arbitrary presheaf F, elements of each $F(U)$ are called *sections* over U, and the image of a section $s \in F(U)$ under $F(j)\colon F(U) \to F(V)$ is called the *restriction* of s and is denoted $s|_V$. If U is an open subset of B, the presheaf $\mathrm{Char}\, U := \sigma\chi_U$ can be described by $(\mathrm{Char}\, U)(V) = \{\mathrm{id}_V\}$ if $V \subset U$ and $(\mathrm{Char}\, U)(V) = \varnothing$ otherwise.

If F is a presheaf on a space B and $b \in B$, a *germ* of F at b is an element of the direct limit (=colimit, inductive limit) $F_b := \varinjlim F(U)$ over all open neighborhoods U of b ordered by inclusion. In other words, a germ of F at b is an equivalence class of sections $s_U \in F(U)$ over open neighborhoods U of b, where $s_U \sim t_V$ if $U \cap V$ contains

an open neighborhood W of b such that $s_U|_W = t_V|_W$. In particular, $F_b = \varnothing$ if and only if $F(U) = \varnothing$ for all open neighborhoods U of b.

A presheaf F on a space B gives rise to the *sheaf of germs* (or sheafafication) of F, denoted γF, whose stalk $(\gamma F)_b$ at b is the set of germs F_b; a base of topology on $E = \bigcup_{b \in B} F_b$ consists of sets of the form $O_{s,U} = \{\varphi_{U,b}(s) \mid b \in U\}$, where $U \subset B$ is open and $\varphi_{U,b} \colon F(U) \to F_b$ is the natural map. The open set $\operatorname{Supp} F := \operatorname{Supp} \gamma F$ can be described as the union of all open subsets U of B such that $F(U) \neq \varnothing$.

6.1.2. Sheaf conditions. Given a sheaf $\mathcal{F}$, it is easy to see that $\gamma \sigma \mathcal{F}$ is isomorphic to $\mathcal{F}$. Given a presheaf F, we have an obvious morphism $F \to \sigma \gamma F$, whose component $F(U) \to \sigma \gamma F(U)$ is a bijection if and only if the following two *sheaf conditions* hold:

- If $\{U_\alpha\}$ is an open cover of U, and $s, t \in F(U)$ are such that $s|_{U_\alpha} = t|_{U_\alpha}$ for all α, then $s = t$.
- If $\{U_\alpha\}$ is an open cover of U, and $s_\alpha \in F(U_\alpha)$ are such that $s_\alpha|_{U_\alpha \cap U_\beta} = s_\beta|_{U_\alpha \cap U_\beta}$ for all α, β, then there exists an $s \in F(U)$ such that each $s_\alpha = s|_{U_\alpha}$.

Thus sheaves on B are in one-to-one correspondence with presheaves on B satisfying the sheaf conditions for every open $U \subset B$. For example, a presheaf of the form σf always satisfies the sheaf conditions, so it is isomorphic to $\sigma \mathcal{F}$, where $\mathcal{F} = \gamma \sigma f$ is called the *sheaf of germs of sections* of f. Moreover, every sheaf morphism $\mathcal{F} \to \mathcal{G}$ clearly determines a presheaf morphism $\sigma \mathcal{F} \to \sigma \mathcal{G}$; and every presheaf morphism $F \to G$ clearly determines a sheaf morphism $\gamma F \to \gamma G$. It follows that the category of sheaves on B can be identified with a full subcategory of the category of presheaves on B.

Example 6.1. If $\pi \colon B \times X \to B$ is the projection, then $\gamma \sigma \pi$ is the *sheaf of germs of continuous functions $B \to X$*. Note that the total space E of this sheaf is non-Hausdorff already for $B = X = \mathbb{R}$, since the germs at 0 of the constant function $f(b) = 0$ and the function $g(b) = \max(b, 0)$ are distinct, but have no disjoint neighborhoods.

If C is a category, a *C-valued presheaf* on a space B is a contravariant functor from the category of open subsets of B and their inclusions to C. When C is a concrete category (for example, the category Ab of abelian groups), each C-valued presheaf on B is also a presheaf of sets on B.

Conversely, if F is a presheaf of sets on a space B, it determines an Ab-valued presheaf $\mathbb{Z}F$ on B by setting $\mathbb{Z}F(U)$ to be the free abelian group $\mathbb{Z}[F(U)]$ with generator set $F(U)$ for every open $U \subset B$, and $\mathbb{Z}F(j) \colon \mathbb{Z}[F(U)] \to \mathbb{Z}[F(V)]$ to be the linear extension of $F(j) \colon F(U) \to F(V)$ for every inclusion $j \colon V \hookrightarrow U$ between open subsets of B. It is clear that F satisfies the sheaf conditions if and only if $\mathbb{Z}F$ does.

On the other hand, the sheaf conditions for an Ab-valued presheaf F are easily seen to be equivalent to exactness of the sequence

$$0 \to F(U) \xrightarrow{\varphi} \prod_\alpha F(U_\alpha) \xrightarrow{\psi} \prod_{(\alpha,\beta)} F(U_\alpha \cap U_\beta) \qquad (*)$$

for any open subset $U \subset B$ and any open cover $\{U_\alpha\}$ of U, where $\varphi(s) = (\alpha \mapsto s|_{U_\alpha})$ and $\psi(\alpha \mapsto s_\alpha) = \big((\alpha, \beta) \mapsto s_\alpha|_{U_\alpha \cap U_\beta} - s_\beta|_{U_\alpha \cap U_\beta}\big)$.

Example 6.2. Let $f\colon E \to B$ be a monotone map between posets. If we endow E and B with the Alexandroff topology, then f is continuous. Every poset endowed with the Alexandroff topology is weakly homotopy equivalent to its order complex [132], so singular (co)homology of a poset, absolute or relative, may be thought of as (co)homology of its order complex. For each $i = 0, 1, \ldots$ let us define presheaves L^i and L_i on B by $L^i(U) = H^i(f^{-1}(U); \mathbb{Z})$ and $L_i(U) = H_i(E, E \setminus f^{-1}(U); \mathbb{Z})$; the restriction maps $L^i(U) \to L^i(V)$ and $L_i(U) \to L_i(V)$ are the usual forgetful homomorphisms. By the Mayer–Vietoris spectral sequence, the sequence $(*)$ is exact for $F = L^0$ and, if the order complex of E is of dimension $n < \infty$, then also for $F = L_n$. Thus L^0 and L_n satisfy the sheaf conditions.

In general, the sheaf of germs $\mathcal{H}^i(f) = \gamma L^i$ is known as the ith Leray sheaf of f, and when $f = \mathrm{id}_X\colon X \to X$, the sheaf of germs $\mathcal{H}_i(X) = \gamma L_i$ is known as the ith local homology sheaf of X. If X is the face poset of a graph, then $\mathcal{H}_1(X)$ has stalk $\mathbb{Z}$ at every edge and stalk $\mathbb{Z}^{d-1}$ at every vertex of degree d. If X is the face poset of a triangulation of a closed n-manifold, then $\mathcal{H}_n(X)$ is locally constant, with stalks $\mathbb{Z}$; it will be constant precisely when the manifold is orientable. If B and E are the face posets of simplicial complexes, and f comes from a simplicial map that triangulates a bundle with fiber a closed orientable k-manifold, then $\mathcal{H}^k(f)$ is locally constant, with stalks $\mathbb{Z}$; it will be constant precisely when the bundle is orientable.

6.1.3. Sheaves and problems. Let $f\colon X \to B$ be a continuous map (for instance, this could be a real polynomial $f\colon \mathbb{R} \to \mathbb{R}$ or a complex polynomial $f\colon \mathbb{C} \to \mathbb{C}$, $f(x) = a_n x^n + \cdots + a_0$), and consider the following parametric problem $\Gamma_f(b)$, $b \in B$:

Find a solution of the equation $f(x) = b$.

Thus each point-inverse $f^{-1}(b)$ is nothing but the set of solutions of $\Gamma_f(b)$.

If the parameter b represents experimental data (which inevitably contains noise), then any talk about the exact value of b (e.g. whether b is a rational or irrational number when $B = \mathbb{R}$) is pointless. It would be not so unreasonable, however, to assume that the value of b can in principle be determined up to arbitrary precision. Indeed, the point of continuous maps is that if a point in the domain is known only up to a sufficient accuracy, it still makes sense to speak of its image — it will be known up to a desired accuracy. This leads us to consider only *stable solutions* of the equation $f(x) = b$, that is, such $x_0 \in X$ that $f(x_0) = b$ and there exists a neighborhood U of b in the space B over which f has a section. Indeed, if the parameter b is only known to us up to a certain degree of precision, we should only stick with a solution x_0 of $f(x) = b$ as long as we can be certain that it would not disappear when our knowledge of b improves.[83] Thus we will consider the following parametric problem Δ_b, $b \in B$:

Find a stable solution of the equation $f(x) = b$.

---

[83]This may remind the reader of Brouwer's idea that "all functions are continuous", as well as some of Poincaré's writings; a more rigorous exposition of these ideas can be found in M. Escardo's notes [51] and in P. S. Novikov's physical interpretation of intuitionisic logic in terms of weighting of masses [149; §III.4].

If is easy to see that the set $\mathcal{F}_b \subset f^{-1}(b)$ of all stable solutions of $f(x) = b$ is nothing but the stalk of the sheaf $\mathcal{F} = \gamma\sigma f$ of germs of sections of the map f.

For example, if $f \colon \mathbb{R} \to \mathbb{R}$ is a polynomial with real coefficients, then stable solutions of $f(x) = b$ are all the non-repeated real roots of the polynomial $f(x) - b$, along with its other roots of odd multiplicity. Indeed, f is a local homeomorphism at each of those roots, but has a local extremum at every root x_0 of even multiplicity. Hence the perturbed equation $f(x) = b \pm \varepsilon$ (with $+$ chosen in the case of a local maximum, and $-$ in the case of a local minimum) has no real solutions near x_0 as long as $\varepsilon > 0$ is sufficiently small.

If $f \colon \mathbb{C} \to \mathbb{C}$ is a polynomial with complex coefficients, then stable solutions of the equation $f(z) = b$, where $b \in \mathbb{C}$, are precisely all the non-repeated complex roots of the polynomial $f(z) - b$. Indeed, at a root z_0 of multiplicity $d > 1$ we can write $f(z) = b + k(z - z_0)^d + o\left((z - z_0)^d\right)$. So f is locally a d-fold branched covering, and therefore admits no section over any neighborhood of $b = f(z_0)$. In other words, even though the perturbed equation $f(z) = b + \varepsilon$ does have solutions near z_0 whenever $|\varepsilon|$ is sufficiently small, we cannot choose such a solution continuously depending on ε; but with discontinuous choice, we can never be sure that our chosen explicit solution will persist when our knowledge of b improves.

Similar behaviour occurs if we consider the problem of finding a root of a complex polynomial $f_b = a_n z^n + \cdots + a_0$ as a function of all its coefficients $b = (a_0, \ldots, a_n)$. Set $\Sigma_n = \{(a_0, \ldots, a_n, x) \in \mathbb{C}^{n+2} \mid a_n x^n + \cdots + a_0 = 0\}$, and let $\varphi \colon \Sigma_n \to \mathbb{C}^{n+1}$ be the projection along the last coordinate. Then $\varphi^{-1}(b)$ is the set of roots of f_b, and the stalk $(\sigma\varphi)_b$ is the set of its stable roots. By varying just a_0 like above we see that no multiple root is stable; clearly, all non-repeated roots are stable.

6.2. Operations on sheaves

6.2.1. Product and coproduct. Given a collection of presheaves F^i, $i \in I$, their *product* $\prod F^i$ and *coproduct* $\bigsqcup F^i$ are defined respectively by $U \mapsto \prod_i F^i(U)$ and by $U \mapsto \bigsqcup_i F^i(U)$, with the obvious restriction maps.

If the presheaves F^i satisfy the sheaf conditions, then clearly so does $\prod F^i$. On the other hand, $\bigsqcup F^i$ generally does not satisfy the second sheaf condition, since for a disconnected open set $U \sqcup V$ the sheaf conditions for each F^i imply $F^i(U \sqcup V) \simeq F^i(U) \times F^i(V)$, so that $\left(\bigsqcup_i F^i\right)(U \sqcup V) \simeq \bigsqcup_i F^i(U) \times F^i(V)$, which is generally not the same as $\left(\bigsqcup_i F^i\right)(U) \times \left(\bigsqcup_i F^i\right)(V) = \left(\bigsqcup_i F^i(U)\right) \times \left(\bigsqcup_i F^i(V)\right)$.

Given a collection of sheaves $\mathcal{F}^i$, $i \in I$, their *product* $\prod \mathcal{F}^i := \gamma \prod \sigma\mathcal{F}^i$ satisfies $\sigma \prod \mathcal{F}^i \simeq \prod \sigma\mathcal{F}^i$ by the above. Their *coproduct* is defined by $\bigsqcup \mathcal{F}^i = \gamma\bigsqcup \sigma\mathcal{F}^i$.

It is not hard to see that the operations just introduced are precisely the product and the coproduct in the categories of sheaves and presheaves.

The coproduct of sheaves $\mathcal{F}^i \colon E_i \to B$ can be alternatively described as the sheaf $\bigsqcup \mathcal{F}^i \colon \bigsqcup E_i \to B$, where $\bigsqcup E_i$ is the disjoint union of spaces, and $\bigsqcup \mathcal{F}^i$ is defined by $\mathcal{F}^i$ on the ith summand. Thus $\left(\bigsqcup \mathcal{F}^i\right)_b \simeq \bigsqcup(\mathcal{F}^i)_b$ (isomorphism in the category of sets) for

each $b \in B$. More generally, $\left(\bigsqcup F^i\right)_b \simeq \bigsqcup (F^i)_b$ for any presheaves F^i since direct limit commutes with coproducts.

The product of *two* sheaves $\mathcal{F} \colon E \to B$ and $\mathcal{F}' \colon E' \to B$ can be alternatively described as the fiberwise product $\mathcal{F} \times \mathcal{F}' \colon E \times_B E' \to B$, where $E \times_B E'$ is the subspace of the Cartesian product $E \times E'$ consisting of all pairs (e, e') such that $\mathcal{F}(e) = \mathcal{F}'(e')$. Thus $(\mathcal{F} \times \mathcal{F}')_b \simeq \mathcal{F}_b \times \mathcal{F}'_b$. More generally, $(F \times F')_b \simeq F_b \times F'_b$ for any presheaves F and F' since direct limit commutes with finite products.

The fiberwise product $\left(\prod_B E_i\right) \to B$ of the sheaves $\mathcal{F}^i \colon E_i \to B$ need not be a sheaf when I is infinite. For, although every $x_i \in (\mathcal{F}^i)_b$ has a neighborhood in E_i that maps homeomorphically onto a neighborhood U_i of b in B, the fiberwise product may fail to be a local homemorphism at the point $(i \mapsto x_i)$ if $\bigcap_i U_i$ is not a neighborhood of b.

Given presheaves F^i, $i \in I$, for each $n \in I$, the projection $\prod_i F^i \to F^n$ induces a map $\left(\prod_i F^i\right)_b \to (F^n)_b$ for each $b \in B$. These in turn yield a map

$$\mathscr{G}_b \colon \left(\prod F^i\right)_b \to \prod (F^i)_b.$$

An element of the right hand side is a collection of germs of local sections $s_i \colon U_i \to E_i$, whereas an element of the left hand side is a germ of a collection of local sections $U \to E_i$. The germs of s_i's may fail to define such a single germ if $\bigcap U_i$ is not a neighborhood of b; thus $\mathscr{G}_b$ is not surjective in general. On the other hand, given two collections of local sections, $\sigma_i \colon U \to E_i$ and $\tau_i \colon V \to E_i$ such that each σ_i equals τ_i on some open set $W_i \subset U \cap V$, there might be no single open set $W \subset U \cap V$ such that each $\sigma_i|_W = \tau_i|_W$; thus $\mathscr{G}_b$ is not injective in general.[84]

It is clear, however, that $\mathscr{G}_b$ is a bijection when B is an Alexandroff space.

Let us note that the existence of $\mathscr{G}_b$ implies that if $\left(\prod F^i\right)_b \neq \varnothing$, then each $(F^i)_b \neq \varnothing$. The converse fails:

Proposition 6.3. *There exist sheaves $\mathcal{F}^1, \mathcal{F}^2, \ldots$ on $\mathbb{R}$ such that each $(\mathcal{F}^i)_b \neq \varnothing$ for each $b \in \mathbb{R}$ but $\prod \mathcal{F}^i = \chi_\varnothing$.*

Proof. It suffices to consider, for each i, the cover $C_i = \{(r - \frac{1}{i}, r + \frac{1}{i}) \mid r \in \mathbb{R}\}$ of $\mathbb{R}$ by all open intervals of length $2/i$, and the sheaf $\mathcal{F}^i = \bigsqcup_r \chi_{(r-1/i, r+1/i)}$. $\square$

6.2.2. The support of a product. A topological space X is of *covering dimension zero* if every open cover of X has a refinement consisting of disjoint open sets; and of *inductive dimension zero* if it has a base of topology consisting of clopen sets. For T_4 spaces (i.e. T_1 spaces where disjoint closed subsets are separated by neighborhoods), covering dimension zero implies inductive dimension zero (see [49; 6.2.5, 6.2.6]). The converse holds for spaces whose topology has a countable base (see [49; 6.2.7]). In particular, the two notions are equivalent for separable metrizable spaces, so we may speak of *zero-dimensional separable metrizable spaces*. Examples of such spaces include the Cantor set and the Baire space, as well as their arbitrary subspaces. Indeed, it is easy to see that if a space is of inductive dimension zero, then so is every its subspace.

---

[84]Of course, it is easy to provide a specific example, thus refuting an assertion in [66; §II.1.10].

Lemma 6.4. *If $\mathcal{F}$ is a sheaf on a zero-dimensional separable metrizable space B, and U is an open subset of B, then $\mathcal{F}$ has a section over U if and only if $U \subset \operatorname{Supp} \mathcal{F}$.*

Proof. The "only if" implication follows from the definitions. Given an open $U \subset \operatorname{Supp} \mathcal{F}$, each $x \in U$ is contained in an open set V_x over which $\mathcal{F}$ has a section. Without loss of generality $V_x \subset U$. Since U is zero-dimensional, the cover $\{V_x \mid x \in U\}$ of U has a refinement $\{U_\alpha\}$ consisting of disjoint open sets. Since $\mathcal{F}$ has a section over each U_α and is a sheaf (not just a presheaf), by combining these sections we get a section over U. $\square$

Proposition 6.5. *If F^i are presheaves on a space B, then $\operatorname{Supp} \prod F^i \subset \operatorname{Int} \left(\bigcap \operatorname{Supp} F^i \right)$. The reverse inclusion holds in each of the following cases:*

    (a) B is an Alexandroff space;
    (b) B is a zero-dimensional separable metrizable space, and $F^i = \sigma \mathcal{F}^i$ for some $\mathcal{F}^i$.

Proof. The inclusion follows from the existence of the map $\mathscr{G}_b$ and from the fact that the left hand side is an open set. The reverse inclusion for holds for Alexandroff spaces since their $\mathscr{G}_b$ is bijective. If B is a zero-dimensional separable metrizable space and x lies in an open set U contained in $\operatorname{Supp} \mathcal{F}^i$ for each i, then by 6.4 each $\mathcal{F}^i$ has a section over U. Hence so does their product, whose support therefore contains x. $\square$

6.2.3. Hom-sheaf. If F is a presheaf on a space B, and $A \subset B$ is open, then open subsets of A are open in B, so F determines the restriction presheaf $F|_A$ on A. In the case of a sheaf $\mathcal{F}\colon E \to B$, the corresponding construction generalizes to the case of an arbitrary continuous map $f\colon A \to B$, which induces the inverse image $f^*\mathcal{F}\colon f^*E \to A$, where $f^*E = \{(a, e) \in A \times E \mid f(a) = \mathcal{F}(e)\}$ and $(f^*\mathcal{F})(a, e) = a$.

    If F and G are presheaves on B, let us consider the set $\operatorname{Hom}(F, G)(U)$ of morphisms $F|_U \to G|_U$. If V is an open subset of U, every such morphism φ restricts to a morphism $\varphi|_V\colon F|_V \to G|_V$, and thus we get the *exponential* presheaf $\operatorname{Hom}(F, G)$. If G satisfies the sheaf conditions, then clearly so does $\operatorname{Hom}(F, G)$. This enables one to define the sheaf $\operatorname{Hom}(\mathcal{F}, \mathcal{G}) := \gamma \operatorname{Hom}(\sigma \mathcal{F}, \sigma \mathcal{G})$, called the *sheaf of germs of morphisms* between the sheaves $\mathcal{F}$ and $\mathcal{G}$. In general, for any presheaves F and G, the stalk $\operatorname{Hom}(F, G)_b$ consists of germs of morphisms $F \to G$, that is, of equivalence classes of local morphisms $\varphi_U\colon F|_U \to G|_U$ over open neighborhoods U of b, where $\varphi_U \sim \varphi_V$ if $U \cap V$ contains an open neighborhood W of b such that $\varphi_U|_W = \varphi_V|_W$. A germ of morphisms determines a map between germs, so we get a map

$$\mathscr{F}_b\colon \operatorname{Hom}(F, G)_b \to \operatorname{Hom}(F_b, G_b),$$

where $\operatorname{Hom}(S, T)$ denotes the set of all maps between the sets S and T. This map is neither injective nor surjective in general:

Proposition 6.6. *For any set S there exist sheaves $\mathcal{F}$, $\mathcal{G}$ on $\mathbb{R}$ such that $\mathcal{F}_0 = \mathcal{G}_0 = \varnothing$ and $\operatorname{Hom}(\mathcal{F}, \mathcal{G})_0 \simeq S$.*

Proof. Let $U = \mathbb{R} \setminus \{0\}$ and let $\mathcal{F} = \chi_U$. Let $\mathcal{G}$ be the disjoint union of copies of $\mathcal{F}$ indexed by S. Then for every open neighborhood V of 0 the set of sheaf morphisms $\mathcal{F}|_V \to \mathcal{G}|_V$ is in a bijection with S, and its elements persist under restriction. $\square$

Let us mention two more peculiar examples:

Example 6.7. There exist sheaves $\mathcal{F}$, $\mathcal{G}$ on $\mathbb{R}^2$ such that $\operatorname{Supp}\mathcal{F} = \operatorname{Supp}\mathcal{G} = \mathbb{R}^2 \setminus \{0\}$, yet $\operatorname{Hom}(\mathcal{F}, \mathcal{G})_0 = \varnothing$. Indeed, let $U = \mathbb{R}^2 \setminus \{0\}$ and let $\mathcal{F} = \chi_U$. Let $\mathcal{G} \colon \tilde{U} \to U \subset \mathbb{R}^2$ be the nontrivial double covering over U. Then for every open neighborhood V of 0 there exist no sheaf morphisms $\mathcal{F}|_V \to \mathcal{G}|_V$.

Example 6.8. There exist sheaves $\mathcal{F}^n$ and $\mathcal{G}$ on $\mathbb{R}$, $n \in \mathbb{N}$, such that each $\operatorname{Hom}(\mathcal{F}^n, \mathcal{G}) \neq \chi_\varnothing$ but $\operatorname{Hom}(\bigsqcup_n \mathcal{F}^n, \mathcal{G}) = \chi_\varnothing$. Indeed, let $\mathcal{F}^n = \chi_{(-\infty, -1/n) \cup (1/n, \infty)}$ and let $\mathcal{G} = \chi_\varnothing$. Then $\operatorname{Hom}(\mathcal{F}^n, \mathcal{G}) \simeq \chi_{(-1/n, 1/n)}$. It follows that $\operatorname{Hom}(\bigsqcup_n \mathcal{F}^n, \mathcal{G}) \simeq \prod_{n \in \mathbb{N}} \chi_{(-1/n, 1/n)} \simeq \chi_\varnothing$.

6.2.4. The support of a Hom-sheaf.

Proposition 6.9. $\operatorname{Supp}\operatorname{Hom}(F, G) \subset \operatorname{Int}\big(\operatorname{Supp}G \cup (B \setminus \operatorname{Supp}F)\big)$ *for presheaves F, G on a space B, and the reverse inclusion holds in each of the following cases:*
  (a) G is the characteristic presheaf of an open set, in which case so is $\operatorname{Hom}(F, G)$;
  (b) B is a zero-dimensional separable metrizable space, and $G = \sigma\mathcal{G}$ for some $\mathcal{G}$.

Proof. The inclusion follows from the existence of the map $\mathscr{F}_b$ and from the fact that the left hand side is an open set. Suppose that x lies in an open set U contained in $\operatorname{Supp}G \cup (B \setminus \operatorname{Supp}F)$. Thus if $y \in U$ and $y \in \operatorname{Supp}F$, then $y \in \operatorname{Supp}G$; that is, $U \cap \operatorname{Supp}F \subset U \cap \operatorname{Supp}G$. Then for each open $V \subset U$ such that $V \not\subset \operatorname{Supp}G$ we have $V \not\subset \operatorname{Supp}F$, hence $F(V) = \varnothing$ and there is a unique map $F(V) \to G(V)$. Let us now consider an open $V \subset U$ such that $V \subset \operatorname{Supp}G$.

If $G = \operatorname{Char}(\operatorname{Supp}G)$, then $G(V) = \{\operatorname{id}_V\}$ and there is a unique map $F(V) \to G(V)$.

If B is a zero-dimensional separable metrizable space, and $G = \sigma\mathcal{G}$, then by [6.4] $\mathcal{G}$ has a section s over $U \cap \operatorname{Supp}G$, and we have the constant map $F(V) \to G(V)$ onto $\{s|_V\}$.

In either case, the constructed maps $F(V) \to G(V)$ commute with the restriction maps $G(V) \to G(W)$ and $F(V) \to F(W)$ for each open $W \subset V \subset U$, and so determine a morphism of presheaves $F|_U \to G|_U$. The germ at x of this morphism is an element of the stalk $\operatorname{Hom}(F, G)_x$, so x is contained in $\operatorname{Supp}\operatorname{Hom}(F, G)$.

Moreover, in the case (a) we have shown that for every open U contained in $\operatorname{Supp}G \cup (B \setminus \operatorname{Supp}F)$, we have a unique morphism $F|_U \to G|_U$; thus $\operatorname{Hom}(F, G)(U)$ contains precisely one element. If U is not contained in $\operatorname{Supp}G \cup (B \cup \operatorname{Supp}F)$, then U contains a point y such that $y \notin \operatorname{Supp}G$ and $y \in \operatorname{Supp}F$. The latter implies that y lies in an open set V such that $F(V) \neq \varnothing$. Without loss of generality $V \subset U$. On the other hand, since $y \notin \operatorname{Supp}G$, we have $G(V) = \varnothing$. Hence there exists no map $F(V) \to G(V)$, and consequently no morphism $F|_U \to G|_U$. Thus $\operatorname{Hom}(F, G)(U) = \varnothing$. It then follows that $\operatorname{Hom}(F, G)$ is isomorphic to $\operatorname{Char}\big(\operatorname{Supp}G \cup (B \cup \operatorname{Supp}F)\big)$. $\qquad\square$

Corollary 6.10. $\operatorname{Hom}\big(F, \operatorname{Char}\varnothing\big) \simeq \operatorname{Char}\big(\operatorname{Int}(B \setminus \operatorname{Supp}F)\big)$.

