Theoretical Computer Science 102 (1992) 1-133
Elsevier

Fundamental Study

Map theory

Klaus Grue

DIKU, The University of Copenhagen, Universitetsparken 1, DK-2100 Copenhagen, Denmark

Communicated by M. Nivat
Received May 1990
Revised November 1990

Abstract

Grue, K., Map theory, Theoretical Computer Science 102 (1992) 1-133.

Map theory is a foundation of mathematics based on λ-calculus instead of logic and sets, and thereby fulfills Church's original aim of introducing λ-calculus. Map theory can do anything set theory can do. In particular, all of classical mathematics is contained in may theory. In addition, and contrary to set theory, map theory has unlimited abstraction and contains a computer programming language as a natural subset. This makes map theory more suited to deal with mechanical procedures than set theory. In addition, the unlimited abstraction allows definition, e.g., of the notion of truth and the category of all categories. This paper introduces map theory, gives a number of applications and gives a relative consistency proof. To demonstrate the expressive power of map theory, the paper develops *ZFC* set theory within map theory.

Contents

Correspondence to: Klaus Grue, DIKU, The University of Copenhagen, Universitetsparken 1, DK-2100 Copenhagen, Denmark. Email: grue@diku.dk.

Part I. Introduction to map theory

1. Overview

1.1. Properties of map theory

This paper introduces a theory — Map Theory — which has five important properties:

(1) map theory is a rigorously-defined formal theory;
(2) map theory has enough expressive power to serve as a foundation of all of classical mathematics;
(3) map theory has unlimited abstraction and a computer programming language as a natural subset;
(4) map theory is as simple as set theory;
(5) relative consistency proofs for map theory exist.

It is hard to find other theories with all of these properties. Set theory [17, 22] fails to satisfy (3). This makes map theory more suited as a foundation of computer science than set theory. Theories like intuitionistic type theory [20], λ-calculus [3], and Meta-IV [8] fail to satisfy (2). Category theory [18, 4] is the common name of a host of theories, each of which fails to satisfy at least one of the points above. All versions of category theory in practical use fail to satisfy (1). Attempts have been made to make set theory satisfy (3) [1, 6], but set theory is highly nonconstructive and inherently unsuited to support computer science.

Set theory is suited as a foundation of all of classical mathematics, and map theory is intended to enlarge the scope to include algorithms, metalogic and computer science. It is easy to move from set to map theory since all concepts of set theory are definable in map theory and all theorems of set theory (*ZFC* to be precise) are also theorems of map theory. *ZFC* as stated in [17] has four elementary concepts: membership ($\in$), negation ($\neg$), implication ($\Rightarrow$) and universal quantification ($\forall$). To demonstrate the expressive power of map theory, this paper defines all four in map theory and proves that any theorem of *ZFC* is provable in map theory.

1.2. Comparison with earlier work

Map theory is considerably different from set theory, but compares well to set theory with respect to formal rigor, expressive power and relative consistency. Hence, map theory has finally turned λ-calculus into the alternative foundation it was originally intended to be.

References [23, 24] reviews earlier attempts to turn λ-calculus into a foundation and make their own contribution. Reference [24] concludes, however, that the attempts did not succeed well.

Like earlier attempts, map theory is based on *functions*. Functions are termed *maps* in map theory. Reference [23] identifies certain functions to be particularly well-behaved and calls them *definite*. Likewise, map theory identifies certain maps to be *well-founded*. However, the definition of well-foundedness differs from any of the earlier attempts and is crucial to the expressive power of map theory.

Earlier attempts let a function f represent the set

$$\{x \mid f(x) = \mathsf{true}\},$$

whereas map theory let a map f represent the set

$$\{f(x) \mid x \in \Phi\}$$

where Φ is the collection of well-founded maps. The difference is like the difference between co- and contravariance.

As a consequence of the representation of sets, earlier attempts have used $f(x)$ to stand for $x \in f$ and $\lambda x.\mathscr{A}$ to stand for $\{x \mid \mathscr{A}\}$. Matters are slightly more complicated in map theory due to the different representation. In particular, λ-abstraction and set abstraction are two different things in map theory. The paradoxes of set theory have shown that unrestricted set abstraction is unattainable, so in hindsight it is no surprise that unrestricted λ-abstraction differs from set abstraction.

Map theory also differs from earlier attempts in that it is based on a type-free but not completely pure λ-calculus: In addition to functions and the inevitable $\perp$ element, the maps of map theory also comprise a single object which is not a function. This is no dramatic innovation but crucial to the expressive power of map theory and deserves mention here.

1.3. Three descriptions

This paper gives three descriptions of map theory: an intuitive, an axiomatic, and a model theoretic one. Part I, II and III take these three points of view, respectively.

Part I uses words like *function*, *set* and *tree* in the naive sense. Sets of *ZFC* and classes of *NBG* [22] are referred to as *ZFC*-sets and *NBG*-classes, respectively. Part I gives the intuition behind map theory and sketches a number of applications of map theory.

Part II presents an axiomatization of map theory and develops *ZFC* within that axiom system. This verifies the expressive power of map theory.

Part III proves the consistency of the axiomatization from Part II. To do so, Part III has to assume the existence of a strongly inaccessible ordinal [22, 17]. Part III also considers the consistency of various weaker and stronger axiomatizations of map theory.

Part II and III together demonstrate the consistency and expressive power of map theory and thereby verify that map theory deserves to be called a foundation.

2. The intuition behind maps

2.1. *Black boxes*

Think of a map as a black box with an input, an output and some hidden mechanical computing machinery inside. Whenever a black box f receives an input x, its machinery starts working and delivers an output after a while. Let $(f x)$ denote this output. The term $(f x)$ reads "f applied to x".

A black box f merely accepts black boxes as input and merely produces black boxes as output. The simplest black box is the identity I which, whenever it receives a black box x as input, delivers this x unchanged as output. This I satisfies $(I x) = x$ for all black boxes x. Other notations read $I : x \mapsto x$ and $I = \lambda x.x$.

Another example is the black box $I' = \lambda y.\lambda x.x$. Whenever it inputs a y, it outputs $\lambda x.x$ regardless of the value of y. In the other notations we have $(I' y) = I$ and $I' : x \mapsto I$. Since $(I' y) = I$ and $(I x) = x$, we have $((I' y) x) = x$. Let $(f x_1 \ldots x_n)$ be shorthand for $(\ldots((f x_1) x_2) \ldots x_n)$. With this convention, $(I' y x) = x$.

Two further examples are $K = \lambda x.\lambda y.x$ and $S = \lambda x.\lambda y.\lambda z(x z (y z))$ which satisfy $(K x y) = x$ and $(S x y z) = (x z (y z))$.

In essence, the machinery inside a map can do anything a real world computer can do, and in addition it can run infinitely many processes in parallel. Even though maps have infinite computational power, there are, as we shall see, operations they cannot perform.

A particularly interesting map is $R = \lambda x.(x x)$ which is closely related to Russell's set $\{x \mid x \notin x\}$. When R receives a black box A as input, it takes two copies of A and enters one A as input to the other A. When the output of the latter A appears, R takes it and outputs it as its own output. Hence, computation of $(R A)$ causes computation of $(A A)$. If a copy of R is entered as input to R, then computation of $(R R)$ causes computation of $(R R)$, which in turn causes computation of $(R R)$ and so on indefinitely. Hence, when R receives a copy of itself, it will work indefinitely without producing output.

Let $\perp$ (bottom) denote "no output". With this convention, $(R R) = \perp$. As another example, let $R' = \lambda x.(x x x)$. Since R' produces no output when it receives R as input, $(R' R) = \perp = (R R)$.

Even though $\perp$ denotes "no map", it is considered to be a map. At this moment a map is either $\perp$ or a black box, a black box accepts both $\perp$ and black boxes as input, and it may produce both $\perp$ and black boxes as output.

Since $\perp$ has no input and output, $(\perp x)$ does not make sense. Nevertheless it is convenient to define that $(\perp x) = \perp$ for all x. This convention ensures that $(f x)$ is defined and is a map for all maps f and x.

2.2. *White boxes*

All black boxes look the same from the outside, but their inner machinery may react differently to input.

The limited output facilities of black boxes make them useless, so it is necessary to enhance them slightly. As an example, consider a black box f which, given an input x, decides whether or not x has a certain property $p(x)$. How should f communicate its findings to the outside world?

One possibility is to let f output one black box if $p(x)$ is true and another if $p(x)$ is false, but that would not be very helpful since all black boxes look the same.

Another possibility is to let f output a black box if $p(x)$ is true and loop indefinitely if $p(x)$ is false. Then if (fx) produced output, $p(x)$ would be known to be true. However, if (fx) had not produced output after a while, then this could either indicate that $p(x)$ is false or that $p(x)$ is true but f needs more time to find out.

To remedy for this, we now introduce the third and last kind of map: a *white box* with no input, no output and no machinery inside. The central property of a white box is that it is immediately distinguishable from any black box. Any two white boxes are equal.

Now a map is either $\perp$ or a white box or a black box. Black boxes accept $\perp$, white and black boxes as input and produce $\perp$, white and black boxes as output.

Most theories are based on a logic, but map theory is not. In map theory, by convention, the white box is used to represent truth and the black boxes are used to represent falsehood. The map $\perp$ represents undefinedness.

Let T denote the white box. The similarity of the symbols T and $\perp$ is incidental: T denotes *truth* whereas $\perp$ denotes the *bottom* element of an ordering to be defined later.

Since T has no input and output, $(\mathsf{T}\,x)$ does not make sense. Nevertheless, it is convenient to define that $(\mathsf{T}\,x) = \mathsf{T}$ for all x. This convention ensures that (fx) is defined and is a map for all maps f and x.

Let $\tilde{\mathsf{T}}$, $\tilde{\lambda}$, and $\tilde{\perp}$ denote "white", "black" and "no color", respectively, and let $r(f)$ denote the color of f. With these conventions,

$$r(\mathsf{T}) = \tilde{\mathsf{T}},$$

$$r(\perp) = \tilde{\perp},$$

$$r(x) = \tilde{\lambda} \quad \text{for black boxes } x.$$

From now on, we refer to black boxes as *proper maps*, to $\tilde{\mathsf{T}}$, $\tilde{\lambda}$ and $\tilde{\perp}$ as *labels*, and to $r(f)$ as the *root* of f.

Three-valued logic [16] has three truth values *true*, *false* and *undefined*. Map theory represents them by $\tilde{\mathsf{T}}$, $\tilde{\lambda}$ and $\tilde{\perp}$, respectively. More precisely, T represents *true*, $\perp$ represents *undefined*, any any proper map represents *false*.

Let M denote the (naive) set of all maps. The set M is neither a *ZFC*-set nor an *NBG*-class.

2.3. Equality of maps

Two maps f and g are equal iff

$$r(f\,x_1 \ldots x_n) = r(g\,x_1 \ldots x_n)$$

for all $n \geq 0$ and all maps $x_1, \ldots, x_n$. This defines equality of maps from equality of labels. The definition is close in spirit to the definition of equality of sets: two sets A and B are equal iff $x \in A \Leftrightarrow x \in B$ for all sets x. This defines equality $=$ of sets from equality $\Leftrightarrow$ of truth values.

Theorem 2.3.1. *Two proper maps f and g are equal iff $(f\,x) = (g\,x)$ for all maps x.*

Proof. Assume f and g are proper, i.e., $r(f) = r(g) = \tilde{\lambda}$. If $f = g$ and if x is a map then

$$\begin{aligned} r((f\,x)\,x_1 \ldots x_n) &= r(f\,x\,x_1 \ldots x_n) \\ &= r(g\,x\,x_1 \ldots x_n) \\ &= r((g\,x)\,x_1 \ldots x_n) \end{aligned}$$

for all $n \geq 0$ and all maps $x_1, \ldots, x_n$, so $(f\,x) = (g\,x)$. On the contrary, if $(f\,x) = (g\,x)$ for all maps x, then $r(f\,x\,x_1 \ldots x_n) = r(g\,x\,x_1 \ldots x_n)$ for all $n \geq 0$ and all maps $x, x_1, \ldots, x_n$. Combined with $r(f) = r(g)$ this gives $r(f\,x_1 \ldots x_n) = r(g\,x_1 \ldots x_n)$ for all $n \geq 0$ and all maps $x_1, \ldots, x_n$, so $f = g$. $\square$

2.4. Well-founded maps

Map theory uses T and proper maps to represent truth and falsehood, respectively. Map theory also uses T to represent the empty set $\emptyset$ and certain proper maps to represent the nonempty *ZFC*-sets. Those maps that represent *ZFC*-sets are going to be termed *well-founded*; the others will be termed *ill-founded*.

A map f is said to be well-founded w.r.t. a set G of maps if

$$\forall x_1, x_2, \ldots \in G\ \exists n\colon\ (f\,x_1 \ldots x_n) = \mathsf{T}.$$

Let G° denote the set of maps that are well-founded w.r.t. G. In particular, let $\emptyset^\circ$ be the set of all maps except $\perp$.

The set Φ of well-founded maps is the least set such that

- $\mathsf{T} \in \Phi$, and
- if $G \subseteq \Phi$ is a set of "limited size" and $\forall x \in G^\circ\colon (f\,x) \in \Phi$, then $f \in \Phi$, where "limited size" will be defined shortly.

The definition builds up Φ in *stages*. The first state Φ_0 merely contains T, so $\Phi_0 = \{\mathsf{T}\}$. Stage α contains all maps whose well-foundedness can be verified, knowing that all maps on stages before stage α are well-founded. For example, $\lambda x.\mathsf{T} \in \Phi_1$ because $\forall x \in \emptyset^\circ\colon ((\lambda x.\mathsf{T})\ x) = \mathsf{T} \in \Phi_0$ and because $\emptyset$ happens to be a set of limited size.

A well-founded map f is said to be *introduced* at stage Φ_α if it belongs to that stage but does not belong to any previous stage. A well-founded map g is *introduced before* a well-founded map f if the stage where g is introduced comes before the one where f is introduced.

The *stock* of a well-founded f is the set of all those well-founded maps g that have to be introduced before f can be introduced. However, it is difficult to formalize this definition, so we shall use another: The stock of a well-founded f is the set of all well-founded maps introduced before f. Let f^s denote the stock of f.

A set of well-founded maps is of *limited size* it it is (or is a subset of) the stock of some well-founded map. Hence, Φ is the least set such that

- $\mathsf{T} \in \Phi$, and
- $g \in \Phi \wedge \forall x \in g^{so}: (f\,x) \in \Phi \Rightarrow f \in \Phi$.

The process of introducing well-founded maps comes to a halt by itself. The size limitation in the definition of Φ prohibits new maps to be introduced at stage σ or later where σ is the first strongly inaccessible ordinal (if such a one exists). As a consequence, the "standard" model for map theory has no inaccessible ordinals (but map theory has "nonstandard" models in which strongly inaccessible ordinals exist).

The universe of *ZFC* is constructed in a similar manner. However, "limited size" is not mentioned in the construction of the *ZFC* universe even though that concept is central in *ZFC*. As a consequence, the construction is imprecise in that it does not specify when the process of introducing sets stops.

One stage approach to *ZFC* is stated in [25]. In that approach, a set S may only contain sets introduced before S. Another approach is the transfinite iteration of the power and union operations described in [22]. In both cases, the constructions provide support for the axiom of foundation more than they explain how the paradoxes are avoided.

2.5. Operations on maps

Maps can perform a few basic operations which they can combine in numerous ways. Among other, they can refer to their input and to the maps T and $\perp$, they can apply one map to another, and they can make abstractions.

Sections 2.1 and 2.2 gave examples of these operations, but the examples are repeated here for emphasis. The map $\lambda x.x$ outputs its input x unchanged, so $\lambda x.x$ refers to its input. The map $\lambda x.\mathsf{T}$ outputs T irrespective of its input, so $\lambda x.\mathsf{T}$ refers to T. Likewise, $\lambda x.\perp$ refers to $\perp$. The map $\lambda x.(x\,\mathsf{T})$ refers to its input as well as T, and applies the former to the latter. The map $\lambda x.\lambda y.x$ inputs x and outputs a black box $\lambda y.x$, which in turn outputs x for any input y. Hence, $\lambda x.\lambda y.x$ builds an abstraction.

Consider $R = \lambda z.(z\,z)$, $R' = \lambda y.(R\,R)$, and $R'' = \lambda x.\lambda y.(R\,R)$. When R'' receives an input x, it outputs R'. When R' receives an input y, it loops indefinitely in the attempt to compute $(R\,R)$. Hence, computation of $(R''\,x)$ does not in itself cause $(R\,R)$ to be computed even though it is part of R''. Rather, computation of $(R\,R)$ is *delayed* until R' receives input. Black boxes are *lazy* [14] in that they delay all computations they are not forced to perform. Rather, they stop computing immediately when they have determined the color (black or white) of their output, and leaves it to their output to continue the computation if necessary.

Black boxes can perform three more operations: they can select, choose and classify.

Selection (if $x\,y\,z$) is an operation on three maps x, y and z. If x is a white box, then y is selected, and if x is black, then z is selected. If x is $\perp$, then one must wait forever for x, and it is impossible to choose between y and z. Hence, $(\mathsf{if}\ \perp y\,z) = \perp$. In short,

$$(\mathsf{if}\ x\,y\,z) = \begin{cases} y & \text{if } r(x) = \tilde{\mathsf{T}}, \\ z & \text{if } r(x) = \tilde{\lambda}, \\ \perp & \text{if } r(x) = \tilde{\perp}. \end{cases}$$

This construct is the McCarthy conditional ([21, p. 54]), and is well-known in computer science.

The operations stated so far are all machine executable, and they form a Turing-complete [22] computer programming language. The last two operations—choice and classification—are not executable on any real machine, and they are the ones that make map theory a powerful theory rather than another programming language.

Choice εf is an operation on a single map f. When a black box performs a choice εf, it first computes $(f\,x)$ for all well-founded x, and waits for all these infinitely many computations to terminate. If $(f\,x)$ loops indefinitely for any well-founded x, then εf never terminates, i.e.,

$$\exists x \in \Phi\colon\ (f\,x) = \perp\ \Rightarrow\ \varepsilon f = \perp. \tag{1}$$

If $(f\,x)$ terminates for all well-founded x, then εf chooses a well-founded x such that $(f\,x) = \mathsf{T}$ if such an x exists. Otherwise, εf chooses a well-founded x such that $(f\,x) \neq \mathsf{T}$ in lack of better. In other words,

$$\forall x \in \Phi\colon\ (f\,x) \neq \perp\ \Rightarrow\ \varepsilon f \in \Phi, \tag{2}$$

$$\forall x \in \Phi\colon\ (f\,x) \neq \perp \wedge \exists x \in \Phi\colon\ (f\,x) = \mathsf{T}\ \Rightarrow\ (f\,\varepsilon f) = \mathsf{T}. \tag{3}$$

The choice operator ε chooses in a deterministic rather than random way as expressed by Ackermann's axiom ([10, p. 244]):

$$\forall x \in \Phi\colon\ r(f\,x) = r(g\,x)\ \Rightarrow\ \varepsilon f = \varepsilon g. \tag{4}$$

The construct εf corresponds to Hilbert's epsilon operator [15].

A map p represents a predicate in the following sense: The predicate is true for a map x if $(p\,x) = \mathsf{T}$. It is false if $(p\,x)$ is proper, and undefined if $(p\,x) = \perp$. The construct εp attempts to choose a well-founded x that satisfies the predicate p, but may fail if no such x exists or if $(p\,x) = \perp$ for some well-founded x. Ackermann's axiom states that if p and q express the same three-valued predicate, then εp and εq pick the same x.

Classification ϕx is the last operation. It classifies x as well- or ill-founded as follows:

$$\phi x = \begin{cases} \mathsf{T} & \text{if } x \in \Phi, \\ \perp & \text{otherwise.} \end{cases}$$

The predicate ϕ corresponds to the "menge" predicate $M(x)$ in *NBG* [22].

The terms of map theory have the following syntax. Any term denotes a map.

$$\mathsf{variable} ::= x \mid y \mid z \mid \ldots,$$

$$\mathsf{term} ::= \mathsf{variable} \mid \lambda\mathsf{variable.term} \mid (\mathsf{term\ term})$$
$$\mid \mathsf{T} \mid \bot \mid (\mathsf{if\ term\ term\ term}) \mid \phi\mathsf{term} \mid \varepsilon\mathsf{term}.$$

As usual, the λx in $\lambda x.\mathscr{A}$ is said to bind all unbound occurrences of x in $\mathscr{A}$. A term is *closed* (or is a *combinator* [3]) if it has no free variables. To be very precise, a term $\mathscr{B}$ merely denotes a *specific* map if $\mathscr{B}$ has no free variables.

There are only countably many terms and, intuitively, there are more maps than there are *ZFC*-sets. Hence, the terms merely denote a small fraction of the maps. Further, different terms may denote the same map. For example, $\lambda x.x = \lambda y.y$ according to Lemma 2.3.1.

2.6. *Well-foundedness theorems*

Some terms of map theory are said to be *simple* and some are said to be *dual.* Let Σ and $\bar{\Sigma}$ denote the syntax classes of simple and dual terms, respectively. Let $x_0, x_1, \ldots$ and $y_0, y_1, \ldots$ denote distinct variables, and let $\Sigma^{\#}$ denote the syntax class of simple terms in which $x_0, x_1, \ldots$ do not occur free. A definition of Σ and $\bar{\Sigma}$ is given by

$$\Sigma ::= y_i \mid \lambda x_i.\Sigma \mid (\Sigma\,\bar{\Sigma}) \mid \mathsf{T} \mid \varepsilon y_i.\Sigma^{\#} \mid \phi\Sigma \mid (\mathsf{if}\,\bar{\Sigma}\,\Sigma\,\Sigma) \mid ((\lambda y_i.\Sigma)\,\Sigma),$$

$$\bar{\Sigma} ::= x_i \mid (\bar{\Sigma}\,\Sigma) \mid \Sigma,$$

where $\varepsilon y_i.\mathscr{A}$ is shorthand for $\varepsilon(\lambda y_i.\mathscr{A})$ for all terms $\mathscr{A}$.

As shown in Section 7, some important consequences of the definition of Φ are given in the following theorems.

Theorem 2.6.1 (Totality). *If $\mathscr{A}$ is a simple term whose free variables occur among $y_0, \ldots, y_n$, then $\mathscr{A}$ denotes a well-founded map for all well-founded $y_0, \ldots, y_n$.*

Theorem 2.6.2 (Well-foundedness). *If f and $a_1, a_2, \ldots$ are well-founded, then there is an n such that $(f\,a_1 \ldots a_n) = \mathsf{T}$.*

Theorem 2.6.3 (Induction). *If*

$$\mathscr{P}(\mathsf{T}) \quad \textit{and} \quad \forall x \in \Phi\colon\ (\forall y \in \Phi\colon\ \mathscr{P}(x\,y) \Rightarrow \mathscr{P}(x))$$

then

$$\forall x \in \Phi\colon\ \mathscr{P}(x).$$

Theorem 2.6.4 (Primitive recursion). *If a, $b \in \Phi$, if $\forall x \in \Phi$: $(g\,x) \in \Phi$, and if*

$$(f\,x) = (\mathsf{if}\ x\ a\ (g\,\lambda u.(f\,(x\,(b\,u)))))$$

then $f \in \Phi$.

These theorems allow to decide the well-foundedness of a wide class of terms and to prove properties about well-foundedness. As an example of use of Theorem 2.6.1, $\lambda x.(\mathsf{if}\ x\ a\ b)$ is well-founded if a and b are well-founded. As an example of use of Theorem 2.6.2, $\bot$ is ill-founded (i.e. not well-founded) since $(\bot\ a_1 \ldots a_n) = \bot \neq \mathsf{T}$ for all n and all well-founded $a_1, \ldots, a_n$.

As a more subtle application of Theorem 2.6.2, $I = \lambda x.x$ is not well-founded, for if I was well-founded, then $(I\,I \ldots I) = \mathsf{T}$ should hold for sufficiently many I's in succession, but $(I\,I \ldots I) = I \neq \mathsf{T}$.

2.7. Maps as trees

Black and white boxes provide one mental picture of the maps. This "box picture" considers maps as mechanical procedures (or algorithms).

The present section presents another mental picture of the maps in which maps are thought of as trees. This "tree picture" considers maps as data structures, and the two views complement each other.

The tree representation of a map f is a tree whose nodes are labelled by the labels $\tilde{\mathsf{T}}$, $\tilde{\lambda}$ and $\tilde{\bot}$, and whose edges are labelled by maps. The tree picture f' of the map f is constructed as follows.

- If $f = \mathsf{T}$ then the root node of f' is labelled $\tilde{\mathsf{T}}$. If $f = \bot$ then the root node is labelled $\tilde{\bot}$, and if f is proper then the root node is labelled $\tilde{\lambda}$.
- If $f = \mathsf{T}$ or $f = \bot$, then f' has no other nodes than the root node, and f' has no edges. Figure 1 shows the tree pictures of T and $\bot$.
- If f is proper, then the tree picture f' of f is constructed recursively: For each map x, the tree picture $(f\,x)'$ of the map $(f\,x)$ is constructed. Then $(f\,x)'$ is attached to an edge labelled x descending from the root node of f' as shown in Fig. 2. Figure 3 shows $\lambda x.\mathsf{T}$, Fig. 4 shows $\lambda x.(\mathsf{if}\ x\ (\lambda y.\mathsf{T})\ \mathsf{T})$, and Fig. 5 shows $\lambda x.x$. The

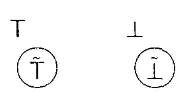

Fig. 1. The maps T and $\bot$.

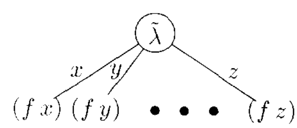

Fig. 2. The graphical representation of a proper map f.

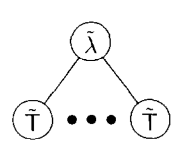

Fig. 3. The map $\lambda x.\mathsf{T}$.

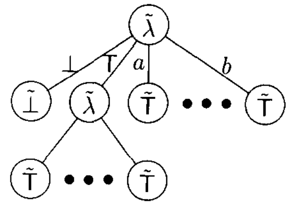

Fig. 4. The map $\lambda x.(\text{if } x\ (\lambda y.\mathrm{T})\ \mathrm{T})$.

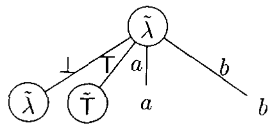

Fig. 5. The map $\lambda x.x$

figures have to be incomplete since the tree picture of any proper map is infinitely large.

The root $r(f)$ of the map f was introduced in Section 2.2. The root $r(f)$ is the label of the root node of f.

As mentioned in Section 2.3, two maps f and g are equal if

$$(f x_1 \ldots x_n) = (g\, x_1 \ldots x_n)$$

for all $n \geqslant 0$ and all maps $x_1, \ldots, x_n$. Hence, two maps are equal if their tree pictures are equal.

When computing $(f x)$, it is often convenient to use the box picture for f and the tree picture for x, i.e., f is considered as an algorithm and x as data on which f acts. While f is computing $(f x)$, it has access to the labels of the tree picture of x through the operations of selection, choice and classification. The map f cannot access any further information about x since it has no operations for doing so (cf. Section 2.5). This justifies the important property

$$x = y \Rightarrow (f x) = (f y).$$

2.8. *Order and monotonicity*

The expression $r(f x_1 \ldots x_n)$ denotes the label of the node reached when traveling from the root node of f downwards along the path $x_1, \ldots, x_n$. To find the tree picture of a map f, it is necessary to compute $r(f x_1 \ldots x_n)$ for various $n \geqslant 0$ and maps $x_1, \ldots, x_n$.

Let $n \geqslant 0$, let f, $x_1, \ldots, x_n$ be maps, and let $u = r(f x_1 \ldots x_n)$. Computation of u either yields $u = \tilde{\mathrm{T}}$ or $u = \tilde{\lambda}$ within "finite" time or loops indefinitely, in which case $r(f x_1 \ldots x_n) = \tilde{\perp}$. If computation of u has given no result after a while, this either means that more time is needed or that $u = \tilde{\perp}$. Hence, absence of a result after a while provides no information. As a consequence, $u = \tilde{\perp}$ means that computation of u gives no information ever.

For all labels $u, v \in \{\tilde{\mathsf{T}}, \tilde{\lambda}, \tilde{\bot}\}$ define

$$u \leqslant_L v \Leftrightarrow u = \tilde{\bot} \vee u = v.$$

The relation $u \leqslant_L v$ states that any information present in u is also present in v, but v may contain more information than u. We have $\tilde{\bot} \leqslant_L \tilde{\mathsf{T}}$ and $\tilde{\bot} \leqslant_L \tilde{\lambda}$ because $\tilde{\bot}$ contains no information, so any information present in $\tilde{\bot}$ is also present in $\tilde{\mathsf{T}}$ and $\tilde{\lambda}$. We neither have $\tilde{\mathsf{T}} \leqslant_L \tilde{\lambda}$ nor $\tilde{\lambda} \leqslant_L \tilde{\mathsf{T}}$ because $\tilde{\mathsf{T}}$ and $\tilde{\lambda}$ contain different information. Figure 6 illustrates the partial order $\leqslant_L$. The L in $\leqslant_L$ refers to *label*.

For all maps f and g define that $f \leqslant g$ iff

$$r(f x_1 \ldots x_n) \leqslant_L r(g x_1 \ldots x_n)$$

for all $n \geqslant 0$ and all maps $x_1, \ldots, x_n$. Like $u \leqslant_L v$, $f \leqslant g$ states that any information present in f is also present in g. However, $\leqslant$ applies to maps whereas $\leqslant_L$ applies to labels. As an example, $\bot$ contains "less information" than any other map. One may think of $\bot$ as a map containing "no information".

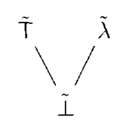

Fig. 6. The partial order $\leqslant_L$ on labels.

Maps have infinite computational power because their choice operation allow them to do infinitely many computations in parallel. However, maps still resemble physical computers in that there are operations they cannot perform. As an example of such an operation, consider

$$\mathscr{G}(x) = \begin{cases} \mathsf{T} & \text{if } x = \bot \\ \mathsf{F} & \text{otherwise.} \end{cases}$$

Suppose we want the map g to satisfy $(g\,x) = \mathscr{G}(x)$. Now $x = \bot$ iff $r(x) = \tilde{\bot}$, so to compute $(g\,x)$, g needs to compute $r(x)$. If $r(x)$ is computable in finite time then $x \neq \bot$ so g can output F. Hence, g can satisfy $(g\,x) = \mathscr{G}(x)$ for $x \neq \bot$. However, if computation of $r(x)$ loops indefinitely, then g sits and waits forever for the value of $r(x)$, so g produces no output. Hence, $(g\,x) = \bot$ so g cannot satisfy $(g\,x) = \mathscr{G}(x)$ for all x.

The operation $\mathscr{G}$ is not computable by any map; the operation

$$\mathscr{G}'(x) = \begin{cases} \bot & \text{if } x = \bot, \\ \mathsf{F} & \text{otherwise,} \end{cases}$$

on the other hand, is. The difference is that $\mathscr{G}'$ is monotonic in $\leqslant$, i.e., $x \leqslant y \Rightarrow \mathscr{G}'(x) \leqslant \mathscr{G}'(y)$. All maps f are monotonic, i.e.

$$x \leqslant y \;\Rightarrow\; (f\,x) \leqslant (f\,y).$$

This is a property that maps inherit from physical computers in general and lazy functional programs [14] in particular.

As shown later, the monotonicity property and the existence of $\perp$ give simple explanations of Russell's and Cantor's paradoxes and allow the notion of truth to be definable in map theory.

Some immediate consequences of the definition of $\leqslant$ are given in the following lemma.

Lemma 2.8.1. *For all maps x, y and z,*

$$\perp \leqslant x,$$

$$\mathsf{T} \leqslant x \;\Rightarrow\; x = \mathsf{T},$$

$$x \leqslant y \wedge x \textit{ is proper} \;\Rightarrow\; y \textit{ is proper},$$

$$x \leqslant y \wedge y \leqslant x \;\Rightarrow\; x = y,$$

$$x \leqslant y \wedge y \leqslant z \;\Rightarrow\; x \leqslant z,$$

$$x \leqslant y \;\Rightarrow\; (x\,z) \leqslant (y\,z).$$

The map T is maximal w.r.t. $\leqslant$, but there are other maximal maps, so T is no top element. As mentioned earlier, T stands for *truth* rather than *top.*

2.9. Problems in defining Φ

The definition of Φ is crucial to the expressive power of map theory. This section considers a few alternative definitions to the authorized one given in Section 2.4 and shows how these alternatives lead to considerably weaker theories. As shown below, if the set Φ is chosen too small or too large, then the expressive power of map theory becomes less than that of *ZFC.*

As an example of a too small Φ, consider the definition

$$\Phi = \{\mathsf{T}\}.$$

With this definition,

$$\phi x = (\mathsf{if}\ x\ \mathsf{T}\ \perp), \qquad \varepsilon x = (\mathsf{if}\ (x\ \mathsf{T})\ \mathsf{T}\ \mathsf{T}),$$

hence, if $\Phi = \{\mathsf{T}\}$, then ϕ and ε are expressible in terms of the other operations, and map theory reduces to a computer programming language and, in particular, becomes weaker than *ZFC.* In general, if Φ has fewer elements than the universe of *ZFC,* then map theory is bound to be weaker than *ZFC.*

As an example of a too large Φ, consider the definition

$$\Phi = M\backslash\{\perp\}.$$

With this definition,

$$\phi x = (\mathsf{if}\ x\ \mathsf{T}\ \mathsf{T}).$$

Now define $\hat{\Phi} = \{\mathsf{T}, \lambda x.\perp\}$. We have

$$x \in \Phi \;\Leftrightarrow\; \exists y \in \hat{\Phi}\colon y \leqslant x.$$

Hence, using the monotonicity of maps, one possible definition of ε reads

$$\varepsilon x = (\mathsf{if}\ (x\ \mathsf{T})\ (\mathsf{if}\ (x\ \lambda x.\bot)\ \mathsf{T}\ \mathsf{T})\ (\mathsf{if}\ (x\ \lambda x.\bot)\ (\lambda x.\bot)\ \mathsf{T})).$$

Again, ϕ and ε are expressible in terms of the other operations. In general, if a set $\hat{\Phi}$ of maps has fewer elements than the universe of ZFC, and if $x \in \Phi \Leftrightarrow \exists y \in \hat{\Phi}: y \leq x$, then map theory is bound to be weaker than set theory.

The monotonicity of maps dictates $x \leq y \Rightarrow ((\lambda z.\phi z)\ x) \leq ((\lambda z.\phi z)\ y) \Rightarrow \phi x \leq \phi y$ so $x \in \Phi \wedge x \leq y \Rightarrow y \in \Phi$, which restricts the possible choices of Φ. As an example, if $\lambda x.\bot \in \Phi$, then all proper maps belong to Φ, and map theory becomes trivial.

On this background it would be obvious to define Φ to be the set of maximal elements of M. This possible definition, however, requires further work to investigate.

3. Uses of map theory

3.1. Logical connectives

As mentioned in Section 2.2, T represents truth, $\bot$ represents undefinedness, and any proper map represents falsehood. Define

$$\mathsf{F} = \lambda x.\mathsf{T}.$$

The map F is one of those that represent falsehood.

Define

$$\begin{aligned}
\dot{\neg} x &= (\mathsf{if}\ x\ \mathsf{F}\ \mathsf{T}),\\
\approx x &= (\mathsf{if}\ x\ \mathsf{T}\ \mathsf{F}),\\
!x &= (\mathsf{if}\ x\ \mathsf{T}\ \mathsf{T}),\\
¡x &= (\mathsf{if}\ x\ \mathsf{F}\ \mathsf{F}),\\
x \mathbin{\dot{\wedge}} y &= (\mathsf{if}\ x \approx y\ ¡y),\\
x \mathbin{\dot{\vee}} y &= (\mathsf{if}\ x\ !y \approx y),\\
x \mathbin{\dot{\Rightarrow}} y &= (\mathsf{if}\ x \approx y\ !y),\\
x \mathbin{\dot{\Leftrightarrow}} y &= (\mathsf{if}\ x \approx y\ \dot{\neg} y).
\end{aligned}$$

The dots in $\dot{\neg}$, $\dot{\wedge}$, $\dot{\vee}$, $\dot{\Rightarrow}$ and $\dot{\Leftrightarrow}$ are introduced to distinguish these terms of map theory from the logical connectives $\neg$, $\wedge$, $\vee$, $\Rightarrow$ and $\Leftrightarrow$ they emulate. Part III uses dots in a slightly different way.

Figure 7 shows the truth tables of $x \mathbin{\dot{\wedge}} y$ and $y \mathbin{\dot{\wedge}} x$. As an example of use, the fourth line of the table states that if $r(x) = \tilde{\lambda}$ and $r(y) = \tilde{\mathsf{T}}$, then $x \mathbin{\dot{\wedge}} y = \mathsf{F}$ and $y \mathbin{\dot{\wedge}} x = \mathsf{F}$. In other words, if x is false and y is true, then both $x \mathbin{\dot{\wedge}} y$ and $y \mathbin{\dot{\wedge}} x$ are false.

Since $x \mathbin{\dot{\wedge}} y$ and $y \mathbin{\dot{\wedge}} x$ have identical truth tables, $x \mathbin{\dot{\wedge}} y = y \mathbin{\dot{\wedge}} x$. A *tautology* of map theory is an equation $\mathcal{A} = \mathcal{B}$ where $\mathcal{A}$ and $\mathcal{B}$ have identical truth tables.

$r(x)$	$r(y)$	$x \dot{\wedge} y$	$y \dot{\wedge} x$
$\tilde{\mathsf{T}}$	$\tilde{\mathsf{T}}$	T	T
$\tilde{\mathsf{T}}$	$\tilde{\lambda}$	F	F
$\tilde{\mathsf{T}}$	$\tilde{\bot}$	$\bot$	$\bot$
$\tilde{\lambda}$	$\tilde{\mathsf{T}}$	F	F
$\tilde{\lambda}$	$\tilde{\lambda}$	F	F
$\tilde{\lambda}$	$\tilde{\bot}$	$\bot$	$\bot$
$\tilde{\bot}$	$\tilde{\mathsf{T}}$	$\bot$	$\bot$
$\tilde{\bot}$	$\tilde{\lambda}$	$\bot$	$\bot$
$\tilde{\bot}$	$\tilde{\bot}$	$\bot$	$\bot$

Fig. 7. Truth tables of $x \dot{\wedge} y$ and $y \dot{\wedge} x$.

Examples of tautologies are

$$x \dot{\wedge} y = y \dot{\wedge} x, \qquad (x \dot{\wedge} y) \dot{\wedge} z = x \dot{\wedge} (y \dot{\wedge} z), \qquad \dot{\neg}(x \dot{\wedge} y) = \dot{\neg}x \dot{\vee} \dot{\neg}y.$$

The formula $x \dot{\vee} \dot{\neg}x = \mathsf{T}$ is no tautology since, for $x = \bot$, $x \dot{\vee} \dot{\neg}x = \bot \neq \mathsf{T}$. However,

$$x \dot{\vee} \dot{\neg}x = {!x}$$

is a tautology. Likewise, $x \dot{\wedge} x = x$ is no tautology, for if $x = \lambda y.\bot$ then $x \dot{\wedge} x = \mathsf{F} \neq x$ (actually, the right-hand side of $x \dot{\wedge} x = x$ does not even have a truth table, which disqualifies $x \dot{\wedge} x = x$ as a tautology). However,

$$x \dot{\wedge} x = {\approx}x$$

is a tautology.

3.2. Quantifiers

Define

$$\dot{\exists}\mathscr{A} = {\approx}(\mathscr{A}\, \varepsilon\mathscr{A}),$$

$$\dot{\exists}x.\mathscr{A} = \dot{\exists}(\lambda x.\mathscr{A}),$$

$$\dot{\forall}x.\mathscr{A} = \dot{\neg}\dot{\exists}x.\dot{\neg}\mathscr{A}.$$

Assume $\exists x \in \Phi\colon (fx) = \bot$. In this case, $\varepsilon f = \bot$. Now let $x \in \Phi$ satisfy $(fx) = \bot$. By monotonicity, $\dot{\exists}f = (f \varepsilon f) = {\approx}(f\,\bot) \leq {\approx}(fx) = \bot$, so $\dot{\exists}f = \bot$.

Now assume $\forall x \in \Phi\colon (fx) \neq \bot$. If $\exists x \in \Phi\colon (fx) = \mathsf{T}$, then $\varepsilon f \in \Phi$ and $(f \varepsilon f) = \mathsf{T}$, so $\dot{\exists}f = \mathsf{T}$. If $\forall x \in \Phi\colon (fx) \neq \mathsf{T}$, then $\varepsilon f \in \Phi$, so $(f \varepsilon f) \neq \mathsf{T}$. Since $(f \varepsilon f) \neq \mathsf{T}$ and $(f \varepsilon f) \neq \bot$, $\dot{\exists}f = {\approx}(f \varepsilon f) = \mathsf{F}$.

As a result, $\dot{\exists}f = \bot$ if $\exists x \in \Phi\colon (fx) = \bot$ and

$$\exists x \in \Phi\colon (fx) = \mathsf{T} \Rightarrow \dot{\exists}f = \mathsf{T}, \qquad \forall x \in \Phi\colon (fx) \neq \mathsf{T} \Rightarrow \dot{\exists}f = \mathsf{F},$$

otherwise. The construct $\dot{\exists}x.\mathscr{A}$ states that $\mathscr{A}$ holds for some $x \in \Phi$ and $\dot{\forall}x.\mathscr{A}$ states that $\mathscr{A}$ holds for all $x \in \Phi$. However, $\dot{\exists}x.\mathscr{A}$ and $\dot{\forall}x.\mathscr{A}$ equal $\bot$ if $\mathscr{A} = \bot$ for some $x \in \Phi$.

3.3. Recursion

Define

$$S_f = (\lambda x.(f\ (x\ x))),$$
$$\mathsf{Y} = \lambda f.(S_f\ S_f),$$
$$\mathsf{Y}f.\mathcal{A} = (\mathsf{Y}\ \lambda f.\mathcal{A}).$$

For any map f,

$$(\mathsf{Y}\ f) = (S_f\ S_f) = ((\lambda x.(f\ (x\ x)))\ S_f) = (f\ (S_f\ S_f))$$
$$= (f\ (\mathsf{Y}\ f)).$$

Hence, for any f, $(\mathsf{Y}\ f)$ is a g such that $(f g) = g$. For this reason, Y is termed the *fixed point operator* [3].

Now consider a recursive definition such as

$$(g\ y) = (\mathsf{if}\ y\ \mathsf{T}\ (g\ (y\ \mathsf{F}))).$$

This recursive definition is satisfied by the g given by

$$g = \mathsf{Y}f.\lambda y.(\mathsf{if}\ y\ \mathsf{T}\ (f\ (y\ \mathsf{F}))).$$

As another example,

$$(g\ x\ y) = (\mathsf{if}\ y\ \mathsf{T}\ \lambda z.(g\ (y\ z)\ x))$$

is satisfied by

$$g = \mathsf{Y}f.\lambda x.\lambda y.(\mathsf{if}\ y\ \mathsf{T}\ \lambda z.(f\ (y\ z)\ x)).$$

In what follows, recursive definitions are shorthand for the corresponding definitions using Y. Further, a recursive definition like

$$x|y = (\mathsf{if}\ x\ \mathsf{T}\ (\mathsf{if}\ y\ \mathsf{F}\ ((x\ \mathsf{T})|(y\ \mathsf{T}))))$$

is shorthand for

$$x|y = ((\mathsf{Y}f.\lambda x.\lambda y.(\mathsf{if}\ x\ \mathsf{T}\ (\mathsf{if}\ y\ \mathsf{F}\ (f\ (x\ \mathsf{T})\ (y\ \mathsf{T})))))\ x\ y)$$

where $(f x y)$ plays the role of $x|y$ within the scope of $\mathsf{Y}f$.

3.4. Programming

Define

$$hd\, x = (x\ \mathsf{T}),$$
$$tl\, x = (x\ \mathsf{F}),$$
$$\mathsf{nil} = \mathsf{T},$$
$$x :: y = \lambda z.(\mathsf{if}\ z\ x\ y),$$
$$\langle\rangle = \mathsf{nil},$$
$$\langle x\rangle = x :: \mathsf{nil},$$
$$\langle x_1, \ldots, x_n\rangle = x_1 :: \langle x_2, \ldots, x_n\rangle.$$

This defines the concepts *head* or *car* (hd), *tail* or *cdr* (tl), the empty list (nil), *cons* ($::$), and lists $\langle x_1, \ldots, x_n \rangle$ that are typical ingredients of functional programming languages [14].

As an example of a small program, consider the *append* function $\cdot$ which satisfies

$$\langle x_1, \ldots, x_m \rangle \cdot \langle y_1, \ldots, y_n \rangle = \langle x_1, \ldots, x_m, y_1, \ldots, y_n \rangle.$$

A machine executable definition reads

$$x \cdot y = (\mathsf{if}\; x\, y\; ((hd\, x)::((tl\, x) \cdot y))).$$

As mentioned in Section 3.3, this is shorthand for a definition involving Y.

The definition is machine executable in the sense that a computer can compute the value of $x \cdot y$ from no other input than x, y and the definition of $\cdot$.

A definition in map theory is machine executable if it makes no reference to ϕ and ε. Assuming Church's thesis [22], map theory can express any computable function by a machine executable definition. In other words, map theory without ϕ and ε is a general (or Turing complete) computer programming language.

The definitions above introduce lists and list operations. It is also possible to introduce integers, real numbers, arrays, files, exceptions and all other data types used in programming. It is possible but requires a trick to make map theory as efficient as contemporary programming languages. These issues are interesting in themselves, but outside the scope of the present paper.

3.5. Sets

By recursion, define $s(f)$ for all $f \in \Phi$ by

$$s(\mathsf{T}) = \emptyset,$$

$$s(f) = \{s(f\, x) \mid x \in \Phi\} \quad \text{if } f \text{ is proper.}$$

For each $f \in \Phi$, $s(f)$ happens to be a set of *ZFC*, and the well-founded map f is said to *represent* the set $s(f)$. Any set of *ZFC* happens to be representable by at least one $f \in \Phi$.

As an example, $\mathsf{F} = \lambda x.\mathsf{T}$ represents $\{\emptyset\}$:

$$s(\mathsf{F}) = \{s(\mathsf{F}\, x) \mid x \in \Phi\} = \{s(\mathsf{T}) \mid x \in \Phi\} = \{\emptyset \mid x \in \Phi\}$$
$$= \{\emptyset\}.$$

Now define $P = \lambda x.\lambda y.\lambda z.(\mathsf{if}\; z\, x\, y)$. If a, $b \in \Phi$ represent the sets A and B, respectively, then $(P\, a\, b)$ represents $\{A, B\}$:

$$s(P\, a\, b) = \{s(P\, a\, b\, x) \mid x \in \Phi\} = \{s(\mathsf{if}\; x\, a\, b) \mid x \in \Phi\} = \{s(a), s(b)\}$$
$$= \{A, B\}.$$

Hence, P represents the pairing operator. Accidentally, $(P\, a\, b) = a::b$ (cf. Section 3.4).

3.6. Set equality and membership

Let the relation $x \doteq y$ denote that the well-founded maps x and y represent the same set. Further, let a and b be proper, well-founded maps that represent A and B, respectively, i.e.

$$A = \{s(a\ u) \mid u \in \Phi\}, \qquad B = \{s(b\ v) \mid v \in \Phi\}.$$

The maps a and b represent nonempty sets and T represents $\emptyset$, so

$$\mathsf{T} \doteq \mathsf{T} = \mathsf{T}, \qquad \mathsf{T} \doteq b = \mathsf{F}, \qquad a \doteq \mathsf{T} = \mathsf{F}, \quad \text{and}$$

$$a \doteq b = \mathsf{T} \Leftrightarrow A = B.$$

Now rewrite $A = B$ as follows:

$$\begin{aligned} A = B &\Leftrightarrow A \subseteq B \wedge B \subseteq A \\ &\Leftrightarrow \forall U \in A\colon U \in B \wedge \forall V \in B\colon V \in A \\ &\Leftrightarrow \forall U \in A\ \exists V \in B\colon U = V \wedge \forall V \in B\ \exists U \in A\colon U = V \\ &\Leftrightarrow \forall u \in \Phi\ \exists v \in \Phi\colon (a\ u) \doteq (b\ v) \wedge \forall v \in \Phi\ \exists u \in \Phi\colon (a\ u) \doteq (b\ v). \end{aligned}$$

This allows a recursive definition of set equality in map theory:

$$x \doteq y = (\mathsf{if}\ x\ (\mathsf{if}\ y\ \mathsf{T}\ \mathsf{F})\ (\mathsf{if}\ y\ \mathsf{F}\ (\dot{\forall} u\ \dot{\exists} v.(x\ u) \doteq (y\ v))\ \dot{\wedge}\ (\dot{\forall} v\ \dot{\exists} u.(x\ u) \doteq (y\ v)))).$$

Further,

$$\begin{aligned} s(a) \in s(b) &\Leftrightarrow \exists V \in s(b)\colon s(a) = V \\ &\Leftrightarrow \exists v \in \Phi\colon a \doteq (b\ v). \end{aligned}$$

This allows a definition of set membership in map theory:

$$a \mathrel{\dot{\in}} b \;=\; (\mathsf{if}\ b\ \mathsf{F}\ \dot{\exists} v.a \doteq (b\ v)).$$

Now that $\dot{\neg}$, $\dot{\Rightarrow}$, $\dot{\forall}$ and $\dot{\in}$ are all defined, any well-formed formula of *ZFC* is also a well-formed formula of map theory. Section 8 goes further by showing that any theorem of *ZFC* is provable in axiomatic map theory.

Part III goes even further. It proves that for any consistent extension ZFC^{+} of *ZFC* there is an axiomatization $Map^{\circ +}$ of map theory so that any theorem of ZFC^{+} is provable in $Map^{\circ +}$.

3.7. Further set operators

All the usual set operators are treated formally in Section 8. However, the definitions of the set operators are stated below to show how short they are.

In this section, let $\emptyset$ and $\{a, b\}$ be shorthand for T and $(P\ a\ b)$, respectively.

A map p is *set extensive* if $x \doteq y$ implies $(p\ x) \Leftrightarrow (p\ y)$; or, stated more formally, $(x \doteq y) = \mathsf{T}$ implies $((p\ x) \Leftrightarrow (p\ y)) = \mathsf{T}$. Any well-formed formula of *ZFC* is set extensive. Now define

$$(\mathit{Subset}\ a\ p) = (\mathsf{if}\ \dot{\forall} x.\dot{\neg}(p\ (a\ x))\ \emptyset\ \lambda x.(\mathsf{if}\ (p\ (a\ x))\ (a\ x)\ (a\ \varepsilon x.(p\ (a\ x))))).$$

If a represents A and p is set extensive, then $(Subset\ a\ p)$ represents $\{x \in A \mid p(x)\}$. Let $\{x \dot{\in} \mathscr{A} \mid \mathscr{B}\}$ be shorthand for $(Subset\ \mathscr{A}\ \lambda x.\mathscr{B})$.

Define

$$(Power'\ a) = \lambda x.(\mathsf{if}\ x\ \emptyset\ \lambda y.(a\ (x\ (a\ y)))),$$

$$(Union'\ a) = \lambda x.(a\ (x\ \mathsf{T})\ (x\ \mathsf{F})).$$

If a represents A, then $(Power'\ a)$ and $(Union'\ a)$ represent supersets of the power and union sets of A, respectively. Hence, the power and union set operators are definable by

$$x \dot{\subseteq} y = \dot{\forall} z.(z \dot{\in} x \dot{\Rightarrow} z \dot{\in} y),$$

$$(Power\ a) = \{x \dot{\in} (Power'\ a) \mid x \dot{\subseteq} a\},$$

$$(Union\ a) = \{x \dot{\in} (Union'\ a) \mid \dot{\exists} y.(x \dot{\in} y \dot{\wedge} y \dot{\in} a)\}.$$

Define

$$(Choice\ a) = (\mathsf{if}\ a\ \emptyset\ \lambda x.\varepsilon y.(y \dot{\in} (a\ x))).$$

If a represents a set A of disjoint, nonempty sets, then $(Choice\ a)$ represents a choice set of A.

Define

$$\{a\} = \{a, a\},$$

$$a \cup b = (Union\ \{a, b\}),$$

$$(Suc\ a) = a \cup \{a\},$$

$$(\omega\ x) = (\mathsf{if}\ x\ \emptyset\ (Suc\ (\omega\ (x\ \mathsf{T})))).$$

The (recursively defined) map ω represents the least infinite ordinal.

3.8. Beyond set theory

Section 3.5 defined $s(x)$ for all $x \in \Phi$, but the definition makes sense for a much wider range of maps. As an example,

$$s(\lambda x.x) = \{s((\lambda x.x)\ y) \mid y \in \Phi\} = \{s(y) \mid y \in \Phi\},$$

so $\lambda x.x$ represents the class V of all sets. Further,

$$s(\lambda x.\lambda y.y) = \{((\lambda x.\lambda y.y)\ z) \mid z \in \Phi\} = \{s(\lambda y.y)\},$$

so $\lambda x.\lambda y.y$ represents the class $\{V\}$ whose sole element is the class V of all sets. This is not only beyond *ZFC* but also beyond *NBG* in which classes merely contain sets. Now define

$$w = \mathsf{Y} f.\lambda x.f.$$

This w satisfies $w = \lambda x.w$, so

$$s(w) = \{s((\lambda x.w)\ y) \mid y \in \Phi\} = \{s(w)\}.$$

Hence, w represents a non-wellfounded set [2] $W = \{W\}$ which contains itself and nothing else. The given representation provides exactly one such set (i.e. all sets with the property $W = \{W\}$ are equal in the given representation).

Next, define

$$a = \{Power,\ Union\}.$$

The map a represents the class containing the power and unions set operators. Such a construction is totally beyond set theory and is syntactically impossible in *ZFC*.

This shows that map theory has the ability to define more complex structures than set theory.

Map theory allows quantification over arbitrary classes: Define

$$\dot{\forall} x \mathrel{\dot{\in}} y.\mathscr{A} = \dot{\forall} x.((\lambda x.\mathscr{A})\ (y\, x)).$$

For example,

$$\dot{\forall} f \mathrel{\dot{\in}} \{Power,\ Union\}.\dot{\forall} y,\ z.(y \mathrel{\dot{\subseteq}} z \mathrel{\dot{\Rightarrow}} (f\, y) \mathrel{\dot{\subseteq}} (f\, z))$$

states that both the power and the union set operator are monotonic w.r.t. $\dot{\subseteq}$.

3.9. Map versus set theory

Formally, map and set theory are of approximately the same power. More precisely, map theory can prove the consistency of *ZFC*, and *ZFC* plus the existence of a strongly inaccessible ordinal can prove the consistency of map theory.

At first sight, *ZFC* is a very simple theory which has only one concept: the set. This is not correct, however. Set theory has

- sets,
- truth values,
- the membership relation,
- logical connectives and quantifiers.

Further, in practical work, *ZFC* is enriched with

- set operators,
- defined relation symbols (and classes [17]).

Map theory can represent them all, but only has one concept: the map.

Previous sections have represented sets and truth values by maps. The membership relation is representable by the map $\lambda x.\lambda y.(x \mathrel{\dot{\in}} y)$, and a logical connective such as $\dot{\Rightarrow}$ is representable by the map $\lambda x.\lambda y.(x \mathrel{\dot{\Rightarrow}} y)$. Section 3.7 represented the set operator $\bigcup$ by the map *Union*, and a defined relation symbol like $\dot{\subseteq}$ is representable by the map $\lambda x.\lambda y.(x \mathrel{\dot{\subseteq}} y)$.

Even the basic operators if, ε and ϕ of map theory are representable by the maps $\lambda x.\lambda y.\lambda z.(\text{if}\ x\ y\ z)$, $\lambda x.\varepsilon x$ and $\lambda x.\phi x$, respectively. Lambda abstraction as such is hard to represent, but as is well known [3], lambda abstraction may be eliminated by use of the maps $S = \lambda x.\lambda y.\lambda z.((x\, z)(y\, z))$ and $K = \lambda x.\lambda y.x$.

As shown in Section 3.8, map theory is more coherent than set theory in that it is possible, e.g., to form classes that contain set operators. This is a syntactic impossibility of *ZFC*. The difference is due to the fact that map theory only has maps whereas *ZFC* has several, incompatible concepts like sets, truth values and set operators.

3.10. Specification

Programmers occasionally *specify* a program before they write it down [11, 8]. An informal specification of the append program $x \cdot y$ in Section 3.4 may read

$$\text{The program } x \cdot y \text{ shall append the lists } x \text{ and } y. \tag{5}$$

A more formal specification is given by

$$\langle x_1, \ldots, x_m \rangle \cdot \langle y_1, \ldots, y_n \rangle = \langle x_1, \ldots, x_m, y_1, \ldots, y_n \rangle. \tag{6}$$

(The specifications are *incomplete* since they merely specify $x \cdot y$ when x and y are tuples.) Finally, the *program* $x \cdot y$ reads

$$x \cdot y = (\mathsf{if}\; x\, y\; ((hd\, x) :: ((tl\, x) \cdot y))). \tag{7}$$

In general, a specification is a predicate and a program is a machine executable definition. In the example above, the program (7) is said to *satisfy* the specifications (5) and (6).

Formal specifications usually involve quantifiers, so when proving that a given program satisfies a given specification, it is convenient to work in a theory that supports both quantifiers and machine executable definitions. This is exactly what map theory does—and does more coherently than other theories around.

Specification followed by implementation is one approach to programming. Another is as follows: The programmer first writes down the program. However, the programmer allows himself to use quantifiers and Hilbert's ε operator, so the program is not really a program, and definitely not machine executable. Next, the programmer *refines* the program step by step, and in each refinement step he replaces quantifiers and ε operators by executable code that perform the same operations (which is not always possible). If the programmer succeeds, the final program is a genuine, machine executable program. The initial and final programs are mathematically equal, but the former is short and easy to comprehend whereas the latter is larger but machine executable. Again, map theory is a suitable environment for such development activity.

Meta IV [8] also supports various development methods. Meta IV supports countable infinity, quantifiers, and the ε operator (the *such that* operator). It does not support uncountable infinities. Hence, contrary to map theory, Meta IV is unsuited to consider, e.g., topology and its applications to numerical analysis. It is straightforward to define Meta IV within map theory.

The specification language Z [1] is *ZFC* set theory plus a number of defined concepts that are useful in computer science. In comparison with map theory, Z supports quantifiers and anything set theory supports, but it is impossible to state machine executable definitions directly in Z.

$ZFC + (\lambda\text{-calculus})$ [6] supports anything ZFC set theory supports and also supports executable definitions, but $ZFC + (\lambda\text{-calculus})$ is more like a disjoint union than a Cartesian product: it is not possible to mix ZFC and λ-calculus arbitrarily as in map theory.

At this place, unfortunately, it is impossible to discuss the wealth of existing specification languages and compare them all to map theory.

3.11. Russell's and Cantor's paradoxes

Define

$$S = \{x \mid x \notin x\}, \qquad P \Leftrightarrow S \in S.$$

Now

$$P \Leftrightarrow S \in S \Leftrightarrow S \in \{x \mid x \notin x\} \Leftrightarrow S \notin S$$
$$\Leftrightarrow \neg P.$$

This is Russell's paradox for Frege set theory [13]. Now define

$$S' = \lambda x.\dot{\neg}(x\,x), \qquad P' = (S'\,S').$$

In map theory,

$$P' = (S'\,S') = ((\lambda x.\dot{\neg}(x\,x))\,S') = \dot{\neg}(S'\,S')$$
$$= \dot{\neg} P'.$$

However, $P' = \dot{\neg} P'$ is no paradox in map theory; it merely shows $P' = \perp$ since $\perp$ is the unique map in map theory that equals its own negation. Frege set theory is inconsistent because it has unlimited abstraction but no "undefined" truth value $\perp$. Map theory avoids Russell's paradox by having $\perp$.

Cantor's paradox is almost the same as Russell's paradox. To see this, proceed as follows. Let $\mathcal{P}a$ denote the power set of the set a.

Lemma 3.11.1 (Cantor). *For all sets a, a and $\mathcal{P}a$ have different cardinalities.*

Proof. Assume $f: a \to \mathcal{P}a$ is one-to-one. Define $S = \{x \in a \mid x \notin f(x)\}$. Define $S'' \in a$ such that $f(S'') = S$. Now,

$$S'' \in f(S'') \;\Leftrightarrow\; S'' \in S$$
$$\Leftrightarrow\; S'' \notin f(S'').$$

Hence, f cannot be one-to-one so a and $\mathcal{P}a$ have different cardinalities. $\square$

However, let V be the set of all sets. Since $\mathcal{P}V$ is a set of sets, $\mathcal{P}V \subseteq V$. Further (if we do not have uhr-elements), $V \subseteq \mathcal{P}V$, so $V = \mathcal{P}V$ and V and $\mathcal{P}V$ must have equal cardinalities, contradicting Cantor's lemma. This is Cantor's paradox.

The identity function $i: V \to V$ is a one-to-one function from V onto $V = \mathcal{P}V$. Using this for f in the proof of Cantor's lemma yields the definitions $S'' = S = \{x \in V \mid x \notin x\}$ and the contradiction $S'' \in S'' \Leftrightarrow S'' \notin S''$, which is Russell's paradox.

A translation of Cantor's paradox to map theory is omitted since Cantor's paradox is essentially equal to Russell's paradox.

3.12. Burali-Forti's paradox and stages

Burali-Forti's paradox says: The set *On* of all ordinals is an ordinal, and *On* is the largest ordinal there is, but $On \cup \{On\}$ is even bigger, yielding a contradiction. The dilemma is that we cannot allow being able to form the largest ordinal as well as constructing successors of any ordinal. Hence, the problem is to know when to stop: it is convenient to be able to construct as many ordinals as possible, but destructive to form the last one (which is *On*).

One approach to avoid the paradoxes in set theory is outlined in [25]. In this approach, following [25], a set z can have as members only those sets which are formed *before* z. Sets are formed in *stages*, and at each stage, each collection of sets formed at previous stages is formed into a set. There are no other sets than those formed at the stages.

Suppose x is a collection of sets and S is a collection of stages such that each member of x is formed at a stage which is a member of S. If there is a stage after all the members of S, then we can form x at this stage.

It would be convenient if any collection S of stages were followed by a stage T, but then there should be a stage U following all stages. In particular, U should come after itself, which is impossible (as a minimum, *before* and *after* must be partial orders for the stage approach to make sense).

Again, the problem is to know when to stop. It is convenient to have as many stages as possible but destructive to have a stage after all stages.

The introduction of stages does not really avoid the paradoxes. Rather, the stages justify the axiom of foundation that says that $x_1 \ni x_2 \ni x_3 \ni \cdots$ holds for no infinite sequence of x's. The axiom of foundation simplifies set theory, but it is not essentially needed for any practical application. Further, theories of non-well-founded sets [2] may happen to be useful.

The theory *ZFC* is not explicit about when the formation of stages stops, but it does prevent the formation of a stage after all stages. The lack of explicitness makes it undecidable, e.g., whether or not *ZFC* has a strongly inaccessible ordinal.

The theory *NBG* has the same stages as *ZFC*, but also has a "semistage" after all stages. Collections formed at the semistage become *classes* rather than sets. Collections that can be formed at the semistage but not at any previous stage are *proper classes*. Proper classes are not members of any set or class since there is no stage or semistage in *NBG* where collections containing proper classes can be formed.

The approach of map theory is somewhat different. Sections 2.1 and 2.2 introduce the notion of a map by a fairly simple definition, and Section 2.4 introduces the collection of well-founded maps as the least collection such that

$$\mathsf{T} \in \Phi,$$

$$g \in \Phi \wedge \forall x \in g^{\mathrm{so}}\colon (fx) \in \Phi \Rightarrow f \in \Phi.$$

The "least collection ... closed under ..."-construct tacitly introduces the well-founded maps in stages but, contrary to *ZFC* and *NBG*, the minimality requirement and the built in size limitation determines the collection of well-founded maps uniquely. As an example, the natural model of map theory contains no strongly inaccessible ordinal. (But, as shown in Part III, there are axiomatizations of map theory that allow strongly inaccessible ordinals to exist.)

