EXAMENSARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

Paradoxes in Intuitionistic Type Theory

av

Susanne Thon

2007 - No 20

Paradoxes in Intuitionistic Type Theory

Susanne Thon

Examensarbete i matematik 30 högskolepoäng, fördjupningskurs

Handledare: Jesper Carlström

2007

Paradoxes in Intuitionistic Type Theory

Susanne Thon

Abstract

One of the most famous paradoxes in mathematics is Russell's paradox. It considers the set of all sets that do not contain themselves. Does this set contain itself or not? Whatever answer we give, we get a contradiction.
Paradoxes in mathematical systems indicate inconsistency. An early version of Intuitionistic type theory developed by Martin-Löf turned out to be inconsistent and had to be revised. In this paper we will present formalizations of Girard's paradox and Coquand's paradox of trees, which can be seen as versions of Russell's paradox, in the inconsistent version of Intuitionistic type theory.

Contents

1 Introduction

Imagine a village barber who shaves those and only those men in the village who do not shave themselves. Does the barber shave himself or not?

Whatever answer we give, it leads to a contradiction.

This paradox can be seen as an everyday version of the famous paradox that was discovered by Bertrand Russell in 1902 and shook the logician's world. Cantor's set theory, that served as a basis of all mathematics, did not hold anymore, neither did Frege's just completed work "Foundations of Arithmetic".[1]

Several attempts were made to solve paradoxes like the one of Russell. One approach was taken by Russell himself by introducing a theory of types.

A type theory that is based on intuitionistic ideas, the Intuitionistic type theory, was developed in the 1970's by Martin-Löf. An early version of this type theory however turned out to be inconsistent; paradoxes could be formalized in it.

In this paper we will present formalizations of two paradoxes that can be seen as versions of Russell's paradox in the inconsistent version of Martin-Löf's type theory: Girard's paradox and Coquand's paradox of trees.
We will start by giving an introduction to these paradoxes and the different attempts to solve them in chapter 2. In chapter 3 we then present type theory in the version of Martin-Löf. A formalization of Girard's paradox is presented in chapter 4, Coquand's paradox of trees in chapter 5.

[1]van Heijenoort 1967, p. 124

2 Paradoxes

The encyclopedia of mathematics defines a paradox as "a situation in which two mutual contradictory statements are demonstrated, each one having been deduced by means that are convincing from the point of view of the same theory."[2]

What does it mean for a system or a theory to contain a paradox?

Because it is known from classical logic that from a contradiction anything can be deduced, proofs of theorems in a system that contains a paradox are not convincing. A paradox indicates that one or several basic assumptions made in the theory in question are wrong and often leads to a fundamental revision of the system as a whole.[3]

In the latter part of the nineteenth century Cantor introduced set theory, which soon came to be seen as a fundamental branch of mathematics. The work of Dedekind, Weierstrass, and others showed that from the arithmetic of natural numbers much of classical mathematics can be derived. The concept of natural numbers in its turn can be obtained from the basic principles of set theory.[4] The following is a modern way to do this:

0 is defined to be the empty set, 1 is defined to be the set that contains the empty set as its only element, 2 is defined to be the set that contains the empty set and the set that contains the empty set and so on. Using these definitions, we can prove propositions of natural numbers.[5]

Set theory however turned out to contain a number of contradictions, which were discovered between 1895 and 1910. The Burali-Forti paradox, which was discovered in 1897, is one of them. It deals with the existence of the greatest ordinal number.
An ordinal number is a ranking number; it designates the rank or position of an object in a linearly ordered array.[6] Ordinal numbers are represented by well-ordered sets. Thus the above concept of natural numbers describes ordinal numbers.
We consider an ordinal number as being the set of all preceding ordinals. For each ordinal a we can construct a strictly greater one by forming the set that contains all preceding ordinals of a together with a itself. The paradox is derived if we consider the greatest ordinal number, that is, the set of all ordinal numbers. Such a set would be an ordinal number greater than or equal to every ordinal number. But for every ordinal number there exists a strictly greater one and we get a contradiction.

At first it was not paid much attention to the paradoxes because it was hoped that they could be corrected by some alternation of the basic definitions. But when in 1902 Russell's paradox involved elementary aspects of set theory they could no longer be ignored.
Russell's paradox can be seen as a version of the Burali-Forti paradox and can be described as follows: We can consider two types of sets: sets that contain themselves and sets that do not contain themselves. Most sets we use do not contain themselves, but the set of all sets must be a member of itself.
We now define M as the set of all sets that do not contain themselves. Does M contain itself? Suppose M is not a member of itself, then it satisfies the condition to be a member of M and hence contains itself. If we suppose that M is a member of itself it does not satisfy the condition to be a member of M and hence does not contain itself. Thus M contains itself if and only if it does not contain itself, which is a contradiction of the most fundamental sort.[7]

When these paradoxes were discovered, the conception of a set as represented in Cantor's definition did no longer provide a satisfactory basis neither for set theory nor for mathematics as a whole.[8] New approaches were needed.

One approach was made by Zermelo in 1908.[9] Zermelo's system, modified by Skolem and Fraenkel, is widely used still today.
Zermelo conceived a restriction that is called the axiom of selection. It states that a property of objects can only be used to select objects from a set that already exists, but not to form a new set. This means

[2] Encyclopedia of mathematics 1993, p. 205
[3] ibid. p. 205
[4] Pinter 1971, p. 2
[5] ibid. p. 2
[6] ibid. p. 166
[7] ibid. p. 3
[8] ibid. p. 3
[9] ibid. pp. 10–11

that if A is a set and $S(x)$ is a property, we can form a set that consists of exactly those elements x in A that satisfy $S(x)$. But we cannot form a set that consists of all x that satisfy $S(x)$. Thus Russell's construction of a paradox cannot be carried out.