The inclusion in [6.9] cannot be reversed for sheaves over an Alexandroff space:

Proposition 6.11. *There exist sheaves $\mathcal{F}$, $\mathcal{G}$ over a certain Alexandroff space B such that $\operatorname{Supp}\mathcal{F} = \operatorname{Supp}\mathcal{G}$, yet $\operatorname{Supp}\operatorname{Hom}(\mathcal{F}, \mathcal{G}) \neq B$.*

For future purposes we fix some terminology regarding sheaves on Alexandroff spaces. If B in an Alexandroff space and $p \in B$, let $\langle p \rangle = \{q \in B \mid q \geq p\}$, the minimal open set containing p. Given a sheaf $\mathcal{F} \colon E \to B$, for each $p, q \in B$ with $p \leq q$ let $\mathcal{F}_{pq} \colon \mathcal{F}_p \to \mathcal{F}_q$ denote the restriction map $(\sigma \mathcal{F})(j) \colon (\sigma \mathcal{F})(\langle p \rangle) \to (\sigma \mathcal{F})(\langle q \rangle)$ corresponding to the inclusion $j \colon \langle q \rangle \hookrightarrow \langle p \rangle$. Let us note that the topology on E is an Alexandroff topology; its corresponding preorder is given by $x \leq y$ if $x \in \mathcal{F}_p$, $y \in \mathcal{F}_q$, $p \leq q$ and $y = \mathcal{F}_{pq}(x)$.

Proof. Let B be the product of two copies of the poset $0 < 1$. Let $\mathcal{F}_{(0,0)} = \varnothing$ and $\mathcal{F}_{(1,0)} = \mathcal{F}_{(0,1)} = \mathcal{F}_{(1,1)} = \{z\}$. Let $\mathcal{G}_{(0,0)} = \varnothing$, $\mathcal{G}_{(1,0)} = \{x\}$, $\mathcal{G}_{(0,1)} = \{y\}$ and $\mathcal{G}_{(1,1)} = \{x, y\}$, where the maps $\mathcal{G}_{pq}$ are treated as inclusions. Clearly, $\mathrm{Hom}(\mathcal{F}, \mathcal{G})_{(0,0)} = \varnothing$. $\qquad \square$

In the general case, the inclusion in 6.9 can be partially reversed:

Proposition 6.12. $\mathrm{Supp}\, G \cup \mathrm{Int}(B \setminus \mathrm{Supp}\, F) \subset \mathrm{Supp}\, \mathrm{Hom}(F, G)$.

Proof. If $b \in \mathrm{Supp}\, G$, then $G(U)$ is nonempty for some open neighborhood U of b. Then we can pick some $s_U \in G(U)$ and define a morphism $F|_U \to G|_U$ by sending every $F(V)$, $V \subset U$, onto $\{s_U|_V\} \subset G(V)$. If $B \setminus \mathrm{Supp}\, F$ contains an open neighborhood U of b, then $F(V) = \varnothing$ for all open $V \subset U$. Hence a morphism $F|_U \to G|_U$ is defined by observing that all objects in the domain are empty. $\qquad \square$

6.3. Sheaf-valued models

Let us summarize the behavior of stalks under operations on presheaves:

Proposition 6.13. *If F, G and F^i, $i \in I$, are presheaves on a space B, $b \in B$, then*

- $(F \times G)_b \simeq F_b \times G_b$;
- $(F \sqcup G)_b \simeq F_b \sqcup G_b$;
- *there is a map* $\mathscr{F}_b \colon \mathrm{Hom}(F, G)_b \to \mathrm{Hom}(F_b, G_b)$;
- $(\mathrm{Char}\, \varnothing)_b = \varnothing$;
- $\left(\bigsqcup F^i \right)_b \simeq \bigsqcup (F^i)_b$;
- *there is a map* $\mathscr{G}_b \colon \left(\prod F^i \right)_b \to \prod (F^i)_b$.

This looks suspiciously similar to the BHK interpretation (see, in particular, §**3.9**), which suggests that sheaves might have something to do with intuitionistic logic.

6.3.1. Sheaf-valued structures. Let B be a topological space and $\mathscr{D}$ a set. The set $\mathcal{O}$ will consist of all sheaves over B, and the function $! \colon \mathcal{O} \to \{\top, \bot\!\!\!\bot\}$ will send a sheaf $\mathcal{F}$ to $\top$ if and only if it has a global section, i.e. $(\sigma \mathcal{F})(B) \neq \varnothing$.

Thus, to specify a valuation, each problem variable α of arity n must be interpreted by a $\mathscr{D}^n$-indexed family of sheaves on B (i.e., by a function $|\alpha| \colon \mathscr{D}^n \to \mathcal{O}$). Upon a choice of a valuation v, a closed formula will be interpreted by a sheaf on B, and more generally a closed n-formula Φ will be interpreted by a $\mathscr{D}^n$-indexed family of sheaves on B, to be denoted $|\Phi|$ or in more detail $|\Phi|^v$, or in still more detail $|\Phi|^v_B$. (The valuation will be suppressed in the notation when it is clear from context.) Upon a further choice of a variable assignment u, an arbitrary formula will be interpreted by a sheaf on B, and

more generally an arbitrary n-formula Φ will be interpreted by a $\mathscr{D}^n$-indexed family of sheaves on B, to be denoted $|\Phi|$ or in more detail $|\Phi|_u$ or in still more detail $|\Phi|_u^v$. (The assignment will be suppressed in the notation when it is clear from context.) Validity of an n-formula Φ in the model, $\vDash \Phi$, means that the sheaf $|\Phi|_u^v(t_1, \ldots, t_n)$ has a global section for each tuple $(t_1, \ldots, t_n) \in \mathscr{D}^n$ for an arbitrary valuation v and assignment u.

Intuitionistic connectives and quantifiers are interpreted by the basic operations on sheaves:

- $|\Phi \vee \Psi| = |\Phi| \sqcup |\Psi|$;
- $|\Phi \wedge \Psi| = |\Phi| \times |\Psi|$;
- $|\Phi \to \Psi| = \mathrm{Hom}(|\Phi|, |\Psi|)$;
- $|\times| = \mathrm{Char}\,\varnothing$;
- $|\exists\mathsf{x}\,\Xi(\mathsf{x})| = \bigsqcup_{d \in D} |\Xi|(d)$;
- $|\forall\mathsf{x}\,\Xi(\mathsf{x})| = \prod_{d \in D} |\Xi|(d)$,

where Φ and Ψ are formulas and Ξ is a 1-formula, and u and v are fixed.

In exactly the same way we define *pre*sheaf-valued structures. The only essential difference between sheaf- and presheaf-valued structures is in the interpretations of $\vee$ and $\exists$, due to the fact that $\sigma\mathcal{F} \sqcup \sigma\mathcal{G} \neq \sigma(\mathcal{F} \sqcup \mathcal{G})$ in general.

Theorem 6.14. *Sheaf- and presheaf-valued structures are models of intuitionisitic logic.*

Since (pre)sheaves form a topos, this is a special case of Palmgren's models of intuitionistic logic in locally cartesian closed categories with finite sums [156].

Proof. We need to show that all laws and inference rules of the modified Spector's derivation system in §4.6 are valid in the two kinds of structures. It suffices to show that these laws and inference rules are valid with respect to an arbitrary predicate valuation v. Thus we fix v, and $|\cdot|$ will mean $|\cdot|^v$. For those laws and inference rules that involve neither $\vee$ nor $\exists$, the case of sheaves reduces to the case of presheaves, and so need not be considered separately.

(I) (modus ponens). We are given a global section s of $|\alpha|$ and a global section of $|\alpha \to \beta|$, that is, a presheaf morphism $f \colon |\alpha| \to |\beta|$. Then $f(B) \colon |\alpha|(B) \to |\beta|(B)$ sends s to a global section of $|\beta|$.

(II), (III) The identity is a natural transformation, and composition of natural transformations is a natural transformation.

(VI), (VII) The projections of the product of presheaves onto its factors as well as the inclusions of summands into the coproduct of presheaves are presheaf morphisms. Also, the inclusions of summands into the coproduct of sheaves are sheaf morphisms.

(VIII), (IX) (Co)product of presheaves is their (co)product in the category of presheaves. Also, coproduct of sheaves is their coproduct in the category of sheaves.

(IV) (exponential law) Given a presheaf morphism $f \colon |\alpha| \times |\beta| \to |\gamma|$, we need to construct a presheaf morphism $g \colon |\alpha| \to \mathrm{Hom}(|\beta|, |\gamma|)$. Let $g(U)$ send an $s \in |\alpha|(U)$ to the following presheaf morphism $h \colon |\beta||_U \to |\gamma||_U$. If $V \subset U$, define $h(V) \colon |\beta|(V) \to |\gamma|(V)$ by $t \mapsto f(V)(s|_V, t)$. If $V' \subset V$, then $f(V')(s|_{V'}, t|_{V'}) = f(V)(s|_V, t)|_{V'}$ since $s|_{V'} =$

$(s|_V)|_{V'}$ and f is a natural transformation. Hence $h = g(U)(s)$ is a natural transformation. If $U' \subset U$, then $g(U')(s|_{U'}) = g(U)(s)|_{U'}$ as presheaf morphisms $|\beta||_{U'} \to |\gamma||_{U'}$ due to $(s|_{U'})|_V = s|_V$ for each $V \subset U'$. Thus g is a natural transformation.

Conversely, given a g, we need to construct an f. Let $f(U) \colon |\alpha|(U) \times |\beta|(U) \to |\gamma|(U)$ send a pair (s,t) to $h(U)(t)$, where $h = g(U)(s)$. If $U' \subset U$, then $f(U')(s|_{U'}, t|_{U'}) = h'(U')(t|_{U'})$, where $h' = g(U')(s|_{U'})$. Since g is a natural transformation, $h' = h|_{U'}$ as presheaf morphisms $|\beta||_{U'} \to |\gamma||_{U'}$, where again $h = g(U)(s)$. In particular, $h'(U') = h(U')$ as maps $|\beta|(U') \to |\gamma|(U')$. Since h is a natural transformation, $h(U')(t|_{U'}) = h(U)(t)|_{U'}$. Thus $f(U')(s|_{U'}, t|_{U'}) = f(U)(s,t)|_{U'}$, whence f is a natural transformation.

(V) (explosion principle) $\chi_\varnothing$ is the initial object in the category of sheaves.

(X), (XI) For each $t \in D$, the projection onto the tth factor is a presheaf morphism $\prod_{d \in D} |\alpha|(d) \to |\alpha|(t)$, and the inclusion onto the tth summand is a (pre)sheaf morphism $|\alpha|(t) \to \bigsqcup_{d \in D} |\alpha|(t)$.

(XII$_\checkmark$) (generalization rule) We need to show that if $\alpha(x)$ is valid in a presheaf-valued structure for every variable assignment, then $\forall x\, \alpha(x)$ is valid in the structure.

Indeed, the hypothesis says that for each $d \in D$ we have a global section of $|\alpha|(d)$. These yield a global section of the presheaf $\prod_{d \in D} |\alpha|(d)$.

(XII'), (XIII') We need to prove that if $\forall x\, \big(\beta \to \alpha(x)\big)$ is valid in a (pre)sheaf structure, then so is $\beta \to \forall x\, \alpha(x)$; and if $\forall x\, \big(\alpha(x) \to \beta\big)$ is valid, then so is $\exists x\, \alpha(x) \to \beta$.

The hypothesis implies that for each $d \in D$ we have a presheaf morphism $|\beta| \to |\alpha|(d)$ (resp. $|\alpha|(d) \to |\beta|$). Since (co)product of presheaves is their (co)product in the category of presheaves, we get a presheaf morphism $|\beta| \to \prod_{d \in D} |\alpha|(d)$ (resp. $\bigsqcup_{d \in D} |\alpha|(d) \to |\beta|$). This works similarly for coproduct of sheaves. $\qquad\square$

6.3.2. Completeness. We note the following consequence of 6.13.

Proposition 6.15. *If F, G and F^i, $i \in I$, are presheaves on a space B, then*

(1) $\operatorname{Supp}(F \times G) = \operatorname{Supp} F \cap \operatorname{Supp} G$;

(2) $\operatorname{Supp}(F \sqcup G) = \operatorname{Supp} F \cup \operatorname{Supp} G$;

(3) $\operatorname{Supp} \operatorname{Hom}(F, G) \subset \operatorname{Int} \big(\operatorname{Supp} G \cup (B \setminus \operatorname{Supp} F)\big)$;

(4) $\operatorname{Supp}(\operatorname{Char} \varnothing) = \varnothing$;

(5) $\operatorname{Supp} \big(\bigsqcup F^i\big) = \bigcup \operatorname{Supp} F^i$;

(6) $\operatorname{Supp} \big(\prod F^i\big) \subset \operatorname{Int} \big(\bigcap \operatorname{Supp} F^i\big)$.

This is reminiscent of Tarski models, except that we have seen in the previous sections that the inclusions in (3) and (6) are strict in general, with (3) being strict even for some Alexandroff spaces.

However, by 6.5 and 6.9, the inclusions in (3) and (6) can be reversed if the presheaves satisfy the sheaf conditions, and B is a zero-dimensional separable metrizable space. Also, by 6.4 a sheaf over such a space B has a global section if and only if its support is the entire B. Thus from Dragalin's and Kramer's completeness results for Tarski models (see §5.3) we obtain

Theorem 6.16. *Intuitionistic logic is strongly complete with respect to its sheaf-valued model with countable domain over the Baire space $\mathbb{N}^{\mathbb{N}}$, and complete with respect to its sheaf-valued model with countable domain over any given zero-dimensional separable metrizable space with no isolated points.*

Although over zero-dimensional separable metrizable spaces sheaf-valued models can be seen as merely a proof-relevant conservative extension of Tarski models, in general they tend to be very different from Tarski models.

In particular, from 6.3 it is easy to see that the Tarski model of intuitionistic logic in $\mathbb{R}$ with $\mathscr{D} = \mathbb{N}$ is not obtainable by taking supports from the sheaf-valued model in $\mathbb{R}$ with $\mathscr{D} = \mathbb{N}$. This example can be improved:

Proposition 6.17. *The Tarski model of zero-order intuitionistic logic in some Alexandroff space B is not obtainable by taking supports from the sheaf-valued model in B.*

Proof. Let $B = \{x, y, xy, yx, \hat{1}\}$ with partial order generated by $x, y < xy$; $x, y < yx$ and $xy, yx < \hat{1}$. Then $\langle x \rangle \cap \langle y \rangle = \langle xy \rangle \cup \langle yx \rangle$, and therefore if α, β, γ, δ are interpreted respectively by $\langle x \rangle$, $\langle y \rangle$, $\langle xy \rangle$, $\langle yx \rangle$, then $(\alpha \wedge \beta) \to (\gamma \vee \delta)$ is interpreted by B.

On the other hand, if $\mathcal{F}$, $\mathcal{F}'$, $\mathcal{G}$, $\mathcal{G}'$ are sheaves with supports $\langle x \rangle$, $\langle y \rangle$, $\langle xy \rangle$ and $\langle yx \rangle$ respectively, then we have sheaf sections $\mathrm{Char}\,\langle x \rangle \to \mathcal{F}$ and $\mathrm{Char}\,\langle y \rangle \to \mathcal{F}'$ and sheaf morphisms $\mathcal{G} \to \mathrm{Char}\,\langle xy \rangle$ and $\mathcal{G}' \to \mathrm{Char}\,\langle yx \rangle$. Consequently we have a sheaf morphism $\mathrm{Hom}(\mathcal{F} \times \mathcal{F}', \mathcal{G} \sqcup \mathcal{G}') \to \mathrm{Hom}(\mathrm{Char}\,\langle x \rangle \times \mathrm{Char}\,\langle y \rangle, \mathrm{Char}\,\langle xy \rangle \sqcup \mathrm{Char}\,\langle yx \rangle)$. Therefore from 6.11, $\mathrm{Hom}(\mathcal{F} \times \mathcal{F}', \mathcal{G} \sqcup \mathcal{G}')_x = \varnothing = \mathrm{Hom}(\mathcal{F} \times \mathcal{F}', \mathcal{G} \sqcup \mathcal{G}')_y$. Thus if α, β, γ, δ are interpreted respectively by $\mathcal{F}$, $\mathcal{F}'$, $\mathcal{G}$, $\mathcal{G}'$, then $(\alpha \wedge \beta) \to (\gamma \vee \delta)$ is interpreted by a sheaf with support $\{xy, yx, \hat{1}\}$ and not B. $\qquad\square$

Example 6.18. Let us discuss the Kreisel–Putnam principle and Harrop's rule in the light of sheaf-valued models. Let us fix a valuation in a sheaf-valued model of intuitionistic logic in a space B, and set $|\beta| = \chi_{U_1}$, $|\gamma| = \chi_{U_2}$ and $|\neg\alpha| = \chi_U$. Then by 6.9(a), each $\mathrm{Hom}(\chi_U, \chi_{U_i}) \simeq \chi_{V_i}$, where $V_i = \mathrm{Int}\left(U_i \cup (X \setminus U)\right)$. On the other hand, $|\beta \vee \gamma| \simeq \chi_{U_1} \sqcup \chi_{U_2}$.

In our previous discussion of the Kreisel–Putnam principle and Harrop's rule, which was based on Tarski models (see §5.2.2), we had $B = \mathbb{R}^2$, $U_1 = \{(x, y) \mid x > 0\}$, $U_2 = \{(x, y) \mid y > 0\}$ and $U = U_1 \cup U_2$. It is not hard to see that in this case, $|\neg\alpha \to \beta \vee \gamma| = \mathrm{Hom}(\chi_U, \chi_{U_1} \sqcup \chi_{U_2})$ is isomorphic to the amalgamated union $\chi_{V_1} \cup_{\chi_V} \chi_{V_2}$, where $V = \{(x, y) \mid x < 0 \wedge y < 0\}$. The support of this amalgamated union is $\mathbb{R}^2 \setminus \{(0,0)\}$, so it does not have a global section (in fact, it does not even have a section over its support). So this example no longer works to refute Harrop's rule. However, it still applies to refute the Kreisel–Putnam principle (even though in this case $\mathrm{Supp}\,|\neg\alpha \to \beta \vee \gamma|$ coincides with $\mathrm{Supp}\,|(\neg\alpha \to \beta) \vee (\neg\alpha \to \beta)|$), since there are no sheaf morphisms $(\chi_{V_1} \cup_{\chi_V} \chi_{V_2})|_O \to (\chi_{V_1} \sqcup \chi_{V_2})|_O$ for any neighborhood O of $(0,0)$.

To refute Harrop's rule with a sheaf-valued model we can, however, take $B = \mathbb{R}^2$, $U_1 = \{(x, y) \mid x > 0 \wedge y > 0\}$, $U_2 = \{(x, y) \mid x < 0 \wedge y < 0\}$ and $U = U_1 \cup U_2$. Then

$V_1 = \{(x, y) \mid x > 0 \lor y > 0\}$ and $V_2 = \{(x, y) \mid x < 0 \lor y < 0\}$, so that the support of $|(\neg\alpha \to \beta) \lor (\neg\alpha \to \gamma)|$ is $V_1 \cup V_2 = \mathbb{R}^2 \setminus \{(0, 0)\}$. On the other hand, $|\neg\alpha \to \beta \lor \gamma| \simeq \chi_{\mathbb{R}^2}$.

Let us now analyze the difference between the two models just considered. If we write $U_0 = \mathrm{Int}((U_1 \cup U_2) \setminus (U_1 \cap U_2))$, then the sheaf morphism $\chi_{U_0} \to \chi_{U_1} \sqcup \chi_{U_2}$ induces a sheaf morphism $\mathrm{Hom}(\chi_U, \chi_{U_0}) \to \mathrm{Hom}(\chi_U, \chi_{U_1} \sqcup \chi_{U_2})$. By 6.9(a), $\mathrm{Hom}(\chi_U, \chi_{U_0}) \simeq \chi_{V_0}$, where $V_0 = \mathrm{Int}(U_0 \cup (X \setminus U))$. Thus $|\neg\alpha \to \beta \lor \gamma|$ will have a global section as long as V_0 is the entire space, and in general, $\mathrm{Supp}\,|\neg\alpha \to \beta \lor \gamma|$ contains V_0. However, V_0 need not be contained in $\mathrm{Supp}\,|(\neg\alpha \to \beta) \lor (\neg\alpha \to \gamma)| = V_1 \cup V_2$ as shown by the second example. On the other hand, if $V_0 \subset V_1 \cup V_2$, then $\mathrm{Hom}(\chi_U, \chi_{U_1} \sqcup \chi_{U_2})$ is isomorphic to the amalgamated union $\chi_{V_1} \cup_{\chi_V} \chi_{V_2}$, where $V = \mathrm{Int}(X \setminus U)$. In general, one can show that $\mathrm{Hom}(\chi_U, \chi_{U_1} \sqcup \chi_{U_2}) \simeq (\chi_{V_1} \cup_{\chi_V} \chi_{V_2}) \cup_{\chi_W} \chi_{V_0}$, where $W = V_0 \cap (V_1 \cup V_2)$.

Example 6.19. Let us show that the Negative Constant Domain principle is not valid in some sheaf-valued model. Let $\mathscr{D} = \mathbb{N}$ and let $B = (-\infty, 0] \cup N \subset \mathbb{R}$, where $N = \{\frac{1}{n} \mid n \in \mathbb{N}\}$. Then N is a regular open set in B, so $\mathcal{F} = \chi_N$ could interpret a negated formula. Let $\mathcal{F}^i = \chi_{B \setminus \{1/i\}}$. Then $\prod_i \mathcal{F}^i \simeq \chi_{(-\infty, 0)}$, so $\mathcal{F} \sqcup \prod_i \mathcal{F}^i \simeq \chi_{B \setminus \{0\}}$, which has no global section. On the other hand, each $\mathcal{F} \sqcup \mathcal{F}^i$ has a global section, hence so does their product. Thus there exist no sheaf morphisms $\prod_i (\mathcal{F} \sqcup \mathcal{F}^i) \to \mathcal{F} \sqcup \prod_i \mathcal{F}^i$.

Presheaf-valued models are not as good as sheaf-valued models:

Proposition 6.20. *Presheaf-valued models satisfy the Negative Constant Domain Principle.*

Proof. We need to construct a morphism of presheaves $\prod_i (F \sqcup F^i) \to F \sqcup \prod_i F^i$, where F is of the form $|\neg\Phi|$. Given an open set U and a collection $s = (s_i)$ of sections $s_i \in F(U) \sqcup F^i(U)$, we consider two cases. If none of the s_i actually belongs to $F(U)$, then we send s via the inclusion $\prod_i F^i(U) \subset F(U) \sqcup \prod_i F^i(U)$. If some s_i belongs to $F(U)$, we send s to the image of this s_i under the inclusion $F(U) \subset F(U) \sqcup \prod_i F^i(U)$. Since $F = |\neg\Phi|$ is the characteristic presheaf of some open set (in fact, of a regular open set), $F(U)$ contains at most one element, so the choice of i is irrelevant. This construction is clearly natural in U. $\qquad\square$

Proposition 6.21. *Presheaf-valued models satisfy the Rule of Omniscience and the Disjunction Property.*

In sheaf-valued models, every formula interpreted by the characteristic sheaf of a clopen set is decidable. Hence the Disjunction Property fails if B is disconnected; and the Rule of Omniscience can be refuted similarly to §5.5.

Proof. To prove that the Disjunction Property is valid in presheaf-valued models it suffices to note that $F \sqcup G$ has a global section if and only if at least one of the presheaves F, G has a global section.

In particular, if Φ is a formula such that $\Phi \lor \neg\Phi$ is valid with respect to some valuation ν and some assignment u in a presheaf-valued model, then either $|\Phi|$ has a global section or $|\neg\Phi|$ has a global section; the latter amounts to saying that $|\Phi| = \mathrm{Char}\,\varnothing$. If each

$|\alpha|(n)$, $n \in \mathbb{N}$, coincides with $\mathrm{Char}\,\varnothing$, then so does $|\exists x\,\alpha(x)| = \bigsqcup_n |\alpha|(n)$; and if some $|\alpha|(n)$ has a global section, then so does $\bigsqcup_n |\alpha|(n)$. Thus $\exists x\,\alpha(x) \vee \neg\exists x\,\alpha(x)$ is valid with respect to v and u as long as $\alpha(x) \vee \neg\alpha(x)$ is. $\qquad\square$

It is clear from the proofs of 6.20 and 6.21 that when the space B is connected, sheaf-valued models also satisfy the Negative Constant Domain Principle and the Disjunction Property — in contrast to Tarski models.

6.3.3. Relation to Medvedev–Skvortsov models. Let B be the two-element poset $0 < 1$ endowed with its Alexandroff topology. A presheaf (of sets) F over B such that $F(\varnothing)$ is a singleton is always a sheaf. (In general, sheafafication of presheaves over B has the effect of identifying the points of $F(\varnothing)$ with each other.)

A sheaf $\mathcal{F}$ over B consists of two stalks $\mathcal{F}_0 = \sigma\mathcal{F}(\{0,1\})$ and $\mathcal{F}_1 = \sigma\mathcal{F}(\{0\})$ and a map $\mathcal{F}_{01}\colon \mathcal{F}_0 \to \mathcal{F}_1$ corresponding to the inclusion $\{0\} \subset \{0,1\}$. We have $\mathrm{Hom}(\mathcal{F},\mathcal{G})_1 = \mathrm{Hom}(\mathcal{F}_1, \mathcal{G}_1)$, whereas $\mathrm{Hom}(\mathcal{F}, \mathcal{G})_0$ consists of commutative diagrams

$$
\begin{array}{ccc}
\mathcal{F}_0 & \xrightarrow{\ \mathcal{F}_{01}\ } & \mathcal{F}_1 \\
\downarrow & & \downarrow \\
\mathcal{G}_0 & \xrightarrow{\ \mathcal{G}_{01}\ } & \mathcal{G}_1.
\end{array}
$$

If $\mathcal{F}_{01}$ and $\mathcal{G}_{01}$ are inclusions, such a diagram amounts to a map $\varphi\colon \mathcal{F}_0 \to \mathcal{G}_0$ such that $\varphi(\mathcal{F}_1) \subset \mathcal{G}_1$.

In the case where the restriction maps of the sheaves are inclusions, by 6.13 we further have $\left((\bigsqcup_i \mathcal{F}^i)_1, (\bigsqcup_i \mathcal{F}^i)_0\right) = (\bigsqcup_i (\mathcal{F}^i)_1, \bigsqcup_i (\mathcal{F}^i)_0)$ and, using that B is an Alexandroff space, also $\left((\prod_i \mathcal{F}^i)_1, (\prod_i \mathcal{F}^i)_0\right) = (\prod_i (\mathcal{F}^i)_1, \prod_i (\mathcal{F}^i)_0)$.

This is reminiscent of the interpretation of connectives and quantifiers in Medvedev–Skvortsov models (see §3.13). To recover an actual version of Medvedev–Skvortsov models, let us fix a domain $\mathscr{D}$, and one set X_{rk} corresponding to each problem variable $\varkappa_k^{0^r \to \mathbb{1}}$. Let $X = \prod_{r \in \omega} \prod_{k \in \mathbb{N}} \mathrm{Hom}(\mathscr{D}^r, 2^{X_{rk}})$, where $\omega = \{0, 1, \dots\}$ and $\mathbb{N} = \{1, 2, \dots\}$. Here 2^S denotes the set of all subsets of S and $\mathrm{Hom}(Y, Z)$ denotes the set of all maps $Y \to Z$. Let us consider the Alexandroff space $X^+ = X \cup \{\hat{1}\}$, where X is regarded as a discrete space, or alternatively as a poset with no comparable pair of elements, and $\hat{1}$ is an additional element that is greater than all elements of X_Φ. Each $\vec{Y} = (Y_{rk}) \in X$ determines the two-element subspace $\{\vec{Y}, \hat{1}\}$ of X^+, which is homeomorphic to B.

Let us interpret the kth r-ary problem variable, $\varkappa_k^{0^r \to \mathbb{1}}$, by the following r-ary family $\mathcal{F}^{rk}(t_1, \dots, t_r)$ of sheaves over X^+. Let $\mathcal{F}^{rk}(\vec{t})_{\hat{1}} = X_{rk}$ and $\mathcal{F}^{rk}(\vec{t})_{\vec{Y}} = Y_{rk}(\vec{t})$, and let $\mathcal{F}^{rk}(\vec{t})_{\vec{Y}\hat{1}}\colon Y_{rk}(\vec{t}) \to X_{rk}$ be the inclusion map.