Among other things, map theory is an attempt to form a theory with a simpler and more intuitively appealing universe than the various theories of sets. Even though some of the goal has been achieved, map theory is not completely successful on this point. To see this, consider the map

$$S = \lambda x.\dot{\neg}(x\,x).$$

Next define $S' \in M \to M$ by

$$(S'\,x) = \begin{cases} \mathsf{T} & \text{if } x \geq S, \\ (S\,x) & \text{otherwise.} \end{cases}$$

Now, $x \leq y \Rightarrow (S'\,x) \leq (S'\,y)$, so S' is monotonic and $S' \in (M \xrightarrow{\text{m}} M)_{\mathsf{T}\perp}$. However, $S' \notin M$, for if $S' \in M$ then $S' \geq S$, $(S'\,S') = \mathsf{T}$ and

$$\begin{aligned} \mathsf{T} &= (S'\,S') \\ &\geq (S\,S') \\ &= ((\lambda x.\dot{\neg}(x\,x))\,S') \\ &= \dot{\neg}(S'\,S') \\ &= \dot{\neg}\mathsf{T} \\ &= \mathsf{F}. \end{aligned}$$

Hence, $\mathsf{T} \geq \mathsf{F}$ which contradicts the definition of $\geq$. In conclusion, S' is monotonic but not a map.

The problem here is that there is no obvious reason why S' should not be a map. Rather, the assumption that S' is a map leads to a contradiction. This is the same situation as in *ZFC*: The assumption that the set of all sets exists leads to a contradiction, so there is no such set.

The set of all sets would be very convenient to have at hand, whereas there are probably no uses of S' above. In general, the closure properties of map theory are strong, and it is hard to think of a useful, monotonic pre-map which is not a map. However, it is still annoying that map theory offers no precise, positive characterization of maps.

Note, that no term of map theory denotes S', so S' does not give rise to a paradox.

It would be a great enhancement if all monotonic, *continuous* operations were maps, where *continuity* would have to be defined somehow. Reasonable definitions of continuity exist within computer science, but these definitions rule out noncomputable constructs like the ε operator and universal and existential quantifiers.

3.13. *The notion of truth*

In *ZFC*, a *notion of truth* is a predicate $p(x)$ such that $p(\lceil \mathscr{A} \rceil) \Leftrightarrow \mathscr{A}$ for closed, well-formed formulas $\mathscr{A}$ where $\lceil \mathscr{A} \rceil$ denotes the *Gödel number* of $\mathscr{A}$. According to Tarski's theorem [22, 19], the notion of truth of *ZFC* is not definable in *ZFC* (provided *ZFC* is consistent). In contrast, as shown below, map theory is capable of defining its own notion of truth.

For each nonnegative integer i define

$$\lceil i \rceil = \overbrace{\langle \mathsf{T}, \ldots, \mathsf{T} \rangle}^{i}.$$

In particular, $\lceil 0 \rceil = \mathsf{T}$, $\lceil 1 \rceil = \langle \mathsf{T} \rangle$ and $\lceil 2 \rceil = \langle \mathsf{T}, \mathsf{T} \rangle$. The equations below assign a Gödel number $\lceil \mathscr{A} \rceil$ to each term $\mathscr{A}$ of map theory.

$$\lceil \mathsf{T} \rceil = \mathsf{T}.$$

$$\lceil \perp \rceil = \langle \mathsf{T} \rangle.$$

$$\lceil x_i \rceil = \langle \lceil i \rceil, \mathsf{T} \rangle.$$

$$\lceil (\mathscr{A}\,\mathscr{B}) \rceil = \langle \lceil \mathscr{A} \rceil, \lceil \mathscr{B} \rceil, \mathsf{T} \rangle.$$

$$\lceil \lambda x_i . \mathscr{A} \rceil = \langle \lceil i \rceil, \lceil \mathscr{A} \rceil, \mathsf{T}, \mathsf{T} \rangle.$$

$$\lceil (\text{if } \mathscr{A}\,\mathscr{B}\,\mathscr{C}) \rceil = \langle \lceil \mathscr{A} \rceil, \lceil \mathscr{B} \rceil, \lceil \mathscr{C} \rceil, \mathsf{T}, \mathsf{T} \rangle.$$

$$\lceil \varepsilon \mathscr{A} \rceil = \langle \lceil \mathscr{A} \rceil, \mathsf{T}, \mathsf{T}, \mathsf{T}, \mathsf{T}, \mathsf{T} \rangle.$$

$$\lceil \phi \mathscr{A} \rceil = \langle \lceil \mathscr{A} \rceil, \mathsf{T}, \mathsf{T}, \mathsf{T}, \mathsf{T}, \mathsf{T}, \mathsf{T} \rangle.$$

Define

$$x[y] = (\text{if } y\ (hd\,x)\ (tl\,x)[tl\,y])$$

(this is a recursive definition of the two-place construct $\bullet[\bullet]$). Now,

$$\langle x_0, \ldots, x_n \rangle [\lceil i \rceil] = x_i$$

for $0 \leqslant i \leqslant n$. Define

$$[x/y := z] = (\text{if } y\ (z::(tl\,x))\ (hd\,x::[tl\,x/tl\,y := z])).$$

Now,

$$[\langle x_0, \ldots, x_n \rangle / \lceil i \rceil := y] = \langle x_0, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n \rangle$$

for $0 \leqslant i \leqslant n$. Define

$$
\begin{aligned}
tl^n\, x &= \overbrace{tl \ldots tl}^{n}\, x, \\
x' &= hd\, x, \\
x'' &= hd\; tl\, x, \\
x''' &= hd\; tl\; tl\, x, \\
x_y &= (\text{if } (tl^0\, x) \;\; \mathsf{T} \\
&\qquad (\text{if } (tl^1\, x) \;\; \perp \\
&\qquad (\text{if } (tl^2\, x) \;\; y[x'] \\
&\qquad (\text{if } (tl^3\, x) \;\; (x'_y\; x''_y) \\
&\qquad (\text{if } (tl^4\, x) \;\; \lambda z.x''_{[y/x':=z]} \\
&\qquad (\text{if } (tl^5\, x) \;\; (\text{if } x'_y\; x''_y\; x'''_y) \\
&\qquad (\text{if } (tl^6\, x) \;\; \varepsilon(x'_y) \\
&\qquad\qquad\qquad\quad \phi(x'_y)))))))).
\end{aligned}
$$

As an example,

$$\lceil \lambda x_2.(\text{if } x_2\; x_1\; x_0) \rceil_{\langle a,b,c,d \rangle} = \lambda x_2.(\text{if } x_2\; b\; a).$$

In general, $\lceil \mathscr{A} \rceil_{\langle a_0,\ldots,a_n \rangle}$ is the value of $\mathscr{A}$ for $x_0 = a_0, \ldots, x_n = a_n$ and $x_{n+1}, x_{n+2}, \ldots =$ T. If $\mathscr{A}$ has no free variables, then $\lceil \mathscr{A} \rceil_{\mathsf{T}} = \mathscr{A}$.

Now define $p(x) = x_{\mathsf{T}}$. Terms—among others—serve as predicates in map theory, and $p(\lceil \mathscr{A} \rceil) = \mathscr{A}$ for closed terms $\mathscr{A}$, so p is a notion of truth in map theory defined within map theory.

3.14. Category theory

A *category* $\mathscr{C}$ is a structure consisting of

- a collection $\bar{\mathscr{C}}$ of objects,
- a ternary relation $f : A \to B$,
- a binary operation $\circ$, and
- a unary operation id_A (defined for all $A \in \bar{\mathscr{C}}$),

which for $A, B, C, D \in \bar{\mathscr{C}}$, $f : A \to B$, $g : B \to C$ and $h : C \to D$ satisfies

- $g \circ f : A \to C$,
- $(h \circ g) \circ f = h \circ (g \circ f)$,
- $id_A : A \to A$, and
- $f \circ id_A = id_B \circ f = f$.

If $f : A \to B$ for some $A, B \in \bar{\mathscr{C}}$, then f is a *morphism* of $\mathscr{C}$.

Category theory is the theory of categories [4] but, contrary to set theory, category theory has remained informal.

Within set theory, various formal theories like *ZFC* and *NBG* have been proposed, and each formal theory attempts to cover all of set theory (see [13] and [22] for an overview of theories of sets). Each formal theory of sets contains ways of forming sets that are ensured to exist by the theory. Further, for each theory of sets it makes sense to ask whether or not the theory is consistent.

No generally accepted formal theory attempts to account for all of category theory. Generally accepted theories of categories do not allow mechanical checking of categorical proofs, they do not offer ways of forming categories that are ensured to exist, and it does not make sense to ask whether or not a theory is consistent. Workers in category theory may encounter paradoxes, but try to avoid them ad hoc.

Section 8 develops all of *ZFC* set theory within map theory, and Part III proves that for any consistent extension of *ZFC* there is a matching axiomatization of map theory with the same expressive power. In conclusion, map theory can do anything set theory can do.

It is impossible to do the same for category theory since "all of category theory" is not sufficiently well defined. In the following, categories are represented in map theory and the category of *all* categories is introduced. Formalization of the various constructions of category theory in map theory would be a substantial task, but the award would be that category theory became formalized. The ability to define the category of all categories indicates that map theory is a reasonable basis for such a formalization. The ability to define the category of all categories is due to the unlimited abstraction of map theory.

Define

$$\begin{aligned}
p_i\, x &= hd\, \overbrace{tl \ldots tl}^{i-1}\, x \\
x \stackrel{\star}{=} y &= \langle x, y \rangle \\
x \stackrel{\star}{\wedge} y &= \langle \langle p_1\, x, p_1\, y \rangle, \langle p_2\, x, p_2\, y \rangle \rangle \\
\stackrel{\star}{\forall} x.\mathscr{A} &= \langle \lambda x.(p_1\, \mathscr{A}), \lambda x.(p_2\, \mathscr{A}) \rangle \\
x \stackrel{w}{\Rightarrow} y &= (p_1\, y) \stackrel{\star}{=} (w\, x\, y\, (p_1\, x)\, (p_2\, x)) \stackrel{\star}{\wedge} (p_2\, y) = (w\, x\, y\, (p_2\, x)\, (p_1\, x))
\end{aligned}$$

and let $\star\mathscr{A}$ be shorthand for $(p_1\, \mathscr{A}) = (p_2\, \mathscr{A})$. With these conventions

$$\begin{aligned}
&p_i \langle x_1, \ldots x_n \rangle = x_i \quad \text{for } 1 \leq i \leq n, \\
&\star(\mathscr{A} \stackrel{\star}{=} \mathscr{B}) \Leftrightarrow \mathscr{A} = \mathscr{B}, \\
&\star(\mathscr{A} \stackrel{\star}{\wedge} \mathscr{B}) \Leftrightarrow (\star\mathscr{A}) \wedge (\star\mathscr{B}), \\
&\star(\stackrel{\star}{\forall} x.\mathscr{A}) \Leftrightarrow \forall x \in M : \star\mathscr{A}, \\
&\star(\mathscr{A} \stackrel{w}{\Rightarrow} \mathscr{B}) \Rightarrow ((\star\mathscr{A}) \Rightarrow (\star\mathscr{B})).
\end{aligned}$$

In $\mathscr{A} \stackrel{w}{\Rightarrow} \mathscr{B}$, w is a *witness* that testifies that $\star\mathscr{A}$ implies $\star\mathscr{B}$.

Map theory may represent a category $\mathscr{C}$ as a map $\langle \bar{\mathscr{C}}, \bar{\mathscr{M}}, c, id, w \rangle$ such that $\star(\bar{\mathscr{C}}\, A)$ states that A is an object of $\mathscr{C}$, $\star(\bar{\mathscr{M}}\, f\, A\, B)$ states that $f : A \rightarrow B$, $(c\, f\, g)$ denotes $f \circ g$,

$(id\,A)$ denotes id_A, and w witnesses that $\mathscr{C}$ is a category. For all maps $\mathscr{C}$, A, B and f define

$$\begin{aligned}
A \overset{\star}{\in} \mathscr{C} &= ((p_1\,\mathscr{C})\,A),\\
f : A \overset{\mathscr{C}}{\to} B &= ((p_2\,\mathscr{C})\,f\,A\,B),\\
f \overset{\mathscr{C}}{\circ} g &= ((p_3\,\mathscr{C})\,f\,g),\\
id_A^{\mathscr{C}} &= ((p_4\,\mathscr{C})\,A),\\
A_1, \ldots, A_n \overset{\star}{\in} \mathscr{C} &= A_1 \overset{\star}{\in} \mathscr{C} \overset{\star}{\wedge} \cdots \overset{\star}{\wedge} A_n \overset{\star}{\in} \mathscr{C}.
\end{aligned}$$

Further define

$$\begin{aligned}
(cat\,\mathscr{C}) = \overset{\star}{\forall} &A, B, C, D, f, g, h.\\
&(A, B, C, D \overset{\star}{\in} \mathscr{C} \overset{\star}{\wedge} f : A \overset{\mathscr{C}}{\to} B \overset{\star}{\wedge} g : B \overset{\mathscr{C}}{\to} C \overset{\star}{\wedge} h : C \overset{\mathscr{C}}{\to} D\\
&\xRightarrow{(p_5\,\mathscr{C})} g \overset{\mathscr{C}}{\circ} f : A \overset{\mathscr{C}}{\to} C \overset{\star}{\wedge} (h \overset{\mathscr{C}}{\circ} g) \overset{\mathscr{C}}{\circ} f \overset{\star}{=} h \overset{\mathscr{C}}{\circ} (g \overset{\mathscr{C}}{\circ} f)\\
&\qquad \overset{\star}{\wedge} id_A^{\mathscr{C}} : A \overset{\mathscr{C}}{\to} A \overset{\star}{\wedge} f \overset{\mathscr{C}}{\circ} id_A^{\mathscr{C}} \overset{\star}{=} f \overset{\star}{\wedge} id_B^{\mathscr{C}} \overset{\mathscr{C}}{\circ} f \overset{\star}{=} f).
\end{aligned}$$

Now, $\star(cat\,\mathscr{C})$ holds iff $\mathscr{C}$ represents a category.

Map theory may represent a functor $F : \mathscr{C} \to \mathscr{D}$ as a map $\langle \bar{F}, \bar{\bar{F}}, w \rangle$ such that $\bar{F}$ is a mapping of objects, $\bar{\bar{F}}$ is a mapping of morphisms, and w witnesses that F is a functor. Define

$$\begin{aligned}
\bar{F} &= (p_1\,F),\\
\bar{\bar{F}} &= (p_2\,F),\\
(func\,F\,\mathscr{C}\,\mathscr{D}) &= \overset{\star}{\forall} A, B, C, f, g.\\
&(A, B, C \overset{\star}{\in} \mathscr{C} \overset{\star}{\wedge} f : A \overset{\mathscr{C}}{\to} B \overset{\star}{\wedge} g : B \overset{\mathscr{C}}{\to} C\\
&\xRightarrow{((p_3\,F)\,\mathscr{C}\,\mathscr{D})} (\bar{F}A) \overset{\star}{\in} \mathscr{D} \overset{\star}{\wedge} (\bar{\bar{F}} f) : (\bar{F}A) \overset{\mathscr{D}}{\to} (\bar{F}B) \overset{\star}{\wedge} (\bar{\bar{F}}\,id_A^{\mathscr{C}}) \overset{\star}{=} id_{(\bar{F}A)}^{\mathscr{D}}\\
&\qquad \overset{\star}{\wedge} (\bar{\bar{F}}\,(g \overset{\mathscr{C}}{\circ} f)) \overset{\star}{=} (\bar{\bar{F}} g) \overset{\mathscr{D}}{\circ} (\bar{\bar{F}} f)).
\end{aligned}$$

Now, $\star(func\,F\,\mathscr{C}\,\mathscr{D})$ expresses that $F : \mathscr{C} \to \mathscr{D}$ is a functor. The identity functor $id_{\mathscr{C}}$ and functor composition are expressible by

$$\begin{aligned}
&\overline{id}_{\mathscr{C}} = \langle \lambda x.x, \lambda x.x, \mathscr{A}' \rangle,\\
&G \bar{\circ} F = \langle \lambda x.(\bar{G}\,(\bar{F}x)), \lambda x.(\bar{\bar{G}}\,(\bar{\bar{F}}x)), \mathscr{A}'' \rangle,
\end{aligned}$$

for suitable terms $\mathscr{A}'$ and $\mathscr{A}''$. The category Cat of all categories can be written as

$$Cat = \langle cat, func, \lambda x.\lambda y.(x \bar{\circ} y), \lambda x.\overline{id}_x, \mathscr{A} \rangle$$

for a suitable term $\mathscr{A}$.

The above is merely a very rough outline of one approach to formalize category theory in map theory. It is a substantial task to make the formalization of category theory fluent and thorough.

In particular, the "suitable" terms $\mathscr{A}'$, $\mathscr{A}''$ and $\mathscr{A}$ above have to be specified. There is no point in specifying them directly, since they are large and complex. Rather, it is reasonable to state and prove metatheorems like

$$\star(A_1 \overset{\star}{\wedge} \cdots \overset{\star}{\wedge} A_n \overset{\mathscr{A}}{\Rightarrow} A_i) \quad \text{for a suitable } \mathscr{A},$$

$$\star(A \overset{\mathscr{A}}{\Rightarrow} C) \quad \text{for a suitable } \mathscr{A} \text{ if } \star(A \overset{\mathscr{A}'}{\Rightarrow} B) \text{ and } \star(B \overset{\mathscr{A}''}{\Rightarrow} C),$$

$$\star(A \overset{\mathscr{A}}{\Rightarrow} B \overset{\star}{\wedge} C) \quad \text{for a suitable } \mathscr{A} \text{ if } \star(A \overset{\mathscr{A}'}{\Rightarrow} B) \text{ and } \star(A \overset{\mathscr{A}''}{\Rightarrow} C).$$

Part II. Axiomatization of map theory

4. Elementary axioms

4.1. Presentation of axioms

The definition of (if $\mathscr{A}\,\mathscr{B}\,\mathscr{C}$) combined with $r(\mathsf{T}) = \tilde{\mathsf{T}}$, $r(\lambda x.\mathscr{A}) = \tilde{\lambda}$ and $r(\perp) = \tilde{\perp}$ gives the following axioms

(Select1) $\quad \vdash(\text{if } \mathsf{T}\ \mathscr{B}\ \mathscr{C}) = \mathscr{B}$.

(Select2) $\quad \vdash(\text{if } (\lambda x.\mathscr{A})\ \mathscr{B}\ \mathscr{C}) = \mathscr{C}$.

(Select3) $\quad \vdash(\text{if } \perp\ \mathscr{B}\ \mathscr{C}) = \perp$.

(Script letters $\mathscr{A}$, $\mathscr{B}$, $\mathscr{C}$, etc. denote arbitrary terms.)

In axioms, $\vdash \mathscr{A} = \mathscr{B}$ states that $\mathscr{A} = \mathscr{B}$ is an axiom. In inference rules, $\mathscr{A}_1 = \mathscr{B}_1; \cdots; \mathscr{A}_n = \mathscr{B}_n \vdash \mathscr{A} = \mathscr{B}$ states that $\mathscr{A}_1 = \mathscr{B}_1, \ldots, \mathscr{A}_n = \mathscr{B}_n$ directly infers $\mathscr{A} = \mathscr{B}$. In metatheorems, $\mathscr{A}_1 = \mathscr{B}_1; \cdots; \mathscr{A}_n = \mathscr{B}_n \vdash \mathscr{A} = \mathscr{B}$ states that if $\mathscr{A}_1 = \mathscr{B}_1, \ldots, \mathscr{A}_n = \mathscr{B}_n$ are all provable, then $\mathscr{A} = \mathscr{B}$ is provable. Section 5.4 defines a fourth use of $\vdash$.

For terms $\mathscr{A}$ and $\mathscr{B}$ and a variable x, $[\mathscr{A}/x := \mathscr{B}]$ denotes the result of replacing all free x's in $\mathscr{A}$ by $\mathscr{B}$. The term $\mathscr{B}$ is said to be *free for x in $\mathscr{A}$* [22] if no free variable in $\mathscr{B}$ becomes bound in $[\mathscr{A}/x := \mathscr{B}]$.

The conventions $(\mathsf{T}\ x) = \mathsf{T}$ and $(\perp\ x) = \perp$ and the definition of lambda abstraction gives three further axioms.

(Apply1) $\quad \vdash(\mathsf{T}\ \mathscr{B}) = \mathsf{T}$.

(Apply2) $\quad \vdash((\lambda x.\mathscr{A})\ \mathscr{B}) = [\mathscr{A}/x := \mathscr{B}]$ if $\mathscr{B}$ is free for x in $\mathscr{A}$.

(Apply3) $\quad \vdash(\perp\ \mathscr{B}) = \perp$.

The definition of $\lambda x.\mathscr{E}$ shows that the names of bound variables are insignificant:

(Rename) $\quad \vdash \lambda x.[\mathscr{A}/y := x] = \lambda y.[\mathscr{A}/x := y]$

if x is free for y in $\mathscr{A}$ and y is free for x in $\mathscr{A}$.

The following inference rules provide the axiomatic description of equality in map theory. They describe the transitivity (trans) and substitutivity (sub1 and sub2) of equality.

(trans) $\mathscr{A} = \mathscr{B}\ ;\ \mathscr{A} = \mathscr{C}\ \vdash\ \mathscr{B} = \mathscr{C}$.

(sub1) $\mathscr{A} = \mathscr{B}\ ;\ \mathscr{C} = \mathscr{D}\ \vdash\ (\mathscr{A}\ \mathscr{C}) = (\mathscr{B}\ \mathscr{C})$.

(sub2) $\mathscr{A} = \mathscr{B}\ \vdash\ \lambda x.\mathscr{A} = \lambda x.\mathscr{B}$.

A *formula* (or *well-formed formula*) of map theory has the form $\mathscr{A} = \mathscr{B}$ where $\mathscr{A}$ and $\mathscr{B}$ are terms. Free variables of $\mathscr{A}$ and $\mathscr{B}$ implicitly range over all maps.

Formulas of set theory are more complex; they are composed from atomic formulas $x \in y$, negation $\neg\mathscr{A}$, implication $\mathscr{A} \Rightarrow \mathscr{B}$, and quantification $\forall x\colon \mathscr{A}$. As shown later, map theory treats membership, negation, implication, and quantification as defined concepts at term level.

Let $\mathscr{A}$ denote some term. Here is an example of a formal proof of $\mathscr{A} = \mathscr{A}$ in map theory.

Proof of $\mathscr{A} = \mathscr{A}$.

1. Select1 $(\mathsf{if}\ \mathsf{T}\ \mathscr{A}\ \mathscr{A}) = \mathscr{A}$
2. Select1 $(\mathsf{if}\ \mathsf{T}\ \mathscr{A}\ \mathscr{A}) = \mathscr{A}$
3. 1,2,trans $\mathscr{A} = \mathscr{A}$. □

A more terse proof reads as follows.

Proof of $\mathscr{A} = \mathscr{A}$.

1. Select1 $(\mathsf{if}\ \mathsf{T}\ \mathscr{A}\ \mathscr{A}) = \mathscr{A}$
2. 1,1,trans $\mathscr{A} = \mathscr{A}$. □

4.2. Metatheorems of equality

Theorem 4.2.1. *For all terms $\mathscr{A}$, $\mathscr{B}$ and $\mathscr{C}$,*

$$\vdash\ \mathscr{A} = \mathscr{A},$$

$$\mathscr{A} = \mathscr{B}\ \vdash\ \mathscr{B} = \mathscr{A},$$

$$\mathscr{A} = \mathscr{B}\,;\,\mathscr{B} = \mathscr{C}\ \vdash\ \mathscr{A} = \mathscr{C},$$

The theorem states that $\mathscr{A} = \mathscr{A}$ is provable in axiomatic map theory for any term $\mathscr{A}$. Further, if $\mathscr{A} = \mathscr{B}$ is provable, then $\mathscr{B} = \mathscr{A}$ is provable, and if both $\mathscr{A} = \mathscr{B}$ and $\mathscr{B} = \mathscr{C}$ are provable, then $\mathscr{A} = \mathscr{C}$ is provable. The statement $\mathscr{A} = \mathscr{A}$ was proved in Section 4.1.

Proof of $\mathscr{A} = \mathscr{B} \vdash \mathscr{B} = \mathscr{A}$.

1. hyp $\quad \mathscr{A} = \mathscr{B}$
2. Th.4.2.1 $\quad \mathscr{A} = \mathscr{A}$
3. 1,2,trans $\quad \mathscr{B} = \mathscr{A}$. □

Proof of $\mathscr{A} = \mathscr{B}\,;\mathscr{B} = \mathscr{C} \vdash \mathscr{A} = \mathscr{C}$.

1. hyp $\quad \mathscr{A} = \mathscr{B}$
2. 1,Th.4.2.1 $\quad \mathscr{B} = \mathscr{A}$
3. hyp $\quad \mathscr{B} = \mathscr{C}$
4. 1,3,trans $\quad \mathscr{A} = \mathscr{C}$. □

Lemma 4.2.2. *If* $\vdash \mathscr{A} = \mathscr{A}'$, $\vdash \mathscr{B} = \mathscr{B}'$ *and* $\vdash \mathscr{C} = \mathscr{C}'$, *then* $\vdash (\mathsf{if}\ \mathscr{A}\ \mathscr{B}\ \mathscr{C}) = (\mathsf{if}\ \mathscr{A}'\ \mathscr{B}'\ \mathscr{C}')$, $\vdash \varepsilon\mathscr{A} = \varepsilon\mathscr{A}'$ *and* $\vdash \phi\mathscr{A} = \phi\mathscr{A}'$.

Proof of $\mathscr{A} = \mathscr{A}' \vdash \varepsilon\mathscr{A} = \varepsilon\mathscr{A}'$.

1. Th.4.2.1 $\quad \lambda x.\varepsilon x = \lambda x.\varepsilon x$
2. hyp $\quad \mathscr{A} = \mathscr{A}'$
3. 1,2,sub1 $\quad ((\lambda x.\varepsilon x)\mathscr{A}) = ((\lambda x.\varepsilon x)\mathscr{A}')$
4. Apply2 $\quad ((\lambda x.\varepsilon x)\mathscr{A}) = \varepsilon\mathscr{A}$
5. 3,4,trans $\quad ((\lambda x.\varepsilon x)\mathscr{A}') = \varepsilon\mathscr{A}$
6. Apply2 $\quad ((\lambda x.\varepsilon x)\mathscr{A}') = \varepsilon\mathscr{A}'$
7. 5,6,trans $\quad \varepsilon\mathscr{A} = \varepsilon\mathscr{A}'$. □

The proof of the remainder of Lemma 4.2.2 is analogous.

Theorem 4.2.3. $\mathscr{A} = \mathscr{A}' \vdash \mathscr{B} = \mathscr{C}$ *if* $\mathscr{C}$ *arises from* $\mathscr{B}$ *by replacing* $\mathscr{A}$ *by* $\mathscr{A}'$ *any number of times.*

Proof. The lemma follows by structural induction (or by induction in the number of connectives) in $\mathscr{B}$ from the following fact: If $\vdash \mathscr{D} = \mathscr{D}'$, $\vdash \mathscr{E} = \mathscr{E}'$ and $\vdash \mathscr{F} = \mathscr{F}'$ then

$$\vdash \lambda x.\mathscr{D} = \lambda x.\mathscr{D}'$$
$$\vdash (\mathscr{D}\ \mathscr{E}) = (\mathscr{D}'\ \mathscr{E}')$$
$$\vdash (\mathsf{if}\ \mathscr{D}\ \mathscr{E}\ \mathscr{F}) = (\mathsf{if}\ \mathscr{D}'\ \mathscr{E}'\ \mathscr{F}')$$
$$\vdash \varepsilon\mathscr{D} = \varepsilon\mathscr{D}'$$
$$\vdash \phi\mathscr{D} = \phi\mathscr{D}'$$
$$\vdash \mathsf{T} = \mathsf{T}$$
$$\vdash \bot = \bot$$
$$\vdash x = x. \qquad \square$$

Theorem 4.2.4. *If $\mathscr{A}$ and $\mathscr{A}'$ are identical except for naming of bound variables, then* $\vdash \mathscr{A} = \mathscr{A}'$.

The theorem follows from the *Rename* axiom.

4.3. Metatheorems of reduction

If $\mathscr{A} = \mathscr{B}$ is an axiom according to one of the schemes (Select 1-3) or (Apply 1-3), and if $\mathscr{C}$ and $\mathscr{D}$ are identical except that one subterm of $\mathscr{C}$ of form $\mathscr{A}$ is replaced by $\mathscr{B}$ in $\mathscr{D}$, then $\mathscr{D}$ is said to be a *reduct* of $\mathscr{C}$. Let $\mathscr{C} \xrightarrow{r} \mathscr{D}$ denote that either $\mathscr{D}$ is a reduct of $\mathscr{C}$ or $\mathscr{D}$ is identical to $\mathscr{C}$ except possibly for renaming of bound variables. Let $\mathscr{A} \stackrel{\bullet}{=} \mathscr{B}$ denote that the terms $\mathscr{A}$ and $\mathscr{B}$ are identical.

Let $\mathscr{A}_0 \xrightarrow{r} \mathscr{A}_1 \xrightarrow{r} \cdots \xrightarrow{r} \mathscr{A}_n$ stand for $\mathscr{A}_{i-1} \xrightarrow{r} \mathscr{A}_i$ for all $i \in \{1, \ldots, n\}$. If $\mathscr{A}_0 \xrightarrow{r} \mathscr{A}_1 \xrightarrow{r} \cdots \xrightarrow{r} \mathscr{A}_n$ for some $\mathscr{A}_1, \ldots, \mathscr{A}_{n-1}$, then $\mathscr{A}_0 \xrightarrow{*} \mathscr{A}_n$. Obviously, if $\mathscr{A} \xrightarrow{r} \mathscr{B}$, then $\vdash \mathscr{A} = \mathscr{B}$. Also, if $\mathscr{A} \xrightarrow{*} \mathscr{B}$, then $\vdash \mathscr{A} = \mathscr{B}$. Hence, we have the following result.

Theorem 4.3.1 (Reduction). *If $\mathscr{A} \xrightarrow{*} \mathscr{C}$ and $\mathscr{B} \xrightarrow{*} \mathscr{C}$ for some $\mathscr{C}$, then $\vdash \mathscr{A} = \mathscr{B}$.*

Define

$$\mathsf{F} = \lambda x.\mathsf{T}, \qquad \dot{\neg} x = (\mathsf{if}\ x\ \mathsf{F}\ \mathsf{T}).$$

The reductions

$$\dot{\neg}\mathsf{T} \stackrel{\bullet}{=} (\mathsf{if}\ \mathsf{T}\ \mathsf{F}\ \mathsf{T}) \xrightarrow{r} \mathsf{F},$$

$$\dot{\neg}\mathsf{F} \stackrel{\bullet}{=} (\mathsf{if}\ (\lambda x.\mathsf{T})\ \mathsf{F}\ \mathsf{T}) \xrightarrow{r} \mathsf{F} \quad \text{and}$$

$$\dot{\neg}\bot \stackrel{\bullet}{=} (\mathsf{if}\ \bot\ \mathsf{F}\ \mathsf{T}) \xrightarrow{r} \bot,$$

show that

$$\vdash \dot{\neg}\mathsf{T} = \mathsf{F}, \qquad \vdash \dot{\neg}\mathsf{F} = \mathsf{T}, \qquad \vdash \dot{\neg}\bot = \bot.$$

Define

$$cons = \lambda x.\lambda y.\lambda z.(\mathsf{if}\ z\ x\ y),$$

$$hd = \lambda x.(x\ \mathsf{T}),$$

$$tl = \lambda x.(x\ \mathsf{F}).$$

The reduction

$$\begin{aligned}(hd\ (cons\ x\ y)) &\stackrel{\bullet}{=} ((\lambda x.(x\ \mathsf{T}))\ (cons\ x\ y))\\ &\xrightarrow{r} (cons\ x\ y\ \mathsf{T})\\ &\stackrel{\bullet}{=} ((\lambda x.\lambda y.\lambda z.(\mathsf{if}\ z\ x\ y))\ x\ y\ \mathsf{T})\\ &\xrightarrow{r} ((\lambda y.\lambda z.(\mathsf{if}\ z\ x\ y))\ y\ \mathsf{T})\\ &\xrightarrow{r} ((\lambda z.(\mathsf{if}\ z\ x\ y))\ \mathsf{T})\\ &\xrightarrow{r} (\mathsf{if}\ \mathsf{T}\ x\ y)\\ &\xrightarrow{r} x\end{aligned}$$

shows that

$$\vdash (hd\ (cons\ x\ y)) = x.$$

Likewise

$$\vdash (tl\ (cons\ x\ y)) = y.$$

Like in Section 3.3 define

$$S_f = \lambda x.(f\ (x\ x))$$

$$\mathsf{Y} = \lambda f.(S_f\ S_f),$$

$$\mathsf{Y}f.\mathscr{A} = (\mathsf{Y}\ \lambda f.\mathscr{A}).$$

The reduction theorem gives

$$\vdash (\mathsf{Y}\ \mathscr{A}) = (\mathscr{A}\ (\mathsf{Y}\ \mathscr{A})).$$

Further, if $\mathsf{Y}f.\mathscr{A}$ is free for f in $\mathscr{A}$ then

$$\vdash \mathsf{Y}f.\mathscr{A} = [\mathscr{A}/f := \mathsf{Y}f.\mathscr{A}].$$

The map Y is the "fixed point operator" [3]. To define, e.g., a map *mirror* which satisfies

$$(mirror\ x) = (\mathsf{if}\ x\ \mathsf{T}\ (cons\ (mirror\ (tl\ x))\ (mirror\ (hd\ x)))),$$

the following definition will do:

$$mirror = \mathsf{Y}f.\lambda x.(\mathsf{if}\ x\ \mathsf{T}\ (cons\ (f\ (tl\ x))\ (f\ (hd\ x)))).$$

It is possible to strengthen the reduction theorem: $\mathscr{A} \xrightarrow{*} \mathscr{C}$ and $\mathscr{B} \xrightarrow{*} \mathscr{C}$ for some $\mathscr{C}$ if and only if $\mathscr{A} = \mathscr{B}$ is provable from the axioms and inference rules stated until now. The *only if* part is similar to the reduction theorem whereas the *if* part follows from the Church–Rosser Theorem [3].

Part III defines a model of map theory in *ZFC*. Let $\models \mathscr{A} = \mathscr{B}$ denote that $\mathscr{A} = \mathscr{B}$ holds in that particular model. For all terms $\mathscr{A}$ and $\mathscr{B}$, $\vdash \mathscr{A} = \mathscr{B}$ implies $\models \mathscr{A} = \mathscr{B}$, but the opposite is not always true.

A term $\mathscr{E}$ is a *program* if $\mathscr{E}$ is built up from application $(\mathscr{A}\ \mathscr{B})$, abstraction $\lambda x.\mathscr{A}$, selection $(\mathsf{if}\ \mathscr{A}\ \mathscr{B}\ \mathscr{C})$, truth T and bound variables x. If $\mathscr{E}$ is a program and $\mathscr{E} \xrightarrow{*} \mathscr{F}$, then $\mathscr{F}$ is a program.

It is fairly easy [26, 3] to define a mechanical procedure $\mathscr{P}(\mathscr{E})$ which, given a program $\mathscr{E}$, answers *yes* if $\mathscr{E} \xrightarrow{*} \mathsf{T}$, answers *no* if $\dot\neg\mathscr{E} \xrightarrow{*} \mathsf{T}$, and loops indefinitely otherwise. An important property of the model reads: If $\mathscr{E}$ is a program and if neither $\mathscr{E} \xrightarrow{*} \mathsf{T}$ nor $\dot\neg\mathscr{E} \xrightarrow{*} \mathsf{T}$, then $\models \mathscr{E} = \bot$. Hence,

- $\models \mathscr{E} = \mathsf{T}$ iff $\mathscr{P}(\mathscr{E})$ returns *yes*,
- $\models \dot\neg\mathscr{E} = \mathsf{T}$ iff $\mathscr{P}(\mathscr{E})$ returns *no*,
- $\models \mathscr{E} = \bot$ iff $\mathscr{P}(\mathscr{E})$ loops indefinitely.

This gives the intended meaning of $\bot$: $\bot$ stands for infinite looping.

5. Quartum non datur

5.1. Presentation of QND′

The rule *tertium non datur* of classical theories states that a formula is true or false—there is no third possibility. In contrast, Gödel's incompleteness theorem [12] states that a formula may be provable, disprovable or *undecidable*. Classical theories have no truth value to match undecidability.

The rule *quartum non datur* of map theory states that the root of a map is $\tilde{\mathsf{T}}$, $\tilde{\lambda}$ or $\tilde{\bot}$—there is no fourth possibility. Here, $\tilde{\mathsf{T}}$, $\tilde{\lambda}$ and $\tilde{\bot}$ correspond to provability, disprovability and undecidability, or to truth, falsehood and nontermination.

To express quartum non datur (QND') formally, define

$$\mathsf{F}' = \lambda x.\lambda y.(x\,y).$$

For any x, $(\mathsf{F}'\,x)$ is proper and, for any proper x, $x = (\mathsf{F}'\,x)$ according to Lemma 2.3.1. Hence, if

$$[\mathscr{A}/x := (\mathsf{F}'\,x)] = [\mathscr{B}/x := (\mathsf{F}'\,x)],$$

then $\mathscr{A} = \mathscr{B}$ holds for all proper x.

Now, the QND' inference rule is given by

$$(QND') \qquad \text{If} \quad \begin{array}{l} \vdash\ [\mathscr{A}/x := \mathsf{T}] = [\mathscr{B}/x := \mathsf{T}], \\ \vdash\ [\mathscr{A}/x := (\mathsf{F}'\,x)] = [\mathscr{B}/x := (\mathsf{F}'\,x)] \quad \text{and} \\ \vdash\ [\mathscr{A}/x := \bot] = [\mathscr{B}/x := \bot], \end{array}$$

$$\text{then} \ \vdash\ \mathscr{A} = \mathscr{B}.$$

In other words, if $\mathscr{A} = \mathscr{B}$ holds when x is T, $\bot$ or proper, then $\mathscr{A} = \mathscr{B}$ holds for all x.

5.2. Tautologies

Let $\mathscr{A} = \mathscr{B}$ be a formula whose free variables are exactly $x_1, \ldots, x_n$. A *logical instance* of $\mathscr{A} = \mathscr{B}$ is a formula

$$[\mathscr{A}/x_1 := \mathscr{C}_1/\cdots/x_n := \mathscr{C}_n] = [\mathscr{B}/x_1 := \mathscr{C}_1/\cdots/x_n := \mathscr{C}_n],$$

where each $\mathscr{C}_i$ is one of the terms T, $\bot$ or $(\mathsf{F}'\,x_i)$. As can be seen, $\mathscr{A} = \mathscr{B}$ has exactly 3^n logical instances. Repeated application of inference QND' gives the following theorem.

Theorem 5.2.1. *If all logical instances of a formula $\mathscr{A} = \mathscr{B}$ are provable, then $\mathscr{A} = \mathscr{B}$ itself is provable.*

A formula $\mathscr{A} = \mathscr{B}$ is a *tautology* if each logical instance is provable by the reduction theorem. Hence, we obtain the following result.

Theorem 5.2.2. *Each tautology is provable.*

If $\mathscr{A} = \mathscr{B}$ is a tautology with free variables $x_1, \ldots, x_n$, if $\mathscr{C}_1, \ldots, \mathscr{C}_n$ are terms, and $\mathscr{C}_1, \ldots, \mathscr{C}_n$ are free for $x_1, \ldots, x_n$ in $\mathscr{A}$ and $\mathscr{B}$, then the formula

$$[\mathscr{A}/x_1 := \mathscr{C}_1/ \cdots /x_n := \mathscr{C}_n] = [\mathscr{B}/x_1 := \mathscr{C}_1/ \cdots /x_n := \mathscr{C}_n]$$

is an *instance* of $\mathscr{A} = \mathscr{B}$. Since $\mathscr{A} = \mathscr{B}$ is a tautology, $\mathscr{A} = \mathscr{B}$ is provable so, by Theorem 4.2.3,

$$\vdash \ ((\lambda x_1 \ldots \lambda x_n.\mathscr{A})\mathscr{C}_1 \ldots \mathscr{C}_n) = ((\lambda x_1 \ldots \lambda x_n.\mathscr{B})\mathscr{C}_1 \ldots \mathscr{C}_n),$$

which entails

$$\vdash \ [\mathscr{A}/x_1 := \mathscr{C}_1/ \cdots /x_n := \mathscr{C}_n] = [\mathscr{B}/x_1 := \mathscr{C}_1/ \cdots /x_n := \mathscr{C}_n].$$

This proves the following theorem.

Theorem 5.2.3. *Any instance of a tautology is provable.*

For instance, using the above results one immediately proves

$$\vdash \ \mathscr{A} \mathbin{\dot\wedge} \mathscr{B} = \mathscr{B} \mathbin{\dot\wedge} \mathscr{A},$$

$$\vdash \ \mathscr{A} \mathbin{\dot\wedge} \mathscr{A} = \approx \mathscr{A},$$

$$\vdash \ \mathscr{A} \mathbin{\dot\vee} \dot\neg \mathscr{A} = !\mathscr{A},$$

for all terms $\mathscr{A}$ and $\mathscr{B}$.

5.3. Nonmonotonic implication

Whenever a term $\mathscr{A}$ occurs in a position where a formula is expected, $\mathscr{A}$ is shorthand for $\mathscr{A} = \mathsf{T}$. For example, if a line of a proof reads $\dot\neg\mathsf{F}$, then that line states that $\dot\neg\mathsf{F} = \mathsf{T}$.

Equality of map theory is nonmonotonic in the sense that a map f would be nonmonotonic if it satisfies $(f x y) = \mathsf{T} \Leftrightarrow x = y$. To see this, note that $\perp = \perp$, so $(f \perp \perp) = \mathsf{T}$. If f is monotonic, then $\perp \leqslant x$ and $\perp \leqslant y$ implies $\mathsf{T} = (f \perp \perp) \leqslant (f x y)$ which implies $(f x y) = \mathsf{T}$ for all x and y, contradicting $(f x y) = \mathsf{T} \Leftrightarrow x = y$.

Consequently, no map f satisfies $(f x y) = \mathsf{T}$ iff $x = y$ since all maps are monotonic according to (map 11.6.1).

The nonmonotonicity of equality allows to express an implication concept more powerful than $x \mathbin{\dot\Rightarrow} y$ where $x \mathbin{\dot\Rightarrow} y$ is monotonic in x and y like any other term of map theory. To do so, define the *guard* $x{:}y$ by

$$x{:}y = (\mathsf{if}\ x\ y\ \mathsf{T})$$

and define $\mathscr{A} \rightarrow (\mathscr{B} = \mathscr{C})$ to be shorthand for

$$\mathscr{A}{:}\mathscr{B} = \mathscr{A}{:}\mathscr{C}.$$

If $\mathscr{A}$ is proper, then $\mathscr{A}:\mathscr{B} = \mathsf{T} = \mathscr{A}:\mathscr{C}$, and if $\mathscr{A} = \perp$, then $\mathscr{A}:\mathscr{B} = \perp = \mathscr{A}:\mathscr{C}$. If $\mathscr{A} = \mathsf{T}$ then $\mathscr{A}:\mathscr{B} = \mathscr{B}$ and $\mathscr{A}:\mathscr{C} = \mathscr{C}$, so $\mathscr{A}:\mathscr{B} = \mathscr{A}:\mathscr{C}$ iff $\mathscr{B} = \mathscr{C}$. Hence, $\mathscr{A} \to (\mathscr{B} = \mathscr{C})$ expresses the statement "if $\mathscr{A}$ is true, then $\mathscr{B}$ equals $\mathscr{C}$".

The expression $\mathscr{A} \to \mathscr{B}$ is shorthand for $\mathscr{A} \to (\mathscr{B} = \mathsf{T})$ and expresses "if $\mathscr{A}$ is true, then $\mathscr{B}$ is true". This relation is nonmonotonic just like equality. For comparison, $(\mathscr{A} \dot{\Rightarrow} \mathscr{B}) = \mathsf{T}$ expresses "$\mathscr{A}$ and $\mathscr{B}$ both differ from $\perp$, and $\mathscr{A}$ implies $\mathscr{B}$".

The following lemma is easy.

Lemma 5.3.1. *For all terms $\mathscr{A}$, $\mathscr{B}$ and $\mathscr{C}$,*

$$\mathscr{A}\,;\mathscr{A} \to (\mathscr{B} = \mathscr{C}) \;\vdash\; \mathscr{B} = \mathscr{C},$$

$$\mathscr{A}\,;\mathscr{A} \to \mathscr{B} \;\vdash\; \mathscr{B}.$$

In other words, if $\mathscr{A} = \mathsf{T}$ and $\mathscr{A}:\mathscr{B} = \mathscr{A}:\mathsf{T}$ are provable, then so is $\mathscr{B} = \mathsf{T}$.

The formula $\mathscr{A} \to \mathscr{A}$ is a tautology (i.e. $\mathscr{A}:\mathscr{A} = \mathscr{A}:\mathsf{T}$ is a tautology), which implies the next lemma.

Lemma 5.3.2. $\vdash \mathscr{A} \to \mathscr{A}$

If $\vdash \mathscr{B} = \mathscr{C}$ then $\vdash \mathscr{A}:\mathscr{B} = \mathscr{A}:\mathscr{C}$ by Theorem 4.2.3, which proves the following lemma.

Lemma 5.3.3. $\mathscr{B} = \mathscr{C} \vdash \mathscr{A} \to (\mathscr{B} = \mathscr{C})$.

Since $(x:y):z = x:(y:z)$ is a tautology, there is no need to put parentheses in expressions like $\mathscr{A}_1:\mathscr{A}_2:\cdots:\mathscr{A}_n$.

Since the formula $\mathscr{B} \to (\mathscr{C} = \mathscr{D})$ is shorthand for $\mathscr{B}:\mathscr{C} = \mathscr{B}:\mathscr{D}$, $\mathscr{A} \to (\mathscr{B} \to (\mathscr{C} = \mathscr{D}))$ is shorthand for $\mathscr{A}:(\mathscr{B}:\mathscr{C}) = \mathscr{A}:(\mathscr{B}:\mathscr{D})$. In general, $\mathscr{A}_1 \to \cdots \to \mathscr{A}_n \to (\mathscr{C} = \mathscr{D})$ is shorthand for $\mathscr{A}_1:\cdots:\mathscr{A}_n:\mathscr{C} = \mathscr{A}_1:\cdots:\mathscr{A}_n:\mathscr{D}$.

Let $\mathscr{A}_1, \ldots, \mathscr{A}_n \to (\mathscr{C} = \mathscr{D})$ be shorthand for $\mathscr{A}_1 \to \cdots \to \mathscr{A}_n \to (\mathscr{C} = \mathscr{D})$ which in turn is shorthand for $\mathscr{A}_1:\cdots:\mathscr{A}_n:\mathscr{C} = \mathscr{A}_1:\cdots:\mathscr{A}_n:\mathscr{D}$. The following lemma is trivial but nevertheless important.

Lemma 5.3.4. *The formulas*

$$\mathscr{A}_1, \ldots, \mathscr{A}_n, \mathscr{B}_1, \ldots, \mathscr{B}_m \to (\mathscr{C} = \mathscr{D})$$

and

$$\mathscr{A}_1, \ldots, \mathscr{A}_n \to (\mathscr{B}_1, \ldots, \mathscr{B}_m \to (\mathscr{C} = \mathscr{D}))$$

are shorthand for the same term.

As an example, $\mathscr{A}, \mathscr{B}_1, \ldots, \mathscr{B}_n \to (\mathscr{C} = \mathscr{D})$ is the same term as $\mathscr{A} \to (\mathscr{B}_1, \ldots, \mathscr{B}_n \to (\mathscr{C} = \mathscr{D}))$, so, by Lemma 5.3.1

$$\mathscr{A}\,;\mathscr{A}, \mathscr{B}_1, \ldots, \mathscr{B}_m \to (\mathscr{C} = \mathscr{D}) \;\vdash\; \mathscr{B}_1, \ldots, \mathscr{B}_m \to (\mathscr{C} = \mathscr{D}).$$

Note how semicolons separate antecedents of $\vdash$ whereas commas separate antecedents of $\rightarrow$.

Since $\mathscr{A}:\mathscr{A}:\mathscr{C} = \mathscr{A}:\mathscr{C}$ and $\mathscr{A}:\mathscr{B}:\mathscr{C} = \mathscr{B}:\mathscr{A}:\mathscr{C}$ are tautologies, the following holds.

Lemma 5.3.5. *If* $\{\mathscr{A}_1, \ldots, \mathscr{A}_n\}$ *and* $\{\mathscr{B}_1, \ldots, \mathscr{B}_m\}$ *are the same sets of terms, then*

$$\mathscr{A}_1, \ldots, \mathscr{A}_n \rightarrow (\mathscr{C} = \mathscr{D})$$

iff

$$\mathscr{B}_1, \ldots, \mathscr{B}_n \rightarrow (\mathscr{C} = \mathscr{D}).$$

In other words, repetition and ordering among the antecedents of $\rightarrow$ are insignificant.

The following theorem will be used extensively in proofs.

Theorem 5.3.6. (QND). *If*

$$\mathscr{A}_1, \ldots, \mathscr{A}_n \rightarrow (\mathscr{C} = \mathscr{D})$$

is an instance of a tautology, then

$$\mathscr{A}_1; \cdots; \mathscr{A}_n \ \vdash \ \mathscr{B} = \mathscr{C}.$$

The theorem can be used, e.g., to prove

$$\mathscr{A}; \mathscr{A} \Rightarrow \mathscr{B} \ \vdash \ \mathscr{B}.$$

The QND theorem expresses quartum non datur whereas TND below expresses tertium non datur.

Theorem 5.3.7 (TND).

$$!\mathscr{A}; \mathscr{A} \rightarrow (\mathscr{B} = \mathscr{C}); \dot{\neg}\mathscr{A} \rightarrow (\mathscr{B} = \mathscr{C}) \ \vdash \ \mathscr{B} = \mathscr{C}.$$

Since $!\mathscr{A} = \mathsf{T}$ iff $\mathscr{A} \neq \bot$, the first antecedent of TND rules out the possibility that $\mathscr{A} = \bot$. Once $\mathscr{A} = \bot$ is ruled out, tertium non datur holds.

Proof of TND. (See below for an explanation of lines 6 and 7.)

1. hyp	$!\mathscr{A}$
2. hyp	$\mathscr{A} \rightarrow (\mathscr{B} = \mathscr{C})$
3. hyp	$\dot{\neg}\mathscr{A} \rightarrow (\mathscr{B} = \mathscr{C})$
4. 1,QND	$(\mathsf{if}\ \mathscr{A}\ (\mathscr{A}:\mathscr{B})\ ((\dot{\neg}\mathscr{A}):\mathscr{B})) = \mathscr{B}$
5. 1,QND	$(\mathsf{if}\ \mathscr{A}\ (\mathscr{A}:\mathscr{C})\ ((\dot{\neg}\mathscr{A}):\mathscr{C})) = \mathscr{C}$
6. 2,3,triv	$(\mathsf{if}\ \mathscr{A}\ (\mathscr{A}:\mathscr{B})\ ((\dot{\neg}\mathscr{A}):\mathscr{B})) = (\mathsf{if}\ \mathscr{A}\ (\mathscr{A}:\mathscr{C})\ ((\dot{\neg}\mathscr{A}):\mathscr{C}))$
7. 4,5,6,triv	$\mathscr{B} = \mathscr{C}$. $\square$

In the proof, line 7 follows "trivially" from lines 4, 5 and 6. An equation $\mathscr{C} = \mathscr{D}$ follows "trivially" from $\mathscr{A}_1 = \mathscr{B}_1, \ldots, \mathscr{A}_n = \mathscr{B}_n$ if

(1) $\mathscr{C} = \mathscr{D}$ is deducible from $\mathscr{A}_1 = \mathscr{B}_1, \ldots, \mathscr{A}_n = \mathscr{B}_n$ using only the inferences trans, sub1 and sub2, and the axioms Select 1-3, Apply 1-3, and Rename;

(2) it is expected to be easy for the reader to fill in the missing details.

From this point on there will be few references to the metatheorems on equality and reduction since they are "trivial".

5.4. The deduction theorem

A formula $\mathscr{A} = \mathscr{B}$ follows *hypothetically* from the formulas $\mathscr{A}_1 = \mathscr{B}_1, \ldots, \mathscr{A}_n = \mathscr{B}_n$ if there is a sequence $\mathscr{C}_1 = \mathscr{D}_1, \ldots, \mathscr{C}_m = \mathscr{D}_m$ of formulas such that

(a) $\mathscr{A} = \mathscr{B}$ is the last formula in the sequence;

(b) Each formula of the sequence

 (1) is an axiom of map theory,

 (2) is one of the *hypotheses* $\mathscr{A}_1 = \mathscr{B}_1, \ldots, \mathscr{A}_n = \mathscr{B}_n$, or

 (3) follows from previous formulas in the sequence by one of the inference rules of map theory;

(c) The sub2 inference is never used to verify $\lambda x.\mathscr{C}_i = \lambda x.\mathscr{D}_i$ from $\mathscr{C}_i = \mathscr{D}_i$ if the latter depends on a hypothesis $\mathscr{A}_j = \mathscr{B}_j$ in which x occurs free.

The statement $\mathscr{A}_1 = \mathscr{B}_1; \cdots; \mathscr{A}_n = \mathscr{B}_n \overset{h}{\vdash} \mathscr{A} = \mathscr{B}$ states that $\mathscr{A} = \mathscr{B}$ follows hypothetically from $\mathscr{A}_1 = \mathscr{B}_1, \ldots, \mathscr{A}_n = \mathscr{B}_n$.

Theorem 5.4.1 (Deduction). *If*

$$\mathscr{H} \overset{h}{\vdash} \mathscr{A} = \mathscr{B}$$

then

$$\mathscr{H} \rightarrow (\mathscr{A} = \mathscr{B}).$$

The deduction theorem is close in spirit to the deduction theorem of first-order predicate calculus [22]. In particular, that deduction theorem has a requirement similar to (c) above.

The deduction theorem validates proofs of the form "Assume that $\mathscr{H}$ is true. Given that $\mathscr{H}$ is true, $\mathscr{A} = \mathscr{B}$ holds. Hence, $\mathscr{H} \rightarrow (\mathscr{A} = \mathscr{B})$". This will be used later in this paper when formal deductions are replaced by a more conversational style of proofs.

From the deduction theorem the following corollary can be obtained.

Corollary 5.4.2. *If*

$$\mathscr{H}_1; \cdots; \mathscr{H}_n \overset{h}{\vdash} \mathscr{A} = \mathscr{B}$$

then

$$\mathscr{H}_1, \ldots, \mathscr{H}_n \rightarrow \mathscr{A} = \mathscr{B}.$$

There are two different ways to prove the corollary. One way is to apply the deduction theorem n times. Another way is to apply the deduction theorem once with hypothesis $\mathcal{H}_1 : \cdots : \mathcal{H}_n$. If the goal is actually to construct a proof of $\mathcal{H}_1, \ldots, \mathcal{H}_n \to \mathcal{A} = \mathcal{B}$, the second approach is best since it produces a considerably shorter proof. This may be critical in computer-based proof systems.

Proof of the deduction theorem. Map theory has five inference rules: trans, sub1, sub2, QND' and ind, where ind will be introduced in Section 7.4. The meaning of the ind inference is unimportant for now, but its syntax is needed for the proof of the deduction theorem. It says

$$\mathcal{A}, x \to \mathcal{B} ; \mathcal{A}, \dot{\neg} x, \phi x, \dot{\forall} y.[\mathcal{B}/x := (x\, y)] \to \mathcal{B} \ \vdash\ \mathcal{A}, \phi x \to \mathcal{B}$$

if x is not free in $\mathcal{A}$ and y does not occur (free or bound) in $\mathcal{B}$.

Seven auxiliary lemmas are:

(trans′) $\quad \mathcal{H} \to (\mathcal{A} = \mathcal{B}) ; \mathcal{H} \to (\mathcal{A} = \mathcal{C}) \ \vdash\ \mathcal{H} \to (\mathcal{B} = \mathcal{C})$.

(sub1′) $\quad \mathcal{H} \to (\mathcal{A} = \mathcal{B}) ; \mathcal{H} \to (\mathcal{C} = \mathcal{D}) \ \vdash\ \mathcal{H} \to ((\mathcal{A}\, \mathcal{C}) = (\mathcal{B}\, \mathcal{D}))$.

(sub2′) $\quad \mathcal{H} \to (\mathcal{A} = \mathcal{B}) \ \vdash\ \mathcal{H} \to (\lambda x.\mathcal{A} = \lambda x.\mathcal{B})$ if x is not free in $\mathcal{H}$.

(QND'') $\quad \mathcal{H} \to ([\mathcal{A}/x := \mathsf{T}] = [\mathcal{B}/x := \mathsf{T}])$;

$\mathcal{H} \to ([\mathcal{A}/x := \mathsf{F}'(x)] = [\mathcal{B}/x := \mathsf{F}'(x)])$;

$\mathcal{H} \to ([\mathcal{A}/x := \bot] = [\mathcal{B}/x := \bot])$

$\vdash\ \mathcal{H} \to (\mathcal{A} = \mathcal{B})$.

(ind′) $\quad \mathcal{H}, \mathcal{A}, x \to \mathcal{B} ; \mathcal{H}, \mathcal{A}, \dot{\neg} x, \phi x, \dot{\forall} y.[\mathcal{B}/x := (x\, y)] \to \mathcal{B} \ \vdash\ \mathcal{H}, \mathcal{A} \to \mathcal{B}$.

(axiom′) $\quad \mathcal{H} \ \vdash\ (\mathcal{A} = \mathcal{B})$ if $\mathcal{A} = \mathcal{B}$ is an axiom of map theory.

(hyp′) $\quad \mathcal{H} \ \vdash\ \mathcal{H}$.

Proof of trans′. The trans′ statement holds since it is a special case of the trans inference.

Proof of sub1′.

1. hyp	$\mathcal{H}:\mathcal{A} = \mathcal{H}:\mathcal{B}$
2. hyp	$\mathcal{H}:\mathcal{C} = \mathcal{H}:\mathcal{D}$
3. QND	$\mathcal{H}:(\mathcal{A}\, \mathcal{C}) = \mathcal{H}:((\mathcal{H}:\mathcal{A})\, (\mathcal{H}:\mathcal{C}))$
4. QND	$\mathcal{H}:(\mathcal{B}\, \mathcal{D}) = \mathcal{H}:((\mathcal{H}:\mathcal{B})\, (\mathcal{H}:\mathcal{D}))$
5. 1,2,triv	$\mathcal{H}:((\mathcal{H}:\mathcal{A})\, (\mathcal{H}:\mathcal{C})) = \mathcal{H}:((\mathcal{H}:\mathcal{B})\, (\mathcal{H}:\mathcal{D}))$
6. 3,4,5,triv	$\mathcal{H}:(\mathcal{A}\, \mathcal{C}) = \mathcal{H}:(\mathcal{B}\, \mathcal{D})$.

Proof of sub2′.

1. hyp	$\mathcal{H}:\mathcal{A} = \mathcal{H}:\mathcal{B}$
2. *QND*	$\mathcal{H}:\lambda x.\mathcal{A} = \mathcal{H}:\lambda x.\mathcal{H}:\mathcal{A}$
3. *QND*	$\mathcal{H}:\lambda x.\mathcal{B} = \mathcal{H}:\lambda x.\mathcal{H}:\mathcal{B}$
4. 1,2,3,triv	$\mathcal{H}:\lambda x.\mathcal{A} = \mathcal{H}:\lambda x.\mathcal{H}:\mathcal{B}$.

Note that line 2 is an instance of the tautology $y:\lambda x.z = y:\lambda x.y:z$ because x does not occur free in $\mathcal{H}$. The same is true for line 3.

The proof of QND'' is left as an exercise. For ind′, replace $\mathcal{A}$ by $\mathcal{H}:\mathcal{A}$ in the ind inference to obtain ind′ (the sole purpose of having $\mathcal{A}$ in the ind inference is to prove the deduction theorem). The axiom′ statement follows from $\mathcal{A} = \mathcal{B}$ as mentioned in Lemma 5.3.3. The hyp′ statement is an instance of a tautology as mentioned in Lemma 5.3.2.

Now assume that $\mathcal{C}_1 = \mathcal{D}_1, \ldots, \mathcal{C}_m = \mathcal{D}_m$ deduces $\mathcal{A} = \mathcal{B}$ hypothetically from $\mathcal{H} = \mathsf{T}$. By induction in i and using trans′, sub1′, sub2′, QND'', ind′, axiom′ and hyp′,

$$\mathcal{H} \rightarrow (\mathcal{C}_i = \mathcal{D}_i).$$

In particular, for $i = m$,

$$\mathcal{H} \rightarrow (\mathcal{A} = \mathcal{B}).$$

This proves the deduction theorem. □

The following theorem justifies the method of indirect proof.

Theorem 5.4.3 (Contra). $!\mathcal{A},\ \dot{\neg}\mathcal{A} \rightarrow \mathsf{F} \vdash \mathcal{A}$.

Proof.

1. hyp	$!\mathcal{A}$
2. hyp	$(\dot{\neg}\mathcal{A}):\mathsf{F} = (\dot{\neg}\mathcal{A}):\mathsf{T}$
3. *QND*	$\approx\mathcal{A} = (\dot{\neg}\mathcal{A}):\mathsf{F}$
4. 2,3,triv	$\approx\mathcal{A} = (\dot{\neg}\mathcal{A}):\mathsf{T}$
5. 1,*QND*	$(\dot{\neg}\mathcal{A}):\mathsf{T} = \mathsf{T}$
6. 4,5,triv	$\approx\mathcal{A} = \mathsf{T}$
7. 6,*QND*	$\mathcal{A}$. □

The following theorem can be used in connection with the deduction theorem.

Theorem 5.4.4. (Monotonic deduction).

$$!\mathcal{A}_1; \cdots; !\mathcal{A}_n; !\mathcal{B}; \mathcal{A}_1, \ldots, \mathcal{A}_n \rightarrow \mathcal{B} \ \vdash\ \mathcal{A}_1 \dot{\wedge} \cdots \dot{\wedge} \mathcal{A}_n \dot{\Rightarrow} \mathcal{B}.$$

Proof. (See below for an explanation of line 1.)

1. hyp	$!\mathscr{A}_1; \cdots; !\mathscr{A}_n$
2. hyp	$!\mathscr{B}$
3. hyp	$\mathscr{A}_1, \ldots, \mathscr{A}_n \to \mathscr{B}$
4. 1,*QND*	$\mathscr{A}_1 \dot{\wedge} \cdots \dot{\wedge} \mathscr{A}_n \dot{\Rightarrow} (\mathscr{A}_1 : \cdots : \mathscr{A}_n : \mathsf{T})$
5. 3,4,triv	$\mathscr{A}_1 \dot{\wedge} \cdots \dot{\wedge} \mathscr{A}_n \dot{\Rightarrow} (\mathscr{A}_1 : \cdots : \mathscr{A}_n : \mathscr{B})$
6. 2,*QND*	$\mathscr{A}_1 \dot{\wedge} \cdots \dot{\wedge} \mathscr{A}_n \dot{\Rightarrow} (\mathscr{A}_1 : \cdots : \mathscr{A}_n : \mathscr{B}) = \mathscr{A}_1 \dot{\wedge} \cdots \dot{\wedge} \mathscr{A}_n \dot{\Rightarrow} \mathscr{B}$
7. 5,6,triv	$\mathscr{A}_1 \dot{\wedge} \cdots \dot{\wedge} \mathscr{A}_n \dot{\Rightarrow} \mathscr{B}$. $\square$

In a proof line,

$$\mathscr{A}_1 = \mathscr{B}_1; \cdots; \mathscr{A}_n = \mathscr{B}_n$$

is shorthand for n proof lines that are all verified the same way. Likewise,

$$\mathscr{A}_1 = \mathscr{B}_1; \cdots; \mathscr{A}_m = \mathscr{B}_m \;\vdash\; \mathscr{C}_1 = \mathscr{D}_1; \cdots; \mathscr{C}_n = \mathscr{D}_n$$

is shorthand for n statements of the form

$$\mathscr{A}_1 = \mathscr{B}_1; \cdots; \mathscr{A}_m = \mathscr{B}_m \;\vdash\; \mathscr{C}_i = \mathscr{D}_i$$

for $i \in \{1, \ldots, n\}$.