Another approach was made by von Neumann.
Von Neumann noted that the paradoxes are caused by two facts:

1. The crucial sets, like the set of all sets that do not contain themselves, are too large, they include too much.

2. These large sets are allowed to be elements of sets, as for example elements of themselves.

In Zermelo's approach the known constructions of paradoxes are blocked by prohibiting to form these large sets. Von Neumann's approach permits the existence of large sets but does not allow them to be elements of sets.[10]
Both Zermelo's and von Neumann's systems reached their aim of providing a firm foundation for a system of set theory which include Cantor's basic results. But as they had to give up a considerable amount of naïve set theory, many mathematicians pointed out that they should be seen as provisional solutions.[11]

In the early 1900's Russell developed a theory of types from the idea of considering sets to be ordered in a hierarchy of levels, which means that if we have two sets A and B and A is an element of B then B is one level higher than A. Every level represents a type or a range of significance. The theory can be described as follows:[12]
Every set has a level, i.e. a natural number. The simplest sets are of level 0, they are called individuals and do not have elements. A set of level 1 contains such individuals, a set of level 2 contains sets of level 1 and so on. An expression like $a \in A$ is only meaningful if a is of level n and A is of level $n+1$ for some $n \in N$.
Thus a statement like $x \in x$ is not meaningful in the theory of types and as a result, Russell's paradox cannot even be formulated.

Another quite different approach was made by the intuitionists. In intuitionistic mathematics every proof must be constructive, which means that to prove that a mathematical object exists a method for actually constructing the object must be given.
Cantor assumed that if we can name a property of objects, there exists a set of all objects with this property. In intuitionistic set theory a set exists only if we can describe how to build it. Therefore a set of all sets cannot be formed.[13]

Based on the principles of constructive mathematics, Marin-Löf developed intuitionistic type theory, in the 1970's. In his theory every mathematical object has a certain type and is always given together with this type. This can be seen as a both simpler and more general version of Russell's formulation.[14]
An early version of type theory was, however, inconsistent. This was discovered by Girard, who formulated a type theoretic version of the Burali-Forti paradox in it. As a consequence, the theory was modified several times.
We will present Girard's paradox in chapter 4.
Some years later, Coquand formalized a less extensive version of the paradox in the inconsistent version of Martin-Löf's type theory. Coquand's paradox of tree will be presented in chapter 5.
We will start with a presentation of Martin-Löf's type theory in the following chapter.

[10] Pinter 1971, p. 13
[11] ibid. p. 15
[12] ibid. pp. 16–17
[13] ibid. pp. 17–19
[14] Martin-Löf 1998, p. 127

3 Intuitionistic type theory

Martin-Löf developed Intuitionistic type theory (from now on type theory) with the aim to provide a constructive foundation of mathematics as an alternative to the usual foundation of mathematics based on classical set theory. Type theory is not based on first order predicate logic; predicate logic is interpreted within type theory.[15]
A type is defined by prescribing how to construct an object of that type. This corresponds to the definition of a set, and from now on we will refer to types as sets. In an analogous way, a proposition is defined by prescribing how it can be proved. This leads us to one of the basic ideas in type theory: the interpretation of propositions as sets.
A proposition is proved to be true by constructing an element of the corresponding set. An empty set corresponds to a false proposition.
If we decide whether a proposition is true, we make a judgement.
We can make judgements of four forms with the following notation in type theory:

(1)	A is a set	$A : \text{set}$
(2)	A and B are equal sets	$A = B : \text{set}$
(3)	a is an element in the set A	$a : A$
(4)	a and b are equal elements in the set A	$a = b : A$

These judgements are explained in the following way:

(1) We know that A is a set if we know how a canonical element of A is formed and under what conditions two canonical elements of A are equal.
(2) Two sets A and B are equal if an arbitrary canonical element of A is also a canonical element of B and vice versa.
(3) If A is a set then a is an element in A if a yields a canonical element in A as value when evaluated.
(4) Two arbitrary elements a and b in a set A are equal if a and b yield equal canonical elements of A as values when they are evaluated.

Identifying sets with propositions and elements of sets with proofs of propositions, $A : \text{set}$ is interpreted as A *is a proposition* and $a : A$ is interpreted as a *is a proof of* A i.e. A *is true*.

Families of sets

In general judgements are made under assumptions.
An expression of the form

$$B(x_1, ..., x_n) : \text{set} \ \ [x_1 : A_1, ..., x_n : A_n(x_1, ..., x_n)]$$

is called a hypothetical judgement or a family of sets with n indices.
Judgements with an empty context, that is, judgements without assumptions, are called categorical judgements.
We will take a closer look at what the four forms of judgements mean if they are made under one assumption.