Let connectives and quantifiers be interpreted as in a sheaf-valued model, except for $\times$, which is interpreted by the sheaf $\mathcal{F}$ with $\mathcal{F}_{\vec{Y}} = \varnothing$ for each $\vec{Y} \in X$ and $\mathcal{F}_{\hat{1}} = \{\varnothing\}$. As long as all the sets X_{rk} are non-empty, it is easy to see that this yields a model of intuitionistic logic with respect to the valuation field generated by the above valuation.

Given a closed formula Φ, let $\mathbf{x}_{k_i}^{\mathbb{0}^{r_i} \to \mathbb{1}}$, $i = 1, \ldots, m$, be all the problem variables that occur in Φ. Let $X_\Phi = \prod_{i=1}^m \mathrm{Hom}(\mathscr{D}^{r_i}, 2^{X_{r_i k_i}})$ and let $\pi_\Phi \colon X \to X_\Phi$ be the projection onto the subproduct, $\vec{Y} \mapsto (Y_{r_1 k_1}, \ldots, Y_{r_m k_m})$.

Then Φ is interpreted by the sheaf $\mathcal{G}$, where, in the notation of §**3.13**, the stalk $\mathcal{G}_{\hat{1}} = S$ and the stalk $\mathcal{G}_{\vec{Y}'} = T(\vec{Y})$ for each $\vec{Y} \in X_\Phi$ and each $\vec{Y}' \in \pi_\Phi^{-1}(\vec{Y})$. Thus the set of global sections of $\mathcal{G}$ can be identified with $\bigcap_{\vec{Y} \in X_\Phi} T(\vec{Y}) \subset S$, and in particular $\mathcal{G}$ has a global section if and only if this intersection is nonempty.

Remark 6.22. By adding the least element $\hat{0}$ to the poset X^+ and considering sheaves over the resulting Alexandroff space $X^\pm = X \cup \{\hat{0}, \hat{1}\}$ one can also model a modification of Medvedev–Skvortsov models (a non-equivariant version of Läuchli models [120]) where all X_{ij} are taken to be a fixed set U, and $\times$ is interpreted not by $(\{\varnothing\}, \varnothing)$, but by (U, Z) for some fixed $Z \subset U$ (possibly nonempty!), and only Y_{rk} containing Z are considered.

6.3.4. Modified BHK interpretation. Kolmogorov's interpretation of $\cdot\gamma \vee \neg\gamma$ as "the problem of finding a general method of solving the problem $\Gamma \vee \neg\Gamma$ for every contentful problem Γ" contains *constructive* quantification over all problems; it does not seem to be compatible with usual, Tarski-style model theory, which only deals with *classical* quantification over all valuations. (It is compatible with a certain alternative model theory constructed below; see §*7.3.6*.) The following modification of Kolmogorov's interpretation overcomes this difficulty by replacing the informal (i.e., contentual) constructive quantifier with a combination of an informal classical quantifier (enabling compatibility with standard model theory) and a potentially formal constructive quantifier (preventing intuitionistic logic from degeneration into classical). Related ideas are also found in [122; Definition 4.2] (see also [167; Ch. 7]).

In addition to a domain of discourse $\mathscr{D}$, we fix a "hidden parameter" domain $\mathscr{E}$ (just like in the modified Platonist interpretation of classical logic). Now $\mathcal{O}$ is taken to be a class of problems with parameter in $\mathscr{E}$, consisting of a prescribed class of contentful (e.g. mathematical) primitive parametric problems and of composite parametric problems, which are obtained inductively from the primitive ones by using contentual intuitionistic connectives and quantifiers over arbitrary $\mathscr{D}$-indexed families of parametric problems. For instance, if $\mathscr{E} = P \times P$, where P is the set of all finite presentations of groups, primitive parametric problems could be of the form "Find an isomorphism between the groups given by presentations p and q". In this case composite parametric problems would include "Find an isomorphism between the groups given by p and q or find a proof that it is impossible". Formal intuitionistic connectives and quantifiers are interpreted straightforwardly by the contentual ones.

The value of $! \colon \mathcal{O} \to \mathcal{Q} = \{\top, \bot\}$ on a parametric problem Θ will be $\top$ if and only if there exists a general method of solving the problem $\Theta(e)$ for all $e \in \mathscr{E}$. What is meant by "solving" $\Theta(e)$ is defined explicitly for primitive $\Theta \in \mathcal{O}$ and all $e \in \mathscr{E}$ and is extended inductively to composite $\Theta \in \mathcal{O}$ and all $e \in \mathscr{E}$ by the usual BHK interpretation (see §**3.7**), separately for each $e \in \mathscr{E}$. This determines a contentual interpretation of intuitionistic logic and its meta-logic, which we will call the *modified BHK* interpretation.

Let us note that the arguments of §**3.12** and §**3.13** still work to show that $\cdot\gamma \vee \neg\gamma$ is not valid under the modified BHK interpretation.

Example 6.23. Let us record the modified BHK interpretations of the judgements of admissibility, stable admissibility and derivability of a purely logical rule $\Gamma \,/\, \Delta$. The tuples $\vec{\varphi}$ and $\vec{\Phi}$ have the same meaning as in Example 4.84, and e is again understood to run over $\mathscr{E}$ as in Example 4.85. Like before, $?\Theta$ denotes the proposition "The problem Θ has a solution".

- Derivability: $\forall \mathscr{D}, \mathscr{E}, \mathscr{O} \; \forall v \; \big(\forall u \; ?\forall e \,|\Gamma|^v_u(e) \to \forall u \; ?\forall e \,|\Delta|^v_u(e)\big)$;
- Stable admissibility:
  $\forall \vec{\Phi} \, \forall \mathscr{D}, \mathscr{E}, \mathscr{O} \; \big(\forall v, u \; ?\forall e \,|\Gamma[\vec{\varphi}/\vec{\Phi}]|^v_u(e) \; \longrightarrow \; \forall v, u \; ?\forall e \,|\Delta[\vec{\varphi}/\vec{\Phi}]|^v_u(e)\big)$;
- Admissibility:
  $\forall \vec{\Phi} \, \big(\forall \mathscr{D}, \mathscr{E}, \mathscr{O} \; \forall v, u \; ?\forall e \,|\Gamma[\vec{\varphi}/\vec{\Phi}]|^v_u(e) \; \longrightarrow \; \forall \mathscr{D}, \mathscr{E}, \mathscr{O} \; \forall v, u \; ?\forall e \,|\Delta[\vec{\varphi}/\vec{\Phi}]|^v_u(e)\big)$.

Remark 6.24. Let us observe that sheaf-valued models provide a formalization of the modified BHK semantics.

7. Semantics of the meta-logic

7.1. Generalized models

All the models considered so far are based on classical, moreover, two-valued interpretation of the meta-logic, in accordance with traditional model theory. But the meta-logic is intuitionistic, and so is severely collapsed in such models.

Given a logic L based on a language $\mathscr{L}$ and given by a derivation system $\mathcal{D}$, a meta-$\mathscr{L}$-structure M will be called a *generalized model* (resp. a *semi-generalized model*) of L if it is a model (resp. a semi-standard model) of the meta-logic (see §*4.7.1*) and the meta-formula $\mathcal{D}$ is valid in M. When M is an $\mathscr{L}$-structure (i.e., the interpretation of the meta-logic is two-valued), this is the same as a model in the usual sense (see §*4.7.3*).

7.1.1. Generalized Tarski models. Let $\mathscr{D}$ be a set, let X and Y be topological spaces, and let $f\colon X \to Y$ be a continuous map. Let $\mathscr{O}$ be the collection of all open sets in X, $\mathcal{Q}$ the collection of all open sets in Y, and $!\colon \mathscr{O} \to \mathcal{Q}$ be defined by $!(U) = \mathrm{Int}\big(Y \setminus f(X \setminus U)\big)$.[85] Let us interpret the intuitionistic connectives and quantifiers in X in the usual way and the meta-connectives and the first-order meta-quantifier in Y in the similar way (that is, in the way the corresponding connectives and the universal quantifier are interpreted in the Tarski model in Y with domain $\mathscr{D}$). The second-order meta-quantifier $\mathfrak{q}^n$ is also interpreted similarly; namely, its interpretation $|\mathfrak{q}^n|$ associates to a $\mathrm{Hom}(\mathscr{D}^n, \mathscr{O})$-indexed family of open sets in Y the interior of their intersection. Let $?\colon \mathcal{Q} \to \{\top, \bot\}$ be defined by sending Y to $\top$ and any its proper subset to $\bot$.

This yields a meta-structure for the language of intuitionistic logic, which will be called the *generalized Tarski meta-structure* in the map $f\colon X \to Y$ with domain $\mathscr{D}$.

---

[85]Let us note that if X is compact, or more generally f is a closed map, then $Y \setminus f(X \setminus U)$ is open.

Theorem 7.1. *Generalized Tarski meta-structures are semi-generalized models of intuitionistic logic.*

Proof. It is clear that all the laws of the Hilbert-style derivation system of intuitionistic logic (see §*4.6.2*) are valid in this meta-structure (because they are valid in the two-valued structure given by the Tarski model in X with domain $\mathscr{D}$), and that all the meta-laws of the Hilbert-style formulation of the meta-logic (see §*4.4.3*) are also valid in it (because they are similar to the said laws of intuitionistic logic). Also, the $\Rightarrow$-elimination meta-rule and the conditions (A) and (B) of §*4.7.1* are satisfied for the same reasons that the *modus ponens* and the generalization rule of intuitionistic logic are valid in Tarski models. Thus it remains to check that the *modus ponens* and the generalization rule of intuitionistic logic are valid in our meta-structure.

Let us check the validity of the *modus ponens* rule. Given two open sets $U, V \subset X$, let $W = \text{Int}\big(V \cup (X \setminus U)\big)$. Since $\cdot\alpha \wedge (\alpha \to \beta) \to \beta$ is derivable in intuitionistic logic (or by direct reasoning), $U \cap W \subset V$. Then $(X \setminus U) \cup (X \setminus W) = X \setminus (U \cap W) \supset X \setminus V$. Hence $f(X \setminus U) \cup f(X \setminus W) = f\big((X \setminus U) \cup (X \setminus V)\big) \supset f(X \setminus V)$. Since $\text{Int}\, A \cap \text{Int}\, B = \text{Int}(A \cap B)$, we get $!(U) \cap !(W) \subset !(V)$.

Let us check the validity of the generalization rule. Given a family of open sets U_x, $x \in \mathscr{D}$, we need to show that $\text{Int} \bigcap_{x \in \mathscr{D}} !(U_x) \subset !(\text{Int} \bigcap_{x \in \mathscr{D}} U_x)$. Since $\text{Int} \bigcap_x \text{Int}\, S_x = \text{Int} \bigcap_x S_x$ (see §*5.7.3*), and the left hand side is open, this is equivalent to

$$\text{Int} \bigcap_{x \in \mathscr{D}} \big(Y \setminus f(X \setminus U_x)\big) \subset Y \setminus f\left(X \setminus \text{Int} \bigcap_{x \in \mathscr{D}} U_x\right).$$

Since the complement to the interior is the closure of the complement, and the complement to an intersection is the union of the complements, this is equivalent to

$$Y \setminus \text{Cl} \bigcup_{x \in \mathscr{D}} f(X \setminus U_x) \subset Y \setminus f\left(\text{Cl} \bigcup_{x \in \mathscr{D}} (X \setminus U_x)\right).$$

But since the image of a union is the union of the images, the latter is the same as

$$\text{Cl}\, f\left(\bigcup_{x \in \mathscr{D}} (X \setminus U_x)\right) \supset f\left(\text{Cl} \bigcup_{x \in \mathscr{D}} (X \setminus U_x)\right),$$

which follows from the continuity of f. (In fact, $f \colon X \to Y$ is continuous if and only if $f(\text{Cl}\, A) \subset \text{Cl}\, f(A)$ for each $A \subset X$.) $\square$

Example 7.2. Let c be the constant map from a topological space X to a point. It is easy to see that the generalized Tarski model of intuitionistic logic in c with domain $\mathscr{D}$ is nothing but the usual Tarski model in X with domain $\mathscr{D}$.

Example 7.3. If X and Y are topological spaces and X is discrete, any map $f \colon X \to Y$ is continuous. Also, since usual Tarski models in X satisfy the principle $\cdot\alpha \vee \neg\alpha$, this principle is also valid in generalized Tarski models in f. So they are in fact semi-generalized models of classical logic.

Example 7.4. If Y is a topological space and Y^d is the same underlying set re-endowed with the discrete topology, the previous example (Example 7.3) applies to the identity map $f\colon Y^d \to Y$. Let us note that $!$ is interpreted by $S \mapsto \operatorname{Int} S$ in this case.

Example 7.5. As a special case of the previous example (Example 7.4), we have semi-generalized models of classical logic in any set X, with $\mathscr{O} = \mathscr{Q} = 2^Y$ and $! = \mathrm{id}$.

Example 7.6. There exists a semi-generalized model W of classical logic with respect to a valuation field, and principles $\cdot F$ and $\cdot G$ such that neither $\cdot F$ nor $\cdot G$ nor $\cdot F \Rightarrow \cdot G$ is valid in W.

Indeed, let us consider $X = \{\mu, \nu\}$ and let $F = \neg\big(\exists x\, p(x) \wedge \exists x\, \neg p(x)\big)$ and $G = \bot$. The usual Leibniz–Euler model V in X with respect to a certain valuation field K, constructed in Example 4.83, yields a semi-generalized Leibniz–Euler model W in X with respect to K by re-interpreting $!$ and the meta-connectives and the meta-quantifiers as in Example 7.5. Since $|\neg F|_V^v$ does not contain ν for each v in K, but contains μ for some v in K, $|F|_V^v$ contains ν for each v in K and does not contain μ for some v in K. Hence $|!F|_W^v$ also contains ν for each v in K and does not contain μ for some v in K. Consequently $|\cdot F|_W^K = \{\nu\}$. Then $|\cdot F \Rightarrow \bot|_W^K = \{\mu\}$. Thus neither $\cdot F$ nor $\cdot F \Rightarrow \bot$ is valid in W with respect to K.

7.1.2. Tarski/sheaf-valued models. Let us note a slightly different type of generalized models. Let $\mathscr{D}$ be a set, let X and Y be topological spaces, and let $f\colon X \to Y$ be a continuous map. Let $\mathscr{O}$ be the collection of all open sets in X. Let $\mathscr{Q}$ the collection of sheaves on Y. Let $!\colon \mathscr{O} \to \mathscr{Q}$ be defined by setting $!(U)$ to be the characteristic sheaf of the open set $\operatorname{Int}\big(Y \setminus f(X \setminus U)\big)$. Let us interpret the intuitionistic connectives and quantifiers in X in the usual way (as in a Tarski model) and the meta-connectives and the meta-quantifiers in Y in the similar way (as in a sheaf-valued model). Let us interpret the intuitionistic connectives and quantifiers in X in the usual way (as in a Tarski model) and the meta-connectives and the first-order meta-quantifier in Y in the similar way (as in a sheaf-valued model). The second-order meta-quantifier $\mathfrak{q}^n$ is also interpreted similarly; namely, its interpretation $|\mathfrak{q}^n|$ associates to a $\operatorname{Hom}(\mathscr{D}^n, \mathscr{O})$-indexed family of sheaves on Y their product (in the category of sheaves). Let $?\colon \mathscr{Q} \to \{\top, \bot\!\!\!\bot\}$ be defined by sending Y to $\top$ and any its proper subset to $\bot\!\!\!\bot$.

This yields a meta-structure for the language of intuitionistic logic. The arguments of the previous subsection (§7.1.1) apply without changes to show that this is a semi-generalized model of intuitionistic logic. Indeed, the verification of the *modus ponens* and generalization rules really involves only products of characteristic sheaves (a product of two of them and a $\mathscr{D}$-indexed product) which are the same as the characteristic sheaves of the intersection. Let us call this semi-generalized model of intuitionistic logic the *Tarski/sheaf-valued model* in the map $f\colon X \to Y$ with domain $\mathscr{D}$.

Example 7.7. If X is a topological space and X^d is the same underlying set re-endowed with the discrete topology, the identity map $f\colon X^d \to X$ is continuous. But since usual Tarski models in X^d satisfy the principle $\cdot \alpha \vee \neg \alpha$, this principle is also valid in

Tarski/sheaf-valued models in f. So they are in fact semi-generalized models of classical logic. Let us note that $!$ is interpreted by $S \mapsto \chi_{\mathrm{Int}\,S}$ in this case.

7.1.3. Verificationist interpretation. We fix a domain of discourse $\mathscr{D}$. Like in the Platonist interpretation, $\mathscr{O}$ is taken to be a class of propositions, consisting of a prescribed class of contentful (e.g. mathematical) primitive propositions and composite propositions, which are obtained inductively from the primitive ones by using contentual classical connectives and quantifiers over arbitrary $\mathscr{D}$-indexed families of propositions. Formal classical connectives and quantifiers are interpreted straightforwardly by the contentual ones. We also have the notion of "truth", defined explicitly for primitive $P \in \mathscr{O}$, and extended inductively to composite $P \in \mathscr{O}$ by the usual truth tables.

Now $\mathscr{Q}$ is taken to be a class of problems, consisting of primitive problems of the form $!(P) = $ "Prove P", which are the values of $!\colon \mathscr{O} \to \mathscr{Q}$, and of composite problems, which are obtained inductively from the primitive ones by using contentual intuitionistic connectives and quantifiers over arbitrary $\mathscr{D}$-indexed and $\mathrm{Hom}(\mathscr{D}^i, \mathscr{O})$-indexed ($i = 0, 1, \dots$) families of problems.

The value of $?\colon \mathscr{Q} \to \{\top, \bot\}$ on a problem Θ will be $\top$ if and only if Θ has a solution. To "solve" a primitive $\Theta = !(P)$ means to prove that P is true; this notion of "solving" is extended inductively to composite $\Theta \in \mathscr{Q}$ by the usual BHK interpretation (see §**3.7**). This determines a contentual interpretation of classical logic and its meta-logic, which we will call the *modified Verificationist* interpretation.

The Verificationist interpretation is very close to the sketch of an interpretation in §*3.10.3*, which however suggests fully classical interpretation of the meta-logic.

Example 7.8. Let us record the modified Verificationist interpretations of the judgements of admissibility, stable admissibility and derivability of a purely logical rule $F\,/\,G$. The tuples $\vec{\varphi}$ and $\vec{\Phi}$ have the same meaning as in Example 4.84, and e is again understood to run over $\mathscr{E}$ as in Example 4.85. Like before, $?\Theta$ denotes the proposition "The problem Θ has a solution", and $!P$ denotes the problem "Prove the proposition P".

- Derivability: $\forall \mathscr{D}, \mathscr{E}, \mathscr{O}\ ?\forall v \left(\forall u\ !|F|_u^v \to \forall u\ !|G|_u^v\right)$;
- Stable admissibility:
  $$\forall \vec{\Phi}\, \forall \mathscr{D}, \mathscr{E}, \mathscr{O}\ ?\!\left(\forall v, u\ !|F[\vec{\varphi}/\vec{\Phi}]|_u^v \ \longrightarrow\ \forall v, u\ !|F[\vec{\varphi}/\vec{\Phi}]|_u^v\right);$$
- Admissibility:
  $$\forall \vec{\Phi}\left(\forall \mathscr{D}, \mathscr{E}, \mathscr{O}\ ?\forall v, u\ !|F[\vec{\varphi}/\vec{\Phi}]|_u^v \ \longrightarrow\ \forall \mathscr{D}, \mathscr{E}, \mathscr{O}\ ?\forall v, u\ !|F[\vec{\varphi}/\vec{\Phi}]|_u^v\right).$$

Remark 7.9. The Verificationist interpretation can be straightforwardly generalized using a "hidden parameter" domain $\mathscr{E}$, so that $\mathscr{O}$ consists of predicates on $\mathscr{E}$ (similarly to the modified Platonist interpretation) and $\mathscr{Q}$ consists of parametric problems with parameter in $\mathscr{E}$ (similarly to the modified BHK interpretation).

Let us note that this modified Verificationist semantics is formalized by the generalized models of Example 7.7.

7.1.4. Absolute realizability. In 1945, Kleene introduced the notion of realizability for arithmetical formulas, which is weaker than provability in intuitionistic arithmetic (see

[103; §82], [161]). A parallel notion of realizability for predicate formulas, called "abstract realizability", was introduced by Plisko in the 1970s (see [160], [161]). However, traditionally, realizability for predicate formulas has been more extensively studied in an indirect way, by substituting arithmetical formulas for predicate variables, which leads to several inequivalent definitions of realizability for predicate formulas (see [161; §4, §8] for a comparison of these definitions with each other and [209; §1, Assertion 1 and subsequent remarks][86] concerning their relationship with abstract realizability).

All notions of realizability for predicate formulas are weaker than provability in intuitionistic logic. Thus one can consider them as models of intuitionistic logic. However, these are not models in the sense of the traditional, Tarski-style model theory.

We will now see that abstract realizability can in fact be regarded as a generalized model of intuitionistic logic.

Let $\mathscr{D} = \mathbb{N}$ and $\mathscr{O} = 2^{\mathbb{N}}$, the set of all subsets of $\mathbb{N}$. The connectives and quantifiers are interpreted as follows. For $m, n \in \mathbb{N}$ let $\#(x, y) = 2^m \cdot 3^n$. Let $\mathfrak{R}$ denote the set of all (partial) recursive functions $\mathbb{N} \supset \mathrm{Dom}\, f \xrightarrow{f} \mathbb{N}$, and for any $f \in \mathfrak{R}$ let $\#f \in \mathbb{N}$ be the Gödel number of f.[87] Then for any subsets $X, Y \subset \mathbb{N}$ and any $P \colon \mathbb{N} \to 2^{\mathbb{N}}$ we set:

- $X \,|\vee|\, Y = \{\#(0, x) \mid x \in X\} \cup \{\#(1, x) \mid y \in Y\}$;
- $X \,|\wedge|\, Y = \{\#(x, y) \mid x \in X,\, y \in Y\}$;
- $X \,|\to|\, Y = \{\#f \mid f \in \mathfrak{R},\, X \subset \mathrm{Dom}\, f \text{ and } f(X) \subset Y\}$;
- $|\times| = \varnothing$;
- $|\exists|(P) = \{\#(x, y) \mid y \in P(x)\}$;
- $|\forall|(P) = \{\#f \mid \mathrm{Dom}\, f = \mathbb{N} \text{ and } f(x) \in P(x) \text{ for each } x \in \mathbb{N}\}$.

Let us note that $\mathrm{Dom}\, f$ may be strictly larger than X in the definition of $X \,|\to|\, Y$.[88] Thus, in fact, $|\times \to \times|$ consists of *all* partial recursive functions.

It should be noted that this interpretation of the connectives is nothing but a rather computational reading of the BHK interpretation: the set of solutions of every problem is identified with a set of natural numbers, and "general method" is interpreted as "algorithm". This at once guarantees that, in the notation of §3.9, $\mathscr{F}$ is neither injective nor surjective, and $\mathscr{G}$ is not surjective.

Then it is natural (cf. the meta-clarified BHK interpretation in §7.3.6 below) to set $\mathcal{Q} = \mathscr{O} = 2^{\mathbb{N}}$, $! = \mathrm{id} \colon \mathscr{O} \to \mathcal{Q}$, and interpret the meta-connectives and the first-order meta-quantifier similarly to the corresponding connectives and the universal quantifier. Thus for any subsets $X, Y \subset \mathbb{N}$ and any $P \colon \mathbb{N} \to 2^{\mathbb{N}}$ and $Q_n \colon \mathrm{Hom}(\mathbb{N}^n, 2^{\mathbb{N}}) \to 2^{\mathbb{N}}$ we set

- $X \,|\&|\, Y = X \,|\wedge|\, Y$;
- $X \,|\Rightarrow|\, Y = X \,|\to|\, Y$;
- $|\mathtt{q}|(P) = |\forall|(P)$;
- $|\mathtt{q}_n|(Q_n) = \bigcap_{P \colon \mathbb{N}^n \to 2^{\mathbb{N}}} Q_n(P)$.

---

[86]Not found in the 2002 conference proceedings paper by the same authors with the same title.

[87]It is essential here that f need not be total. The set of all total recursive functions is not recursively enumerable.

[88]This freedom is essential. Domains of recursive functions coincide with recursively enumerable sets.

Here the second-order meta-quantifier cannot be interpreted directly similarly to the first-order one, since the set $\mathrm{Hom}(\mathbb{N}^n, 2^{\mathbb{N}})$ is uncountable, so we cannot speak of recursive functions defined on this set (as long as "recursive" is understood in the original sense). However, a constant function would arguably qualify as "recursive"; so one straightforward option is to be content with constant functions; this yields the above definition of $|\mathbb{q}_n|$. Finally, it is natural to define $?\colon \mathcal{Q} \to \{\top, \bot\}$ by $?(X) = \top$ if and only if $X \neq \varnothing$.

Thus for a closed formula Φ, if $\vec{\gamma}$ is the tuple of all problem variables that occur in Φ, and $r_1, \ldots, r_m$ are their arities, then we have $?(|\cdot \Phi|) = \top$ if and only if there exists a natural number e such that $e \in |\vec{\gamma} \mapsto \Phi|(|\vec{\gamma}|)$ for all tuples $|\vec{\gamma}|$ of functions $|\gamma_i|\colon \mathbb{N}^{r_i} \to 2^{\mathbb{N}}$. This is the precisely one of the two equivalent definitions ("absolute uniform realizability") of Φ being abstractly realizable in the sense of Plisko [160], [161].

For this reason we will call the meta-structure just described the *abstract realizability meta-structure*.

Theorem 7.10. *The abstract realizability meta-structure is a generalized model of intuitionistic logic.*

Proof. It is well-known that intuitionistically valid formulas are abstractly realizable (see [160], [161]), so all the laws of the Hilbert-style derivation system of intuitionistic logic (see §4.6.2) must be valid in this meta-structure. It follows that the meta-laws of the Enderton-style formulation of the meta-logic (see §4.4.4) that do not involve the second-order quantifier are also valid in it.

There remain three meta-laws involving the second-order quantifier. Let us check the validity of $(\vec{z})(\vec{\beta})\big((\gamma)(\mathcal{F} \Rightarrow \mathcal{G}) \Rightarrow ((\gamma)\mathcal{F} \Rightarrow (\gamma)\mathcal{G})\big)$. Given an individual assignment u and a predicate valuation v, the set $|\mathcal{F} \Rightarrow \mathcal{G}|^{uv}$ consists of the Gödel numbers of all total recursive functions f such that $f(|\mathcal{F}|^{uv}) \subset |\mathcal{G}|^{uv}$. By definition, $|(\gamma)\mathcal{H}|^{uv}$ is the intersection of the sets $|\mathcal{H}|^{uv'}$ over all $v' \in V$, where V consists of all predicate valuations that agree with v on all predicate variables except possibly γ. Thus $|(\gamma)(\mathcal{F} \Rightarrow \mathcal{G})|^{uv}$ consists of the Gödel numbers of all total recursive f such that $f(|\mathcal{F}|^{uv'}) \subset |\mathcal{G}|^{uv'}$ for every $v' \in V$. This is clearly a subset of the set $|(\gamma)\mathcal{F} \Rightarrow (\gamma)\mathcal{G}|^{uv}$ of the Gödel numbers of all total recursive f such that $f(\bigcap_{v' \in V} |\mathcal{F}|^{uv'}) \subset \bigcap_{v' \in V} |\mathcal{G}|^{uv'}$. Thus $|(\gamma)(\mathcal{F} \Rightarrow \mathcal{G}) \Rightarrow (\mathcal{F} \Rightarrow (\gamma)\mathcal{G})|^{uv}$ contains the Gödel number of the identity function. Consequently so does $|(\vec{z})(\vec{\beta})\big((\gamma)(\mathcal{F} \Rightarrow \mathcal{G}) \Rightarrow ((\gamma)\mathcal{F} \Rightarrow (\gamma)\mathcal{G})\big)|^{uv}$.