6. Quantification

6.1. Quantification axioms

If $\phi\mathscr{A} = \mathsf{T}$ and $(\dot{\forall} x.\mathscr{B}) = \mathsf{T}$, then $\mathscr{A}$ is well-founded and $\mathscr{B} = \mathsf{T}$ for all well-founded x. In particular, $\mathscr{B} = \mathsf{T}$ for $x = \mathscr{A}$, so $((\lambda x.\mathscr{B})\ \mathscr{A}) = \mathsf{T}$. Hence,

(Quantify1) $\quad \vdash\ \phi\mathscr{A}, \dot{\forall} x.\mathscr{B} \to ((\lambda x.\mathscr{B})\ \mathscr{A})$.

For all well-founded x, $\phi x = \mathsf{T}$, so $r(\mathscr{A}) = r(\phi x \dot{\wedge} \mathscr{A})$. Hence, by Ackermanns axiom (4) (Section 2.5),

(Quantify2) $\quad \vdash\ \varepsilon x.\mathscr{A} = \varepsilon x.(\phi x \dot{\wedge} \mathscr{A})$.

The next axiom expresses (1) and (2) of Section 2.5:

(Quantify3) $\quad \vdash\ \phi\varepsilon x.\mathscr{A} = \dot{\forall} x.!\mathscr{A}$.

A verification requires two cases.

- If $\forall x \in \Phi$: $(\mathscr{A} \neq \bot)$ then $\forall x \in \Phi$: $(!\mathscr{A} = \mathsf{T})$ so $(\dot{\forall} x.!\mathscr{A}) = \mathsf{T}$. Further, according to (2), $\varepsilon x.\mathscr{A} \in \Phi$ so $\phi\varepsilon x.\mathscr{A} = \mathsf{T}$. Hence, $\phi\varepsilon x.\mathscr{A} = \mathsf{T} = \dot{\forall} x.!\mathscr{A}$.
- If $\exists x \in \Phi$: $(\mathscr{A} = \bot)$ then $\exists x \in \Phi$: $(!\mathscr{A} = \bot)$ so $(\dot{\forall} x.!\mathscr{A}) = \bot$. Further, according to (1), $\varepsilon x.\mathscr{A} = \bot$. As mentioned in Section 2.6, $\bot \notin \Phi$, so $\phi\varepsilon x.\mathscr{A} = \bot = \dot{\forall} x.!\mathscr{A}$.

The axioms Quantify 4 and 5 follow directly from Quantify 3 and 2, respectively, using the monotonicity of maps. It would be more satisfactory to express the monotonicity directly as an axiom and omit Quantify 4 and 5. A formulation of monotonicity is complicated, but should be included in next iteration of map theory.

(Quantify4) $\quad \vdash \dot{\exists} x.\mathscr{A} \rightarrow \phi \varepsilon x.\mathscr{A}.$

(Quantify5) $\quad \vdash \dot{\forall} x.\mathscr{A} = \dot{\forall} x.(\phi x \mathbin{\dot{\wedge}} \mathscr{A}).$

6.2. Metatheorems of quantification

A direct consequence of the Quantify2 axiom is the following theorem.

Theorem 6.2.1. *If $\mathscr{A}$ is free for x in $\mathscr{B}$, then*

$$\phi.\mathscr{A}; \dot{\forall} x.\mathscr{B} \vdash [\mathscr{B}/x := \mathscr{A}].$$

Section 8 states proofs in a conversational style as opposed to the more formal proofs that consist of numbered proof lines. A typical construct in conversational proofs reads "if x is well-founded, then $\mathscr{A}$ is true—hence, $\dot{\forall} x.\mathscr{A}$ holds". The theorem below justifies this conversational construct.

Theorem 6.2.2. $\phi x \rightarrow \mathscr{A} \vdash \dot{\forall} x.\mathscr{A}.$

Proof.

1. hyp	$\phi x : \mathscr{A} = \phi x : \mathsf{T}$
2. *QND*	$\phi x \mathbin{\dot{\wedge}} \mathscr{A} = \phi x \mathbin{\dot{\wedge}} (\phi x : \mathscr{A})$
3. *QND*	$\phi x \mathbin{\dot{\wedge}} \mathsf{T} = \phi x \mathbin{\dot{\wedge}} (\phi x : \mathsf{T})$
4. 1,2,3,triv	$\phi x \mathbin{\dot{\wedge}} \mathscr{A} = \phi x \mathbin{\dot{\wedge}} \mathsf{T}$
5. 4,triv	$\dot{\forall} x.\phi x \mathbin{\dot{\wedge}} \mathscr{A} = \dot{\forall} x.\phi x \mathbin{\dot{\wedge}} \mathsf{T}$
6. 5,triv,Quantify5	$\dot{\forall} x.\mathscr{A} = \dot{\forall} x.\mathsf{T}$
7. 6,triv	$\dot{\forall} x.\mathscr{A}.$

(In line 7, note that $\dot{\forall} x.\mathsf{T} = \mathsf{T}$ follows trivially from the definition of $\dot{\forall}$ and $(\mathsf{T}\, x) = \mathsf{T}$). □

Another conversational construct reads "$\mathscr{A}$ is well-founded, $\dot{\exists} x.\mathscr{B} \neq \bot$, and $\mathscr{B}$ holds for $x = \mathscr{A}$, so $\dot{\exists} x.\mathscr{A} = \mathsf{T}$" (Chapter 8.3 considers conversational proofs systematically). In many situations in conversational proofs, it will be obvious that $\mathscr{A}$ is well-founded and $\dot{\exists} x.\mathscr{B} \neq \bot$, in which case the construct reduces to "$\mathscr{B}$ holds for $x = \mathscr{A}$, so $\dot{\exists} x.\mathscr{A}$". The construct is justified by the following theorem.

Theorem 6.2.3. *If $\mathscr{A}$ is free for x in $\mathscr{B}$, then*

$$\phi\mathscr{A}\,;\,!\dot{\exists}x.\mathscr{B}\,;\,[\mathscr{B}/x := \mathscr{A}] \vdash \dot{\exists}x.\mathscr{A}.$$

Proof.

1. hyp	$\phi\mathscr{A}$
2. hyp	$!\dot{\exists}x.\mathscr{B}$
3. hyp	$[\mathscr{B}/x := \mathscr{A}]$
4. triv,Quantify1	$\phi\mathscr{A},\ \dot{\neg}\dot{\exists}x.\dot{\neg}\dot{\neg}\mathscr{B} \to \dot{\neg}[\mathscr{B}/x := \mathscr{A}]$
5. 1,3,4,triv	$\dot{\neg}\dot{\exists}x.\mathscr{B} \to \mathsf{F}$
6. 2,5,indir.pf.	$\dot{\exists}x.\mathscr{B}.$ □

In conversational proofs, if $\dot{\exists}x.\mathscr{R}(x)$ has been proved, the construct "let u satisfy $\mathscr{R}(u)$" means "let u be shorthand for $\varepsilon x.\mathscr{R}(x)$ within this proof or until u is redefined". Further, "let u satisfy $\mathscr{R}(u)$" tacitly establishes the fact that $\mathscr{R}(u)$ actually holds for $u = \varepsilon x.\mathscr{R}(x)$ and that u is well-founded as justified by the next theorem.

Theorem 6.2.4. *If u is shorthand for $\varepsilon x.\mathscr{A}$ and u is free for x in $\mathscr{A}$, then*

$$\dot{\exists}x.\mathscr{A} \vdash \phi u\,;\,[\mathscr{A}/x := u].$$

Proof.

1. hyp	$\dot{\exists}x.\mathscr{A}$
2. 1,triv,Quantify4	$\phi\varepsilon x.\mathscr{A}$
3. 1,triv	$\approx[\mathscr{A}/x := \varepsilon x.\mathscr{A}]$
4. 3,*QND*	$[\mathscr{A}/x := \varepsilon x.\mathscr{A}].$ □

The construct "let u satisfy $\mathscr{R}(u)$" becomes clumsy when $\mathscr{R}(u)$ is a large, well-formed formula. In particular, it is clumsy first to state $\dot{\exists}x.\mathscr{R}(x)$ and then "let u satisfy $\mathscr{R}(u)$" for large $\mathscr{R}$ since $\mathscr{R}$ is stated twice in a row. In such situations, "let u be such an x" is shorthand for "let u satisfy $\mathscr{R}(u)$" since it is obvious what $\mathscr{R}$ is. Further, "pick x" is shorthand for "let all free occurrences of x be shorthand for $\varepsilon x.\mathscr{R}(x)$ within this proof or until x is redefined" when it is obvious what $\mathscr{R}$ is.

The Quantify2 axiom was justified from Ackermann's axiom (4). The following theorem proves Ackermanns axiom from the Quantify2 axiom.

Theorem 6.2.5 (Ackermann). $\phi x \to \mathscr{A} \Leftrightarrow \mathscr{B} \vdash \varepsilon x.\mathscr{A} = \varepsilon x.\mathscr{B}$.

Proof.

1. hyp	$\phi x \to \mathscr{A} \Leftrightarrow \mathscr{B}$
2. 1,triv	$\phi x \dot\wedge \mathscr{A} \dot\wedge (\phi x{:}(\mathscr{A} \Leftrightarrow \mathscr{B})) = \phi x \dot\wedge \mathscr{A} \dot\wedge (\phi x{:}\mathsf{T})$
3. 1,triv	$\phi x \dot\wedge \mathscr{B} \dot\wedge (\phi x{:}(\mathscr{A} \Leftrightarrow \mathscr{B})) = \phi x \dot\wedge \mathscr{B} \dot\wedge (\phi x{:}\mathsf{T})$
4. *QND*	$\phi x \dot\wedge \mathscr{A} = \phi x \dot\wedge \mathscr{A} \dot\wedge (\phi x{:}\mathsf{T})$
5. *QND*	$\phi x \dot\wedge \mathscr{B} = \phi x \dot\wedge \mathscr{B} \dot\wedge (\phi x{:}\mathsf{T})$
6. *QND*	$\phi x \dot\wedge \mathscr{A} \dot\wedge (\phi x{:}(\mathscr{A} \Leftrightarrow \mathscr{B})) = \phi x \dot\wedge \mathscr{B} \dot\wedge (\phi x{:}(\mathscr{A} \Leftrightarrow \mathscr{B}))$
7. 2,3,4,5,6,triv	$\phi x \dot\wedge \mathscr{A} = \phi x \dot\wedge \mathscr{B}$
8. 7,triv	$\varepsilon x.\phi x \dot\wedge \mathscr{A} = \varepsilon x.\phi x \dot\wedge \mathscr{B}$
9. 8,triv,Quantify2	$\varepsilon x.\mathscr{A} = \varepsilon x.\mathscr{B}$. $\square$

7. Well-foundedness

7.1. Properties of well-foundedness

As mentioned in Section 2.4, Φ is the least set such that

$$\mathsf{T} \in \Phi,$$
$$g \in \Phi \wedge \forall x \in g^{s\circ}{:}\ (fx) \in \Phi \Rightarrow f \in \Phi.$$

Here, g^s is the set of well-founded maps introduced before g, and G° is defined by

$$G^\circ = \{f \in M \mid \forall x_1, x_2, \ldots \in G\ \exists n \geq 0{:}\ (fx_1 \ldots x_n) = \mathsf{T}\}.$$

In particular, $\emptyset^\circ$ is the set of all maps except $\perp$.

Let $f <_s g$ denote that f is introduced before g. The relation $<_s$ is a well-order and $f \in g^s \Leftrightarrow f <_s g$. Since $f <_s g \wedge g <_s h \Rightarrow f <_s h$ we have $g <_s h \Rightarrow g^s \subseteq h^s$.

From the definition of G° one easily verifies $G \subseteq H \Rightarrow H^\circ \subseteq G^\circ$. Hence, if $f <_s g$ then $g^{s\circ} \subseteq f^{s\circ}$.

When a proper well-founded map f is introduced, it is introduced by verifying $\forall x \in g^{s\circ}{:}\ (fx) \in \Phi$ for some $g \in \Phi$. To verify this, g and (fx) have to be introduced before f. Hence, for all proper, well-founded f,

$$\exists g \in f^s\ \forall x \in g^{s\circ}{:}\ (fx) \in f^s.$$

In particular,

$$\exists g \in f^s\ \forall x \in g^{s\circ}{:}\ (fx) \in \Phi.$$

The latter statement happens to hold also for $f = \mathsf{T}$, so it holds for all well-founded f.

If $g \in f^s$ then $g <_s f$ and $f^{s\circ} \subseteq g^{s\circ}$, so

$$\forall x \in f^{s\circ}{:}\ (fx) \in f^s.$$

If $f <_s h$ then $f^s \subseteq h^s$ and $h^{s\circ} \subseteq g^{s\circ}$, so

$$\forall f \in h^s\ \forall x \in h^{s\circ}{:}\ \ (fx) \in h^s.$$

Lemma 7.1.1. *If a, $b \in \Phi$ then $\lambda x.(\mathsf{if}\ x\ a\ b) \in \Phi$.*

Proof. If $x \in \mathsf{T}^{so} = \emptyset^{\circ}$ then $x \neq \perp$ and $((\lambda x.(\mathsf{if}\ x\ a\ b))\ x) \in \{a, b\} \subseteq \Phi$. Hence, $\forall x \in \mathsf{T}^{so}$: $((\lambda x.(\mathsf{if}\ x\ a\ b))\ x) \in \Phi$, so $\lambda x.(\mathsf{if}\ x\ a\ b) \in \Phi$. □

Lemma 7.1.2. *If a, $b \in \Phi$ then there is a $c \in \Phi$ such that $a <_s c$ and $b <_s c$.*

Proof. Let $c = \lambda x.(\mathsf{if}\ x\ a\ b)$. □

Define $\langle u, v\rangle <_{s'} \langle x, y\rangle \Leftrightarrow u <_s x \wedge v = y \vee v <_s x \wedge u = y$. The relation $<_{s'}$ is easily shown to be well-founded.

Lemma 7.1.3. *If f, $g \in \Phi$ and $g^s \subseteq a^{so}$ and $a^s \subseteq g^{so}$ for all $a \in f^s$, then $f^s \subseteq g^{so}$.*

Proof. Assume $a \in f^s$. From $\forall x \in a^{so}$: $(a\ x) \in a^s$, $g^s \subseteq a^{so}$, and $a^s \subseteq g^{so}$ we have $\forall x \in g^s$: $(a\ x) \in g^{so}$ which entails $a \in g^{so}$ (by the definition of G°). Hence, $a \in f^s \Rightarrow a \in g^{so}$. □

Lemma 7.1.4. *If f, $g \in \Phi$ then $f^s \subseteq g^{so}$.*

Proof. Follows from the previous lemma by transfinite induction on $<_{s'}$. □

Lemma 7.1.5. *If f, $g \in \Phi$ then $f \in g^{so}$.*

Proof. Choose $h \in \Phi$ such that $f <_s h$. Now $f \in h^s \subseteq g^{so}$. □

Lemma 7.1.6. *If f, $g \in \Phi$ then $(f\,g) \in \Phi$.*

Proof. Follows from $g \in f^{so}$ and $\forall x \in f^{so}$: $(f\,x) \in \Phi$. □

Corollary 7.1.7. *If f, $g \in \Phi$ and $f \neq \mathsf{T}$ then $(f\,g) <_s f$.*

Lemma 7.1.8. $\Phi \subseteq \Phi^{\circ}$.

Proof. Let $f, x_1, x_2, \ldots \in \Phi$ and let $g_n = (f\,x_1 \ldots x_n)$. For all $n \geqslant 0$ we have $g_n \in \Phi$ and $g_n <_s g_{n+1} \vee g_n = \mathsf{T}$. Since $<_s$ is well-founded, $g_n <_s g_{n+1}$ cannot hold for all n, so $g_n = \mathsf{T}$ must hold for some n. Hence, $\forall x_1, x_2, \ldots \in \Phi\ \exists n \geqslant 0$: $(f\,x_1 \ldots x_n) = \mathsf{T}$ which proves $f \in \Phi^{\circ}$. From $f \in \Phi \Rightarrow f \in \Phi^{\circ}$ we have $\Phi \subseteq \Phi^{\circ}$. □

7.2. Well-foundedness axioms

Well-foundedness is described by ten axiom schemes and one inference rule in map theory. None of these express the intuition behind well-foundedness. It would be more satisfactory to have a single axiom that expresses the intuition. The ten axiom schemes and the inference rule correspond to some extent to the union and power set axioms, etc., of *ZFC* that point out that certain sets exist.

In what follows, the ten axioms and the inference rule are presented, and they are proved to follow from the intuitive notion of well-foundedness stated in Section 2.4 and repeated in Section 7.1.

Three axioms describe some elementary properties of Φ like Apply 1–3 described functional application and Select 1–3 described selection.

(Well1) $\phi\mathsf{T}$.

(Well2) $\phi\lambda x.\mathscr{A} = \phi\lambda x.\phi\mathscr{A}$.

(Well3) $\phi\bot = \bot$.

The Well1 axiom is shorthand for $\phi\mathsf{T} = \mathsf{T}$. It follows directly from $\mathsf{T} \in \Phi$. Since $\bot \notin G^{\circ}$ for all sets G of maps, we have $\bot \notin \Phi$ which verifies Well3. The Well2 axiom is more complicated.

Lemma 7.2.1. *If* $\lambda x.\mathscr{A} \in \Phi$ *then* $\lambda x.\phi\mathscr{A} \in \Phi$.

Proof. Choose $g \in \Phi$ such that $\forall x \in g^{s\circ}$: $((\lambda x.\mathscr{A})\ x) \in \Phi$. For all $x \in g^{s\circ}$, $\mathscr{A} \in \Phi$ so $\phi\mathscr{A} = \mathsf{T} \in \Phi$. Hence, $\forall x \in g^{s\circ}$: $((\lambda x.\phi\mathscr{A})\ x) \in \Phi$ so $\lambda x.\phi\mathscr{A} \in \Phi$. □

Lemma 7.2.2. *If* $\lambda x.\phi\mathscr{A} \in \Phi$ *then* $\lambda x.\mathscr{A} \in \Phi$.

Proof. Choose $g \in \Phi$ such that $\forall x \in g^{s\circ}$: $((\lambda x.\phi\mathscr{A})\ x) \in \Phi$. For all $x \in g^{s\circ}$, $\phi\mathscr{A} \in \Phi$ so $\phi\mathscr{A} \neq \bot$ and $\mathscr{A} \in \Phi$. Hence, $\forall x \in g^{s\circ}$: $((\lambda x.\mathscr{A})\ x) \in \Phi$ so $\lambda x.\mathscr{A} \in \Phi$. □

Lemma 7.2.3. $\phi\lambda x.\mathscr{A} = \phi\lambda x.\phi\mathscr{A}$.

Proof. If $\lambda x.\mathscr{A} \in \Phi$ then $\lambda x.\phi\mathscr{A} \in \Phi$ according to Lemma 7.2.1 so $\phi\lambda x.\mathscr{A} = \mathsf{T} = \lambda x.\phi\mathscr{A}$. If $\lambda x.\mathscr{A} \notin \Phi$ then $\lambda x.\phi\mathscr{A} \notin \Phi$ according to Lemma 7.2.2 so $\phi\lambda x.\mathscr{A} = \bot = \phi\lambda x.\phi\mathscr{A}$. □

7.3. Construction axioms

Let $\phi x.\mathscr{A}$ be shorthand for $\phi\lambda x.\mathscr{A}$ and define

$$P = \lambda a.\lambda b.\lambda x.(\mathsf{if}\ x\ a\ b),$$

$$Curry = \lambda a.\lambda x.\lambda y.(a\ (P\,x\,y)),$$

$$Prim = \lambda f.\lambda a.\lambda b.\mathsf{Y} g.\lambda x.(\mathsf{if}\ x\ a\ (f\,\lambda u.(g\ (x\ (b\,u))))).$$

The definition of P (*Pair*) is identical to the definition of *cons* in Section 4.3, and $(P\,a\,b)$ equals $a::b$ defined in Section 3.4.

For any map a, $(Curry\ a)$ expresses the inverse of *Currying* of a [7].

If f, a and b are maps, then $g = (Prim\,f\,a\,b)$ satisfies the primitive recursive definition

$$(g\,x) = \begin{cases} a & \text{if } x = \mathsf{T}, \\ \bot & \text{if } x = \bot, \\ (f\,\lambda u.(g\ (x\ (b\,u)))) & \text{otherwise.} \end{cases}$$

The construction axioms are suited to prove the well-foundedness of a wide range of maps. They are as follows.

(C-A) $\phi a, \phi b \rightarrow \phi(a\,b)$.

(C-K′) $\phi x.\mathsf{T}$.

(C-P′) $\phi x.(\mathsf{if}\ x\ \mathsf{T}\ \mathsf{T})$.

(C-Curry) $\phi a \rightarrow \phi(\mathit{Curry}\ a)$.

(C-Prim) $\forall x.\phi(f\,x), \phi a, \phi b \rightarrow \phi(\mathit{Prim}\ f\,a\,b)$.

(C-M1) $\forall u.\phi x.\mathscr{A} \rightarrow \forall u.\phi x.((\lambda u.\mathscr{A})\ (u\,x))$.

(C-M2) $\forall u.\phi x.\mathscr{A} \rightarrow \forall u.\phi x.((\lambda x.\mathscr{A})\ (x\,u))$.

The remainder of this chapter verifies the construction axioms from the intuition behind well-foundedness.

Lemma 7.1.6 is a verification of the C-A axiom. From $\forall x \in \mathsf{T}^{\mathrm{so}}$: $((\lambda x.\mathsf{T})\ x) = \mathsf{T} \in \Phi$ we have $\lambda x.\mathsf{T} \in \Phi$ which verifies the C-K′ axiom. Lemma 7.1.1 verifies the C-P′ axiom.

Lemma 7.3.1 (The C-Curry axiom). *If* $a \in \Phi$ *then* $\lambda x.\lambda y.(a\ (P\,x\,y)) \in \Phi$.

Proof. Choose $g \in \Phi$ such that $\forall x \in g^{\mathrm{so}}$: $(a\,x) \in \Phi$. Assume $x, y \in g^{\mathrm{so}}$. Now $\forall z_1, \ldots \in g^{\mathrm{s}}\ \exists n$: $(x\,z_1 \ldots z_n) = \mathsf{T}$. If $z \in g^{\mathrm{s}}$ then $z \neq \perp$ so $(P\,x\,y\,z) = x \vee (P\,x\,y\,z) = y$. Hence, $\forall z, z_1, z_2, \ldots \in g^{\mathrm{s}}\ \exists n$: $(P\,x\,y\,z\,z_1 \ldots z_n) = \mathsf{T}$ which proves $(P\,x\,y) \in g^{\mathrm{so}}$ so $(a\ (P\,x\,y)) \in \Phi$. From $\forall y \in g^{\mathrm{so}}$: $(a\ (P\,x\,y)) \in \Phi$ we have $\lambda y.(a\ (P\,x\,y)) \in \Phi$, and from $\forall x \in g^{\mathrm{so}}$: $\lambda y.(a\ (P\,x\,y)) \in \Phi$ we have $\lambda x.\lambda y.(a\ (P\,x\,y)) \in \Phi$. $\square$

Lemma 7.3.2. *If* $z \in \Phi$ *then* $\lambda x.(a\,x\,x) \in \Phi$.

Proof. If $x \in a^{\mathrm{so}}$ then $(a\,x) \in a^{\mathrm{s}}$ and $(a\,x\,x) = ((a\,x)\ x) \in a^{\mathrm{s}}$ so $\forall x \in a^{\mathrm{so}}$: $((\lambda x.(a\,x\,x))\ x) \in a^{\mathrm{s}}$ and $\lambda x.(a\,x\,x) \in \Phi$. $\square$

Lemma 7.3.3. *If* $\forall x \in \Phi$: $(a\,x) \in \Phi$ *and* $b \in \Phi$ *then* $\lambda x.(a\ (b\,x)) \in \Phi$.

Proof. For all $x \in b^{\mathrm{so}}$ we have $(b\,x) \in b^{\mathrm{s}} \subseteq \Phi$, so $(a\ (b\,x)) \in \Phi$. Hence, $\forall x \in b^{\mathrm{so}}$: $((\lambda x.(a\ (b\,x)))\ x) \in \Phi$ and $\lambda x.(a\ (b\,x)) \in \Phi$. $\square$

Lemma 7.3.4 (The C-M1 axiom). *If we have* $\forall u \in \Phi$: $\lambda x.\mathscr{A} \in \Phi$ *then* $\forall v \in \Phi$: $\lambda x.((\lambda u.\mathscr{A})\ (v\,x)) \in \Phi$.

Proof. Assume $\forall u \in \Phi$: $\lambda x.\mathscr{A} \in \Phi$ and $v \in \Phi$, and define $f = \lambda u.\lambda x.\mathscr{A}$. Now $\forall u \in \Phi$: $(f\,u) \in \Phi$, so we have $\lambda y.(f\ (v\,y)) \in \Phi$ according to Lemma 7.3.3 and $\lambda z.((\lambda y.(f\ (v\,y)))\ z\,z) \in \Phi$ by Lemma 7.3.2. By reduction,

$$\begin{aligned}((\lambda y.(f\ (v\,y)))\ z\,z) &= (f\ (v\,z)\ z)\\ &= ((\lambda u.\lambda x.\mathscr{A})\ (v\,z)\ z)\\ &= [\mathscr{A}/u := (v\,z)/x := z],\end{aligned}$$

so $\lambda z.[\mathscr{A}/u := (v\,z)/x := z] \in \Phi$. By renaming z to x, $\lambda x.[\mathscr{A}/u := (v\,x)] \in \Phi$, so $\lambda x.((\lambda u.\mathscr{A})\,(v\,x)) \in \Phi$. □

Lemma 7.3.5 (The C-M2 axiom). *If we have* $\forall u \in \Phi$: $\lambda x.\mathscr{A} \in \Phi$ *then* $\forall u \in \Phi$: $\lambda x.((\lambda x.\mathscr{A})\,(x\,u)) \in \Phi$.

Proof. Assume $u \in \Phi$ and $\lambda x.\mathscr{A} \in \Phi$. Let $f = \lambda x.\mathscr{A}$. Choose $g \in \Phi$ such that $f <_s g$ and $u <_s g$. Assume $x \in g^{so}$. Since $u \in g^s$ we have $(x\,u) \in g^{so}$ and $(f\,(x\,u)) \in g^s$. Hence, $\forall x \in g^{so}$: $((\lambda x.(f\,(x\,u))\,x) \in \Phi$ so $\lambda x.(f\,(x\,u)) \in \Phi$. □

Lemma 7.3.6 (The C-Prim axiom). *If* $\forall x \in \Phi$: $(f\,x) \in \Phi$, $a, b \in \Phi$ *and*

$$(g\,x) = \begin{cases} a & \textit{if}\ x = \mathsf{T}, \\ \bot & \textit{if}\ x = \bot, \\ (f\,\lambda u.(g\,(x\,(b\,u)))) & \textit{otherwise}, \end{cases}$$

then $g \in \Phi$.

Proof. Define $x <_{\text{prim}} y \Leftrightarrow y \neq \mathsf{T} \wedge \exists z \in b^s$: $x = (y\,z)$. The relation $<_{\text{prim}}$ is easily shown to be well-founded on b^{so}. We now prove $(g\,x) \in \Phi$ for all $x \in b^{so}$ by transfinite induction in x and $<_{\text{prim}}$.

if $x = \mathsf{T}$ then $(g\,x) = a \in \Phi$. Now assume $x \neq \mathsf{T}$ and assume (as inductive hypothesis) $\forall y \in b^s$: $(g\,(x\,y)) \in \Phi$. If $u \in b^{so}$ then $(b\,u) \in b^s$ so $(g\,(x\,(b\,u))) \in \Phi$. Hence, $\forall u \in b^{so}$: $((\lambda u.(g\,(x\,(b\,u))))\,u) \in \Phi$ so $\lambda u.(g\,(x\,(b\,u))) \in \Phi$ and $(g\,x) = (f\,\lambda u.(g\,(x\,(b\,u)))) \in \Phi$ as required. □

7.4. The inference of induction

Corollary 7.1.7 states $(f\,g) <_s f$ for all $f, g \in \Phi$, $f \neq \mathsf{T}$. Since $<_s$ is well-founded, this gives rise to an induction principle: If $\mathscr{R}(\mathsf{T})$ is true and $\forall x \in \Phi \backslash \{\mathsf{T}\}$: $(\forall y \in \Phi$: $\mathscr{R}(x\,y) \Rightarrow \mathscr{R}(x))$ then $\forall x \in \Phi$: $\mathscr{R}(x)$. Now let the term $\mathscr{B}$ (where x may occur free) stand for $\mathscr{R}(x)$. We have

$$x \to \mathscr{B}\,; \dot{\neg} x, \phi x, \dot{\forall} y.[\mathscr{B}/x := (x\,y)] \to \mathscr{B} \;\vdash\; \phi x \to \mathscr{B},$$

where y is not allowed to occur in $\mathscr{B}$. To make it possible to prove the deduction theorem (Theorem 5.4.1), the inference is stated as follows:

(induction) If x does not occur free in $\mathscr{A}$ and y does not occur (free or bound) in $\mathscr{B}$, then

$$\mathscr{A}, x \to \mathscr{B}\,; \mathscr{A}, \dot{\neg} x, \phi x, \dot{\forall} y.[\mathscr{B}/x := (x\,y)] \to \mathscr{B} \;\vdash\; \mathscr{A} \to \mathscr{B}.$$

7.5. The metatheorem of totality

We now state and prove Theorem 2.6.1 formally in a slightly generalized form which, among other things includes Theorem 2.6.4.

As in Section 2.6 let Σ and $\bar{\Sigma}$ denote the syntax classes of *simple* and *dual* terms, respectively. Let $x_1, x_2, \ldots$ and $y_1, y_2, \ldots$ denote distinct variables, and let $\Sigma^{\#}$ denote the syntax class of simple terms in which $x_1, x_2, \ldots$ do not occur free. Let Σ^j denote a sequence of exactly j simple terms. Let f_{ij} ($i, j \in \{0, 1, \ldots\}$), denote terms in which $x_1, x_2, \ldots$ and $y_1, y_2, \ldots$ do not occur free. This section defines Σ to be slightly more general than Section 2.6 did:

$$\begin{aligned} \Sigma \ &::= \ y_i \mid \lambda x_i.\Sigma \mid (\Sigma\,\bar{\Sigma}) \mid \mathsf{T} \mid \varepsilon y_i.\Sigma^{\#} \mid \phi\Sigma \mid (\mathsf{if}\ \bar{\Sigma}\,\Sigma\,\Sigma) \\ &\quad \mid ((\lambda y_{k_1} \ldots \lambda y_{k_i}.\Sigma^{\#})\,\Sigma^i) \mid (f_{ij}\,\Sigma^i) \mid (Prim\,(\lambda y_k.\Sigma^{\#})\,\Sigma\,\Sigma), \\ \bar{\Sigma} \ &::= \ x_i \mid (\bar{\Sigma}\,\Sigma) \mid \Sigma. \end{aligned}$$

Theorem 7.5.1 (Totality). *If $\mathscr{A}$ is a simple term, if $x_1, x_2, \ldots$ and $y_1, y_2, \ldots$ do not occur free in $\mathscr{A}$, and if $\phi y_1, \ldots, \phi y_i \to \phi(f_{ij}\,y_1 \ldots y_i)$ for all f_{ij} that occur in $\mathscr{A}$, then $\phi\mathscr{A}$.*

We shall discuss simple terms before proving the totality theorem. If $\mathscr{A}$ and $\mathscr{B}$ are simple terms, then $\approx\mathscr{A}$, $!\mathscr{A}$, $¡\mathscr{A}$, $\dot{\neg}\mathscr{A}$, $\mathscr{A} \dot{\wedge} \mathscr{B}$, $\mathscr{A} \dot{\vee} \mathscr{B}$, $\mathscr{A} \dot{\Rightarrow} \mathscr{B}$ and $\mathscr{A} \dot{\Leftrightarrow} \mathscr{B}$ are simple terms. Further, if $K = \lambda y_1.\lambda x_1.y_1$ and $S = \lambda y_1.\lambda y_2.\lambda x_1.(y_1\,x_1\,(y_2\,x_1))$ then $(K\,\mathscr{A})$ and $(S\,\mathscr{A}\,\mathscr{B})$ are simple terms. If $x_1, x_2, \ldots$ do not occur free in $\mathscr{A}$ then $\dot{\exists} y_i.\mathscr{A}$ and $\dot{\forall} y_1.\mathscr{A}$ are simple terms.

If $\mathscr{A}$ is simple and $\mathscr{A} = \mathscr{B}$ holds according to the theorem of reduction, then we call $\mathscr{B}$ *almost simple*. The totality theorem trivially extends to almost simple terms. This is particularly useful in the production rule

$$\Sigma \ ::= \ (f_{ij}\,\Sigma^i).$$

As an example, we prove later that $\phi y_1, \phi y_2 \to \phi(y_1 \doteq y_2)$. Hence, $\lambda y_1.\lambda y_2.(y_1 \doteq y_2)$ may take the place of f_{20} so that

$$\Sigma \ ::= \ ((\lambda y_1.\lambda y_2.(y_1 \doteq y_2))\,\Sigma\,\Sigma)$$

becomes a production rule. Since $((\lambda y_1.\lambda y_2.(y_1 \doteq y_1))\,\mathscr{A}\,\mathscr{B}) = (\mathscr{A} \doteq \mathscr{B})$ we have by the theorem of reduction that $\mathscr{A} \doteq \mathscr{B}$ is almost simple if $\mathscr{A}$ and $\mathscr{B}$ are almost simple. Hence, the production rule $\Sigma ::= (f_{ij}\,\Sigma^i)$ allows to extend the class of almost simple terms for each theorem of form $\phi y_1, \ldots, \phi y_i \to \phi\mathscr{E}$.

In particular, for $i = 0$, a result of form $\phi\mathscr{E}$ allows to extend the production rules by $\Sigma ::= \mathscr{E}$. As an example of its use, consider the following proof of $\phi y \to \phi\lambda x.(y\,(x\,(y\,x)))$, where line 3 is verified by the deduction theorem.

Proof.

1. assume	ϕy
2. 1,totality	$\phi y \to \phi\lambda x.(y\,(x\,(y\,x)))$
3. 1-2	$\phi y \to \phi y \to \phi\lambda x.(y\,(x\,(y\,x))).$ $\square$

Hence, $f_{00}, f_{01}, \ldots$ in the totality theorem replace $y_0, \ldots, y_n$ in Theorem 2.6.1.

The deduction theorem is used so often, that we introduce a special notation for it. In proofs, a range of lines (e.g. 1–2 in the proof above) denotes that the deduction theorem has been applied to that range. The presence of a dash always refers to the deduction theorem.

The reduction rule

$$\Sigma \ ::= \ ((\lambda y_{k_1} \ldots \lambda y_{k_i}.\Sigma^{\#})\ \Sigma^{i})$$

is primarily intended to stand for

$$\Sigma \ ::= \ [\Sigma^{\#}/y_{k_1} := \Sigma / \cdots / y_{k_i} := \Sigma],$$

but the former formulation was chosen to avoid variable conflicts. As an example of its usage, define

$$\mathscr{A} = \varepsilon y_1.\lambda x_1.\varepsilon y_2.(y_1\ x_1\ y_2),$$

$$\mathscr{B} = \varepsilon y_1.\lambda x_1.((\lambda y_3.\varepsilon y_2.(y_3\ y_2))\ (y_1\ x_1)).$$

Now $\mathscr{B}$ is simple and $\mathscr{A} = \mathscr{B}$ by the reduction theorem, so $\mathscr{A}$ is almost simple. Hence, $\phi\mathscr{A}$ holds. On the other hand, $\mathscr{A}$ is not simple because $\varepsilon y_2.(y_1\ x_1\ y_2)$ is not simple. The latter is not simple because the production rule $\Sigma ::= \varepsilon y_i.\Sigma^{\#}$ merely allows $\varepsilon y_i.\mathscr{A}$ to be simple if $\mathscr{A}$ contains no x_j free. Hence, the reduction rule

$$\Sigma \ ::= \ ((\lambda y_{k_1} \ldots \lambda y_{k_i}.\Sigma^{\#})\ \Sigma^{i})$$

is useful to circumvent the restriction in the production rules

$$\Sigma \ ::= \ \varepsilon y_i.\Sigma^{\#},$$

$$\Sigma \ ::= \ (\mathit{Prim}\ (\lambda y_i.\Sigma^{\#})\ \Sigma\ \Sigma).$$

It is not possible to circumvent the restrictions in all cases. As an example, $\lambda x.\varepsilon y.(y\ x)$ is not well-founded. It is slightly complicated to state exactly when the restrictions can be circumvented, and we shall avoid stating the rule.

The rest of this section proves the totality theorem. In order to prove the totality theorem we first prove some lemmas.

Lemma 7.5.2. *If x is not free in $\mathscr{A}$ then $\phi\mathscr{A} \to \phi x.\mathscr{A}$.*

Proof.

1. assume	$\phi\mathscr{A}$
2. C-K′	$\phi x.\mathsf{T}$
3. 1,2,triv	$\phi x.\phi\mathscr{A}$
4. 3,Well2,triv	$\phi x.\mathscr{A}$
5. 1–4	$\phi\mathscr{A} \to \phi x.\mathscr{A}$. $\square$

Lemma 7.5.3. $\phi\mathscr{A}, \phi\mathscr{B} \to \phi(P\,\mathscr{A}\,\mathscr{B})$.

Proof.

1. assume	$\phi\mathscr{A}, \phi\mathscr{B}$
2. C-P′	$\phi\lambda x.(\mathsf{if}\ x\ \mathsf{T}\ \mathsf{T})$
3. 1,2,triv	$\phi\lambda x.(\mathsf{if}\ x\ \phi\mathscr{A}\ \phi\mathscr{B})$
4. *QND*	$(\mathsf{if}\ x\ \phi\mathscr{A}\ \phi\mathscr{B}) = \phi(\mathsf{if}\ x\ \mathscr{A}\ \mathscr{B})$
5. 3,4,triv	$\phi\lambda x.\phi(\mathsf{if}\ x\ \mathscr{A}\ \mathscr{B})$
6. 5,Well2,triv	$\phi\lambda x.(\mathsf{if}\ x\ \mathscr{A}\ \mathscr{B})$
7. 6,triv	$\phi(P\,\mathscr{A}\,\mathscr{B})$
8. 1–7	$\phi\mathscr{A}, \phi\mathscr{B} \to \phi(P\,\mathscr{A}\,\mathscr{B})$. $\square$

Lemma 7.5.4.

$$\phi y_1, \ldots, \phi y_m \to \mathscr{A} \ \vdash\ \phi y_1, \ldots, \phi y_m \to \dot{\forall} y_i.\mathscr{A} \quad \textit{if } i \in \{1, \ldots, m\}.$$

Proof.

1. assume	$\phi y_1, \ldots, \phi y_m \to \mathscr{A}$
2. assume	$\phi y_1, \ldots, \phi y_{i-1}, \phi y_{i+1}, \ldots, \phi y_m, y_i$
3. 1,2,triv	$\mathscr{A}$
4. 2,3,Lem.6.2.2	$\phi y_1, \ldots, \phi y_{i-1}, \phi y_{i+1}, \ldots, \phi y_m \to \dot{\forall} y_i.\mathscr{A}$
5. 4,*QND*,triv	$\phi y_1, \ldots, \phi y_m \to \dot{\forall} y_i.\mathscr{A}$. $\square$

Lemma 7.5.5.

(a) $\phi y_1, \ldots, \phi y_m \to \phi\mathscr{A} \ \vdash\ \phi y_1, \ldots, \phi y_m \to \phi x.\mathscr{A}$ *if x is not free in $\mathscr{A}$.*

(b) $\phi y_1, \ldots, \phi y_m \to \phi x.\mathscr{A} \ \vdash\ \phi y_1, \ldots, \phi y_m \to \phi x.[\mathscr{A}/y_i := (y_i\, x)]$.

(c) $\phi y_1, \ldots, \phi y_m \to \phi x.\mathscr{A} \ \vdash\ \phi y_1, \ldots, \phi y_m \to \phi x.[\mathscr{A}/x := (x\, y_i)]$.

(d) $\phi y_1, \ldots, \phi y_m \to \phi x.\mathscr{A} \ \vdash\ \phi y_1, \ldots, \phi y_m \to \phi[\mathscr{A}/x := \mathsf{T}]$.

Proof of (a).

1. assume	$\phi y_1, \ldots, \phi y_m \to \phi\mathscr{A}$
2. assume	$\phi y_1, \ldots, \phi y_m$
3. 1,2,triv	$\phi\mathscr{A}$
4. 3,triv,Lem.7.5.2	$\phi x.\mathscr{A}$
5. 2–4	$\phi y_1, \ldots, \phi y_m \to \phi x.\mathscr{A}$. $\square$

Proof of (b).

1. assume	$\phi y_1, \ldots, \phi y_m \to \phi x.\mathscr{A}$
2. 1,Lem.7.5.4	$\phi y_1, \ldots, \phi y_m \to \dot{\forall} y_i.\phi x.\mathscr{A}$
3. 2,C-M1,triv	$\phi y_1, \ldots, \phi y_m \to \dot{\forall} y_i.\phi x.[\mathscr{A}/y_i := (y_i\, x)]$
4. assume	$\phi y_1, \ldots, \phi y_m$
5. 3,4,triv,Lem.6.2.1	$\phi x.[\mathscr{A}/y_i := (y_i\, x)]$
6. 4–5	$\phi y_1, \ldots, \phi y_m \to \phi x.[\mathscr{A}/y_i := (y_i\, x)]$. $\square$

Proof of (c).

1. assume	$\phi y_1, \ldots, \phi y_m \to \phi x.\mathscr{A}$
2. 1,Lem.7.5.4	$\phi y_1, \ldots, \phi y_m \to \dot{\forall} y_i.\phi x.\mathscr{A}$
3. 2,C-M2,triv	$\phi y_1, \ldots, \phi y_m \to \dot{\forall} y_i.\phi x.[\mathscr{A}/x := (x\, y_i)]$
4. assume	$\phi y_1, \ldots, \phi y_m$
5. 3,4,triv,Lem.6.2.1	$\phi x.[\mathscr{A}/x := (x\, y_i)]$
6. 4–5	$\phi y_1, \ldots, \phi y_m \to \phi x.[\mathscr{A}/x := (x\, y_i)]$. $\square$

Proof of (d).

1. assume	$\phi y_1, \ldots, \phi y_m \to \phi x.\mathscr{A}$
2. assume	$\phi y_1, \ldots, \phi y_m$
3. 1,2,triv	$\phi x.\mathscr{A}$
4. 3,Well1,C-A,triv	$\phi((\lambda x.\mathscr{A})\, \mathsf{T})$
5. 4,triv	$\phi[\mathscr{A}/x := \mathsf{T}]$
6. 2–5	$\phi y_1, \ldots, \phi y_m \to \phi[\mathscr{A}/x := \mathsf{T}]$. $\square$

Now let x and z be distinct variables that do occur among $y_1, y_2, \ldots$ and define the syntax class Γ of *unary simple terms* by

$$\begin{aligned} \Gamma ::= {} & y_i \mid \lambda z.((\lambda x.\Gamma)\,(P\,x\,z)) \mid (\Gamma\,\Gamma) \mid (\Gamma\,(x\,\Gamma^*)) \mid \mathsf{T} \mid \varepsilon y_i.[\Gamma/x := \mathsf{T}] \\ & \mid \phi\Gamma \mid (P\,\Gamma\,\Gamma) \mid ((\lambda y_{k_1} \ldots \lambda y_{k_i}.[\Gamma/x := \mathsf{T}])\,\Gamma^i) \\ & \mid (f_{ij}\,\Gamma^i) \mid (\mathit{Prim}\,(\lambda y_k.[\Gamma/x := \mathsf{T}])\,\Gamma\,\Gamma), \end{aligned}$$

where Γ^* denotes a sequence of zero, one, or more unary simple terms and Γ^i denotes a sequence of exactly i unary simple terms.

Any of the variables $x_1, x_2, \ldots$ may occur free in simple terms whereas merely x may occur free in unary simple terms (apart from $y_1, y_2, \ldots$ which may occur free in simple as well as unary simple terms). We shall prove a theorem about the well-foundedness of unary simple terms, and then prove the totality theorem from this special case.

Lemma 7.5.6 (Unary totality). *If $\mathcal{T}$ is a unary simple term, if $y_{m+1}, y_{m+2}, \ldots$ do not occur in $\mathcal{T}$, and if $\phi y_1, \ldots, \phi y_i \to \phi(f_{ij}\, y_1 \ldots y_i)$ for all f_{ij} that occur in $\mathcal{T}$, then $\phi y_1, \ldots, \phi y_m \to \phi x.\mathcal{T}$.*

Proof. We proceed by structural induction in Γ and $\mathcal{T}$.

Case 1. Assume that $\mathcal{T}$ is y_i.

1. *QND*	$\phi y_1, \ldots, \phi y_m \to \phi y_i$
2. 1,Lem.7.5.5	$\phi y_1, \ldots, \phi y_m \to \phi x.y_i$.

Case 2. Assume that $\mathcal{T}$ is $\lambda z.((\lambda x.\mathcal{A})\,(P\,x\,z))$.

1. ind. hyp	$\phi y_1, \ldots, \phi y_m \to \phi x.\mathcal{A}$
2. assume	$\phi y_1, \ldots, \phi y_m$
3. 1,2,triv	$\phi x.\mathcal{A}$
4. C-Curry	$\phi a \to \phi x.\lambda z.(a\,(P\,x\,z))$
5. 3,4,triv	$\phi x.\lambda z.((\lambda x.\mathcal{A})\,(P\,x\,z))$
6. 2–5	$\phi y_1, \ldots, \phi y_m \to \phi x.\lambda z.((\lambda x.\mathcal{A})\,(P\,x\,z))$.

Case 3. Assume that $\mathcal{T}$ is $(\mathcal{A}\,\mathcal{B})$.

1. ind. hyp	$\phi y_1, \ldots, \phi y_m \to \phi x.\mathcal{A}$
2. ind. hyp	$\phi y_1, \ldots, \phi y_m \to \phi x.\mathcal{B}$
3. assume	$\phi y_1, \ldots, \phi y_m$
4. 1,2,3,triv	$\phi x.\mathcal{A},\ \phi x.\mathcal{B}$
5. C-A	$\phi a,\ \phi b \to \phi(a\,b)$
6. 5,Lem.7.5.5	$\phi a,\ \phi b \to \phi x.((a\,x)\,(b\,x))$
7. 4,6,triv	$\phi x.(\mathcal{A}\,\mathcal{B})$
8. 3–7	$\phi y_1, \ldots, y_m \to \phi x.(\mathcal{A}\,\mathcal{B})$.

Case 4. Assume that $\mathcal{T}$ is $(\mathcal{A}\,(x\,\mathcal{B}_1 \ldots \mathcal{B}_n))$.

1. ind. hyp	$\phi y_1, \ldots, \phi y_m \to \phi x.\mathcal{A}$
2. ind. hyp	$\phi y_1, \ldots, \phi y_m \to \phi x.\mathcal{B}_i,\ i \in \{1, \ldots, n\}$
3. assume	$\phi y_1, \ldots, y_m$
4. 1,2,3,triv	$\phi x.\mathcal{A},\ \phi x.\mathcal{B}_1, \ldots, \phi x.\mathcal{B}_n$
5. *QND*	$\phi a,\ \phi b_1, \ldots, \phi b_n \to \phi a$
6. 5,Lem.7.5.5	$\phi a,\ \phi b_1, \ldots, \phi b_n \to \phi x.((a\,x)\,x\,(b_1\,x) \ldots (b_n\,x))$
7. 4,6,triv	$\phi x.(\mathcal{A}\,(x\,\mathcal{B}_1 \ldots \mathcal{B}_n))$
8. 3–7	$\phi y_1, \ldots, \phi y_m \to \phi x.(\mathcal{A}\,(x\,\mathcal{B}_1 \ldots \mathcal{B}_m))$.

Case 5. Assume that $\mathcal{T}$ is T.

1. assume	$\phi y_1, \ldots, \phi y_m$
2. C-K'	$\phi x.\mathsf{T}$
3. 1–2	$\phi y_1, \ldots, \phi y_m \to \phi x.\mathsf{T}$.

Case 6. Assume that $\mathcal{T}$ is $\varepsilon y_k.\mathcal{B}$ where $\mathcal{B}$ is shorthand for $[\mathcal{A}/x := \mathsf{T}]$.

1. ind. hyp	$\phi y_1, \ldots, \phi y_m \to \phi x.\mathcal{A}$
2. 1,Lem.7.5.5	$\phi y_1, \ldots, \phi y_m \to \phi \mathcal{B}$
3. assume	$\phi y_1, \ldots, \phi y_m$
4. 2,3,triv	$\phi \mathcal{B}$
5. 4,*QND*	$!\mathcal{B}$
6. 3–5	$\phi y_1, \ldots, \phi y_m \to !\mathcal{B}$
7. 6,Lem.7.5.4	$\phi y_1, \ldots, \phi y_m \to \forall y_i.!\mathcal{B}$
8. 7,Quantify2,triv	$\phi y_1, \ldots, \phi y_m \to \phi \varepsilon y_i.\mathcal{B}$.

Case 7. Assume that $\mathcal{T}$ is $\phi\mathcal{A}$.

1. ind. hyp	$\phi y_1, \ldots, \phi y_m \to \phi x.\mathcal{A}$
2. assume	$\phi y_1, \ldots, \phi y_m$
3. 1,2,triv	$\phi x.\mathcal{A}$
4. 3,Well2	$\phi x.\phi\mathcal{A}$
5. 2–4	$\phi y_1, \ldots, \phi y_m \to \phi x.\mathcal{A}$.

Case 8. Assume that $\mathcal{T}$ is $(P\,\mathcal{A}\,\mathcal{B})$.

1. ind. hyp	$\phi y_1, \ldots, \phi y_m \to \phi x.\mathcal{A}$
2. ind. hyp	$\phi y_1, \ldots, \phi y_m \to \phi x.\mathcal{B}$
3. assume	$\phi y_1, \ldots, \phi y_m$
4. 1,2,3,triv	$\phi x.\mathcal{A},\ \phi x.\mathcal{B}$
5. C-P	$\phi a,\ \phi b \to \phi(P\,a\,b)$
6. 5,Lem.7.5.5	$\phi a,\ \phi b \to \phi x.(P\ (a\,x)\ (b\,x))$
7. 4,6,triv	$\phi x.(P\,\mathcal{A}\,\mathcal{B})$
8. 3–7	$\phi y_1, \ldots, \phi y_m \to \phi x.(P\,\mathcal{A}\,\mathcal{B})$.

Case 9. Assume that $\mathcal{T}$ is $((\lambda y_{k_1} \ldots \lambda y_{k_i}.\mathcal{B})\ \mathcal{C}_1 \ldots \mathcal{C}_i)$ where $\mathcal{B}$ is shorthand for $[\mathcal{A}/x := \mathsf{T}]$. Let $\mathcal{B}'$ be shorthand for $(\lambda y_{k_1} \ldots \lambda y_{k_i}.\mathcal{B})$.

1. ind. hyp	$\phi y_1, \ldots, \phi y_m \rightarrow \phi x.\mathcal{A}$
2. 1,Lem.7.5.5.	$\phi y_1, \ldots, \phi y_m \rightarrow \mathcal{B}$
3. 2,triv	$\phi y_1, \ldots, \phi y_m \rightarrow (\mathcal{B}'\ y_{k_1} \ldots y_{k_i})$
4. 3,Lem.7.5.5	$\phi y_1, \ldots, \phi y_m \rightarrow \phi x.(\mathcal{B}'\ (y_{k_1}\ x) \ldots (y_{k_i}\ x))$
5. ind. hyp	$\phi y_1, \ldots, \phi y_m \rightarrow \phi x.\mathcal{C}_j, \quad j \in \{1, \ldots, i\}$
6. assume	$\phi y_1, \ldots, \phi y_m$
7. 4,6,triv	$\phi y_{k_1}, \ldots, \phi y_{k_i} \rightarrow \phi x.(\mathcal{B}'\ (y_{k_1}\ x) \ldots (y_{k_i}\ x))$
8. 5,6,triv	$\phi x.\mathcal{C}_1, \ldots, \phi x.\mathcal{C}_i$
9. 7,8,triv	$\phi x.(\mathcal{B}'\ \mathcal{C}_1 \ldots \mathcal{C}_i)$
10. 6–9	$\phi y_1, \ldots, \phi y_m \rightarrow \phi x.(\mathcal{B}'\ \mathcal{C}_1 \ldots \mathcal{C}_i)$.

Case 10. Assume that $\mathcal{T}$ is $(f_{ij}\ \mathcal{A}_1 \ldots \mathcal{A}_i)$.

1. assume	$\phi a_1, \ldots, \phi a_i \rightarrow \phi(f_{ij}\ a_1 \ldots a_i)$
2. ind. hyp	$\phi y_1, \ldots, \phi y_m \rightarrow \phi x.\mathcal{A}_j, \quad j \in \{1, \ldots, i\}$
3. assume	$\phi y_1, \ldots, \phi y_m$
4. 2,3,triv	$\phi x.\mathcal{A}_1, \ldots, \phi x.\mathcal{A}_i$
5. 1,Lem.7.5.5	$\phi a_1, \ldots, \phi a_i \rightarrow \phi x.(f_{ij}\ (a_1\ x) \ldots (a_i\ x))$
6. 4,5,triv	$\phi x.(f_{ij}\ \mathcal{A}_1 \ldots \mathcal{A}_i)$
7. 3–6	$\phi y_1, \ldots, \phi y_m \rightarrow \phi x.(f_{ij}\ \mathcal{A}_1 \ldots \mathcal{A}_i\}$.

Case 11. Assume that $\mathcal{T}$ is $(Prim(\lambda y_i.\mathcal{B})\ \mathcal{C}\ \mathcal{D})$ where $\mathcal{B}$ is shorthand for $[\mathcal{A}/x := \mathsf{T}]$.

1. ind. hyp	$\phi y_1, \ldots, \phi y_m \rightarrow \phi x.\mathcal{A}$
2. 1,Lem.7.5.5	$\phi y_1, \ldots, \phi y_m \rightarrow \phi\mathcal{B}$
3. 2,Lem.7.5.4	$\phi y_1, \ldots, \phi y_m \rightarrow \dot{\forall} y_i.\phi\mathcal{B}$
4. ind. hyp	$\phi y_1, \ldots, \phi y_m \rightarrow \phi x.\mathcal{C}$
5. ind. hyp	$\phi y_1, \ldots, \phi y_m \rightarrow \phi x.\mathcal{D}$
6. assume	$\phi y_1, \ldots, \phi y_m$
7. 3,4,5,6,triv	$\dot{\forall} y_i.\phi((\lambda y_i.\mathcal{B})\ y_i),\ \phi x.\mathcal{C},\ \phi x.\mathcal{D}$
8. 7,C-Prim	$\phi c, \phi d \rightarrow \phi x.(Prim(\lambda y_i.\mathcal{B})\ c\ d)$
9. 8,Lem.7.5.5	$\phi c, \phi d \rightarrow \phi x.(Prim(\lambda y_i.\mathcal{B})\ (c\ x)\ (d\ x))$
10. 7,9,triv	$\phi x.(Prim(\lambda y_i.\mathcal{B})\ \mathcal{C}\ \mathcal{D})$
11. 6–10	$\phi y_1, \ldots, \phi y_m \rightarrow \phi x.(Prim(\lambda y_i.\mathcal{B})\ \mathcal{C}\ \mathcal{D})$. $\square$

We may now prove the Totality Theorem from the Unary Totality Theorem.

Proof of the Totality Theorem. Let $\mathscr{A}$ be a simple term that satisfies the conditions of the Totality Theorem. We may construct a unary simple term $\mathscr{B}$ such that $\mathscr{A}=[\mathscr{B}/x:=\mathsf{T}]$ by reduction and such that $\mathscr{B}$ satisfies the conditions of the Unary Totality Theorem. The construction is as follows: Replace all occurrences of $\lambda x_i.\mathscr{C}$ in $\mathscr{A}$ by $\lambda z.((\lambda x.\mathscr{C})\,(P\,x\,z))$. Replace all occurrences of x_i by

$$(x\,\overbrace{\mathsf{T}\ldots\mathsf{T}}^{j}\,\mathsf{F})$$

where j is the De Bruijn index of the occurrence of x_i. Replace all occurrences of $\varepsilon y_i.\mathscr{C}$ and $(\mathit{Prim}(\lambda y_k.\mathscr{C})\,\mathscr{D}\,\mathscr{E})$ by $\varepsilon y_i.[\mathscr{C}/x:=\mathsf{T}]$ and $(\mathit{Prim}(\lambda y_k.[\mathscr{C}/x:=\mathsf{T}])\,\mathscr{D}\,\mathscr{E})$, respectively. Replace all occurrences of $(\text{if}\ \mathscr{C}\,\mathscr{D}\,\mathscr{E})$ by $(P\,\mathscr{D}\,\mathscr{E}\,\mathscr{C})$. It is left to the reader to see that $\mathscr{B}$ is a unary simple term and that $\mathscr{A}=[\mathscr{B}/x:=\mathsf{T}]$ (note that $\mathscr{B}$ does not depend on x). Since $\mathscr{A}$ satisfies the conditions of the Totality Theorem, it contains no y_i free. Hence, neither does $\mathscr{B}$, so $\phi x.\mathscr{B}$ by the Unary Totality Theorem. From axiom Well1 and C-A we have $\phi((\lambda x.\mathscr{B})\,\mathsf{T})$ which shows $\phi\mathscr{A}$. $\square$

8. Development of *ZFC*

8.1. The syntax and axioms of ZFC

The following axiomatization of *ZFC* is inspired by [17] and [22]. The syntax of terms and well-formed formulas (wff's) of *ZFC* reads:

$$\text{variable} ::= x \mid y \mid z \mid \cdots$$

$$\text{term} ::= \text{variable} \in \text{variable}$$

$$\text{wff} ::= \text{term} \mid \neg\text{wff} \mid (\text{wff} \Rightarrow \text{wff}) \mid \forall\text{variable}.\text{wff}.$$

Let $\mathscr{A}{:}\mathit{ZFC}(x_1,\ldots,x_n)$ denote that $\mathscr{A}$ is a well-formed formula of *ZFC* whose free variables occur among $x_1,\ldots,x_n$.

The logical axioms read ([22])

(ZFC-A1) $\mathscr{A}\Rightarrow(\mathscr{B}\Rightarrow\mathscr{A})$

(ZFC-A2) $(\mathscr{A}\Rightarrow(\mathscr{B}\Rightarrow\mathscr{C}))\Rightarrow((\mathscr{A}\Rightarrow\mathscr{B})\Rightarrow(\mathscr{A}\Rightarrow\mathscr{C}))$

(ZFC-A3) $(\neg\mathscr{B}\Rightarrow\neg\mathscr{A})\Rightarrow((\neg\mathscr{B}\Rightarrow\mathscr{A})\Rightarrow\mathscr{B})$

(ZFC-A4) $\forall x.\mathscr{A}\Rightarrow[\mathscr{A}/x:=t]$ where t is a term free for x in $\mathscr{A}$

(ZFC-A5) $\forall x.(\mathscr{A}\Rightarrow\mathscr{B})\Rightarrow(\mathscr{A}\Rightarrow\forall x.\mathscr{B})$ if x is not free in $\mathscr{A}$.

The inference rules read

(ZFC-MP) $\mathscr{A}; \mathscr{A} \Rightarrow \mathscr{B} \vdash \mathscr{B}$

(ZFC-Gen) $\mathscr{A} \vdash \forall x.\mathscr{A}$.

The axiom of extensionality reads

(ZFC-E) $\forall z.(z \in x \Leftrightarrow z \in y) \Rightarrow \forall z.(x \in z \Leftrightarrow y \in z)$,

where $\Leftrightarrow$ is defined from $\Rightarrow$ and $\neg$ as usual. We shall use defined logical connectives like $\Leftrightarrow$ and the defined quantifier $\exists$ in stating axioms, but we shall avoid defined relations like $\subseteq$, and defined functions since they would complicate matters. The axiom of subsets reads

(ZFC-S) $\forall y \exists z \forall x.(x \in z \Leftrightarrow x \in y \wedge \mathscr{A})$ where y and z do not occur free in $\mathscr{A}$,

where $\mathscr{A}$ is any well-formed formula of *ZFC*.

The construction axioms of *ZFC* are easier to state when we have the axiom of subsets at our disposal [17]. As an example, we state the axiom of pair sets as

(ZFC-P) $\forall x \forall y \exists z.x \in z \wedge y \in z$.

This axiom states that for any sets x and y there is a z containing both which makes z a superset of the pair set $\{x, y\}$. Having a superset of the pair set, the pair set itself may be constructed using the axiom of subsets. The axioms of union and power sets are stated similarly:

(ZFC-U) $\forall x \exists y \forall u \forall v.(u \in v \wedge v \in x \Rightarrow u \in y)$

(ZFC-W) $\forall x \exists y \forall z.(\forall u.(u \in z \Rightarrow u \in x) \Rightarrow z \in y)$.

The axioms of pair and null sets are not strictly necessary, but it is instructive to verify them in Section 8.10. The axiom of null sets reads:

(ZFC-N) $\exists x \forall y. \neg y \in x$.

The axiom of replacement reads

(ZFC-R) $\forall z \exists u \forall x.(x \in z \wedge \exists y.\mathscr{A} \Rightarrow \exists y.y \in u \wedge \mathscr{A})$

where z and u are not free in $\mathscr{A}$.

The interpretation is as follows: The well-formed formula $\mathscr{A}$ may contain x and y free. Consider $\mathscr{A}$ as a multi-valued function that maps x to y iff $\mathscr{A}$ is true. In particular, $\mathscr{A}$ may be a single-valued function, in which case the axiom expresses the usual axiom of replacement. When $\mathscr{A}$ is multi-valued, the axiom states part of what the axiom of choice states. Hence, ZFC-R is stronger than the usual axiom of replacement, but ZFC-R plus the axiom of choice has the same strength as the usual axiom of replacement plus the axiom of choice.

Stated without use of defined functions and relations, the axiom of choice is colossal. We use the formulation that provides a choice set y for any collection x of disjoint, nonempty sets.

$$
\begin{aligned}
(\mathrm{ZFC\text{-}C}) \quad & \forall x.[\forall u.(u \in x \Rightarrow \exists v.v \in u) \\
& \quad \wedge \forall u, v, w, z.(u \in x \wedge v \in x \wedge w \in u \wedge w \in v \wedge z \in u \Rightarrow z \in v) \\
& \Rightarrow \exists y.\forall u.(u \in x \Rightarrow \exists v.v \in u \wedge v \in y) \\
& \quad \wedge \forall u, v, w, z.(u \in x \wedge v \in u \wedge v \in y \wedge w \in u \wedge w \in y \wedge z \in v \Rightarrow z \in w)].
\end{aligned}
$$

The axiom of restriction says that any nonempty set x contains a set y which is disjoint from x. This axiom allows to prove the axiom of foundation [22], i.e., that there is no infinite sequence $x_1, x_2, \ldots$ of sets such that $x_1 \ni x_2 \ni x_3 \ni \cdots$.

$$(\mathrm{ZFC\text{-}D}) \quad \forall x.(\exists y.y \in x \Rightarrow \exists y.y \in x \wedge \neg\exists z.z \in x \wedge z \in y).$$

The axiom of infinity states that there exists an infinite set. Having the axiom of restriction, it is sufficient to assume that there exists a nonempty set y such that whenever y contains a set z, then it also contains a set u which contains z. Without the axiom of restriction, this axiom of infinity would be satisfied, e.g., by any set u that satisfied $u = \{u\}$, and such a u is finite since it merely has one element. The axiom of infinity reads

$$(\mathrm{ZFC\text{-}I}) \quad \exists y.((\exists z.z \in y) \wedge \forall z.(z \in y \Rightarrow \exists u.z \in u \wedge u \in y)).$$

8.2. The strategy of development

We shall prove that any theorem of ZFC is provable in map theory. More specifically we shall prove the following result.

Theorem 8.2.1. *If $\mathscr{A}:ZFC(x_1, \ldots, x_n)$ is a theorem of ZFC, then $\phi x_1, \ldots, \phi x_n \rightarrow \mathscr{A}$ is a theorem of map theory.*

The theorem follows from the following lemmas.