(1) $B(x) : \text{set}\ [x : A]$ expresses that $B(x)$ is a set under the assumption $x : A$, or that $B(x)$ is a family of sets over A. A judgement of this form means that for an arbitrary element a in A, $B(a)$ is a set and if a and b are equal elements in A, then $B(a)$ and $B(b)$ are equal sets.
(2) $B(x) = C(x)\ [x : A]$ expresses that $B(x)$ and $C(x)$ are equal families of sets over the set A. A judgement of this form means that for any arbitrary element $a : A$, $B(a)$ and $C(a)$ are equal sets.
(3) $b(x) : B(x)\ [x : A]$ expresses that $b(x)$ is an element of $B(x)$ under the assumption $x : A$. A judgement of this form means that for an arbitrary element $a : A$, $b(a)$ is an element in $B(a)$, and if a and d are equal elements in A, then $b(a)$ and $b(d)$ are equal elements in $B(a)$.
(4) $b(x) = c(x) : B(x)\ [x : A]$ expresses that $b(x)$ and $c(x)$ are equal elements of $B(x)$ under the assumption $x : A$. A judgement of this form means that, for an arbitrary element $a : A$, $b(a)$ and $c(a)$ are equal elements in the set $B(a)$.

Judgements made under n assumptions are generalizations of judgements made under one assumption. The meanings of these judgements follow by induction.

[15]In the following presentation of type theory we mainly follow Martin-Löf 1984 and the part about polymorphic sets in Nordström, Petersson, Smith 1990

Equality rules

From the meanings of judgements of the form $A = B$, $a : A$ and $a = b : A$ the rules for set equality follow immediately.

Set equality

$$\frac{a : A \quad A = B : \text{set}}{a : B} \qquad \frac{a = b : A \quad A = B : \text{set}}{a = b : B}$$

We have the following general equality rules, which, by definition, hold for canonical elements:

Reflexivity

$$\frac{a : A}{a = a : A} \qquad \frac{A : \text{set}}{A = A : \text{set}}$$

Symmetry

$$\frac{a = b : A}{b = a : A} \qquad \frac{A = B : \text{set}}{B = A : \text{set}}$$

Transitivity

$$\frac{a = b : A \quad b = c : A}{a = c : A} \qquad \frac{A = B : \text{set} \quad B = C : \text{set}}{A = C : \text{set}}$$

Substitution

From the verification of the meanings of judgements for families of sets the substitution rules follow immediately:

Substitution in sets

$$\frac{a : A \quad C(x) : \text{set} \ [x : A]}{C(a) : \text{set}} \qquad \frac{a = b : A \quad C(x) : \text{set} \ [x : A]}{C(a) = C(b) : \text{set}}$$

Substitution in elements

$$\frac{a : A \quad c(x) : C(x) \ [x : A]}{c(a) : C(a)} \qquad \frac{a = b : A \quad c(x) : C(x) \ [x : A]}{c(a) = c(b) : C(a)}$$

Substitution in equal sets

$$\frac{a : A \quad B(x) = C(x) : \text{set} \ [x : A]}{B(a) = C(a) : \text{set}}$$

Substitution in equal elements

$$\frac{a : A \quad b(x) = c(x) : B(x) \ [x : A]}{b(a) = c(a) : B(a)}$$

We will now give the rules for the different set-forming operations we use.
For each operation there are four kinds of rules:

- The *set formation rule* describes under which conditions we can form a set A
- The *introduction rule* defines the set A by describing how the canonical elements of the set are formed
- The *elimination rule* shows how to define functions on the set
- The *equality rule* shows how a function defined by the elimination rule operates on canonical elements of the set

To each rule of set formation, introduction and elimination there are corresponding rules that allow us to substitute equal elements for equal elements and equal sets for equal sets. These rules are left out here since they are of immediate inference.

Each element of a set can be evaluated to canonical form. The value of a canonical element is the canonical element itself, whereas a non-canonical element has a canonical element as its value.

The rules are presented in a natural deduction style

$$\frac{P_1 \quad P_2 \quad ... \quad P_n}{C}$$

where P_1, P_2, ..., P_n are the premises and C the conclusion. In general they are all hypothetical judgements.

We then relate to each set the corresponding proposition it is identified with.

Cartesian Product of a family of sets

If A ist a set and B is a family of sets over A, then $(\Pi x : A)B(x)$ denotes the cartesian product of the sets in the family. The corresponding proposition is the universal quantifier; $(\Pi x : A)B(x)$ is identified with $(\forall x : A)B(x)$.
The canonical elements in $(\Pi x : A)B(x)$ are of the form $(\lambda x)b(x)$, where $b(x) : B(x)$ under the assumption $x : A$. When applied to an element $a : A$, they give an element in the set $B(a)$.

$$\Pi\text{-formation:} \quad \frac{A : \text{set} \quad \begin{array}{c}[x : A] \\ B(x) : \text{set}\end{array}}{(\Pi x : A)B(x) : \text{set}} \qquad \forall\text{-formation:} \quad \frac{A : \text{prop} \quad \begin{array}{c}[x : A] \\ B(x) : \text{prop}\end{array}}{(\forall x : A)B(x) : \text{prop}}$$

$$\Pi\text{-introduction:} \quad \frac{\begin{array}{c}[x : A] \\ b(x) : B(x)\end{array}}{(\lambda x)b(x) : (\Pi x : A)B(x)} \qquad \forall\text{-introduction:} \quad \frac{\begin{array}{c}[x : A] \\ B(x)\ \text{true}\end{array}}{(\forall x : A)B(x)\ \text{true}}$$

$$\Pi\text{-elimination:} \quad \frac{f : (\Pi x : A)B(x) \quad a : A}{\text{ap}(f, a) : B(a)} \qquad \forall\text{-elimination:} \quad \frac{(\forall x : A)B(x)\ \text{true} \quad a : A}{B(a)\ \text{true}}$$

We presuppose the premises $A : \text{set}$ and $B(x) : \text{set}\ [x : A]$, which are not written out explicitly.