The validity of the two remaining meta-laws involving the second-order quantifier, $(\vec{z})(\vec{\beta})\big((\gamma)\mathcal{F} \Rightarrow \mathcal{F}[\gamma/\Phi]\big)$, provided that Φ is free for γ in $\mathcal{F}$; and $(\vec{z})(\vec{\beta})\big(\mathcal{F} \Rightarrow (\gamma)\mathcal{F}\big)$, provided that γ does not occur freely in $\mathcal{F}$, is checked similarly, also using the Gödel number of the identity function.

The validity of the $\Rightarrow$-elimination meta-rule is easy: if $|\mathcal{F}|^{uv}$ is nonempty and there exists a total recursive function f such that $f(|\mathcal{F}|^{uv}) \subset |\mathcal{G}|^{uv}$, then $|\mathcal{G}|^{uv}$ must be nonempty as well.

It remains to check the validity of the two object-level rules: the generalization rule $(\alpha)\big((\mathsf{x})\,\alpha(\mathsf{x}) \Rightarrow \forall \mathsf{x}\,\alpha(\mathsf{x})\big)$ and the *modus ponens* $(\alpha)(\beta)\big(\alpha\ \&\ \alpha \to \beta \Rightarrow \beta\big)$, or equivalently

(using the exponential meta-law) $(\alpha)(\beta)\big(\alpha \to \beta \Rightarrow (\alpha \Rightarrow \beta)\big)$. But these are trivial, since $|\!\to\!| = |\!\Rightarrow\!|$ and $|\mathfrak{q}| = |\forall|$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

7.2. Alternative semantics

7.2.1. Kolmogorov's semantic consequence. The interpretations considered above are traditional, "Tarski-style" interpretations, which are based, in particular, on Tarski's concept of semantic consequence [193]; and the meta-interpretations considered above can be called "generalized Tarski-style".

We will next do something different, and consider alternative, "Kolmogorov-style" interpretations, which are based, in particular, on Kolmogorov's concept of semantic consequence, found in his problem interpretation of intuitionistic logic [108].

Kolmogorov's problem interpretation explains not only logical, but, in disguise, also meta-logical connectives and quantifiers by the same means of solutions of problems, and it is this coincidence of $\mathcal{O}$ and $\mathcal{Q}$ that makes Kolmogorov's concept of semantic consequence particularly apposite to his context. One might argue that the meta-logical constants should rather be interpreted by means of meta-solutions of meta-problems, so that the issue of coincidence of $\mathcal{O}$ with $\mathcal{Q}$ does not arise. But what are the meta-problems, and where does one find them? In fact, Kolmogorov's notion of semantic consequence yields a nice description of meta-problems in terms of problems (see §*7.3.6*) — and one can hardly dream of anything simpler. It also applies to yield a description of meta-propositions in terms of propositions (see §*7.2.5*).

Namely, Kolmogorov offered the following interpretation of the rules used by Heyting in his original axiomatization of intuitionistic logic: $\mathfrak{p} \wedge \mathfrak{q} \,/\, \mathfrak{p}$, the *modus ponens* rule $\mathfrak{p}, \mathfrak{p} \to \mathfrak{q} \,/\, \mathfrak{q}$ and the (non-structural) substitution rule — or rather of the assumption that these are inference rules of the logic:[89]

> "[W]e assume that we are always able (i.e. we possess a general method) to solve, for any given elementary problem functions [=formulas of zero-order intuitionistic logic regarded as functions of their problem variables] $p, q, r, s, \ldots$, the following problems:
> I. If $\cdot(p \wedge q)$ is solved, to solve $\cdot p$;
> II. If $\cdot p$ and $\cdot(p \to q)$ are solved, to solve $\cdot q$;
> III. If $\cdot p(a, b, c, \ldots)$ is solved, to solve $\cdot p(q, r, s, \ldots)$."

Remark 7.11. It should be noted that Kolmogorov's "If ... is solved, to solve ..." is apparently meant to be the same as his interpretation of object-level implication: "$a \to b$ is the problem 'assuming that a solution to a is given, solve b', or, which means the same, 'reduce the solution of b to the solution of a'."

But in our presentation of the BHK interpretation in §**3.7**, we follow Heyting's more nuanced reading of $a \to b$ as the problem "find a general method to transform the given

---

[89]Kolmogorov's formulas are given here in a modern notation (replacing $\supset$ with $\to$ and Peano's dots with the usual parentheses); also, Kolmogorov's symbol "$\vdash$" is replaced with "$\cdot$" since his interpretation of $\vdash$ is precisely the same as our interpretation of $\cdot$ (and different from our interpretation of $\vdash$).

solution of a into a solution of b". Hence it is natural to apply this more nuanced reading here as well.

It is not hard to see that the same approach applies to an arbitrary rule just as well, including those in first-order logic. So here is the natural Kolmogorov-style interpretation of a rule

$$\frac{\Gamma_1, \ldots, \Gamma_k}{\Gamma_0}$$

over intuitionistic logic.

(1) Suppose that all variables that occur freely in the formulas Γ_i form a p-tuple $\vec{t}$ of individual variables and an s-tuple $\vec{\varphi}$ of problem variables of arities $q_1, \ldots, q_s$. Then the rule $\Gamma_1, \ldots, \Gamma_k \, / \, \Gamma_0$ is interpreted as the problem of finding a general method of solving a certain problem

$$\left| \cdot \Gamma_1[\vec{t}/\vec{T}, \, \vec{\varphi}/\vec{\Phi}] \, \& \cdots \& \, \cdot \Gamma_k[\vec{t}/\vec{T}, \, \vec{\varphi}/\vec{\Phi}] \, \Rightarrow \, \cdot \Gamma_0[\vec{t}/\vec{T}, \, \vec{\varphi}/\vec{\Phi}] \right|$$

for any given p-tuple $\vec{T}$ of terms and any given s-tuple $\vec{\Phi}$ of q_i-formulas.

(2) A solution of the problem $|\cdot \Delta_1 \& \cdots \& \cdot \Delta_k \Rightarrow \cdot \Delta_0|$ is in turn a general method of transforming any given solutions of $|\cdot \Delta_1|, \ldots, |\cdot \Delta_k|$ into a solution of $|\cdot \Delta_0|$.

(3) Suppose that all variables that occur freely in the formulas $\Delta_i = \Gamma_i[\vec{t}/\vec{T}, \, \vec{\varphi}/\vec{\Phi}]$ form an n-tuple $\vec{x}$ of individual variables and an m-tuple $\vec{\theta}$ of problem variables of arities $r_1, \ldots, r_m$. Then a solution of each $|\cdot \Delta_i|$ is a general method of solving the contentful problem $|\vec{x}, \vec{\theta} \mapsto \Delta_i|(\vec{X}, \vec{\Theta})$ for any n-tuple $\vec{X}$ of elements of the domain and any m-tuple $\vec{\Theta}$ of contentful problems of arities $r_1, \ldots, r_m$.

(4) What is a solution of each contentful problem $|\vec{x}, \vec{\theta} \mapsto \Delta_i|(\vec{X}, \vec{\Theta})$ is determined by the usual BHK interpretation (see §**3.7**).

An obvious issue with this interpretation of rules is that step (1) applies what in essence is constructive quantification ("to find a general method") to syntactic entities ("for any given formulas and terms"), thereby creating a dangerous mix of syntax and semantics. (Perhaps this was not such an obvious issue back in 1932, when Kolmogorov published his interpretation of rules [108] but, for instance, Tarski's paper [193] on semantic consequence had not appeared yet.)

But it turns out that this mix of syntax and semantics can be disentangled. What is important in Kolmogorov's interpretation of rules is that the formulas $\vec{\Phi}$ and terms $\vec{T}$ being quantified over are treated as functions of the problem variables $\vec{\theta}$ and the individual variables $\vec{x}$. But such a functional dependence can well be understood on the semantic level.

7.2.2. Alternative semantics. Given a Tarski-style meta-interpretation $\jmath$ of a language $\mathscr{L}$ as in §*4.3.4*, we want to define an alternative meta-interpretation of $\mathscr{L}$. We recall that $\jmath$ consists of the following data: an $\mathscr{L}$-structure $\imath$, the set $\mathcal{Q}$, the maps $!\colon \mathcal{O} \to \mathcal{Q}$ and $?\colon \mathcal{Q} \to \{\top, \bot\}$, elements $|\&|$, $|\Rightarrow|$ of $\mathrm{Hom}(\mathcal{Q} \times \mathcal{Q}, \mathcal{Q})$, an element $|\mathfrak{q}|$ of $\mathrm{Hom}(\mathrm{Hom}(\mathscr{D}, \mathcal{Q}), \mathcal{Q})$ and an element $|\mathfrak{q}^n|$ of $\mathrm{Hom}(\mathrm{Hom}(\mathrm{Hom}(\mathscr{D}^n, \mathcal{O}), \mathcal{Q}), \mathcal{Q})$ for each $n \in \mathbb{N}$.

The first step, which is also the crucial one, is to interpret $\lozenge\colon \mathbb{1} \to \mu$. Its Tarski-style interpretation was simply by $!\colon \mathcal{O} \to \mathcal{Q}$, which made it possible to routinely omit $\lozenge$ altogether. The alternative interpretation of $\lozenge$ will be more sophisticated.

Given a λ-closed $(n, \vec{r})$-meta-formula of the "atomic" form $\mathcal{G} = \vec{\gamma} \mapsto (\vec{x} \mapsto \lozenge F)$, where $\vec{r} = (r_1, \ldots, r_m)$, it can be interpreted by a function $\|\mathcal{G}\|\colon \mathscr{F}_{r_1} \times \ldots \times \mathscr{F}_{r_m} \to \mathrm{Hom}(\mathscr{T}^n, \mathcal{Q})$, where $\mathscr{T} = \mathscr{T}(\mathscr{L})$ is the set of terms and $\mathscr{F}_i = \mathscr{F}_i(\mathscr{L})$ is the set of i-formulas of the language $\mathscr{L}$. Namely, given an n-tuple $\vec{T} \in \mathscr{T}^n$ and an m-tuple $\vec{\Phi} \in \mathscr{F}_{r_1} \times \ldots \times \mathscr{F}_{r_m}$, we can find a formula F', α-equivalent to F and such that $\vec{T}$ is free for $\vec{x}$ in F and $\vec{\Phi}$ is free for $\vec{\gamma}$ in F. Let $F_{\vec{\Phi}, \vec{T}} = F'[\vec{\gamma}/\vec{\Phi}, \vec{x}/\vec{T}]$. The principle $\cdot F_{\vec{\Phi}, \vec{T}}$ is Tarski-interpreted by a specific element $|\cdot F_{\vec{\Phi}, \vec{T}}| \in \mathcal{Q}$, and we set $\|\mathcal{G}\|(\vec{F})(\vec{T})$ to be this element. Clearly, the α-equivalence class of $\cdot F_{\vec{\Phi}, \vec{T}}$ does not depend on the choice of F', and hence $\|\mathcal{G}\|$ is well-defined.

However, the set of terms $\mathscr{T}$ and the sets of i-formulas $\mathscr{F}_i$ are not really semantic notions, so it is not desirable to use them in our "alternative semantics". To this end we will replace them with their semantic analogues.

(Incidentally, let us also note that the above definition of $\|\mathcal{G}\|$ does not seem to generalize in a reasonable way to the case where $\mathcal{G}$ is not λ-closed, because same variables could have free occurrences in the original formula and also new occurrences introduced upon the substitution.)

Let ω denote, as usual, the totally ordered set $\{0, 1, \ldots\}$. A *semantic term of degree* $k \in \omega$ is a function $T\colon \mathscr{D}^k \to \mathscr{D}$, and a *semantic n-formula of degree* $(k, l) \in \omega \times \omega$ is a function $\Phi\colon \mathscr{D}^k \times \mathcal{O}^{[l]} \to \mathrm{Hom}(\mathscr{D}^n, \mathcal{O})$, where $\mathcal{O}^{[l]}$ denotes the product $\prod_{i=0}^{l-1} \mathrm{Hom}(\mathscr{D}^i, \mathcal{O})^l$ of $l \times l$ factors. (We allow for the possibility that $k \neq l$ only as a matter of convenience; it is not strictly necessary.) A *semantic term* is an element $[k, T]$ of the disjoint union $|\mathscr{T}| := \bigsqcup_{k \in \omega} \mathrm{Hom}(\mathscr{D}^k, \mathscr{D})$, and a *semantic n-formula* is an element $[k, l, \Phi]$ of the disjoint union $|\mathscr{F}_n| := \bigsqcup_{k, l \in \omega} \mathrm{Hom}(\mathscr{D}^k \times \mathcal{O}^{[l]}, \mathrm{Hom}(\mathscr{D}^n, \mathcal{O}))$. A semantic 0-formula is also called a *semantic formula*.[90]

Given a λ-closed $(n, \vec{r})$-meta-formula of the "atomic" form $\mathcal{G} = \vec{\gamma} \mapsto (\vec{x} \mapsto \lozenge F)$, where $\vec{r} = (r_1, \ldots, r_m)$, its alternative interpretation will be a map $\{\mathcal{G}\}\colon |\mathscr{F}_{\vec{r}}| \to \mathrm{Hom}(|\mathscr{T}|^n, \mathcal{Q})$, defined as follows. Let us start from the usual Tarski-style interpretation $|\mathcal{G}|\colon \prod_{i=1}^{m} \mathrm{Hom}(\mathscr{D}^{r_i}, \mathcal{O}) \to \mathrm{Hom}(\mathscr{D}^n, \mathcal{Q})$. Given an n-tuple $\vec{T}$ of semantic terms $T_i\colon \mathscr{D}^{k_i} \to \mathscr{D}$, we can combine them into a map $T\colon \mathscr{D}^k \to \mathscr{D}^n$, where $k = \max(k_1, \ldots, k_n)$. Given an m-tuple $\vec{\Phi}$ of semantic r_i-formulas $\Phi_k\colon \mathscr{D}^{k_i'} \times \mathcal{O}^{[l_i]} \to \mathrm{Hom}(\mathscr{D}^{r_i}, \mathcal{O})$, we can combine them into a map $\Phi\colon \mathscr{D}^{k'} \times \mathcal{O}^{[l]} \to \prod_{i=1}^{m} \mathrm{Hom}(\mathscr{D}^{r_i}, \mathcal{O})$, where $k' = \max(k_1', \ldots, k_m')$ and $l = \max(l_1, \ldots, l_m)$. Finally, $|\mathcal{G}|$, T and Φ combine into a map $\mathscr{D}^{k''} \times \mathcal{O}^{[l]} \to \mathcal{Q}$, where $k'' = \max(k, k')$, which we will denote, somewhat loosely, by $|\mathcal{G}| \circ (\vec{T} \times \vec{\Phi})$. We can apply to this map the Tarski-interpreted k'-fold first-order and $l \times l$-fold second-order meta-quantifier $|\mathfrak{q}_{k''}^{[l]}|\colon \mathrm{Hom}(\mathscr{D}^{k''} \times \mathcal{O}^{[l]}, \mathcal{Q}) \to \mathcal{Q}$ and obtain an element $|\mathfrak{q}_{k''}^{[l]}|\big(|\mathcal{G}| \circ (\vec{T} \times \vec{\Phi})\big) \in \mathcal{Q}$.

---

[90] Although "semantic terms" and "semantic formulas" are technically oxymorons, this nevertheless seems to be a reasonable terminology.

Thus we have described a map $\{G\} : |\mathscr{F}_{\vec{r}}| \to \mathrm{Hom}(|\mathscr{T}|^n, \mathcal{Q})$. Let us note that the Tarski-interpreted meta-quantifiers $|\mathfrak{q}|$ and $|\mathfrak{q}^n|$ are used here to interpret $\mathfrak{l}$ rather than any meta-quantifiers.

Next, we keep the Tarski-style interpretations $\{\&\} := |\&|$ and $\{\Rightarrow\} := |\Rightarrow|$ for the meta-connectives, whereas the first-order meta-quantifier $\mathfrak{q}$ is re-interpreted by a function $\{\mathfrak{q}\} : \mathrm{Hom}(|\mathscr{T}|, \mathcal{Q}) \to \mathcal{Q}$; and the n-ary second-order meta-quantifier $\mathfrak{q}^n$ by a function $\{\mathfrak{q}^n\} : \mathrm{Hom}(|\mathscr{F}_n|, \mathcal{Q}) \to \mathcal{Q}$.

The above will be called the *alternative meta-interpretation* associated to the given Tarski-style meta-interpretation $\jmath$. In the case of a Tarski-style interpretation $\jmath = \imath_+$, we can define an *alternative interpretation* by providing explicit definitions of the functions $\{\mathfrak{q}\}$ and $\{\mathfrak{q}^n\}$. Namely, we set $\{\mathfrak{q}\}(f) = \top$ if and only $f(T) = \top$ for all $T \in |\mathscr{T}|$ and $\{\mathfrak{q}^n\}(f) = \top$ if and only $f(\Phi) = \top$ for all $\Phi \in |\mathscr{F}_n|$.

Given an alternative meta-interpretation, we can interpret any λ-closed $(n, \vec{r})$-meta-formula $\mathcal{G}$ by a function $\{\mathcal{G}\} : |\mathscr{F}_{\vec{r}}| \to \mathrm{Hom}(|\mathscr{T}|^n, \mathcal{Q})$ by a straightforward induction. In particular, any λ-closed meta-formula $\mathcal{F}$ is interpreted by an element $\{\mathcal{F}\} \in \mathcal{Q}$. We may also write $\{\mathcal{F}\}_{\jmath}$ to emphasize the given meta-interpetation $\jmath$, or $\{\mathcal{F}\}_{\imath}$ in the case of an interpretation $\jmath = \imath_+$.

We will see below that the alternative interpretation turns out to be the same as the original Tarski-style interpretation for principles (§*7.2.3*), but generally different for rules (§*7.2.4*). In order to generalize this to many-valued meta-interpretations, we will need to reformulate alternative interpretations in a different language (§**7.3**).

7.2.3. Principles and rules. A *pre-assignment* $\mathscr{U}$ assigns a semantic term $\mathscr{U}(x) \in |\mathscr{T}|$ to each individual variable x. Clearly, a pre-assignment $\mathscr{U}$ along with an interpretation $\imath$ of function symbols yield an interpretation of each term T by a semantic term $\{T\}_{\mathscr{U}}^{\imath}$, also denoted $\{T\}_{\mathscr{U}}$. Let us note that a usual variable assignment u is a special case of a pre-assignment $\mathscr{U}$ that assigns to each individual variable x a semantic term $\mathscr{U}(x)$ of the form $[0, T]$, $T : \mathscr{D}^0 \to \mathscr{D}$, i.e. an element of $\mathscr{D}$.

A *pre-valuation* $\mathscr{V}$ assigns a semantic n-formula $\mathscr{V}(\gamma) \in |\mathscr{F}_n|$ to each n-ary predicate variable γ, for each $n = 0, 1, \ldots$. Let us note that a usual predicate valuation is a special case of a pre-valuation $\mathscr{V}$ that assigns to each n-ary predicate variable γ a semantic n-formula $\mathscr{V}(\gamma)$ of the form $[0, 0, \Phi]$, $\Phi : \mathscr{D}^0 \times \mathcal{O}^{(0)} \to \mathrm{Hom}(\mathscr{D}^n, \mathcal{O})$, i.e. an element of $\mathrm{Hom}(\mathscr{D}^n, \mathcal{O})$.

Clearly, a pre-valuation $\mathscr{V}$ along with an interpretation $\imath$ of function symbols, predicate constants, connectives and quantifiers yield an interpretation of each closed n-formula G by a semantic n-formula $\{G\}_{\imath}^{\mathscr{V}}$, also denoted $\{G\}^{\mathscr{V}}$. Given additionally a pre-assignment $\mathscr{U}$, then also each formula F is interpreted by a semantic formula $\{F\}_{\mathscr{U}}^{\mathscr{V}}$. Namely, if $\vec{x} = (x_1, \ldots, x_n)$ is the tuple of all free individual variables of F, then $\{F\}_{\mathscr{U}}^{\mathscr{V}} = [\max(k, k'), l, \Phi \circ T]$, $\Phi \circ T : \mathscr{D}^{\max(k, k')} \times \mathcal{O}^{[l]} \to \mathcal{O}$, is obtained by pre-composing the semantic n-formula $|\vec{x} \mapsto F|^{\mathscr{V}} = [k', l, \Phi]$, $\Phi : \mathscr{D}^k \times \mathcal{O}^{[l]} \to \mathrm{Hom}(\mathscr{D}^n, \mathcal{O})$ with the function $T : \mathscr{D}^k \to \mathscr{D}^n$, $k = \max(k_1, \ldots, k_n)$, obtained by combining the semantic terms $\mathscr{U}(x_i) = [k_i, T_i]$, $T_i : \mathscr{D}^{k_i} \to \mathscr{D}$.

Given an assignment u and a pre-assignment $\mathscr{U}$, we can define a new assignment $\mathscr{U}[u]$ by $\mathscr{U}[u](x) = f\big(u(\mathbb{x}_1^{\mathbb{0}}), \ldots, u(\mathbb{x}_k^{\mathbb{0}})\big)$, where $\mathscr{U}(x) = [k, f]$, $f \colon \mathscr{D}^k \to \mathscr{D}$. Then for each term T we have $|T|_{\mathscr{U}[u]} = f(|\mathbb{x}_1^{\mathbb{0}}|_u, \ldots, |\mathbb{x}_k^{\mathbb{0}}|_u)$, where $\{\!|T|\!\}_{\mathscr{U}} = [k, f]$, $f \colon \mathscr{D}^k \to \mathscr{D}$.

Given an assignment u, a valuation v and a pre-valuation $\mathscr{V}$, we define a new valuation $\mathscr{V}[u, v]$ by $\mathscr{V}[u, v](\gamma) = f\Big(u(\mathbb{x}_1^{\mathbb{0}}), \ldots, u(\mathbb{x}_k^{\mathbb{0}}); \big(v(\mathbb{x}_1^{\mathbb{0}^i \to \mathbb{1}}), \ldots, v(\mathbb{x}_l^{\mathbb{0}^i \to \mathbb{1}})\big)_{i=0}^{l-1}\Big)$, where $\mathscr{V}(\gamma) = [k, l, f]$, $f \colon \mathscr{D}^k \times \mathcal{O}^{[l]} \to \mathrm{Hom}(\mathscr{D}^n, \mathcal{O})$. Then for each closed n-formula G we have $|G|^{\mathscr{V}[u,v]} = f\Big(u(\mathbb{x}_1^{\mathbb{0}}), \ldots, u(\mathbb{x}_k^{\mathbb{0}}); \big(v(\mathbb{x}_1^{\mathbb{0}^i \to \mathbb{1}}), \ldots, v(\mathbb{x}_l^{\mathbb{0}^i \to \mathbb{1}})\big)_{i=0}^{l-1}\Big)$, where $\{\!|G|\!\}^{\mathscr{V}} = (k, l, f)$, $f \colon \mathscr{D}^k \times \mathcal{O}^{[l]} \to \mathrm{Hom}(\mathscr{D}^n, \mathcal{O})$.

Consequently, given an assignment u, a valuation v, a pre-assignment $\mathscr{U}$ and a pre-valuation $\mathscr{V}$, for each formula F we again have

$$|F|^{\mathscr{V}[u,v]}_{\mathscr{U}[u]} = f\Big(u(\mathbb{x}_1^{\mathbb{0}}), \ldots, u(\mathbb{x}_k^{\mathbb{0}}); \big(v(\mathbb{x}_1^{\mathbb{0}^i \to \mathbb{1}}), \ldots, v(\mathbb{x}_l^{\mathbb{0}^i \to \mathbb{1}})\big)_{i=0}^{l-1}\Big),$$

where $\{\!|F|\!\}^{\mathscr{V}}_{\mathscr{U}} = (k, l, f)$, $f \colon \mathscr{D}^k \times \mathcal{O}^{[l]} \to \mathcal{O}$.

Example 7.12. Let us discuss the alternative interpretation (i.e. the alternative two-valued meta-interpretation) of a principle, $\cdot F$.

Since $\mathcal{Q} = \{\top, \bot\}$ and $?\colon \mathcal{Q} \to \{\top, \bot\}$ is the identity map, we may identify elements of $\mathcal{Q}$ with judgements about pre-valuations and pre-assignments.

From the definition, we have

$$\{\!|\cdot F|\!\} = \forall \mathscr{V}, \mathscr{U} \ \forall v, u \ !|F|^{\mathscr{V}[u,v]}_{\mathscr{U}[u]}.$$

This is clearly equivalent to the Tarski-style interpretation

$$|\cdot F| = \forall v, u \ !|F|^v_u.$$

Example 7.13. Let us now discuss the alternative interpretation (i.e. the alternative two-valued meta-interpretation) of a rule, F/G.

From the definition, we have

$$\{\!|F \mathbin{/} G|\!\} = \forall \mathscr{V} \ (\forall \mathscr{U} \ \forall v, u \ !|F|^{\mathscr{V}[u,v]}_{\mathscr{U}[u]} \ \longrightarrow \ \forall \mathscr{U} \ \forall v, u \ !|G|^{\mathscr{V}[u,v]}_{\mathscr{U}[u]}).$$

This clearly simplifies as

$$\{\!|F \mathbin{/} G|\!\} = \forall \mathscr{V} \ (\forall v, u \ !|F|^{\mathscr{V}[u,v]}_{u} \ \longrightarrow \ \forall v, u \ !|G|^{\mathscr{V}[u,v]}_{u}).$$

The latter clearly follows from

$$|F \mathbin{/} G| = \forall \mathscr{V} \ \forall v \ (\forall u \ !|F|^{\mathscr{V}[u,v]}_{u} \ \longrightarrow \ \forall u \ !|G|^{\mathscr{V}[u,v]}_{u}),$$

which is clearly equivalent to the Tarski-style interpretation

$$|F \mathbin{/} G| = \forall v \ (\forall u \ !|F|^v_u \ \longrightarrow \ \forall u \ !|G|^v_u).$$

However, we will see in Example 7.16 below that there is no reverse implication, at least, in a minor modification of the setup.

7.2.4. Two-valued alternative models. Example 7.13 (see also Example 7.17 below) suggests that at least with two-valued interpretation of the meta-logic, the alternative interpretation of the derivability of a rule is very similar to the interpretation of its stable

admissibility (either Tarski-style or alternative, these are the same by Example 7.12). Let us show that despite this similarity, they do not coincide.

Example 7.14. The rule $\exists x\, p(x) \wedge \exists x\, \neg p(x) \,/\, \bot$ is stably admissible for classical logic (see Example 4.71). However, this rule is not alternatively-valid already in the two-valued model of classical logic (two-valued both on the meta-level and the object-level, i.e. $\mathcal{O} = \mathcal{Q} = \{\top, \bot\}$). This is established in essence similarly to Example 4.72.