Lemma 8.2.2. *If we have $\phi y_1, \ldots, \phi y_m \rightarrow \mathscr{A}$ and $\phi y_1, \ldots, \phi y_m \rightarrow (\mathscr{A} \dot{\Rightarrow} \mathscr{B})$ then $\phi y_1, \ldots, \phi y_m \rightarrow \mathscr{B}$.*

Lemma 8.2.3. *If $\phi y_1, \ldots, \phi y_m \rightarrow \mathscr{A}$ then $\phi y_1, \ldots, \phi y_m \rightarrow \dot{\forall} y_i.\mathscr{A}$.*

Lemma 8.2.4. *If $\mathscr{A}:ZFC(y_1, \ldots, y_m)$ is an axiom of ZFC then $\phi y_1, \ldots, \phi y_m \rightarrow \mathscr{A}$.*

Lemma 8.2.5. *If we have $\mathscr{A}:ZFC(x_1, \ldots, x_n)$ and $\mathscr{A}:ZFC(y_1, \ldots, y_m)$ then $\phi x_1, \ldots, \phi x_n \rightarrow \mathscr{A}$ iff $\phi y_1, \ldots, \phi y_m \rightarrow \mathscr{A}$.*

Lemmas 8.2.2–8.2.4 are verified in Sections 8.8–8.11, while Lemma 8.2.5 is trivial. The theorem follows from the lemmas as follows.

Assume $\mathscr{A}{:}ZFC(x_1, \ldots, x_n)$ is a theorem of ZFC. Since $\mathscr{A}$ is a theorem, it has a proof $\mathscr{A}_1, \ldots, \mathscr{A}_k$ ending with $\mathscr{A}$ where each $\mathscr{A}_i$ is either an axiom or follows from previous $\mathscr{A}_j$'s by an inference rule. Assume the free variables of $\mathscr{A}_1, \ldots, \mathscr{A}_k$ occur among $y_1, \ldots, y_m$. It is straightforward to verify $\phi y_1, \ldots, \phi y_n \rightarrow \mathscr{A}_i$ by induction in i using Lemmas 8.2.2–8.2.4. In particular, for $i = n$, $\phi y_1, \ldots, \phi y_m \rightarrow \mathscr{A}$, so $\phi x_1, \ldots, \phi x_n \rightarrow \mathscr{A}$ by Lemma 8.2.5.

Theorem 8.2.1 would be true trivially if $\phi x_1, \ldots, \phi x_n \rightarrow \mathscr{A}$ would hold regardless whether or not $\mathscr{A}$ was a theorem. To rule out this possibility it is sufficient to find a single well-formed non-theorem $\mathscr{A}$ such that $\phi x_1, \ldots, \phi x_n \rightarrow \mathscr{A}$ is not provable. For example, let $\mathscr{A}$ be $x \dot{\in} x$. If $\phi x \rightarrow x \dot{\in} x$ is provable in map theory then, in particular, $\phi \mathsf{T} \rightarrow \mathsf{T} \dot{\in} \mathsf{T}$ is provable so, since $\phi \mathsf{T} = \mathsf{T}$, we would have $\mathsf{T} \dot{\in} \mathsf{T} = \mathsf{T}$. However, $\mathsf{T} \dot{\in} \mathsf{T} = \mathsf{F}$ by reduction, so we would have $\mathsf{T} = \mathsf{F}$ and map theory would be inconsistent. Hence, if map theory is consistent then Theorem 8.2.1 is nontrivial.

In Sections 8.8–8.11 we shall not prove Lemma 8.2.4 for every list $y_1, \ldots, y_m$ of variables. We merely prove $\phi z_1, \ldots, \phi z_p \rightarrow \mathscr{A}$ for *one* list $z_1, \ldots, z_p$ of variables after which Lemma 8.2.4 follows from Lemma 8.2.5.

The axioms ZFC-P, ZFC-N, ZFC-U, ZFC-W, ZFC-I, ZFC-S, ZFC-R, ZFC-C and ZFC-D all assert the existence of certain sets. They are all proved in map theory by constructing the set explicitly:

(ZFC-P) (*pairset*)	$(P\,x\,y)$
(ZFC-N) (*emptyset*)	T
(ZFC-U) (*unionset*)	$\lambda z.(x\,(z\,\mathsf{T})\,(z\,\mathsf{F}))$
(ZFC-W) (*powerset*)	$\lambda u.(\mathsf{if}\ u\ \mathsf{T}\ \lambda v.(x\,(u\,(I'\,x\,v))))$
	where $(I'\,x\,y) = \varepsilon z.((x\,z) \dot{=} (x\,y))$
(ZFC-I) (*infinity*)	$(Prim(\lambda w.w)\,\mathsf{T}\,\mathsf{T})$
(ZFC-S) (*subset*)	$(\mathsf{if}\ y\ \mathsf{T}\ (\mathsf{if}\ \dot{\neg}\dot{\exists} u.(\mathscr{B}\,(y\,u))\ \mathsf{T}$
	$\lambda u.(\mathsf{if}\ (\mathscr{B}\,(y\,u))\,(y\,u)\,(y\,\varepsilon u.(\mathscr{B}\,(y\,u))))))$
	where $\mathscr{B} = \lambda x.\mathscr{A}$
(ZFC-R) (*replacement*)	$\lambda x.(u'\,(z\,x))$ where $u' = \lambda x.\varepsilon y.\mathscr{A}$
(ZFC-C) (*choice*)	$\lambda z.(y'\,(x\,z))$ where $y' = \lambda z'.\varepsilon z''.z'' \dot{\in} z'$
(ZFC-D) (*restriction*)	$D(x, \varepsilon w.w \dot{\in} x)$ where $D(x, y)$
	$= (\mathsf{if}\ (\dot{\neg} y \dot{\wedge} \dot{\exists} v.(y\,v) \dot{\in} x)\ D(x, (y\,\varepsilon v.(y\,v) \dot{\in} x))\ y)$.

8.3. Conversational proofs

In the previous sections, the axioms and inference rules of axiomatic map theory have been introduced. In principle, any theorem of map theory is provable using only these axioms and inference rules. In practice, however, proofs become exceedingly long and unreadable if they use only axioms and inference rules. We shall refer to proofs as "formal" if they only use axioms and inference rules.

Proofs become considerably more manageable when use of metatheorems is allowed, but they still remain long and difficult to read.

In this section we shall state proofs in the "conversational" style commonly used between mathematicians. However, since map theory is a new theory such a conversational style has to be developed first.

The aim of a "conversational proof" of a theorem is to present the idea behind the proof in order to convince the skilled reader that it is straightforward to write out a formal proof in all detail.

Trivial details are omitted from conversational proofs. As an example, suppose $\mathscr{A} = \mathscr{B}$ is established in the middle of a proof. Further suppose that $\mathscr{B} = \mathscr{A}$ is needed to continue the proof. In a formal proof, $\mathscr{B} = \mathscr{A}$ has to be proven from $\mathscr{A} = \mathscr{B}$ using a few proof lines. A conversational proof would neither prove $\mathscr{B} = \mathscr{A}$ nor even mention that this is necessary for the proof to proceed.

As an example of a conversational proof, consider the following proof of $x \rightarrow \dot{\forall} y.\phi(x \doteq y)$:

Proof. Assume x and ϕy. Now $x \doteq y = (\text{if } y \text{ T F}) \in \Phi$ so $\phi(x \doteq y)$. □

This proof starts out saying "assume x" which indicates that either $x \rightarrow \mathscr{A}$ or $x \dot{\Rightarrow} \mathscr{A}$ will be proved for some term $\mathscr{A}$ using the deduction theorem and Theorem 5.4.4. Since the proof is going to prove $x \rightarrow \dot{\forall} y.\phi(x \doteq y)$, it is obvious that the goal is to prove $x \rightarrow \mathscr{A}$ where $\mathscr{A}$ is $\dot{\forall} y.\phi(x \doteq y)$.

The proof continues with "assume ϕy" which indicates that $\phi y \rightarrow \mathscr{B}$, $\phi y \dot{\Rightarrow} \mathscr{B}$ or $\dot{\forall} y.\mathscr{B}$ will be proved for some term $\mathscr{B}$ using the deduction theorem, Theorem 5.4.4 or Theorem 6.2.2. Again, it is obvious that the goal is to prove $\dot{\forall} y.\mathscr{B}$ where $\mathscr{B}$ is $\phi(x \doteq y)$. The assumption ϕy could also be stated $y \in \Phi$ since $y \in \Phi$ is shorthand for ϕy in conversational proofs.

Next, the proof states $x \doteq y = (\text{if } y \text{ T F}) \in \Phi$. This is shorthand for $x \doteq y = (\text{if } y \text{ T F})$ and $\phi(\text{if } y \text{ T F})$. The former follows by the reduction theorem if x is replaced by T, and the latter follows from the totality theorem.

The proof concludes $\phi(x \doteq y)$; this follows from $x \doteq y = (\text{if } y \text{ T F})$ and $\phi(\text{if } y \text{ T F})$ by substitutivity, but substitutivity is not mentioned since this is considered trivial.

The proof ends after the conclusion $\phi(x \doteq y)$ even though the goal was to prove $x \rightarrow \dot{\forall} y.\phi(x \doteq y)$. Hence, it is left to the reader to verify $\phi y \rightarrow \phi(x \doteq y)$ by the deduction theorem, then $\dot{\forall} y.\phi(x \doteq y)$ by Theorem 6.2.2, and then $x \rightarrow \dot{\forall} y.\phi(x \doteq y)$ be the deduction theorem. However, the proof started by assuming x and ϕy which implies that this "post processing" is needed.

As another example, consider the following proof of $\dot{\forall} x \dot{\forall} y \dot{\exists} z.(\text{if } (x\,y)\ z\ \dot{\neg} z)$.

Proof. Assume $x, y \in \Phi$. The proof is by *TND* from the following two cases:

Case 1. Assume $(x\,y)$ and let $z = \text{T}$. We have $(\text{if } (x\,y)\ z\ \dot{\neg} z) = z = \text{T}$ as required.

Case 2. Assume $\dot{\neg}(x\,y)$ and let $z = \text{F}$. We have $(\text{if } (x\,y)\ z\ \dot{\neg} z) = \dot{\neg} z = \text{T}$ as required. □

This proof starts out assuming ϕx and ϕy in order to prove $\dot\forall x \dot\forall y.\mathcal{A}$ where $\mathcal{A}$ is $\dot\exists z.(\mathsf{if}\ (x\,y)\ z\ \dot\neg z)$. Next, it is stated that $\dot\exists z.(\mathsf{if}\ (x\,y)\ z\ \dot\neg z)$ will be proved from

$$!(x\,y)$$

$$(x\,y) \to \dot\exists z.(\mathsf{if}\ (x\,y)\ z\ \dot\neg z)$$

$$\dot\neg(x\,y) \to \dot\exists z.(\mathsf{if}\ (x\,y)\ z\ \dot\neg z)$$

using the *TND* theorem. The statement $!(x\,y)$ is not proved explicitly since $\phi(x\,y)$ follows directly from the totality theorem and since $\phi\mathcal{A} \to\ !\mathcal{A}$. In general, verification of premises of the form $\phi\mathcal{A}$ and $!\mathcal{A}$ are omitted whenever the proof is straightforward (but remember these "side conditions" both when reading and writing proofs).

The proofs of $(x\,y) \to \dot\exists z.(\mathsf{if}\ (x\,y)\ z\ \dot\neg z)$ and $\dot\neg(x\,y) \to \dot\exists z.(\mathsf{if}\ (x\,y)\ z\ \dot\neg z)$ are stated under the headings *Case 1* and *Case 2*. In Case 1, $(x\,y)$ is assumed. Then $\dot\exists z.\mathcal{B}$ where $\mathcal{B}$ is $(\mathsf{if}\ (x\,y)\ z\ \dot\neg z)$ is proved by giving a z explicitly. This is the usual way to prove existence in map theory. To make the intention clear, z is locally defined to stand for T and then $(\mathsf{if}\ (x\,y)\ z\ \dot\neg z)$ is verified. It is then left to the reader to apply Theorem 6.2.3 to conclude $\dot\exists z.(\mathsf{if}\ (x\,y)\ z\ \dot\neg z)$. Theorem 6.2.3 requires ϕz and $!\dot\exists z.(\mathsf{if}\ (x\,y)\ z\ \dot\neg z)$, but the proofs of these two requirements are omitted as before.

In Case 2, z is locally defined to stand for F and then the proof is similar to Case 1. The two local definitions $z = \mathsf{T}$ and $z = \mathsf{F}$ do not conflict with each other since they apply to different parts of the proof. It is up to the reader's judgement to understand the scope of each local definition, and it is up to the writer of the proof to allow the reader to guess the scope.

The two examples give the flavor of conversational proofs. Conversational proofs make use of theorems (metatheorems, to be precise) like the deduction theorem and *QND*, but they do not always identify the theorems explicitly. Conversational proofs almost never refer to the reduction, *QND* and totality theorems since they are considered "trivial". Conversational proofs do not refer to the deduction theorem and related theorems either, since the use of "assume" indicate their use.

There are a few constructs of conversational proofs that have not been mentioned. First, "let $x \in \Phi$ satisfy $\mathcal{A}$" does not choose an arbitrary x but chooses the x chosen by ε. In other words, if we say "let $x, y \in \Phi$ satisfy $\mathcal{A}$" then $x = y$. The construct "let $x \in \Phi$ satisfy $\mathcal{A}$" is shorthand for the local definition "let $x = \varepsilon x.\mathcal{A}$" and should only be used when $\dot\exists x.\mathcal{A}$ has been established. When $\dot\exists x.\mathcal{A}$ has been established, then $\mathcal{A} = \mathsf{T}$ follows trivially from $x = \varepsilon x.\mathcal{A}$ and the definition of $\dot\exists$.

Another construct is "The proof is by *QND′* from the following three cases". This indicates that $\mathcal{A} = \mathcal{B}$ is going to be proved from $[\mathcal{A}/x := \mathsf{T}] = [\mathcal{B}/x := \mathsf{T}]$, $[\mathcal{A}/x := \bot] = [\mathcal{B}/x := \bot]$ and $[\mathcal{A}/x := \mathsf{F}'(x)] = [\mathcal{B}/x := \mathsf{F}'(x)]$ for given $\mathcal{A}$ and $\mathcal{B}$.

A third construct is "the proof is by induction in x" where $\dot\forall x.\mathcal{A}$ or $\phi x \to \mathcal{A}$ is proved from $x \to \mathcal{A}$ and $\dot\neg x$, ϕx, $\dot\forall y.((\lambda x.\mathcal{A})\,(x\,y)) \to \mathcal{A}$. The two cases x and $\dot\neg x$, ϕx, $\dot\forall y.((\lambda x.\mathcal{A})\,(x\,y))$ are usually not stated under two headings like *Case 1* and *Case 2*. Rather, the proof looks like: "The proof is by induction in x. First assume $x \ldots$ Now, as inductive hypothesis assume $\dot\neg x$, ϕx, $\dot\forall y.((\lambda x.\mathcal{A})\,(x\,y)) \ldots$"

A fourth and very common construct is to deduce $[\mathscr{A}/x := \mathscr{B}]$ from $\dot{\forall} x.\mathscr{A}$ using Theorem 6.2.1. This requires $\phi\mathscr{B}$ to be proved, but the proof of $\phi\mathscr{B}$ is omitted if it is trivial. Use of this construct also requires $\mathscr{B}$ to be free for x in $\mathscr{A}$.

A fifth construct is the method of indirect proof. If assumption of $\dot{\neg}\mathscr{A}$ leads to a proof of F (i.e., of $\mathsf{F} = \mathsf{T}$), then $\mathscr{A}$ holds according to Theorem 5.4.3 (provided $!\mathscr{A}$ holds, but the proof of $!\mathscr{A}$ is omitted if it is trivial).

The list of constructs that can be used in conversational proofs is open. The sole purpose of a conversational proof is to convince the reader, so whatever convinces the reader may be used in conversational proofs.

8.4. *Trivial lemmas*

The following statements are trivial to prove, and we shall use them without reference in conversational proofs.

Theorem 8.4.1.

(a) $x \dot{\in} y \vdash \dot{\neg} y, \phi u, x \doteq (y\ u)$ *where* $u = \varepsilon u.(x \doteq (y\ u))$.

(b) $y, \phi x \vdash \dot{\neg}(x \dot{\in} y)$.

(c) $\dot{\neg} x, x \doteq y, \phi u \vdash \phi v, (x\ u) \doteq (y\ u)$ *where* $v = \varepsilon v.(x\ u) \doteq (y\ v)$.

(d) $\dot{\neg} x, x \doteq y, \phi v \vdash \phi u, (x\ u) \doteq (y\ v)$ *where* $u = \varepsilon u.(x\ u) \doteq (y\ v)$.

(e) $\dot{\neg} y, x \doteq y, \phi u \vdash \phi v, (x\ u) \doteq (y\ v)$ *where* $v = \varepsilon v.(x\ u) \doteq (y\ v)$.

(f) $\dot{\neg} y, x \doteq y, \phi v \vdash \phi u, (x\ u) \doteq (y\ v)$ *where* $u = \varepsilon u.(x\ u) \doteq (y\ v)$.

(g) $\dot{\forall} x.\dot{\neg}(x \dot{\in} y) \vdash y$.

(h) $\dot{\neg} x, \phi x, \phi u \vdash (x\ u) \dot{\in} x$.

8.5. *The totality of* $\doteq$

Theorem 8.5.1. $\phi x, \phi y \rightarrow \phi(x \doteq y)$.

Proof. The proof is by induction on x. First assume $x = \mathsf{T}$. Further assume $y \in \Phi$. Now $x \doteq y = (\text{if } y\ \mathsf{T}\ \mathsf{F}) \in \Phi$ so $\phi y \rightarrow \phi(x \doteq y)$.

Next, as inductive hypothesis assume ϕx, $\dot{\neg} x$ and $\dot{\forall} u.\dot{\forall} y.\phi((x\ u) \doteq y)$. Further assume $y, u, v \in \Phi$. Since $(y\ v) \in \Phi$ we have $\phi((x\ u) \doteq (y\ u))$. Having $\dot{\forall} u.\dot{\forall} v.\phi((x\ u) \doteq (y\ v))$ the totality theorem gives $\phi\dot{\forall} u.\dot{\exists} v.((x\ u) \doteq (y\ v))$ and $\phi\dot{\forall} v.\dot{\exists} u.((x\ u) \doteq (y\ v))$. Hence, by the definition of $\doteq$ and the totality theorem, $\phi(x \doteq y)$ so $\phi y \rightarrow \phi(x \doteq y)$. The theorem now follows by induction. $\square$

From now on, we regard the above theorem as part of the totality theorem, i.e., the definition of the syntax class Σ of simple terms is extended to

$$\Sigma ::= \cdots \mid \Sigma \doteq \Sigma.$$

As a consequence we have ϕx, $\phi y \to \phi(x \dot{\in} y)$ by the totality theorem. Furthermore, the totality theorem entails the following corollary.

Corollary 8.5.2. *For each* $\mathscr{A}{:}ZFC(x_1, \ldots, x_n)$ *we have* $\phi x_1, \ldots, \phi x_n \to \phi\mathscr{A}$.

8.6. *Equality properties of* $\doteq$

Theorem 8.6.1 (Reflexivity). *If* ϕx *then* $x \doteq x$.

Proof. The proof is by induction on x. For $x = \mathsf{T}$, $x \doteq x$ holds trivially. Now, as inductive hypothesis assume ϕx, $\dot{\neg} x$ and $\dot{\forall} u.(x\,u) \doteq (x\,u)$.

If ϕu then $(x\,u) \doteq (x\,u)$. Hence, $\dot{\forall} u.\dot{\exists} v.(x\,u) \doteq (x\,v)$ and $\dot{\forall} v.\dot{\exists} u.(x\,u) \doteq (x\,v)$. Combined with $\dot{\neg} x$ this gives $x \doteq x$. The theorem now follows by induction. $\square$

Theorem 8.6.2 (Transitivity). *If* ϕx, ϕy, ϕz, $x \doteq y$ *and* $x \doteq z$ *then* $y \doteq z$.

Proof. We first prove $\dot{\forall} x.\dot{\forall} y.\dot{\forall} z.(x \doteq y \mathbin{\dot{\wedge}} x \doteq z \mathbin{\dot{\Rightarrow}} y \doteq z)$ by induction on x. For $x = \mathsf{T}$ assume $y, z \in \Phi$. Now $x \doteq y \mathbin{\dot{\wedge}} x \doteq z \mathbin{\dot{\Rightarrow}} y \doteq z$ by *TND* so $\dot{\forall} y.\dot{\forall} z.(x \doteq y \mathbin{\dot{\wedge}} x \doteq z \mathbin{\dot{\Rightarrow}} y \doteq z)$.

As inductive hypothesis assume ϕx, $\dot{\neg} x$ and $\dot{\forall} u.\dot{\forall} y.\dot{\forall} z.((x\,u) \doteq y \mathbin{\dot{\wedge}} (x\,u) \doteq z \mathbin{\dot{\Rightarrow}} y \doteq z)$. Further assume $u, y, z \in \Phi$, $x \doteq y$ and $x \doteq z$. Choose $u' \in \Phi$ such that $(x\,u') \doteq (y\,u)$ and choose $u'' \in \Phi$ such that $(x\,u') \doteq (z\,u'')$. As a special case of the inductive hypothesis we have $(x\,u') \doteq (y\,u) \mathbin{\dot{\wedge}} (x\,u') \doteq (z\,u'') \mathbin{\dot{\Rightarrow}} (y\,u) \doteq (z\,u'')$, so $(y\,u) \doteq (z\,u'')$. Hence, $\dot{\forall} u.\dot{\exists} v.(y\,u) \doteq (z\,v)$. Likewise, $\dot{\forall} v.\dot{\exists} u.(y\,u) \doteq (z\,v)$. Hence, from the definition of $\doteq$, $y \doteq z$.

Now, $\dot{\forall} x.\dot{\forall} y.\dot{\forall} z.(x \doteq y \mathbin{\dot{\wedge}} x \doteq z \mathbin{\dot{\Rightarrow}} y \doteq z)$ follows by induction on x. The theorem follows easily. $\square$

The theorems of reflexivity and transitivity allow us to treat $\doteq$ as an equivalence relation.

Theorem 8.6.3 (Substitutivity). *If* ϕx, ϕy, ϕz *and* $x \doteq y$ *then* $x \dot{\in} z \mathbin{\dot{\Leftrightarrow}} y \dot{\in} z$ *and* $z \dot{\in} x \mathbin{\dot{\Leftrightarrow}} z \dot{\in} y$.

Proof. Assume ϕx, ϕy, ϕz and $x \doteq y$.

- Proof of $x \dot{\in} z \mathbin{\dot{\Leftrightarrow}} y \dot{\in} z$. Assume $x \dot{\in} z$. From $x \dot{\in} z$ we have $\dot{\neg} z$. Choose $u \in \Phi$ such that $x \doteq (z\,u)$. By transitivity of $\doteq$ we have $y \doteq (z\,u)$, so $y \dot{\in} z$. Hence, $x \dot{\in} z \mathbin{\dot{\Rightarrow}} y \dot{\in} z$. Likewise, $y \dot{\in} z \mathbin{\dot{\Rightarrow}} x \dot{\in} z$, so $x \dot{\in} z \mathbin{\dot{\Leftrightarrow}} y \dot{\in} z$.
- Proof of $z \dot{\in} x \mathbin{\dot{\Leftrightarrow}} z \dot{\in} y$. Assume $z \dot{\in} x$. From $z \dot{\in} x$ we have $\dot{\neg} x$, and from $x \doteq y$ we have $\dot{\neg} y$. Choose $u \in \Phi$ such that $z \doteq (x\,u)$. Choose $v \in \Phi$ such that $(x\,u) \doteq (y\,v)$. By transitivity of $\doteq$ we have $z \doteq (y\,v)$, so $z \dot{\in} y$. Hence, $z \dot{\in} x \mathbin{\dot{\Rightarrow}} z \dot{\in} y$. Likewise, $z \dot{\in} y \mathbin{\dot{\Rightarrow}} z \dot{\in} x$, so $z \dot{\in} x \mathbin{\dot{\Leftrightarrow}} z \dot{\in} y$. $\square$

Corollary 8.6.4 (Substitutivity in *ZFC*). *If* $\mathscr{A}:ZFC(x, x_1, \ldots, x_n)$, ϕx, ϕy, $\phi x_1, \ldots, \phi x_n$ *and* $x \doteq y$ *then* $\mathscr{A} \Leftrightarrow [\mathscr{A}/x := y]$.

The corollary allows us to treat $\doteq$ as an equality relation when dealing with well-formed formulas $\mathscr{A}$ of *ZFC*.

8.7. Extensionality

Theorem 8.7.1 (Extensionality). *If* ϕx, ϕy *and* $\dot\forall z.(z \mathrel{\dot\in} x \mathrel{\dot\Leftrightarrow} z \mathrel{\dot\in} y)$ *then* $x \doteq y$.

Proof. Assume ϕx, ϕy and $\dot\forall z.(z \mathrel{\dot\in} x \mathrel{\dot\Leftrightarrow} z \mathrel{\dot\in} y)$. The proof is by *TND*.

Case 1. Assume x. Further assume ϕz. We have $z \mathrel{\dot\in} x \mathrel{\dot\Leftrightarrow} z \mathrel{\dot\in} y$. Since $x = \mathsf{T}$ we have $\dot\neg z \mathrel{\dot\in} x$ so $\dot\neg z \mathrel{\dot\in} y$. Hence, $\dot\forall z.\dot\neg z \mathrel{\dot\in} y$ so $\dot\neg\dot\exists z.z \mathrel{\dot\in} y$ and $y = \mathsf{T}$ according to Theorem 8.4.1. From $x = y = \mathsf{T}$ we have $x \doteq y$.

Case 2. Assume $\dot\neg x$. Further assume $u \in \Phi$. We have $(x\,u) \mathrel{\dot\in} x$. Further, $(x\,u) \in \Phi$ so $(x\,u) \mathrel{\dot\in} x \mathrel{\dot\Leftrightarrow} (x\,u) \mathrel{\dot\in} y$ and $(x\,u) \mathrel{\dot\in} y$. From $(x\,u) \mathrel{\dot\in} y$ we have $\dot\exists v.(x\,u) \doteq (y\,v)$. Hence, $\dot\forall u.\dot\exists v.(x\,u) \doteq (y\,v)$. Likewise, $\dot\forall v.\dot\exists u.(x\,u) \doteq (y\,v)$, so $x \doteq y$. □

8.8. Logical axioms and inference rules

Theorem 8.8.1. *Let* $\mathscr{A}$, $\mathscr{B}$, $\mathscr{C}:ZFC(x_1, \ldots, x_n)$. *We have*

(ZFC-A1) $\quad \phi x_1, \ldots, \phi x_n \to \mathscr{A} \mathrel{\dot\Rightarrow} (\mathscr{B} \mathrel{\dot\Rightarrow} \mathscr{A})$

(ZFC-A2) $\quad \phi x_1, \ldots, \phi x_n \to (\mathscr{A} \mathrel{\dot\Rightarrow} (\mathscr{B} \mathrel{\dot\Rightarrow} \mathscr{C})) \mathrel{\dot\Rightarrow} ((\mathscr{A} \mathrel{\dot\Rightarrow} \mathscr{B}) \mathrel{\dot\Rightarrow} (\mathscr{A} \mathrel{\dot\Rightarrow} \mathscr{C}))$

(ZFC-A3) $\quad \phi x_1, \ldots, \phi x_n \to (\dot\neg\mathscr{B} \mathrel{\dot\Rightarrow} \dot\neg\mathscr{A}) \mathrel{\dot\Rightarrow} ((\dot\neg\mathscr{B} \mathrel{\dot\Rightarrow} \mathscr{A}) \mathrel{\dot\Rightarrow} \mathscr{B})$.

Proof. Assume $\phi x_1, \ldots, \phi x_n$. Now $\phi\mathscr{A}$, $\phi\mathscr{B}$ and $\phi\mathscr{C}$ according to Corollary 8.5.2, so $\mathscr{A} \mathrel{\dot\Rightarrow} (\mathscr{B} \mathrel{\dot\Rightarrow} \mathscr{A})$ and the other two statements are trivial to prove by *TND*. □

Theorem 8.8.2 (ZFC-A4). *Let* $\mathscr{A}:ZFC(x_1, \ldots, x_n)$. *Let* $i, j \in \{1, \ldots, n\}$, *and assume that* x_i *is free for* x_j *in* $\mathscr{A}$. *We have*

$$\phi x_1, \ldots, \phi x_n \to \dot\forall x_i.\mathscr{A} \mathrel{\dot\Rightarrow} [\mathscr{A}/x_i := x_j].$$

Proof. Assume $\phi x_1, \ldots, \phi x_n$, and $\dot\forall x_i.\mathscr{A}$. Now, $[\mathscr{A}/x_i := x_j]$ follows by Theorem 6.2.1. □

Theorem 8.8.3 (ZFC-A5). *Let* $\mathscr{A}:ZFC(x_1, \ldots, x_n)$ *and let* $\mathscr{B}:ZFC(x, x_1, \ldots, x_n)$ *where* x *does not occur among* $x_1, \ldots, x_n$. *We have*

$$\phi x_1, \ldots, \phi x_n \to \dot\forall x.(\mathscr{A} \mathrel{\dot\Rightarrow} \mathscr{B}) \mathrel{\dot\Rightarrow} (\mathscr{A} \mathrel{\dot\Rightarrow} \dot\forall x.\mathscr{B}).$$

Proof. Assume $\phi x_1, \ldots, \phi x_n$. Corollary 8.5.2 gives $\phi\mathscr{A}$. Further, if ϕx, then $\phi\mathscr{B}$ so $\phi\dot{\forall}x.\mathscr{B}$. The proof proceeds by *TND* in $\mathscr{A}$ and $\dot{\forall}x.\mathscr{B}$.

- If $\mathscr{A}$ and $\dot{\forall}x.\mathscr{B}$ then $\dot{\forall}x.(\mathscr{A} \dot{\Rightarrow} \mathscr{B}) \dot{\Rightarrow} (\mathscr{A} \dot{\Rightarrow} \dot{\forall}x.\mathscr{B}) = \dot{\forall}x.\mathscr{B} \dot{\Rightarrow} \dot{\forall}x.\mathscr{B} = \mathsf{T}$.
- If $\mathscr{A}$ and $\dot{\neg}\dot{\forall}x.\mathscr{B}$ then $\dot{\forall}x.(\mathscr{A} \dot{\Rightarrow} \mathscr{B}) \dot{\Rightarrow} (\mathscr{A} \dot{\Rightarrow} \dot{\forall}x.\mathscr{B}) = \dot{\forall}x.\mathscr{B} \dot{\Rightarrow} \dot{\forall}x.\mathscr{B} = \mathsf{T}$.
- If $\dot{\neg}\mathscr{A}$ and $\dot{\forall}x.\mathscr{B}$ then $\dot{\forall}x.(\mathscr{A} \dot{\Rightarrow} \mathscr{B}) \dot{\Rightarrow} (\mathscr{A} \dot{\Rightarrow} \dot{\forall}x.\mathscr{B}) = \dot{\forall}x.\mathsf{F} \dot{\Rightarrow} \dot{\forall}x.\mathsf{F} = \mathsf{T}$.
- If $\dot{\neg}\mathscr{A}$ and $\dot{\neg}\dot{\forall}x.\mathscr{B}$ then $\dot{\forall}x.(\mathscr{A} \dot{\Rightarrow} \mathscr{B}) \dot{\Rightarrow} (\mathscr{A} \dot{\Rightarrow} \dot{\forall}x.\mathscr{B}) = \dot{\forall}x.\mathsf{F} \dot{\Rightarrow} \dot{\forall}x.\mathsf{F} = \mathsf{T}$.

Theorem 8.8.4 (ZFC-MP). *Let $\mathscr{A}$, $\mathscr{B}$:$ZFC(x_1, \ldots, x_n)$. We have*

$$\phi x_1, \ldots, \phi x_n \to \mathscr{A}\,; \phi x_1, \ldots, \phi x_n \to \mathscr{A} \dot{\Rightarrow} \mathscr{B} \;\vdash\; \phi x_1, \ldots, \phi x_n \to \mathscr{B}.$$

Proof. Assume $\phi x_1, \ldots, \phi x_n \to \mathscr{A}$, $\phi x_1, \ldots, \phi x_n \to \mathscr{A} \dot{\Rightarrow} \mathscr{B}$, and $\phi x_1, \ldots, \phi x_n$. Now, $\mathscr{A}$ and $\mathscr{A} \dot{\Rightarrow} \mathscr{B}$ holds, which logically implies $\mathscr{B}$. Hence, $\phi x_1, \ldots, \phi x_n \to \mathscr{B}$. □

Theorem 8.8.5 (ZFC-Gen). *Let $\mathscr{A}$:$ZFC(x_1, \ldots, x_n)$ and let $i \in \{1, \ldots, n\}$. We have*

$$\phi x_1, \ldots, \phi x_n \to \mathscr{A} \;\vdash\; \phi x_1, \ldots, \phi x_n \to \dot{\forall}x_i.\mathscr{A}.$$

Proof. This theorem was previously stated and proved in Lemma 7.5.4. Here is a conversational proof: Assume $\phi x_1, \ldots, \phi x_{i-1}$, ϕy, $\phi x_{i+1}, \ldots, \phi x_n$, ϕx_i. From $\phi x_1, \ldots, \phi x_n$ we have $\mathscr{A}$, so $\dot{\forall}x_i.\mathscr{A}$ and $\phi x_1, \ldots, \phi x_{i-1}$, ϕy, $\phi x_{i+1}, \ldots, \phi x_n \to \dot{\forall}x_i.\mathscr{A}$. The theorem follows by renaming y into x_i. □

8.9. The axiom of extensionality

Theorem 8.9.1 (ZFC-E). ϕx, $\phi y \to \dot{\forall}z.(z \mathbin{\dot{\in}} x \dot{\Leftrightarrow} z \mathbin{\dot{\in}} y) \dot{\Rightarrow} \dot{\forall}z.(x \mathbin{\dot{\in}} z \dot{\Leftrightarrow} y \mathbin{\dot{\in}} z)$.

Proof. Assume ϕx, ϕy, $\dot{\forall}z.(z \mathbin{\dot{\in}} x \dot{\Leftrightarrow} z \mathbin{\dot{\in}} y)$ and ϕz. From Theorem 8.7.1 we have $x \doteq y$. Assume $x \mathbin{\dot{\in}} z$ and choose $u \in \Phi$ such that $x \doteq (z\, u)$. We now have $y \doteq x \doteq (z\, u)$ so $y \mathbin{\dot{\in}} z$. Hence, $x \mathbin{\dot{\in}} z \dot{\Rightarrow} y \mathbin{\dot{\in}} z$. Likewise, $y \mathbin{\dot{\in}} z \dot{\Rightarrow} x \mathbin{\dot{\in}} z$ so $x \mathbin{\dot{\in}} z \dot{\Leftrightarrow} y \mathbin{\dot{\in}} z$ as required. □

8.10. Construction axioms of ZFC

Theorem 8.10.1 (ZFC-P). $\dot{\forall}x.\dot{\forall}y.\dot{\exists}z.x \mathbin{\dot{\in}} z \mathbin{\dot{\wedge}} y \mathbin{\dot{\in}} z$.

Proof. Assume $x, y \in \Phi$ and let $z = (P\,x\,y)$. From $(z\,\mathsf{T}) = (P\,x\,y\,\mathsf{T}) = x$ and $(x\,\mathsf{F}) = (P\,x\,y\,\mathsf{F}) = y$ we have $x \doteq (z\,\mathsf{T})$ and $y \doteq (z\,\mathsf{F})$, so $x \mathbin{\dot{\in}} z$ and $y \mathbin{\dot{\in}} z$. □

Theorem 8.10.2 (ZFC-N). $\dot{\exists}x.\dot{\forall}y.\dot{\neg}(y \mathbin{\dot{\in}} x)$.

Proof. Take $x = \mathsf{T}$. □

Theorem 8.10.3 (ZFC-U). $\dot{\forall}x.\dot{\exists}y.\dot{\forall}u.\dot{\forall}v.(u \mathbin{\dot{\in}} v \mathbin{\dot{\wedge}} v \mathbin{\dot{\in}} x \dot{\Rightarrow} u \mathbin{\dot{\in}} y)$.

Proof. Assume ϕx and let $y = \lambda z.(x\,(z\,\mathsf{T})\,(z\,\mathsf{F}))$. Assume ϕu, ϕv, $u \mathbin{\dot{\in}} v$ and $v \mathbin{\dot{\in}} x$. From $u \mathbin{\dot{\in}} v$ and $v \mathbin{\dot{\in}} x$ we have $\dot{\neg}v$ and $\dot{\neg}x$. Choose v' such that $v \doteq (x\,v')$. From $u \mathbin{\dot{\in}} v$

and $v \doteq (x\, v')$ we have $u \mathrel{\dot\in} (x\, v')$. Choose u' such that $u \doteq (x\, v'\, u')$ and let $z = (P\, v'\, u')$. Now ϕz and $(y\, z) = (x\, (z\, \mathsf{T})\, (z\, \mathsf{F})) = (x\, v'\, u') \doteq u$ so $u \mathrel{\dot\in} y$ as required. □

Intuitively, $\lambda u.(\text{if } u\ \mathsf{T}\ \lambda v.(x\, (u\, (x\, v))))$ represents the power set of x. However, the given axiomatization of map theory seems to be insufficient to prove this. Instead, we represent the power set of x by

$$\lambda u.(\text{if } u\ \mathsf{T}\ \lambda v.(x\, (u\, (I'\, x\, v))))$$

where I' is "almost" an identity relation. We define I' (conditional identity) by

$$I' = \lambda x.\lambda y.\varepsilon z.(x\, z) \doteq (x\, y).$$

The following lemma expresses the "almost identity" property of I'.

Lemma 8.10.4 (I'). *If ϕx and ϕy then $(x\, (I'\, x\, y)) \doteq (x\, y)$.*

Hence, $(I'\, x\, y)$ behaves like y when occurring as the argument of x.

Proof. Assume ϕx and ϕy and let $z = (I'\, x\, y) = \varepsilon z.(x\, z) \doteq (x\, y)$. From $(x\, y) \doteq (x\, y)$ we have $\dot\exists u.(x\, u) \doteq (x\, y)$, so $(x\, z) \doteq (y\, z)$ as required. □

Note that $\phi x \rightarrow \phi(I'\, x)$. This is important when using the totality theorem.

Theorem 8.10.5 (ZFC-W). $\dot\forall x.\dot\exists y.\dot\forall z.(\dot\forall u.(u \mathrel{\dot\in} z \mathrel{\dot\Rightarrow} u \mathrel{\dot\in} x) \mathrel{\dot\Rightarrow} z \mathrel{\dot\in} y)$.

Proof. Assume ϕx and let $y = \lambda u.(\text{if } u\ \mathsf{T}\ \lambda v.(x\, (u\, (I'\, x\, v))))$. Assume ϕz and $\dot\forall u.(u \mathrel{\dot\in} z \mathrel{\dot\Rightarrow} u \mathrel{\dot\in} x)$. We shall prove $z \mathrel{\dot\in} y$ and do so by *TND* from the following two cases.

Case 1. Assume z. From $(y\, \mathsf{T}) = \mathsf{T}$ and $z = \mathsf{T}$ we have $z \doteq (y\, \mathsf{T})$, so $z \mathrel{\dot\in} y$ as required.

Case 2. Assume $\dot\neg z$. Define $u = \lambda v.(\text{if } ((x\, v) \mathrel{\dot\in} z)\ (I'\, x\, v)\ (\varepsilon v.(x\, v) \mathrel{\dot\in} z))$. To prove $z \mathrel{\dot\in} y$ it is sufficient to prove $z \doteq (y\, u)$, and to prove $z \doteq (y\, u)$ it is sufficient to prove $v \mathrel{\dot\in} z \mathrel{\dot\Leftrightarrow} v \mathrel{\dot\in} (y\, u)$ for all $v \in \Phi$. Hence, assume ϕv. Note that the definitions of y and u gives

$$(y\, u) = \lambda v.(x\, (\text{if } ((x\, v) \mathrel{\dot\in} z)\ (I'\, x\, v)\ (\varepsilon v.(x\, v) \mathrel{\dot\in} z))).$$

Proof of $v \mathrel{\dot\in} z \mathrel{\dot\Rightarrow} v \mathrel{\dot\in} (y\, u)$. Assume $v \mathrel{\dot\in} z$. From $\dot\forall u.(u \mathrel{\dot\in} z \mathrel{\dot\Rightarrow} u \mathrel{\dot\in} x)$ we have $v \mathrel{\dot\in} x$ so $\dot\neg x$. Choose v' such that $v \doteq (x\, v')$. From $v \doteq (x\, v')$ and $v \mathrel{\dot\in} z$ we have $(x\, v') \mathrel{\dot\in} z$ so $(y\, u\, v') = (x\, (I'\, x\, v')) \doteq (x\, v') \doteq v$ and $v \mathrel{\dot\in} (y\, u)$.

Proof of $v \mathrel{\dot\in} (y\, u) \mathrel{\dot\Rightarrow} v \mathrel{\dot\in} z$. Assume $v \mathrel{\dot\in} (y\, u)$ and choose $v' \in \Phi$ such that $v \doteq (y\, u\, v')$. We now prove $v \mathrel{\dot\in} z$ by *TND* from two cases.

Case a. Assume $(x\ v') \dot{\in} z$. In this case $v \doteq (y\ u\ v') = (x\ (I'\ x\ v')) \doteq (x\ v') \dot{\in} z$ so $v \dot{\in} z$.

Case b. Assume $\dot{\neg}(x\ v') \dot{\in} z$. In this case $v \doteq (y\ u\ v') \doteq (x\ \varepsilon v.(x\ v) \dot{\in} z)$. From $\dot{\neg} z$ we have $(z\ \mathsf{T}) \dot{\in} z$, and from $\dot{\forall} u.(u \dot{\in} z \dot{\Rightarrow} u \dot{\in} x)$ we have $(z\ \mathsf{T}) \dot{\in} x$. Choose v'' such that $(z\ \mathsf{T}) \doteq (x\ v'')$. From $(x\ v'') \doteq (z\ \mathsf{T})$ we have $(x\ v'') \dot{\in} z$ so $\dot{\exists} v.(x\ v) \dot{\in} z$. Hence, $v \doteq (x\ \varepsilon v.(x\ v) \dot{\in} z) \dot{\in} z$ which proves $v \dot{\in} z$ as required. □

Theorem 8.10.6 (ZFC-I). $\dot{\exists} y.(\dot{\exists} z.z \dot{\in} y \dot{\wedge} \dot{\forall} z.(z \dot{\in} y \dot{\Rightarrow} \dot{\exists} u.z \dot{\in} u \dot{\wedge} u \dot{\in} y))$.

Proof. Let $y = (\mathit{Prim}(\lambda w.w)\ \mathsf{T}\ \mathsf{T})$. We have ϕy, $\dot{\neg} y$, and

$$(y\ v) = (\mathsf{if}\ v\ \mathsf{T}\ \lambda w.(y\ (v\ \mathsf{T}))).$$

Since $(y\ \mathsf{T}) = \mathsf{T}$ we have $\mathsf{T} \dot{\in} y$ so $\dot{\exists} z.z \dot{\in} y$. Now assume ϕz and $z \dot{\in} y$. Choose $v \in \Phi$ such that $z \doteq (y\ v)$ and define $u = \lambda w.(y\ v)$ and $v' = \lambda w.v$. From $(y\ v') = \lambda w.(y\ (v'\ \mathsf{T})) = \lambda w.(y\ v) = u$ we have $u \dot{\in} y$ and from $z \doteq (y\ v) = (u\ \mathsf{T})$ we have $z \dot{\in} u$, so $\dot{\exists} u.(z \dot{\in} u \dot{\wedge} u \dot{\in} y)$. □

Theorem 8.10.7 (ZFC-S). *If* $\mathscr{A}: ZFC(x, x_1, \ldots, x_n)$ *then*

$$\phi x_1, \ldots, \phi x_n \rightarrow \dot{\forall} y.\dot{\exists} z.\dot{\forall} x.(x \dot{\in} z \dot{\Leftrightarrow} x \dot{\in} y \dot{\wedge} \mathscr{A}).$$

Proof. Assume $\phi x_1, \ldots, \phi x_n$ and ϕy. Let $\mathscr{B}$ be shorthand for $\lambda x.\mathscr{A}$ and define

$$z = (\mathsf{if}\ y\ \mathsf{T}\ (\mathsf{if}\ \dot{\neg}\dot{\exists} u.(\mathscr{B}\ (y\ u))\ \mathsf{T}\ \lambda y.(\mathsf{if}\ (\mathscr{B}\ (y\ u))\ (y\ u)\ (y\ \varepsilon u.(\mathscr{B}\ (y\ u)))))).$$

We have ϕz. Assume ϕx. We shall prove $x \dot{\in} z \dot{\Leftrightarrow} x \dot{\in} y \dot{\wedge} \mathscr{A}$.

Proof of $x \dot{\in} z \dot{\Rightarrow} x \dot{\in} y \dot{\wedge} \mathscr{A}$. Assume $x \dot{\in} z$, choose $u \in \Phi$ such that $x \doteq (z\ u)$, and define $v = (\mathsf{if}\ (\mathscr{B}\ (y\ u))\ u\ \varepsilon u.(\mathscr{B}\ (y\ u)))$. Since $\mathscr{A}$ is a term of *ZFC* we have $\phi x, \phi x_1, \ldots, \phi x_n \rightarrow \phi \mathscr{A}$. In particular, since $\phi x_1, \ldots, \phi x_n$ and ϕy holds, we have $\phi w \rightarrow \phi(\mathscr{B}\ (y\ w))$. If $y = \mathsf{T}$ or if $\dot{\neg}\dot{\exists} x.(\mathscr{B}\ (y\ x))$ then $z = \mathsf{T}$ contradicting $x \dot{\in} z$. Hence, $\dot{\neg} y$ and $\dot{\exists} x.(\mathscr{B}\ (y\ x))$. If $(\mathscr{B}\ (y\ u)) = \mathsf{T}$ then $x \doteq (z\ u) = (y\ u)$ so $x \dot{\in} y$. Further, $\mathscr{A} = (\mathscr{B}\ x) = (\mathscr{B}\ (z\ u)) = (\mathscr{B}\ (y\ u)) = \mathsf{T}$. If $\dot{\neg}(\mathscr{B}\ (y\ u))$ then $x \doteq (z\ u) = (y\ \varepsilon u.(\mathscr{B}\ (y\ u)))$ so $x \dot{\in} y$. Further, since $\dot{\exists} u.(\mathscr{B}\ (y\ u))$ we have $\mathscr{A} = (\mathscr{B}\ x) = (\mathscr{B}\ (z\ u)) = (\mathscr{B}\ (y\ \varepsilon u.(\mathscr{B}\ (y\ u)))) = \mathsf{T}$. In any case, $x \dot{\in} y \dot{\wedge} \mathscr{A}$ as required.

Proof of $x \dot{\in} y \dot{\wedge} \mathscr{A} \dot{\Rightarrow} x \dot{\in} z$. Assume $x \dot{\in} y$ and $\mathscr{A}$. Choose $u \in \Phi$ such that $x \doteq (y\ u)$. We have $(\mathscr{B}\ (y\ u)) = (\mathscr{B}\ x) = \mathscr{A} = \mathsf{T}$. Hence, $\dot{\exists} u.(\mathscr{B}\ (y\ u))$. The definition of z combined with $\dot{\neg} y$, $\dot{\exists} u.(\mathscr{B}\ (y\ u))$ and $(\mathscr{B}\ (y\ u))$ gives $(z\ u) = (y\ u) \doteq x$ so $x \dot{\in} z$ as required. □

Theorem 8.10.8 (ZFC-R). *If* $\mathscr{A}: ZFC(x, y, x_1, \ldots, x_n)$ *then*

$$\phi x_1, \ldots, \phi x_n \rightarrow \dot{\forall} z.\dot{\exists} u.\dot{\forall} x.(x \dot{\in} z \dot{\wedge} \dot{\exists} y.\mathscr{A} \dot{\Rightarrow} \dot{\exists} y.y \dot{\in} u \dot{\wedge} \mathscr{A}).$$

Proof. Assume $\phi x_1, \ldots, \phi x_n$ and ϕz. Define $u' = \lambda x.\varepsilon y.\mathscr{A}$ and $u = \lambda x.(u'\,(z\,x))$. We have ϕu. Assume ϕx, $x \dot{\in} z$ and $\dot{\exists} y.\mathscr{A}$. Choose $v, y \in \Phi$ such that $x \doteq (z\,v)$ and $\mathscr{A} = \mathsf{T}$. Now $(u\,v) = (u'\,(z\,v)) = \varepsilon y.((\lambda x.\mathscr{A})\,(z\,v)) = \varepsilon y.((\lambda x.\mathscr{A})\,x) = \varepsilon y.\mathscr{A} = y$ so $y \dot{\in} u$. Hence, $y \dot{\in} u \mathrel{\dot{\wedge}} \mathscr{A}$ as required. (Recall that "choose $y \in \Phi$ such that $\mathscr{A} = \mathsf{T}$" means "let $y = \varepsilon y.\mathscr{A}$". □

Theorem 8.10.9 (ZFC-C).

$$\begin{aligned}
&\dot{\forall} x.(\dot{\forall} u.(u \dot{\in} x \mathrel{\dot{\Rightarrow}} \dot{\exists} v.v \dot{\in} u)\\
&\qquad \mathrel{\dot{\wedge}} \dot{\forall} u, v, w, z.(u \dot{\in} x \mathrel{\dot{\wedge}} v \dot{\in} x \mathrel{\dot{\wedge}} w \dot{\in} u \mathrel{\dot{\wedge}} w \dot{\in} v \mathrel{\dot{\wedge}} z \dot{\in} u \mathrel{\dot{\Rightarrow}} z \dot{\in} v)\\
&\quad \mathrel{\dot{\Rightarrow}} \dot{\exists} y.\dot{\forall} u.(u \dot{\in} x \mathrel{\dot{\Rightarrow}} \dot{\exists} v.v \dot{\in} u \mathrel{\dot{\wedge}} v \dot{\in} y)\\
&\qquad \mathrel{\dot{\wedge}} \dot{\forall} u, v, w, z.(u \dot{\in} x \mathrel{\dot{\wedge}} v \dot{\in} u \mathrel{\dot{\wedge}} v \dot{\in} y \mathrel{\dot{\wedge}} w \dot{\in} u \mathrel{\dot{\wedge}} w \dot{\in} y \mathrel{\dot{\wedge}} z \dot{\in} v \mathrel{\dot{\Rightarrow}} z \dot{\in} w)).
\end{aligned}$$

Proof. Assume ϕx, $\dot{\forall} u.(u \dot{\in} x \mathrel{\dot{\Rightarrow}} \dot{\exists} v.v \dot{\in} u)$ and $\dot{\forall} u, v, w, z.(u \dot{\in} x \mathrel{\dot{\wedge}} v \dot{\in} x \mathrel{\dot{\wedge}} w \dot{\in} u \mathrel{\dot{\wedge}} w \dot{\in} v \mathrel{\dot{\wedge}} z \dot{\in} u \mathrel{\dot{\Rightarrow}} z \dot{\in} v)$. Define $y' = \lambda z'.\varepsilon z''.z'' \dot{\in} z'$ and $y = \lambda z.(y'\,(x\,z))$. We have ϕy. To prove the theorem we have to prove $\dot{\forall} u.(u \dot{\in} x \mathrel{\dot{\Rightarrow}} \dot{\exists} v.v \dot{\in} u \mathrel{\dot{\wedge}} v \dot{\in} y)$ and $\dot{\forall} u, v, w, z.(u \dot{\in} x \mathrel{\dot{\wedge}} v \dot{\in} u \mathrel{\dot{\wedge}} v \dot{\in} y \mathrel{\dot{\wedge}} w \dot{\in} u \mathrel{\dot{\wedge}} w \dot{\in} y \mathrel{\dot{\wedge}} z \dot{\in} v \mathrel{\dot{\Rightarrow}} z \dot{\in} w)$. Note that

$$(y\,\mathscr{A}) = (y'\,(x\,\mathscr{A})) = \varepsilon z''.z'' \dot{\in} (x\,\mathscr{A})$$

for any term $\mathscr{A}$ that does not contain z'' free.

Proof of $\dot{\forall} u.(u \dot{\in} x \mathrel{\dot{\Rightarrow}} \dot{\exists} v.v \dot{\in} u \mathrel{\dot{\wedge}} v \dot{\in} y)$. Assume ϕu and $u \dot{\in} x$. Choose $u' \in \Phi$ such that $u \doteq (x\,u')$ and define $v = (y\,u') = \varepsilon z''.z'' \dot{\in} (x\,u')$. From $v = (y\,u')$ we have $v \dot{\in} y$ as required. We now prove $v \dot{\in} u$. From $(x\,u') \dot{\in} x$ and $\dot{\forall} u.(u \dot{\in} x \mathrel{\dot{\Rightarrow}} \dot{\exists} v.v \dot{\in} u)$ we have $\dot{\exists} v.v \dot{\in} (x\,u')$. Since $v = \varepsilon z''.z'' \dot{\in} (x\,u')$ we have $v \dot{\in} (x\,u')$ which combined with $u \doteq (x\,u')$ gives $v \dot{\in} u$.

Proof of $\dot{\forall} u, v, w, z.(u \dot{\in} x \mathrel{\dot{\wedge}} v \dot{\in} u \mathrel{\dot{\wedge}} v \dot{\in} y \mathrel{\dot{\wedge}} w \dot{\in} u \mathrel{\dot{\wedge}} w \dot{\in} y \mathrel{\dot{\wedge}} z \dot{\in} v \mathrel{\dot{\Rightarrow}} z \dot{\in} w)$. Assume ϕu, ϕv, ϕw, ϕz, $u \dot{\in} x$, $v \dot{\in} u$, $v \dot{\in} y$, $w \dot{\in} u$, $w \dot{\in} y$ and $z \dot{\in} v$. We shall prove $z \dot{\in} w$. Choose $u', v', w' \in \Phi$ such that $u \doteq (x\,u')$, $v \doteq (y\,v')$ and $w \doteq (y\,w')$.

Let $v'' = (y\,v') = \varepsilon z''.z'' \dot{\in} (x\,v')$. From $(x\,v') \dot{\in} x$ and the assumption $\dot{\forall} u.(u \dot{\in} x \mathrel{\dot{\Rightarrow}} \dot{\exists} v.v \dot{\in} u)$ we have $\dot{\exists} z''.z'' \dot{\in} (x\,v')$ so $v'' \dot{\in} (x\,v')$. Combined with $v'' = (y\,v') \doteq v$ this gives $v \dot{\in} (x\,v')$. Likewise, $w \dot{\in} (x\,w')$.

We now prove $z' \dot{\in} (x\,v') \mathrel{\dot{\Leftrightarrow}} z' \dot{\in} (x\,w')$ for all $z' \in \Phi$. Assume $\phi z'$ and $z' \dot{\in} (x\,v')$. The assumption $\dot{\forall} u, v, w, z.(u \dot{\in} x \mathrel{\dot{\wedge}} v \dot{\in} x \mathrel{\dot{\wedge}} w \dot{\in} u \mathrel{\dot{\wedge}} w \dot{\in} v \mathrel{\dot{\wedge}} z \dot{\in} u \mathrel{\dot{\Rightarrow}} z \dot{\in} v)$ combined with $(x\,v') \dot{\in} x$, $u \dot{\in} x$, $v \dot{\in} (x\,v')$, $v \dot{\in} u$ and $z' \dot{\in} (x\,v')$ yields $z' \dot{\in} u$. The same assumption combined with $u \dot{\in} x$, $(x\,w') \dot{\in} x$, $w \dot{\in} u$, $w \dot{\in} (x\,w')$ and $z' \dot{\in} u$ yields $z' \dot{\in} (x\,w')$. Hence, $z' \dot{\in} (x\,v') \mathrel{\dot{\Rightarrow}} z' \dot{\in} (x\,w')$. Likewise, $z' \dot{\in} (x\,w') \mathrel{\dot{\Rightarrow}} z' \dot{\in} (x\,v')$ so $\phi z' \rightarrow (z' \dot{\in} (x\,v') \mathrel{\dot{\Leftrightarrow}} z' \dot{\in} (x\,w))$.

By Ackermann's axiom we have $v \doteq \varepsilon z''.z'' \dot{\in} (x\,v') = \varepsilon z''.z'' \dot{\in} (x\,w') \doteq w$ which combined with $z \dot{\in} v$ gives $z \dot{\in} w$ as required. □

8.11. The axiom of restriction

The axiom of restriction states that for any non-empty set x there is a set $z \in x$ such that x and z have no elements in common. In order to prove this statement, we introduce a function $D(x, y)$ with the following property: If $y \in x$, then $z = D(x, y)$ satisfies $z \in x$ and has no elements in common with x. We define $D(x, y)$ as follows: If x and y have no elements in common, then $D(x, y) = y$. Otherwise, $D(x, y) = D(x, y')$ where y' is a common element of x and y. The definition $D(x, y) = D(x, y')$ is recursive. The well-foundedness of y ensures that the recursion terminates (this is expressed formally in Lemma 8.11.1 below). The formal definition reads

$$D(x, y) = (\text{if } (\dot\neg y \mathbin{\dot\wedge} \dot\exists v.(y\, v) \mathrel{\dot\in} x)\ D(x, (y\, \varepsilon v.(y\, v) \mathrel{\dot\in} x))\ y).$$

Lemma 8.11.1. $\dot\forall x.\dot\forall y.\phi D(x, y)$.

Proof. Assume ϕx. The proof is by induction in y. For $y = \mathsf{T}$ we have $D(x, y) = \mathsf{T} \in \Phi$ so $\phi D(x, y)$ holds. Now assume as inductive hypothesis ϕy, $\dot\neg y$ and $\dot\forall u.\phi D(x, (y\, u))$. We shall prove $D(x, y)$ by *TND* from the following two cases.

Case 1. Assume $\dot\exists v.(y\, v) \mathrel{\dot\in} x$. Let $u = \varepsilon v.(y\, v) \mathrel{\dot\in} x$. We have $D(x, y) = D(x, (y\, u)) \in \Phi$ so $\phi D(x, y)$ holds.

Case 2. Assume $\dot\neg\dot\exists v.(y\, v) \mathrel{\dot\in} x$. We have $D(x, y) = y \in \Phi$ so $\phi D(x, y)$ holds. □

From now on, we regard the above lemma as part of the totality theorem, i.e., the definition of the syntax class Σ of simple terms is extended to

$$\Sigma \ ::= \ \cdots \mid D(\Sigma, \Sigma).$$

Lemma 8.11.2. $\phi x, \phi y \rightarrow y \mathrel{\dot\in} x \mathrel{\dot\Rightarrow} D(x, y) \mathrel{\dot\in} x$.

Proof. Assume ϕx. The proof is by induction in y. Assume $y = \mathsf{T}$ and $y \mathrel{\dot\in} x$. We have $D(x, y) = y$ and $y \mathrel{\dot\in} x$ so $y \mathrel{\dot\in} x \mathrel{\dot\Rightarrow} D(x, y) \mathrel{\dot\in} x$ as required. Now assume as inductive hypothesis ϕy, $\dot\neg y$ and $\dot\forall u.D(x, (y\, u)) \mathrel{\dot\in} x$. Further assume $y \mathrel{\dot\in} x$. We shall prove $D(x, y) \mathrel{\dot\in} x$ by *TND* from the following two cases.

Case 1. Assume $\dot\exists v.(y\, v) \mathrel{\dot\in} x$. Let $u = \varepsilon v.(y\, v) \mathrel{\dot\in} x$. We have $(y\, u) \mathrel{\dot\in} x$ so $D(x, y) = D(x, (y\, u))$ and $D(x, (y\, u)) \mathrel{\dot\in} x$.

Case 2. Assume $\dot\neg\dot\exists v.(y\, v) \mathrel{\dot\in} x$. We have $D(x, y) = y$ and $y \mathrel{\dot\in} x$ so $D(x, y) \mathrel{\dot\in} x$ as required. □

Lemma 8.11.3. $\phi x, \phi y \rightarrow \dot\forall z.(z \mathrel{\dot\in} D(x, y) \mathrel{\dot\Rightarrow} \dot\neg z \mathrel{\dot\in} x)$.

Proof. Assume ϕx. The proof is by induction y. Assume $y = \mathsf{T}$ and ϕz. From $D(x, y) = \mathsf{T}$ we have $\dot\neg z \mathrel{\dot\in} D(x, y)$ so $z \mathrel{\dot\in} D(x, y) \mathrel{\dot\Rightarrow} \dot\neg z \mathrel{\dot\in} x$ as required. Now assume as inductive hypothesis ϕy, $\dot\neg y$ and $\dot\forall u.\dot\forall z.(z \mathrel{\dot\in} D(x, (y\, u)) \mathrel{\dot\Rightarrow} \dot\neg z \mathrel{\dot\in} x)$. Further assume ϕz and $z \mathrel{\dot\in} D(x, y)$. We shall prove $\dot\neg z \mathrel{\dot\in} x$ by *TND* from the following two cases.

Case 1. Assume $\dot{\exists} v.(y\, v) \dot{\in} x$. Let $u = \varepsilon v.(y\, v) \dot{\in} x$. From $z \dot{\in} D(x, y)$ and $D(x, y) = D(x, (y\, u))$ we have $z \dot{\in} D(x, (y\, u))$ so, by the inductive hypothesis, $\dot{\neg} x \dot{\in} x$.

Case 2. Assume $\dot{\neg}\dot{\exists} v.(y\, v) \dot{\in} x$. From $z \dot{\in} D(x, y)$ and $D(x, y) = y$ we have $z \dot{\in} y$. Choose $u \in \Phi$ such that $z \doteq (y\, u)$. From $\dot{\neg}\dot{\exists} v.(y\, v) \dot{\in} x$ we have $\dot{\forall} v.\dot{\neg}(y\, v) \dot{\in} x$ so $\dot{\neg} z \dot{\in} x$. □

Theorem 8.11.4 (ZFC-D). $\dot{\forall} x.(\dot{\exists} y.y \dot{\in} x \dot{\Rightarrow} \dot{\exists} y.y \dot{\in} x \dot{\wedge} \dot{\neg}\dot{\exists} z.z \dot{\in} x \dot{\wedge} z \dot{\in} y)$.

Proof. Assume ϕx and $\dot{\exists} y.y \dot{\in} x$. Choose $u \in \Phi$ such that $u \dot{\in} x$ and let $y = D(x, u)$. From Lemma 8.11.1 we have ϕy, from Lemma 8.11.2 we have $y \dot{\in} x$, and from Lemma 8.11.3 we have $\dot{\neg}\dot{\exists} z.z \dot{\in} x \dot{\wedge} z \dot{\in} y$. □

Part III. The consistency of map theory

9. General concepts and notations

For all axiomatizations of ZFC' of set theory, let $Con(ZFC')$ be the statement that there is no proof of $\forall x: x \in x$ in ZFC'. For all axiomatizations Map' of map theory let $Con(Map')$ be the statement that there is no proof of $\mathsf{T} = \lambda x.\mathsf{T}$ in Map'. Let SI be the statement that there exists a strongly inaccessible ordinal.

We shall prove

$$SI \Rightarrow Con(Map) \tag{8}$$

where Map is the axiomatization of map theory stated in Part II. It is an open question whether or not $Con(ZFC) \Rightarrow Con(Map)$ is provable in ZFC.

Let Map° be the axiom system Map where Well2 and the construction axioms are omitted. We shall prove

$$Con(ZFC) \Rightarrow Con(Map^\circ). \tag{9}$$

Actually, it is merely necessary to exclude Well2, C-Prim and C-M1 from Map° to prove the above result. Each of these axioms has a strength similar to the axiom scheme of replacement in ZFC. The other construction axioms can be verified from $Con(ZFC)$, but the verification on basis of SI gives a better understanding of the intuition behind the axioms.

Further, for any extension ZFC^+ of ZFC we shall prove

$$Con(ZFC^+) \Rightarrow Con(Map^{\circ +}) \tag{10}$$

in ZFC^+ where $Map^{\circ +}$ is Map° extended with all theorems of ZFC^+, i.e., the translation of any theorem of ZFC^+ into map theory is an axiom of $Map^{\circ +}$. The system $Map^{\circ +}$ is not interesting in itself, but (10) ensures that for any consistent extension of ZFC there is a corresponding consistent extension of Map° with at least the same strength, so ZFC cannot outsmart map theory by additional axioms.