$\text{ap}(f, a)$ is evaluated as follows:
First evaluate f.
If the value of f is $(\lambda x)b(x)$, then the value of $\text{ap}(f, a)$ is the value of $b(a)$.
We then get the rule

$$\Pi\text{-equality:} \quad \frac{\begin{array}{c}[x : A] \\ b(x) : B(x)\end{array} \quad a : A}{\text{ap}((\lambda x)b(x), a) = b(a) : B(a)}$$

The function set

If B does not depend on x, $A \to B \equiv (\Pi x : A)B$ is the set of functions from A to B. The corresponding proposition is implication; $A \to B$ is identified with $A \supset B$.
The canonical elements of $A \to B$ are of the form $(\lambda x)b(x)$, where $b(x) : B$ under the assumption $x : A$ and x must not occur free in B. When applied to an element $a : A$, they give an element in the set B.
We get, as special cases

$$\to\text{-formation:} \quad \frac{A : \text{set} \quad \begin{array}{c}[x : A] \\ B : \text{set}\end{array}}{A \to B : \text{set}} \qquad \supset\text{-formation:} \quad \frac{A : \text{prop} \quad \begin{array}{c}[A\ \text{true}] \\ B : \text{prop}\end{array}}{A \supset B : \text{prop}}$$

where x must not occur free in B

$$\to\text{-introduction:} \quad \frac{\begin{array}{c}[x : A] \\ b(x) : B\end{array}}{(\lambda x)b(x) : A \to B} \qquad \supset\text{-introduction:} \quad \frac{\begin{array}{c}[A\ \text{true}] \\ B\ \text{true}\end{array}}{A \supset B\ \text{true}}$$

where x must not occur free in B

$$\to\text{-elimination:} \quad \frac{f : A \to B \quad a : A}{\text{ap}(f, a) : B} \qquad \supset\text{-elimination:} \quad \frac{A \supset B\ \text{true} \quad A\ \text{true}}{B\ \text{true}}$$

$$\to\text{-equality:} \quad \frac{\begin{array}{c}[x : A] \\ b(x) : B\end{array} \quad a : A}{\text{ap}((\lambda x)b(x), a) = b(a) : B}$$

where x must not occur free in B or f

Disjoint union of a family of sets

If A is a set and $B(x)$ is a family of sets over A, then $(\Sigma x : A)B(x)$ denotes the disjoint union of the sets in the family $B(x)$. The corresponding proposition is the existential quantifier; $(\Sigma x : A)B(x)$ is identified with $(\exists x : A)B(x)$.
The canonical elements in $(\Sigma x : A)B(x)$ are pairs of the form $\langle a, b\rangle$, where $a : A$ and $b : B(a)$.

$$\Sigma\text{-formation:}\quad \frac{A : \mathsf{set} \quad \begin{array}{c}[x : A]\\ B(x) : \mathsf{set}\end{array}}{(\Sigma x : A)B(x) : \mathsf{set}} \qquad\qquad \exists\text{-formation:}\quad \frac{A : \mathsf{prop} \quad \begin{array}{c}[x : A]\\ B(x) : \mathsf{prop}\end{array}}{(\exists x : A)B(x) : \mathsf{prop}}$$

$$\Sigma\text{-introduction:}\quad \frac{a : A \quad b : B(a)}{\langle a, b\rangle : (\Sigma x : A)B(x)} \qquad\qquad \exists\text{-introduction:}\quad \frac{a : A \quad B(a)\ \mathsf{true}}{(\exists x : A)B(x)\ \mathsf{true}}$$

$$\Sigma\text{-elimination:}\quad \frac{c : (\Sigma x : A)B(x) \quad \begin{array}{c}[v : (\Sigma x : A)B(x)]\\ C(v) : \mathsf{set}\end{array} \quad \begin{array}{c}[x : A, \quad y : B(x)]\\ d(x, y) : C(\langle x, y\rangle)\end{array}}{E(c, (x, y)d(x, y)) : C(c)} \qquad \exists\text{-elimination:}\quad \frac{(\exists x : A)B(x)\ \mathsf{true} \quad \begin{array}{c}[x : A, \quad B(x)\ \mathsf{true}]\\ C\ \mathsf{true}\end{array}}{C\ \mathsf{true}}$$

$E(c, (x, y)d(x, y))$ is evaluated in the following way:
First evaluate c.
If the value of c is $\langle a, b\rangle$, then the value of $E(c, (x, y)d(x, y))$ is the value of $d(a, b)$, which is a canonical element of $C(\langle a, b\rangle)$.
This gives the rule

$$\Sigma\text{-equality:}\quad \frac{a : A \quad b : B(a) \quad \begin{array}{c}[v : (\Sigma x : A)B(x)]\\ C(v) : \mathsf{set}\end{array} \quad \begin{array}{c}[x : A, \quad y : B(x)]\\ d(x, y) : C(\langle x, y\rangle)\end{array}}{E(\langle a, b\rangle, (x, y)d(x, y)) = d(a, b) : C(\langle a, b\rangle)}$$

The projections p and q, which select the left and the right component of a pair, respectively, are special cases of $E(c, (x, y)d(x, y))$:

$$p(c) \equiv E(c, (x, y)x) : A$$
$$q(c) \equiv E(c, (x, y)y) : B(p(c))$$

We then have the elimination rules

$$\frac{c : (\Sigma x : A)B(x)}{p(c) : A} \qquad \frac{c : (\Sigma x : A)B(x)}{q(c) : B(p(c))}$$

and the equality rules

$$\frac{a : A \quad b : B(a)}{p(\langle a, b\rangle) = a : A} \qquad \frac{a : A \quad b : B(a)}{q(\langle a, b\rangle) = b : B(a)}$$