In more detail, we need to interpret p by a semantic 1-formula $\mathscr{V}(p) = [l, \Phi] \in |\mathscr{F}_1|$, $\Phi \colon \mathscr{D}^k \times \mathcal{O}^{[l]} \to \mathrm{Hom}(\mathscr{D}, \mathcal{O})$. Let us choose $k = 1$, $l = 0$ and $\Phi \colon \mathscr{D} \to \mathrm{Hom}(\mathscr{D}, \mathcal{O})$ defined by $\Phi(y, \gamma)(x) = \top$ if and only if $x = y$. Then the semantic formula $\{\exists x\, p(x) \wedge \exists x\, \neg p(x)\}^{\mathscr{V}}$ is $[1, 0, \Psi] \in |\mathscr{F}_0|$, where $\Psi \colon \mathscr{D} \to \mathcal{O}$ is the Tarski-style interpretation of the λ-closed 1-formula $y \mapsto \exists x\, x = y \wedge \exists x\, \neg x = y$ of the language extended by the binary predicate constant $=$ with the obvious interpretation.

Therefore the alternative interpretation of the λ-closed $(0, (1))$-meta-formula $p \mapsto \mathbb{!}\big(\exists x\, p(x) \wedge \exists x\, \neg p(x)\big) \Rightarrow \bot$ is by a function $f \colon |\mathscr{F}_1| \to \mathcal{Q}$ whose value on $\mathscr{V}(p)$ is the Tarski-style interpretation of the λ-closed meta-formula $((y)\, \exists x\, x = y \wedge \exists x\, \neg x = y) \Rightarrow \bot$ of the extended language. As long as $\mathscr{D}$ contains at least two elements, this value $f\big(\mathscr{V}(p)\big) \in \mathcal{Q}$ is $\bot$. (see Example 4.72). Hence the alternative interpretation of the rule in question, i.e. the λ-closed meta-formula $(p)\left(\mathbb{!}\big(\exists x\, p(x) \wedge \exists x\, \neg p(x)\big) \Rightarrow \bot\right)$, is also $\bot$.

Let us call a semantic n-formula $(k, l, \Phi) \in |\mathscr{F}_k|$, $\Phi \colon \mathscr{D}^k \times \mathcal{O}^{[l]} \to \mathrm{Hom}(\mathscr{D}^n, \mathcal{O})$, *pure* if it uses the given $n + k$ elements of $\mathscr{D}$ only by feeding them into the given elements of the factors $\mathrm{Hom}(\mathscr{D}^i, \mathcal{O})$ of $\mathcal{O}^{[l]}$; or more formally if there exists a semantic formula $(0, l, \Phi') \in |\mathscr{F}_0|$ such that $\Phi(\vec{x}, \vec{\gamma})(\vec{y}) = \Phi'(\vec{\gamma}')$, where $\vec{\gamma} = (\gamma_{01}, \ldots, \gamma_{0l}; \ldots; \gamma_{l-1,1}, \ldots, \gamma_{l-1,l})$, $\vec{\gamma}' = (\gamma_{01} \circ \varphi_{01}, \ldots, \gamma_{0l} \circ \varphi_{0l}; \ldots; \gamma_{l-1,1} \circ \varphi_{l-1,1}, \ldots, \gamma_{l-1,l} \circ \varphi_{l-1,l})$ and each $\varphi_{ij} \colon \mathscr{D}^i \to \mathscr{D}^i$ replaces some of the given elements of $\mathscr{D}$ with some entries of the tuples $\vec{x}$ and $\vec{y}$.

Let $|\mathscr{F}_k|^p$ denote the subset of $|\mathscr{F}_k|$ consisting of pure semantic k-formulas. If the language contains no function symbols and predicate constants, then we define a *pure alternative interpretation* and a *pure pre-valuation* by repeating the definitions of an alternative interpretation and a pre-valuation with $|\mathscr{F}_k|^p$ in place of $|\mathscr{F}_k|$.

Remark 7.15. One can define "pre-valuation fields" similarly to valuation fields (either of the two definitions in §4.7.4), and check that the family $|\mathscr{F}_k|^p$ forms a pre-valuation field.

Example 7.16. The rule $\exists x\, p(x) \wedge \exists x\, \neg p(x) \,/\, \bot$ is not derivable in classical logic, and is not valid in the Tarski-style two-valued model of classical logic with any non-singleton domain. However, this rule is purely alternatively valid in the two-valued model of classical logic with every domain. This is established in essence similarly to Example 4.71.

In more detail, suppose that p is interpreted by a pure semantic 1-formula $\mathscr{V}(p) = (k, l, \Phi) \in |\mathscr{F}_1|^p$, $\Phi \colon \mathscr{D}^k \times \mathcal{O}^{[l]} \to \mathrm{Hom}(\mathscr{D}, \mathcal{O})$. Let $f_i \colon \mathscr{D}^i \to \mathcal{O}$ send every input to $\bot$, and let $\vec{f} = (f_0, \ldots, f_0; \ldots; f_{l-1}, \ldots, f_{l-1}) \in \mathcal{O}^{[l]}$. Since $\mathscr{V}(p)$ is pure, $\Phi(\vec{y}, \vec{f})(x)$ does not depend on $\vec{y}$ and x. Then the semantic formula $\{\exists x\, p(x) \wedge \exists x\, \neg p(x)\}^{\mathscr{V}}$ is the same

as $\{p(x) \wedge \neg p(x)\}_u^{\mathscr{V}}$ for any individual assignment u. But the latter is $\perp\!\!\!\perp$. It follows that the rule in question is purely alternatively valid.

7.2.5. Platonist interpretation II (alternative). We now present a new contentual interpretation of the meta-logic of classical logic based on the alternative semantics of §*7.2.2* (two-valued, both on the meta-level and the object-level).

Like in the Platonist interpretation of §*4.8.3*, a domain $\mathscr{D}$ is fixed, and $\mathscr{O}$ is taken to be a class of propositions, consisting of a prescribed class of contentful (e.g. mathematical) primitive propositions and of composite propositions, which are obtained inductively from the primitive ones by using contentual classical connectives and quantifiers over arbitrary $\mathscr{D}$-indexed families of propositions. Formal classical connectives and quantifiers are interpreted straightforwardly by the contentual ones.

$\mathcal{Q}$ is taken to be a larger class of propositions, consisting of those in $\mathscr{O}$, called "meta-primitive" (thus the function $!\colon \mathscr{O} \to \mathcal{Q}$ is the inclusion) and of "meta-composite" propositions, obtained inductively from the meta-primitive ones by using contentual classical connectives and quantifiers over arbitrary $\mathscr{D}$-indexed, $\mathrm{Hom}(\mathscr{D}^i, \mathscr{O})$-indexed ($i = 0, 1, \dots$), $|\mathscr{T}|$-indexed and $|\mathscr{F}_i|$-indexed ($i = 0, 1, 2, \dots$) families of propositions. (We recall that $|\mathscr{T}|$ and $|\mathscr{F}_i|$ are defined in terms of $\mathscr{D}$ and $\mathscr{O}$, see §*7.2.2*.) Meta-connectives and meta-quantifiers are Tarski-interpreted by contentual classical connectives and $\mathscr{D}$-indexed and $\mathrm{Hom}(\mathscr{D}^i, \mathscr{O})$-indexed quantifiers in the straightforward way; and meta-quantifiers are also alternatively-interpreted by contentual classical $|\mathscr{T}|$-indexed and $|\mathscr{F}_i|$-indexed quantifiers.

The function $?\colon \mathcal{Q} \to \{\top, \perp\!\!\!\perp\}$ sends a proposition P to $\top$ if and only if P is true, where "true" is defined explicitly for primitive $P \in \mathscr{O}$ and is extended inductively to composite $P \in \mathscr{O}$ and then further to all meta-composite $P \in \mathcal{Q}$ by the usual truth tables. This determines a contentual interpretation of classical logic and its meta-logic, which we will call the *alternative Platonist interpretation.*

Example 7.17. Let us record the alternative Platonist interpretations of the judgements of admissibility, stable admissibility and derivability of a purely logical rule $F\,/\,G$. Let $\vec{\varphi}$ be the tuple of all predicate variables that occur freely in F or in G, and let $\vec{r}$ be the tuple of their arities. Below $\vec{\Phi}$ is understood to run over $\mathscr{F}_{\vec{r}}$.

- Derivability: $\forall \mathscr{D}, \mathscr{O}\ \forall \mathscr{V}\ (\forall v, u\ |F|_u^{\mathscr{V}[u,v]} \ \longrightarrow\ \forall v, u\ |G|_u^{\mathscr{V}[u,v]})$;
- Stable admissibility: $\forall \vec{\Phi}\, \forall \mathscr{D}, \mathscr{O}\ (\forall v, u\ |F[\vec{\varphi}/\vec{\Phi}]|_u^v \ \longrightarrow\ \forall v, u\ |F[\vec{\varphi}/\vec{\Phi}]|_u^v)$;
- Admissibility: $\forall \vec{\Phi}\ (\forall \mathscr{D}, \mathscr{O}\ \forall v, u\ |F[\vec{\varphi}/\vec{\Phi}]|_u^v \ \longrightarrow\ \forall \mathscr{D}, \mathscr{O}\ \forall v, u\ |F[\vec{\varphi}/\vec{\Phi}]|_u^v)$.

7.3. Generalized alternative semantics

Our next goal is to better understand alternative meta-interpretations, so as to be able to define alternative meta-models. In particular, we have seen above that principles are interpreted in the same way in Tarski-style and in alternative two-valued semantics. But so far we have no way to see that this is also so in the many-valued case.

To this end we will first look at an extension of the meta-logic, which besides the original copy of the meta-logic will contain another, "unintended" embedding. The point is that the straightforward, generalized-Tarski-style interpretation of the extended meta-logic will restrict to the Kolmogorov-style interpretation on the re-embedded meta-logic (and, unsurprisingly, to the Tarski-style interpretation on the original copy).

This extension of the meta-logic goes somewhat beyond being another special case of Paulson's formalism in `Isabelle`'s metalogic: it employs λ-calculus with $\mathbb{N}$-indexed sums (i.e., countably infinite disjoint unions), and also assumes a generalized form of $\Rightarrow$-elimination.

7.3.1. Extended λ-calculus. We define an extension of the λ-calculus described in §**4.2** which involves infinite sums and certain variable lists.

Let $\omega = \{0, 1, \dots\}$ and for any $n \in \omega$ let $[n] = \{1, \dots, n\}$; thus $[0] = \varnothing$. The extended λ-calculus has the following type forming rules:

(1) base types are types;
(2) if Γ and Δ are types, then so is $\Gamma \to \Delta$;
(3) if $n \in \omega$ and Γ_i is a type for each $i \in [n]$, then so is $\prod_{i \in [n]} \Gamma_i$;
(4) if Γ_i is a type for each $i \in \omega$, then so is $\sum_{i \in \omega} \Gamma_i$.

We write $\Gamma^n = \prod_{i \in [n]} \Gamma$. We often shorten $\prod_{i \in [n]} \Gamma_i$ to $\prod_{i=1}^n \Gamma_i$ and $\sum_{i \in \mathbb{N}} \Gamma_i$ to $\sum_{i=1}^\infty \Gamma_i$.

While n-ary products as such could be defined in terms of binary products (see §*4.2.4*), we will need them already to properly describe the expression forming rules (cf. Remark 7.18 below.)

Constants and variables are denoted as before (see §*4.2.1*). Expression forming rules are as follows, where $n \in \omega$:

(1) a constant of type Γ is an expression of type Γ;
(2) a variable of type Γ is an expression of type Γ;
(3) (abstraction) if $T : \Delta$ on the assumption that x is a variable of type Γ, then $x \mapsto T : \Gamma \to \Delta$;
(4) (application) if $F : \Delta \to \Gamma$, and $T : \Delta$, then $F(T) : \Gamma$;
(5) (projections) if $k \in [n]$, then $\mathbb{p}_k : \prod_{i \in [n]} \Gamma_i \to \Gamma_k$;
(6) (function tuples) if $F_i : \Delta \to \Gamma_i$ on the assumption that $i \in [n]$, then $(F_i)_{i \in [n]} : \Delta \to \prod_{i \in [n]} \Gamma_i$;[91]
(7) (inclusions) if $k \in \omega$, then $\mathbb{i}_k : \Gamma_k \to \sum_{i \in \omega} \Gamma_i$;
(8) (definition by cases) if $F_i : \Gamma_i \to \Delta$ on the assumption that $i \in \omega$, then $(F_i)^{i \in \omega} : \sum_{i \in \omega} \Gamma_i \to \Delta$;[92]
(9) (pre-tuple) there is a constant $\vec{\varnothing} : \prod_\varnothing$.

Using the pre-tuple and function tuples, we can define usual tuples (at least, up to α-equivalence): given a $T_i : \Gamma_i$ for each $i \in [n]$, let $(T_i)_{i \in [n]} : \prod_{i \in [n]} \Gamma_i$ denote $(x \mapsto T_i)_{i \in [n]}(\vec{\varnothing})$, where x is a variable of type $\prod_\varnothing$ that does not occur freely in any of the T_i.

---

[91]We use these "function tuples" instead of the traditional n-tuples in order to emphasize symmetry between product and sum.

[92]This is yet another divergence from traditional notation of λ-calculus.

Thus the empty tuple (): $\prod_\varnothing$ is the image of the pre-tuple $\vec{\varnothing} : \prod_\varnothing$ under the empty function tuple (): $\prod_\varnothing \to \prod_\varnothing$. (The latter has the meaning of a constant map, but secretly is also the identity map.)

Remark 7.18. We treat the index i of the variable $\varkappa_i^\Gamma$ as a variable, rather than a meta-variable. This means that we may say, for example, that "$\varkappa_i^\Gamma : \Gamma$ on the assumption that $i \in \omega$". It follows that $(\varkappa_i^\Gamma)_{i \in [n]}$, which can also be written as $(\varkappa_1^\Gamma, \ldots, \varkappa_n^\Gamma)$, is a valid expression of type Γ^n on the assumption that $n \in \omega$. Consequently, $\big(f_i \mapsto f_i(\varkappa_1^\Gamma, \ldots, \varkappa_n^\Gamma)\big)^{i \in \omega}$ is a valid expression of type $\sum_{i \in \omega}(\Gamma^i \to \Gamma) \to \Gamma$. Also, $\big((\varkappa_1^{\Gamma_1}, \ldots, \varkappa_n^{\Gamma_1}), \ldots, (\varkappa_1^{\Gamma_n}, \ldots, \varkappa_n^{\Gamma_n})\big)$ is a valid expression of type $\prod_{i=1}^n \Gamma_i^n$ on the assumption that $n \in \omega$.

Example 7.19. $\big(f \mapsto (\mathbb{p}_i \circ f)\big)_{i \in [n]} : \Big(\Delta \to \prod_{i \in [n]} \Gamma_i\Big) \to \prod_{i \in [n]}(\Delta \to \Gamma_i)$ is a λ-closed expression.

Example 7.20. $t \mapsto (\mathbb{p}_i(t))_{i \in [n]} : \prod_{i \in [n]}(\Delta \to \Gamma_i) \to \Big(\Delta \to \prod_{i \in [n]} \Gamma_i\Big)$ is a λ-closed expression.

The previous two examples show that function tuples amount to a special case of usual tuples (even though usual tuples were defined via a special case of function tuples).

Although we do not really need n-ary sums, they can be defined just like the infinite sums above, and have the following relations with n-ary products:

Example 7.21. $t \mapsto \big(\mathbb{p}_i(t)\big)^{i \in [n]} : \prod_{i \in [n]}(\Gamma_i \to \Delta) \to \Big(\sum_{i \in [n]} \Gamma_i \to \Delta\Big)$ is a λ-closed expression.

Example 7.22. $f \mapsto (f \circ \mathbb{i}_i)_{i \in [n]} : \Big(\sum_{i \in [n]} \Gamma_i \to \Delta\Big) \to \prod_{i \in [n]}(\Gamma_i \to \Delta)$ is a λ-closed expression.

The previous two examples show that $(\cdot)^{i \in [n]}$ is in essence a special case of $(\cdot)_{i \in [n]}$. (But technically they live in different types. Also, their comparison uses both rules, so none of the two rules has been deduced from the other one.)

Example 7.23. $\Big(f_i \mapsto \big(t \mapsto f_i \circ \mathbb{p}_i(t)\big)\Big)^{i \in [n]} : \sum_{i \in [n]}(\Gamma_i \to \Delta) \to \Big(\prod_{i \in [n]} \Gamma_i \to \Delta\Big)$ is a λ-closed expression.

Example 7.24. $f_i \mapsto (\mathbb{i}_i \circ f_i)^{i \in [n]} : \sum_{i \in [n]}(\Delta \to \Gamma_i) \to \Big(\Delta \to \sum_{i \in [n]} \Gamma_i\Big)$ is a λ-closed expression. The same goes for the infinite sums.

The definition of *free occurrence* is extended as follows:
   (1) x occurs freely in $(F_i)_{i \in [n]}$ if it occurs freely in F_i for some $i \in [n]$;
   (2) x occurs freely in $(F_i)^{i \in \omega}$ if it occurs freely in F_i for some $i \in \omega$.

The definition of *substitution* is extended as follows:
   (1) $(F_i)_{i \in [n]}|_{x := T} = (F_i|_{x := T})_{i \in [n]}$;
   (2) $(F_i)^{i \in \omega}|_{x := T} = (F_i|_{x := T})^{i \in \omega}$.

We also define terminal types and their units following [36]:

(1) If each Θ_i is a terminal type with unit U_i, then $\prod_{i \in [n]} \Theta_i$ is a terminal type with unit $(U_i)_{i \in [n]}$. In particular, $\prod_\varnothing$ is a terminal type with unit $()$.
(2) If Γ is a type and Θ is a terminal type with unit U, then $\Gamma \to \Theta$ is a terminal type with unit $x \mapsto U$.

The definition of β-*reduction* is extended as follows:

(1) if $n \in [m]$, every expression of the form $\mathbb{p}_n\big((F_i)_{i \in [m]}(T)\big)$ β-reduces to $F_n(T)$;
(2) if $n \in \omega$, every expression of the form $(F_i)^{i \in \omega}\big(\mathbb{i}_n(T)\big)$ β-reduces to $F_n(T)$.
(3) if Θ is a terminal type with unit U, every $T : \Theta$ β-reduces to U.

Example 7.25. Let us see how to recover the β-reduction for usual tuples: $\mathbb{p}_n\big((T_i)_{i \in [m]}\big)$ β-reduces to T_n. Indeed, $\mathbb{p}_n\big((x \mapsto T_i)_{i \in [m]}(\vec\varnothing)\big)$, where x does not occur freely in any of the T_i, β-reduces to $(x \mapsto T_n)(\vec\varnothing)$, which in turn β-reduces to T_n.

Example 7.26. If Θ is a terminal type with unit U, we have $\varphi : \Gamma \to (\Theta \to \Gamma)$, $\varphi = t \mapsto (x \mapsto t)$, and $\psi : (\Theta \to \Gamma) \to \Gamma$, $\psi = f \mapsto f(U)$. Now $\psi\big(\varphi(t)\big)$ β-reduces to $(x \mapsto t)(U)$, which in turn β-reduces to t. Conversely, $\varphi\big(\psi(f)\big)$ β-reduces to $x \mapsto f(U)$. This does not β-reduce to f, but we will now see that it $\beta\eta$-reduces to f.

The definition of η-*reduction* is extended as follows for $n > 0$:

(1) every expression of the form $\Big(x \mapsto \mathbb{p}_i\big(F(x)\big)\Big)_{i \in [n]}$ η-reduces to F;
(2) every expression of the form $\Big(x_i \mapsto F\big(\mathbb{i}_i(x_i)\big)\Big)^{i \in \omega}$ η-reduces to F;
(3) if Θ is a terminal type with unit U, for every $F : \Theta \to \Gamma$, the expression $x \mapsto F(U)$ η-reduces to F;
(4) if Θ is a terminal type with unit U, for every $F : \Gamma \to \Theta \times \Delta$, the expression $\Big(x \mapsto U, \, x \mapsto \mathbb{p}_1\big(F(x)\big)\Big)$ η-reduces to F;
(5) if Θ is a terminal type with unit U, for every $F : \Theta \sqcup \Gamma \to \Delta$, the expression $\Big(x_1 \mapsto F\big(\mathbb{i}_1(U)\big), \, x_2 \mapsto F\big(\mathbb{i}_1(x_2)\big)\Big)$ η-reduces to F.

Example 7.27. Let us see how to recover the η-reduction for usual tuples: $\big(\mathbb{p}_i(T)\big)_{i \in [n]}$ is $\beta\eta$-equivalent to T. Indeed, $\big(x \mapsto \mathbb{p}_i(T)\big)_{i \in [n]}(\vec\varnothing)$, where x does not occur freely in T, β-expands to $\Big(x \mapsto \mathbb{p}_i\big((x \mapsto T)(x)\big)\Big)_{i \in [n]}(\vec\varnothing)$ which η-reduces to $(x \mapsto T)(\vec\varnothing)$, which β-reduces to T.

Remark 7.28. Let us note some immediate consequences of the β- and η-rules:

- $\mathbb{p}_n \circ (F_i)_{i \in [m]}$ $\beta\eta$-reduces to F_n, as long as $n \in [m]$;
- $(\mathbb{p}_i \circ F)_{i \in [m]}$ η-reduces to F;
- $(F_i)^{i \in \omega} \circ \mathbb{i}_n$ $\beta\eta$-reduces to F_n, as long as $n \in \omega$;
- $(F \circ \mathbb{i}_i)^{i \in \omega}$ η-reduces to F.

Here are some further consequences:

Lemma 7.29. *(a)* $(F_i)_{i \in [n]} \circ G$ *is $\beta\eta$-equivalent to* $(F_i \circ G)_{i \in [n]}$*;*
(b) $G \circ (F_i)^{i \in \omega}$ *is $\beta\eta$-equivalent to* $(G \circ F_i)^{i \in \omega}$*.*

Proof. By the previous remark, $(F_i \circ G)_{i \in [n]}$ $\beta\eta$-expands to $\big(\mathbb{p}_i \circ (F_j)_{j \in [n]} \circ G\big)_{i \in [n]}$, which in turn η-reduces to $(F_j)_{j \in [n]} \circ G$.

Similarly, $(G \circ F_i)^{i \in \omega}$ $\beta\eta$-expands to $\big(G \circ (F_j)^{j \in \omega} \circ \mathring{\mathbb{i}}_i\big)^{i \in \omega}$, which in turn η-reduces to $G \circ (F_j)^{j \in \omega}$. $\qquad\square$

Remark 7.30. Let $\Gamma = \sum_{i \in \omega} \Gamma_i$ and $\Delta = \sum_{i \in \omega} \Delta_i$. Suppose that $F_{ij} : \Gamma_i \times \Delta_j \to \Theta$ for each $i, j \in I$. We will define $(F_{ij})^{(i,j) \in \omega \times \omega} : \Gamma \times \Delta \to \Theta$.

We will use $\mathfrak{F}$ and $\mathfrak{G}$ defined in Example 4.11. We have $\mathfrak{F}(F_{ij}) : \Gamma_i \to (\Delta_j \to \Theta)$. Then $\big(\mathfrak{F}(F_{ij})\big)^{i \in \omega} : \Gamma \to (\Delta_j \to \Theta)$. Let $\mathfrak{H} = \mathfrak{F} \circ \mathfrak{T} \circ \mathfrak{G}$, where $\mathfrak{T} : (\Gamma \times \Delta \to \Theta) \to (\Delta \times \Gamma \to \Theta)$ is defined by $f \mapsto \big(q \mapsto f(\mathbb{p}_2 q, \mathbb{p}_1 q)\big)$. Then $\mathfrak{H}\big(\mathfrak{F}(F_{ij})\big)^{i \in \omega} : \Delta_j \to (\Gamma \to \Theta)$. Hence $\Big(\mathfrak{H}\big(\mathfrak{F}(F_{ij})\big)^{i \in \omega}\Big)^{j \in \omega} : \Delta \to (\Gamma \to \Theta)$. Finally, let

$$(F_{ij})^{(i,j) \in \omega \times \omega} = \mathfrak{G}\Big(\mathfrak{H}\big(\mathfrak{F}(F_{ij})\big)^{i \in \omega}\Big)^{j \in \omega} : \Gamma \times \Delta \to \Theta.$$

More generally, given $F_{k_1, \dots, k_n} : \Gamma_{1, k_1} \times \dots \times \Gamma_{n, k_n} \to \Theta$ for each $k_1, \dots, k_n \in I$, one can similarly define

$$(F_{k_1, \dots, k_n})^{(k_1, \dots, k_n) \in \omega^n} : \sum_{i \in \omega} \Gamma_{1,i} \times \dots \times \sum_{i \in \omega} \Gamma_{n,i} \to \Theta.$$

7.3.2. Extended meta-logic. We use the notation introduced in §7.2.2. Given a $k \in \omega$ (i.e., a nonnegative integer), we denote by $\mathbb{1}^{[k]}$ the product $\prod_{i=0}^{k-1}(\mathbb{0}^i \to \mathbb{1})^k$ of $k \times k$ factors. We will also write $\varkappa_{[k]}^\tau = (\varkappa_1^\tau, \dots, \varkappa_k^\tau) : \tau^k$ and $\varkappa_{[[k]]}^{\mathbb{1}} = (\varkappa_{[k]}^{\mathbb{0}^0 \to \mathbb{1}}, \dots, \varkappa_{[k]}^{\mathbb{0}^{k-1} \to \mathbb{1}})) : \mathbb{1}^{[k]}$ (see Remark 7.18).

It is easy to see that every term T of the language can be obtained as a β-reduction of $T_k(\varkappa_{[k]}^{\mathbb{0}})$ for some λ-closed expression $T_k : \mathbb{0}^k \to \mathbb{0}$, where k is the maximal number such that $\varkappa_k^{\mathbb{0}}$ occurs in T. Similarly, every n-formula F of the language can be obtained as a β-reduction of $F_k(\varkappa_{[k]}^{\mathbb{0}}, \varkappa_{[[k]]}^{\mathbb{1}})$ for some λ-closed $F_k : \mathbb{0}^k \times \mathbb{1}^{[k]} \to (\mathbb{0}^n \to \mathbb{1})$, where k is the maximal number such that either $\varkappa_k^{\mathbb{0}}$ or $\varkappa_k^{\mathbb{0}^{l-1} \to \mathbb{1}}$ or $\varkappa_l^{\mathbb{0}^{k-1} \to \mathbb{1}}$ occurs in F for some l.

Let $\mathbb{0}^* = \sum_{k \in \omega}(\mathbb{0}^k \to \mathbb{0})$ and $\mathbb{1}_n^* = \sum_{k \in \omega}\big(\mathbb{0}^k \times \mathbb{1}^{[k]} \to (\mathbb{0}^n \to \mathbb{1})\big)$. Let us define *term evaluation* $\mathfrak{E} : \mathbb{0}^* \to \mathbb{0}$ by $\mathfrak{E} = (\mathfrak{E}_k)^{k \in \omega}$, where $\mathfrak{E}_k : (\mathbb{0}^k \to \mathbb{0}) \to \mathbb{0}$ is defined by $t_k \mapsto t_k(\varkappa_{[k]}^{\mathbb{0}})$. (See Remark 7.18.) Then every term T of the language can also be obtained as a β-reduction of $\mathfrak{E}(T^*)$ for some λ-closed $T^* : \mathbb{0}^*$. Indeed, let $T^* = \mathring{\mathbb{i}}_k(T_k)$, where T_k is as above; then $\mathfrak{E}\big(\mathring{\mathbb{i}}_k(T_k)\big)$ β-reduces to $\mathfrak{E}_k(T_k)$, which in turn β-reduces to $T_k(\varkappa_{[k]}^{\mathbb{0}})$.