Before giving an overview of the consistency proof, it is necessary to give a condensed introduction to the notation used in the proof.

9.1. Basic concepts

The symbols $\neg$, $\Rightarrow$ and $\Leftrightarrow$ stand for negation, implication and bi-implication, respectively, and

$$p_1 \Rightarrow p_2 \Rightarrow p_3 \Rightarrow \cdots \Rightarrow p_{n-1} \Rightarrow p_n$$

stands for repeated use of $\Rightarrow$, i.e.

$$(p_1 \Rightarrow p_2) \wedge (p_2 \Rightarrow p_3) \wedge \cdots \wedge (p_{n-1} \Rightarrow p_n),$$

and likewise for $\Leftrightarrow$. The order of precedence is $\Leftrightarrow, \Rightarrow, \forall, \exists, \vee, \wedge, \neg$; e.g. $\forall x.\mathscr{A} \wedge \mathscr{B} \Rightarrow \mathscr{C}$ means $(\forall x.(\mathscr{A} \wedge \mathscr{B})) \Rightarrow \mathscr{C}$. The symbols $\bigcup$ and $\mathscr{P}$ stand for union and power set, respectively. We have $x \in \bigcup G \Leftrightarrow \exists y \in G\colon x \in y$ and $x \in \mathscr{P}G \Leftrightarrow x \subseteq G$. Further, $\bigcup_{y \in G} f(y)$ stands for $\bigcup \{f(y) \mid y \in G\}$. Hence, $x \in \bigcup_{y \in G} f(y) \Leftrightarrow \exists y \in G\colon x \in f(y)$.

9.2. Functions

The construct (x, y) stands for the Kuratowski pair $\{\{x\}, \{x, y\}\}$ of x and y, and $G \times H$ stands for the Cartesian product $\{(x, y) \mid x \in G \wedge y \in H\}$.

The construct $fnc(g)$ stands for "g is a function", i.e.

$$\begin{aligned} fnc(g) \ \Leftrightarrow\ & \exists G, H\colon g \subseteq G \times H \\ & \qquad \wedge\ \forall x, y, z\colon ((x, y) \in g \wedge (x, z) \in g \Rightarrow y = z). \end{aligned}$$

For all functions g we define the domain g^{d} and the range g^{r}:

$$x \in g^{\mathrm{d}} \ \Leftrightarrow\ fnc(g) \wedge \exists y\colon (x, y) \in g,$$

$$x \in g^{\mathrm{r}} \ \Leftrightarrow\ fnc(g) \wedge \exists y\colon (y, x) \in g.$$

Note that $g^{\mathrm{d}} = \emptyset$ and $g^{\mathrm{r}} = \emptyset$ if g is not a function. If g is a function and $x \in g^{\mathrm{d}}$, then we let $g(x)$ stand for the unique y such that $(x, y) \in g$. If g is not a function or $x \notin g^{\mathrm{d}}$ we define $g(x) = \emptyset$, and we actually make use of this convention.

We use $G \to H$ to stand for the set of functions from G into H:

$$g \in G \to H \ \Leftrightarrow\ fnc(g) \wedge g^{\mathrm{d}} = G \wedge g^{\mathrm{r}} \subseteq H.$$

For all sets G we define G^{D} and G^{R} as follows:

$$G^{\mathrm{D}} = \bigcup_{g \in G} g^{\mathrm{d}}, \qquad G^{\mathrm{R}} = \bigcup_{g \in G} g^{\mathrm{r}}.$$

For all sets G and H we have $(G \to H)^{\mathrm{D}} = G$ if $H \neq \emptyset$ and $(G \to H)^{\mathrm{R}} = H$ if $G \neq \emptyset$.

For all variables x, sets G and terms Φ of ZFC we introduce $x \in G \mapsto \Phi$ as the function whose domain is G and which maps x to Φ (where x may occur free in Φ). More formally,

$$x \in G \mapsto \Phi = \{(x, \Phi) \mid x \in G\}.$$

As an example, $x \in \mathbf{R} \mapsto x^2$ is the squaring function on the set $\mathbf{R}$ of real numbers. As another example, if $g = x \in \mathbf{R} \mapsto y \in \mathbf{R} \mapsto x^2 + y^2$, then $g(3) = y \in \mathbf{R} \mapsto 9 + y^2$, and $g(3)(4) = 9 + 16 = 25$. We have $g \in \mathbf{R} \to (\mathbf{R} \to \mathbf{R})$.

The constructs $g \circ h$ and $g|h$ stand for functional composition and restriction, respectively:

$$(x, z) \in g \circ h \Leftrightarrow \exists y\colon (x, y) \in h \wedge (y, z) \in g,$$

$$(x, y) \in g|G \Leftrightarrow (x, y) \in g \wedge x \in G.$$

9.3. Ordinals

We use von Neumann ordinals [22]. The term $ord(x)$ stands for 'x is an ordinal', 0 stands for $\emptyset$, x^+ stands for the successor of x, i.e. $x \cup \{x\}$, and ω stands for the least infinite ordinal, i.e. the set of all finite ordinals. The term On stands for the class of all ordinals, i.e., $x \in On$ is shorthand for $ord(x)$.

The variables α, β, etc., implicitly range over ordinals. The variable μ implicitly ranges over limit ordinals. The constructs $<_{\mathrm{o}}$ and $\leqslant_{\mathrm{o}}$ are defined by $\alpha <_{\mathrm{o}} \beta \Leftrightarrow \alpha \in \beta$ and $\alpha \leqslant_{\mathrm{o}} \beta \Leftrightarrow \alpha \in \beta \vee \alpha = \beta$.

9.4. Tuples

We represent tuples like in [17]: x is a tuple if x is a function and $x^{\mathrm{d}} \in \omega$:

$$tpl(x) \Leftrightarrow fnc(x) \wedge x^{\mathrm{d}} \in \omega.$$

We use $(x_0, \ldots, x_{\alpha-1})$ to denote $\{(0, x_0), \ldots, (\alpha - 1, x_{\alpha-1})\}$. Hence, $\langle x_0, \ldots, x_{\alpha-1}\rangle$ is the unique tuple x for which $x^{\mathrm{d}} = \{0, \ldots, \alpha - 1\} = \alpha$ and $x(\beta) = x_\beta$ for $\beta \in \{0, \ldots, \alpha - 1\} = \alpha$. In particular, we use $\langle\ \rangle$ to stand for the empty tuple $\emptyset$ for which $\langle\ \rangle^{\mathrm{d}} = 0 = \emptyset$. G^* stands for the set of tuples of elements of G, i.e. $G^* = \bigcup_{\alpha \in \omega} (\alpha \to G)$ (G^* is written $G^{<\omega}$ in [17]). We shall use the following fact extensively:

$$G^{*\mathrm{R}} = G.$$

The construct $x \cdot y$ stands for tuple concatenation. We have

$$\langle x_1, \ldots, x_\alpha\rangle \cdot \langle y_1, \ldots, y_\beta\rangle = \langle x_1, \ldots, x_\alpha, y_1, \ldots, y_\beta\rangle.$$

Define $G^\omega = \omega \to G$. The elements of G^ω are infinite sequences of elements of G. If $f \in G^\omega$ and $\alpha \in \omega$, then $(f|\alpha) \in G^*$ and $(f|\alpha)$ is the tuple containing the first α elements of the infinite sequence f.

9.5. Rank

For all sets G and H, $G =_\kappa H$ stands for "G has the same cardinality as H". $<_\kappa$, $\leqslant_\kappa$, etc., are defined likewise. We shall use transfinite induction in various well-founded relations and transfinite recursion in various well-founded, set-like relations [17]. As an example, we define the rank $\rho(G)$ by $\rho(G) = \bigcup_{x \in G} \rho(x)^+$ by transfinite recursion in $\in$. For all sets G, $\rho(G)$ is an ordinal. $G <_\rho H$ stands for $\rho(G) <_o \rho(H)$. We define the transitive closure $tc(G)$ by

$$tc(G) = G \cup \bigcup_{x \in G} tc(x).$$

9.6. Relations

In order to have a sufficient supply of names for relations, they are given names like $=_\kappa$, $\leqslant_o$ and $<_\rho$, where the index is part of the name. In general, relations with names like $=_a$ are equivalence relations. Relations with names like $<_a$ generally are strict partial preorders, i.e. they satisfy $x \not<_a x$ and $x <_a y \wedge y <_a z \Rightarrow x <_a z$. Relations with names like $\leqslant_a$ generally are weak partial preorders, i.e. they satisfy $x \leqslant_a x$ and $x \leqslant_a y \wedge y \leqslant_a z \Rightarrow x \leqslant_a z$.

Occasionally we shall need parameterized relations like $\leqslant_z^a$. For each x, y and z, $x \leqslant_z^a y$ is either true or false. The name $\leqslant_z^a$ suggests that $x \leqslant_z^a y$ is a partial preorder in x and y for each, fixed z.

For all relations $<_a$ we define $<_a^*$ as follows:

$$f <_a^* g \Leftrightarrow fnc(f) \wedge fnc(g) \wedge f^{\mathrm{d}} = g^{\mathrm{d}} \wedge \forall x \in f^{\mathrm{d}}: f(x) <_a g(x).$$

For example,

$$\langle x_1, \ldots, x_\alpha \rangle <_a^* \langle y_1, \ldots, y_\beta \rangle \Leftrightarrow \alpha = \beta \wedge x_1 <_a y_1 \wedge \cdots \wedge x_\alpha <_a y_\alpha.$$

Stars may be applied several times:

$$f <_a^{**} g \Leftrightarrow fnc(f) \wedge fnc(g) \wedge f^{\mathrm{d}} = g^{\mathrm{d}} \wedge \forall x \in f^{\mathrm{d}}: f(x) <_a^* g(x).$$

Stars may also be applied to relations with names like $=_a$ and $\leqslant_a$, and to parameterized relations like $\leqslant_z^a$ to form relations like $=_a^*$ and $\leqslant_z^{a**}$.

9.7. Labels

In Part I we have introduced the labels $\tilde{\mathsf{T}}, \tilde{\lambda}$ and $\tilde{\bot}$. We now introduce them formally by the arbitrary definitions $\tilde{\bot} = 0$, $\tilde{\lambda} = 1$ and $\tilde{\mathsf{T}} = 2$. The definition of $\tilde{\bot}$ is not quite arbitrary: If $x \notin g^{\mathrm{d}}$ then $g(x) = \tilde{\bot}$, and we shall use this result. Let $L = \{\tilde{\mathsf{T}}, \tilde{\lambda}, \tilde{\bot}\}$ be the set of labels. We organize L by a partial ordering $\leqslant_L$ as follows:

$$x \leqslant_L y \Leftrightarrow x = \tilde{\bot} \vee x = y.$$

For all $G \subseteq L$ we define

$$\bigsqcup G = \begin{cases} \tilde{\mathsf{T}} & \text{if } \tilde{\mathsf{T}} \in G, \\ \tilde{\lambda} & \text{if } \tilde{\mathsf{T}} / \in G \wedge \tilde{\lambda} \in G, \\ \tilde{\bot} & \text{if } \tilde{\mathsf{T}} \notin G \wedge \tilde{\lambda} \notin G. \end{cases}$$

In what follows, we are going to form $\bigsqcup G$ occasionally, and in each case we have $\forall x \in G: x \leqslant_L \bigsqcup G$. In several cases, however, the proof of $\forall x \in G: x \leqslant_L \bigsqcup G$ cannot be stated immediately after $\bigsqcup G$ is formed.

9.8. Well-founded functions

We say that a function f is "well-founded on G" if $f \in G^\circ$ where

$$G^\circ = \{f \in G^* \to \{\tilde{\mathsf{T}}, \tilde{\lambda}\} \mid \forall x, y \in G^*: (f(x) \neq \tilde{\lambda} \Rightarrow f(x \cdot y) = f(x)) \wedge \forall x \in G^\omega\ \exists \alpha \in \omega: f(x \mid \alpha) = \tilde{\mathsf{T}}\}.$$

The notion G° corresponds to $wf(G)$ in Part I. We shall use such well-founded functions to represent well-founded maps. If $f \in G^\circ$ then $f^{\mathrm{d}} = G^*$ and $f^{\mathrm{dR}} = G$. Hence, if $f \in G^\circ$ and $f \in H^\circ$ then $G = f^{\mathrm{dR}} = H$. For all sets f and g we introduce the relation $<_w$ as follows:

$$f <_w g \Leftrightarrow g(\langle\ \rangle) = \tilde{\lambda} \wedge f \in G^\circ \wedge g \in G^\circ \wedge \exists x \in G\ \forall y \in G^*: f(y) = g(\langle x \rangle \cdot y)$$

where $G = g^{\mathrm{dR}}$. Obviously, $<_w$ is well-founded and set-like.

On several occasions we shall need to apply a function f to an argument x such that $f(x) \neq \tilde{\lambda} \Rightarrow f(x \cdot y) = f(x)$. In order to obtain this, we define $f\langle\langle x \rangle\rangle$ as follows:

(1) $f\langle\langle x \rangle\rangle = f(x)$,

(2) $f\langle\langle x \rangle\rangle \neq \tilde{\lambda} \Rightarrow f\langle\langle x \cdot y \rangle\rangle = f(x)$ and

(3) (2) takes precedence over (1);

or, more precisely, for all y and all tuples x define

$$f\langle\langle \langle\ \rangle \rangle\rangle = f(\langle\ \rangle),$$

$$f\langle\langle x \cdot \langle y \rangle \rangle\rangle = \begin{cases} f(x \cdot \langle y \rangle) & \text{if } f\langle\langle x \rangle\rangle = \tilde{\lambda}, \\ f\langle\langle x \rangle\rangle & \text{otherwise.} \end{cases}$$

If $f \in G^\circ$ and $x \in G^*$ then $f\langle\langle x \rangle\rangle = f(x)$.

9.9. Gödel numbers

We now introduce Gödel numbers for terms of map theory and for well-formed formulas of *ZFC*. To do so, define

- $\dot{A}(x, y) = \langle 0, x, y \rangle$,
- $\dot{S} = \langle 1 \rangle$,
- $\dot{K} = \langle 2 \rangle$,
- $\dot{\mathsf{T}} = \langle 3 \rangle$,
- $\dot{P} = \langle 4 \rangle$,
- $\dot{C} = \langle 5 \rangle$,
- $\dot{W} = \langle 6 \rangle$,
- $\dot{\bot} = \langle 7 \rangle$,
- $\dot{v}_i = \langle 8, i \rangle$,
- $\dot{\lambda} x.y = \langle 9, x, y \rangle$,
- $(\dot{\mathsf{if}}\ x\ y\ z) = \langle 10, x, y, z \rangle$,
- $\dot{\varepsilon} x = \langle 11, x \rangle$,
- $\dot{\phi} x = \langle 12, x \rangle$,

- $\ddot{v}_i = \langle 8, i\rangle$,
- $x \,\ddot{\in}\, y = \langle 13, x, y\rangle$,
- $\ddot{\neg} x = \langle 14, x, y\rangle$,
- $x \ddot{\Rightarrow} y = \langle 15, x, y\rangle$,
- $\ddot{\forall} x\colon y = \langle 16, x, y\rangle$.

Note that $\dot{v}_i = \ddot{v}_i$. We use $\dot{v}_i$ when talking about variables in map theory and $\ddot{v}_i$ when talking about variables in set theory.

As an example, the Gödel number of

$$\forall x\colon (x \in x \Rightarrow \neg x \in y)$$

is

$$\ddot{\forall} \ddot{x}\colon (\ddot{x} \,\ddot{\in}\, \ddot{x} \ddot{\Rightarrow} \ddot{\neg} \ddot{x} \,\ddot{\in}\, \ddot{y})$$

if we let $\ddot{x} = \ddot{v}_0$ and $\ddot{y} = \ddot{v}_1$ represent x and y, respectively.

It is customary in the literature to introduce a notation like $\lceil \Phi \rceil$ for the Gödel number of the well-formed formula Φ. We shall avoid this notation in order to ensure referential transparency since this gives most flexibility in nontrivial handling of Gödel numbers.

We use $\dot{(} x\, y \dot{)}$ as shorthand for $\dot{A}(x, y)$, and $\dot{(} x\, y_1\, y_2 \ldots y_\alpha \dot{)}$ as shorthand for $\dot{(} \ldots \dot{(} \dot{(} x\, y_1 \dot{)}\, y_2 \dot{)} \ldots y_\alpha \dot{)}$.

9.10. Models

In the consistency proofs we shall use a transitive standard model D of ZFC. However, we use different D in different places of the text. In the proof of (8), i.e. in the proof of $SI \Rightarrow Con(Map)$ we define D in the obvious way [17]:

Let σ be strongly inaccessible,

$$\begin{aligned} &\text{define } \Phi(\alpha) = \textstyle\bigcup_{\beta \in \alpha} \mathcal{P}(\Phi(\beta)), \\ &\text{define } D = \Phi(\sigma). \end{aligned} \tag{11}$$

In the proof of (9) and (10), i.e. in the proof of $Con(ZFC) \Rightarrow Con(Map^\circ)$ and $Con(ZFC^+) \Rightarrow Con(Map^{\circ +})$ we proceed like Cohen, i.e. at any time we assume that D satisfies finitely many axioms of ZFC and ZFC^+, respectively, without being explicit about which ones. In other words, we constantly assume that D satisfies sufficiently many axioms for the argument at hand. For a discussion of this see [9, 17, 5].

For any well-formed formula Φ we let $\lfloor \Phi \rfloor$ stand for the relativization of Φ to D, i.e. the expression obtained by replacing each occurrence of $\forall x$ by $\forall x \in D$ in Φ. We say that Φ is absolute if $\forall x_1, \ldots, x_\alpha\colon (\Phi \Leftrightarrow \lfloor \Phi \rfloor)$ where $x_1, \ldots, x_\alpha$ are the free variables of Φ. We use the relativization and absoluteness results of [17] without further reference.

The notation $\lfloor \Phi \rfloor$ is not referentially transparent, i.e. $\eta = \theta \Rightarrow (\lfloor \Phi(\eta) \rfloor \Leftrightarrow \lfloor \Phi(\theta) \rfloor)$ does not hold for all terms η and θ and well-formed formulas $\Phi(\bullet)$. We compromise on referential transparency in this case since we shall only make trivial use of relativization.

10. Overview of the model construction

10.1. The semantic model

In this section we give an overview of the model construction rather than stating formal definitions. The purpose of doing so is to give an intuitive understanding of the model before defining it. As a starting point we take the description of map theory from Part I, and then we elaborate the description. In particular, we shall go into the details of well-founded maps. Later, when it comes to the formal definition, we start with the details and gradually build up the model.

In Part I, functional application (fx) and the combinators S, K, T, P, C, W and $\perp$ were introduced. In addition to functional application and the combinators, we shall introduce a number of concepts that are explained below.

The set of maps is denoted M and the set of well-founded maps is denoted Φ. The function $A \in M \times M \to M$ is defined by $A(f, x) = (fx)$, so $A(f, x)$ is just another notation for functional application in map theory. The function $m \in M \times M^* \to M$ is defined by $m(f, \langle x_1, \dots, x_\alpha\rangle) = (f x_1 \dots x_\alpha)$. Hence, $A(f, z)$ denotes f applied to one argument whereas $m(f, \langle x_1, \dots, x_\alpha\rangle)$ denotes f applied to a list of arguments. The names A and m stand for "application" and "multiple application", respectively.

The function $r \in M \to L$ is defined by

$$r(x) = \begin{cases} \tilde{\mathsf{T}} & \text{if } x = \mathsf{T}, \\ \tilde{\perp} & \text{if } x = \perp, \\ \tilde{\lambda} & \text{otherwise.} \end{cases}$$

Hence, $r(x)$ denotes the label of the root of x (r stands for "root"). The function $a \in M \to (M^* \to L)$ is defined by $a(f)(\langle x_1, \dots, x_\alpha\rangle) = r(f x_1 \dots x_\alpha) = r(m(f, \langle x_1, \dots, x_\alpha\rangle))$. Hence, $a(f)(\langle x_1, \dots, x_\alpha\rangle)$ stands for the label of the node reached by traveling from the root node of f along the path $\langle x_1, \dots, x_\alpha\rangle$ (a stands for "application" as does A).

According to the extensionality of maps we have $f = g$ iff

$$r(f x_1 \dots x_\alpha) = r(g x_1 \dots x_\alpha)$$

for all $x_1, \dots, x_\alpha \in M$. In other words,

$$f = g \Leftrightarrow a(f) = a(g).$$

We define $f \leq g$ iff

$$r(f x_1 \dots x_\alpha) \leq_L r(g x_1 \dots x_\alpha)$$

for all $x_1, \dots, x_\alpha \in M$. In other words,

$$f \leq g \Leftrightarrow a(f) \leq_L^* a(g).$$

We have

$$f \leqslant g \wedge g \leqslant f \Leftrightarrow f = g.$$

Figure 8 shows the universe M and four of its elements: T, $\perp$, $\lambda x.\mathsf{T}$ and $\lambda x.\perp$. The relation $\perp \leqslant \mathsf{T}$ is illustrated by drawing T above $\perp$ and interconnecting the two by a line. The same is done for the relation $\perp \leqslant \lambda x.\perp \leqslant \lambda x.\mathsf{T}$.

Since $x \leqslant y \wedge y \leqslant z \Rightarrow x \leqslant z$, the same information could be represented as in Fig. 9. The only difference is that Fig. 9 displays $\perp \leqslant \lambda x.\mathsf{T}$ in addition to $\perp \leqslant \lambda x.\perp \leqslant \lambda x.\mathsf{T}$.

The monotonicity of maps states

$$f \leqslant g \wedge x \leqslant y \Rightarrow (f x) \leqslant (g y),$$

$$f \leqslant g \wedge x \leqslant^* y \Rightarrow m(f, x) \leqslant m(g, y),$$

$$f \leqslant g \wedge x \leqslant^* y \Rightarrow a(f)(x) \leqslant_L a(g)(y).$$

For all well-founded $f, x_1, x_2, \ldots \in \Phi$ there exists an $\alpha \in \omega$ such that $(f x_1 \ldots x_\alpha) = \mathsf{T}$. In other words,

$$\forall f \in \Phi\ \forall x \in \Phi^\omega\ \exists \alpha \in \omega: m(f, x|\alpha) = \mathsf{T}. \tag{12}$$

For all well-founded $f, g \in \Phi$ define

$$f <_A g \Leftrightarrow f \neq \mathsf{T} \wedge \exists x \in \Phi: f = (g x).$$

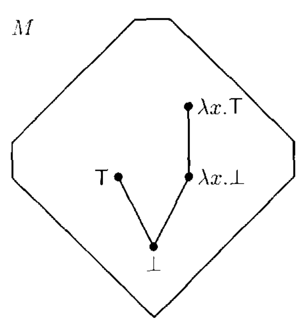

Fig. 8. Illustration of the $\leqslant$ relation.

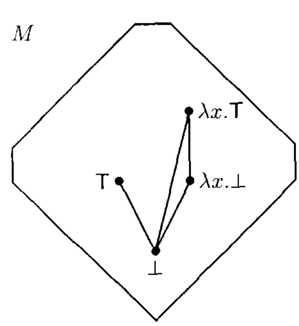

Fig. 9. Another illustration of the $\leqslant$ relation.

This relation is well-founded according to (12). The function $s(x)$ is defined for all $x \in \Phi$ by

$$s(x) = \begin{cases} \emptyset & \text{if } x = \mathsf{T}, \\ \{s(A(x, y)) | y \in \Phi\} & \text{otherwise.} \end{cases}$$

This can also be written

$$s(x) = \{s(y) \mid y <_A x\}$$

which is a valid definition since $<_A$ is well-founded (and set-like since it is restricted to the set Φ). We have that $s(x)$ denotes the set represented by $x \in \Phi$. The model will be constructed such that

$$s^{\mathrm{r}} = D$$

holds, i.e. such that the representable sets are exactly the elements of the model D. This is a key result in proving $Con(ZFC^+) \Rightarrow Con(Map^{o+})$.

10.2. Types and observational equivalence

Let $f \in M$. For any $x_1, \ldots, x_\alpha \in M$ we say that the root of $(f x_1 \ldots x_\alpha)$ is an "observable property" of f. In other words, the value of $a(f)(x)$ is an observable property of f for each $x \in M^*$. If two maps $f, g \in M$ are observably equivalent, i.e. if $a(f)(x) = a(g)(x)$ for all $x \in M^*$, then they are equal according to the extensionality of maps. In other words, $a(f) = a(g) \Leftrightarrow f = g$.

If $G \subseteq M$ and $x_1, \ldots, x_\alpha \in G$, then we say that the root of $(f x_1 \ldots x_\alpha)$ is observable within G. Furthermore, we say that f and g are observably equivalent within G and write $f =^a_G g$ if $\forall x \in G^*$: $a(f)(x) = a(g)(x)$. Observational equivalence is an equivalence relation for each, fixed G.

Define the "type" $t_G(f)$ of f w.r.t. G by

$$t_G(f) = x \in G^* \mapsto a(f)(x).$$

We have $t_G(f) \in G^* \mapsto L$ and

$$f =^a_G g \Leftrightarrow t_G(f) = t_G(g).$$

If $f, g \in M$, $G \subseteq M$ and $f =^a_G g$, then there exists an $h \in M$ such that $h \leq f$, $h \leq g$, and $h =^a_G f =^a_G g$. To see this, define

$$\begin{aligned} H = \mathsf{Y} h'.\lambda f'.\lambda g'.& \\ & (\mathsf{if}\ f'\ (\mathsf{if}\ g'\ \mathsf{T}\ \perp)) \\ & (\mathsf{if}\ g'\ \perp\ \lambda x.(h'\ (f'\ x)\ (g'\ x))), \\ h = (H\,f\,g).& \end{aligned}$$

10.3. Well-founded maps

Let $G \subseteq M$. The map $f \in M$ is said to be well-founded w.r.t. G iff, for all $x_1, x_2, \ldots \in G$ there exists an $\alpha \in \omega$ such that $(f x_1 \ldots x_\alpha) = \mathsf{T}$. In other words, f is well-founded w.r.t. G iff $\forall x \in G^\omega\ \exists \alpha \in \omega$: $m(f, x|\alpha) = \mathsf{T}$, which can also be written

$$t_G(f) \in G^\circ.$$

In particular, any $f \in \Phi$ is well-founded w.r.t. Φ.

Let $wf(G)$ denote the set of $f \in M$ that are well-founded w.r.t. G, i.e.

$$wf(G) = \{f \in M \mid t_G(f) \in G^\circ\}.$$

Now define

$$\Phi'(\alpha) = wf(Q'(\alpha)),$$

$$\Phi''(\alpha) = \bigcup_{\beta \in \alpha} \Phi'(\beta),$$

$$Q'(\alpha) = wf(\Phi''(\alpha)).$$

$$\Phi = \bigcup_{\beta \in On} \Phi'(\beta).$$

In the definition of Φ, β ranges over On which may cause Φ to become a proper class. This is avoided by a modified definition later on.

If f is well-founded w.r.t. G, then f is also well-founded w.r.t. any $H \subseteq G$. Hence, $H \subseteq G \Rightarrow wf(G) \subseteq wf(H)$. Furthermore, if $\alpha \leqslant_{\mathrm{o}} \beta$ then

$$\Phi''(\alpha) \subseteq \Phi'(\alpha) \subseteq \Phi''(\beta) \subseteq \Phi'(\beta) \subseteq \Phi \subseteq Q'(\beta) \subseteq Q'(\alpha).$$

If $u \in G^\circ$ and $v \in G^* \to L$ then $u \leqslant_L^* v \Leftrightarrow u = v$. If $f \leqslant g$ and $f \in wf(G)$ then $t_G(f) \leqslant_L^* t_G(g)$ and $t_G(g) \in G^\circ$, so $t_G(f) = t_G(f) \in G^\circ$ which proves $g \in wf(G)$. Hence

$$f \leqslant g \wedge f \in wf(G) \Rightarrow g \in wf(G) \wedge f =_G^{\mathrm{a}} g,$$

$$f \leqslant g \wedge f \in \Phi'(\alpha) \Rightarrow g \in \Phi'(\alpha),$$

$$f \leqslant g \wedge f \in \Phi''(\alpha) \Rightarrow g \in \Phi''(\alpha),$$

$$f \leqslant g \wedge f \in Q'(\alpha) \Rightarrow g \in Q'(\alpha),$$

$$f \leqslant g \wedge f \in \Phi \Rightarrow g \in \Phi.$$

Figures 10, 11 and 12 give a picture of how $\Phi'(\alpha)$, $\Phi''(\alpha)$, $Q'(\alpha)$ and Φ relate to each other. Figures 10–12 display $\Phi'(2)$, Φ and $Q'(2)$, respectively. The figures illustrate statements like $\Phi'(2) \subseteq \Phi \subseteq Q'(2)$ and $f \leqslant g \wedge f \in \Phi \Rightarrow g \in \Phi$.

Figures 10–12 do not capture the close relationship between $\Phi''(\alpha)$, $Q'(\alpha)$ and $\Phi'(\alpha)$. This is done in Fig. 13. However, Fig. 13 does not illustrate statements like $\Phi'(2) \subseteq Q'(2)$ and $f \leqslant g \wedge f \in \Phi'(2) \Rightarrow g \in \Phi'(2)$.

10.4. Minimal well-founded maps

We say that the map $f \in G$ is minimal in G if $\forall g \in G$: $(g \leqslant f \Rightarrow g = f)$. Define the boundary ∂G to be the set of minimal elements of G. We say that G is closed if

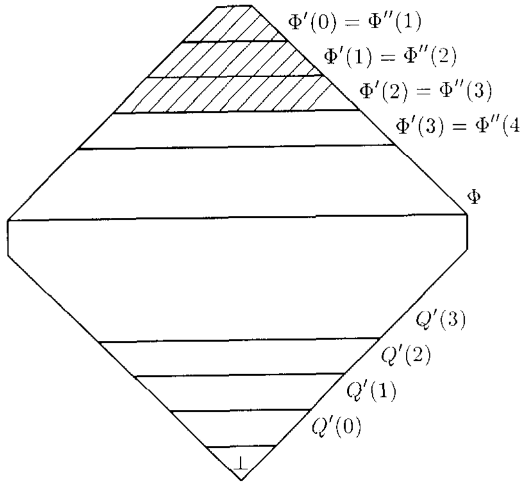

Fig. 10. Illustration of $\Phi'(2)$.

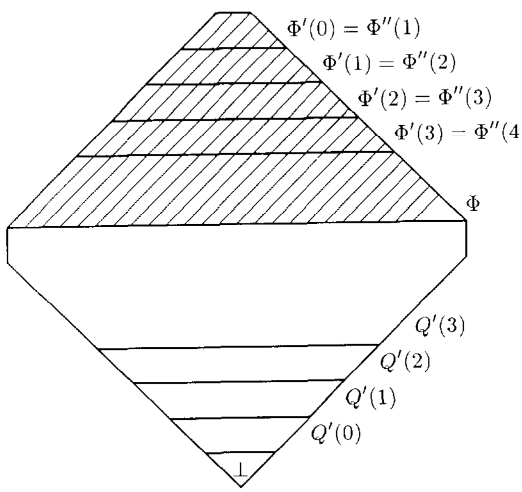

Fig. 11. Illustration of Φ.

$\forall f \in G\ \exists g \in \partial G$: $g \leqslant f$. The model M is going to be defined such that $\Phi'(\alpha)$ and $Q'(\alpha)$ are closed for all α. Define

$$\check{\Phi}'(\alpha) = \partial\Phi'(\alpha),$$

$$\check{Q}'(\alpha) = \partial Q'(\alpha),$$

$$\check{\Phi}''(\alpha) = \bigcup_{\beta \in \alpha} \check{\Phi}'(\beta),$$

$$\check{\Phi} = \bigcup_{\beta \in On} \check{\Phi}'(\beta),$$

$$\check{Q} = \bigcup_{\beta \in On} \check{Q}'(\beta).$$

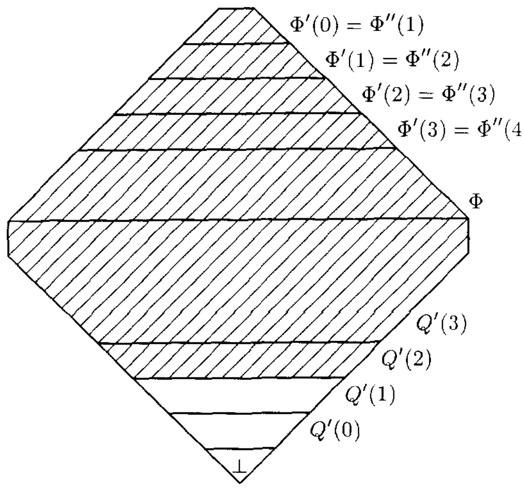

Fig. 12. Illustration of $Q'(2)$.

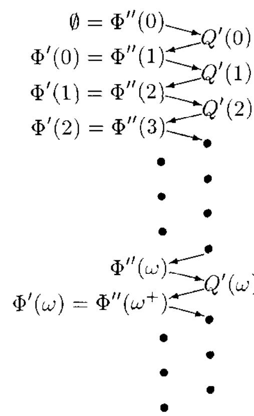

Fig. 13. Illustration of who is defined using whom.

We have

$$f \in \Phi'(\alpha) \;\Leftrightarrow\; \exists g \in \check{\Phi}'(\alpha)\colon g \leq f,$$

$$f \in Q'(\alpha) \;\Leftrightarrow\; \exists g \in \check{Q}'(\alpha)\colon g \leq f,$$

$$f \in \Phi''(\alpha) \;\Leftrightarrow\; \exists g \in \check{\Phi}''(\alpha)\colon g \leq f,$$

$$f \in \Phi \;\Leftrightarrow\; \exists g \in \check{\Phi}\colon g \leq f.$$

Fig. 14 shows $\check{\Phi}'(\alpha)$ and $\check{Q}'(\alpha)$, and illustrates the relationship between $\check{\Phi}'(2)$ and $\Phi'(2)$. $\check{\Phi}'(2)$ is the vertical line which forms the lower boundary of $\Phi'(2)$.

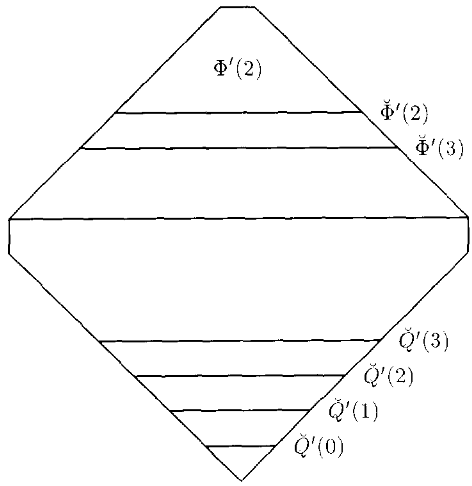

Fig. 14. The relation between $\breve{\Phi}'(2)$ and $\Phi'(2)$.

As we shall see later, $\breve{Q}'(0) <_\kappa \breve{\Phi}'(0) <_\kappa \breve{Q}'(1) <_k \breve{\Phi}'(1) <_\kappa \ldots$ (actually, we merely prove $\breve{Q}'(0) <_\kappa \breve{\Phi}'(0) \leqslant_\kappa \breve{Q}'(1) <_\kappa \breve{\Phi}'(1) \leqslant_\kappa \cdots$). These relative sizes of the $\breve{Q}'(\alpha)$ and $\breve{\Phi}'(\beta)$ are illustrated by the length of the lines in Fig. 14.

For all $x \in \Phi'(\beta)$ there is a $y \in \breve{\Phi}'(\beta)$ such that $y \leqslant x$. In particular, if $\alpha \leqslant_o \beta$ and $x \in \breve{\Phi}'(\alpha)$ then there is a $y \in \breve{\Phi}'(\beta)$ such that $y \leqslant x$. Fig. 15 displays some typical relations $y \leqslant x$ using the conventions from Fig. 8.

Let $H \subseteq M$ be closed. Assume $f \in wf(\partial H)$ and $x_1, x_2, \ldots \in H$. Choose $y_1, y_2, \ldots \in \partial H$ such that $y_i \leqslant x_i$ for all $i \in \{1, 2, \ldots\}$. Choose $\alpha \in \omega$ such that $(fy_1 \ldots y_\alpha) = \mathsf{T}$. From $\mathsf{T} = (fy_1 \ldots y_\alpha) \leqslant (fx_1 \ldots x_\alpha)$ we have $(fx_1 \ldots x_\alpha) = \mathsf{T}$. Hence, $\forall x_1, x_2, \ldots \in H\ \exists \alpha \in \omega\colon (fx_1 \ldots x_\alpha) = \mathsf{T}$, so $f \in wf(H)$. Hence, $wf(\partial H) \subseteq wf(H)$. From $\partial H \subseteq H$ we have $wf(H) \subseteq wf(\partial H)$ so $wf(\partial H) = wf(H)$. Hence,

$$\Phi'(\alpha) = wf(Q'(\alpha)) = wf(\partial \breve{Q}'(\alpha)),$$

$$Q'(\alpha) = wf(\Phi''(\alpha)) = wf(\partial \breve{\Phi}''(\alpha)),$$

We now have

$$\breve{\Phi}'(\alpha) = \partial wf(\breve{Q}'(\alpha)), \qquad \breve{Q}'(\alpha) = \partial wf(\breve{\Phi}''(\alpha)),$$

so $\breve{\Phi}'(\alpha)$, $\breve{Q}'(\alpha)$, $\breve{\Phi}''(\alpha)$, $\breve{\Phi}$ and $\breve{Q}$ are definable without reference to $\Phi'(\alpha)$ and $Q'(\alpha)$.

10.5. Types of well-founded maps

Assume that $f \in \partial wf(G)$ and $f =^a_G g$. Choose $h \in M$ such that $h \leqslant f$, $h \leqslant g$ and $h =^a_G f =^a_G g$. From $h =^a_G f$ and $f \in wf(G)$ we have $h \in wf(G)$. From the minimality

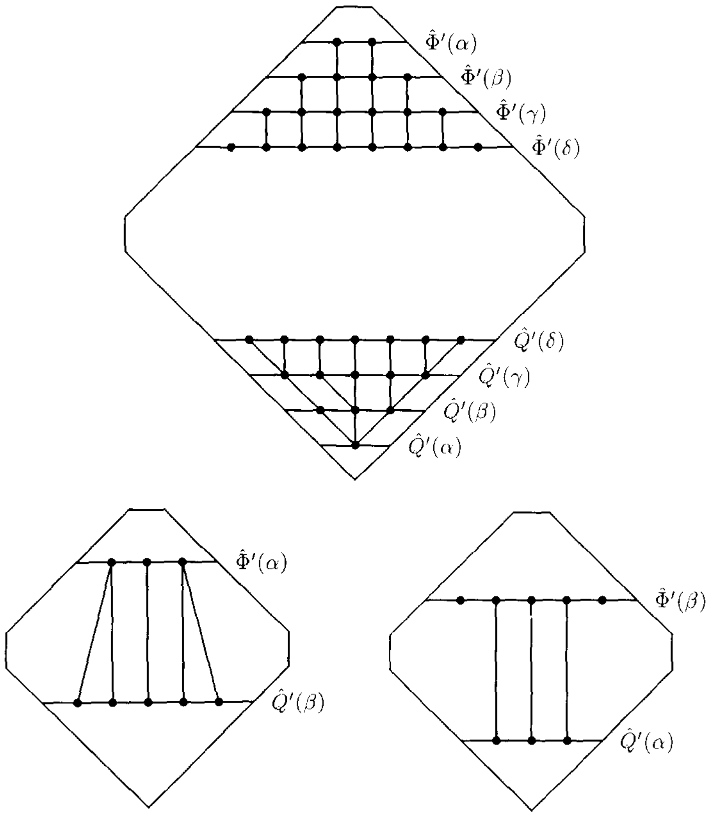

Fig. 15. Illustrations of the $\leqslant$ relation.

of f and $h \leqslant f$ we have $h = f$. Hence, $f = h \leqslant g$, so if $f \in \partial wf(G)$ then $f =^a_G g \Rightarrow f \leqslant g$. We have previously proved $f \leqslant g \Rightarrow f =^a_G g$ if $f \in wf(G)$. Hence, if $f \in \partial wf(G)$ then

$$f =^a_G g \;\Leftrightarrow\; f \leqslant g.$$

In particular,

$$f =^a_{\check{Q}'(\alpha)} h \;\Leftrightarrow\; f \leqslant h, \qquad g =^a_{\check{\Phi}''(\alpha)} h \;\Leftrightarrow\; g \leqslant h,$$

for $f \in \check{\Phi}'(\alpha)$, $g \in \check{Q}'(\alpha)$ and $h \in M$.

If $f, g \in \partial wf(G)$ then $t_G(f) = t_G(g) \Leftrightarrow f =^a_G g \Leftrightarrow f \leqslant g \wedge g \leqslant f \Leftrightarrow f = g$, so the elements of $\partial wf(G)$ are uniquely determined by their type $t_G(f)$. In particular, the elements f of $\check{\Phi}'(\alpha)$ are uniquely determined by $t_{\check{Q}'(\alpha)}(f)$ and the elements of $\check{Q}'(\alpha)$ by $t_{\check{\Phi}''(\alpha)}(f)$.

We now identify $f \in \check{\Phi}'(\alpha)$ with $t_{\check{Q}'(\alpha)}(f)$ and $g \in \check{Q}'(\alpha)$ with $t_{\check{\Phi}''(\alpha)}(g)$. We have to be slightly cautious because $\mathsf{T} \in \check{\Phi}'(\alpha)$ and $\mathsf{T} \in \check{Q}'(\alpha)$ for all α, so T is identified with many different types.

As a consequence of the identification we have

$$\check{\Phi}'(\alpha) \subseteq \check{Q}'(\alpha)^\circ, \qquad \check{Q}'(\alpha) \subseteq \check{\Phi}''(\alpha)^\circ.$$

If $f \in \check{\Phi}'(\alpha)$, then $f \in \check{Q}'(\alpha)^\circ$, so $f^{\mathrm{d}} = \check{Q}'(\alpha)^*$ and $f^{\mathrm{dR}} = \check{Q}'(\alpha)$, so

$$f \in \check{\Phi}'(\alpha) \Rightarrow f^{\mathrm{dR}} = \check{Q}'(\alpha),$$

$$f \in \check{Q}'(\alpha) \Rightarrow f^{\mathrm{dR}} = \check{\Phi}''(\alpha).$$

As mentioned earlier, $\Phi'(\alpha) \subseteq Q'(\beta)$ for all α and β. If $f \in \check{\Phi}'(\alpha) \subseteq \Phi'(\alpha) \subseteq Q'(\beta)$ then there is a unique $g \in \check{Q}'(\beta)$ such that $g \leqslant f$. From $g \leqslant f$ we have $g =^{a}_{\Phi''(\beta)} f$, so $t_{\Phi'(\beta)}(g) = t_{\Phi'(\beta)}(f)$. Since we have identified g and $t_{\Phi''(\beta)}(g)$, we have $g = t_{\Phi''(\beta)}(f)$. From the definition of $t_{\Phi''(\beta)}(f)$ we obtain $g = v \in \check{\Phi}''(\beta)^* \mapsto a(f)(v)$. Hence, if $f \in \check{\Phi}'(\alpha)$ then

$$[v \in \check{\Phi}''(\beta)^* \mapsto a(f)(v)] \in \check{Q}'(\beta),$$

$$[v \in \check{\Phi}''(\beta)^* \mapsto a(f)(v)] \leqslant f.$$

If $f \in \check{\Phi}'(\alpha)^*$ then

$$[u \in f^{\mathrm{d}} \mapsto v \in \check{\Phi}''(\beta)^* \mapsto a(f(u))(v)] \in \check{Q}'(\beta)^*,$$

$$[u \in f^{\mathrm{d}} \mapsto v \in \check{\Phi}''(\beta)^* \mapsto a(f(u))(v)] \leqslant^* f.$$

If $f \in \check{\Phi}'(\beta)$ and $x \in \check{Q}'(\beta)^*$, then, by the definition of $t_{\check{Q}'(\beta)}$ we have $t_{\check{Q}'(\beta)}(f)(x) = a(f)(x)$. Since we have identified f and $t_{\check{Q}'(\beta)}(f)$ we have

$$a(f)(x) = f(x).$$

If $x \in \check{\Phi}'(\beta)$ and $y \in \check{\Phi}^*$, then let $y' = u \in y^{\mathrm{d}} \mapsto v \in \check{\Phi}''(\beta)^* \mapsto a(y(u))(v)$. We have $y' \in \check{Q}'(\beta)^*$ and $y' \leqslant^* y$, so $x(y') = a(x)(y') \leqslant_L a(x)(y)$. Since $x(y') \neq \tilde{\bot}$ we have $x(y') = a(x)(y)$. Since $x^{\mathrm{dR}} = \check{Q}'(\beta)$ and $x^{\mathrm{dRD}} = \check{\Phi}''(\beta)^*$ we may write this result as

$$a(x)(y) = x(u \in y^{\mathrm{d}} \mapsto v \in x^{\mathrm{dRD}} \mapsto a(y(u))(v)).$$

This equation is essential since it makes no reference to yet undefined concepts like M and it uniquely determines $a(x)(y)$ for all x and y. In a slightly modified form, it is going to be the first formal definition in the construction of M. The function defined by the equation gives the correct value for $a(x)(y)$ whenever $x \in \check{\Phi}$ and $y \in \check{\Phi}^*$ (but we have not yet defined $\check{\Phi}$). Since the function defined in the equation does not give the correct value for $a(x)(y)$ for all $x \in M$ and $y \in M^*$, we shall refer to it as $\hat{a}$.

10.6. The construction of $\check{\Phi}$

If $f \in \check{\Phi}'(\alpha)$ then we have seen that $v \in \check{\Phi}''(\beta)^* \mapsto a(f)(v) \in \check{Q}'(\beta)$. Let $J = f \in \check{\Phi}'(\alpha) \mapsto v \in \check{\Phi}''(\beta)^* \mapsto a(f)(v)$. We have $J \in \check{\Phi}'(\alpha) \to \check{Q}'(\beta)$. We shall define M

such that J is surjective (or onto or epimorphic) if $\beta \leqslant_{\mathrm{o}} \alpha$ and injective (or one-to-one or monomorphic) if $\alpha <_{\mathrm{o}} \beta$. This gives rise to the following cardinalities:

$$\emptyset = \Phi''(0) \leqslant_{\kappa} \check{Q}'(0) \leqslant_{\kappa} \Phi'(0) \leqslant_{\kappa} \check{\Phi}''(1) \leqslant_{\kappa} \cdots \leqslant_{\kappa} \Phi''(\omega)$$
$$\leqslant_{\kappa} \check{Q}'(\omega) \leqslant_{\kappa} \check{\Phi}'(\omega) \leqslant_{\kappa} \cdots .$$

Further, since J is surjective for $\beta \leqslant_{\mathrm{o}} \alpha$, we have

$$\forall x \in \check{Q}'(\beta)\, \exists y \in \check{\Phi}''(\alpha): x \leqslant y$$

for all $\beta <_{\mathrm{o}} \alpha$.

Lemma 10.6.1. *Let* $f, g \in \check{\Phi}''(\alpha)$. *We have*

$$f =^{a}_{\Phi''(\alpha)} g \Leftrightarrow f \leqslant g \vee g \leqslant f.$$

Proof. Since we have not yet formally defined the concepts involved in the lemma, the proof is merely based on the assumptions we have made about these concepts. Assume $f \leqslant g \vee g \leqslant f$. Without loss of generality assume $g \leqslant f$ (see Fig. 16). Let h be the unique element of $\check{Q}'(\alpha)$ for which $h \leqslant g$. From $h \leqslant g \leqslant f$ we have $h \leqslant f$. From $h \in Q'(\alpha)$, $h \leqslant f$, and $h \leqslant g$ we have $h = t_{\Phi''(\alpha)}(f)$ and $h = t_{\Phi''(\alpha)}(g)$. Hence, $t_{\Phi''(\alpha)}(f) = t_{\Phi''(\alpha)}(g)$ which proves $f =^{a}_{\Phi''(\alpha)} g$.

Now assume $f =^{a}_{\Phi''(\alpha)} g$. Let β and γ be such that $f \in \check{\Phi}'(\beta)$ and $g \in \check{\Phi}'(\gamma)$. Without loss of generality assume $\beta \leqslant_{\mathrm{o}} \gamma$. Let $x_\gamma \in \check{Q}'(\gamma)^*$. Choose $x_\alpha \in \check{\Phi}''(\alpha)^*$ such that $x_\gamma \leqslant^* x_\alpha$. We have $x_\gamma \in \check{Q}'(\gamma)^* \subseteq Q'(\gamma)^* \subseteq Q'(\beta)^*$. Choose $x_\beta \in \check{Q}'(\beta)^*$ such that $x_\beta \leqslant^* x_\gamma$. We have $\tilde{\bot} \neq a(f)(x_\beta) \leqslant_L a(f)(x_\gamma) \leqslant_L a(f)(x_\alpha) = a(g)(x_\alpha) \geqslant_L a(f)(x_\gamma) \neq \tilde{\bot}$ which proves $a(f)(x_\gamma) = a(g)(x_\gamma)$. Since this holds for all $x_\gamma \in \check{Q}'(\gamma)^*$ we have $f =^{a}_{\check{Q}'(\gamma)} g$ which entails $g \leqslant f$. This concludes the proof. $\square$

If $f \in \check{Q}'(\alpha)$ and $x, y \in \check{\Phi}''(\alpha)^*$, $x =^{a}_{\Phi''(\alpha)} y$, then $a(f)(x) = a(f)(y)$ follows from the lemma. Hence, $\check{Q}'(\alpha) \subseteq \check{\Phi}''(\alpha)^{\circ}$ can be narrowed down to

$$\check{Q}'(\alpha) \subseteq \{f \in \check{\Phi}''(\alpha)^{\circ} \mid \forall x, y \in \check{\Phi}''(\alpha)^*: (x =^{a}_{\Phi''(\alpha)} y \Rightarrow f(x) = f(y))\}.$$

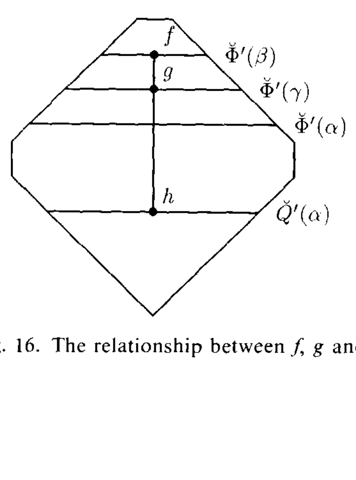

Fig. 16. The relationship between f, g and h.

To sum up we have

$$\check{\Phi}'(\alpha) \subseteq \check{Q}'(\alpha)^{\circ}, \tag{13}$$

$$\check{Q}'(\alpha) \subseteq \{f \in \check{\Phi}''(\alpha)^{\circ} \mid \forall x, y \in \check{\Phi}''(\alpha)^{*}: (x =^{\alpha *}_{\check{\Phi}''(\alpha)} y \Rightarrow f(x) = f(y))\}, \tag{14}$$

$$\check{\Phi}''(\alpha) = \bigcup_{\beta \in \alpha} \check{\Phi}'(\beta), \tag{15}$$

$$\check{\Phi} = \bigcup_{\beta \in On} \check{\Phi}'(\beta), \tag{16}$$

$$\check{Q} = \bigcup_{\beta \in On} \check{Q}'(\beta). \tag{17}$$

Again, these statements are independent of M, but they cannot be used as a definition of $\check{\Phi}$ because (13) and (14) do not determine $\check{\Phi}'(\alpha)$ and $\check{Q}'(\alpha)$ uniquely.

The idea behind the model construction is to replace $\subseteq$ by $=$ in (13) and (14). This certainly produces a universe $\check{\Phi}$ large enough to represent any set, but at the same time $\check{\Phi}$ becomes a class, and the model construction later on requires $\check{\Phi}$ to be a set. To get around this problem, we define

$$\hat{\Phi}'(\alpha) = \hat{Q}'(\alpha)^{\circ},$$

$$\hat{Q}'(\alpha) = \{f \in \hat{\Phi}''(\alpha)^{\circ} \mid \forall x, y \in \hat{\Phi}''(\alpha)^{*}: (x =^{\alpha *}_{\hat{\Phi}''(\alpha)} y \Rightarrow f(x) = f(y))\},$$

$$\hat{\Phi}''(\alpha) = \bigcup_{\beta \in \alpha} \hat{\Phi}'(\beta),$$

$$\hat{\Phi} = \bigcup_{\beta \in On} \hat{\Phi}'(\beta),$$

$$\hat{Q} = \bigcup_{\beta \in On} \hat{Q}'(\beta),$$

and then we define $\check{\Phi}'(\alpha)$, $\check{Q}'(\alpha)$, $\check{\Phi}''(\alpha)$, $\check{\Phi}$ and $\check{Q}$ as the relativization to D of $\hat{\Phi}'(\alpha)$, $\hat{Q}'(\alpha)$, $\hat{\Phi}''(\alpha)$, $\hat{\Phi}$ and $\hat{Q}$, respectively. This ensures that $\check{\Phi}$ gets the "richness" of D and a manageable size at the same time.

Define $\hat{s}(x) = \{\hat{s}(x') \mid x' <_{w} x\}$. We say that $x \in \hat{\Phi}$ represents the set $\hat{s}(x)$. The role of $\hat{\Phi}$ in the model construction is to ensure that the model of map theory becomes rich enough to allow representation of any set of set theory within the model. A central property of $\hat{\Phi}$ is the theorem $\forall y\, \exists x \in \hat{\Phi}: y = \hat{s}(x)$, which we shall refer to as the Adequacy Theorem. Even though it is central, the Adequacy Theorem is only used when verifying the consistency of $Map^{\circ +}$.

10.7. The syntactic model

In addition to the semantic model M and the models $\hat{\Phi}$ and $\check{\Phi}$ of well-founded maps we shall construct a syntactic model $\acute{M}$.

The definition of $\acute{M}$ is simple in itself: It is the least set such that

$$\{\dot{S}, \dot{K}, \dot{\mathsf{T}}, \dot{P}, \dot{C}, \dot{W}, \dot{\perp}\} \subseteq \acute{M},$$

$$\forall x, y \in \acute{M}: \dot{A}(x, y) \in \acute{M},$$

$$\check{\Phi} \subseteq \acute{M},$$

$$\check{Q} \subseteq \acute{M}.$$

The elements of $\acute{M}$ are Gödel numbers of closed terms of map theory (i.e. terms with no free variables). However, we allow elements of $\check{\Phi}$ and $\check{Q}$ to occur in the terms.

Each term $x \in \acute{M}$ denotes a map $c(x) \in M$. We define $\acute{r}(x)$ to be the root of the map denoted by x, i.e.

$$\acute{r}(x) = r(c(x)).$$

Likewise, we define

$$\acute{a}(x)(\langle y_1, \ldots, y_\alpha \rangle) = a(c(x))(\langle c(y_1), \ldots, c(y_\alpha) \rangle),$$

$$x \mathrel{\acute{\leqslant}} y \Leftrightarrow c(x) \leqslant c(y),$$

$$x \mathrel{\acute{=}} y \Leftrightarrow c(x) = c(y).$$

The function $\acute{m} \in \acute{M} \times \acute{M}^* \rightarrow \acute{M}$ is defined by

$$\acute{m}(x, \langle y_1, \ldots, y_\alpha \rangle) = \dot{(} x\, y_1 \ldots y_\alpha \dot{)}.$$

10.8. The actual order of the formal definitions

Above, a number of concepts have been introduced based on the yet undefined set M of all maps. The formal definitions are stated in another order:

- First, $\hat{a}$, $\hat{\Phi}$ and $\hat{Q}$ are defined as outlined above.
- Second, $\check{a}$, $\check{\Phi}$ and $\check{Q}$ are defined by relativization of $\hat{a}$, $\hat{\Phi}$ and $\hat{Q}$, respectively.
- Third, $\acute{r}$ is defined by a fixed point construction, and $\acute{a}$ and $\acute{=}$ are defined from $\acute{r}$.
- Last, M is formed by the quotient construction $M = \acute{M}/\acute{=}$, i.e., the elements of M are equivalence classes of $\acute{M}$ under $\acute{=}$.

11. Construction of the model

11.1. Definition of well-founded maps

Define $<_{\rho\rho}$ by $\langle u, v \rangle <_{\rho\rho} \langle x, y \rangle \Leftrightarrow u \in y^{\mathrm{r}} \wedge v \in x^{\mathrm{dRD}}$. If $\langle u, v \rangle <_{\rho\rho} \langle x, y \rangle$ then $u <_\rho y$ and $v <_\rho x$. Hence, $<_{\rho\rho}$ is well-founded and set-like. As foreseen in Section 10.5 define $\hat{a}$ by transfinite recursion in $<_{\rho\rho}$:

$$\hat{a}(x, y) = x(u \in y^{\mathrm{d}} \mapsto v \in x^{\mathrm{dRD}} \mapsto \hat{a}(y(u), v)).$$

For all sets G define $x \triangleq^a_G y \Leftrightarrow \forall z \in G^*\colon \hat{a}(x, z) = \hat{a}(y, z)$. Following Section 10.6 and by transfinite recursion in α define

$$\hat{\Phi}'(\alpha) = \hat{Q}'(\alpha)^\circ,$$

$$\hat{\Phi}''(\alpha) = \bigcup_{\beta \in \alpha} \hat{\Phi}'(\alpha),$$

$$\hat{Q}'(\alpha) = \{f \in \hat{\Phi}''(\alpha)^\circ \mid \forall x, y \in \hat{\Phi}''(\alpha)^*\colon (x \triangleq^{a*}_{\hat{\Phi}''(\alpha)} y \Rightarrow f(x) = f(y))\}.$$

Define $\hat{\Phi}(x) \Leftrightarrow \exists \alpha\colon x \in \hat{\Phi}'(\alpha)$ and $\hat{Q}(x) \Leftrightarrow \exists \alpha\colon x \in \hat{Q}'(\alpha)$.

We shall work with proper classes like [17], e.g., we introduce $x \in \hat{\Phi}$ as shorthand for $\hat{\Phi}(x)$, $x \subseteq \hat{\Phi}$ as shorthand for $\forall y \in x\colon y \in \hat{\Phi}$ and $x \in \hat{\Phi}^*$ as shorthand for $tpl(x) \wedge x^{\mathrm{r}} \subseteq \hat{\Phi}$ etc.

In accordance with Section 10.2, define the type $\hat{t}_G(x)$ of x w.r.t. G by

$$\hat{t}_G(x) = v \in G^* \mapsto \hat{a}(x, v).$$

Define $\hat{t}^*_G(x)$ as the coordinatewise application of $\hat{t}_G$:

$$\hat{t}^*_G(x) = u \in x^{\mathrm{d}} \mapsto \hat{t}_G(x(u)).$$

For $G \subseteq \hat{\Phi}$ we have

$$x \triangleq^a_G y \;\Leftrightarrow\; \hat{t}_G(x) = \hat{t}_G(y) \qquad (x, y \in \hat{\Phi})$$

$$x \triangleq^{a*}_G y \;\Leftrightarrow\; \hat{t}^*_G(x) = \hat{t}^*_G(y) \qquad (x, y \in \hat{\Phi}^*)$$

$$\hat{t}^*_G(x \cdot y) \;=\; \hat{t}^*_G(x) \cdot \hat{t}^*_G(y) \qquad (x, y \in \hat{\Phi}^*)$$

$$\hat{t}^*_G(x|\alpha) \;=\; (\hat{t}^*_G(x)|\alpha) \qquad (x \in \hat{\Phi}^\omega)$$

$$\hat{a}(x, y) \;=\; x(\hat{t}^*_{x^{\mathrm{dRDR}}}(y)) \qquad (x, y \in \hat{\Phi})$$

$$\hat{a}(x, y) \;=\; x(\hat{t}^*_{\hat{\Phi}''(\alpha)}(y)) \qquad (x \in \hat{\Phi}'(\alpha), y \in \hat{\Phi}).$$

11.2. Properties of well-founded maps

Define

$$R'_{\alpha,\beta} \;\Leftrightarrow\; \forall x \in \hat{\Phi}'(\alpha)\colon \hat{t}_{\hat{\Phi}''(\beta)}(x) \in \hat{Q}'(\beta),$$

$$I'_{\alpha,\beta} \;\Leftrightarrow\; (\alpha <_{\mathrm{o}} \beta \Rightarrow \forall x, y \in \hat{\Phi}'(\alpha)\colon (x \neq y \Rightarrow \hat{t}_{\hat{\Phi}''(\beta)}(x) \neq \hat{t}_{\hat{\Phi}''(\beta)}(y))),$$

$$S'_{\alpha,\beta} \;\Leftrightarrow\; (\alpha \geqslant_{\mathrm{o}} \beta \Rightarrow \forall x \in \hat{Q}'(\beta)\, \exists y \in \hat{\Phi}'(\alpha)\colon x = \hat{t}_{\hat{\Phi}''(\beta)}(y)).$$

The formula $R'_{\alpha,\beta}$ states that $\hat{t}_{\hat{\Phi}'(\beta)}$ maps $\hat{\Phi}'(\alpha)$ into $\hat{Q}'(\beta)$, $I'_{\alpha,\beta}$ states that $\hat{t}_{\hat{\Phi}'(\beta)}$ is injective (i.e. one-to-one, or monomorphic) if $\alpha \leqslant_{\mathrm{o}} \beta$ and $S'_{\alpha,\beta}$ states that $\hat{t}_{\hat{\Phi}'(\beta)}$ is surjective (i.e. onto, or epimorphic) if $\alpha \geqslant_{\mathrm{o}} \beta$.

Theorem 11.2.1. *For all ordinals α and β, $R'_{\alpha,\beta}$, $I'_{\alpha,\beta}$ and $S'_{\alpha,\beta}$ hold.*

Theorem 11.2.1 justifies several claims from Section 10. In particular, it justifies Fig. 15. As we shall see, the relativization of Theorem 11.2.1 justifies $\check{t}_{\Phi''(\beta)}(x) \in \check{Q}'(\beta)$ and the claims in the beginning of Section 10.6.

Corollary 11.2.2. *We have*

$$\forall x \in \hat{\Phi}\ \forall y \in \hat{\Phi}^*\colon \hat{a}(x, y) \in \{\tilde{\mathsf{T}}, \tilde{\lambda}\},$$

$$\forall x \in \hat{\Phi}\ \forall y \in \hat{\Phi}^*\colon \hat{a}(x, y) = x\langle\!\langle \hat{t}^*_{x^{\mathrm{dRDR}}}(\mathrm{y})\rangle\!\rangle,$$

$$\forall x \in \hat{\Phi}'(\alpha)\ \forall y \in \hat{\Phi}^*\colon \hat{a}(x, y) = x\langle\!\langle \hat{t}^*_{\hat{\Phi}''(\alpha)}(y)\rangle\!\rangle,$$

$$\forall x \in \hat{\Phi}\ \forall y \in \hat{\Phi}^{\omega}\ \exists \alpha \in \omega\colon \hat{a}(x, y|\alpha) = \tilde{\mathsf{T}}.$$

The corollary follows from $R'_{\alpha,\beta}$.

Proof of Theorem 11.2.1. Let $\langle \gamma, \delta\rangle <_{\mathrm{oo}} \langle \alpha, \beta\rangle$ stand for

$$(\gamma <_{\mathrm{o}} \alpha \wedge \delta \leqslant_{\mathrm{o}} \beta) \vee (\gamma \leqslant_{\mathrm{o}} \alpha \wedge \delta <_{\mathrm{o}} \beta) \vee (\gamma <_{\mathrm{o}} \beta \wedge \delta \leqslant_{\mathrm{o}} \alpha)$$
$$\vee (\gamma \leqslant_{\mathrm{o}} \beta \wedge \delta <_{\mathrm{o}} \alpha).$$

We have that $<_{\mathrm{oo}}$ is well-founded. Now assume that α and β are ordinals and assume

$$\langle \gamma, \delta\rangle <_{\mathrm{oo}} \langle \alpha, \beta\rangle \Rightarrow R'_{\gamma,\delta}, \tag{18}$$

$$\langle \gamma, \delta\rangle <_{\mathrm{oo}} \langle \alpha, \beta\rangle \Rightarrow S'_{\gamma,\delta}, \tag{19}$$

$$\langle \gamma, \delta\rangle <_{\mathrm{oo}} \langle \alpha, \beta\rangle \Rightarrow I'_{\gamma,\delta}. \tag{20}$$

If we can prove $R'_{\alpha,\beta}$, $I'_{\alpha,\beta}$ and $S'_{\alpha,\beta}$ from these assumptions, then the theorem follows by transfinite induction on $<_{\mathrm{oo}}$.