Cartesian product of two sets

If B does not depend on x, $A \times B \equiv (\Sigma x : A)B$ denotes the cartesian product of A and B. The corresponding proposition is conjunction; $A \times B$ is identified with $A \,\&\, B$.
The canonical elements in $A \times B$ are thus pairs of the form $\langle a, b\rangle$, where $a : A$ and $b : B$.
We get, as special cases

$$\times\text{-formation:}\quad \frac{A : \mathsf{set} \quad B : \mathsf{set}}{A \times B : \mathsf{set}} \qquad\qquad \&\text{-formation:}\quad \frac{A : \mathsf{prop} \quad B : \mathsf{prop}}{A \,\&\, B : \mathsf{prop}}$$

$$\times\text{-introduction:}\quad \frac{a : A \quad b : B}{\langle a, b\rangle : A \times B} \qquad\qquad \&\text{-introduction:}\quad \frac{A\ \mathsf{true} \quad B\ \mathsf{true}}{A \,\&\, B\ \mathsf{true}}$$

$$\times\text{-elimination:}\quad \frac{c : A \times B \quad \begin{array}{c}[v : A \times B]\\ C(v) : \mathsf{set}\end{array} \quad \begin{array}{c}[x : A, \quad y : B]\\ d(x, y) : C(\langle x, y\rangle)\end{array}}{E(c, (x, y)d(x, y)) : C(c)} \qquad \&\text{-elimination:}\quad \frac{A \,\&\, B\ \mathsf{true} \quad \begin{array}{c}[A\ \mathsf{true}, \quad B\ \mathsf{true}]\\ C\ \mathsf{true}\end{array}}{C\ \mathsf{true}}$$

$$\times\text{-equality:}\quad \frac{a : A \quad b : B \quad \begin{array}{c}[v : A \times B]\\ C(v) : \mathsf{set}\end{array} \quad \begin{array}{c}[x : A, \quad y : B]\\ d(x, y) : C(\langle x, y\rangle)\end{array}}{E(\langle a, b\rangle, (x, y)d(x, y)) = d(a, b) : C(\langle a, b\rangle)}$$

Disjoint union of two sets

If A is a set and B is a set, then $A+B$ denotes the disjoint union of A and B. The corresponding proposition is disjunction; $A+B$ is identified with $A \vee B$.
The canonical elements in $A+B$ are of the form $i(a)$ and $j(b)$, where i and j give the information about which of the sets A or B the elements come from.

+-formation: $$\frac{A : \text{set} \quad B : \text{set}}{A+B : \text{set}}$$

$\vee$-formation: $$\frac{A : \text{prop} \quad B : \text{prop}}{A \vee B : \text{prop}}$$

+-introduction 1: $$\frac{a : A}{i(a) : A+B}$$

$\vee$-introduction 1: $$\frac{A \ \text{true}}{A \vee B \ \text{true}}$$

+-introduction 2: $$\frac{b : B}{j(b) : A+B}$$

$\vee$-introduction 2: $$\frac{B \ \text{true}}{A \vee B \ \text{true}}$$

+-elimination: $$\frac{c : A+B \quad \begin{matrix}[v : A+B] \\ C(v) : \text{set}\end{matrix} \quad \begin{matrix}[x : A] \\ d(x) : C(i(x))\end{matrix} \quad \begin{matrix}[y : B] \\ e(y) : C(j(y))\end{matrix}}{D(c, (x)d(x), (y)e(y)) : C(c)}$$

$\vee$-elimination: $$\frac{A \vee B \ \text{true} \quad \begin{matrix}[A \ \text{true}] \\ C \ \text{true}\end{matrix} \quad \begin{matrix}[B \ \text{true}] \\ C \ \text{true}\end{matrix}}{C \ \text{true}}$$

$D(c, (x)d(x), (y)e(y))$ is evaluated as follows:
First evaluate c.
If the value of c is of the form $i(a)$, then the value of $D(c, (x)d(x), (y)e(y))$ is the value of $d(a)$.
If the value of c is of the form $j(b)$, then the value of $D(c, (x)d(x), (y)e(y))$ is the value of $e(b)$.
This gives the rules

+-equality 1: $$\frac{a : A \quad \begin{matrix}[v : A+B] \\ C(v) : \text{set}\end{matrix} \quad \begin{matrix}[x : A] \\ d(x) : C(i(x))\end{matrix} \quad \begin{matrix}[y : B] \\ e(y) : C(j(y))\end{matrix}}{D(i(a), (x)d(x), (y)e(y)) = d(a) : C(i(a))}$$

+-equality 2: $$\frac{b : B \quad \begin{matrix}[v : A+B] \\ C(v) : \text{set}\end{matrix} \quad \begin{matrix}[x : A] \\ d(x) : C(i(x))\end{matrix} \quad \begin{matrix}[y : B] \\ e(y) : C(j(y))\end{matrix}}{D(j(b), (x)d(x), (y)e(y)) = e(b) : C(j(b))}$$

Propositional equality

If A is a set, then the proposition $\text{Id}(A, a, b)$ is similar to the judgement $a = b : A$.
The canonical elements of $\text{Id}(A, a, b)$ have the form $\text{ref}(a)$.
$\text{Id}(A, a, b)$ can be seen as the smallest reflexive relation on A, as its only elements are reflexivity proofs.