Similarly, we have *n-formula evaluation* $\mathfrak{E}^n : \mathbb{1}_n^* \to (\mathbb{0}^n \to \mathbb{1})$, $\mathfrak{E}^n = (\mathfrak{E}_k^n)^{k \in \omega}$, where $\mathfrak{E}_k^n : \big(\mathbb{0}^k \times \mathbb{1}^{[k]} \to (\mathbb{0}^n \to \mathbb{1})\big) \to (\mathbb{0}^n \to \mathbb{1})$ is defined by $\varphi_k \mapsto \varphi_k(\varkappa_{[k]}^{\mathbb{0}}, \varkappa_{[[k]]}^{\mathbb{1}})$. Every n-formula F of the language can also be obtained as a β-reduction of $\mathfrak{E}^n(F^*)$ for some λ-closed $F^* : \mathbb{1}_n^*$. Namely, $F^* = \mathring{\mathbb{i}}_k(F_k)$, where F_k is as above. Then $\mathfrak{E}^n\big(\mathring{\mathbb{i}}_k(F_k)\big)$ β-reduces to $\mathfrak{E}_k^n(F_k)$, which in turn β-reduces to $F_k(\varkappa_{[k]}^{\mathbb{0}}, \varkappa_{[[k]]}^{\mathbb{1}})$.

The variables of type $\mathbb{0}^*$ will be called *term variables*, and the variables of type $\mathbb{1}_n^*$ will be called *n-formula variables* (or *formula variables* when $n = 0$). Corresponding to these two sorts of variables, there are two new types of meta-quantifiers:

- *the term meta-quantifier* $\mathfrak{q}^* : (\mathbb{0}^* \to \mu) \to \mu$;
- *the k-formula meta-quantifier* $\mathfrak{q}_n^* : (\mathbb{1}_n^* \to \mu) \to \mu$.

In line with previous notation, we write $(t)\,\mathcal{F}$ to mean $\mathfrak{q}^*(t \mapsto \mathcal{F})$ where t is a term variable of an n-formula variable and $\mathcal{F}$ is a λ-expression of type μ. We also write $(\vec{t})\,\mathcal{F}$ to mean $(t_1) \dots (t_n)\,\mathcal{F}$, where $\vec{t} = (t_1, \dots, t_n)$ is a tuple of term variables and $(\vec{\varphi})\,\mathcal{F}$ to mean $(\varphi_1) \dots (\varphi_m)\,\mathcal{F}$, where $\vec{\varphi} = (\varphi_1, \dots, \varphi_m)$ is a tuple of r_i-formula variables.

We define a *term scheme* (resp. a *term pre-scheme*) as an expression of type $\mathbb{0}$ (resp. $\mathbb{0}^*$) defined using function application, projections, inclusions, tuples and definition by cases; function symbols of the language; individual and term variables and abstraction over them. We define a *k-formula scheme* (resp. a *k-formula pre-scheme*) as an expression of type $\mathbb{0}^k \to \mathbb{1}$ (resp. $\mathbb{1}_k^*$) defined using all of the above, plus: predicate constants, connectives and quantifiers of the language; predicate and n-formula variables and abstraction over them. Finally, we define a *meta-formula scheme* as an expression of type μ defined using all of the above, plus $\mathbb{0}$, the two meta-connectives and the four sorts of meta-quantifiers. In particular, terms schemes may involve $\mathfrak{E}$, and k-formula schemes and meta-formula schemes may involve $\mathfrak{E}$ and each $\mathfrak{E}^n$.

The inference meta-rules of §*4.4.1* are now applied to term schemes, k-formula schemes and meta-formula schemes in place of terms, k-formulas and meta-formulas, and are augmented by the following new meta-rules, where $\mathcal{F}$, $\mathcal{F}_k$, $\mathcal{G}_k$ are meta-formula schemes, t is a term variable, T is a term pre-scheme, φ is an n-formula variable and Φ is an n-formula pre-scheme.

(1) (meta-generalization)

$$\frac{\begin{array}{c} \vdots \\ \mathcal{F} \end{array}}{(t)\,\mathcal{F}}, \text{ provided that } t \text{ does not occur freely in any of the assumptions;}$$

$$\frac{\begin{array}{c} \vdots \\ \mathcal{F} \end{array}}{(\varphi)\,\mathcal{F}}, \text{ provided that } \varphi \text{ does not occur freely in any of the assumptions.}$$

(2) (meta-specialization)

$$\frac{\begin{array}{c} \vdots \\ (t)\,\mathcal{F} \end{array}}{\mathcal{F}[t/T]}, \text{ provided that } T \text{ is free for } t \text{ in } \mathcal{F};$$

$$\frac{\begin{array}{c} \vdots \\ (\varphi)\,\mathcal{F} \end{array}}{\mathcal{F}[\varphi/\Phi]}, \text{ provided that } \Phi \text{ is free for } \varphi \text{ in } \mathcal{F}.$$

(3) (generalized $\Rightarrow$-elimination)

$$\frac{(t_k \mapsto \mathcal{F}_k)^{k\in\omega}(t) \qquad \mathcal{F}_k \Rightarrow \mathcal{G}_k \text{ for each } k \in \omega}{(t_k \mapsto \mathcal{G}_k)^{k\in\omega}(t)}$$

$$\frac{(\varphi_k \mapsto \mathcal{F}_k)^{k\in\omega}(\varphi) \qquad \mathcal{F}_k \Rightarrow \mathcal{G}_k \text{ for each } k \in \omega}{(\varphi_k \mapsto \mathcal{G}_k)^{k\in\omega}(\varphi)}$$

Let us note that each of the latter two meta-rules has infinitely many premisses.

Remark 7.31. Here is a brief sketch of an equivalent formalism. Instead of meta-quantifiers over term variables and k-formula variables, one could use their pre-limit versions (i.e. meta-quantifiers over variables of types $\mathbb{0}^k \to \mathbb{0}$ and $\mathbb{0}^l \times \mathbb{1}^{[k]} \to (\mathbb{0}^k \to \mathbb{1})$) along with a meta-quantifier over variables running over ω. All three sorts of meta-quantifiers come with natural specialization and generalization rules. Moreover, the first two types of meta-quantifiers subsume the original meta-quantifiers over individual and predicate variables (and similarly for the corresponding rules). Thus only three types of meta-quantifiers are needed, in contrast to the four types in the approach used above; and additionally the generalized $\Rightarrow$-elimination meta-rule is not needed.

So in some sense the alternative formalism is simpler; also it is closer to dependent type theory. On the other hand, the above formalism has precisely those four meta-quantifiers that we actually care about as primitive notions, and also avoids some excessive layers of indices by focusing on limit, rather than pre-limit types.

7.3.3. Enderton-style formulation. The extended meta-logic also admits a straightforward Enderton-style formulation.

Suppose that $\mathcal{F}$, $\mathcal{F}_k$, $\mathcal{G}$, $\mathcal{G}_k$ and $\mathcal{H}$ are meta-formula schemes, x is an individual variable, T is a term scheme, γ is an n-ary predicate variable, Φ is an n-formula scheme, t is a term variable, W is a term pre-scheme, φ is an n-formula variable and Ξ is an n-formula pre-scheme. Further suppose that $\vec{z}$ is a tuple of individual variables, $\vec{\beta}$ is tuple of predicate variables, $\vec{s}$ is a tuple of term variables and $\vec{\psi}$ is a tuple of i-formula variables.

(1) $(\vec{z})\,(\vec{\beta})\,(\vec{s})\,(\vec{\psi})\,\big(\mathcal{F} \Rightarrow \mathcal{G}\big)$, if $\mathcal{F}$ is α-equivalent to $\mathcal{G}$;

(2) $(\vec{z})\,(\vec{\beta})\,(\vec{s})\,(\vec{\psi})\,\big(\mathcal{F} \Rightarrow \mathcal{F}\big)$;

(3) $(\vec{z})\,(\vec{\beta})\,(\vec{s})\,(\vec{\psi})\,\big((\mathcal{F} \Rightarrow \mathcal{G})\,\&\,(\mathcal{G} \Rightarrow \mathcal{H}) \implies (\mathcal{F} \Rightarrow \mathcal{H})\big)$;

(4) $(\vec{z})\,(\vec{\beta})\,(\vec{s})\,(\vec{\psi})\,\Big(\big((\mathcal{F}\,\&\,\mathcal{G}) \Rightarrow \mathcal{H}\big) \iff \big(\mathcal{F} \Rightarrow (\mathcal{G} \Rightarrow \mathcal{H})\big)\Big)$;

(5) $(\vec{z})\,(\vec{\beta})\,(\vec{s})\,(\vec{\psi})\,\big(\mathcal{F}\,\&\,\mathcal{G} \Rightarrow \mathcal{F}\big)$ and $(\vec{z})\,(\vec{\beta})\,(\vec{s})\,(\vec{\psi})\,\big(\mathcal{F}\,\&\,\mathcal{G} \Rightarrow \mathcal{G}\big)$;

(6) $(\vec{z})\,(\vec{\beta})\,(\vec{s})\,(\vec{\psi})\,\big((\mathcal{H} \Rightarrow \mathcal{F})\,\&\,(\mathcal{H} \Rightarrow \mathcal{G}) \implies (\mathcal{H} \Rightarrow \mathcal{F}\,\&\,\mathcal{G})\big)$;

(7) $(\vec{z})(\vec{\beta})(\vec{s})(\vec{\psi})\big((x)\mathcal{F} \Rightarrow \mathcal{F}[x/T]\big)$, provided that T is free for x in $\mathcal{F}$;

(8) $(\vec{z})(\vec{\beta})(\vec{s})(\vec{\psi})\big((\gamma)\mathcal{F} \Rightarrow \mathcal{F}[\gamma/\Phi]\big)$, provided that Φ is free for γ in $\mathcal{F}$;

(9) $(\vec{z})(\vec{\beta})(\vec{s})(\vec{\psi})\big((t)\mathcal{F} \Rightarrow \mathcal{F}[t/W]\big)$, provided that W is free for t in $\mathcal{F}$;

(10) $(\vec{z})(\vec{\beta})(\vec{s})(\vec{\psi})\big((\varphi)\mathcal{F} \Rightarrow \mathcal{F}[\varphi/\Xi]\big)$, provided that Ξ is free for φ in $\mathcal{F}$;

(11) $(\vec{z})(\vec{\beta})(\vec{s})(\vec{\psi})\big((x)(\mathcal{F} \Rightarrow \mathcal{G}) \implies ((x)\mathcal{F} \Rightarrow (x)\mathcal{G})\big)$;

(12) $(\vec{z})(\vec{\beta})(\vec{s})(\vec{\psi})\big((\gamma)(\mathcal{F} \Rightarrow \mathcal{G}) \implies ((\gamma)\mathcal{F} \Rightarrow (\gamma)\mathcal{G})\big)$;

(13) $(\vec{z})(\vec{\beta})(\vec{s})(\vec{\psi})\big((t)(\mathcal{F} \Rightarrow \mathcal{G}) \implies ((t)\mathcal{F} \Rightarrow (t)\mathcal{G})\big)$;

(14) $(\vec{z})(\vec{\beta})(\vec{s})(\vec{\psi})\big((\varphi)(\mathcal{F} \Rightarrow \mathcal{G}) \implies ((\varphi)\mathcal{F} \Rightarrow (\varphi)\mathcal{G})\big)$;

(15) $(\vec{z})(\vec{\beta})(\vec{s})(\vec{\psi})\big(\mathcal{F} \Rightarrow (x)\mathcal{F}\big)$, provided that x does not occur freely in $\mathcal{F}$;

(16) $(\vec{z})(\vec{\beta})(\vec{s})(\vec{\psi})\big(\mathcal{F} \Rightarrow (\gamma)\mathcal{F}\big)$, provided that γ does not occur freely in $\mathcal{F}$;

(17) $(\vec{z})(\vec{\beta})(\vec{s})(\vec{\psi})\big(\mathcal{F} \Rightarrow (t)\mathcal{F}\big)$, provided that t does not occur freely in $\mathcal{F}$;

(18) $(\vec{z})(\vec{\beta})(\vec{s})(\vec{\psi})\big(\mathcal{F} \Rightarrow (\varphi)\mathcal{F}\big)$, provided that φ does not occur freely in $\mathcal{F}$;

(19) $\dfrac{\mathcal{F},\ \mathcal{F} \Rightarrow \mathcal{G}}{\mathcal{G}}$;

(20) $\dfrac{(t_k \mapsto \mathcal{F}_k)^{k\in\omega}(t), \quad \mathcal{F}_k \Rightarrow \mathcal{G}_k \text{ for each } k \in \omega}{(t_k \mapsto \mathcal{G}_k)^{k\in\omega}(t)}$;

(21) $\dfrac{(\varphi_k \mapsto \mathcal{F}_k)^{k\in\omega}(\varphi), \quad \mathcal{F}_k \Rightarrow \mathcal{G}_k \text{ for each } k \in \omega}{(\varphi_k \mapsto \mathcal{G}_k)^{k\in\omega}(\varphi)}$.

7.3.4. Principles and rules.

Proposition 7.32. *Let $\mathcal{F}(x)$ be a meta-formula scheme, where x is an individual variable and no individual variable occurs freely in $\mathcal{F}$, and let t be a term variable that does not occur freely in $\mathcal{F}$. Then the meta-formula scheme*

$$(x)\mathcal{F}(x) \iff (t)\left(t_k \mapsto (\varkappa^{0}_{[k]})\,\mathcal{F}\big(\mathfrak{E}_k(t_k)\big)\right)^{k\in\omega}(t)$$

is deducible.

Let us note that the β-reduction of $\mathcal{F}\big(\mathfrak{E}_k(t_k)\big)$ may require renaming of bound variables; in fact, it *must* require renaming of bound variables if bound variables ever occur in $\mathcal{F}$, since *each* individual variable occurs freely in some $\mathfrak{E}_k$.

Proof. By the $\Rightarrow$-introduction meta-rule, it suffices to deduce the left hand side from the right hand side, and conversely.

Suppose that $x = \varkappa^{0}_m$. By the meta-specialization for term variables, from the right hand side we obtain

$$\left(t_k \mapsto (\varkappa^{0}_{[k]})\,\mathcal{F}\big(\mathfrak{E}_k(t_k)\big)\right)^{k\in\omega}\big(\mathring{\mathfrak{l}}_m(\mathbb{p}_m)\big), \tag{$*$}$$

where $\mathring{\imath}_m \colon (\mathbb{0}^m \to \mathbb{0}) \to \mathbb{0}^*$ is the inclusion of the mth summand of $\mathbb{0}^* = \sum_{k \in \omega} (\mathbb{0}^k \to \mathbb{0})$ and $\mathbb{p}_m \colon \mathbb{0}^m \to \mathbb{0}$ is the projection onto the mth factor. The expression $(*)$ β-reduces to $\left(t_m \mapsto (\varkappa_{[m]}^{\mathbb{0}}) \, \mathcal{F}\big(\mathfrak{E}_m(t_m)\big) \right)(\mathbb{p}_m)$, which in turn β-reduces to $(\varkappa_{[m]}^{\mathbb{0}}) \, \mathcal{F}\big(\mathfrak{E}_m(\mathbb{p}_m)\big)$. Here $\mathfrak{E}_m(\mathbb{p}_m)$ β-reduces to $\mathbb{p}_m(\varkappa_{[m]}^{\mathbb{0}})$, and hence to $\varkappa_m^{\mathbb{0}} = x$. Thus we have deduced $(x)\,\mathcal{F}(x)$.

Conversely, by the meta-specialization for individual variables, from $(x)\,\mathcal{F}(x)$ we obtain $\mathcal{F}\big(\mathfrak{E}(t)\big)$, which β-expands to $\big(\mathcal{F} \circ \mathfrak{E}\big)(t)$. Here $\mathcal{F} \circ \mathfrak{E}$ is $\beta\eta$-equivalent to $\big(\mathcal{F} \circ \mathfrak{E}_k\big)^{k \in \omega}$ by Lemma 7.29. Now $\mathcal{F} \circ \mathfrak{E}_k$ is the same as $t_k \mapsto \mathcal{F}\big(\mathfrak{E}_k(t_k)\big)$, thus we have deduced $\left(t_k \mapsto \mathcal{F}\big(\mathfrak{E}_k(t_k)\big) \right)^{k \in \omega}(t)$. On the other hand, by the meta-generalization for individual variables and $\Rightarrow$-introduction $\mathcal{F}\big(\mathfrak{E}_k(t_k)\big) \Rightarrow (\varkappa_{[k]}^{\mathbb{0}}) \, \mathcal{F}\big(\mathfrak{E}_k(t_k)\big)$ is deducible for each $k \in \omega$. Hence by the generalized $\Rightarrow$-elimination we get $\left(t_k \mapsto (\varkappa_{[k]}^{\mathbb{0}}) \, \mathcal{F}\big(\mathfrak{E}_k(t_k)\big) \right)^{k \in \omega}(t)$. Finally, by the meta-generalization for term variables, from the latter meta-formula scheme we get $(t)\left(t_k \mapsto (\varkappa_{[k]}^{\mathbb{0}}) \, \mathcal{F}\big(\mathfrak{E}_k(t_k)\big) \right)^{k \in \omega}(t)$. $\qquad\square$

Similarly one can handle second-order quantification:

Proposition 7.33. *Let $\mathcal{F}(\gamma)$ be a meta-formula scheme, where γ is an r-ary predicate variable and no individual or predicate variables occur freely in $\mathcal{F}$, and let $\vec{\varphi}$ be an r-formula variable that does not occur freely in $\mathcal{F}$. Then the meta-formula scheme*

$$(\gamma)\,\mathcal{F}(\gamma) \iff (\varphi) \left(\varphi_k \mapsto (\varkappa_{[k]}^{\mathbb{0}}) \, (\varkappa_{[[k]]}^{\mathbb{1}}) \, \mathcal{F}\big(\mathfrak{E}_k^r(\varphi_k)\big) \right)^{k \in \omega}(\varphi)$$

is deducible.

Let us also formulate the multivariable case, using the notation of Remark 7.30:

Proposition 7.34. *Let $\mathcal{F}(\vec{x}, \vec{\gamma})$ be a meta-formula scheme, where $\vec{x}$ is an n-tuple of individual variables, $\vec{\gamma}$ is an m-tuple of predicate variables of arities $(r_1, \ldots, r_m) = \vec{r}$ and no individual or predicate variables occur freely in $\mathcal{F}$, and let $\vec{t}$ be an n-tuple of term variables that do not occur freely in $\mathcal{F}$ and $\vec{\varphi}$ be an m-tuple of r_i-formula variables that do not occur freely in $\mathcal{F}$. Then the meta-formula scheme*

$$(\vec{x})\,(\vec{\gamma})\,\mathcal{F}(\vec{x}, \vec{\gamma}) \iff (\vec{t})\,(\vec{\varphi}) \left(t_{\vec{k}}, \varphi_{\vec{l}} \mapsto (\varkappa_{[\vec{k},\vec{l}]}^{\mathbb{0}}) \, (\varkappa_{[[\vec{l}]]}^{\mathbb{1}}) \, \mathcal{F}\big(\mathfrak{E}_{\vec{k}}(t_{\vec{k}}), \mathfrak{E}_{\vec{l}}^{\vec{r}}(\varphi_{\vec{l}})\big) \right)^{(\vec{k},\vec{l}) \in \omega^n \times \omega^m}(\vec{t}, \vec{\varphi})$$

is deducible.

Here for $\vec{k} = (k_1, \ldots, k_n)$ we write $[\vec{k}] = [\max(k_1, \ldots, k_n)]$, $[[\vec{k}]] = [[\max(k_1, \ldots, k_n)]]$, $t_{\vec{k}} = (t_{k_1}, \ldots, t_{k_n})$ and $\varphi_{\vec{k}} = (\varphi_{k_1}, \ldots, \varphi_{k_n})$. The evaluations $\mathfrak{E}_{\vec{k}} \colon \prod_{i=1}^n (\mathbb{0}^{k_i} \to \mathbb{0}) \to \mathbb{0}^n$ and $\mathfrak{E}_{\vec{l}}^{\vec{r}} \colon \prod_{i=1}^m \big(\mathbb{0}^{l_i} \times \mathbb{1}^{[l_i]} \to (\mathbb{0}^{r_i} \to \mathbb{1})\big) \to \prod_{i=1}^m (\mathbb{0}^{r_i} \to \mathbb{1})$, where $\vec{l} = (l_1, \ldots, l_m)$, are defined respectively by $t_{\vec{k}} \mapsto \big(\mathfrak{E}_{k_1}(t_{k_1}), \ldots, \mathfrak{E}_{k_n}(t_{k_n})\big)$ and $\varphi_{\vec{l}} \mapsto \big(\mathfrak{E}_{l_1}^{r_1}(\varphi_{l_1}), \ldots, \mathfrak{E}_{l_m}^{r_m}(\varphi_{l_m})\big)$.

7.3.5. Semantics of the extended meta-logic. A *co-assignment* $\mathscr{U}$ assigns a semantic term $\mathscr{U}(x) \in |\mathscr{T}|$ to each term variable x. A *co-valuation* $\mathscr{V}$ assigns a semantic k-formula $\mathscr{V}(\gamma) \in |\mathscr{F}_k|$ to each k-formula variable γ, for each $k = 0, 1, \ldots$.

Given a co-assignment $\mathscr{U}$ and an assignment u along with an interpretation i of function symbols, we clearly get an interpretation of each term scheme T by an element

$|T|^i_{u\mathscr{U}} \in \mathcal{O}$, which does not depend on u (resp. on $\mathscr{U}$) if T contains no free occurrences of individual (resp. term) variables.

Given additionally a co-valuation $\mathscr{V}$ and a valuation v along with an interpretation $\imath$ of predicate constants, connectives and quantifiers, we clearly get an interpretation of each formula scheme F by an element $|F|^{v\mathscr{V}u\mathscr{U}}_\imath \in \mathcal{O}$, which does not depend on u (resp. on $\mathscr{U}$) if F contains no free occurrences of individual (resp. term) variables and does not depend on v (resp. on $\mathscr{V}$) if F contains no free occurrences of predicate (resp. k-formula) variables.

Given additionally also an interpretation $\jmath$ of meta-connectives and meta-quantifiers, we clearly also get an interpretation of each meta-formula scheme $\mathcal{F}$ by an element $|\mathcal{F}|^{v\mathscr{V}u\mathscr{U}}_\jmath \in \mathcal{Q}$, with the same dependence properties. In particular, a λ-closed meta-formula scheme $\mathcal{F}$ is interpreted by an element $|\mathcal{F}|_\jmath \in \mathcal{Q}$.

We can now generalize the definitions of §*4.7.1* to the case of the extended meta-logic. An absolute meta-rule is said to be *satisfied* in a meta-structure $\jmath$ if for every its instance

$$\frac{\mathcal{F}_1, \ldots, \mathcal{F}_n}{\mathcal{G}}$$

and every individual assignment u, co-assignment $\mathscr{U}$, predicate valuation v, and co-valuation $\mathscr{V}$, if $?(|\mathcal{F}_i|^{u\mathscr{U}v\mathscr{V}}_\jmath) = \top$ for each i, then $?(|\mathcal{G}|^{u\mathscr{U}v\mathscr{V}}_\jmath) = \top$. In particular, a meta-law is satisfied in $\jmath$ if and only $?(|\mathcal{G}|^{u\mathscr{U}v\mathscr{V}}_\jmath) = \top$ for every its instance $\mathcal{G}$ and arbitrary u, $\mathscr{U}$, v and $\mathscr{V}$.

A meta-interpretation $\jmath$ of the extended meta-logic is called a *model of the extended meta-logic* if it satisfies all meta-rules of the Enderton-style formulation of the extended meta-logic (see §*7.3.3*).

Now we define the *Kolmogorov translation* $\mathscr{F}$ of λ-closed meta-formulas into λ-closed meta-formula schemes:

- $\mathscr{F}$ sends each individual variable $\varkappa^{\mathbb{0}}_i$ to the corresponding term variable $\varkappa^{\mathbb{0}^*}_i$, and each k-ary predicate variable $\varkappa^{\mathbb{0}^k\to\mathbb{1}}_i$ to the corresponding k-formula variable $\varkappa^{\mathbb{1}^*_k}_i$.
- $\mathscr{F}$ sends both meta-connectives to themselves, the first-order meta-quantifier to the term meta-quantifier, and the k-ary second-order meta-quantifier to the k-formula meta-quantifier.
- If $\mathcal{G}$ is a λ-closed atomic $(n, \vec{r})$-meta-formula $\vec{x}, \vec{\gamma} \mapsto \text{!}F$, where F is a formula and $\vec{r} = (r_1, \ldots, r_m)$, then its Kolmogorov-image $\mathscr{F}(\mathcal{G})$ is the following λ-closed expression of type $(\mathbb{0}^*)^n \times \mathbb{1}^*_{r_1} \times \ldots \times \mathbb{1}^*_{r_m} \to \mathbb{\mu}$:

$$\left(t_{\vec{k}}, \varphi_{\vec{l}} \mapsto (\varkappa^{\mathbb{0}}_{[\vec{k},\vec{l}]})(\varkappa^{\mathbb{1}}_{[[\vec{l}]]}) \, \text{!}F\left(\mathfrak{E}_{\vec{k}}(t_{\vec{k}}), \mathfrak{E}^{\vec{r}}_{\vec{l}}(\varphi_{\vec{l}}) \right) \right)^{(\vec{k},\vec{l}) \in \omega^n \times \omega^m}.$$

If $\mathcal{F}$ is a λ-closed meta-formula, it is easy to see that $|\mathscr{F}(\mathcal{F})|_\jmath$ is precisely the alternative meta-interpretation $\{\!\{\mathcal{F}\}\!\}_\jmath$ as defined in §*7.2.2*.

Proposition *7.34* now implies that Examples *7.12* and *7.13* extend to the case of many-valued interpretations of the meta-logic:

Theorem 7.35. *If F, G be formulas and $\jmath$ a model of the extended meta-logic. Then*

(a) $\{\!\cdot F\}\!_{\mathscr{J}} = |\!\cdot F|_{\mathscr{J}}$.
(b) $?(|F \mathbin{/} G|_{\mathscr{J}}) = \top$ *implies* $?(\{F \mathbin{/} G\}_{\mathscr{J}}) = \top$.

An *alternative model of the meta-logic* is defined by applying a model of the extended meta-logic to the Kolmogorov translation from meta-logic to extended meta-logic. Given a logic L based on a language $\mathscr{L}$ and given by a derivation system $\mathcal{D}$, a meta-$\mathscr{L}$-structure M is called a *generalized alternative model* of L if it is an alternative model of the meta-logic and the meta-formula $\mathcal{D}$ is valid in M.

By Theorem 7.35, a principle is valid in a generalized alternative model if and only if it is valid in the corresponding generalized model; this clearly remains so if valuation fields or pre-valuation fields are used. But by Example 7.16, this is not the case for rules, at least if pre-valuation fields are used.

7.3.6. Meta-clarified BHK interpretation. This will be based on the alternative semantics of §7.2.2 and §7.3.5 with many-valued meta-interpretation of the meta-logic.

A domain $\mathscr{D}$ is fixed, and $\mathcal{O}$ is taken to be a class of problems, consisting of a prescribed class of contentful (e.g. mathematical) primitive problems and of composite problems, which are obtained inductively from the primitive ones by using contentual intuitionistic connectives and quantifiers over arbitrary $\mathscr{D}$-indexed families of parametric problems. Formal intuitionistic connectives and quantifiers are interpreted straightforwardly by the contentual ones.

$\mathcal{Q}$ is taken to be a larger class of problems, consisting of those in $\mathcal{O}$, called "meta-primitive" (thus the function $!: \mathcal{O} \to \mathcal{Q}$ is the inclusion) and of "meta-composite" problems obtained inductively from the meta-primitive ones by using contentual intuitionistic connectives and quantifiers over arbitrary $\mathscr{D}$-indexed, $\mathrm{Hom}(\mathscr{D}^i, \mathcal{O})$-indexed $(i = 0, 1, \dots)$, $|\mathscr{T}|$-indexed and $|\mathscr{F}_i|$-indexed $(i = 0, 1, 2, \dots)$ families of problems. (We recall that $|\mathscr{T}|$ and $|\mathscr{F}_i|$ are defined in terms of $\mathscr{D}$ and $\mathcal{O}$, see §7.2.2.) Meta-connectives and meta-quantifiers are Tarski-interpreted by contentual intuitionistic connectives and $\mathscr{D}$-indexed and $\mathrm{Hom}(\mathscr{D}^i, \mathcal{O})$-indexed quantifiers in the straightforward way; and meta-quantifiers are also alternatively-interpreted by contentual intuitionistic $|\mathscr{T}|$-indexed and $|\mathscr{F}_i|$-indexed quantifiers.