For $\langle \gamma, \delta\rangle <_{\mathrm{oo}} \langle \alpha, \beta\rangle$ we have

$$\forall x \in \hat{\Phi}'(\gamma)^*\colon \hat{t}^*_{\hat{\Phi}'(\delta)}(x) \in \hat{Q}'(\delta)^*, \tag{21}$$

$$\gamma <_{\mathrm{o}} \delta \Rightarrow \forall x, y \in \hat{\Phi}'(\gamma)^*\colon (x \neq y \Rightarrow \hat{t}^*_{\hat{\Phi}''(\delta)}(x) \neq \hat{t}^*_{\hat{\Phi}''(\beta)}(y)), \tag{22}$$

$$\gamma \geqslant_{\mathrm{o}} \delta \Rightarrow \forall x \in \hat{Q}'(\delta)^*\ \exists y \in \hat{\Phi}'(\gamma)^*\colon x = \hat{t}^*_{\hat{\Phi}''(\delta)}(y). \tag{23}$$

We now prove

$$\forall x, y \in \hat{\Phi}''(\beta)\colon (x \mathrel{\hat{=}}^{a}_{\hat{\Phi}''(\beta)} y \Rightarrow x \mathrel{\hat{=}}^{a}_{\hat{\Phi}''(\alpha)} y). \tag{24}$$

Assume $x, y \in \hat{\Phi}''(\beta)$ and $x \mathrel{\hat{=}}^{a}_{\hat{\Phi}''(\beta)} y$. If $\alpha \leqslant_{\mathrm{o}} \beta$ then $\hat{\Phi}''(\alpha) \subseteq \hat{\Phi}''(\beta)$ from which $x \mathrel{\hat{=}}^{a}_{\hat{\Phi}''(\alpha)} y$ follows. Now assume $\beta <_{\mathrm{o}} \alpha$. We have $\langle \gamma, \beta\rangle <_{\mathrm{oo}} \langle \alpha, \beta\rangle$ for all $\gamma \leqslant_{\mathrm{o}} \beta$.

Let δ, ε be such that $x \in \hat{\Phi}'(\delta)$ and $y \in \hat{\Phi}'(\varepsilon)$. Without loss of generality assume $\delta \leqslant_{\mathrm{o}} \varepsilon$ (see Fig. 17). Let $u \in \hat{\Phi}''(\alpha)^*$ and define $u' = \hat{t}^*_{\hat{\Phi}''(\varepsilon)}(u)$. From $R'_{\alpha,\varepsilon}$ we have $u' \in \hat{Q}'(\varepsilon)^*$. Choose $u'' \in \hat{\Phi}''(\beta)$ such that $\hat{t}^*_{\hat{\Phi}''(\varepsilon)}(u'') = u'$. This is possible since $S'_{\beta,\varepsilon}$

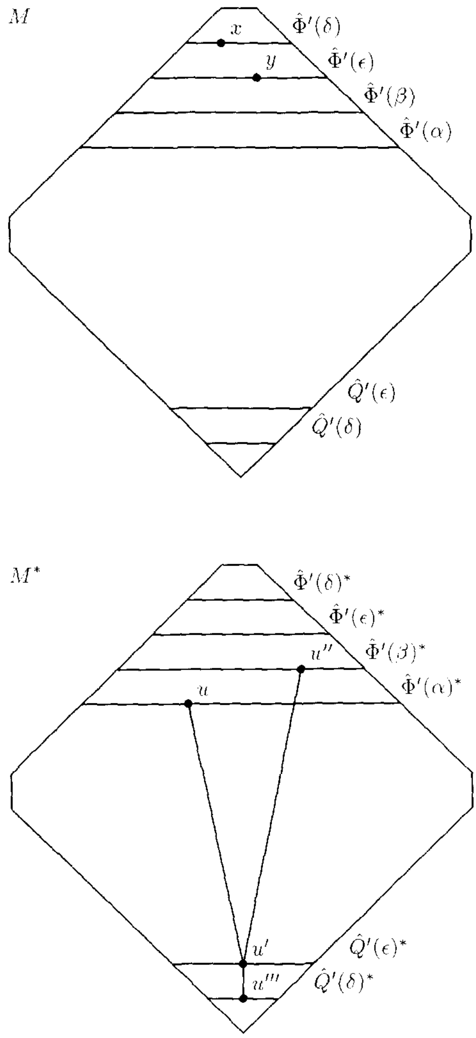

Fig. 17. The relationship between u, u', u'' and u'''.

holds. Let $u''' = \hat{t}^*_{\hat{\Phi}''(\delta)}(u)$. And we have $u''' = \hat{t}^*_{\hat{\Phi}''(\delta)}(u) = (\hat{t}^*_{\hat{\Phi}''(\epsilon)}(u) | \hat{\Phi}''(\delta)^*) = (\hat{t}^*_{\hat{\Phi}''(\epsilon)}(u'') | \hat{\Phi}''(\delta)^*) = \hat{t}^*_{\hat{\Phi}''(\delta)}(u'')$. Therefore, $\hat{a}(x, u) = x(\hat{t}^*_{\hat{\Phi}''(\delta)}(u)) = x(\hat{t}^*_{\hat{\Phi}''(\delta)}(u'')) = \hat{a}(x, u'')$. Likewise, $\hat{a}(y, u) = \hat{a}(y, u'')$. From $x \triangleq^{a}_{\hat{\Phi}'(\beta)} y$ we have $\hat{a}(x, u'') = \hat{a}(y, u'')$. Hence, $\hat{a}(x, u) = \hat{a}(y, u)$. Since this holds for all $u \in \hat{\Phi}''(\alpha)^*$, we have proved (24).

We now prove $R'_{\alpha,\beta}$. Assume $x \in \hat{\Phi}'(\alpha)$. From $\hat{\Phi}'(\alpha) = \hat{Q}'(\alpha)^\circ$ we obtain

$$\forall v' \in \hat{Q}'(\alpha)^*: x(v') \in \{\tilde{\mathsf{T}}, \tilde{\lambda}\},$$

$$\forall v', w' \in \hat{Q}'(\alpha)^*: (x(v') \neq \tilde{\lambda} \Rightarrow x(v' \cdot w') = x(v')),$$

$$\forall v' \in \hat{Q}'(\alpha)^\omega\ \exists \delta \in \omega: x(v'|\delta) = \tilde{\mathsf{T}}.$$

From $\forall\gamma\in\beta\colon R'_{\gamma,\alpha}$ we deduce $\forall v\in\hat{\Phi}''(\beta)\colon \hat{t}_{\hat{\Phi}''(\alpha)}(v)\in\hat{Q}'(\alpha)$. Hence

$$\forall v\in\hat{\Phi}''(\beta)^*\colon x(\hat{t}^*_{\hat{\Phi}''(\alpha)}(v))\in\{\tilde{\mathsf{T}},\tilde{\lambda}\},$$

$$\forall v,w\in\hat{\Phi}''(\beta)^*\colon (x(\hat{t}^*_{\hat{\Phi}''(\alpha)}(v))\neq\tilde{\lambda}\Rightarrow x(\hat{t}^*_{\hat{\Phi}''(\alpha)}(v\cdot w))=x(\hat{t}^*_{\hat{\Phi}''(\alpha)}(v))),$$

$$\forall v\in\hat{\Phi}''(\beta)^\omega\ \exists\delta\in\omega\colon x(\hat{t}^*_{\hat{\Phi}''(\alpha)}(v|\delta))=\tilde{\mathsf{T}}.$$

Define $y=\hat{t}_{\hat{\Phi}''(\beta)}(x)$. For all $v\in\hat{\Phi}''(\beta)^*$ we have $y(v)=\hat{t}_{\hat{\Phi}''(\beta)}(x)(v)=\hat{a}(x,v)=x(\hat{t}^*_{\hat{\Phi}''(\alpha)}(v))$. Hence

$$\forall v\in\hat{\Phi}''(\beta)^*\colon y(v)\in\{\tilde{\mathsf{T}},\tilde{\lambda}\},$$

$$\forall v,w\in\hat{\Phi}''(\beta)^*\colon (y(v)\neq\tilde{\lambda}\Rightarrow y(v\cdot w)=y(v)),$$

$$\forall v\in\hat{\Phi}''(\beta)^\omega\ \exists\delta\in\omega\colon y(v|\delta)=\tilde{\mathsf{T}},$$

which proves $y\in\hat{\Phi}''(\beta)^\circ$.

Now let $u,v\in\hat{\Phi}''(\beta)^*$ satisfy $u\triangleq^{a*}_{\hat{\Phi}''(\beta)}v$. From (24) we have $u\triangleq^{a*}_{\hat{\Phi}''(\alpha)}v$, so $y(u)=x(\hat{t}^*_{\hat{\Phi}''(\alpha)}(u))=x(\hat{t}^*_{\hat{\Phi}''(\alpha)}(v))=y(v)$. Hence, $\forall u,v\in\hat{\Phi}''(\beta)^*\colon (u\triangleq^{a*}_{\hat{\Phi}''(\beta)}v\Rightarrow y(u)=y(v))$ which combined with $y\in\hat{\Phi}''(\beta)^\circ$ yields $y\in\hat{Q}'(\beta)$. Hence, $\forall x\in\hat{\Phi}'(\alpha)\colon \hat{t}^*_{\hat{\Phi}''(\beta)}(x)=y\in\hat{Q}'(\beta)$ which proves $R'_{\alpha,\beta}$.

We now prove $S'_{\alpha,\beta}$. If $\alpha<_\mathrm{o}\beta$ then $S'_{\alpha,\beta}$ holds trivially, so assume $\beta\leqslant_\mathrm{o}\alpha$. Let $x\in\hat{Q}'(\beta)$. Define $f=u\in\hat{\Phi}''(\beta)\mapsto\hat{t}^*_{\hat{\Phi}''(\alpha)}(u)$. From $\forall\gamma\in\beta\colon R'_{\gamma,\alpha}$ we obtain $f\in\hat{\Phi}''(\beta)\to\hat{Q}'(\alpha)$.

Let $u,v\in\hat{\Phi}''(\beta)^*$ satisfy $f(u)=f(v)$. From $f(u)=f(v)$ we obtain $u\triangleq^{a*}_{\hat{\Phi}''(\alpha)}v$ which combined with $\hat{\Phi}''(\beta)\subseteq\hat{\Phi}''(\alpha)$ gives $u\triangleq^{a*}_{\hat{\Phi}''(\beta)}v$. Since $x\in\hat{Q}'(\beta)$ we have $x(u)=x(v)$. Hence,

$$\forall u,v\in\hat{\Phi}''(\beta)^*\colon (f(u)=f(v)\Rightarrow x(u)=x(v)).$$

Now let $G=f^\mathrm{r}\subseteq\hat{Q}'(\alpha)$ and let $y'\in G^*\to\{\tilde{\mathsf{T}},\tilde{\lambda}\}$ be the unique function that satisfies

$$\forall u\in\hat{\Phi}''(\beta)^*\colon x(u)=y'(f\circ u).$$

From $x\in\hat{\Phi}''(\beta)^\circ$ one easily deduces $y'\in G^\circ$. Now define $y\in\hat{Q}'(\alpha)^*\to\{\tilde{\mathsf{T}},\tilde{\lambda}\}$ by

$$y(u)=\begin{cases} y'(u) & \text{if } u\in G^*,\\ \tilde{\mathsf{T}} & \text{otherwise.}\end{cases}$$

We have $y\in\hat{Q}'(\alpha)^\circ=\hat{\Phi}'(\alpha)$. Furthermore, for $u\in\hat{\Phi}''(\beta)^*$ we have $\hat{t}^*_{\hat{\Phi}''(\beta)}(y)(u)=\hat{a}(y,u)=y(\hat{t}^*_{\hat{\Phi}''(\alpha)}(u))=y(f(u))=y'(f(u))=x(u)$ which proves $\hat{t}^*_{\hat{\Phi}''(\beta)}(y)=x$. Hence, $\forall x\in\hat{Q}'(\beta)\ \exists y\in\hat{\Phi}''(\alpha)\colon \hat{t}^*_{\hat{\Phi}''(\beta)}(y)=x$ which proves $S'_{\alpha,\beta}$.

We now prove $I'_{\alpha,\beta}$. If $\beta\leqslant_\mathrm{o}\alpha$ then $I'_{\alpha,\beta}$ holds trivially, so assume $\alpha<_\mathrm{o}\beta$. Further assume $x,y\in\hat{\Phi}'(\alpha)$, $x\neq y$. Choose $z\in\hat{Q}'(\alpha)^*$ such that $x(z)\neq y(z)$. From $\alpha<_\mathrm{o}\beta$ and $\forall\gamma\in\beta\colon S'_{\gamma,\alpha}$ we obtain $S'_{\alpha,\alpha}$. Choose $z'\in\hat{\Phi}'(\alpha)^*$ such that $z=\hat{t}^*_{\hat{\Phi}''(\alpha)}(z')$. We have $\hat{a}(x,z')=x(\hat{t}^*_{\hat{\Phi}''(\alpha)}(z'))=x(z)\neq y(z)=x(\hat{t}^*_{\hat{\Phi}''(\alpha)}(z'))=\hat{a}(y,z')$. From $z'\in\hat{\Phi}'(\alpha)^*\subseteq\hat{\Phi}''(\beta)^*$ and $\hat{a}(x,z')\neq\hat{a}(y,z')$ we obtain $x\not\triangleq^{a}_{\hat{\Phi}''(\beta)}y$. Hence, $x\neq y\Rightarrow x\not\triangleq^{a}_{\hat{\Phi}''(\beta)}y$ which proves $I'_{\alpha,\beta}$. $\square$

11.3. Relativization of well-founded maps

We now relativize the model of well-founded maps to the transitive standard model D:

$$\check{\Phi} = \{x \in D \mid \lfloor \hat{\Phi}(x) \rfloor\},$$
$$\check{Q} = \{x \in D \mid \lfloor \hat{Q}(x) \rfloor\},$$
$$\check{\Phi}'(\alpha) = \lfloor \hat{\Phi}'(\alpha) \rfloor,$$
$$\check{\Phi}''(\alpha) = \lfloor \hat{\Phi}''(\alpha) \rfloor,$$
$$\check{Q}'(\alpha) = \lfloor \hat{Q}'(\alpha) \rfloor,$$
$$\check{a} = x \in \check{\Phi} \mapsto y \in \check{\Phi}^* \mapsto \lfloor \hat{a}(x, y) \rfloor.$$

We use the relativization and absoluteness results of [17, Chapter 4] without further notice. For example, we have

$$x \in \check{\Phi} \Leftrightarrow \lfloor x \in \hat{\Phi} \rfloor \qquad \text{for } x \in D,$$
$$x \subseteq \check{\Phi} \Leftrightarrow \lfloor x \subseteq \hat{\Phi} \rfloor \qquad \text{for } x \in D,$$
$$x \in \check{\Phi}^* \Leftrightarrow \lfloor x \in \hat{\Phi}^* \rfloor \qquad \text{for } x \in D,$$
$$\check{a}(x)(y) = \lfloor \hat{a}(x, y) \rfloor \qquad \text{for } x \in \check{\Phi} \text{ and } y \in \check{\Phi}^*.$$

11.4. Definition of the syntactic model

Let $\acute{M}$ be the least set such that

$$\{\dot{S}, \dot{K}, \dot{\mathsf{T}}, \dot{P}, \dot{C}, \dot{W}, \dot{\bot}\} \subseteq \acute{M},$$
$$\check{\Phi} \subseteq \acute{M},$$
$$\check{Q} \subseteq \acute{M},$$
$$\forall x, y \in \acute{M}: \dot{A}(x, y) \in \acute{M}.$$

By reading the somewhat arbitrary definitions of $\dot{S}$, $\dot{K}$, $\dot{\mathsf{T}}$, $\dot{P}$, $\dot{C}$, $\dot{W}$, $\dot{\bot}$, $\dot{A}(x, y)$, $\check{\Phi}$ and $\check{Q}$ very carefully one verifies that

$$\acute{M} = \{\dot{S}\} \cup \{\dot{K}\} \cup \{\dot{\mathsf{T}}\} \cup \{\dot{P}\} \cup \{\dot{C}\} \cup \{\dot{W}\} \cup \{\dot{\bot}\} \cup \bigcup_{x, y \in \acute{M}} \{\dot{A}(x, y)\} \cup \check{\Phi} \cup \check{Q}$$

is a direct sum.

Define $\acute{m}(x, \langle y_1, \ldots, y_\alpha \rangle) = \dot{(} x\, y_1 \ldots y_\alpha \dot{)}$. Let $\check{q}(G)$ be a choice function that satisfies

$$\check{q}(G) \in G \quad \text{for } G \subseteq \check{\Phi}, G \neq \emptyset,$$
$$\check{q}(\emptyset) = \dot{\mathsf{T}}.$$

For all $v \in \acute{M} \to L$, $x \in \check{\Phi} \cup \check{Q}$ and $y \in \acute{M}$ we define $x \mathrel{\acute{\preceq}}^{\Phi}_{v} y$ so that $x \mathrel{\acute{\preceq}}^{\Phi}_{v} y$ "approximates" $x \mathrel{\acute{\preceq}} y$. According to the isomorphism theorem stated later, the two relations coincide when $v = \acute{r}$. The definition is:

$$x \mathrel{\acute{\preceq}}^{\Phi}_{v} y \Leftrightarrow \forall z \in x^{\mathrm{d}}: x(z) = v(\acute{m}(y, z)).$$

Further, for technical reasons, we introduce $\doteq^{Q}_{v,w}$ which in conjunction with $x \doteq^{\Phi}_{v} y$ gives a "slightly more conservative" approximation to $x \acute{\leqslant} y$ than $x \doteq^{\Phi}_{v} y$ alone:

$$x \doteq^{Q}_{v,w} y \Leftrightarrow \forall z \in w^{\mathrm{d}}\colon \check{a}(x)(z) = v(\acute{m}(y, z)).$$

We are going to define $\acute{a}(x)(y)$ as $\acute{r}(\acute{m}(x, y))$ where $\acute{r}$ is going to be defined as the "minimal fixed point" for a functional $\acute{r}' \in (\acute{M} \to L) \to (\acute{M} \to L)$. For all $x, y, z \in \acute{M}$, $u \in \acute{M}^*$, $v \in \acute{M} \to L$, $w \in \check{\Phi}$ and $w' \in \check{Q}$ define:

$$\acute{r}'(v)(\dot{S}) = \tilde{\lambda}, \tag{25}$$

$$\acute{r}'(v)(\dot{(}\dot{S}\,x\dot{)}) = \tilde{\lambda}, \tag{26}$$

$$\acute{r}'(v)(\dot{(}\dot{S}\,x\,y\dot{)}) = \tilde{\lambda}, \tag{27}$$

$$\acute{r}'(v)(\acute{m}(\dot{(}\dot{S}\,x\,y\,z\dot{)}, u)) = v(\acute{m}(\dot{(}x\,z\,\dot{(}y\,z\dot{)}\dot{)}, u)), \tag{28}$$

$$\acute{r}'(v)(\dot{K}) = \tilde{\lambda}, \tag{29}$$

$$\acute{r}'(v)(\dot{(}\dot{K}\,x\dot{)}) = \tilde{\lambda}, \tag{30}$$

$$\acute{r}'(v)(\acute{m}(\dot{(}\dot{K}\,x\,y\dot{)}, u)) = v(\acute{m}(x, u)), \tag{31}$$

$$\acute{r}'(v)(\acute{m}(\dot{\mathsf{T}}, u)) = \tilde{\mathsf{T}}, \tag{32}$$

$$\acute{r}'(v)(\dot{P}) = \tilde{\lambda}, \tag{33}$$

$$\acute{r}'(v)(\dot{(}\dot{P}\,x\dot{)}) = \tilde{\lambda}, \tag{34}$$

$$\acute{r}'(v)(\dot{(}\dot{P}\,x\,y\dot{)}) = \tilde{\lambda}, \tag{35}$$

$$\acute{r}'(v)(\acute{m}(\dot{(}\dot{P}\,x\,y\,z\dot{)}, u)) = \begin{cases} \tilde{\bot} & \text{if } v(z) = \tilde{\bot}, \\ v(\acute{m}(x, u)) & \text{if } v(z) = \tilde{\mathsf{T}}, \\ v(\acute{m}(y, u)) & \text{if } v(z) = \tilde{\lambda}, \end{cases} \tag{36}$$

$$\acute{r}'(v)(\dot{C}) = \tilde{\lambda}, \tag{37}$$

$$\begin{aligned} &\acute{r}'(v)(\acute{m}(\dot{(}\dot{C}\,x\dot{)}, u)) \\ &\quad = \begin{cases} \tilde{\bot} & \text{if } \exists y \in \check{\Phi}\colon v(\dot{(}x\,y\dot{)}) = \tilde{\bot}, \\ v(\acute{m}(\check{q}(\{y \in \check{\Phi} \mid v(\dot{(}x\,y\dot{)}) = \tilde{\mathsf{T}}\}),)) & \text{otherwise,} \end{cases} \end{aligned} \tag{38}$$

$$\acute{r}'(v)(\dot{W}) = \tilde{\lambda}, \tag{39}$$

$$\acute{r}'(v)(\acute{m}(\dot{(}\dot{W}\,x\dot{)}, u)) = \begin{cases} \tilde{\mathsf{T}} & \text{if } \exists y \in \check{\Phi}\colon y \doteq^{\Phi}_{v} x, \\ \tilde{\bot} & \text{otherwise,} \end{cases} \tag{40}$$

$$\acute{r}'(v)(\acute{m}(\dot{\bot}, u)) = \tilde{\bot}, \tag{41}$$

$$\acute{r}'(v)(\acute{m}(w, u)) = w\langle\!\langle i \in u^{\mathrm{d}} \mapsto j \in w^{\mathrm{dRD}} \mapsto v(\acute{m}(u(i), j))\rangle\!\rangle, \tag{42}$$

$$\begin{aligned} &\acute{r}'(v)(\acute{m}(w', u)) \\ &\quad = [u \in \acute{M}^* \mapsto \bigsqcup\{w'(u') \mid u' \in w'^{\mathrm{d}} \wedge u' \doteq^{\Phi *}_{v} u \wedge u' \doteq^{Q *}_{v,w} u\}]\langle\!\langle u\rangle\!\rangle. \end{aligned} \tag{43}$$

By transfinite recursion in α define $\acute{r}''(\alpha)$ as α iterations of $\acute{r}'$:

$$\acute{r}''(0) = x \in \acute{M} \mapsto \tilde{\bot},$$

$$\acute{r}''(\alpha^{+}) = \acute{r}'(\acute{r}''(\alpha)),$$

$$\acute{r}''(\mu) = x \in \acute{M} \mapsto \bigsqcup\{\acute{r}''(\alpha)(x) \mid \alpha \in \mu\}.$$

Now let ξ be the least ordinal with cardinality greater than $\acute{M} \mapsto L$. We define

$$\acute{r} = \acute{r}''(\xi).$$

From $\acute{m}$ and $\acute{r}$ we define $\acute{a} \in \acute{M} \to (M^{*} \to L)$ as follows:

$$\acute{a}(x)(y) = \acute{r}(\acute{m}(x, y)).$$

For all $x, y \in \acute{M}$ we define the equivalence relation $x \mathrel{\acute{=}} y$ by $x \mathrel{\acute{=}} y \Leftrightarrow \acute{a}(x) = \acute{a}(y)$. This relation expresses "observational equivalence". In terms of Part I we could say that $x \mathrel{\acute{=}} y$ expresses that x and y have the same graphical representation.

Further, we define $x \mathrel{\acute{\leqslant}} y \Leftrightarrow \acute{a}(x) \leqslant_{L}^{*} \acute{a}(y)$ and $\acute{\Phi} = \{x \in \acute{M} \mid \exists y \in \check{\Phi}\colon y \mathrel{\acute{\leqslant}} x\}$.

The intuition behind (43) is as follows: If $v \in \acute{M} \to L$, $w \in \check{Q}'(\alpha)$, $x, y \in \check{\Phi}$ and $z \in \acute{M}$, then

$$x \mathrel{\acute{=}}^{Q}_{v,w} z \wedge y \mathrel{\acute{=}}^{Q}_{v,w} z \;\Rightarrow\; x \mathrel{\breve{=}}^{a}_{\check{\Phi}''(\alpha)} y.$$

Hence, if $v \in \acute{M} \to L$, $w \in \check{Q}$, $x, y \in w^{\mathrm{d}}$, and $z \in \acute{M}$, then

$$x \mathrel{\acute{=}}^{Q}_{v,w} z \wedge y \mathrel{\acute{=}}^{Q}_{v,w} z \;\Rightarrow\; w(x) = w(y), \tag{44}$$

so the set $\{w'(u') \mid u' \in w'^{\mathrm{d}} \wedge u' \mathrel{\acute{=}}^{\Phi *}_{v} u \wedge u' \mathrel{\acute{=}}^{Q*}_{v,w} u\}$ has at most one element. Hence, if $v \in \acute{M} \to L$, $w' \in \check{Q}$, $u \in \acute{M}^{*}$, $u' \in w'^{\mathrm{d}}$, $u' \mathrel{\acute{=}}^{\Phi *}_{v} u$ and $u' \mathrel{\acute{=}}^{Q*}_{v,w} u$, then (43) reduces to

$$\acute{r}'(v)(\acute{m}(w', u)) = w'(u').$$

11.5. The fixed point theorem

If $v, v' \in \acute{M} \to L$, $v \leqslant_{L}^{*} v'$, $w \in \check{Q}$, $x \in \check{\Phi}$ and $z \in \acute{M}$, then

$$x \mathrel{\acute{=}}^{\Phi}_{v} z \;\Rightarrow\; x \mathrel{\acute{=}}^{\Phi}_{v'} z, \tag{45}$$

$$x \mathrel{\acute{=}}^{Q}_{v,w} z \;\Rightarrow\; x \mathrel{\acute{=}}^{Q}_{v',w} z. \tag{46}$$

With (44)-(46), it is easy to verify

$$\forall v, v' \in \acute{M} \to L\colon (v \leqslant_{L}^{*} v' \Rightarrow \acute{r}'(v) \leqslant_{L}^{*} \acute{r}'(v'))$$

by a proof by cases using equations (25)-(43) (i.e. by a trivial proof by 19 cases!). It then becomes easy to verify

$$\forall \alpha, \beta\colon (\alpha \leqslant_{\mathrm{o}} \beta \Rightarrow \acute{r}''(\alpha) \leqslant_{L}^{*} \acute{r}''(\beta)).$$

Since $\acute{r}''(\alpha) \in \acute{M} \to L$ for all α and $\acute{M} \to L <_{\kappa} \xi$, there exist $\alpha, \beta \in \xi$, $\alpha \neq \beta$ such that $\acute{r}''(\alpha) = \acute{r}''(\beta)$. It is now straightforward to verify $\acute{r}''(\gamma) = \acute{r}''(\delta)$ for all $\gamma, \delta \geqslant_{\mathrm{o}} \alpha$. In particular, $\acute{r} = \acute{r}''(\xi) = \acute{r}''(\xi^{+}) = \acute{r}'(\acute{r}''(\xi)) = \acute{r}'(\acute{r})$ which proves the fixed point theorem:

Theorem 11.5.1 (Fixed point theorem). $\acute{r}'(\acute{r}) = \acute{r}$.

11.6. Extensionality and monotonicity

Theorem 11.6.1 (Monotonicity theorem). *If* $u, v, x, y \in \acute{M}$, $u \mathrel{\acute{\leq}} v$ *and* $x \mathrel{\acute{\leq}} y$, *then* $\dot{(}u\,x\dot{)} \mathrel{\acute{\leq}} \dot{(}v\,y\dot{)}$.

Theorem 11.6.2 (Extensionality theorem). *If* $u, v, x, y \in \acute{M}$, $u \doteq v$ *and* $x \doteq y$, *then* $\dot{(}u\,x\dot{)} \doteq \dot{(}v\,y\dot{)}$.

From the definitions of $\doteq$ and $\acute{\leq}$ we have

$$x \doteq y \;\Leftrightarrow\; x \mathrel{\acute{\leq}} y \wedge y \mathrel{\acute{\leq}} x.$$

Hence, the extensionality theorem follows from the monotonicity theorem. Further, from the definition of $\acute{\leq}$ we have

$$u \mathrel{\acute{\leq}} v \;\Rightarrow\; \dot{(}u\,x\dot{)} \mathrel{\acute{\leq}} \dot{(}v\,x\dot{)}.$$

Hence, to prove the monotonicity theorem, it is sufficient to prove

$$x \mathrel{\acute{\leq}} y \;\Rightarrow\; \dot{(}u\,x\dot{)} \mathrel{\acute{\leq}} \dot{(}u\,y\dot{)}. \tag{47}$$

In order to prove this we introduce a set $\acute{E} \subseteq \acute{M} \to \acute{M}$ of "polynomials". For each polynomial $f \in \acute{E}$ we then prove

$$y \mathrel{\acute{\leq}} z \;\Rightarrow\; f(y) \mathrel{\acute{\leq}} f(z). \tag{48}$$

Further, we make sure that $z \in \acute{M} \mapsto \dot{(}x\,z\dot{)}$ is a polynomial for each fixed $x \in \acute{M}$ so that (47) becomes a special case of (48).

We define the polynomials to be those functions that can be written

$$z \in \acute{M} \mapsto \Phi$$

where Φ is an expression built up from z and elements of $\acute{M}$ by repeated application of $\dot{A}$. More formally, we let $\acute{E} \subseteq \acute{M} \to \acute{M}$ be the least set such that

$$(z \in \acute{M} \mapsto z) \in \acute{E}$$

$$\forall x \in \acute{M} \colon (z \in \acute{M} \mapsto x) \in \acute{E}$$

$$\forall x, y \in \acute{E} \colon (z \in \acute{M} \mapsto \dot{A}(x(z), y(z))) \in \acute{E}.$$

Any polynomial $g = x \in \acute{M} \mapsto \Phi$ can be written on the form $g = z \in \acute{M} \mapsto \dot{(}\Phi_0\,\Phi_1 \ldots \Phi_\alpha\dot{)}$ where $\alpha \geq 0$ and where Φ_0 is either an element of $\{\dot{S}, \dot{K}, \dot{\mathsf{T}}, \dot{P}, \dot{C}, \dot{W}, \dot{\bot}\} \cup \breve{\Phi} \cup \breve{Q}$ or the variable "z". Hence, for all $g \in \acute{E}$ there exist $x \in \{\dot{S}, \dot{K}, \dot{\mathsf{T}}, \dot{P}, C, \dot{W}, \dot{\bot}\} \cup \breve{\Phi} \cup \breve{Q}$ and $h_1, \ldots, h_\alpha \in \acute{E}$ such that

$$g(z) = \dot{(}x\,h_1(z) \ldots h_\alpha(z)\dot{)} \tag{49}$$

or

$$g(z) = \dot{(}z\,h_1(z) \ldots h_\alpha(z)\dot{)}. \tag{50}$$

To prove (48) we first prove an auxiliary lemma.

Lemma 11.6.3.

$$\forall v \in \acute{M} \to L\, \forall y, z \in \acute{M}: (y \acute{\leqslant} z \wedge \forall f \in \acute{E}: v(f(y)) \leqslant_L \acute{r}(f(z)) \Rightarrow \forall g \in \acute{E}: \acute{r}'(v)(g(y)) \leqslant_L \acute{r}(g(z))).$$

Proof. Assume $v \in \acute{M} \to L$, $y, z \in \acute{M}$, $y \acute{\leqslant} z$, $\forall f \in \acute{E}: v(f(y)) \leqslant_L \acute{r}(f(z))$ and $g \in \acute{E}$. We split the proof of $\acute{r}'(v)(g(y)) \leqslant_L \acute{r}(g(z))$ in two cases.

Case 1. Assume that (49) holds. From the definition of $\acute{r}'$ and $\forall f \in \acute{E}$: $v(f(y)) \leqslant_L \acute{r}(f(z))$ we deduce

$$\acute{r}'(v)(\dot{(}x\, h_1(y) \ldots h_\alpha(y)\dot{)}) \leqslant_L \acute{r}'(\acute{r})(\dot{(}x\, h_1(z) \ldots h_\alpha(z)\dot{)}).$$

The proof is similar to the proof of $\forall v, v' \in \acute{M} \to L$: $(v \leqslant_L^* v' \Rightarrow \acute{r}'(v) \leqslant_L^* \acute{r}'(v'))$. From $\acute{r}'(\acute{r}) = \acute{r}$ and $g(u) = \dot{(}x\, h_1(u) \ldots h_\alpha(u)\dot{)}$ we deduce $\acute{r}'(v)(g(y)) \leqslant_L \acute{r}(g(z))$.

Case 2. Assume that (50) holds. From $\forall f \in \acute{E}: v(f(y)) \leqslant_L \acute{r}(f(z))$ we deduce

$$\acute{r}'(v)(g(y)) = \acute{r}'(v)(\dot{(}y\, h_1(y) \ldots h_\alpha(y)\dot{)}) \leqslant_L \acute{r}(\dot{(}y\, h_1(z) \ldots h_\alpha(z)\dot{)})$$

as in Case 1. From $y \acute{\leqslant} z$ we obtain

$$\acute{r}(\dot{(}y\, h_1(z) \ldots h_\alpha(z)\dot{)}) \leqslant_L \acute{r}(\dot{(}z\, h_1(z) \ldots h_\alpha(z)\dot{)}) = \acute{r}(g(z)).$$

These together entail $\acute{r}'(v)(g(y)) \leqslant_L \acute{r}(g(z))$. $\square$

For $y, z \in \acute{M}$, $y \acute{\leqslant} z$ we can now prove $\forall f \in \acute{E}: \acute{r}''(\alpha)(f(y)) \leqslant_L \acute{r}(f(z))$ by transfinite induction in α. The $\forall f \in \acute{E}: \acute{r}(f(y)) \leqslant_L \acute{r}(f(z))$ follows as the special case $\alpha = \xi$, which proves (48) from which the theorems follow.

11.7. The root theorem

Theorem 11.7.1 (Root theorem). *For all $x \subseteq \acute{M}$ we have*

$$x \doteq \dot{\top} \Leftrightarrow \acute{r}(x) = \tilde{\top},$$

$$x \doteq \dot{\bot} \Leftrightarrow \acute{r}(x) = \tilde{\bot}.$$

Proof. Define

$$\acute{M}' = \{v \subseteq \acute{M} \to L \mid \forall x \in \acute{M}\ \forall y, z \in \acute{M}^*: (v(\acute{m}(x, y)) \neq \tilde{\lambda} \Rightarrow v(\acute{m}(x, y \cdot z)) = v(\acute{m}(x, y)))\}.$$

Let $v \in \acute{M}'$, $x \in \acute{M}$ and $y \in \acute{M}^*$. One easily verifies $\acute{r}'(v)(x) \neq \tilde{\lambda} \Rightarrow \acute{r}'(v)(\acute{m}(x, y)) = \acute{r}'(v)(x)$ by a proof by cases in x using (25)–(43). Hence, $\acute{r}'(v) \in \acute{M}'$. We now have $\acute{r}''(\alpha) \in \acute{M}'$ by transfinite induction in α and $\acute{r} \in \acute{M}'$ as the special case $\alpha = \xi$. The theorem now follows trivially. $\square$

11.8. Isomorphism

Theorem 11.8.1 (Isomorphism). *We have*

$$\forall x \in \breve{\Phi}\ \forall y \in \breve{\Phi}^*\colon \breve{a}(x)(y) = \acute{a}(x)(y), \tag{51}$$

$$\forall x \in \breve{\Phi}\ \forall y \in \acute{M}\colon (x \mathrel{\acute{\doteq}}^{\Phi}_{\acute{r}} y \Leftrightarrow x \mathrel{\acute{\leqslant}} y). \tag{52}$$

Proof of (51). For $x \in \hat{\Phi}$ and $y \in \hat{\Phi}^*$ we have

$$\begin{aligned} \hat{a}(x, y) &= x(\hat{t}^*_{x^{\mathrm{dRDR}}}(y)) \\ &= x\langle\!\langle \hat{t}^*_{x^{\mathrm{dRDR}}}(y)\rangle\!\rangle \\ &= x\langle\!\langle u \in y^{\mathrm{d}} \mapsto v \in x^{\mathrm{dRD}} \mapsto \hat{a}(y(u), v)\rangle\!\rangle. \end{aligned}$$

By relativization we obtain

$$\breve{a}(x)(y) = x\langle\!\langle u \in y^{\mathrm{d}} \mapsto v \in x^{\mathrm{dRD}} \mapsto \breve{a}(y(u))(v)\rangle\!\rangle$$

for $x \in \breve{\Phi}$ and $y \in \breve{\Phi}^*$. Using the fixed point theorem, $\acute{a}(x)(y) = \acute{r}(\acute{m}(x, y))$ and (42) we obtain

$$\acute{a}(x)(y) = x\langle\!\langle u \in y^{\mathrm{d}} \mapsto v \in x^{\mathrm{dRD}} \mapsto \acute{a}(y(u))(v)\rangle\!\rangle.$$

Equation (51) of the isomorphism theorem now follows by transfinite induction on $\langle x, y\rangle$ using the well-founded relation $<_{\rho\rho}$ defined in Section 11.1. □

Equation (52) requires a considerably longer proof. We prove it by a series of lemmas.

Lemma 11.8.2. *If*

$$\forall x \in \breve{\Phi}''(\alpha)\ \forall y \in x^{\mathrm{d}}\colon x(y) = \acute{a}(x)(y)$$

then

$$\forall x \in \breve{Q}'(\alpha)\ \forall y \in x^{\mathrm{d}}\colon x(y) = \acute{a}(x)(y).$$

Proof. Assume $\forall x \in \breve{\Phi}''(\alpha)\ \forall y \in x^{\mathrm{d}}\colon x(y) = \acute{a}(x)(y)$, $x \in \breve{Q}'(\alpha)$ and $y \in x^{\mathrm{d}} = \breve{\Phi}''(\alpha)^*$. Inspired by (43) let $G = \{x(y') \mid y' \in \breve{\Phi}''(\alpha)^* \wedge y' \mathrel{\acute{\doteq}}^{\Phi*}_{\acute{r}} y \wedge y' \mathrel{\acute{\doteq}}^{Q*}_{\acute{r},x} y\}$. For $z \in \breve{\Phi}''(\alpha)$ we may prove $z \mathrel{\acute{\doteq}}^{\Phi}_{\acute{r}} z$ by $z \mathrel{\acute{\doteq}}^{\Phi}_{\acute{r}} z \Leftrightarrow \forall u \in z^{\mathrm{d}}\colon z(u) = \acute{a}(z)(u) \Leftrightarrow \forall u \in z^{\mathrm{d}}\colon \acute{a}(x)(u) = \acute{a}(z)(u)$. Further, we may prove $z \mathrel{\acute{\doteq}}^{Q}_{\acute{r},x} z$ by $z \mathrel{\acute{\doteq}}^{Q}_{\acute{r},x} z \Leftrightarrow \forall u \in x^{\mathrm{d}}\colon \breve{a}(x)(z) = \acute{r}(\acute{m}(x, z)) \Leftrightarrow \forall u \in x^{\mathrm{d}}\colon \acute{a}(x)(z) = \acute{a}(x)(z)$. Hence, $x(y) \in G$. Since G has at most one element according to (44), we have $G = \{x(y)\}$ and $\bigsqcup G = x(y)$, so $\acute{a}(x)(y) = x(y)$ follows from (43). □

Lemma 11.8.3. *If*

$$\forall x \in \breve{Q}'(\alpha)\ \forall y \in x^{\mathrm{d}}\colon x(y) = \acute{a}(x)(y)$$

then

$$\forall x \in \breve{\Phi}'(\alpha)\ \forall y \in x^{\mathrm{d}}\colon x(y) = \acute{a}(x)(y).$$

Proof. Assume $\forall x \in \breve{Q}'(\alpha)\ \forall y \in x^{\mathrm{d}}$: $x(y) = \acute{a}(x)(y)$, $x \in \breve{\Phi}'(\alpha)$ and $y \in x^{\mathrm{d}} = \breve{Q}'(\alpha)^*$. Using (42) we have

$$\begin{aligned} \acute{a}(x)(y) &= x\langle\!\langle u \in y^{\mathrm{d}} \mapsto v \in x^{\mathrm{dRD}} \mapsto \acute{a}(y(u))(v)\rangle\!\rangle \\ &= x\langle\!\langle u \in y^{\mathrm{d}} \mapsto v \in x^{\mathrm{dRD}} \mapsto y(u)(v)\rangle\!\rangle \\ &= x\langle\!\langle y\rangle\!\rangle \\ &= x(y). \qquad \square \end{aligned}$$

Lemma 11.8.4. $\forall x \in \breve{\Phi}\ \forall y \in x^{\mathrm{d}}$: $x(y) = \acute{a}(x)(y)$.

Proof. Follows from the preceding two lemmas by transfinite induction. $\square$

Lemma 11.8.5. $\forall x \in \breve{\Phi}\ \forall y \in \acute{M}$: $(x \acute{\leqslant} y \Rightarrow x \doteq_r^{\Phi} y)$.

Proof. Assume $x \in \breve{\Phi}$ and $y \in \acute{M}$. We have

$$\begin{aligned} x \acute{\leqslant} y &\Leftrightarrow \forall z \in \acute{M}^*\colon \acute{a}(x)(z) \leqslant_L \acute{a}(y)(z) \\ &\Rightarrow \forall z \in x^{\mathrm{d}}\colon \acute{a}(x)(z) \leqslant_L \acute{a}(y)(z) \\ &\Leftrightarrow \forall z \in x^{\mathrm{d}}\colon x(z) \leqslant_L \acute{a}(y)(z) \\ &\Leftrightarrow \forall z \in x^{\mathrm{d}}\colon x(z) = \acute{a}(y)(z) \\ &\Leftrightarrow x \doteq_r^{\Phi} y. \qquad \square \end{aligned}$$

This establishes half of (52).

Lemma 11.8.6. *If*

$$\forall x \in \breve{\Phi}''(\alpha)\ \forall y \in \acute{M}\colon (x \doteq_r^{\Phi} y \Rightarrow x \acute{\leqslant} y)$$

then

$$\forall x \in \breve{Q}'(\alpha)\ \forall y \in \acute{M}\colon (x \doteq_r^{\Phi} y \Rightarrow x \acute{\leqslant} y).$$

Proof. Assume $\forall x \in \breve{\Phi}''(\alpha)\ \forall y \in \acute{M}$: $(x \doteq_r^{\Phi} y \Rightarrow x \acute{\leqslant} y)$, $x \in \breve{Q}'(\alpha)$, $y \in \acute{M}$, $x \doteq_r^{\Phi} y$ and $z \in \acute{M}^*$. We shall prove $\acute{a}(x)(z) \leqslant_L \acute{a}(y)(z)$. We divide the proof in three cases: $\acute{a}(x)(z) = \tilde{\bot}$, $\acute{a}(x)(z) = \tilde{\lambda}$ and $\acute{a}(x)(z) = \tilde{\mathsf{T}}$.

If $\acute{a}(x)(z) = \tilde{\bot}$ then $\acute{a}(x)(z) \leqslant_L \acute{a}(y)(z)$ is trivial.

If $\acute{a}(x)(z) = \tilde{\lambda}$ then let $z' \in x^{\mathrm{d}} = \breve{\Phi}''(\alpha)^*$ satisfy $z' \doteq_r^{\Phi *} z \wedge z' \doteq_{r,x}^{Q*} z$ (if no such z' exists, then $\acute{a}(x)(z) = \tilde{\bot}$ by (43) and the fixed point theorem). From $z' \doteq_r^{\Phi *} z$ we obtain $z' \acute{\leqslant}^* z$ so, by the monotonicity theorem and $x \doteq_r^{\Phi} y$, $\acute{a}(x)(z) = x(z') = \acute{a}(y)(z') \leqslant_L \acute{a}(y)(z)$.

If $\acute{a}(x)(z) = \tilde{\mathsf{T}}$ then let α be the least ordinal such that $\acute{a}(x)(z|\alpha) = \tilde{\mathsf{T}}$ and let $z' \in x^{\mathrm{d}}$ satisfy $z' \doteq_r^{\Phi *} (z|\alpha) \wedge z' \doteq_{r,x}^{Q*} (z|\alpha)$. From $z' \doteq_r^{\Phi *} (z|\alpha)$ we obtain $z' \acute{\leqslant} (z|\alpha)$. Hence, $\tilde{\mathsf{T}} = \acute{a}(x)(z|\alpha) = x(z') = \acute{a}(y)(z') \leqslant_L \acute{a}(y)(z|\alpha)$, so $\acute{a}(y)(z) = \tilde{\mathsf{T}}$ and $\acute{a}(x)(z) \leqslant_L \acute{a}(y)(z)$. $\square$

Lemma 11.8.7. *If*

$$\forall x \in \check{Q}'(\alpha)\, \forall y \in \acute{M}: (x \mathrel{\dot{=}}^{\Phi}_{r} y \Rightarrow x \mathrel{\acute{\leqslant}} y)$$

then

$$\forall x \in \check{\Phi}'(\alpha)\, \forall y \in \acute{M}: (x \mathrel{\dot{=}}^{\Phi}_{r} y \Rightarrow x \mathrel{\acute{\leqslant}} y).$$

Proof. Assume $\forall x \in \check{Q}'(\alpha)\, \forall y \in \acute{M}: (x \mathrel{\dot{=}}^{\Phi}_{r} y \Rightarrow x \mathrel{\acute{\leqslant}} y)$, $x \in \check{\Phi}'(\alpha)$, $y \in \acute{M}$, $x \mathrel{\dot{=}}^{\Phi}_{r} y$ and $z \in \acute{M}^*$. We shall prove $\acute{a}(x)(z) \leqslant_L \acute{a}(y)(z)$. We divide the proof in three cases: $\acute{a}(x)(z) = \tilde{\bot}$, $\acute{a}(x)(z) = \tilde{\lambda}$ and $\acute{a}(x)(z) = \tilde{\mathsf{T}}$.

If $\acute{a}(x)(z) = \tilde{\bot}$ then $\acute{a}(x)(z) \leqslant_L \acute{a}(y)(z)$ is trivial.

If $\acute{a}(x)(z) = \tilde{\lambda}$ then let $z' = u \in z^{\mathrm{d}} \mapsto v \in x^{\mathrm{dRD}} \mapsto \acute{a}(z(u))(v)$. We have $\acute{a}(x)(z) = x(z')$ and $z' \in x^{\mathrm{d}} = \check{Q}'(\alpha)^*$. From $z' \mathrel{\dot{=}}^{\Phi *}_{r} z$ we obtain $z' \mathrel{\acute{\leqslant}}^* z$. Hence, $\acute{a}(x)(z) = x(z') = \acute{a}(y)(z') \leqslant_L \acute{a}(y)(z)$.

If $\acute{a}(x)(z) = \tilde{\mathsf{T}}$ then let α be the least ordinal such that $\acute{a}(x)(z|\alpha) = \tilde{\mathsf{T}}$. Let $z' = u \in (z|\alpha)^{\mathrm{d}} \mapsto v \in x^{\mathrm{dRD}} \mapsto \acute{a}(z(u))(v)$. We have $\acute{a}(x)(z|\alpha) = x(z')$ and $z' \in x^{\mathrm{d}} = \check{Q}'(\alpha)^*$. From $z' \mathrel{\dot{=}}^{\Phi}_{r} (z|\alpha)$ we obtain $z' \mathrel{\acute{\leqslant}}^* (z|\alpha)$. Hence $\tilde{\mathsf{T}} = \acute{a}(x)(z|\alpha) = x(z') = \acute{a}(y)(z') \leqslant_L \acute{a}(y)(z|\alpha)$, so $\acute{a}(y)(z) = \tilde{\mathsf{T}}$ and $\acute{a}(x)(z) \leqslant_L \acute{a}(y)(z)$. □

Lemma 11.8.8. *If $x \in \check{\Phi}$, $y \in \acute{M}$ and $x \mathrel{\dot{=}}^{\Phi}_{r} y$, then $x \mathrel{\acute{\leqslant}} y$.*

Proof. By transfinite induction using the preceding two lemmas. □

Lemmas 11.8.5 and 11.8.8 together establish (52).

11.9. Definition of the model

Define the equivalence class $c(x)$ of $x \in \acute{M}$ by $c(x) = \{y \in \acute{M} \mid y \mathrel{\dot{=}} x\}$. We now define the quotient $\acute{M}/\dot{=}$:

$$M = \{c(u) \mid u \in \acute{M}\}, \tag{53}$$

$$\Phi = \{c(u) \mid u \in \acute{\Phi}\}, \tag{54}$$

$$A(x, y) = \bigcup \{c(\dot{A}(u, v)) \mid u \in x \wedge v \in y\}, \tag{55}$$

$$S = c(\dot{S}), \tag{56}$$

$$K = c(\dot{K}), \tag{57}$$

$$\mathsf{T} = c(\dot{\mathsf{T}}), \tag{58}$$

$$P = c(\dot{P}), \tag{59}$$

$$C = c(\dot{C}), \tag{60}$$

$$W = c(\dot{W}), \tag{61}$$

$$\bot = c(\dot{\bot}). \tag{62}$$

From the extensionality theorem we have that this construction really is a quotient construction, i.e. we have

$$A(c(x), c(y)) = c(\dot{A}(x, y)). \tag{63}$$

As an analogy to $(\dot{x\, y_1\, y_2 \ldots y_\alpha})$ we introduce $(x\, y_1\, y_2 \ldots y_\alpha)$ as shorthand for

$$A(\ldots A(A(x, y_1), y_2) \ldots, y_\alpha).$$

11.10. The well-foundedness theorem

Lemma 11.10.1. $\mathsf{T} \in \Phi$ *and* $\perp \notin \Phi$.

Proof. Let $\hat{\mathsf{T}} = x \in \hat{Q}'(0)^* \mapsto \tilde{\mathsf{T}}$. We have $\hat{\mathsf{T}} \in \hat{\Phi}'(0)$, so $\hat{\mathsf{T}} \in \hat{\Phi}$. Let $\check{\mathsf{T}} = \lfloor \hat{\mathsf{T}} \rfloor$. We have $\check{\mathsf{T}} \in \check{\Phi}$ and $\forall x \in \check{\mathsf{T}}^{\mathrm{d}}$: $\check{\mathsf{T}}(x) = \tilde{\mathsf{T}}$. Hence, $\check{\mathsf{T}} \doteq^{\Phi}_{r} \dot{\mathsf{T}}$, so $\check{\mathsf{T}} \acute{\leqslant} \dot{\mathsf{T}}$ which entails $\dot{\mathsf{T}} \in \acute{\Phi}$ and $\mathsf{T} = c(\dot{\mathsf{T}}) \in \Phi$.

Now assume $x \in \check{\Phi}'(\alpha) = \check{Q}'(\alpha)^{\circ}$. From the definition of $\check{Q}'(\alpha)^{\circ}$ we have $x(\langle\,\rangle) \in \{\tilde{\mathsf{T}}, \tilde{\lambda}\}$, so $\forall x \in \hat{\Phi}$: $x(\langle\,\rangle) \neq \tilde{\perp}$ which proves $\forall x \in \check{\Phi}$: $x(\langle\,\rangle) \neq \tilde{\perp}$, $\forall x \in \acute{\Phi}$: $x \acute{\neq}^{\Phi}_{r} \dot{\perp}$, $\forall x \in \acute{\Phi}$: $x \not\acute{\leqslant} \dot{\perp}$, $\dot{\perp} \notin \acute{\Phi}$ and $\perp = c(\dot{\perp}) \notin \Phi$. $\square$

We now verify the property of well-founded maps that gave them their name.

Lemma 11.10.2. $\forall x \in \Phi\ \forall y \in \Phi^{\omega}\ \exists \alpha \in \omega$: $m(x, y) = \mathsf{T}$.

Proof. Let $x, y_1, y_2, \ldots \in \Phi$. For all $\alpha \in \omega$ choose $u, u_\alpha \in \acute{\Phi}$ and $v, v_\alpha \in \check{\Phi}$ such that $x = c(u)$, $y_\alpha = c(v_\alpha)$, $v \acute{\leqslant} u$ and $v_\alpha \acute{\leqslant} u_\alpha$. Using the relativization of Corollary 11.2.2, choose $\alpha \in \omega$ such that $\breve{a}(v)(\langle v_1, \ldots, v_\alpha \rangle) = \tilde{\mathsf{T}}$. From the isomorphism and monotonicity theorems we have $\acute{a}(u)(\langle u_1, \ldots, u_\alpha \rangle) = \tilde{\mathsf{T}}$ which entails $\acute{m}(u, \langle u_1, \ldots, u_\alpha \rangle) \doteq \dot{\mathsf{T}}$ and $m(x, \langle y_1, \ldots, y_\alpha \rangle) = \mathsf{T}$ by the root theorem. $\square$

For all $x, y \in \Phi$ define $x <_A y$ by

$$x <_A y \iff y \neq \mathsf{T} \wedge \exists z \in \Phi: x = (y\, z).$$

From the above lemmas we conclude the well-foundedness theorem.

Theorem 11.10.3 (Well-foundedness theorem). $\mathsf{T} \in \Phi$, $\perp \notin \Phi$ *and* $<_A$ *is well-founded on* Φ.

Corollary 11.10.4 (Induction theorem). *Let* $\mathcal{R}(x)$ *be a predicate. If*

$$\mathcal{R}(\mathsf{T}) \quad \textit{and} \quad \forall x \in \Phi \backslash \{\mathsf{T}\}: (\forall y \in \Phi: \mathcal{R}(x\, y) \Rightarrow \mathcal{R}(x))$$

then

$$\forall x \in \Phi: \mathcal{R}(x).$$

There is a much deeper way to well-order the well-founded maps: For all $x \in \Phi$ let $f(x)$ be the least ordinal α such that $\exists z \in \breve{\Phi}'(\alpha)\, \exists y \in \acute{\Phi}\colon z \mathrel{\acute{\leqslant}} y \wedge x = c(y)$ and define $x <_i y \Leftrightarrow f(x) <_o f(y)$. Even though it is not explicitly mentioned in the proof, the well-foundedness of $<_i$ on Φ is central in proving the consistency of *Map*. In the proofs, induction in $<_i$ has been replaced by induction in the ordinals.

The $<_i$ relation corresponds to the "introduced before" relation in Section 2.4. None of the axioms of map theory expresses the well-foundedness of $<_i$, which is clearly unsatisfactory.

The well-foundedness of $<_A$ corresponds to the well-foundedness of $\in$ in *ZFC* as expressed by the axiom of foundation. The well-foundedness of $<_i$ also relates to the well-foundedness of $\in$, but in a less clear way. It seems that $<_A$ and $<_i$ factor out two distinct sides of the well-foundedness of $\in$. Non-wellfounded sets [2] are well-founded w.r.t. $<_i$ but non-wellfounded w.r.t. $<_A$.

12. Terms and their values

12.1. Representation of terms

Define the set $\dot{V}$ of syntactic variables by

$$\dot{V} = \{\dot{v}_i \mid i \in \omega\}.$$

In what follows, $\dot{x}$, $\dot{y}$, $\dot{z}$, $\dot{u}$, $\dot{v}$, $\dot{w}$, $\dot{f}$, $\dot{g}$, $\dot{h}$, etc. stand for arbitrary, distinct variables, i.e., $\dot{x} = \dot{v}_i$, $\dot{y} = \dot{v}_j$, $\dot{z} = \dot{v}_k$, etc. where $i \neq j$, $i \neq k$, $j \neq k$, etc.

Let the set $\dot{M}'$ of combinator terms be the least set such that

$$\{\dot{S}, \dot{K}, \dot{\mathsf{T}}, \dot{P}, \dot{C}, \dot{W}, \dot{\bot}\} \subseteq \dot{M}',$$

$$\forall x, y \in \dot{M}'\colon \dot{A}(x, y) \in \dot{M}',$$

$$\dot{V} \subseteq \dot{M}' \quad \text{and}$$

$$M \subseteq \dot{M}'.$$

Let the set $\dot{M}$ of terms of map theory be the least set such that

$$\dot{M}' \subseteq \dot{M},$$

$$\forall \dot{x} \in \dot{V}\ \forall f \in \dot{M}\colon \dot{\lambda}\dot{x}.f \in \dot{M},$$

$$\forall x, y \in \dot{M}\colon \dot{A}(x, y) \in \dot{M},$$

$$\forall x, y, z \in \dot{M}\colon (\dot{\text{if}}\ x\ y\ z) \in \dot{M},$$

$$\forall x \in \dot{M}\colon \dot{\varepsilon} x \in \dot{M} \quad \text{and}$$

$$\forall x \in \dot{M}\colon \dot{\phi} x \in \dot{M}.$$

12.2. Structural induction

The principle of structural induction for $\grave{M}'$ can be stated as follows.

Lemma 12.2.1. *Let $\mathcal{R}(x)$ be a predicate. If*

$$\forall x \in \grave{V}: \mathcal{R}(x),$$

$$\forall x \in \{\dot{S}, \dot{K}, \dot{\mathsf{T}}, \dot{P}, \dot{C}, \dot{W}, \dot{\bot}\}: \mathcal{R}(x),$$

$$\forall x, y \in \grave{M}': (\mathcal{R}(x) \wedge \mathcal{R}(y) \Rightarrow \mathcal{R}(\dot{A}(x, y))),$$

$$\forall x \in M: \mathcal{R}(x),$$

then $\forall x \in \grave{M}': \mathcal{R}(x)$.

The principle of structural induction for $\grave{M}$ is given in the next lemma.

Lemma 12.2.2. *Let $\mathcal{R}(x)$ be a predicate. If*

$$\forall x \in \grave{V}: \mathcal{R}(x),$$

$$\forall \{\dot{S}, \dot{K}, \dot{\mathsf{T}}, \dot{P}, \dot{C}, \dot{W}, \dot{\bot}\}: \mathcal{R}(x),$$

$$\forall \dot{x} \in \grave{V}\ \forall f \in \grave{M}: (\mathcal{R}(f) \Rightarrow \mathcal{R}(\dot{\lambda}\dot{x}.f)),$$

$$\forall x, y \in \grave{M}: (\mathcal{R}(x) \wedge \mathcal{R}(y) \Rightarrow \mathcal{R}(\dot{A}(x, y))),$$

$$\forall x, y, z \in \grave{M}: (\mathcal{R}(x) \wedge \mathcal{R}(y) \wedge \mathcal{R}(z) \Rightarrow \mathcal{R}((\dot{\mathsf{if}}\ x\ y\ z))),$$

$$\forall x \in \grave{M}: (\mathcal{R}(x) \Rightarrow \mathcal{R}(\dot{\varepsilon} x)),$$

$$\forall x \in \grave{M}: (\mathcal{R}(x) \Rightarrow \mathcal{R}(\dot{\phi} x)),$$

then $\forall x \in \grave{M}: \mathcal{R}(x)$.

Proofs by structural induction tend to be long and trivial in that they tend to consist of a long list of trivial cases. For this reason, we shall omit the details of most proofs that use structural induction.

12.3. Freeness and substitution

For all $\dot{u} \in \grave{V}$ and $x \in \grave{M}$ we define the predicate $free(\dot{u}, x)$ to stand for "u occurs free in x". The definition is standard: For all $\dot{u}, \dot{v} \in \grave{V}$, $\dot{u} \neq \dot{v}$ and $x, y, z \in \grave{M}$ define

$$free(\dot{u}, \dot{u}),$$

$$\neg free(\dot{u}, \dot{v}),$$

$$\neg free(\dot{u}, x) \quad \text{if } x \in \{\dot{S}, \dot{K}, \dot{\mathsf{T}}, \dot{P}, \dot{C}, \dot{W}, \dot{\bot}\},$$

$$\neg free(\dot{u}, \dot{\lambda}\dot{u}.x),$$

$$free(\dot{u}, \dot{\lambda}\dot{v}.x) \Leftrightarrow free(\dot{u}, x),$$

$$free(\dot{u}, \dot{A}(x, y)) \Leftrightarrow free(\dot{u}, x) \vee free(\dot{u}, y),$$

$$free(\dot{u}, (\dot{\mathsf{if}}\; x\; y\; z)) \Leftrightarrow free(\dot{u}, x) \vee free(\dot{u}, y) \vee free(\dot{u}, z),$$

$$free(\dot{u}, \dot{\varepsilon} x) \Leftrightarrow free(\dot{u}, x),$$

$$free(\dot{u}, \dot{\phi} x) \Leftrightarrow free(\dot{u}, x),$$

$$\neg free(\dot{u}, x) \quad \text{if } x \in M.$$

For all $\dot{u} \in \dot{V}$ and $x, y \in \dot{M}$ we define the predicate $freefor(x, \dot{u}, y)$ to stand for "y is free for $\dot{u}$ in x" [22]. The definition is standard: For all $\dot{u}, \dot{v} \in \dot{V}$, $\dot{u} \neq \dot{v}$ and $x, y, z, w \in \dot{M}$ define

$$freefor(\dot{u}, \dot{u}, w),$$

$$freefor(\dot{v}, \dot{u}, w),$$

$$freefor(x, \dot{u}, w) \quad \text{if } x \in \{\dot{S}, \dot{K}, \dot{\mathsf{T}}, \dot{P}, \dot{C}, \dot{W}, \dot{\perp}\},$$

$$freefor(\dot{\lambda} \dot{u}.x, \dot{u}, w),$$

$$freefor(\dot{\lambda} \dot{v}.x, \dot{u}, w) \Leftrightarrow \neg free(\dot{v}, w) \vee \neg free(\dot{u}, x),$$

$$freefor(\dot{A}(x, y), \dot{u}, w) \Leftrightarrow freefor(x, \dot{u}, w) \wedge freefor(y, \dot{u}, w),$$

$$freefor((\dot{\mathsf{if}}\; x\; y\; z), \dot{u}, w)$$

$$\Leftrightarrow freefor(x, \dot{u}, w) \wedge freefor(y, \dot{u}, w) \wedge freefor(z, \dot{u}, w),$$

$$freefor(\dot{\varepsilon} x, \dot{u}, w) \Leftrightarrow freefor(x, \dot{u}, w),$$

$$freefor(\dot{\phi} x, \dot{u}, w) \Leftrightarrow freefor(x, \dot{u}, w),$$

$$freefor(x, \dot{u}, w) \quad \text{if } x \in M.$$

For all $\dot{u} \in \dot{V}$ and $x, y \in M$ we define $[x/\dot{u} := y]$ to be the result of substituting y for all free occurrences of $\dot{u}$ in x. Also this definition is standard: For all $\dot{u}, \dot{v} \in \dot{V}$, $\dot{u} \neq \dot{v}$ and $x, y, z, w \in \dot{M}$ define

$$[\dot{u}/\dot{u} := w] = w,$$

$$[\dot{v}/\dot{u} := w] = \dot{v},$$

$$[x/\dot{u} := w] = x \quad \text{if } x \in \{\dot{S}, \dot{K}, \dot{\mathsf{T}}, \dot{P}, \dot{C}, \dot{W}, \dot{\perp}\},$$

$$[\dot{\lambda} \dot{u}.x/\dot{u} := w] = \dot{\lambda} \dot{u}.x,$$

$$[\dot{\lambda} \dot{v}.x/\dot{u} := w] = \dot{\lambda} \dot{v}.[x/\dot{u} := w],$$

$$[\dot{A}(x, y)/\dot{u} := w] = \dot{A}([x/\dot{u} := w], [y/\dot{u} := w]),$$

$$[(\dot{\mathsf{if}}\; x\; y\; z)/\dot{u} := w] = (\dot{\mathsf{if}}\; [x/\dot{u} := w]\; [y/\dot{u} := w]\; [z/\dot{u} := w]),$$

$$[\dot{\varepsilon} x/\dot{u} := w] = \dot{\varepsilon}[x/\dot{u} := w],$$

$$[\dot{\phi} x/\dot{u} := w] = \dot{\phi}[x/\dot{u} := w],$$

$$[x/\dot{u} := w] = x \quad \text{if } x \in M.$$

12.4. Translation of terms

For all $\dot{u} \in \dot{V}$ and $x \in \dot{M}'$ we now define $\grave{\lambda} \dot{u}.x$ such that $\grave{\lambda} \dot{u}.x$ becomes the combinator equivalent of $\dot{\lambda} \dot{u}.x$. The definition is standard (cf. [3]). For $\dot{u} \in \dot{V}$ and $x, y \in \dot{M}'$ define

$$
\begin{aligned}
&\grave{\lambda} \dot{u}.\dot{u} = \grave{(}\dot{S}\, \dot{K}\, \dot{K}\grave{)}, \\
&\grave{\lambda} \dot{u}.x = \grave{(}\dot{K}\, x\grave{)} && \text{if } \neg \mathit{free}(\dot{u}, x), \\
&\grave{\lambda} \dot{u}.\grave{(}x\, y\grave{)} = \grave{(}\dot{S}\, \grave{\lambda} \dot{u}.x\, \grave{\lambda} \dot{u}.y\grave{)} && \text{if } \mathit{free}(\dot{u}.\, \grave{(}x\, y\grave{)}).
\end{aligned}
$$

For all terms $x \in \dot{M}$ we now define the translation $[\![x]\!] \in \dot{M}'$. For all terms x, $[\![x]\!]$ is the translation of x into combinator form ([3]). For $\dot{u} \in \dot{V}$ and $x, y, z \in \dot{M}$ define

$$
\begin{aligned}
&[\![\dot{u}]\!] = \dot{u}, \\
&[\![x]\!] = x \quad \text{if } x \in \{\dot{S}, \dot{K}, \dot{\mathsf{T}}, \dot{P}, \dot{C}, \dot{W}, \dot{\perp}\}, \\
&[\![\dot{\lambda} \dot{u}.x]\!] = \grave{\lambda} \dot{u}.[\![x]\!], \\
&[\![\dot{A}(x, y)]\!] = \dot{A}([\![x]\!], [\![y]\!]), \\
&[\![(\dot{\text{if}}\ x\, y\, z)]\!] = \grave{(}\dot{P}\, [\![y]\!]\, [\![z]\!]\, [\![x]\!]\grave{)}, \\
&[\![\dot{\varepsilon} x]\!] = \grave{(}\dot{C}\, [\![x]\!]\grave{)}, \\
&[\![\dot{\phi} x]\!] = \grave{(}\dot{W}\, [\![x]\!]\grave{)}, \\
&[\![x]\!] = x \quad \text{if } x \in M.
\end{aligned}
$$

Note that $\dot{M}' \subseteq \dot{M}$ and $[\![x]\!] = x$ for $x \in \dot{M}'$.