Id-formation: $$\frac{A : \text{set} \quad a : A \quad b : A}{\text{Id}(A, a, b) : \text{set}}$$

Id-introduction: $$\frac{a : A}{\text{ref}(a) : \text{Id}(A, a, a)}$$

Id-elimination: $$\frac{a : A \quad b : A \quad c : \text{Id}(A, a, b) \quad \begin{matrix}[x : A \quad y : A \quad z : \text{Id}(A, x, y)] \\ C(x, y, z) : \text{set}\end{matrix} \quad \begin{matrix}[x : A] \\ d(x) : C(x, x, \text{ref}(x))\end{matrix}}{\text{idpeel}(c, (x)d(x)) : C(a, b, c)}$$

$\text{idpeel}(c, (x)d(x))$ is evaluated as follows:
First evaluate c.
If c has the value $\text{ref}(a)$, then the value of $\text{idpeel}(c, (x)d(x))$ is the value of $d(a)$.
This gives the rule

Id-equality: $$\frac{a : A \quad \begin{matrix}[x : A \quad y : A \quad z : \text{Id}(A, x, y)] \\ C(x, y, z) : \text{set}\end{matrix} \quad \begin{matrix}[x : A] \\ d(x) : C(x, x, \text{ref}(x))\end{matrix}}{\text{idpeel}(\text{ref}(a), (x)d(x)) = d(a) : C(a, a, \text{ref}(a))}$$

Enumeration sets

Given n canonical constants $i_1, ..., i_n$ we introduce the enumeration set $\{i_1, ..., i_n\}$.
The canonical elements of $\{i_1, ..., i_n\}$ are i_1, i_2, ... and i_n.

$\{i_1, ..., i_n\}$-formation: $\dfrac{}{\{i_1, ..., i_n\} : \text{set}}$

We have n introduction rules ($n \geq 0$):

$\{i_1, ..., i_n\}$-introduction 1: $\dfrac{}{i_1 : \{i_1, ..., i_n\}}$

$\vdots$

$\{i_1, ..., i_n\}$-introduction n: $\dfrac{}{i_n : \{i_1, ..., i_n\}}$

$$\{i_1, ..., i_n\}\text{-elimination:} \quad \frac{a : \{i_1, ..., i_n\} \qquad \begin{array}{c}[x : \{i_1, ..., i_n\}] \\ C(x) : \text{set}\end{array} \qquad b_1 : C(i_1), \ ... \ , b_n : C(i_n)}{\text{case}(a, b_1, ..., b_n) : C(a)}$$

$\text{case}(a, b_1, ..., b_n)$ is evaluated in the following way:
First evaluate a.
If the value of a is i_k $(1 \leq k \leq n)$ then the value of $\text{case}(a, b_1, ..., b_n)$ is the value of b_k.
This gives the rule

$$\{i_1, ..., i_n\}\text{-equality:} \quad \frac{\begin{array}{c}[x : \{i_1, ..., i_n\}] \\ C(x) : \text{set}\end{array} \qquad b_1 : C(i_1), ..., b_n : C(i_n)}{\text{case}(i_k, b_1, ..., b_n) = b_k : C(i_k)}$$

The empty set

For $n = 0$ we get the empty set $\{\}$ as a special case. The corresponding proposition is absurdity; $\{\}$ is identified with $\perp$.
The empty set has no introduction rule as it has no elements.

$\{\}$-formation: $\dfrac{}{\{\} : \text{set}}$ $\qquad$ $\perp$-formation: $\dfrac{}{\perp : \text{prop}}$

$$\{\}\text{-elimination:} \quad \frac{a : \{\} \quad \begin{array}{c}[x : \{\}] \\ C(x) : \text{set}\end{array}}{\text{case}(a) : C(a)} \qquad \perp\text{-elimination:} \quad \frac{\perp \ \text{true} \quad C : \text{prop}}{C \ \text{true}}$$

For any proposition A we define a negation by $\neg A \equiv A \to \perp$

Natural numbers

N denotes the set of natural numbers.
0 is a canonical element in N and if a is an element in N, then $s(a)$ is a canonical element in N.

N-formation: $\dfrac{}{N : \text{set}}$

N-introduction 1: $\dfrac{}{0 : N}$

N-introduction 2: $\dfrac{a : N}{s(a) : N}$

$$N\text{-elimination:}\quad \frac{c : N \quad \begin{array}{c}[v : N]\\ C(v) : \text{set}\end{array} \quad d : C(0) \quad \begin{array}{c}[x : N,\ y : C(x)]\\ e(x,y) : C(s(x))\end{array}}{R(c, d, (x,y)e(x,y)) : C(c)}$$

$R(c, d, (x,y)e(x,y))$ is evaluated in the following way:
First evaluate c.
If the value of c is 0 then the value of $R(c, d, (x,y)e(x,y))$ is the value of d.
If the value of c is $s(a)$ for some $a : N$ then the value of $R(c, d, (x,y)e(x,y))$ is the value of $e(a, R(a, d, (x,y)e(x,y)))$.
This gives the rules

$$N\text{-equality 1:}\quad \frac{\begin{array}{c}[v : N]\\ C(v) : \text{set}\end{array} \quad d : C(0) \quad \begin{array}{c}[x : N,\ y : C(x)]\\ e(x,y) : C(s(x))\end{array}}{R(0, d, (x,y)e(x,y)) = d : C(0)}$$

$$N\text{-equality 2:}\quad \frac{a : N \quad \begin{array}{c}[v : N]\\ C(v) : \text{set}\end{array} \quad d : C(0) \quad \begin{array}{c}[x : N,\ y : C(x)]\\ e(x,y) : C(s(x))\end{array}}{R(s(a), d, (x,y)e(x,y)) = e(a, R(a, d, (x,y)e(x,y))) : C(s(a))}$$