The function $?: \mathcal{Q} \to \{\top, \bot\}$ sends a problem Θ to $\top$ if and only if there exists a solution of Θ. What is meant by solving a problem Θ is defined explicitly for primitive $\Theta \in \mathcal{O}$ and is extended inductively to composite $\Theta \in \mathcal{O}$ and then further to all meta-composite $\Theta \in \mathcal{Q}$ by the usual BHK-interpretation (see §3.7). This determines a contentual interpretation of intuitionistic logic and its meta-logic, which we will call the *meta-clarified BHK* interpretation.

Example 7.36. Let us record the meta-clarified BHK interpretations of the judgements of admissibility, stable admissibility and derivability of a purely logical rule $\Gamma \mathbin{/} \Delta$. The tuples $\vec{\varphi}$ and $\vec{\Phi}$ have the same meaning as in Example (4.84). Like before, $?\Theta$ denotes the proposition "The problem Θ has a solution".

- Derivability: $\forall \mathscr{D}, \mathcal{O} \ ?\forall \mathscr{V} \ (\forall v, u \ |\Gamma|_u^{\mathscr{V}[u,v]} \ \longrightarrow \ \forall v, u \ |\Delta|_u^{\mathscr{V}[u,v]})$;

- Stable admissibility: $\forall\vec{\Phi}\,\forall\mathscr{D}, \mathscr{O}\ ?(\forall v, u\ |\Gamma[\vec{\varphi}/\vec{\Phi}]|_u^v \longrightarrow \forall v, u\ |\Delta[\vec{\varphi}/\vec{\Phi}]|_u^v)$;
- Admissibility: $\forall\vec{\Phi}\ (\forall\mathscr{D}, \mathscr{O}\ ?\forall v, u\ |\Gamma[\vec{\varphi}/\vec{\Phi}]|_u^v \longrightarrow \forall\mathscr{D}, \mathscr{O}\ ?\forall v, u\ |\Delta[\vec{\varphi}/\vec{\Phi}]|_u^v)$.

Remark 7.37. Let us note that there is an obvious variation of the Verificationist interpretation with alternative interpretation of the meta-level, similar to the one in the meta-clarified BHK interpretation.

References

[1] S. A. Melikhov, *A Galois connection between classical and intuitionistic logics. I: Syntax.* arXiv:1312.2575. ↑**1.2**, **3.3**, 19, 21, *3.8.4*, 31, 5.1, 82

[2] ———, *A Galois connection between classical and intuitionistic logics. II: Semantics.* arXiv:1504.03379. ↑**1.2**, 21, 31, 5.1

[3] ———, *A Galois connection between classical and intuitionistic logics. III: Euclid's "Elements" as a theory over QHC.* In preparation; see arXiv:1504.03379v2 (not the current v3) §1A, §3 and §4 for a preliminary version. ↑**1.2**, 15, **3.7**, 31, 62

[4] C. Hog-Angeloni, W. Metzler, and A. Sieradski (eds.), *Two-Dimensional Homotopy and Combinatorial Group Theory*, London Math. Soc. Lect. Note Ser., vol. 197, 1993. ↑6

[5] P. Aczel, *The Russell–Prawitz modality*, Math. Structures Comput. Sci. **11** (2001), 541–554. Author's homepage. ↑67, *4.6.3*

[6] S. N. Artëmov, *Kolmogorov and Gödel's approach to intuitionistic logic, and papers in this direction in the last decade*, Uspekhi Mat. Nauk **59** (2004), no. 2, 9–36; English transl., Russian Math. Surveys **59** (2004), no. 2, 203–229. Mathnet. ↑

[7] S. Artëmov and E. Nogina, *The topology of justification*, Logic Log. Philos. **17** (2008), no. 1-2, 59–71. Journal. ↑

[8] A. Avron, *Simple consequence relations*, Inform. and Comput. **92** (1991), no. 1, 105–139. Author's homepage; Journal. ↑*4.1.1*, *4.1.4*, *4.1.4*

[9] ———, *What is a logical system?*, What is a logical system?, Stud. Logic Comput., vol. 4, Oxford Univ. Press, New York, 1994, pp. 217–238. ↑*4.1.4*, *4.1.4*

[10] ———, *The rule of generalization in FOL* (2004). FOM mailing list. ↑*4.1.4*, *4.1.4*

[11] S. Awodey, *Topological representation of the λ-calculus*, Math. Structures Comput. Sci. **10** (2000), 81–96. Carnegie Mellon Univ. Repository. ↑**1.1**, 33

[12] ———, *Type theory and homotopy.* arXiv:1010.1810. ↑

[13] S. Awodey and C. Butz, *Topological Completeness for Higher-Order Logic*, Journal of Symbolic Logic **65** (2000), 1168–1182. arXiv:math.LO/9707206. ↑**1.1**

[14] S. Awodey and A. Bauer, *Propositions as [types]*, J. Logic Comput. **14** (2004), no. 4, 447–471. Institut Mittag-Leffler, Preprint No. 34 (2000/01); Bauer's homepage. ↑

[15] S. Awodey and M. A. Warren, *Homotopy theoretic models of identity types*, Math. Proc. Cambridge Philos. Soc. **146** (2009), no. 1, 45–55. arXiv:0709.0248. ↑

[16] M. E. Baron, *A note on the historical development of logic diagrams: Leibniz, Euler and Venn*, Math. Gazette **53** (1969), 113–125. ↑*4.8.2*

[17] M. J. Beeson, *Problematic principles in constructive mathematics*, Logic Colloquium '80 (Prague, 1980), Stud. Logic Foundations Math., vol. 108, North-Holland, Amsterdam-New York, 1982. ↑**3.7**, 17

[18] M. J. Beeson, *Foundations of Constructive Mathematics*, Ergebn. der Math. (3), vol. 6, Springer-Verlag, Berlin, 1985. ↑

[19] L. D. Beklemishev, *Induction rules, reflection principles, and provably recursive functions*, Ann. Pure Appl. Logic **85** (1997), no. 3, 193–242. ScienceDirect, Utrecht Univ. Logic Group Preprint Series, vol. 168. ↑5.13

[20] L. Beklemishev and A. Visser, *Problems in the logic of provability*, Mathematical Problems from Applied Logic. I, Int. Math. Ser. (N. Y.), vol. 4, Springer, New York, 2006, pp. 77–136. Utrecht Univ. Logic Group Preprint Series, vol. 235. ↑**2.1**, **2.1**

[21] S. Berardi, M. Bezem, and T. Coquand, *On the computational content of the axiom of choice*, J. Symbolic Logic **63** (1998), 600–622. Utrecht Univ. Logic Group Preprint Series, vol. 116. ↑**3.3**

[22] E. Bishop, *Foundations of constructive analysis*, McGraw-Hill, New York, 1967. ↑80

[23] A. Blass, N. Dershowitz, and Yu. Gurevich, *When are two algorithms the same?*, Bull. Symbolic Logic **15** (2009), no. 2, 145–168. `arXiv:0811.0811`. ↑**2.1**

[24] W. J. Blok and D. Pigozzi, *Abstract algebraic logic and the deduction theorem* (2001). Pigozzi's homepage. ↑*4.1.1*, *4.1.4*

[25] G. E. Bredon, *Sheaf Theory*, 2nd ed., Grad. Texts in Math., vol. 170, Springer-Verlag, New York, 1997; Russian transl. in *Теория пучков*, Наука, М., 1988. ↑6

[26] L. E. J. Brouwer, *Over de grondslagen der wiskunde (PhD Dissertation)*, Maas & van Suchtelen, Amsterdam/Leipzig, 1907. archive.org. English transl., *On the foundations of mathematics*, Collected Works, Vol. 1: Philosophy and Foundations of Mathematics (A. Heyting, ed.), North-Holland/Elsevier, Amsterdam/New York, 1975, pp. 13–101. ↑**3.5**, 11

[27] _____, *De onbetrouwbaarheid der logische principes*, Tijdschrift voor Wijsbegeerte **2** (1908), 152–158. archive.org; English transl., *The unreliability of the logical principles*, Collected Works, Vol. 1: Philosophy and Foundations of Mathematics (A. Heyting, ed.), North-Holland/Elsevier, Amsterdam/New York, 1975, pp. 107–111. ↑**3.5**, **3.5**

[28] L. E. J. Brouwer, *Intuitionistische Betrachtungen über den Formalismus*, K. Ned. Akad. Wet. Proc. Sect. Sci. **31** (1928), 374–379. KNAW; Also published in Sitz.ber. Preuss. Akad. Wiss., Phys.-math. Kl. (1928), 48–52. English transl.: *Intuitionist reflections on formalism*, From Brouwer to Hilbert: The debate on the foundations of mathematics in the 1920s (P. Mancosu, ed.), Oxford University Press, New York, 1998, pp. 40–44; §1 also translated in: From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931 (J. van Heijenoort, ed.), Harvard Univ. Press, Cambridge, MA, 1967, pp. 490–492. ↑**3.5**, **3.5**, **3.11**, 28, **3.12**

[29] Chad E. Brown and C. Rizkallah, *Glivenko and Kuroda for simple type theory*. Brown's homepage. ↑**5.6**

[30] Richard A. Brown, *Generalized group presentation and formal deformations of CW complexes*, Trans. Amer. Math. Soc. **334** (1992), no. 2, 519–549. Journal. ↑6

[31] A. V. Chernavsky and V. P. Leksine, *Unrecognizability of manifolds*, Ann. Pure Appl. Logic **141** (2006), no. 3, 325–335. Journal. ↑32

[32] A. Church, *Introduction to Mathematical Logic*, 2nd ed., Princeton Univ. Press, Princeton, NJ, 1956; English transl. in ИЛ, М., 1960. ↑**3.3**, *4.1.3*, *4.1.3*, *4.1.4*, *4.4.1*

[33] A. I. Citkin, *Admissible rules of intuitionistic propositional logic*, Mat. Sb. **102** (1977), no. 2, 314–323; Mathnet; English transl., Math. USSR – Sbornik, 1977 **31** (1977), no. 2, 279–288. ↑*4.5.3*, *5.4.1*

[34] T. Coquand, *La contribution de Kolmogorov en logique intuitionniste*, L'héritage de Kolmogorov en mathématiques (É. Charpentier, A. Lesne, and N. Nikolski, eds.), Belin, Paris, 2003; English transl., *Kolmogorov's contribution to intuitionistic logic*, Kolmogorov's heritage in mathematics, Springer, Berlin, 2007, pp. 19–40. ↑**3.3**, **3.3**, **3.3**, **3.7**, *3.8.4*, **5.6**

[35] J. Conant and O. Thistlethwaite, *Boolean formulae, hypergraphs and combinatorial topology*, Topology Appl. **157** (2010), no. 16, 2449–2461. `arXiv:0808.0739`. ↑**2.1**

[36] P.-L. Curien and R. Di Cosmo, *A confluent reduction for the λ-calculus with surjective pairing and terminal object*, Automata, Languages and Programming (Madrid, 1991), Lecture Notes in Comput. Sci., vol. 510, Springer, Berlin, 1991, pp. 291–302. Di Cosmo's website. ↑4.18, *7.3.1*

[37] D. de Jongh, *The Disjunction Property according to Kreisel and Putnam* (2009). Author's homepage. ↑*5.4.1*, *5.4.1*

[38] R. Di Cosmo and D. Kesner, *A confluent reduction for the extension-typed λ-calculus with pairs, sums, recursion and terminal object*, Automata, Languages and Programming (Lund, 1993), Lecture Notes in Comput. Sci., vol. 700, Springer, Berlin, 1993, pp. 645–656. Di Cosmo's website. ↑4.18

[39] J. Diller and A. S. Troelstra, *Realizability and intuitionistic logic*, Synthese **60** (1984), no. 2, 253–282. ↑**3.7**, **3.7**

[40] K. Došen, *The first axiomatization of relevant logic*, J. Philos. Logic **21** (1992), no. 4, 339–356. Author's homepage. ↑**3.3**

[41] ──────, *Logical consequence: a turn in style*, Logic and Scientific Methods (Florence, 1995), Synthese Lib., vol. 259, Kluwer, Dordrecht, 1997, pp. 289–311. Author's homepage. ↑

[42] ──────, *Identity of proofs based on normalization and generality*, Bull. Symbolic Logic **9** (2003), no. 4, 477–503. arXiv:math.LO/0208094. ↑**2.1**, 2.1, **2.2**

[43] K. Došen and Z. Petrić, *Proof-Theoretical Coherence*, King's College Publications, London, 2004. Revised version (2007), Došen's homepage. ↑**2.2**

[44] ──────, *Equality of proofs for linear equality*, Arch. Math. Logic **47** (2008), no. 6, 549–565. arXiv:math.LO/0601490. ↑**2.2**

[45] G. Dowek, *On the definition of the classical connectives and quantifiers*. Author's homepage. ↑**5.6**

[46] A. G. Dragalin, *Mathematical Intuitionism. Introduction to Proof Theory*, Nauka, Moscow, 1979; English transl. in Transl. Math. Mono., vol. 67, Amer. Math. Soc., Providence, RI, 1988. ↑4.64, **5.3**, 75

[47] W. Dzik, *Chains of structurally complete predicate logics with the application of Prucnal's substitution*, Reports on Math. Logic **38** (2004), 37–48. ↑*4.1.2*, 4.70

[48] H. B. Enderton, *A Mathematical Introduction to Logic*, 2nd ed., Harcourt/Academic Press, Burlington, MA, 2001. ↑*4.1.4*, 57

[49] R. Engelking, *General Topology*, 2nd ed., Heldermann Verlag, Berlin, 1989; Russian transl. in Mir, Moscow, 1986. Translated from 1st ed. and a draft of 2nd ed.; cited theorem numbers refer to this translation. ↑**5.4**, **5.4**, *6.2.2*, *6.2.2*

[50] L. Esakia, *Intuitionistic logic and modality via topology*, Ann. Pure Appl. Logic **127** (2004), no. 1-3, 155–170. Journal. ↑*5.7.2*, *5.7.2*

[51] M. Escardo, *Synthetic topology of data types and classical spaces*, Electronic Notes in Theoretical Computer Science **87** (2004), 21-156. Author's homepage. ↑83

[52] ──────, *Infinite sets that satisfy the principle of omniscience in any variety of constructive mathematics*, J. Symbolic Logic **78** (2013), 764–784. Author's homepage. ↑5.16

[53] S. Feferman, *Reflecting on incompleteness*, J. Symbolic Logic **56** (1991), 1–49. ↑*4.1.3*

[54] G. Fernández Díez, *Kolmogorov, Heyting and Gentzen on the intuitionistic logical constants*, Crítica, Revista Hispanoamericana de Filosofía **32** (2000), 48–57. Journal. ↑**3.7**

[55] G. Ferreira and P. Oliva, *On various negative translations*, Electronic Proc. in Theoret. Comp. Sci. **47** (2011), 21–33. ↑81

[56] R. C. Flagg and H. Friedman, *Epistemic and intuitionistic formal systems*, Ann. Pure Appl. Logic **32** (1986), no. 1, 53–60. Journal. ↑*5.7.3*

[57] G. Frege, *Begriffsschrift*, Halle, 1879. English translations: (i) From Frege to Gödel. A Source Book in Mathematical Logic, 1879–1931 (J. van Heijenoort, ed.), Harvard Univ. Press, Cambridge, MA, 1967, pp. 1–82; (ii) G. Frege, Conceptual Notation and Related Articles (T. W. Bymun, ed.), Oxford Univ. Press, 1972. ↑

[58] H. Friedman, *The consistency of classical set theory relative to a set theory with intuitionistic logic*, J. Symbolic Logic **38** (1973), 315–319. ↑**3.3**

[59] D. M. Gabbay, *Semantical investigations in Heyting's intuitionistic logic*, Synthese Library, vol. 148, D. Reidel Publishing Co., Dordrecht-Boston, 1981. ↑*5.3.3*

[60] D. M. Gabbay, V. B. Shehtman, and D. P. Skvortsov, *Quantification in Nonclassical Logic*, Elsevier, Amsterdam, 2009. Shehtman's homepage. ↑

[61] M. J. Gabbay and D. P. Mulligan, *Two-level lambda-calculus*, Proc. 17th Intern. Workshop on Functional and (Constraint) Logic Programming (WFLP 2008) (R. Echahed, ed.), Electr. Notes in Theoret. Computer Sci., vol. 246, Elsevier, Amsterdam, 2008. M. Gabbay's website, Journal. ↑*4.1.2*

[62] J. Geanakoplos, *Common knowledge*, Handbook of Game Theory with Economic Applications, Vol. II, Handbooks in Econom., vol. 11, North-Holland, Amsterdam, 1994, pp. 1437–1496. Author's homepage. ↑

[63] F. Gilda and P. Oliva, *On various negative translations*, Electr. Proc. in Theor. Comp. Sci. **47** (2011), 21–33. arXiv:1101.5200. ↑

[64] J.-Y. Girard, P. Taylor, and Y. Lafont, *Proofs and Types*, Cambridge Univ. Press, Cambridge, 1989. Online version. ↑**2.1**, **2.2**, **2.2**, 2.1, **2.3**, **2.3**, **3.7**, **3.7**, **4.2**, *4.4.1*, *4.6.3*

[65] G. Gentzen, *Untersuchungen über das logische Schließen*, Math. Z. **39** (1935), 176–210, 405–431. English transl., *Investigations into logical deduction*, The Collected Papers of Gerhard Gentzen, M. E. Szabo, ed., North-Holland, Amsterdam, 1969, pp. 68–131; Russian transl. in *Математическая теория логического вывода,* сб. пер. под ред. А. В. Идельсона и Г. Е. Минца, Наука, М., 1967, стр. 9–74. ↑**2.2**, **2.2**, **2.2**, **3.7**

[66] R. Godement, *Topologie algébrique et théorie des faisceaux*, Actualités Sci. Ind., vol. 1252, Hermann, Paris, 1958 (French); Russian transl. in *Алгебраическая топология и теория пучков*, ИЛ, М., 1961. ↑6, 84

[67] P. G. Goerss and J. F. Jardine, *Simplicial Homotopy Theory*, Progress in Math., vol. 174, Birkhäuser, Basel, 1999. ↑5, 6

[68] K. Gödel, *Lecture at Zilsel's*, Collected Works. Vol. III, The Clarendon Press, Oxford Univ. Press, New York, 1995, pp. 86–113. ↑*5.7.2*

[69] ———, *Eine Interpretation des intuitionistischen Aussagenkalküls*, Ergebn. eines math. Kolloq. **4** (1933), 39–40. English transl., *An interpretation of the intuitionistic propositional calculus,* The philosophy of mathematics (J. Hintikka, ed.), Oxford University Press, London, 1969, pp. 128–129; Collected Works. Vol. I, The Clarendon Press, Oxford Univ. Press, New York, 1995, pp. 300–303. ↑82

[70] R. Goldblatt, *Topoi. The Categorial Analysis of Logic*, 2nd ed., Studies in Logic and the Found. of Math., vol. 98, North-Holland, Amsterdam, 1984; Russian transl. in *Топосы. Категорный анализ логики*, Мир, М., 1983. ↑

[71] W. Goldfarb, *Frege's conception of logic*, Future pasts (J. Floyd and S. Shieh, eds.), Oxford Univ. Press, 2001. The Analytic Tradition in Twentieth-Century Philosophy, pp. 25–41. ↑17

[72] T. Gowers *et. al., When are two proofs essentially the same?*. Gowers's Weblog. ↑**2.1**

[73] A. Guglielmi, *On Lafont's counterexample* (2002). Author's homepage. ↑**2.2**

[74] ———, *Normalisation without cut elimination* (2003). Author's homepage. ↑**2.2**

[75] Y. Guiraud, *The three dimensions of proofs*, Ann. Pure Appl. Logic **141** (2006), no. 1-2, 266–295. arXiv:math.CT/0612089. ↑**2.2**

[76] R. Hakli and S. Negri, *Does the deduction theorem fail for modal logic?*, Synthese **187** (2012), 849–867. Negri's homepage. ↑*4.1.4*, *4.1.4*

[77] G. Havas and C. Ramsay, *On proofs in finitely presented groups*, Groups St. Andrews 2005. Vol. 2, London Math. Soc. Lecture Note Ser., vol. 340, Cambridge Univ. Press, Cambridge, 2007, pp. 457–474. Havas' Homepage. ↑**2.1**

[78] G. Helman, *On the equivalence of proofs involving identity*, Notre Dame J. Formal Logic **28** (1987), no. 3, 297–321. ProjectEuclid. ↑**2.1**

[79] A. Heyting, *Sur la logique intuitionniste*, Bull. Acad. Royale Belgique **16** (1930), 957–963. English transl., *On intuitionistic logic*, in From Brouwer to Hilbert: The Debate on the Foundations of Mathematics in the 1920s (P. Mancosu, ed.), Oxford University Press, New York, 1998, pp. 306–310. ↑**3.3**, 29

[80] ———, *Die intuitionistische Grundlegung der Mathematik*, Erkenntnis **2** (1931), 106–115. English transl., *The intuitionist foundations of mathematics*, Philosophy of Mathematics: Selected Readings, 2nd. ed., P. Benacerraf and H. Putnam (eds.), Cambridge Univ. Press, New York, 1983, pp. 52–65. ↑**3.2**, **3.3**, **3.5**, **3.7**, 14, **3.9**, **3.9**

[81] ———, *Mathematische Grundlagenforschung. Intuitionismus. Beweistheorie*, Springer, Berlin, 1934, reprinted 1974 (German); Expanded French transl. in *Les fondements des mathematiques. Intuitionnisme. Theorie de la demonstration*, Gauthier-Villars, Paris and E. Nauwelaerts, Louvain, 1955. Russian transl., *Обзор исследований по основаниям математики. Интуиционизм — теория доказательства,* ОНТИ НКТП, М., 1936. ↑**3.7**, **3.7**, **3.7**, *3.8.4*, 3.4

[82] ———, *Intuitionism. An Introduction*, North-Holland Publishing Co., Amsterdam, 1956; Russian transl. in *Интуиционизм*, Мир, М., 1965. ↑**3.7**

[83] ———, *Die formalen Regeln der intuitionistischen Logik*, Sitzungsber. Preuß. Akad. Wiss., Phys.-Math. Kl. (1930), 42–56. English transl., *The formal rules of intuitionistic logic*, in From Brouwer to Hilbert: The Debate on the Foundations of Mathematics in the 1920s (P. Mancosu, ed.), Oxford University Press, New York, 1998, pp. 311–327. ↑3.4, *4.6.1*

[84] D. Hilbert, *Theorie der algebraischen Invarianten nebst Anwendungen auf Geometrie,* Lecture notes from Summer 1897 prepared by Sophus Marxsen; English transl. in *Theory of Algebraic Invariants* (R. C. Laubenbacher and B. Sturmfels, eds.), Cambridge Univ. Press, Cambridge, 1993. ↑**2.1**

[85] ———, *Mathematische Probleme*, Nachrichten von der Königl. Gesellschaft der Wiss. zu Göttingen (1900), 253–297. GDZ. Also published in: Archiv der Math. Phys. (3) **1** (1901), 44–63, 213–237; English transl. *Mathematical problems*, Bull. Amer. Math. Soc. **8** (1902), 437–479, Journal. ↑**2.1**, **3.5**, **3.5**, 12

[86] ———, *Neubegründung der Mathematik. Erste Mitteilung*, Abh. Math. Seminar Hamburg. Univ. **1** (1922), 157–177. English transl., *The New Grounding of Mathematics. First Report*, From Kant to Hilbert: Readings in the Foundations of Mathematics (W. B. Ewald, ed.), Oxford Univ. Press, Oxford, 1996, pp. 1115–1133; reprinted in, From Brouwer to Hilbert: The Debate on the Foundations of Mathematics in the 1920s (P. Mancosu, ed.), Oxford Univ. Press, New York, 1998, pp. 198–214. ↑

[87] ———, *Die Grundlagen der Mathematik*, Abh. Math. Seminar Hamburg. Univ. **6** (1928), 65–85. English transl., *The Foundations of Mathematics*, From Frege to Gödel. A Source Book in Mathematical Logic, 1879–1931 (J. van Heijenoort, ed.), Harvard Univ. Press, Cambridge, MA, 1967, pp. 464–479. ↑

[88] ———, *Die Grundlagen Der Elementaren Zahlentheorie*, Math. Ann. **104** (1931), 485–94. English transl. *The Grounding of Elementary Number Theory*, From Brouwer to Hilbert: The Debate on the Foundations of Mathematics in the 1920s (P. Mancosu, ed.), Oxford University Press, New York, 1998, pp. 266–273. ↑79

[89] D. Hilbert and W. Ackermann, *Grundzüge der Theoretischen Logik*, 2nd ed., Springer, Berlin, 1928/38 (German). English transl., *Principles of mathematical logic,* Chelsea, New York, 1950; Russian transl., Д. Гильберт, В. Аккерман, *Основы теоретической логики,* ИЛ, Москва, 1947. ↑**1.1**, *3.10.1*, 3.3, *4.1.3*, 57

[90] D. Hilbert and P. Bernays, *Grundlagen der Mathematik. I*, 2nd ed., Springer, Berlin, 1934/68 (German); Russian transl. in *Основания математики. Логические исчисления и формализация арифметики*, Наука, Москва, 1979. ↑**1.1**, *3.10.1*, **3.3**, *4.1.3*, 57

[91] B. P. Hilken, *Towards a proof theory of rewriting: the simply typed* 2λ*-calculus*, Theoret. Comput. Sci. **170** (1996), no. 1-2, 407–444. ↑**2.1**

[92] The Univalent Foundations Program, *Homotopy Type Theory: Univalent Foundations of Mathematics*, Institute for Advanced Study, 2013. Online edition at HomotopyTypeTheory.org. ↑**2.1**, 5

[93] W. A. Howard, *The formulae-as-types notion of construction*, To H. B. Curry: essays on combinatory logic, lambda calculus and formalism, Academic Press, London–New York, 1980, pp. 480–490. Reprint of a 1969 manuscript. ↑**3.7**, 33

[94] L. Humberstone, *The connectives*, MIT Press, Cambridge, MA, 2011. ↑**5.7**

[95] R. Iemhoff, *On the admissible rules of intuitionistic propositional logic*, J. Symbolic Logic **66** (2001), 281–294. Author's homepage. ↑**5.8**, *5.4.1*, *5.4.1*

[96] _____, *Intermediate logics and Visser's rules*, Notre Dame J. Formal Logic **46** (2005), 65–81. Author's homepage. ↑*5.4.1*

[97] _____, *On the rules of intermediate logics*, Arch. Math. Logic **45** (2006), 581–599. Author's homepage. ↑*5.4.1*

[98] _____, *On rules*, J. Philos. Logic (to appear). Author's homepage. ↑*5.3.2*

[99] _____, *A note on consequence*, J. Philos. Logic (to appear). Author's homepage. ↑*4.1.2*, *5.4.1*

[100] J. R. Isbell, *Uniform Spaces*, Amer. Math. Soc., Providence, RI, 1964. ↑

[101] G. Japaridze and D. de Jongh, *The logic of provability*, Handbook of Proof Theory (S. R. Buss, ed.), Studies in Logic and the Found. of Math., vol. 137, North-Holland, Amsterdam, 1998, pp. 475–545. Japaridze's homepage. ↑8, *5.7.2*, *5.7.2*, *5.7.2*

[102] E. Jeřábek, *Canonical rules*, J. Symbolic Logic **74** (2009), 1171–1205. Author's homepage. ↑*4.5.3*, *4.5.4*