12.5. Interpretation

Define $M^V = \dot{V} \to M$. Elements d of M^V assigns a value $d(\dot{u})$ in M to each variable $\dot{u} \in \dot{V}$, and we refer to elements d of M^V as "assignments". For all combinator terms $x \in \dot{M}'$ and assignments $d \in M^V$ we define the "interpretation" ${}_d x \in M$ to be the value of x when the free variables $\dot{u}$ of x are assigned the values $d(\dot{u})$. More precisely, for all $\dot{u} \in \dot{V}$, $x, y \in \dot{M}'$ and $d \in M^V$ we define

$$
\begin{aligned}
&{}_d \dot{u} = d(\dot{u}), \\
&{}_d x = x \quad \text{if } x \in \{\dot{S}, \dot{K}, \dot{\mathsf{T}}, \dot{P}, \dot{C}, \dot{W}, \dot{\perp}\}, \\
&{}_d \grave{(}x\, y\grave{)} = ({}_d x\, {}_d y), \\
&{}_d x = x \quad \text{if } x \in M.
\end{aligned}
$$

For all terms $x \in \dot{M}$ and assignments $d \in M^V$ we define

$$
{}_d x = {}_d [\![x]\!]
$$

i.e. the interpretation of a term is found by first translating to combinators. Since $\dot{M}' \subseteq \dot{M}$ and $[\![x]\!] = x$ for $x \in \dot{M}'$, the definitions of ${}_d x$ for $x \in \dot{M}'$ and $x \in \dot{M}$ do not conflict. For all combinator terms $x, y \in \dot{M}$ define

$$
x \doteq y \Leftrightarrow \forall d \in M^V \colon {}_d x = {}_d y.
$$

If we let T represent truth, F falsehood and $\bot$ undefinedness, then we may define

$$
\begin{array}{lll}
{}^{\mathrm{D}}_{d}x \Leftrightarrow {}_{d}x \neq \bot & & (x \text{ is defined for assignment } d),\\
{}^{\mathrm{T}}_{d}x \Leftrightarrow {}_{d}x = \mathsf{T} & & (x \text{ is true for assignment } d),\\
{}^{\mathrm{B}}_{d}x \Leftrightarrow {}_{d}x \in \{\mathsf{T}, \mathsf{F}, \bot\} & & (x \text{ is three-valued Boolean for assignment } d).
\end{array}
$$

13. The consistency of *Map°*

13.1. Overview

In this section we prove (9), i.e. we prove that the consistency of *Map°* follows from the consistency of *ZFC*. To do so we prove that the model established above satisfies each axiom and inference rule of *Map°*. For example, to prove that the model satisfies (Ap2) which reads

$$(\lambda u.x\, y) = [x/u := y] \quad \text{if } y \text{ is free for } u \text{ in } x,$$

we prove

$$\mathit{freefor}(x, \acute{u}, y) \Rightarrow \dot{(}\dot{\lambda}\dot{u}.x\, y\dot{)} \doteq [x/\acute{u} := y]$$

for all $\acute{u} \in \grave{V}$ and $x, y \in \grave{M}$.

As mentioned in Section 9.10 we assume at any time that the transitive standard model D satisfies finitely many axioms of *ZFC* without being explicit about which ones. We just assume that D satisfies the axioms necessary for the argument at hand.

13.2. Semantics of basic concepts

If we let $v = \acute{r}$ in (25)-(43) and use the fixed point theorem and $\acute{r}(\acute{m}(x, y)) = \acute{a}(x)(y)$, we obtain

$$\acute{r}(\dot{(}\dot{S}\, x\, y\dot{)}) = \tilde{\lambda},$$

$$\acute{a}(\dot{(}\dot{S}\, x\, y\, z\dot{)})(u) = \acute{a}(\dot{(}x\, z\, \dot{(}y\, z\dot{)}\dot{)})(u),$$

$$\acute{r}(\dot{(}\dot{K}\, x\dot{)}) = \tilde{\lambda},$$

$$\acute{a}(\dot{(}\dot{K}\, x\, y\dot{)})(u) = \acute{a}(x)(u),$$

$$\acute{a}(\dot{\mathsf{T}})(u) = \tilde{\mathsf{T}},$$

$$\acute{a}(\dot{(}\dot{P}\, x\, y\, z\dot{)})(u) = \begin{cases} \tilde{\bot} & \text{if } \acute{r}(z) = \tilde{\bot},\\ \acute{a}(x)(u) & \text{if } \acute{r}(z) = \tilde{\mathsf{T}},\\ \acute{a}(y)(u) & \text{if } \acute{r}(z) = \tilde{\lambda}, \end{cases}$$

$$\acute{a}(\dot{(}\dot{C}\, x\dot{)})(u) = \begin{cases} \tilde{\bot} & \text{if } \exists y \in \check{\Phi}\colon \acute{r}(\dot{(}x\, y\dot{)}) = \tilde{\bot},\\ \acute{a}(\check{q}(\{y \in \check{\Phi} \mid \acute{r}(\dot{(}x\, y\dot{)}) = \tilde{\mathsf{T}}\}))(u) & \text{otherwise}, \end{cases}$$

$$\acute{a}(\dot{(}\dot{W}\, x\dot{)})(u) = \begin{cases} \tilde{\mathsf{T}} & \text{if } \exists y \in \check{\Phi}\colon y \mathrel{\dot{=}}^{\Phi}_{\acute{r}} x,\\ \tilde{\bot} & \text{otherwise}, \end{cases}$$

$$\acute{a}(\dot{\bot})(u) = \tilde{\bot}.$$

Using the root theorem and the definition of $\acute{=}$ we obtain

$$\dot{(}\dot{S}\,x\,y\dot{)} \acute{\neq} \dot{\top},$$

$$\dot{(}\dot{S}\,x\,y\dot{)} \acute{\neq} \dot{\bot},$$

$$\dot{(}\dot{S}\,x\,y\,x\dot{)} \acute{=} \dot{(}x\,z\,\dot{(}y\,z\dot{)}\dot{)},$$

$$\dot{(}\dot{K}\,x\dot{)} \acute{\neq} \dot{\top},$$

$$\dot{(}\dot{K}\,x\dot{)} \acute{\neq} \dot{\bot},$$

$$\dot{(}\dot{K}\,x\,y\dot{)} \acute{=} x,$$

$$\dot{(}\dot{P}\,x\,y\,z\dot{)} \acute{=} \begin{cases} \dot{\bot} & \text{if } z \acute{=} \dot{\bot} \\ x & \text{if } z \acute{=} \dot{\top}, \\ y & \text{otherwise,} \end{cases}$$

$$\dot{(}\dot{C}\,x\dot{)} \acute{=} \begin{cases} \dot{\bot} & \text{if } \exists y \in \check{\Phi}\colon \dot{(}x\,y\dot{)} \acute{=} \dot{\bot}, \\ \check{q}(\{y \in \check{\Phi} \mid \dot{(}x\,y\dot{)} \acute{=} \dot{\top}\}) & \text{otherwise,} \end{cases}$$

$$\dot{(}\dot{W}\,x\dot{)} \acute{=} \begin{cases} \dot{\top} & \text{if } \exists y \in \check{\Phi}\colon y \acute{=}^{\Phi}_{r} x, \\ \dot{\bot} & \text{otherwise.} \end{cases}$$

Using $\acute{a}(\dot{(}x\,y\dot{)})(u) = \acute{a}(x)(\langle y \rangle \cdot u)$ and the root theorem we obtain

$$\dot{(}\dot{\top}\,x\dot{)} \acute{=} \dot{\top} \quad \text{and} \quad \dot{(}\dot{\bot}\,x\dot{)} \acute{=} \dot{\bot}.$$

Using the isomorphism theorem and the definition of $\acute{\Phi}$ we obtain

$$\dot{(}\dot{W}\,x\dot{)} \acute{=} \begin{cases} \dot{\top} & \text{if } x \in \acute{\Phi}, \\ \dot{\bot} & \text{otherwise.} \end{cases}$$

We now investigate the choice combinator $\dot{C}$ and the choice function $\check{q}(\bullet)$. We first prove $\exists y \in \check{\Phi}\colon \dot{(}x\,y\dot{)} \acute{=} \dot{\bot} \Leftrightarrow \exists y \in \acute{\Phi}\colon \dot{(}x\,y\dot{)} \acute{=} \dot{\bot}$. The $\Rightarrow$-direction follows from $\check{\Phi} \subseteq \acute{\Phi}$. Now assume that $y' \in \acute{\Phi}$ satisfies $\dot{(}x\,y'\dot{)} \acute{=} \dot{\bot}$. Choose $y \in \check{\Phi}$ such that $y \acute{\leqslant} y'$ (this is possible due to the definition of $\acute{\Phi}$). From the monotonicity theorem we have $\dot{(}x\,y\dot{)} \acute{\leqslant} \dot{(}x\,y'\dot{)} \acute{=} \dot{\bot}$ which proves $\dot{(}x\,y\dot{)} \acute{=} \dot{\bot}$ and $\exists y \in \check{\Phi}\colon \dot{(}x\,y\dot{)} \acute{=} \dot{\bot}$. From $\exists y \in \check{\Phi}\colon \dot{(}x\,y\dot{)} \acute{=} \dot{\bot} \Leftrightarrow \exists y \in \acute{\Phi}\colon \dot{(}x\,y\dot{)} \acute{=} \dot{\bot}$ we deduce

$$\dot{(}\dot{C}\,x\dot{)} \acute{=} \begin{cases} \dot{\bot} & \text{if } \exists y \in \acute{\Phi}\colon \dot{(}x\,y\dot{)} \acute{=} \dot{\bot}, \\ \check{q}(\{y \in \check{\Phi} \mid \dot{(}x\,y\dot{)} \acute{=} \dot{\top}\}) & \text{otherwise.} \end{cases}$$

Define

$$\acute{q}(G) = \begin{cases} \check{q}(G \cap \check{\Phi}) & \text{if } G \cap \check{\Phi} \neq \emptyset, \\ \check{q}(G) & \text{otherwise.} \end{cases}$$

We have

$$\acute{q}(G) \in G \quad \text{for } G \subseteq \acute{M},\ G \neq \emptyset,$$

$$\acute{q}(\emptyset) = \dot{\top},$$

$$\acute{q}(\{y \in \acute{\Phi} \mid \dot{(}x\,y\dot{)} \acute{=} \dot{\top}\}) = \check{q}(\{y \in \check{\Phi} \mid \dot{(}x\,y\dot{)} \acute{=} \dot{\top}\}).$$

Let $x \in \acute{M}$ and assume $\forall y \in \acute{\Phi}\colon \dot{(}x\,y\dot{)} \not\doteq \dot{\perp}$. Let $G = \{y \in \breve{\Phi} \mid \dot{(}x\,y\dot{)} \doteq \dot{\top}\}$ and $G' = \{y \in \acute{\Phi} \mid \dot{(}x\,y\dot{)} \doteq \dot{\top}\}$. From $\breve{\Phi} \subseteq \acute{\Phi}$ we have $G = G' \cap \breve{\Phi}$. If $G' = \emptyset$ then $\breve{q}(G) = \dot{\top} = \acute{q}(G')$. If $G' \neq \emptyset$ then let $y' \in G' \subseteq \acute{\Phi}$ and choose $y \in \breve{\Phi}$ such that $y \mathrel{\acute{\leqslant}} y'$. From $\dot{\perp} \mathrel{\acute{\neq}} \dot{(}x\,y\dot{)} \mathrel{\acute{\leqslant}} \dot{(}x\,y'\dot{)} \doteq \dot{\top}$ we obtain $\dot{(}x\,y\dot{)} \doteq \dot{\top}$, so $y \in G$ and $G = G' \cap \breve{\Phi} \neq \emptyset$. Hence, $\acute{q}(G) = \breve{q}(G)$ holds whether $G = \emptyset$ or $G \neq \emptyset$, so

$$\dot{(}\dot{C}\,x\dot{)} \doteq \begin{cases} \dot{\perp} & \text{if } \exists y \in \acute{\Phi}\colon \dot{(}x\,y\dot{)} \doteq \dot{\perp}, \\ \acute{q}(\{y \in \acute{\Phi} \mid \dot{(}x\,y\dot{)} \doteq \dot{\top}\}) & \text{otherwise.} \end{cases}$$

If $G \subseteq M$ then $\bigcup G \subseteq \acute{M}$, $\acute{q}(\bigcup G) \in \acute{M}$ and $c(\acute{q}(\bigcup G)) \in M$. Define $q(G) = c(\acute{q}(\bigcup G)) \in M$. We have

$$q(G) \in G \quad \text{for } G \subseteq M,\ G \neq \emptyset, \tag{64}$$

$$q(\emptyset) = \top, \tag{65}$$

$$q(\{y \in \Phi \mid (c(x)\ y) = \top\}) = c(\acute{q}(\{y \in \acute{\Phi} \mid \dot{(}x\,y\dot{)} \doteq \dot{\top}\})).$$

This ends the special investigation of $\dot{C}$ and $\breve{q}(\bullet)$.

From $x \doteq y \Leftrightarrow c(x) = c(y)$ and $\dot{(}\dot{S}\,x\,y\,z\dot{)} \doteq \dot{(}x\,z\,\dot{(}y\,z\dot{)}\dot{)}$ we obtain $c(\dot{(}\dot{S}\,x\,y\,z\dot{)}) = c(\dot{(}x\,z\,\dot{(}y\,z\dot{)}\dot{)})$. From (53)-(63) we obtain $(S\,c(x)\ c(y)\ c(z)) = (c(x)\ c(z)\ (c(y)\ c(z)))$. Since $\forall x' \in M\ \exists x'' \in \acute{M}\colon c(x'') = x'$ we have $(S\,x\,y\,z) = (x\,z\,(y\,z))$ for all $x, y, z \in M$. Likewise,

$$(S\,x\,y) \notin \{\top, \perp\},$$

$$(S\,x\,y\,z) = (x\,z\,(y\,z)),$$

$$(K\,x) \notin \{\top, \perp\},$$

$$(K\,x\,y) = x,$$

$$(\top\,x) = \top,$$

$$(P\,x\,y\,z) = \begin{cases} \perp & \text{if } z = \perp, \\ x & \text{if } z = \top, \\ y & \text{otherwise,} \end{cases}$$

$$(C\,x) = \begin{cases} \perp & \text{if } \exists y \in \Phi\colon (x\,y) = \perp, \\ q(\{y \in \Phi \mid (x\,y) = \top\}) & \text{otherwise,} \end{cases}$$

$$(W\,x) = \begin{cases} \top & \text{if } x \in \Phi, \\ \perp & \text{otherwise,} \end{cases}$$

$$(\perp\,x) = \perp.$$

For all $x, y, z \in \grave{M}$ and $d \in M^V$, the definition of ${}_d x$ yields

$$ {}_d(\dot{S}\,x\,y) \notin \{\mathsf{T}, \bot\}, \tag{66}$$

$$ {}_d(\dot{S}\,x\,y\,z) = {}_d(x\,z\,(y\,z)), \tag{67}$$

$$ {}_d(\dot{K}\,x) \notin \{\mathsf{T}, \bot\}, \tag{68}$$

$$ {}_d(\dot{K}\,x\,y) = {}_d x, \tag{69}$$

$$ {}_d(\dot{\mathsf{T}}\,x) = {}_d\dot{\mathsf{T}}, \tag{70}$$

$$ {}_d(\dot{\text{if}}\,x\,y\,z) = \begin{cases} {}_d\dot{\bot} & \text{if } {}_d x = \bot, \\ {}_d y & \text{if } {}_d x = \mathsf{T}, \\ {}_d z & \text{otherwise}, \end{cases} \tag{71}$$

$$ {}_d\dot{\varepsilon} x = \begin{cases} \bot & \text{if } \exists y \in \Phi\colon ({}_d x\, y) = \bot, \\ q(\{y \in \Phi \mid ({}_d x\, y) = \mathsf{T}\}) & \text{otherwise}, \end{cases} \tag{72}$$

$$ {}_d\dot{\phi} x = \begin{cases} \mathsf{T} & \text{if } x \in \Phi, \\ \bot & \text{otherwise}, \end{cases} \tag{73}$$

$$ {}_d(\dot{\bot}\,x) = {}_d\dot{\bot}. \tag{74}$$

13.3. Binding

We now state a series of well-known lemmas about λ-abstraction and substitution. Each lemma is proved by structural induction and, hence, is a proof involving many cases to check. Since each proof is trivial, we shall omit the details.

Lemma 13.3.1. *Let $u, v \in \grave{V}$, $u \neq v$ and $x, y \in \grave{M}'$. Assume $\neg \mathit{free}(v, y)$. We have*

$$\grave{\lambda} v.[x/u := y] = [\grave{\lambda} v.x/u := y].$$

Proof. By structural induction on x. $\square$

Lemma 13.3.2. *Let $u \in \grave{V}$ and $x, y \in \grave{M}$. Assume $\mathit{freefor}(x, u, y)$. We have*

$$[\![[x/u := y]]\!] = [[\![x]\!]/u := [\![y]\!]].$$

Proof. By structural induction on x. The only nontrivial case in the proof is the case where x has the form $\grave{\lambda} v.x'$ where $v \in \grave{V}$, $v \neq u$ and $x' \in \grave{M}$ is assumed to satisfy the lemma. In this case we shall prove

$$[\![[\grave{\lambda} v.x'/u := y]]\!] = [[\![\grave{\lambda} v.x']\!]/u := [\![y]\!]].$$

From $x = \grave{\lambda} v.x'$ and $\mathit{freefor}(x, u, y)$ we deduce $\neg\mathit{free}(v, y) \vee \neg\mathit{free}(u, x')$. If $\neg\mathit{free}(u, x')$ then

$$[\![[\grave{\lambda} v.x'/u := y]]\!] = [\![\grave{\lambda} v.x']\!] = [[\![\grave{\lambda} v.x']\!]/u := [\![y]\!]].$$

If $\neg free(v, y)$ then we use the definitions of $\langle\bullet\rangle$ and $[\bullet/\bullet := \bullet]$, the previous lemma and the inductive hypothesis that x' satisfies the present lemma:

$$\begin{aligned}\langle[\dot\lambda v.x'/u := y]\rangle &= \langle\dot\lambda v.[x'/u := y]\rangle\\ &= \grave\lambda v.\langle[x'/u := y]\rangle\\ &= \grave\lambda v.[\langle x'\rangle/u := \langle y\rangle]\\ &= [\grave\lambda v.\langle x'\rangle/u := \langle y\rangle]\\ &= [\langle\dot\lambda v.x'\rangle/u := \langle y\rangle]. \qquad \Box\end{aligned}$$

Lemma 13.3.3. *Let $u \in \grave V$ and $x, y \in \grave M'$. We have*

$$\grave(\grave\lambda u.x\, y\grave) \doteq [x/u := y].$$

Proof. By structural induction on x using the definition of $\grave\lambda$ and the equations (67) and (69). In particular, we have

$$\grave(\grave\lambda u.u\, y\grave) = \grave(\dot S\, \dot K\, \dot K\, y\grave) \doteq \grave(\dot K\, y\, \grave(\dot K\, y\grave)\grave) \doteq y = [u/u := y]. \qquad \Box$$

Lemma 13.3.4. *Let $u \in \grave V$, $x \in \grave M$ and $d \in M^V$. We have ${}_d\grave\lambda u.x \notin \{\mathsf{T}, \bot\}$.*

Proof. This lemma is *not* proved by structural induction. Using (66) and (68) we prove the lemma by three cases.

- If $x = u$ then ${}_d\grave\lambda u.x = {}_d\grave(\dot S\, \dot K\, \dot K\grave) \notin \{\mathsf{T}, \bot\}$.
- If $\neg free(u, x)$ then ${}_d\grave\lambda u.x = {}_d\grave(\dot K\, x\grave) \notin \{\mathsf{T}, \bot\}$.
- If $x = \grave(x'\, x''\grave)$ and $free(u, x)$ then ${}_d\grave\lambda u.x = {}_d\grave(\dot S\, \grave\lambda u.x'\, \grave\lambda u.x''\grave) \notin \{\mathsf{T}, \bot\}$. $\Box$

13.4. Computation axioms and QND'

Lemma 13.4.1 (Extensionality). *Let $d \in M^V$ and $x, y \in \grave M$. Assume ${}_dx \notin \{\mathsf{T}, \bot\}$, ${}_dy \notin \{\mathsf{T}, \bot\}$ and $\forall z \in M$: ${}_d\grave(x\, z\grave) = {}_d\grave(y\, z\grave)$. We have ${}_dx = {}_dy$.*

Proof. Let $x, y \in \acute M$ and assume $x \,\acute{\neq}\, \dot{\mathsf{T}}$, $x \,\acute{\neq}\, \dot\bot$, $y \,\acute{\neq}\, \dot{\mathsf{T}}$, $y \,\acute{\neq}\, \dot\bot$ and $\forall z \in \acute M$: $\grave(x\, z\grave) \,\acute{=}\, \grave(y\, z\grave)$. From the root theorem we have $\acute r(x) = \tilde\lambda = \acute r(y)$, so $\acute a(x)(\langle\,\rangle) = \acute a(x)(\langle\,\rangle)$. Let $z \in \acute M$ and $u \in \acute M^*$. From $\grave(x\, z\grave) \,\acute{=}\, \grave(y\, z\grave)$ we obtain $\acute a(x)(\langle z\rangle \cdot u) = \acute a(\grave(x\, z\grave))(u) = \acute a(\grave(y\, z\grave))(u) = \acute a(y)(\langle z\rangle \cdot u)$. We now have $\acute a(x)(u) = \acute a(y)(u)$ for all $u \in \acute M^*$, so $x \,\acute{=}\, y$. Hence,

$$x \,\acute{\neq}\, \dot{\mathsf{T}} \wedge x \,\acute{\neq}\, \dot\bot \wedge y \,\acute{\neq}\, \dot{\mathsf{T}} \wedge y \,\acute{\neq}\, \dot\bot \wedge \forall z \in \acute M\colon \grave(x\, z\grave) \,\acute{=}\, \grave(y\, z\grave) \Rightarrow x \,\acute{=}\, y.$$

This entails that

$$x \notin \{\mathsf{T}, \bot\} \wedge y \notin \{\mathsf{T}, \bot\} \wedge \forall z \in M\colon (x\, z) = (y\, z) \Rightarrow x = y$$

for all $x, y \in M$, and

$${}_dx \notin \{\mathsf{T}, \bot\} \wedge {}_dy \notin \{\mathsf{T}, \bot\} \wedge \forall z \in M\colon {}_d\grave(x\, z\grave) = {}_d\grave(y\, z\grave) \Rightarrow {}_dx = {}_dy$$

for all $x, y \in \grave M$ and $d \in M^V$. $\Box$

Theorem 13.4.2 (Computation axioms). *For $\dot u, \dot v \in \dot V$ and $u, v, x, y, z \in \grave M$ we have*

(Trans) $x \doteq y \wedge x \doteq z \;\Rightarrow\; y \doteq z$

(Sub1) $u \doteq v \wedge x \doteq y \;\Rightarrow\; \dot{(}u\, x\dot{)} \doteq \dot{(}v\, y\dot{)}$

(Sub2) $x \doteq y \;\Rightarrow\; \dot\lambda \dot u.x \doteq \dot\lambda \dot u.y$

(Ren) $\mathit{freefor}(x, \dot u, \dot v) \wedge \mathit{freefor}(x, \dot v, \dot u) \;\Rightarrow\; \dot\lambda \dot u.[x/\dot v := \dot u] \doteq \dot\lambda \dot v.[x/\dot u := \dot v]$

(Ap1) $\dot{(}\dot{\mathsf{T}}\, x\dot{)} \doteq \dot{\mathsf{T}}$

(Ap2) $\mathit{freefor}(x, \dot u, y) \;\Rightarrow\; \dot{(}\dot\lambda \dot u.x\, y\dot{)} \doteq [x/\dot u := y]$

(Ap3) $\dot{(}\dot\perp\, x\dot{)} \doteq \dot\perp$

(Sel1) $(\dot{\mathsf{if}}\; \dot{\mathsf{T}}\; y\, z) \doteq y$

(Sel2) $(\dot{\mathsf{if}}\; \dot\lambda \dot u.x\; y\, z) \doteq z$

(Sel3) $(\dot{\mathsf{if}}\; \dot\perp\; y\, z) \doteq \dot\perp$.

Proof. (Trans) follows from ${}_dx = {}_dy \wedge {}_dx = {}_dz \Rightarrow {}_dy = {}_dz$. (Sub1) follows from ${}_du = {}_dv \wedge {}_dx = {}_dy \Rightarrow ({}_du\; {}_dx) = ({}_dv\; {}_dy) \Rightarrow {}_d\dot{(}u\, x\dot{)} = {}_d\dot{(}v\, y\dot{)}$. (Sub2) Assume $x \doteq y$. We have $\forall z \in M$: ${}_d[x/\dot u := z] = {}_d[y/\dot u := z]$. Hence, ${}_d\dot{(}\dot\lambda \dot u.x\, z\dot{)} = {}_d\dot{(}\dot\lambda \dot u.y\, z\dot{)}$ by Lemma 13.3.3. Further, ${}_d\dot\lambda \dot u.x \notin \{\mathsf{T}, \perp\}$ and ${}_d\dot\lambda \dot u.y \notin \{\mathsf{T}, \perp\}$ by Lemma 13.3.4, so ${}_d\dot\lambda \dot u.x = {}_d\dot\lambda \dot u.y$ by Lemma 13.4.1. (Ren) By structural induction in x one easily proves $\grave\lambda \dot u.[x/\dot v := \dot u] = \grave\lambda \dot v.[x/\dot u := \dot v]$ for all $x \in \grave M'$. Assume $\mathit{freefor}(x, \dot u, \dot v)$ and $\mathit{freefor}(x, \dot v, \dot u)$. Using Lemma 13.3.2 we have $[\![\dot\lambda \dot u.[x/\dot v := \dot u]]\!] = \grave\lambda \dot u.[\![[x/\dot v := \dot u]]\!] = \grave\lambda \dot u.[[\![x]\!]/\dot v := [\![\dot u]\!]] = \grave\lambda \dot v.[[\![x]\!]/\dot u := [\![\dot v]\!]] = \grave\lambda \dot v.[\![[x/\dot u := \dot v]]\!] = [\![\dot\lambda \dot v.[x/\dot u := \dot v]]\!]$ from which Ren follows. (Ap1) follows from (70). (Ap2) Assume $\mathit{freefor}(x, \dot u, y)$. From Lemmas 13.3.2 and 13.3.3 we have ${}_d\dot{(}\dot\lambda \dot u.x\; y\dot{)} = {}_d[\![\dot{(}\dot\lambda \dot u.x\; y\dot{)}]\!] = {}_d\dot{(}\grave\lambda \dot u.[\![x]\!]\; [\![y]\!]\dot{)} = {}_d[[\![x]\!]/\dot u := [\![y]\!]] = {}_d[\![[x/\dot u := y]]\!] = {}_d[x/\dot u := y]$. (Ap3) follows from (74). (Sel1) follows from (71). (Sel2) follows from (71) and Lemma 13.3.4 (Sel3) follows from (71). $\square$

Lemma 13.4.3. *Let $d \in M^V$, $u, v \in \grave V$, $u \neq v$ and $x \in \grave M$. Assume ${}_dx \notin \{\mathsf{T}, \perp\}$. We have ${}_d\dot{(}\dot\lambda u.\dot\lambda v.\dot{(}u\, v\dot{)}\; x\dot{)} = {}_dx$.*

Proof. Let $v' \in \dot V$ satisfy $\neg\mathit{free}(v', x)$ and $v' \neq u$. We have

$$\dot{(}\dot\lambda u.\dot\lambda v.\dot{(}u\, v\dot{)}\; x\dot{)} \doteq \dot{(}\dot\lambda u.\dot\lambda v'.\dot{(}u\, v'\dot{)}\; x\dot{)} \doteq \dot\lambda v'.\dot{(}x\, v'\dot{)}.$$

For all $y \in M$ we have

$$({}_d\dot{(}\dot\lambda u.\dot\lambda v.\dot{(}u\, v\dot{)}\; x\dot{)}\; y) = ({}_d\dot\lambda v'.\dot{(}x\, v'\dot{)}\; y) = {}_d\dot{(}\dot\lambda v'.\dot{(}x\, v'\dot{)}\; y\dot{)} = {}_d\dot{(}x\, y\dot{)} = ({}_dx\; y).$$

From Lemma 13.3.4 we obtain ${}_d\dot{(}\dot\lambda u.\dot\lambda v.\dot{(}u\, v\dot{)}\; x\dot{)} = {}_d\dot\lambda v'.\dot{(}x\, v'\dot{)} \notin \{\mathsf{T}, \perp\}$. Hence, ${}_d\dot{(}\dot\lambda u.\dot\lambda v.\dot{(}u\, v\dot{)}\; x\dot{)} = {}_dx$ by Lemma 13.4.1. $\square$

Theorem 13.4.4 (*QND'*). *Let $u, v, w \in \dot{V}$, $v \neq w$ and $x, y \in \dot{M}$. We have*

$$[x/u := \dot{\mathsf{T}}] \doteq [y/u := \dot{\mathsf{T}}]$$
$$\wedge [x/u := \dot{(}\dot{\lambda} v.\dot{\lambda} w.\dot{(}v\, w\dot{)}\ u\dot{)}] \doteq [y/u := \dot{(}\dot{\lambda} v.\dot{\lambda} w.\dot{(}v\, w\dot{)}\ u\dot{)}]$$
$$\wedge [x/u := \dot{\perp}] \doteq [y/u := \dot{\perp}]$$
$$\Rightarrow x \doteq y.$$

Proof. We obviously have ${}_d x = {}_d[x/u := d(u)]$. If $d(u) = \mathsf{T}$ then $[x/u := \dot{\mathsf{T}}] \doteq [y/u := \dot{\mathsf{T}}]$ yields

$${}_d x = {}_d[x/u := d(u)] = {}_d[y/u := d(u)] = {}_d y.$$

The proof of ${}_d x = {}_d y$ for $d(u) = \perp$ is analogous. If $d(u) \notin \{\mathsf{T}, \perp\}$ then

$${}_d x = {}_d[x/u := d(u)] = {}_d[x/u := \dot{(}\dot{\lambda} v.\dot{\lambda} w.\dot{(}v\, w\dot{)}\ d(u)\dot{)}]$$
$$= {}_d[x/u := \dot{(}\dot{\lambda} v.\dot{\lambda} w.\dot{(}v\, w\dot{)}\ u\dot{)}]$$
$$= {}_d[y/u := \dot{(}\dot{\lambda} v.\dot{\lambda} w.\dot{(}v\, w\dot{)}\ u\dot{)}] = {}_d[y/u := \dot{(}\dot{\lambda} v.\dot{\lambda} w.\dot{(}v\, w\dot{)}\ d(u)\dot{)}]$$
$$= {}_d[y/u := d(u)] = {}_d y.$$

Hence, ${}_d x = {}_d y$ for all $d \in M^V$. □

13.5. Semantics of defined concepts

In Part I we defined T, F, etc. in map theory. When translated into the notation used for the consistency proof, the definitions are as follows.

$$\dot{\mathsf{T}} = \dot{\mathsf{T}}.$$
$$\dot{\mathsf{F}} = \dot{\lambda} x.\dot{\mathsf{T}}.$$
$$\dot{\neg} x = \dot{(}\mathsf{if}\ x\ \dot{\mathsf{F}}\ \dot{\mathsf{T}}\dot{)}.$$
$$\dot{\approx} x = \dot{(}\mathsf{if}\ x\ \dot{\mathsf{T}}\ \dot{\mathsf{F}}\dot{)}.$$
$$\dot{!} x = \dot{(}\mathsf{if}\ x\ \dot{\mathsf{T}}\ \dot{\mathsf{T}}\dot{)}.$$
$$\dot{\mathsf{i}} x = \dot{(}\mathsf{if}\ x\ \dot{\mathsf{F}}\ \dot{\mathsf{F}}\dot{)}.$$
$$x \mathbin{\dot{\wedge}} y = \dot{(}\mathsf{if}\ x\ \dot{(}\mathsf{if}\ y\ \dot{\mathsf{T}}\ \dot{\mathsf{F}}\dot{)}\ \dot{(}\mathsf{if}\ y\ \dot{\mathsf{F}}\ \dot{\mathsf{F}}\dot{)}\dot{)}.$$
$$x \mathbin{\dot{\vee}} y = \dot{(}\mathsf{if}\ x\ \dot{(}\mathsf{if}\ y\ \dot{\mathsf{T}}\ \dot{\mathsf{T}}\dot{)}\ \dot{(}\mathsf{if}\ y\ \dot{\mathsf{T}}\ \dot{\mathsf{F}}\dot{)}\dot{)}.$$
$$x \mathbin{\dot{\Rightarrow}} y = \dot{(}\mathsf{if}\ x\ \dot{(}\mathsf{if}\ y\ \dot{\mathsf{T}}\ \dot{\mathsf{F}}\dot{)}\ \dot{(}\mathsf{if}\ y\ \dot{\mathsf{T}}\ \dot{\mathsf{T}}\dot{)}\dot{)}.$$
$$x \mathbin{\dot{\Leftrightarrow}} y = \dot{(}\mathsf{if}\ x\ \dot{(}\mathsf{if}\ y\ \dot{\mathsf{T}}\ \dot{\mathsf{F}}\dot{)}\ \dot{(}\mathsf{if}\ y\ \dot{\mathsf{F}}\ \dot{\mathsf{T}}\dot{)}\dot{)}.$$
$$x \mathbin{\dot{:}} y = \dot{(}\mathsf{if}\ x\ y\ \dot{\mathsf{T}}\dot{)}.$$
$$\dot{\mathsf{Y}} = \dot{\lambda} f.\dot{(}\dot{\lambda} \dot{x}.\dot{(}f\ \dot{(}\dot{x}\ \dot{x}\dot{)}\dot{)}\ \dot{\lambda} \dot{x}.\dot{(}f\ \dot{(}\dot{x}\ \dot{x}\dot{)}\dot{)}\dot{)}.$$
$$\dot{\mathsf{Y}} u.x = \dot{(}\dot{\mathsf{Y}}\ \dot{\lambda} u.x\dot{)}.$$
$$\dot{\exists} = \dot{\lambda} \dot{x}.\dot{\approx}\dot{(}\dot{x}\ \dot{\varepsilon}\dot{x}\dot{)}.$$
$$\dot{\exists} u.x = \dot{(}\dot{\exists}\dot{\lambda} u.x\dot{)}.$$
$$\dot{\forall} u.x = \dot{\neg}\dot{\exists} u.\dot{\neg} x.$$

We have ${}_d\dot{\mathsf{T}}=\mathsf{T}$ and ${}_d\dot{\mathsf{F}}=\mathsf{F}\notin\{\mathsf{T},\bot\}$. From (71) we obtain the following lemma.

Lemma 13.5.1. *Let* $d\in M^V$. *Assume* $x, y, u, v\in \grave{M}$, ${}^{\mathrm{D}}_{d}u$ *and* ${}^{\mathrm{D}}_{d}v$. *We have*

$$\begin{array}{lll}
{}^{\mathrm{B}}_{d}\dot{\mathsf{T}} & {}^{\mathrm{D}}_{d}\dot{\mathsf{T}} \Leftrightarrow 0=0 & {}^{\mathrm{T}}_{d}\dot{\mathsf{T}} \Leftrightarrow 0=0 \\
{}^{\mathrm{B}}_{d}\dot{\mathsf{F}} & {}^{\mathrm{D}}_{d}\dot{\mathsf{F}} \Leftrightarrow 0=0 & {}^{\mathrm{T}}_{d}\dot{\mathsf{F}} \Leftrightarrow 0=1 \\
{}^{\mathrm{B}}_{d}\dot{\neg}x & {}^{\mathrm{D}}_{d}\dot{\neg}x \Leftrightarrow {}^{\mathrm{D}}_{d}x & {}^{\mathrm{T}}_{d}\dot{\neg}u \Leftrightarrow \neg{}^{\mathrm{T}}_{d}u \\
{}^{\mathrm{B}}_{d}\dot{\approx}x & {}^{\mathrm{D}}_{d}\dot{\approx}x \Leftrightarrow {}^{\mathrm{D}}_{d}x & {}^{\mathrm{T}}_{d}\dot{\approx}u \Leftrightarrow {}^{\mathrm{T}}_{d}u \\
{}^{\mathrm{B}}_{d}\dot{!}x & {}^{\mathrm{D}}_{d}\dot{!}x \Leftrightarrow {}^{\mathrm{D}}_{d}x & {}^{\mathrm{T}}_{d}\dot{!}u \Leftrightarrow 0=0 \\
{}^{\mathrm{B}}_{d}\dot{¡}x & {}^{\mathrm{D}}_{d}\dot{¡}x \Leftrightarrow {}^{\mathrm{D}}_{d}x & {}^{\mathrm{T}}_{d}\dot{¡}u \Leftrightarrow 0=1 \\
{}^{\mathrm{B}}_{d}x\mathbin{\dot{\wedge}}y & {}^{\mathrm{D}}_{d}x\mathbin{\dot{\wedge}}y \Leftrightarrow {}^{\mathrm{D}}_{d}x\wedge{}^{\mathrm{D}}_{d}y & {}^{\mathrm{T}}_{d}u\mathbin{\dot{\wedge}}v \Leftrightarrow {}^{\mathrm{T}}_{d}u\wedge{}^{\mathrm{T}}_{d}v \\
{}^{\mathrm{B}}_{d}x\mathbin{\dot{\vee}}y & {}^{\mathrm{D}}_{d}x\mathbin{\dot{\vee}}y \Leftrightarrow {}^{\mathrm{D}}_{d}x\wedge{}^{\mathrm{D}}_{d}y & {}^{\mathrm{T}}_{d}u\mathbin{\dot{\vee}}v \Leftrightarrow {}^{\mathrm{T}}_{d}u\vee{}^{\mathrm{T}}_{d}v \\
{}^{\mathrm{B}}_{d}x\mathrel{\dot{\Rightarrow}}y & {}^{\mathrm{D}}_{d}x\mathrel{\dot{\Rightarrow}}y \Leftrightarrow {}^{\mathrm{D}}_{d}x\wedge{}^{\mathrm{D}}_{d}y & {}^{\mathrm{T}}_{d}u\mathrel{\dot{\Rightarrow}}v \Leftrightarrow {}^{\mathrm{T}}_{d}u\Rightarrow{}^{\mathrm{T}}_{d}v \\
{}^{\mathrm{B}}_{d}x\mathrel{\dot{\Leftrightarrow}}y & {}^{\mathrm{D}}_{d}x\mathrel{\dot{\Leftrightarrow}}y \Leftrightarrow {}^{\mathrm{D}}_{d}x\wedge{}^{\mathrm{D}}_{d}y & {}^{\mathrm{T}}_{d}u\mathrel{\dot{\Leftrightarrow}}v \Leftrightarrow {}^{\mathrm{T}}_{d}u\Leftrightarrow{}^{\mathrm{T}}_{d}v.
\end{array}$$

Define

$$x\leqslant y \;\Leftrightarrow\; \exists x'\in x\;\exists y'\in y\colon x'\mathrel{\acute{\leqslant}}y'.$$

Lemma 13.5.2. *If* $x', y'\in\acute{M}$ *then*

$$c(x')\leqslant c(y') \;\Leftrightarrow\; x'\mathrel{\acute{\leqslant}}y'.$$

Proof. Let $x', y'\in\acute{M}$. Assume $x'\mathrel{\acute{\leqslant}}y'$. From $x'\in c(x')$ and $y'\in c(y')$ we deduce $\exists x''\in c(x')\;\exists y''\in c(y')\colon x'\mathrel{\acute{\leqslant}}y'$ which entails $c(x')\leqslant c(y')$. Hence, $x'\mathrel{\acute{\leqslant}}y'\Rightarrow c(x')\leqslant c(y')$. Now assume $c(x')\leqslant c(y')$. Choose $x''\in c(x)$ and $y''\in c(y)$ such that $x''\mathrel{\acute{\leqslant}}y''$. From $x''\in c(x')$ we have $x'\doteq x''$ and from $y''\in c(y')$ we have $y'\doteq y''$ which combined with $x''\mathrel{\acute{\leqslant}}y''$ gives $x'\mathrel{\acute{\leqslant}}y'$. $\square$

From the properties of $\acute{\leqslant}$ we obtain the following result.

Lemma 13.5.3. *For all* $u, v, x, y, z\in M$ *we have*

$$\begin{array}{l}
x\leqslant x,\\
x\leqslant y\wedge y\leqslant x \;\Rightarrow\; x=y,\\
x\leqslant y\wedge y\leqslant z \;\Rightarrow\; x\leqslant z,\\
u\leqslant v\wedge x\leqslant y \;\Rightarrow\; (u\,x)\leqslant(v\,y),\\
\bot\leqslant x,\\
x\notin\{\mathsf{T},\bot\} \;\Rightarrow\; x\not\leqslant\mathsf{T}\wedge\mathsf{T}\not\leqslant x,\\
x, y\notin\{\mathsf{T},\bot\} \;\Rightarrow\; (x\leqslant y\Leftrightarrow\forall z\in M\colon(x\,z)\leqslant(y\,z)).
\end{array}$$

Lemma 13.5.4. *Let $d \in M^V$ and $x \in \dot{M}$. We have*

$$ {}_d\dot{\exists} x = \begin{cases} \bot & \text{if } \exists y \in \Phi\colon {}_d\dot{(}x\,y\dot{)} = \bot, \\ \mathsf{F} & \text{if } \forall y \in \Phi\colon {}_d\dot{(}x\,y\dot{)} \notin \{\mathsf{T}, \bot\}, \\ \mathsf{T} & \text{otherwise.} \end{cases} $$

Proof. Assume $d \in M^V$ and $x \in \dot{M}$. Assume $\exists y \in \Phi\colon {}_d\dot{(}x\,y\dot{)} = \bot$. Choose $y \in \Phi$ such that ${}_d\dot{(}x\,y\dot{)} = \bot$. From (72) we have ${}_d\dot{\varepsilon} x = \bot$. Hence, ${}_d\dot{\varepsilon} x \leqslant y$ so ${}_d\dot{(}x\,\dot{\varepsilon} x\dot{)} \leqslant {}_d\dot{(}x\,y\dot{)} = \bot$ which entails ${}_d\dot{(}x\,\dot{\varepsilon} x\dot{)} = \bot$ and ${}_d\dot{\exists} x = {}_d\dot{\approx}\dot{(}x\,\dot{\varepsilon} x\dot{)} = \bot$.

Now assume $\forall y \in \Phi\colon {}_d\dot{(}x\,y\dot{)} \notin \{\mathsf{T}, \bot\}$. From this assumption, (72), and (65) we have ${}_d\dot{\varepsilon} x = \mathsf{T} \in \Phi$. Using the assumption $\forall y \in \Phi\colon {}_d\dot{(}x\,y\dot{)} \notin \{\mathsf{T}, \bot\}$ once more we obtain ${}_d\dot{(}x\,\dot{\varepsilon} x\dot{)} \notin \{\mathsf{T}, \bot\}$ from which ${}_d\dot{\exists} x = {}_d\dot{\approx}\dot{(}x\,\dot{\varepsilon} x\dot{)} = \mathsf{F}$ follows.

Now assume $\neg\exists y \in \Phi\colon {}_d\dot{(}x\,y\dot{)} = \bot$ and $\neg\forall y \in \Phi\colon {}_d\dot{(}x\,y\dot{)} \notin \{\mathsf{T}, \bot\}$. From the assumptions we obtain $\exists y \in \Phi\colon {}_d\dot{(}x\,y\dot{)} = \mathsf{T}$. Hence, from (72) and (64) we have ${}_d\dot{\varepsilon} x \in \{y \in \Phi \mid {}_d\dot{(}x\,y\dot{)} = \mathsf{T}\}$ which proves ${}_d\dot{(}x\,\dot{\varepsilon} x\dot{)} = \mathsf{T}$ and ${}_d\dot{\exists} x = {}_d\dot{\approx}\dot{(}x\,\dot{\varepsilon} x\dot{)} = \mathsf{T}$. □

From this lemma we immediately conclude Lemma 13.5.5.

Lemma 13.5.5. *Let $d \in M^V$ and $x, y \in \dot{M}$. Assume $\forall z \in \Phi\colon {}^{\mathsf{D}}_{d}\dot{(}y\,z\dot{)}$. We have*

$$ {}^{\mathsf{B}}_{d}\dot{\exists} x \quad {}^{\mathsf{D}}_{d}\dot{\exists} x \Leftrightarrow \forall z \in \Phi\colon {}^{\mathsf{D}}_{d}\dot{(}x\,z\dot{)} \quad {}^{\mathsf{T}}_{d}\dot{\exists} y \Leftrightarrow \exists z \in \Phi\colon {}^{\mathsf{T}}_{d}\dot{(}y\,z\dot{)}. $$

Using this lemma and Ap2 in Theorem 13.4.2 and Lemma 13.5.1 we obtain the following result.

Lemma 13.5.6. *Let $d \in M^V$, $x, y \in \dot{M}$* and *$u \in \dot{V}$. Assume $\forall z \in M\colon {}^{\mathsf{D}}_{d}[y/u := z]$. We have*

$$ {}^{\mathsf{B}}_{d}\dot{\exists} u.x \quad {}^{\mathsf{D}}_{d}\dot{\exists} u.x \Leftrightarrow \forall z \in \Phi\colon {}^{\mathsf{D}}_{d}[x/u := z] \quad {}^{\mathsf{T}}_{d}\dot{\exists} u.y \Leftrightarrow \exists z \in \Phi\colon {}^{\mathsf{T}}_{d}[y/u := z], $$

$$ {}^{\mathsf{B}}_{d}\dot{\forall} u.x \quad {}^{\mathsf{D}}_{d}\dot{\forall} u.x \Leftrightarrow \forall z \in \Phi\colon {}^{\mathsf{D}}_{d}[x/u := z] \quad {}^{\mathsf{T}}_{d}\dot{\forall} u.y \Leftrightarrow \forall z \in \Phi\colon {}^{\mathsf{T}}_{d}[y/u := z]. $$

Using Ap2 in Theorem 13.4.2 several times we obtain Lemma 13.5.7.

Lemma 13.5.7. *Let $d \in M^V$, $x, y \in \dot{M}$ and $u \in \dot{V}$. Assume freefor$(y, u, \dot{\mathsf{Y}} u.y)$. We have*

$$ {}_d\dot{\mathsf{Y}} x = {}_d\dot{(}x\;\dot{\mathsf{Y}} x\dot{)}, $$

$$ {}_d\dot{\mathsf{Y}} u.y = {}_d[y/u := \dot{\mathsf{Y}} u.y]. $$

As stated in Part I, $x_1, \ldots, x_\alpha \to y$ stands for $x_1 : \cdots : x_\alpha : y = x_1 : \cdots : x_\alpha : \mathsf{T}$, and this construct is used in several axioms. Define $x_1, \ldots, x_\alpha \mathrel{\dot{\to}} y \Leftrightarrow x_1 \mathrel{\dot{:}} \cdots \mathrel{\dot{:}} x_\alpha \mathrel{\dot{:}} y \doteq x_1 \mathrel{\dot{:}} \cdots \mathrel{\dot{:}} x_\alpha \mathrel{\dot{:}} \dot{\mathsf{T}}$. The following lemma is useful for verifying axioms involving $\to$.

Lemma 13.5.8 (Deduction lemma). *Let $x_1, \ldots, x_\alpha, y \in \dot{M}$ and $d \in M^V$.*

If ${}^{\mathsf{T}}_{d}x_1 \wedge \cdots \wedge {}^{\mathsf{T}}_{d}x_\alpha \Rightarrow {}^{\mathsf{T}}_{d}y$ then ${}_d x_1 \mathrel{\dot{:}} \cdots \mathrel{\dot{:}} x_n \mathrel{\dot{:}} y = {}_d x_1 \mathrel{\dot{:}} \cdots \mathrel{\dot{:}} x_n \mathrel{\dot{:}} \dot{\mathsf{T}}$.

If $\forall d \in M^V\colon ({}^{\mathsf{T}}_{d}x_1 \wedge \cdots \wedge {}^{\mathsf{T}}_{d}x_\alpha \Rightarrow {}^{\mathsf{T}}_{d}y)$ then $x_1, \ldots, x_\alpha \mathrel{\dot{\to}} y$.

Proof. From the definition $x \dot{:} y = (\dot{\text{if}}\; x\; y\; \dot{\mathsf{T}})$ we deduce ${}_d x = \bot \Rightarrow {}_d(x \dot{:} y) = \bot$ and ${}_d x \notin \{\mathsf{T}, \bot\} \Rightarrow {}_d(x \dot{:} y) = \mathsf{T}$. Hence, ${}_d x_1 \dot{:} \cdots \dot{:} x_\alpha \dot{:} y = {}_d x_1 \dot{:} \cdots \dot{:} x_\alpha \dot{:} \dot{\mathsf{T}}$ holds if ${}_d x_1 \neq \mathsf{T} \vee \cdots \vee {}_d x_\alpha \neq \mathsf{T}$. Now assume ${}_d x_1 = \cdots = {}_d x_\alpha = \mathsf{T}$. From the assumption we have ${}_d y = \mathsf{T}$ so ${}_d x_1 \dot{:} \cdots \dot{:} x_\alpha \dot{:} y = {}_d x_1 \dot{:} \cdots \dot{:} x_\alpha \dot{:} \dot{\mathsf{T}}$. The second claim follows from the definition of $\dot{\rightarrow}$ $\square$

13.6. Quantification axioms

Lemma 13.6.1 (Quan1). *Let $u \in \dot{V}$ and $x, y \in \dot{M}$. We have*

$$\dot{\phi} x, \dot{\forall} u.y \dot{\rightarrow} (\dot{\lambda} u.y\, x\dot{)}.$$

Proof. Let $d \in M^V$. Assume ${}^{\mathsf{T}}_d \dot{\phi} x$ and ${}^{\mathsf{T}}_d \dot{\forall} u.y$. We have ${}_d x \in \Phi$ and $\forall z \in \Phi$: ${}^{\mathsf{T}}_d[y/u := z]$. Hence, $\forall z \in \Phi$: $({}_d\dot{\lambda} u.y\; z) = \mathsf{T}$ so ${}_d(\dot{\lambda} u.y\, x\dot{)} = ({}_d\dot{\lambda} u.y\; {}_d x) = \mathsf{T}$. Now the lemma follows from the deduction lemma. $\square$

Lemma 13.6.2 (Quan2). *Let $u \in \dot{V}$ and $x \in \dot{M}$. We have*

$$\dot{\varepsilon} u.x \doteq \dot{\varepsilon} u.(\dot{\phi} u \mathbin{\dot{\wedge}} x).$$

Proof. Let $d \in M^V$. If ${}_d u \in \Phi$ then ${}^{\mathsf{D}}_d \dot{\phi} u$ and ${}^{\mathsf{T}}_d \dot{\phi} u$, so ${}^{\mathsf{D}}_d x \Leftrightarrow {}^{\mathsf{D}}_d \dot{\phi} u \mathbin{\dot{\wedge}} x$ and ${}^{\mathsf{T}}_d x \Leftrightarrow {}^{\mathsf{T}}_d \dot{\phi} u \mathbin{\dot{\wedge}} x$. The lemma now follows from (72). $\square$

Lemma 13.6.3 (Quan3). *Let $u \in \dot{V}$ and $x \in \dot{M}$. We have*

$$\dot{\phi} \dot{\varepsilon} u.x \doteq \dot{\forall} u.\dot{!}x.$$

Proof. Let $d \in M^V$. If $\exists z \in \Phi$: $\neg {}^{\mathsf{D}}_d[x/u := z]$ then $\exists z \in \Phi$: $\neg {}^{\mathsf{D}}_d[\dot{!}x/u := z]$ so $\neg {}^{\mathsf{D}}_d \dot{\varepsilon} u.x$ and $\neg {}^{\mathsf{D}}_d \dot{\forall} u.\dot{!}x$. Hence, ${}_d \dot{\phi} \dot{\varepsilon} u.x = \bot = {}_d \dot{\forall} u.\dot{!}x$. Now assume $\forall z \in \Phi$: ${}^{\mathsf{D}}_d[x/u := z]$. We have ${}_d \dot{\varepsilon} u.x \in \Phi$ so ${}^{\mathsf{T}}_d \dot{\phi} \dot{\varepsilon} u.x$. Further, $\forall z \in \Phi$: ${}^{\mathsf{T}}_d[\dot{!}x/u := z]$ so ${}^{\mathsf{T}}_d \dot{\forall} u.\dot{!}x$. Hence, $\dot{\phi} \dot{\varepsilon} u.x \doteq \dot{\forall} u.\dot{!}x$. $\square$

Lemma 13.6.4 (Quan4). *Let $u \in \dot{V}$ and $x \in \dot{M}$. We have*

$$\dot{\exists} u.x \dot{\rightarrow} \dot{\phi} \dot{\varepsilon} u.x.$$

Proof. Let $d \in M^V$ and assume ${}^{\mathsf{T}}_d \dot{\exists} u.x$. From Lemma 13.5.6 we have $\forall z \in \Phi$: ${}^{\mathsf{D}}_d(\dot{\lambda} u.x\, z\dot{)}$ and $\exists z \in \Phi$: ${}^{\mathsf{T}}_d(\dot{\lambda} u.x\, z\dot{)}$. Hence, by (72), ${}_d \dot{\varepsilon} u.x \in \Phi$ so ${}^{\mathsf{T}}_d \dot{\phi} \dot{\varepsilon} u.x$ by (73). Now Quan4 follows from the deduction lemma. $\square$

Lemma 13.6.5 (Quan5). *Let $u \in \dot{V}$ and $x \in \dot{M}$. We have*

$$\dot{\forall} u.x \doteq \dot{\forall} u.(\dot{\phi} u \mathbin{\dot{\wedge}} x).$$

Proof. Let $d \in M^V$. If ${}_d u \in \Phi$ then ${}^{\mathrm{D}}_{d}\dot{\phi} u$ and ${}^{\mathrm{T}}_{d}\dot{\phi} u$, so ${}^{\mathrm{D}}_{d}x = {}^{\mathrm{D}}_{d}\dot{\phi} u \mathbin{\dot{\wedge}} x$ and ${}^{\mathrm{T}}_{d}x = {}^{\mathrm{T}}_{d}\dot{\phi} u \mathbin{\dot{\wedge}} x$. The lemma now follows from Lemma 13.5.6. □

Lemma 13.6.6 (Ind). *Let $u, v \in \grave{V}$, $u \neq v$ and $x_1, \ldots, x_\alpha, y \in \grave{M}$. Assume*

$$\mathit{freefor}(y, u, \dot{(}u\, v\dot{)}) \quad \textit{and} \quad \neg \mathit{free}(u, x_1), \ldots, \neg \mathit{free}(u, x_\alpha).$$

Assume

$$x_1, \ldots, x_\alpha \mathrel{\dot{\rightarrow}} [y/u := \mathsf{T}],$$

$$x_1, \ldots, x_\alpha, \dot{\phi} u, \dot{\neg} u, \dot{\forall} v.[y/u := \dot{(}u\, v\dot{)}] \mathrel{\dot{\rightarrow}} y. \tag{75}$$

We have

$$x_1, \ldots, x_\alpha, \dot{\phi} u \mathrel{\dot{\rightarrow}} y.$$

Proof. Let $d \in M^V$. Assume ${}^{\mathrm{T}}_{d}x_1, \ldots, {}^{\mathrm{T}}_{d}x_\alpha$. From $x_1, \ldots, x_\alpha \mathrel{\dot{\rightarrow}} [y/u := \mathsf{T}]$ we obtain ${}^{\mathrm{T}}_{d}[y/u := \mathsf{T}]$. Now assume $z \in \Phi \backslash \{\mathsf{T}\}$ and $\forall w \in \Phi$: ${}^{\mathrm{T}}_{d}[y/u := \dot{(}z\, w\dot{)}]$. From (75) we obtain

$$\begin{aligned} {}_d[y/u := z] &= {}_d[x_1 \mathbin{\dot{:}} \cdots \mathbin{\dot{:}} x_\alpha \mathbin{\dot{:}} \dot{\phi} u \mathbin{\dot{:}} \dot{\neg} u \mathbin{\dot{:}} \dot{\forall} v.[y/u := \dot{(}u\, v\dot{)}] \mathbin{\dot{:}} y/u := z] \\ &= {}_d[x_1 \mathbin{\dot{:}} \cdots \mathbin{\dot{:}} x_\alpha \mathbin{\dot{:}} \dot{\phi} u \mathbin{\dot{:}} \dot{\neg} u \mathbin{\dot{:}} \dot{\forall} v.[y/u := \dot{(}u\, v\dot{)}] : \dot{\mathsf{T}}/u := z] \\ &= \mathsf{T} \end{aligned}$$

Hence, ${}^{\mathrm{T}}_{d}[y/u := z]$ so

$$\forall z \in \Phi \backslash \{\mathsf{T}\}: (\forall w \in \Phi: {}^{\mathrm{T}}_{d}[y/u := \dot{(}z\, w\dot{)}] \Rightarrow {}^{\mathrm{T}}_{d}[y/u := z]).$$

Now $\forall z \in \Phi$: ${}^{\mathrm{T}}_{d}[y/u := z]$ follows from the induction lemma.

Assume ${}^{\mathrm{T}}_{d}\dot{\phi} u$. We have ${}_d u \in \Phi$ so ${}^{\mathrm{T}}_{d}y$. Hence, ${}^{\mathrm{T}}_{d}x_1, \ldots, {}^{\mathrm{T}}_{d}x_\alpha, {}^{\mathrm{T}}_{d}\dot{\phi} u \Rightarrow {}^{\mathrm{T}}_{d}y$ so the lemma follows from the deduction lemma. □

13.7. Well-foundedness axioms

By the well-foundedness theorem we have $\mathsf{T} \in \Phi$ and $\bot \notin \Phi$. Hence, from (73), we obtain the following lemmas.

Lemma 13.7.1 (Well1). $\dot{\phi} \dot{\mathsf{T}} \mathrel{\dot{=}} \dot{\mathsf{T}}$.

Lemma 13.7.2 (Well3). $\dot{\phi} \dot{\bot} \mathrel{\dot{=}} \dot{\bot}$.

We have now proved (9), i.e. we have proved the consistency of *Map°* assuming the consistency of *ZFC*.

14. The consistency of *Map*

14.1. Assuming SI

In this section, and this section only, we assume *SI*, i.e. the existence of a strongly inaccessible ordinal σ, and we assume that the transitive standard model D is defined by (11). The central consequence of this assumption is

$$x \subseteq D \wedge x <_\kappa \sigma \Rightarrow x \in D.$$

In this section, the variables α, β, γ, etc. range of σ, i.e. we tacitly assume $\alpha, \beta, \gamma \in \sigma$. We have the following absoluteness results for all $G, H, \alpha, x, y \in D$:

$$\lfloor \mathcal{P}G \rfloor = \mathcal{P}G,$$

$$\lfloor G \to H \rfloor = G \to H,$$

$$\lfloor G \to V \rfloor = G \to D \quad \text{where } V \text{ is the class of all sets,}$$

$$\lfloor G^\circ \rfloor = G^\circ,$$

$$\breve{\Phi}'(\alpha) = \hat{\Phi}'(\alpha),$$

$$\breve{\Phi}''(\alpha) = \hat{\Phi}''(\alpha),$$

$$\breve{Q}'(\alpha) = \hat{Q}'(\alpha).$$

14.2. Some properties of well-foundedness

Define $c''G = \{c(x) \mid x \in G\}$ and $c^*(f) = u \in f^{\mathrm{d}} \mapsto c(f(u))$. For $G \subseteq \acute{M}$ and $f \in \acute{M}^*$ we have $c''G \subseteq M$ and $c^*(f) \in M^*$.

Define $\acute{\nabla} G = \{x \in \acute{M} \mid \exists y \in G : y \acute{\leqslant} x\}$ and $\nabla G = \{x \in M \mid \exists y \in G : y \leqslant x\}$. We have $\nabla c''G = c''\acute{\nabla} G$ for all $G \subseteq \acute{M}$. The ∇ operation is sort of the inverse operation of the boundary operation ∂ in Section 10.4. For some $G \subseteq M$ we have $\partial \nabla G = G$ or $\nabla \partial G = G$ or both. We shall not define ∂ formally.

Define $\acute{t}_G(x) = v \in G^* \mapsto \acute{a}(x)(v)$ and $t_G(x) = v \in G^* \mapsto a(x)(v)$. From the isomorphism theorem we have $t_{c''G}(c(x))(c^*(v)) = \acute{t}_G(x)(v)$ for all $G \subseteq \acute{M}$, $x \in \acute{M}$ and $v \in G^*$. We have that $t_G(x)$ is the type of x w.r.t. G, cf. Section 10.2. Define the coordinatewise application of $\acute{t}_G$ and t_G by $\acute{t}^*_G(f) = u \in f^{\mathrm{d}} \mapsto \acute{t}_G(f(u))$ and $t^*_G(f) = u \in f^{\mathrm{d}} \mapsto t_G(f(u))$.

Like in Section 10.3 define $\acute{wf}(G) = \{f \in \acute{M} \mid \acute{t}_G(f) \in G^\circ\}$ and $wf(G) = \{f \in M \mid t_G(f) \in G^\circ\}$. We have $wf(c''G) = c''\acute{wf}(G)$ for all $G \subseteq \acute{M}$. From the monotonicity theorem we have $\acute{wf}(G) = \acute{wf}(\acute{\nabla} G)$ for all $G \subseteq \acute{M}$. Define $\acute{\Phi}'(\alpha) = \acute{\nabla} \breve{\Phi}'(\alpha)$, $\acute{\Phi}''(\alpha) = \acute{\nabla} \breve{\Phi}''(\alpha)$ and $\acute{Q}'(\alpha) = \acute{\nabla} \breve{Q}'(\alpha)$.