Well-orderings

If A is a set and if B is a family of sets over A, then $(W\ x : A)B(x)$ is a well-ordering, or a well-founded tree. The constructor set A represents the different ways to form a tree and the selector family $B(x)$ represents the parts of a tree formed by $x : A$.
The canonical elements in $(W\ x : A)B(x)$ are of the form $\sup(a, (t)b(t))$, where $a : A$ is a particular form we want the tree to have and b is a function from $B(a)$ to $(W\ x : A)B(x)$, i.e. a collection of subtrees.

$$W\text{-formation:}\quad \frac{A : \text{set} \quad \begin{array}{c}[x : A]\\ B(x) : \text{set}\end{array}}{(W\ x : A)B(x) : \text{set}}$$

$$W\text{-introduction:}\quad \frac{a : A \quad \begin{array}{c}[t : B(a)]\\ b(t) : (W\ x : A)B(x)\end{array}}{\sup(a, (t)b(t)) : (W\ x : A)B(x)}$$

$$W\text{-elimination:}\quad \frac{a : (W\ x : A)B(x) \quad \begin{array}{c}[v : (W\ x : A)B(x)]\\ C(v) : \text{set}\end{array} \quad \begin{array}{c}[y : A,\ z(t) : (W\ x : A)B(x)[t : B(y)],\ u(t) : C(z(t))[t : B(y)]]\\ b(y,z,u) : C(\sup(y, (t)z(t)))\end{array}}{\text{wrec}(a, (y,z,u)b(y,z,u)) : C(a)}$$

$\text{wrec}(a, (y,z,u)b(y,z,u))$ is evaluated as follows:
First evaluate a.
If a has the value $\sup(d, (t)e(t))$ then the value of $\text{wrec}(a, (y,z,u)b(y,z,u))$ is the value of $b(d, (t)e(t), (t)\,\text{wrec}(e(t), (y,z,u)b(y,z,u)))$.
This gives the rule

$$W\text{-equality:}\quad \frac{d : A \quad \begin{array}{c}[t : B(d)]\\ e(t) : (W\ x : A)B(x)\end{array} \quad \begin{array}{c}[v : (W\ x : A)B(x)]\\ C(v) : \text{set}\end{array} \quad \begin{array}{c}[y : A,\ z(t) : (W\ x : A)B(x)[t : B(y)],\ u(t) : C(z(t))[t : B(y)]]\\ b(y,z,u) : C(\sup(y, (t)z(t)))\end{array}}{\text{wrec}(\sup(d, (t)e(t)), (y,z,u)b(y,z,u)) = b(d, (t)e(t), (t)\,\text{wrec}(e(t), (y,z,u)b(y,z,u))) : C(\sup(d, (t)e(t)))}$$

Universes

We define a universe as the least set closed under the representatives of certain set-forming operators. The first universe, or the set of small sets, is the least set that is closed under the operations we have introduced so far. These are Π, Σ, $+$, Id, $\{i_1, ..., i_n\}$, N and W.
The canonical elements of the first universe are codes for the sets that are formed by the set-forming operators.
We further define a familiy of sets $T(x)$: set $[x : U]$, which decodes the elements in the universe to the sets they represent.

U-formation: $$\frac{}{U : \text{set}}$$

T-elimination: $$\frac{a : U}{T(a) : \text{set}}$$

U-introduction π: $$\frac{a : U \quad \begin{array}{c}[x : T(a)]\\ b(x) : U\end{array}}{\pi(a, (x)b(x)) : U}$$

T-equality π: $$\frac{a : U \quad \begin{array}{c}[x : T(a)]\\ b(x) : U\end{array}}{T(\pi(a, (x)b(x))) = (\Pi x : T(a))T(b(x)) : \text{set}}$$

U-introduction σ: $$\frac{a : U \quad \begin{array}{c}[x : T(a)]\\ b(x) : U\end{array}}{\sigma(a, (x)b(x)) : U}$$

T-equality σ: $$\frac{a : U \quad \begin{array}{c}[x : T(a)]\\ b(x) : U\end{array}}{T(\sigma(a, (x)b(x))) = (\Sigma x : T(a))T(b(x)) : \text{set}}$$

U-introduction $\hat{+}$: $$\frac{a : U \quad b : U}{a \hat{+} b : U}$$

T-equality $\hat{+}$: $$\frac{a : U \quad b : U}{T(a \hat{+} b) = T(a) + T(b) : \text{set}}$$

U-introduction $\hat{\text{Id}}$: $$\frac{a : U \quad x : T(a) \quad y : T(a)}{\hat{\text{Id}}(a, x, y) : U}$$

T-equality $\hat{\text{Id}}$: $$\frac{a : U \quad x : T(a) \quad y : T(a)}{T(\hat{\text{Id}}(a, x, y)) = \text{Id}(T(a), x, y) : \text{set}}$$

U-introduction $\widehat{\{i_1, ..., i_n\}}$: $$\frac{}{\widehat{\{i_1, ..., i_n\}} : U}$$

T-equality $\widehat{\{i_1, ..., i_n\}}$: $$\frac{}{T(\widehat{\{i_1, ..., i_n\}}) = \{i_1, ..., i_n\} : \text{set}}$$