[103] S. C. Kleene, *Introduction to Metamathematics*, D. Van Nostrand Co., New York, 1952; Russian transl. in *Введение в метаматематику*, ИЛ, М., 1957. ↑25, *4.1.2*, *4.1.4*, *4.1.4*, *4.3.3*, *4.4.1*, 4.48, *5.2.1*, *5.3.1*, *7.1.4*

[104] _____, *Disjunction and existence under implication in elementary intuitionistic formalisms*, J. Symbolic Logic **27** (1962), 11–18. ↑*4.6.2*

[105] J. W. Klop and R. C. de Vrijer, *Unique normal forms for lambda calculus with surjective pairing*, Inform. and Comput. **80** (1989), 97–113. ScienceDirect. ↑

[106] T. Koetsier, *Hilbert's twenty-fourth problem*, Nieuw Arch. Wiskd. (5) **2** (2001), no. 1, 65–67 (Dutch). ↑**2.1**

[107] A. N. Kolmogorov, *On the principle tertium non datur*, Mat. Sbornik **32** (1925), 646–667. Mathnet; English translations: (i) From Frege to Gödel. A Source Book in Mathematical Logic, 1879–1931 (J. van Heijenoort, ed.), Harvard Univ. Press, Cambridge, MA, 1967, pp. 414–447; (ii) Selected Works of A. N. Kolmogorov. Vol. I, Mathematics and its Applications (Soviet Series), vol. 25, Kluwer, Dordrecht, 1991, pp. 151–158. ↑**3.3**, *3.8.4*, *4.6.1*, **5.6**

[108] A. Kolmogoroff, *Zur Deutung der intuitionistischen Logik*, Math. Z. **35** (1932), no. 1, 58–65 (German). GDZ; Russian transl. in *Колмогоров А. Н. Избранные труды. Математика и механика,* Наука, М. (1985), стр. 142–148; English translations: (i) *On the interpretation of intuitionistic logic*, Selected Works of A. N. Kolmogorov. Vol. I, Mathematics and its Applications (Soviet Series), vol. 25, Kluwer, Dordrecht, 1991, pp. 151–158; (ii) From Brouwer to Hilbert: The debate on the foundations of mathematics in the 1920s (P. Mancosu, ed.), Oxford University Press, New York, 1998, pp. 328–334; (iii) by J. McKinna (2014). ↑**1.1**, 2, **3.4**, **3.4**, **3.7**, 14, 19, *3.8.4*, **3.9**, **3.11**, 3.4, 29, *4.3.4*, *7.2.1*, *7.2.1*

[109] *Letters from A. N. Kolmogorov to A. Heyting*, Uspekhi Mat. Nauk **43** (1988), no. 6, 75–77 (Russian); English transl., Russian Math. Surveys **43** (1988), no. 6, 89–93. Mathnet; The original (in French and German) is found in: A. Troelstra, *On the early history of intuitionistic logic*, in Mathematical Logic (P. Petkov, ed.), New York, Plenum Press, 1990, pp. 3—17; Univ. of Amsterdam Preprint ML-88-04. ↑2

[110] A. N. Kolmogorov, *Contemporary debates on the nature of mathematics*, Nauchnoe slovo (1929), no. 6, 41–54 (Russian). Reprinted in: Problemy Peredachi Informatsii **42** (2006), no. 4, 129–141; Mathnet; English transl., Probl. Inf. Transm. **42** (2006), 379–389. ↑**3.5**

[111] A. N. Kolmogorov and A. G. Dragalin, *Introduction to Mathematical Logic*, Izdat. Mosk. Univ., Moscow, 1982 (Russian). ↑*4.1.4*

[112] G. Kreisel, *A survey of proof theory. II*, Proc. Second Scandinavian Logic Sympos. (Oslo, 1970), Stud. Logic Found. Math., vol. 63, North-Holland, Amsterdam, 1971, pp. 109–170; Russian transl., Исследования по теории доказательств, Мир, М., 1981, pp. 114–182. ↑**2.1**

[113] P. Kremer, *Strong completeness of S4 wrt the real line* (2012). Author's homepage. ↑**5.3**

[114] _____, *Strong completeness of S4 for any dense-in-itself metric space*, Rev. Symb. Log. **6** (2013), no. 3, 545–570. Author's homepage. ↑**5.3**

[115] _____, *Quantified modal logic on the rational line*, Rev. Symb. Log. **7** (2014), no. 3, 439–454. Author's homepage. ↑*5.7.1*

[116] _____, *Quantified intuitionistic logic over metrizable spaces* (2014). Author's homepage. ↑**5.3**, **5.3**

[117] R. Kuznets, *Proof identity for classical logic: generalizing to normality*, Logical Foundations of Computer Science, Lecture Notes in Comput. Sci., vol. 4514, Springer, Berlin, 2007, pp. 332–348. ↑**2.2**

[118] _____, *Self-referential justifications in epistemic logic*, Theory Comput. Syst. **46** (2010), 636–661. Author's homepage. ↑*5.7.2*

[119] H. Läuchli, *An abstract notion of realizability for which intuitionistic predicate calculus is complete*, Intuitionism and Proof Theory (Proc. Conf., Buffalo, N.Y., 1968), North-Holland, Amsterdam, 1970, pp. 227–234. ↑**3.7**, **3.13**, 33

[120] Y. Lafont, *Algebra and geometry of rewriting*, Appl. Categ. Structures **15** (2007), no. 4, 415–437. Author's homepage. ↑**2.1**, 6.22

[121] J. Lambek and P. J. Scott, *Introduction to Higher Order Categorical Logic*, Cambridge Univ. Press, Cambridge, 1986. ↑**1.1**, **2.1**, **4.2**, *4.2.5*

[122] J. Lipton and M. J. O'Donnell, *Some intuitions behind realizability semantics for constructive logic: tableaux and Läuchli countermodels*, Ann. Pure Appl. Logic **81** (1996), 187–239. Journal, O'Donnell's homepage. ↑*6.3.4*

[123] J.-L. Loday, *Homotopical syzygies*, Une dégustation topologique: Homotopy Theory in the Swiss Alps (Arolla, 1999), Contemp. Math., vol. 265, Amer. Math. Soc., Providence, RI, 2000, pp. 99–127. Author's homepage. ↑6

[124] P. Lorenzen, *Einführung in die operative Logik und Mathematik*, The Basic Teachings of the Math. Sci., vol 78, Springer-Verlag, Berlin-New York, 1969. See [177] for a summary in English. ↑*4.5.4*

[125] P. Martin-Löf, *About models for intuitionistic type theories and the notion of definitional equality*, Proc. Third Scandinavian Logic Symposium (Univ. Uppsala, 1973), Stud. Logic Found. Math., vol. 82, North-Holland, Amsterdam, 1975, pp. 81–109. The collected works of Martin-Löf at Github. ↑**2.1**

[126] _____, *Intuitionistic Type Theory*, Studies in Proof Theory. Lecture Notes, vol. 1, Bibliopolis, Naples, 1984. Intuitionistic Type Theory blog by J. G. Granström; The collected works of Martin-Löf at Github. ↑**3.4**, **3.7**

[127] _____, *Truth of a proposition, evidence of a judgement, validity of a proof*, Synthese **73** (1987), 407–420. The collected works of Martin-Löf at Github. ↑**3.2**, **3.7**

[128] _____, *On the meanings of the logical constants and the justifications of the logical laws*, Nordic J. Philos. Logic **1** (1996), no. 1, 11–60. The collected works of Martin-Löf at Github. ↑

[129] _____, *The Hilbert–Brouwer controversy resolved?*, One Hundred Years of Intuitionism (1907–2007), Publ. des Archives H. Poincaré, Birkhäuser, 2008, pp. 243–256. The collected works of Martin-Löf at Github. ↑**3.3**, **3.3**

[130] _____, *100 years of Zermelo's axiom of choice: what was the problem with it?*, Logicism, Intuitionism, and Formalism, Synthese Library, vol. 341, Springer, Dordrecht, 2009, pp. 209–219. The collected works of Martin-Löf at Github. ↑**3.3**

[131] _____, *Verificationism then and now*, Judgement and the Epistemic Foundation of Logic (M. van der Schaar, ed.), Logic, Epistemology, and the Unity of Science, vol. 31, 2012, pp. 3–14. The collected works of Martin-Löf at Github. ↑**3.2**

[132] M. C. McCord, *Singular homology groups and homotopy groups of finite topological spaces*, Duke Math. J. **33** (1966), 465–474. Andrew Ranicki's archive. ↑6.2

[133] J. C. C. McKinsey and A. Tarski, *The algebra of topology*, Ann. of Math. **45** (1944), 141–191. ↑

[134] _____, *On closed elements in closure algebras*, Ann. of Math. **47** (1946), 122–162. ↑

[135] _____, *Some theorems about the sentential calculi of Lewis and Heyting*, J. Symbolic Logic **13** (1948), 1–15. ↑**5.3**, *5.3.1*, 82

[136] Yu. T. Medvedev, *Finite problems*, Dokl. Akad. Nauk SSSR **142** (1962), 1015–1018; English transl., Soviet Math. Dokl. **3** (1962), 227–230. ↑**3.13**

[137] _____, *Interpretation of logical formulas by means of finite problems*, Dokl. Akad. Nauk SSSR **169** (1966), no. 1, 20–23; English transl., Soviet Math. Dokl. **7** (1966), 857–860. ↑**3.13**

[138] S. A. Melikhov, *Metrizable uniform spaces*. `arXiv:1106.3249`. ↑

[139] _____, *Infinite-dimensional uniform polyhedra*. `arXiv:1109.0346`. ↑*5.3.3*

[140] _____, *Combinatorics of combinatorial topology*. `arXiv:1208.6309`. ↑*5.3.3*, *5.3.3*

[141] E. Mendelson, *Introduction to Mathematical Logic*, D. Van Nostrand Co., Princeton, N.J., 1964; Russian transl. in Введение в математическую логику, Наука, М., 1971. ↑*4.1.4*, *4.72*

[142] G. E. Minc, *Derivability of admissible rules*, Zap. Nauchn. Sem. LOMI **32** (1972), 85–89; Mathnet; English transl., J. Soviet Math. **6** (1976), 417–421. ↑

[143] G. Mints, *The Gödel-Tarski translations of intuitionistic propositional formulas*, Correct Reasoning: Essays on Logic-Based AI in Honour of Vladimir Lifschitz, Lecture Notes in Comp. Sci., vol. 7265, Springer, Berlin, 2012, pp. 487–491. ↑*5.7.3*

[144] M. Morillon, *Les compacts extrêmement discontinus sont finis!*, Séminaire d'Analyse, 1993–1994 (Aubière), Sémin. Anal. Univ. Blaise Pascal (Clermont II), vol. 9, Univ. Blaise Pascal (Clermont II), Clermont-Ferrand, 1996, pp. Exp. No. 11, 15 (French). ↑**5.4**

[145] Y. N. Moschovakis, *What is an algorithm?*, Mathematics unlimited—2001 and beyond, Springer, Berlin, 2001, pp. 919–936. Author's homepage. ↑**2.1**

[146] A. Mutlu and T. Porter, *Free crossed resolutions from simplicial resolutions with given CW-basis*, Cahiers Topologie Géom. Différentielle Catég. **40** (1999), no. 4, 261–283. Numdam. ↑6

[147] B. Nordström, K. Petersson, and J. M. Smith, *Programming in Martin-Löf's Type Theory*, Internat. Series of Mono. on Computer Sci., vol. 7, The Clarendon Press, Oxford Univ. Press, New York, 1990. Online version. ↑46

[148] T. Nipkow, L. C. Paulson, and M. Wenzel, *Isabelle/HOL: A Proof Assistant for Higher-Order Logic*, Lecture Notes in Computer Sci., vol. 2283, 2002. Isabelle website (an updated version). ↑*4.1.3*

[149] P. S. Novikov (П. С. Новиков), *Constructive Mathematical Logic from the Point of View of Classical One (Конструктивная математическая логика с точки зрения классической)*, Nauka, Moscow, 1977 (Russian). ↑3.3, *4.1.3*, **5.3**, 5.10, *5.7.3*, 83

[150] P. S. Novikoff, *On the consistency of certain logical calculus*, Rec. Math. [Mat. Sbornik] N.S. **12** (1943), 231–261. Mathnet. ↑79

[151] П. С. Новиков, *Элементы математической логики*, 2nd ed., Наука, Москва, 1959/73; English transl. in *Elements of mathematical logic*, Oliver & Boyd, Edinburgh-London; Addison-Wesley Publishing Co., Inc., Reading, Mass., 1964. ↑3.3, *4.1.3*

[152] I. Németi, *On varieties of cylindric algebras with applications to logic*, Ann. Pure Appl. Logic **36** (1987), 235–277. Journal. ↑*4.1.2*

[153] P. Oliva and T. Streicher, *On Krivine's realizability interpretation of classical second-order arithmetic*, Fund. Inform. **84** (2008), 207–220. Oliva's homepage. ↑**3.3**

[154] I. E. Orlov (И. Е. Орлов), *Calculus of compatibility of propositions (Исчисление совместности предложений)*, Mat. Sbornik **35** (1928), 263–286 (Russian). ↑**3.3**, 82

[155] E. Palmgren, *Constructive logic and type theory*. Author's homepage. ↑

[156] _____, *A categorical version of the Brouwer–Heyting–Kolmogorov interpretation*, Math. Structures Comput. Sci. **14** (2004), 57–72. Instutut Mittag-Leffler, Report No. 16 (2000/01). ↑**1.1**, *6.3.1*

[157] L. C. Paulson, *The foundation of a generic theorem prover*, J. Automat. Reason. **5** (1989), no. 3, 363–397. Univ. of Cambridge Computer Lab. Tech. Report UCAM-CL-TR-130. ↑**1.1**, *4.1.3*, *4.1.3*, *4.1.3*, *4.1.3*, *4.1.4*, **4.4**

[158] _____, *Isabelle: A Generic Theorem Prover (With contributions by T. Nipkow)*, Lecture Notes in Computer Sci., vol. 828, Springer-Verlag, Berlin, 1994. Isabelle website (an updated version). ↑*4.1.3*, 45

[159] V. E. Plisko, *A. N. Kolmogorov's calculus as a fragment of minimal calculus*, Uspekhi Mat. Nauk **43** (1988), no. 6, 79–91. Mathnet; English transl., Russian Math. Surveys **43** (1988), no. 6, 95–110. ↑*4.6.1*, **5.6**

[160] _____, *Absolute realizability of predicate formulas*, Izv. Akad. Nauk SSSR, Ser. Mat. **47** (1983), no. 2, 315–334; English transl., Math. USSR-Izv. **22** (1984), no. 2, 291–308. Mathnet. ↑*7.1.4*, *7.1.4*, *7.1.4*

[161] _____, *A survey of predicate realizability logic*, Tr. MIAN **274** (2011), 222–251; English transl., Proc. Steklov Inst. Math. **274** (2011), 204–230. ↑*7.1.4*, *7.1.4*, *7.1.4*, *7.1.4*, *7.1.4*

[162] W. A. Pogorzelski and T. Prucnal, *Structural completeness of the first-order predicate calculus*, Z. Math. Logik Grundlagen Math. **21** (1975), no. 4, 315–320. ↑*4.1.2*

[163] W. C. Powell, *Extending Gödel's negative interpretation to ZF*, J. Symbolic Logic **40** (1975), 221–229. ↑**3.3**

[164] D. Prawitz, *Natural deduction. A proof-theoretical study*, Acta Universitatis Stockholmiensis. Stockholm Studies in Philosophy, No. 3, Almqvist & Wiksell, Stockholm, 1965. ↑*4.4.1*, 67, **5.8**

[165] _____, *Ideas and results in proof theory*, Proc. of the Second Scandinavian Logic Symposium (Univ. Oslo, Oslo, 1970), Stud. Logic Found. Math., vol. 63, North-Holland, Amsterdam, 1971, pp. 235–307. ↑**2.1**

[166] T. Prucnal, *On two problems of Harvey Friedman*, Studia Logica **38** (1979), 247–262. ↑*5.3.2*

[167] A. Ranta, *Type Theoretical Grammar*, Oxford Univ. Press, 1994. ↑*6.3.4*

[168] H. Rasiowa and R. Sikorski, *The Mathematics of Metamathematics*, Monografie Matematyczne, vol. 41, Państwowe Wydawnictwo Naukowe, Warsaw, 1963; Russian transl. in *Математика метаматематики*, Наука, М., 1972. ↑*4.8.1*, *4.8.1*, 71, *4.8.2*, *4.8.2*, **5.1**, **5.3**, 74, *5.3.1*, *5.7.1*

[169] H. Rogers Jr., *Theory of Recursive Functions and Effective Computability*, McGraw-Hill, New York, 1967; Russian transl. in *Теория рекурсивных функций и эффективная вычислимость*, Мир, М., 1972. ↑

[170] P. Rozière, *Règles admissibles en calcul propositionnel intuitionniste*. Thèse de Doctorat, Université Paris VII, 1992; Author's homepage (in French; partial English transl.) ↑*5.4.1*

[171] W. Ruitenburg, *The unintended interpretations of intuitionistic logic*, Perspectives on the history of mathematical logic, Birkhäuser Boston, Boston, MA, 1991, pp. 134–160. Author's homepage. ↑**3.7**

[172] V. V. Rybakov, *Admissibility of logical inference rules*, Studies in Logic and the Foundations of Mathematics, vol. 136, North-Holland Publishing Co., Amsterdam, 1997. ↑*4.1.2*, *4.5.3*, *4.7.3*, *5.4.1*

[173] W. de C. Sanz and T. Piecha, *A critical remark on the BHK interpretation of implication*. Hypothesis project. ↑**3.7**

[174] K. S. Sarkaria, *Formal theories are acyclic*, Z. Math. Logik Grundlag. Math. (Math. Logic Quarterly) **31** (1985), no. 4, 363–368. Author's homepage. ↑**2.1**

[175] M. Sato, *Classical Brouwer–Heyting–Kolmogorov interpretation*, Algorithmic Learning Theory (Sendai, 1997), Lecture Notes in Comput. Sci., vol. 1316, Springer, Berlin, 1997, pp. 176–196. Kyoto University Research Information Repository. ↑**3.9**

[176] M. Sato, T. Sakurai, Y. Kameyama, and A. Igarashi, *Calculi of meta-variables*, Computer Science Logic, Lecture Notes in Comput. Sci., vol. 2803, Springer, Berlin, 2003, pp. 484–497. Kameyama's homepage. ↑*4.1.2*

[177] P. Schröder-Heister, *Lorenzen's operative justification of intuitionistic logic*, One Hundred Years of Intuitionism (1907–2007), Publ. des Archives H. Poincaré, Birkhäuser, 2008, pp. 214–240. Author's homepage. ↑*4.5.4*, 124

[178] R. A. G. Seely, *Modelling computations: a 2-categorical framework*, Symposium on Logic in Computer Science, IEEE Computer Society Press, 1987. ↑**2.1**

[179] V. B. Shehtman and D. P. Skvortsov, *Logics of some Kripke frames connected with Medvedev notion of informational types*, Studia Logica **45** (1986), no. 1, 101–118. ↑**3.13**

[180] A. Shen and N. Vereshchagin, *Logical operations and Kolmogorov complexity*, Theoret. Comput. Sci. **271** (2002), no. 1-2, 125–129. Shen's homepage. ↑

[181] J. R. Shoenfield, *Mathematical Logic*, Addison-Wesley, Reading, Mass., 1967; English transl. in *Математическая логика*, Наука, М., 1975. ↑*4.1.4*

[182] S. G. Simpson, *Mass problems associated with effectively closed sets*, Tohoku Math. J. **63** (2011), no. 4, 489–517. Author's homepage. ↑

[183] D. P. Skvortsov, *Two generalizations of the notion of a finite problem (Два обобщения понятия финитной задачи)*, Studies in nonclassical logics and set theory (Исследования по неклассическим логикам и теории множеств), Nauka, Moscow, 1979, pp. 201–240 (Russian). ↑**3.13**

[184] S. Slavnov, *On completeness of dynamic topological logic*, Mosc. Math. J. **5** (2005), no. 2, 477–492, 495–496. Journal. ↑**5.3**

[185] C. Smorynski, *The incompleteness theorems*, Handbook of Mathematical Logic (J. Barwise, ed.), Studies in Logic and the Found. of Math., vol. 90, North-Holland, Amsterdam, 1977. ↑20, *5.7.2*, *5.7.2*, *5.7.2*

[186] G. Sommaruga, *History and Philosophy of Constructive Type Theory*, Kluwer, Dordrecht, 2000. ↑

[187] A. Sorbi and S. A. Terwijn, *Intermediate logics and factors of the Medvedev lattice*, Ann. Pure Appl. Logic **155** (2008), no. 2, 69–85. Terwijn's homepage. ↑

[188] M. H. Sørensen and P. Urzyczyn, *Lectures on the Curry–Howard Isomorphism*, Studies in Logic and Found. of Math., vol. 149, Elsevier, Amsterdam, 2006. ↑**2.3**, **3.7**, **3.7**, 33, **4.2**, *4.2.2*, *4.2.5*, **5.3**

[189] M. H. Stone, *Topological representations of distributive lattices and Brouwerian logics*, Časopis pro pěstování matematiky a fysiky **67** (1938), 1–25. DML-CZ. ↑**5.2**

[190] T. Streicher, *Introduction to Constructive Logic and Mathematics*, 2001. Author's Homepage. ↑**3.7**

[191] G. Takeuti, *Proof theory*, North-Holland, Amsterdam, 1975. Stud. in Logic and Found. of Math., Vol. 81. ↑

[192] Tsao-Chen Tang, *Algebraic postulates and a geometric interpretation for the Lewis calculus of strict implication*, Bull. Amer. Math. Soc. **44** (1938), no. 10, 737–744. Journal, ProjectEuclid. ↑**5.2**

[193] A. Tarski, *O pojoiu wynikania logicznego*, Przegląd Filozoficzny **89** (1936), 58–68; English transl., *On the concept of following logically*, Hist. Philos. Logic **23** (2002), no. 3, 155–196. Author's German version, *Über den Begriff der logischen Folgerung,* Actes du Congrès International de Philosophie Scientifique, vol. 7 (Actualités Scientifiques et Industrielles, vol. 394), Paris, 1936, pp. 1-11; English transl.: *On the concept of logical consequence*, in Logic, Semantics, Metamathematics: Papers from 1923 to 1938 by Alfred Tarski (Corcoran, J., ed.), Hackett, Indianapolis, IN, 1956, pp. 421–454. ↑52, *7.2.1*, *7.2.1*

[194] _____, *Der Aussagenkalkül und die Topologie*, Fund. Math. **31** (1938), 103–134; English transl., *Sentential calculus and topology*, Logic, Semantics, Metamathematics: Papers from 1923 to 1938 by Alfred Tarski (Corcoran, J., ed.), Hackett, Indianapolis, IN, 1956, pp. 421–454. ↑**5.2**, **5.3**

[195] R. Thiele, *Hilbert's twenty-fourth problem*, Amer. Math. Monthly **110** (2003), no. 1, 1–24. Publisher. ↑**2.1**, **2.1**

[196] _____, *Hilbert and his twenty-four problems*, Mathematics and the Historian's Craft, CMS Books Math., vol. 21, Springer, New York, 2005, pp. 243–295. ↑**2.1**

[197] A. S. Troelstra, *Aspects of Constructivity*, Handbook of Mathematical Logic. ↑**3.7**

[198] _____, *Introductory note to "An interpretation of the intuitionistic propositional calculus" by K. Gödel*. ↑*5.7.2*, *5.7.3*, *5.7.3*

[199] _____, *Principles of Intuitionism*, Lecture Notes in Math., vol. 95, Springer, Berlin, 1969. ↑**3.7**, *4.6.1*

[200] A. S. Troelstra *et al., Metamathematical Investigation of Intuitionistic Arithmetic and Analysis*, Lecture Notes in Math., vol. 344, Springer, Berlin, 1973. ↑*4.1.4*, *4.6.1*, *4.6.2*

[201] A. S. Troelstra, *From constructivism to computer science*, Theoret. Comput. Sci. **211** (1999), 233–252. Univ. of Amsterdam Preprint CT-1996-02. ↑**3.5**

[202] A. S. Troelstra and D. van Dalen, *Constructivism in Mathematics. An introduction. (Vols. I, II)*, Studies in Logic and the Found. of Math., vol. 121, 123, North-Holland Publishing Co., Amsterdam, 1988. ↑**1.1**, 1, **3.2**, **3.7**, **3.7**, **3.9**, **3.9**, **3.9**, **3.9**, **3.9**, **3.9**, **3.9**, 25, 33, 5.15

[203] T. Umezawa, *On logics intermediate between intuitionistic and classical predicate logic*, J. Symb. Logic **24** (1959), 141–153. ↑*4.5.3*

[204] V. A. Uspenskii and V. E. Plisko, *Intuitionistic logic* (Comments on Kolmogorov's papers), Колмогоров А. Н. Избранные труды. Математика и механика, Наука, М., 1985, pp. 394–404; English transl., Selected Works of A. N. Kolmogorov. Vol. I, Mathematics and its Applications (Soviet Series), vol. 25, Kluwer, Dordrecht, 1991, pp. 452–465. Publisher (Downloadable Back Matter). ↑**5.6**

[205] M. van Atten, *The Development of Intuitionistic Logic*, The Stanford Encyclopedia of Philosophy (Winter 2012 Edition) (E. N. Zalta, ed.) Online Edition. ↑**3.7**

[206] D. van Dalen, *Intuitionistic logic*, Handbook of Philosophical Logic. Vol. III: Alternatives to Classical Logic, Synthese Library, vol. 166, D. Reidel Publishing Co., Dordrecht, 1986. Edited by D. Gabbay and F. Guenthner, pp. 225–339. ↑**3.7**

[207] _____, *Intuitionistic logic*, The Blackwell Guide to Philosophical Logic, Blackwell, Oxford, 2001. Edited by L. Gobble, pp. 224—257. Author's homepage. ↑*4.1.4*

[208] _____, *Kolmogorov and Brouwer on constructive implication, and the contradiction rule*, Uspekhi Mat. Nauk **59** (2004), no. 2, 53–64; Mathnet (Russian); English transl., Russian Math. Surveys **59** (2004), 247–257; Author's Homepage. ↑**3.7**, *3.8.4*

[209] A. V. Chernov, D. P. Skvortsov, E. Z. Skvortsova, and N. K. Vereshchagin, *Variants of Realizability for Propositional Formulas and the Logic of the Weak Law of Excluded Middle*, Tr. Mat. Inst. Steklova **242** (2003), 77–97; Mathnet; English transl., Proc. Steklov Inst. Math. **242** (2003), 67–85. ↑*7.1.4*

[210] J. von Plato, *Skolem's discovery of Gödel-Dummett logic*, Studia Logica **73** (2003), 153–157. ↑68, *5.2.2*

[211] S. Weinstein, *The intended interpretation of intuitionistic logic*, J. Philos. Logic **12** (1983), no. 2, 261–270. ↑**3.7**

[212] R. Wójcicki, *Theory of Logical Calculi. Basic Theory of Consequence Operations*, Synthese Library, vol. 199, Kluwer, Dordrecht, 1988. ↑*4.1.1*, *4.1.1*

[213] S. A. Yanovskaya (С. А. Яновская), *Mathematical logic and foundations of mathematics (Математическая логика и основания математики)*, Mathematics in the USSR for 40 Years (Математика в СССР за 40 лет). 1917–1957. Volume 1, Gos. izd. fiz.-mat. lit., Moscow, 1959, pp. 13–120. ↑1

Steklov Mathematical Institute of the Russian Academy of Sciences, ul. Gubkina 8, Moscow, 119991 Russia

E-mail address: melikhov@mi.ras.ru