Lemma 14.2.1. *We have*

$$\acute{Q}'(\alpha) = \acute{wf}(\breve{\Phi}''(\alpha)) = \acute{wf}(\acute{\Phi}''(\alpha)),$$

$$\acute{\Phi}'(\alpha) = \acute{wf}(\breve{Q}'(\alpha)) = \acute{wf}(\acute{Q}'(\alpha)).$$

Proof.

$$
\begin{aligned}
x \in \acute{Q}'(\alpha) &\Leftrightarrow x \in \acute{\nabla} \breve{Q}'(\alpha) \\
&\Leftrightarrow \exists y \in \breve{Q}'(\alpha)\colon y \leqslant x \\
&\Leftrightarrow \exists y \in \breve{Q}'(\alpha)\colon y \mathrel{\acute{\doteq}}^{\Phi}_{\acute{r}} x \\
&\Leftrightarrow \exists y \in \breve{Q}'(\alpha) \forall z \in y^{\mathrm{d}}\colon y(z) = \acute{a}(x)(z) \\
&\Leftrightarrow \exists y \in \breve{Q}'(\alpha)\colon y = \acute{t}_{\breve{\Phi}''(\alpha)}(x) \\
&\Leftrightarrow \acute{t}_{\breve{\Phi}''(\alpha)}(x) \in \breve{Q}'(\alpha) \\
&\Leftrightarrow \acute{t}_{\breve{\Phi}''(\alpha)}(x) \in \breve{Q}''(\alpha)^{\circ} \\
&\Leftrightarrow x \in \acute{wf}(\breve{\Phi}''(\alpha)) \\
&\Leftrightarrow x \in \acute{wf}(\nabla \breve{\Phi}''(\alpha)) \\
&\Leftrightarrow x \in \acute{wf}(\acute{\Phi}''(\alpha)).
\end{aligned}
$$

The statement $\acute{\Phi}'(\alpha) = \acute{wf}(\breve{Q}'(\alpha)) = \acute{wf}(\acute{Q}'(\alpha))$ is proved in a similar way. □

Define $\Phi'_\partial(\alpha) = c''\breve{\Phi}'(\alpha)$, $\Phi''_\partial(\alpha) = c''\breve{\Phi}''(\alpha)$, $Q'_\partial(\alpha) = c''\breve{Q}'(\alpha)$, $\Phi'(\alpha) = c''\acute{\Phi}'(\alpha)$, $\Phi''(\alpha) = c''\acute{\Phi}''(\alpha)$ and $Q'(\alpha) = c''\acute{Q}'(\alpha)$. We have $\Phi'(\alpha) = \nabla \Phi'_\partial(\alpha)$, $\Phi''(\alpha) = \nabla \Phi''_\partial(\alpha)$ and $Q'(\alpha) = \nabla Q'_\partial(\alpha)$. Further, we have

$$\Phi'(\alpha) = wf(Q'(\alpha)) = wf(Q'_\partial(\alpha)),$$

$$Q'(\alpha) = wf(\Phi''(\alpha)) = wf(\Phi''_\partial(\alpha)),$$

$$\Phi''(\alpha) = \textstyle\bigcup_{\beta \in \alpha} \Phi'(\beta),$$

$$\Phi''_\partial(\alpha) = \textstyle\bigcup_{\beta \in \alpha} \Phi'_\partial(\beta) \quad \text{and}$$

$$\Phi = \Phi''(\sigma) = \textstyle\bigcup_{\beta \in \sigma} \Phi'(\beta).$$

From $\breve{\Phi}'(\alpha)$, $\breve{\Phi}''(\alpha)$, $\breve{Q}'(\alpha) \in D$ we obtain $\Phi'_\partial(\alpha) <_\kappa \sigma$, $\Phi''_\partial(\alpha) <_\kappa \sigma$, and $Q'_\partial(\alpha) <_\kappa \sigma$.

For all $G \subseteq H \subseteq \acute{M}$ we have $\acute{wf}(H) \subseteq \acute{wf}(G)$. Hence, for all $G \subseteq H \subseteq M$ we have $wf(H) \subseteq wf(G)$. For $\alpha \leqslant_{\mathrm{o}} \beta$ we have $\Phi''(\alpha) \subseteq \Phi''(\beta)$ which entails $Q'(\beta) \subseteq Q'(\alpha)$ and $\Phi'(\alpha) \subseteq \Phi'(\beta)$. Further, we have $\Phi'(\alpha) \subseteq Q'(\beta)$, $\Phi''(\alpha) \subseteq Q'(\beta)$ and $\Phi \subseteq Q'(\beta)$ by transfinite induction on α and β.

If $G \subseteq \acute{M}$, $x \in \acute{wf}(G)$ and $y \in G$, then $\dot{(}x\,y\dot{)} \in \acute{wf}(G)$ follows from the definition of $\acute{wf}(G)$. Hence, if $x \in \Phi'(\alpha)$ and $y \in Q'(\alpha)$ then $(x\,y) \in \Phi'(\alpha)$, and if $y \in Q'(\alpha)$ and $x \in \Phi''(\alpha)$, then $(y\,x) \in Q'(\alpha)$. Hence,

$$\alpha \leqslant_{\mathrm{o}} \beta \wedge x \in \Phi''(\alpha) \wedge y \in Q'(\beta) \Rightarrow (x\,y) \in \Phi''(\alpha),$$

$$\alpha \leqslant_{\mathrm{o}} \beta \wedge x \in \Phi''(\alpha) \wedge y \in Q'(\beta) \Rightarrow (y\,x) \in Q'(\alpha),$$

$$x \in \Phi''(\alpha) \wedge y \in \Phi''(\beta) \Rightarrow (x\,y) \in \Phi''(\alpha).$$

If $G \subseteq \acute{M}$, $G \neq \emptyset$, $x \in \acute{M}$ and $\forall y \in G\colon \dot{(}x\,y\dot{)} \in \acute{wf}(G)$, then $x \in \acute{wf}(G)$. Hence,

$$\alpha \neq 0 \wedge x \in M \wedge \forall y \in \Phi''(\alpha)\colon (x\,y) \in Q'(\alpha) \Rightarrow x \in Q'(\alpha),$$

$$x \in M \wedge \forall y \in Q'(\alpha)\colon (x\,y) \in \Phi'(\alpha) \Rightarrow x \in \Phi'(\alpha).$$

Lemma 14.2.2. *Let $\alpha \in \sigma$. If $x \in M$ and $(x\,y) \in \Phi$ for all $y \in Q'(\alpha)$, then $x \in \Phi$.*

Proof. Assume $x \in M$ and $\forall y \in Q'(\alpha)$: $(x\,y) \in \Phi$. For all $y \in Q'_\partial(\alpha)$ choose β_y such that $(x\,y) \in \Phi'(\beta_y)$ and define $\gamma = \alpha \cup \bigcup_{y \in Q'_\partial(\alpha)} \beta_y$. From $Q'_\partial(\alpha) <_\kappa \sigma$ we have $\gamma \in \sigma$. Further, we have $\forall y \in Q'_\partial(\alpha)$: $(x\,y) \in \Phi''(\gamma)$ so, by the monotonicity theorem, $\forall y \in Q'(\gamma)$: $(x\,y) \in \Phi'(\gamma)$ which entails $x \in \Phi''(\gamma) \subseteq \Phi$. $\square$

Lemma 14.2.3. *Let $\alpha \in \sigma$ and let $\mathscr{R}(x)$ be a predicate. If*

$$\mathscr{R}(\mathsf{T})$$

is true, and if

$$\forall y \in \Phi''(\alpha): \mathscr{R}(x\,y) \Rightarrow \mathscr{R}(x)$$

for all $x \in Q'(\alpha)$, $x \neq \mathsf{T}$, then

$$\mathscr{R}(x)$$

holds for all $x \in Q'(\alpha)$.

Proof. For all $x \in Q'(\alpha)$, $x \neq \mathsf{T}$ and $y \in \Phi''(\alpha)$ we have $(x\,y) \in Q'(\alpha)$, $t_{\Phi''(\alpha)}(x) \in \Phi''(\alpha)^\circ$, $t_{\Phi''(\alpha)}(x\,y) \in \Phi''(\alpha)^\circ$ and $t_{\Phi''(\alpha)}(x\,y) <_w t_{\Phi''(\alpha)}(x)$. The lemma now follows from the well-foundedness of $<_w$ on $\Phi''(\alpha)^\circ$. $\square$

14.3. Consistency proof for Map

We now verify the consistency of *Map*.

Lemma 14.3.1 (Well2). *For $\dot u \in \grave V$ and $x \in \grave M$ we have $\dot\phi\dot\lambda\dot u.x \doteq \dot\phi\dot\lambda\dot u.\dot\phi x$.*

Proof. Assume ${}^{\mathsf{T}}_{d}\dot\lambda\dot u.x$. From (73) we have ${}_d\dot\lambda\dot u.x \in \Phi$. Choose $\alpha \in \sigma$ such that ${}_d\dot\lambda\dot u.x \in \Phi''(\alpha)$. Let $u \in Q'(\alpha)$. We have ${}_d\dot(\dot\lambda\dot u.x\,u\dot) \in \Phi''(\alpha)$. Hence, ${}_d[x/\dot u := u] \in \Phi''(\alpha)$, so ${}_d\dot\phi[x/\dot u := u] = \mathsf{T} \in \Phi$ and ${}_d[\dot\phi x/\dot u := u] = \mathsf{T} \in \Phi$ which entails ${}_d\dot(\dot\lambda\dot u.\dot\phi x\,u\dot) \in \Phi$. Hence, ${}_d\dot\lambda\dot u.\dot\phi x \in \Phi$ and ${}^{\mathsf{T}}_{d}\dot\phi\dot\lambda\dot u.\dot\phi x$. This proves ${}^{\mathsf{T}}_{d}\dot\phi\dot\lambda\dot u.x \Rightarrow {}^{\mathsf{T}}_{d}\dot\phi\dot\lambda\dot u.\dot\phi x$.

We now prove ${}^{\mathsf{T}}_{d}\dot\phi\dot\lambda\dot u.\dot\phi x \Rightarrow {}^{\mathsf{T}}_{d}\dot\phi\dot\lambda\dot u.x$. Assume ${}^{\mathsf{T}}_{d}\dot\phi\dot\lambda\dot u.\dot\phi x$. We have ${}_d\dot\lambda\dot u.\dot\phi x \in \Phi$. Choose $\alpha \in \sigma$ such that ${}_d\dot\lambda\dot u.\dot\phi x \in \Phi''(\alpha)$. Let $u \in Q'(\alpha)$. We have ${}_d\dot(\dot\lambda\dot u.\dot\phi x\,u\dot) \in \Phi''(\alpha)$, so ${}_d[\dot\phi x/\dot u := u] \in \Phi''(\alpha)$. Since $\perp \notin \Phi''(\alpha)$ by the Well-foundedness Theorem we have ${}_d[\dot\phi x/\dot u := u] = {}_d\dot\phi[x/\dot u := u] \neq \perp$, so ${}_d[x/\dot u := u] \in \Phi$ by (73). Hence, ${}_d\dot(\dot\lambda\dot u.x\,u\dot) \in \Phi$ so ${}_d\dot\lambda\dot u.x \in \Phi$ and ${}^{\mathsf{T}}_{d}\dot\phi\dot\lambda\dot u.x$. This proves ${}^{\mathsf{T}}_{d}\dot\phi\dot\lambda\dot u.\dot\phi x \Rightarrow {}^{\mathsf{T}}_{d}\dot\phi\dot\lambda\dot u.x$.

We now have ${}^{\mathsf{T}}_{d}\dot\phi\dot\lambda\dot u.x \Leftrightarrow {}^{\mathsf{T}}_{d}\dot\phi\dot\lambda\dot u.\dot\phi x$. From (73) we have ${}_d\dot\phi\dot\lambda\dot u.x \in \{\mathsf{T}, \perp\}$ and ${}_d\dot\phi\dot\lambda\dot u.\dot\phi x \in \{\mathsf{T}, \perp\}$, so ${}_d\dot\phi\dot\lambda\dot u.x = {}_d\dot\phi\dot\lambda\dot u.\dot\phi x$ as stated in the lemma. $\square$

Lemma 14.3.2 (C-A). $\dot\phi\dot x$, $\dot\phi\dot y \mathrel{\dot\rightarrow} \dot\phi\dot(\dot x\,\dot y\dot)$.

Proof. Let $d \in M^V$. Assume ${}^{\mathsf{T}}_{d}\dot\phi x$ and ${}^{\mathsf{T}}_{d}\dot\phi y$. We have ${}_d x \in \Phi$ and ${}_d y \in \Phi$. Choose $\alpha, \beta \in \sigma$ such that ${}_d x \in \Phi''(\alpha)$ and ${}_d y \in \Phi''(\beta)$. We have ${}_d\dot(x\,y\dot) \in \Phi''(\alpha)$. Hence, ${}_d\dot(x\,y\dot) \in \Phi$ and ${}^{\mathsf{T}}_{d}\dot\phi\dot(x\,y\dot)$. The lemma now follows from the deduction lemma (Lemma 13.5.8). $\square$

Lemma 14.3.3 (C-K′). $\dot\phi\dot\lambda\dot u.\dot{\mathsf{T}} \doteq \dot{\mathsf{T}}$.

Proof. Let $d \in M^V$. For all $u \in Q'(0)$ we have ${}_d\dot{(}\dot\lambda\dot u.\dot{\mathsf{T}}\ u\dot{)} = {}_d\dot{\mathsf{T}} = \mathsf{T} \in \Phi$, so ${}_d\dot\lambda\dot u.\dot{\mathsf{T}} \in \Phi$ and ${}_d\dot\phi\dot u.\dot{\mathsf{T}} = \mathsf{T} = {}_d\dot{\mathsf{T}}$. $\square$

Lemma 14.3.4 (C-P′). $\dot\phi\dot\lambda\dot u.(\dot{\mathsf{if}}\ \dot u\ \dot{\mathsf{T}}\ \dot{\mathsf{T}}) \doteq \dot{\mathsf{T}}$.

Proof. Let $d \in M^V$. We have $Q'(0) = M\backslash\{\bot\}$. Hence, for all $u \in Q'(0)$ we have ${}_d\dot{(}\dot\lambda\dot u.(\dot{\mathsf{if}}\ \dot u\ \dot{\mathsf{T}}\ \dot{\mathsf{T}})\ u\dot{)} = \mathsf{T} \in \Phi$, so ${}_d\dot\lambda\dot u.(\dot{\mathsf{if}}\ \dot u\ \dot{\mathsf{T}}\ \dot{\mathsf{T}}) \in \Phi$ and ${}_d\dot\phi\dot\lambda\dot u.(\dot{\mathsf{if}}\ \dot u\ \dot{\mathsf{T}}\ \dot{\mathsf{T}}) = \mathsf{T} = {}_d\dot{\mathsf{T}}$. $\square$

Define

$$\grave P = \dot\lambda\dot x.\dot\lambda\dot y.\dot\lambda\dot z.(\dot{\mathsf{if}}\ \dot z\ \dot x\ \dot y),$$

$$\dot{Curry} = \dot\lambda\dot x.\dot\lambda\dot y.\dot\lambda\dot z.\dot{(}\dot x\ \dot{(}\grave P\ \dot y\ \dot z\dot{)}\dot{)}.$$

We have $\grave P \doteq \dot P$ and $\grave P \neq \dot P$, so $\grave P$ and $\dot P$ are semantically but not syntactically equal. The introduction of $\grave P$ is necessary since axiom C-Curry mentions $\grave P$ rather than $\dot P$.

Lemma 14.3.5 (C-Curry). *Let $x \in \acute M$. We have*

$$\dot\phi x \dot\rightarrow \dot\phi\dot{(}\dot{Curry}\ x\dot{)}.$$

Proof. Let $d \in M^V$. Assume ${}^{\mathsf{T}}_d\dot\phi x$. We have ${}_d x \in \Phi$. Choose $\alpha \in \sigma$ such that ${}_d x \in \Phi''(\alpha)$. Let $y, z \in Q'(\alpha)$. We have

$${}_d\dot{(}\dot{Curry}\ x\ y\ z\dot{)} = {}_d\dot{(}x\ \dot{(}\grave P\ y\ z\dot{)}\dot{)}.$$

Let $u \in \Phi''(\alpha)$. We have $u \neq \bot$. If $u = \mathsf{T}$ then

$${}_d\dot{(}\grave P\ y\ z\ u\dot{)} = y \in Q'(\alpha).$$

If $u \neq \mathsf{T}$ then $u \notin \{\mathsf{T}, \bot\}$ and

$${}_d\dot{(}\grave P\ y\ z\ u\dot{)} = z \in Q'(\alpha).$$

Hence, $\forall u \in \Phi''(\alpha)$: ${}_d\dot{(}\grave P\ y\ z\ u\dot{)} \in Q'(\alpha)$ so ${}_d\dot{(}\grave P\ y\ z\dot{)} \in Q'(\alpha)$ and ${}_d\dot{(}x\ \dot{(}\grave P\ y\ z\dot{)}\dot{)} \in \Phi''(\alpha)$, so ${}^{\mathsf{T}}_d\dot\phi\dot{(}\dot{Curry}\ x\dot{)}$. Now the lemma follows from the deduction lemma (Lemma 13.5.8). $\square$

Define

$$\dot\phi_1 f = \dot{(}\dot\lambda\dot f.\dot\forall\dot x.\dot\phi\dot{(}\dot f\dot x\dot{)}\ f\dot{)}$$

$$\dot{Prim} = \dot\lambda\dot f.\dot\lambda\dot x.\dot\lambda\dot y.\dot{\mathsf{Y}}\dot g.\dot\lambda\dot z.(\dot{\mathsf{if}}\ \dot z\ \dot x\ \dot{(}\dot f\ \dot\lambda\dot u.\dot{(}\dot g\ \dot{(}\dot z\ \dot{(}\dot y\ \dot u\dot{)}\dot{)}\dot{)}\dot{)}).$$

Lemma 14.3.6 (C-Prim). *We have*

$$\dot\phi_1\dot f, \dot\phi\dot x, \dot\phi\dot y \dot\rightarrow \dot\phi\dot{(}\dot{Prim}\ \dot f\dot x\dot y\dot{)}.$$

Proof. Let $d \in M^V$. Assume ${}^{\mathrm{T}}_{d}\dot{\phi}_1\dot{f}$, ${}^{\mathrm{T}}_{d}\dot{\phi}x$ and ${}^{\mathrm{T}}_{d}\dot{\phi}\dot{y}$. We have $\forall u \in \Phi$: ${}_{d}\dot{(}\dot{f}\, u\dot{)} \in \Phi$, ${}_{d}\dot{x} \in \Phi$, and ${}_{d}\dot{y} \in \Phi$. Define $h = \dot{(}\dot{Prim}\, f\, \dot{x}\, \dot{y}\dot{)}$. We have

$$ {}_{d}h = {}_{d}\dot{\lambda}\dot{z}.(\dot{\text{if}}\ \dot{z}\ \dot{x}\ \dot{(}\dot{f}\ \dot{\lambda}\dot{u}.\dot{(}h\ \dot{(}\dot{z}\ \dot{(}\dot{y}\ \dot{u}\dot{)}\dot{)}\dot{)}\dot{)}), $$

$$ {}_{d}\dot{(}h\ z\dot{)} = {}_{d}(\dot{\text{if}}\ z\ \dot{x}\ \dot{(}\dot{f}\ \dot{\lambda}\dot{u}.\dot{(}h\ \dot{(}z\ \dot{(}\dot{y}\ \dot{u}\dot{)}\dot{)}\dot{)}\dot{)}). $$

Choose α such that ${}_{d}\dot{y} \in \Phi''(\alpha)$. We now prove $\forall z \in Q'(\alpha)$: ${}_{d}\dot{(}h\ z\dot{)} \in \Phi$ by induction on z.

If $z = \mathsf{T}$ then ${}_{d}\dot{(}h\ z\dot{)} = {}_{d}x \in \Phi$. Now assume $z \in Q'(\alpha)$. $z \neq \mathsf{T}$ and $\forall v \in \Phi''(\alpha)$: ${}_{d}\dot{(}h\ \dot{(}z\ v\dot{)}\dot{)} \in \Phi$. If $u \in Q'(\alpha)$ then ${}_{d}\dot{(}\dot{y}\ u\dot{)} \in \Phi''(\alpha)$ and ${}_{d}\dot{(}h\ \dot{(}z\ \dot{(}\dot{y}\ u\dot{)}\dot{)}\dot{)} \in \Phi$. Hence, $\forall u \in Q'(\alpha)$: ${}_{d}\dot{(}\dot{\lambda}\dot{u}.\dot{(}h\ \dot{(}z\ \dot{(}\dot{y}\ \dot{u}\dot{)}\dot{)}\dot{)}\ u\dot{)} \in \Phi$, so we have ${}_{d}\dot{\lambda}\dot{u}.\dot{(}h\ \dot{(}z\ \dot{(}\dot{y}\ \dot{u}\dot{)}\dot{)}\dot{)} \in \Phi$ and ${}_{d}\dot{(}\dot{f}\ \dot{\lambda}\dot{u}.\dot{(}h\ \dot{(}z\ \dot{(}\dot{y}\ \dot{u}\dot{)}\dot{)}\dot{)}\dot{)} \in \Phi$, which entails ${}_{d}\dot{(}h\ z\dot{)} \in \Phi$.

From $\forall z \in Q'(\alpha)$: ${}_{d}\dot{(}h\ z\dot{)} \in \Phi$ we have $h \in \Phi$ as required. $\square$

Lemma 14.3.7 (C-M). *For $\dot{u}, \dot{v} \in \grave{V}$, $\dot{u} \neq \dot{v}$ and $x \in \grave{M}$ we have*

$$ \dot{\forall}\dot{u}.\dot{\phi}\dot{v}.x \mathrel{\dot{\rightarrow}} \dot{\forall}\dot{u}.\dot{\phi}\dot{v}.\dot{(}\dot{\lambda}\dot{u}.x\ \dot{(}\dot{u}\ \dot{v}\dot{)}\dot{)}, \tag{76} $$

$$ \dot{\forall}\dot{u}.\dot{\phi}\dot{v}.x \mathrel{\dot{\rightarrow}} \dot{\forall}\dot{u}.\dot{\phi}\dot{v}.\dot{(}\dot{\lambda}\dot{v}.x\ \dot{(}\dot{v}\ \dot{u}\dot{)}\dot{)}. \tag{77} $$

Proof. Assume $d \in M^V$ and ${}^{\mathrm{T}}_{d}\dot{\forall}\dot{u}.\dot{\phi}\dot{v}.x$. From Lemma 13.5.6 we have $\forall u \in \Phi$: ${}^{\mathrm{T}}_{d}[\dot{\phi}\dot{\lambda}\dot{v}.x/\dot{u} := u]$, so $\forall u \in \Phi$: ${}_{d}[\dot{\lambda}\dot{v}.x/\dot{u} := u] \in \Phi$ by (73).

Now let $\alpha \in \sigma$. For each $u \in \Phi''_{\partial}(\alpha)$ choose $\beta_u \in \sigma$ such that ${}_{d}[\dot{\lambda}\dot{v}.x/\dot{u} := u] \in \Phi''(\beta_u)$. Define $\gamma = \alpha \cup \bigcup_{u \in \Phi''_{\partial}(\alpha)} \beta_u$. Since $\Phi''_{\partial}(\alpha) \leqslant_{\kappa} \sigma$ and σ is strongly inaccessible, we have $\gamma \in \sigma$. Further, we have $\forall u \in \Phi''(\alpha)$: ${}_{d}[\dot{\lambda}\dot{v}.x/\dot{u} := u] \in \Phi''(\gamma)$, so $\forall u \in \Phi''(\alpha)\ \forall v \in Q'(\gamma)$: ${}_{d}\dot{(}[\dot{\lambda}\dot{v}.x/\dot{u} := u]v\dot{)} \in \Phi''(\gamma)$, which entails

$$ \forall u \in \Phi''(\alpha)\ \forall v \in Q'(\gamma)\colon {}_{d}[[x/\dot{v} := v]/\dot{u} := u] \in \Phi''(\gamma). $$

We obviously have $[[x/\dot{v} := v]/\dot{u} := u] = [[x/\dot{u} := u]/\dot{v} := v]$ for all $u, v \in M \subseteq \grave{M}$ since elements of M have no free variables (i.e. $\forall u \in M\ \forall \dot{u} \in \grave{V}$: $\neg \mathit{free}(\dot{u}, u)$).

Now let $u \in \Phi''(\alpha)$ and $v \in Q'(\gamma)$. We have $(u\ v) \in \Phi''(\alpha)$ and $(v\ u) \in Q'(\gamma)$. Hence,

$$ {}_{d}[[x/\dot{v} := (v\ u)]/\dot{u} := u] \in \Phi''(\gamma), \tag{78} $$

$$ {}_{d}[[x/\dot{v} := v]/\dot{u} := (u\ v)] \in \Phi''(\gamma). \tag{79} $$

We have

$$ \begin{aligned} {}_{d}[[x/\dot{v} := (v\ u)]/\dot{u} := u] \in \Phi &\Leftrightarrow {}_{d}[\dot{(}\dot{\lambda}\dot{v}.x\ (v\ u)\dot{)}/\dot{u} := u] \in \Phi \\ &\Leftrightarrow {}_{d}[[\dot{(}\dot{\lambda}\dot{v}.x\ \dot{(}\dot{v}\ \dot{u}\dot{)}\dot{)}/\dot{v} := v]/\dot{u} := u] \in \Phi \\ &\Leftrightarrow {}_{d}[\dot{(}\dot{\lambda}\dot{v}.\dot{(}\dot{\lambda}\dot{v}.x\ \dot{(}\dot{v}\ \dot{u}\dot{)}\dot{)}\ v\dot{)}/\dot{u} := u] \in \Phi. \end{aligned} $$

Since this holds for all $v \in Q'(\gamma)$ we have

$$ {}_{d}[\dot{\lambda}\dot{v}.\dot{(}\dot{\lambda}\dot{v}.x\ \dot{(}\dot{v}\ \dot{u}\dot{)}\dot{)}/\dot{u} := u] \in \Phi $$

so

$$ {}^{\mathrm{T}}_{d}[\dot{\phi}\dot{v}.\dot{(}\dot{\lambda}\dot{v}.x\ \dot{(}\dot{v}\ \dot{u}\dot{)}\dot{)}/\dot{u} := u]. $$

Since this holds for all $\alpha \in \sigma$ and $u \in \Phi''(\alpha)$ we have

$$\dot{\forall} u \in \Phi : {}^{\mathrm{T}}_{d}[\dot{\phi}\dot{v}.\dot{(}\dot{\lambda}\dot{v}.x\ \dot{(}\dot{v}\ \dot{u}\dot{)}\dot{)}/\dot{u} := u]$$

so

$$^{\mathrm{T}}_{d}\dot{\forall}\dot{u}.\dot{\phi}\dot{v}.\dot{(}\dot{\lambda}\dot{v}.x\ \dot{(}\dot{v}\ \dot{u}\dot{)}\dot{)}.$$

Now (77) follows by the deduction lemma (Lemma 13.5.8). Equation (76) follows from (79) in a similar way. □

15. The consistency of $Map^{\circ +}$

15.1. Overview

We now prove (10) which states

$$Con(ZFC^{+}) \Rightarrow Con(Map^{\circ +}).$$

Throughout this section we assume that the transitive standard model D satisfies finitely many axioms of ZFC^{+}. We shall not be explicit about which axioms D satisfies. Rather, we constantly assume that D satisfies sufficiently many axioms for the argument at hand (cf. Section 9.10).

In Section 15.2 we introduce $s \in \Phi \to D$ in such a way that

$$z \in s(x) \Leftrightarrow x \neq \mathsf{T} \wedge \exists y \in \Phi : z = s((x\, y))$$

and

$$s^{\mathrm{r}} = D.$$

In Section 15.3 we prove

$$^{\mathrm{D}}_{d}x \mathbin{\dot{\in}} y, \qquad ^{\mathrm{B}}_{d}x \mathbin{\dot{\in}} y \quad \text{and} \quad ^{\mathrm{T}}_{d}x \mathbin{\dot{\in}} y \Leftrightarrow s(x) \in s(y).$$

for all $x, y \in \Phi$ and $d \in M^{V}$. In other words we prove that $x \mathbin{\dot{\in}} y$ is $\dot{\mathsf{T}}$ if $s(x) \in s(y)$ and $\dot{\mathsf{F}}$ otherwise.

From these results we may deduce that $\dot{\in}$ inherits the properties of $\in$ in D. Further, since $s \in \Phi \to D$ and $s^{\mathrm{r}} = D$, we have $\forall x \in D : \mathscr{R}(x) \Leftrightarrow \forall y \in \Phi : \mathscr{R}(s(y))$ for all predicates $\mathscr{R}(x)$, so $\dot{\forall}$ may be used to represent the universal quantifier of ZFC^{+}.

Having made these observations, it is obvious that any statement true in D is also true in $Map^{\circ +}$. However, to prove this formally, it is necessary to do some bookkeeping. In particular, it is necessary to be cautious concerning the handling of abstraction in set and map theory.

The consistency proofs for *Map* and $Map^{\circ +}$ may be combined as follows: If we define D as in (11), then any statement true in D is also true in the model of *Map*, so *Map* may consistently be extended by any statement true in this D.

15.2. Representation of sets

Define $\hat{s}(x)=\{\hat{s}(x') \mid x' <_w x\}$. For $x \in \Phi$ define $s(x)=\{s(x') \mid x' <_A x\}$. We say that $x \in \hat{\Phi}$ represents the set $\hat{s}(x)$ and that $x \in \Phi$ represents the set $s(x)$. In this section we prove $\forall y \exists x \in \hat{\Phi}: y=\hat{s}(x)$, $\forall y \in D\ \exists x \in \Phi: y=s(x)$ and $\forall x \in \Phi: s(x) \in D$. From the definition of s and $<_A$ we immediately obtain

$$y \in s(x) \Leftrightarrow x \neq \top \wedge \exists u \in \Phi: y=s(x\, u).$$

Lemma 15.2.1. *Let $f \in H \to G$, $z \in G^* \to L$ and $z' \in H^* \to L$. If $f^{\mathrm{r}}=G$ and $z'(v)=z(f \circ v)$ for all $v \in H^*$ then $z \in G^\circ \Leftrightarrow z' \in H^\circ$ and $\hat{s}(z)=\hat{s}(z')$.*

Proof. Let f, z, and z' satisfy the assumptions of the lemma. From $f^{\mathrm{r}}=G$ we have

$$\forall v \in G: \mathscr{R}(v) \Leftrightarrow \forall v \in H: \mathscr{R}(f(v)),$$

$$\forall v \in G^*: \mathscr{R}(v) \Leftrightarrow \forall v \in H^*: \mathscr{R}(f \circ v),$$

$$\forall v \in G^\omega: \mathscr{R}(v) \Leftrightarrow \forall v \in H^\omega: \mathscr{R}(f \circ v).$$

For all predicates $\mathscr{R}(v)$. Hence, $z \in G^\circ \Leftrightarrow z' \in H^\circ$ follows trivially from the definitions of G° and H°.

In what follows, let z'' and w'' be shorthand for $v \in H^* \mapsto z(f \circ v)$ and $v \in H^* \mapsto w(f \circ v)$, respectively (note that z is free in z'' and w is free in w''). We now prove $\forall z \in G^\circ: \hat{s}(z)=\hat{s}(z'')$ by induction in z and $<_w$. If $z(\langle\rangle)=\tilde{\top}$ then $\hat{s}(z)=\emptyset=\hat{s}(z'')$. Now assume $z(\langle\rangle)=\tilde{\lambda}$ and $\forall w: (w <_w z \Rightarrow \hat{s}(w)=\hat{s}(w''))$.

Assume $x \in \hat{s}(z)$. Choose $w <_w z$ such that $x=\hat{s}(w)$. Choose $u \in G$ such that $\forall v \in G^*: w(v)=z(\langle u\rangle \cdot v)$. Choose $u' \in H$ such that $f(u')=u$. We have $\forall v \in H^*: w''(v)=w(f \circ v)=z(\langle u\rangle \cdot (f \circ v))=z''(\langle u'\rangle \cdot v)$, so $w'' <_w z''$ and $x=\hat{s}(w'') \in \hat{s}(z'')$.

Now assume $x \in \hat{s}(z'')$. Choose $w' <_w z''$ such that $x=\hat{s}(w')$. Choose $u \in H$ such that $\forall v \in H^*: w'(v)=z''(\langle u\rangle \cdot v)$. Let $u'=f(u)$ and $w=v \in G^* \mapsto z(\langle u'\rangle \cdot v)$. We have $w <_w z$ and $\forall u \in H^*: w''(v)=w(f \circ v)=z(\langle u'\rangle \cdot (f \circ v))=z''(\langle u\rangle \cdot v)=w'(v)$. Hence, $x=\hat{s}(w')=\hat{s}(w'')=\hat{s}(w) \in \hat{s}(z)$.

We may now conclude $\hat{s}(z)=\hat{s}(z'')$ for all $z \in G^\circ$. From the assumption of the lemma we have $z'=z''$, so $\hat{s}(z)=\hat{s}(z')$ holds. $\square$

Lemma 15.2.2. *Let $\breve{x} \in \breve{\Phi}$, $\acute{x} \in \acute{\Phi}$ and $x \in \Phi$. If $\breve{x} \mathrel{\acute{\leqslant}} \acute{x}$ and $x=c(\acute{x})$ then $s(x)=\hat{s}(\breve{x})$.*

Proof. Define

$$\breve{z}=y \in \breve{\Phi}^* \mapsto \breve{a}(\breve{x})(y),$$

$$\breve{z}'=y \in \breve{\Phi}^* \mapsto \acute{a}(\breve{x})(y),$$

$$\acute{z}=y \in \acute{\Phi}^* \mapsto \acute{a}(\breve{x})(y),$$

$$\acute{z}'=y \in \acute{\Phi}^* \mapsto \acute{a}(\acute{x})(y),$$

$$z=y \in \Phi^* \mapsto a(x)(y),$$

$$\breve{f}=u \in \breve{\Phi} \mapsto v \in \breve{x}^{\mathrm{dRD}} \mapsto \breve{a}(u)(v),$$

$$f=u \in \acute{\Phi} \mapsto c(u).$$

Let $\acute{f} \in \acute{\Phi} \to \breve{\Phi}$ satisfy $\acute{f}(u) \leqslant u$ for $u \in \acute{\Phi} \backslash \breve{\Phi}$ and $\acute{f}(u) = u$ for $u \in \breve{\Phi}$. Choose α such that $\breve{x} \in \breve{\Phi}'(\alpha)$. According to the relativization of Theorem 11.2.1 we have

$$\breve{f} \in \breve{\Phi} \to \breve{Q}'(\alpha), \qquad \breve{f}^{\mathrm{r}} = \breve{Q}'(\alpha).$$

Further,

$$\acute{f} \in \acute{\Phi} \to \breve{\Phi}, \qquad \acute{f}^{\mathrm{r}} = \breve{\Phi},$$

$$f \in \acute{\Phi} \to \Phi, \qquad f^{\mathrm{r}} = \Phi.$$

We now prove $\breve{z}, \breve{z}' \in \breve{\Phi}^{\circ}$, $\acute{z}, \acute{z}' \in \acute{\Phi}^{\circ}$, $z \in \Phi^{\circ}$ and $\hat{s}(\breve{x}) = \hat{s}(\breve{z}) = \hat{s}(\breve{z}') = \hat{s}(\acute{z}) = \hat{s}(\acute{z}') = \hat{s}(z) = s(x)$.

For $y \in \breve{\Phi}^{*}$ we have $\breve{z}(y) = \breve{a}(\breve{x})(y) = \breve{x}(u \in y^{\mathrm{d}} \mapsto v \in x^{\mathrm{dRD}} \mapsto \breve{a}(y(u))(v)) = \breve{x}(u \in y^{\mathrm{d}} \mapsto \breve{f}(y(u))) = \breve{x}(\breve{f} \circ y)$. Hence, $\breve{z} \in \breve{\Phi}^{\circ}$ and $\hat{s}(\breve{z}) = \hat{s}(\breve{x})$ by Lemma 15.2.1.

From the isomorphism theorem we have $\breve{z} = \breve{z}'$, so $\breve{z}' \in \breve{\Phi}^{\circ}$ and $\hat{s}(\breve{z}') = \hat{s}(\breve{z})$.

Let $y \in \acute{\Phi}^{*}$ from the definition of $\acute{f}$ we obtain $\acute{f} \circ y \in \breve{\Phi}^{*}$ and $\acute{f} \circ y \mathrel{\acute{\leqslant}^{*}} y$, so $\tilde{\bot} \neq \acute{a}(\breve{x})(\acute{f} \circ y) \leqslant_{L} \acute{a}(\breve{x})(y)$ which entails $\acute{a}(\breve{x})(\acute{f} \circ y) = \acute{a}(\breve{x})(y)$. Hence, $\breve{z}'(\acute{f} \circ y) = \acute{a}(\breve{x})(\acute{f} \circ y) = \acute{a}(\breve{x})(y) = \acute{z}(y)$ which gives $\acute{z} \in \acute{\Phi}^{\circ}$ and $\hat{s}(\acute{z}) = \hat{s}(\breve{z}')$ by Lemma 15.2.1.

Let $y \in \acute{\Phi}^{*}$. From $\breve{x} \mathrel{\acute{\leqslant}} \acute{x}$ we obtain $\tilde{\bot} \neq \acute{a}(\breve{x})(y) \leqslant_{L} \acute{a}(\acute{x})(y)$, so $\acute{z}(y) = \acute{a}(\breve{x})(y) = \acute{a}(\acute{x})(y) = \acute{z}'(y)$. Hence, $\acute{z} = \acute{z}'$, $\acute{z}' \in \acute{\Phi}^{\circ}$ and $\hat{s}(\acute{z}') = \hat{s}(\acute{z})$.

Let $y \in \acute{\Phi}^{*}$. We have $\acute{z}'(y) = \acute{a}(\acute{x})(y) = a(c(\acute{x}))(c^{*}(y)) = a(x)(f \circ y) = z(f \circ y)$. Hence, $z \in \Phi^{\circ}$ and $\hat{s}(z) = \hat{s}(\acute{z}')$ by Lemma 15.2.1.

In what follows let z and z' be shorthand for $y \in \Phi^{*} \mapsto a(x)(y)$ and $y \in \Phi^{*} \mapsto a(x')(y)$, respectively. We now prove $\forall x \in \Phi$: $s(x) = \hat{s}(z)$ by induction in x and $<_{A}$.

If $x = \mathsf{T}$ then $s(x) = \emptyset = \hat{s}(z)$. Now assume $x \neq \mathsf{T}$, $x \in \Phi$ and $\forall x' \in \Phi$: $(x' <_{A} x \Rightarrow s(x') = \hat{s}(z'))$. From $x \neq \mathsf{T}$ we have $z(\langle\,\rangle) = \tilde{\lambda}$.

Assume $w \in s(x)$. Choose $x' <_{A} x$ such that $w = s(x')$. Choose $u \in \Phi$ such that $x' = (x\,u)$. For all $v \in \Phi^{*}$ we have $z'(v) = a(x')(v) = a(x\,u)(v) = a(x)(\langle u \rangle \cdot v) = z(\langle u \rangle \cdot v)$, so $z' <_{w} z$. From $w = s(x') = \hat{s}(z')$ and $z' <_{w} z$ we conclude $w \in s(z)$. Hence, $w \in s(x) \Rightarrow w \in \hat{s}(z)$.

Now assume $w \in \hat{s}(z)$. Choose $z'' <_{w} z$ such that $w = \hat{s}(z'')$. Choose $u \in \Phi$ such that $\forall v \in F^{*}$: $z''(v) = z(\langle u \rangle \cdot v)$. Let $x' = (x\,u)$. We have $x' <_{A} x$ and $\forall v \in F^{*}$: $z'(v) = z''(v)$, so $w = \hat{s}(z'') = \hat{s}(z') = s(x') \in s(x)$. Hence, $w \in \hat{s}(z) \Rightarrow w \in s(x)$.

We may now conclude $s(x) = \hat{s}(z)$. Hence, $s(x) = \hat{s}(z) = \hat{s}(z') = \hat{s}(\acute{z}) = \hat{s}(\breve{z}') = \hat{s}(\breve{z}) = \hat{s}(\breve{x})$. $\square$

Lemma 15.2.3. $tc(H) \leqslant_{\kappa} G \Rightarrow \exists x \in G^{\circ}$: $H = \hat{s}(x)$.

Proof. By induction in H and $\in$: If $H = \emptyset$ then let $x = v \in G^{*} \mapsto \tilde{\mathsf{T}}$. If $H \neq \emptyset$ then for each $h \in H$ we have $tc(h) \leqslant_{\kappa} G$. For each $h \in H$ assume by the inductive hypothesis that $k_{h} \in G^{\circ}$ satisfies $h = \hat{s}(k_{h})$. Let $g \in G \mapsto H$ be surjective (i.e. $g^{\mathrm{r}} = H$)

(this is possible because $H \leqslant_\kappa G$). Define $x \in G^\circ$ by $x(\langle\,\rangle) = \dot{\lambda}$ and $\forall u \in G\ \forall v \in G^*\colon x(\langle u\rangle \cdot v) = k_{g(u)}(v)$. The lemma follows from $H = \hat{s}(x)$, which we may verify as follows:

$$
\begin{aligned}
\hat{s}(x) &= \{\hat{s}(x') \mid x' <_w x\} \\
&= \{\hat{s}(x') \mid \exists u \in G\ \forall v \in G^*\colon x'(v) = x(\langle u\rangle \cdot v)\} \\
&= \{\hat{s}(x') \mid \exists u \in G\ \forall v \in G^*\colon x'(v) = k_{g(u)}(v)\} \\
&= \{\hat{s}(x') \mid \exists u \in G\colon x' = k_{g(u)}\} \\
&= \{\hat{s}(k_{g(u)}) \mid u \in G\} \\
&= \{\hat{s}(k_u) \mid u \in H\} \\
&= \{u \mid u \in H\} \\
&= H. \qquad \square
\end{aligned}
$$

Let $\alpha <_0 \beta$. From Theorem 11.2.1 we have $\hat{\Phi}'(\alpha) \leqslant_\kappa \hat{Q}'(\beta)$ (this follows from $I'_{\alpha,\beta}$ in Theorem 11.2.1). Further, from the definition of $\hat{\Phi}'(\alpha)$ we have $\hat{Q}'(\alpha) <_\kappa \hat{\Phi}'(\alpha)$, so $\hat{Q}'(\alpha) <_\kappa \hat{Q}'(\beta)$. Hence, there are $\hat{Q}'(\beta)$ of arbitrarily large cardinalities. From the definition of $\hat{\Phi}$ we have $x \in \hat{\Phi} \Leftrightarrow \exists \beta\colon x \in \hat{Q}'(\beta)^\circ$. This combined with the previous lemma gives the Adequacy Theorem.

Theorem 15.2.4 (Adequacy Theorem). $\forall y \exists x \in \hat{\Phi}\colon y = \hat{s}(x)$.

As mentioned in Section 10.6, the Adequacy Theorem is central in understanding the role of $\hat{\Phi}$ in the model construction.

Corollary 15.2.5. *We have* $\forall y \in D\ \exists x \in \Phi\colon y = s(x)$ *and* $\forall x \in \Phi\colon s(x) \in D$.

Proof of the corollary. The definition of $<_w$ is stated such that $<_w$ is absolute. Hence, $\hat{s}$ is absolute, so we may conclude $\forall x \in \check{\Phi}\colon \hat{s}(x) \in D$. The relativization of the Adequacy Theorem gives $\forall y \in D\ \exists x \in \check{\Phi}\colon y = \hat{s}(x)$. Now the corollary follows from Lemma 15.2.2. $\square$

15.3. Semantics of membership

In Section 13.5 we stated lemmas about the semantics of $\dot{\mathsf{if}}$, $\dot{\mathsf{T}}$, $\dot{\mathsf{F}}$, $\dot{\neg}$, $\dot{\approx}$, $\dot{!}$, $\dot{\mathsf{i}}$, $\dot{\wedge}$, $\dot{\vee}$, $\dot{\Rightarrow}$, $\dot{\Leftrightarrow}$, $\dot{\exists}$, $\dot{\forall}$ and $\dot{\mathsf{Y}}$. We now proceed by $\doteq$ and $\dot{\in}$. Define

$$
\begin{aligned}
\mathit{eq\dot{u}al} &= \dot{\mathsf{Y}}\dot{f}.\dot{\lambda}\dot{x}.\dot{\lambda}\dot{y}.\,(\dot{\mathsf{if}}\ \dot{x}\ (\dot{\mathsf{if}}\ \dot{y}\ \dot{\mathsf{T}}\ \dot{\mathsf{F}})\ (\dot{\mathsf{if}}\ y\ \dot{\mathsf{F}} \\
&\qquad \dot{\forall}\dot{u}\dot{\exists}\dot{v}.\dot{(}\dot{f}\dot{(}\dot{x}\ \dot{u}\dot{)}\ \dot{(}\dot{y}\ \dot{v}\dot{)}\dot{)}\ \dot{\wedge}\ \dot{\forall}\dot{v}\dot{\exists}\dot{u}.\dot{(}\dot{f}\ \dot{(}\dot{x}\ \dot{u}\dot{)}\ \dot{(}\dot{y}\ \dot{v}\dot{)}\dot{)}))), \\
x \doteq y &= \dot{(}\mathit{eq\dot{u}al}\ x\ y\dot{)}, \\
\mathit{bel\dot{o}ng} &= \dot{\lambda}\dot{x}.\dot{\lambda}\dot{y}.(\dot{\mathsf{if}}\ \dot{y}\ \dot{\mathsf{F}}\ \dot{\exists}\dot{u}.\dot{x} \doteq \dot{(}\dot{y}\ \dot{u}\dot{)}), \\
x \mathbin{\dot{\in}} y &= \dot{(}\mathit{bel\dot{o}ng}\ x\ y\dot{)}.
\end{aligned}
$$

In what follows, let $d \in M^V$ and $x, y \in \Phi$. Since $x \neq \bot$ and $y \neq \bot$ we have ${}^{\mathrm{D}}_{d}x$ and ${}^{\mathrm{D}}_{d}y$. From the semantics of $\dot{\mathsf{Y}}$ we have

$$\begin{aligned} {}_{d}x \doteq y \;=\; {}_{d}(\dot{\mathsf{if}}\; x\; (\dot{\mathsf{if}}\; y\; \dot{\mathsf{T}}\; \dot{\mathsf{F}})\; (\dot{\mathsf{if}}\; y\; \dot{\mathsf{F}} \\ \dot{\forall}\dot{u}\dot{\exists}\dot{v}.\dot{(}x\,\dot{u}\dot{)} \doteq \dot{(}y\,\dot{v}\dot{)} \mathbin{\dot{\wedge}} \dot{\forall}\dot{v}\dot{\exists}\dot{u}.\dot{(}x\,\dot{u}\dot{)} \doteq \dot{(}y\,\dot{v}\dot{)})). \end{aligned}$$

From the semantics of $\dot{\mathsf{if}}, \dot{\mathsf{T}}, \dot{\mathsf{F}}$ and $\dot{\wedge}$ we have

$${}^{\mathrm{B}}_{d}x \doteq y.$$

Now assume $\forall u,\, z \in \Phi\colon {}^{D}_{d}\dot{(}x\,u\dot{)} \doteq z$. From the semantics of $\dot{\mathsf{if}}$, $\dot{\mathsf{T}}$, $\dot{\mathsf{F}}$, $\dot{\wedge}$, $\dot{\forall}$ and $\dot{\exists}$ we conclude $\dot{\forall} y \in \Phi\colon {}^{\mathrm{D}}_{d}x \doteq y$. Hence, by the Induction Theorem (Corollary 11.10.4) we have

$${}^{\mathrm{D}}_{d}x \doteq y.$$

From the previous section we obtain

$$z \in s(x) \;\Leftrightarrow\; x \neq \mathsf{T} \wedge \exists u \in \Phi\colon z = s(x\,u).$$

Hence

$$\begin{aligned} s(x) = s(y) \;\Leftrightarrow\; & x = \mathsf{T} \wedge y = \mathsf{T} \vee x \neq \mathsf{T} \wedge y \neq \mathsf{T} \\ & \wedge \forall u \in \Phi\; \exists v \in \Phi\colon s(x\,u) = s(y\,v) \\ & \wedge \forall v \in \Phi\; \exists v \in \Phi\colon s(x\,u) = s(y\,v). \end{aligned}$$

From the semantics of $\dot{\mathsf{if}}, \dot{\mathsf{T}}, \dot{\mathsf{F}}, \dot{\wedge}, \dot{\forall}$ and $\dot{\exists}$ we have

$$\begin{aligned} {}^{\mathrm{T}}_{d}x \doteq y \;\Leftrightarrow\; & x = \mathsf{T} \wedge y = \mathsf{T} \vee x \neq \mathsf{T} \wedge y \neq \mathsf{T} \\ & \wedge \forall u \in \Phi\; \exists v \in \Phi\colon {}^{\mathrm{T}}_{d}\dot{(}x\,u\dot{)} \doteq \dot{(}y\,v\dot{)} \wedge \forall v \in \Phi\; \exists u \in \Phi\colon {}^{\mathrm{T}}_{d}\dot{(}x\,u\dot{)} \doteq \dot{(}y\,v\dot{)}. \end{aligned}$$

Hence, by the Induction Theorem we have

$${}^{\mathrm{T}}_{d}x \doteq y \;\Leftrightarrow\; s(x) = s(y).$$

From the semantics of $\dot{if}$, $\dot{\mathsf{F}}$, $\dot{\exists}$ and $\doteq$ we obtain

$${}^{\mathrm{B}}_{d}x \mathbin{\dot{\in}} y, \qquad {}^{\mathrm{D}}_{d}x \mathbin{\dot{\in}} y \quad \text{and}$$

$${}^{\mathrm{T}}_{d}x \mathbin{\dot{\in}} y \;\Leftrightarrow\; y \neq \mathsf{T} \wedge \exists u \in \Phi\colon s(x) = s(y\,u) \;\Leftrightarrow\; s(x) \in s(y).$$

15.4. Terms of ZFC and their values

In analogy to the definition of $\grave{M}$ define the set Z of well-formed formulas of *ZFC* and related concepts as follows: Let Z be the least set such that

$$\forall u,\, v \in \grave{V}\colon u \mathbin{\ddot{\in}} v \in Z,$$

$$\forall x \in Z\colon \ddot{\neg} x \in Z,$$

$$\forall x,\, y \in Z\colon x \mathbin{\ddot{\Rightarrow}} y \in Z,$$

$$\forall u \in \grave{V}\; \forall x \in Z\colon \ddot{\forall} u\colon x \in Z.$$

For u, v, $w \in \dot{V}$ and $x, y \in Z$ define

$$free(u, v \ddot{\in} w) \Leftrightarrow u = v \vee u = w,$$
$$free(u, \ddot{\neg} x) \Leftrightarrow free(u, x),$$
$$free(u, x \ddot{\Rightarrow} y) \Leftrightarrow free(u, x) \vee free(u, y),$$
$$free(u, \ddot{\forall} v\colon x) \Leftrightarrow u \neq v \wedge free(u, x).$$

Define $D^V = \dot{V} \rightarrow D$. For all $x \in Z$ and $f \in D^V$ define the interpretation ${}_f x$ as follows:

$${}_f u \ddot{\in} v \Leftrightarrow f(u) \in f(v),$$
$${}_f \ddot{\neg} x \Leftrightarrow \neg {}_f x,$$
$${}_f x \ddot{\Rightarrow} y \Leftrightarrow {}_f x \Rightarrow {}_f y,$$
$${}_f \ddot{\forall} u\colon x \Leftrightarrow \forall g \in D^V\colon (\forall v \in \dot{V} \setminus \{u\}\colon f(v) = g(v) \Rightarrow {}_g x).$$

For all $x \in Z$ define the translation $\downarrow x$ of x into map theory as follows:

$$\downarrow(u \ddot{\in} v) = u \dot{\in} v,$$
$$\downarrow \ddot{\neg} x = \dot{\neg} \downarrow x,$$
$$\downarrow(x \ddot{\Rightarrow} y = (\downarrow x) \dot{\Rightarrow} (\downarrow y),$$
$$\downarrow \ddot{\forall} u\colon x = \dot{\forall} u. \downarrow x.$$

Define $F^V = \dot{V} \rightarrow \Phi$, let $d \in F^V$, and let $f = s \circ d$. We have $f \in D^V$. From the semantics of $\dot{\in}$, $\dot{\neg}$, $\dot{\Rightarrow}$ and $\dot{\forall}$, and by structural induction in $x \in Z$ we obtain

$${}^{\mathrm{B}}_{d}\!\downarrow x, \qquad {}^{\mathrm{D}}_{d}\!\downarrow x \quad \text{and} \quad {}^{\mathrm{T}}_{d}\!\downarrow x \Leftrightarrow {}_f x.$$

15.5. Consistency proof

As stated previously, $Map^{\circ +}$ is Map° extended by the translation of any theorem of ZFC^+. Now that we have defined $\downarrow x$, we may define $Map^{\circ +}$ more formally: $Map^{\circ +}$ contains all axioms and inference rules of Map°. Further, if x is a theorem of ZFC^+ whose free variables occur among $u_1 \ldots u_\alpha$, then we include

$$\dot{\phi} u_1, \ldots, \dot{\phi} u_\alpha \dot{\rightarrow} \downarrow x$$

as an axiom of $Map^{\circ +}$. $Map^{\circ +}$ is axiomatic since its axioms and hence its theorems are recursively enumerable. To ensure definedness in $Map^{\circ +}$ of all well-formed formulas of ZFC, we also include

$$\dot{\phi} u_1, \ldots, \dot{\phi} u_\alpha \dot{\rightarrow} \dot{!} \downarrow x$$

for any well-formed formula (theorem or not) of ZFC.

Theorem 15.5.1 (Semantic Adequacy). *Let $x \in Z$. Assume that the free variables of x occur among $u_1, \ldots, u_\alpha \in \dot{V}$. We have*

$$\dot{\phi} u_1, \ldots, \dot{\phi} u_\alpha \dot{\rightarrow} \dot{!} \downarrow x,$$
$$\dot{\phi} u_1, \ldots, \dot{\phi} u_\alpha \dot{\rightarrow} \downarrow x \Leftrightarrow \forall f \in D^V\colon {}_f x.$$

Proof. Let $d \in M^V$. Assume ${}^{\mathsf{T}}_{d}\dot{\phi} u_1, \ldots, {}^{\mathsf{T}}_{d}\dot{\phi} u_\alpha$. We have ${}_{d}u_1, \ldots, {}_{d}u_\alpha \in \Phi$. Define $e \in M^V$ by

$$e(v) = \begin{cases} d(v) & \text{if } v \in \{u_1, \ldots, u_\alpha\}, \\ \mathsf{T} & \text{otherwise.} \end{cases}$$

We have $e \in F^V$ and ${}_{d}{\downarrow}x = {}_{e}{\downarrow}x$.

Since $e \in F^V$ we have ${}^{\mathsf{D}}_{e}{\downarrow}x$ which entails ${}^{\mathsf{D}}_{d}{\downarrow}x$ and ${}^{\mathsf{T}}_{d}!{\downarrow}x$. Hence, $\dot{\phi} u_1, \ldots, \dot{\phi} u_\alpha \dot{\rightarrow} !{\downarrow}x$ follows from the deduction lemma.

Now assume $\forall f \in D^V: {}_{f}x$. Let $f = s \circ e$. We have $f \in D^V$, so ${}_{f}x$ holds. Further, ${}_{f}x \Leftrightarrow {}^{\mathsf{T}}_{e}{\downarrow}x \Leftrightarrow {}^{\mathsf{T}}_{d}{\downarrow}x$, so $\dot{\phi} u_1, \ldots, \dot{\phi} u_\alpha \dot{\rightarrow} {\downarrow}x$ follows from the deduction lemma. Hence, $\forall f \in D^V: {}_{f}x \Rightarrow \dot{\phi} u_1, \ldots, \dot{\phi} u_\alpha \dot{\rightarrow} {\downarrow}x$.

Now assume $\dot{\phi} u_1, \ldots, \dot{\phi} u_\alpha \dot{\rightarrow} {\downarrow}x$ and $f \in D^V$. Choose $e \in F^V$ such that $f = s \circ e$ (this is possible since $s \in \Phi \rightarrow D$ is surjective, i.e. $s^{\mathrm{r}} = D$). We have ${}^{\mathsf{T}}_{e}\dot{\phi} u_1, \ldots, {}^{\mathsf{T}}_{e}\dot{\phi} u_\alpha$, so ${}^{\mathsf{T}}_{e}{\downarrow}x$ follows from $\dot{\phi} u_1, \ldots, \dot{\phi} u_\alpha \dot{\rightarrow} {\downarrow}x$. Since ${}^{\mathsf{T}}_{e}{\downarrow}x \Leftrightarrow {}_{f}x$ we have verified that $\dot{\phi} u_1, \ldots, \dot{\phi} u_\alpha \dot{\rightarrow} {\downarrow}x \Rightarrow \forall f \in D^V: {}_{f}x$. $\square$

We have now proved that the transformation of any *ZFC* statement which is true in *D*, holds in *M*. Taking *D* to satisfy (a finite set of axioms of) ZFC^+ we obtain the consistency of $Map^{\circ +}$ as stated in (10).

Taking *D* to be defined as in (11) we see that we may also consistently extend *Map* by the translation of any statement of *ZFC* true in this *D*.

16. Conclusion

16.1. Summary of results

Part II documented the expressive power of map theory by developing set theory in it, and Part III has verified the consistency. Furthermore, the notion of truth has been defined in Section 3 and we have presented a few other constructions that are beyond the capability of set theory. In conclusion, map theory is an alternative foundation to set theory.

16.2. Further work

Further work is needed to improve the axiomatization of map theory. In particular, it is unsatisfactory that well-foundedness is expressed by ten axion schemes and one inference rule, none of which explain the intuition behind well-foundedness.

It is also not satisfactory that no axiom expresses the monotonicity of maps. Define

$$x \sqcup y = (\text{if } x\ (\text{if } y\ \mathsf{T}\ \bot)\ (\text{if } y\ \bot\ \lambda u.(x\,u) \sqcup (y\,u))).$$

We have $x \leq y$ iff $x = x \sqcup y$, so one formulation of monotonicity could be $x \leq y \vdash (f x) \leq (f y)$ where $x \leq y$ is shorthand for $x = x \sqcup y$.

The Quantify4 and Quantify5 axioms ought to be superfluous once monotonicity is stated as an axiom.

16.3. Equality is better than truth

The author strongly prefers to express theories using equality rather than truth as the basic concept, i.e. without using logic, logical connectives and quantifiers at the level of axioms. One reason is that the basic laws of equality such as

$$x = y \,;\, x = z \;\vdash\; y = z$$

are simpler and more appealing than the basic laws of logic such as

$$(\mathcal{A} \Rightarrow (\mathcal{B} \Rightarrow \mathcal{C})) \Rightarrow ((\mathcal{A} \Rightarrow \mathcal{B}) \Rightarrow (\mathcal{A} \Rightarrow \mathcal{C})).$$

Another reason is that theories based on equality immediately suggest to form term models and thereby give a better understanding of the nature of Skolem's paradox [22]. A third reason is that using equality does not favor one particular kind of logic, e.g., classical or intuitionistic. One could say that there are many kinds of logic but only one kind of equality (unfortunately, however, λ-calculus traditionally messes up equality by calling various relations "equality" even though they are really just equivalence relations on syntactic domains). A fourth reason is that use of equality as the basic concept eases the definition of "the notion of truth" which reduces to a self-interpreter for theories based on equality. The formalization of map theory using equality as the basic concept demonstrates the expressive power of this approach.

Index to Part III

In the index, references in parentheses refer to informal definitions.

Note that in Part III accents are used differently from Parts I and II.

Relations involving stars are omitted from the index. See Section 9.6 for an explanation.

Example of use: The construct $x \hat{=}^{a}_{G} y$ has parameters x, y and G. The construct itself consists of an equal sign, an accent, and the letter a. In the index, the construct could be located under $=$ or under a. It is located under $=$ because that character comes before a in the index.

The reader can obtain a combined index/glossary directly from the author. The index/glossary extends the index below with a short explanation of each construct.

References

[1] J-R. Abrial, The specification language Z: Syntax and "semantics". Technical report, Programming Research Group, Oxford, April 1980 (out of print).

[2] P. Aczel, Non-well-founded sets, CSLI Lecture Notes, Vol. 14, CSLI, 333 Ravenswood Avenue, Menlo Park, CA 94025, USA, 1988.

[3] H.P. Barendregt, *The Lambda Calculus, Its Syntax and Semantics*, Studies in Logic and The Foundation of Mathematics, Vol. 103 (North-Holland, Amsterdam, 1984).

[4] M. Barr and C. Wells, *Category Theory for Computing Science*, Series in Computer Science (Prentice-Hall, Englewood Cliffs, NJ, 1990).

[5] J. Barwise, ed., *Handbook of Mathematical Logic*, Studies in Logic and The Foundation of Mathematics, Vol. 90 (North-Holland, Amsterdam, 1977).

[6] M.J. Beeson, Towards a computational system based on set theory, *Theoret. Comput. Sci.* **60** (1988) 297–340.

[7] R. Bird and P. Wadler, *Introduction to Functional Programming* (Prentice-Hall, Englewood Cliffs, NJ, 1988).

[8] D. Bjørner and C.B. Jones, *The Vienna Development Method: The Meta-Language*, Lecture Notes in Computer Science, Vol. 61 (Springer, Berlin, 1978).

[9] P.J. Cohen, *Set Theory and the Continuum Hypothesis* (Benjamin, New York, 1966).

[10] U. Felgner, Choice functions on sets and classes, in: *Sets and Classes: On the works by Paul Bernays* (North-Holland, Amsterdam, 1976) 217–255.

[11] Board for Software Standardisation and Control, Esa software engineering standards, PSS-05-0, issue 2, Technical report, European Space Agency, Noordwijk, February 1991.

[12] K. Gödel, Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I, *Monatsh für Math. u Phys.* **12**(XXXVIII) (1931) 173–198.

[13] W.S. Hatcher, *Foundations of Mathematics* (W.B. Saunders, 1968).

[14] P. Henderson, *Functional Programming* (Prentice-Hall, Englewood Cliffs, NJ, 1980).

[15] D. Hilbert and P. Bernays, *Grundlagen der Mathematik, Vol. 2* (Springer, Berlin, 1939).

[16] S.C. Kleene, *Introduction to Metamathematics*, Bibliotheca Mathematica, Vol. 1 (North-Holland, Amsterdam, 1964).

[17] K. Kunen, *Set Theory, An Introduction to Independence Proofs*, Studies in Logic and The Foundation of Mathematics, Vol. 102 (North-Holland, Amsterdam 1980).

[18] F.W. Lawvere, The category of categories as a foundation for mathematics, in: *Proc. Conf. on Categorical Algebra, La Jolla 1965* (Springer, Berlin, 1966) 1–21.

[19] R.C. Lyndon, *Notes on Logic*, Mathematical Studies, Vol. 6 (Van Nostrand, New York, 1966).

[20] P. Martin-Löf, An intuitionistic theory of types: Predicative part, in: *Logic Colloquium '73* Studies in Logic and The Foundation of Mathematics, Vol. 80 (North-Holland, Amsterdam, 1975) 74–118.

[21] J. McCarthy, A basis for a mathematical theory of computation, in: P. Brafford and D. Hirschberg, eds., *Computer Programming and Formal Systems* (North-Holland, Amsterdam 1963) 33–70.

[22] E. Mendelson, *Introduction to Mathematical Logic* (Wadsworth and Brooks, Belmont, CA, 3rd edn., 1987).

[23] D. Scott, *λ-Calculus and Computer Science Theory*, Lecture Notes in Computer Science, Vol. 37, Chapter "Combinators and classes" (Springer, Berlin. 1975) 1–26.

[24] D. Scott, *λ-Calculus and Computer Science Theory*, Lecture Notes in Computer Science, Vol. 37, Chapter "Some philosophical issues concerning theories of combinators" (Springer, Berlin 1975) 346–366.

[25] J.R. Schoenfield, Axioms of set theory, in: J. Barwise, ed, *Handbook of Mathematical Logic* (North-Holland, Amsterdam, 1977) 321–344.

[26] D.A. Turner, A new implementation technique for applicative languages, *Software—Practice and Experience*, **9** (1979) 31–49.