U-introduction $\hat{N}$: $$\frac{}{\hat{N} : U}$$

T-equality $\hat{N}$: $$\frac{}{T(\hat{N}) = N : \text{set}}$$

U-introduction w: $$\frac{a : U \quad \begin{array}{c}[x : T(a)]\\ b(x) : U\end{array}}{w(a, (x)b(x)) : U}$$

T-equality w: $$\frac{a : U \quad \begin{array}{c}[x : T(a)]\\ b(x) : U\end{array}}{T(w(a, (x)b(x))) = (W\ x : T(a))T(b(x)) : \text{set}}$$

A first version of Martin-Löf's type theory contained the additional rules

$$\frac{}{u : U} \qquad \frac{}{T(u) = U : \text{set}}$$

If we add these rules, the system becomes inconsistent, as we will show in the following chapters: In chapter 4 we will formalize the paradox that was discovered by Girard and in chapter 5 we will explain Coquand's paradox of trees.

4 Girard's paradox

Recall the Burali-Forti paradox. We consider an ordinal number as the well-ordered set that contains all preceding ordinal numbers. To each ordinal number a strictly greater one can be constructed. We then form the set of all ordinal numbers. This set is an ordinal number that is greater than all ordinal numbers. But to each ordinal number a strictly greater one can be constructed. This gives the contradiction.

Girard's paradox may be considered as a type-theoretic version of this argument.[16] Ordinal numbers are replaced by well-ordered sets. An ordered set is a set M with an order relation $<$ which is transitive, i.e. $x < y$ and $y < z$ implies $x < z$ for all $x, y, z : M$. An ordered set is well-ordered if there is no infinitely descending chain $x_1 > x_2 > ... > x_n > ...$ of elements in M.[17]

We will use the notation A^* to denote a well-ordered set, where A^* consists of the name $a : U$ of a set $T(a)$ that has a binary relation $<$ on it, a proof of transitivity of $<$ and a proof of $T(a)$ having no infinitely descending chains. The greatest ordinal number, that is, the set of all ordinal numbers, is represented by Ω. We will show that such a set can indeed be defined using $u : U$ with $T(u) = U$. The empty set will be denoted by $\perp$.

The paradox is constructed as follows:

Let $a : U$ and suppose there is a binary relation $<$ on $T(a)$ such that $<: T(a) \to T(a) \to U$. Applying this relation to any two elements $x : T(a)$ and $y : T(a)$, we get the name $ap(ap(<, x), y)$ in U and define the corresponding set as $x < y \equiv T(ap(ap(<, x), y))$.

We formulate the properties

$$P(a, <) = (\forall x, y, z : T(a))[x < y \to (y < z \to x < z)]$$

and

$$Q(a, <) = (\forall f : N \to T(a))\neg[(\forall n : N)(\mathrm{ap}(f, s(n)) < \mathrm{ap}(f, n))]$$

$P(a, <)$ expresses transitivity of $<$, $Q(a, <)$ expresses that there are no infinitely descending chains in $T(a)$.

We then form the set

$$\Omega = (\exists a : U, <: T(a) \to T(a) \to U)[P(a, <) \,\&\, Q(a, <)]$$

as the set of all sets that have a name a in U and a relation $<$ with the properties $P(a, <)$ and $Q(a, <)$.

For $A^* : \Omega$ we define $a \equiv p(A^*)$ as the name of a set $T(a)$, $<_a \equiv p(q(A^*))$ as the binary relation on $T(a)$ and $p_a \equiv p(q(q(A^*)))$ and $q_a \equiv q(q(q(A^*)))$ as proofs of the properties $P(a, <_a)$ and $Q(a, <_a)$.

We define a binary relation $<_\omega: \Omega \to \Omega \to U$ such that applying this relation to any two elements $A^* : \Omega$ and $B^* : \Omega$ gives the name of the set

$$A^* <_\omega B^* \equiv (\exists f : T(a) \to T(b), z : T(b))[(\forall x, y : T(a))(x <_a y \to \mathrm{ap}(f, x) <_b \mathrm{ap}(f, y)) \& (\forall x : T(a))(\mathrm{ap}(f, x) <_b z)]$$

which expresses that there exists an order-preserving bounded function $f : T(a) \to T(b)$.

We will show that Ω has a name ω in U and the properties $P(\omega, <_\omega)$ and $Q(\omega, <_\omega)$. Here we essentially need the assumptions $u : U$, $T(u) = U$. Then we can show that $\Omega^* \equiv \langle \omega, \langle <_\omega, \langle p_\omega, q_\omega \rangle\rangle\rangle$ is itself an element of Ω.

Further we can show that $A^* <_\omega \Omega^*$ for any $A^* : \Omega$. As $\Omega^* : \Omega$, we get $\Omega^* <_\omega \Omega^*$. But this gives an infinitely descending chain in Ω, which yields a contradiction.

---

[16] We will formalize the idea that is sketched in Troelstra, v.Dalen 1988, vol II, p. 614 and Martin-Löf 1998 pp. 133–135
[17] v. Mangoldt/Knopp 1973, pp. 53–54

We start by showing in U^1 and U^2 that $P(a,<)$ and $Q(a,<)$ are sets with names in U for $a : U$ and $<: T(a) \to T(a) \to U$. In U^3 we show that Ω is a set with a name in U under the assumption $u : U$ with $T(u) = U$. In U^4 we define the binary relation $<_\omega$ on Ω.

We then construct a proof of absurdity.

We will sometimes carry out several steps at the same time, for example when we use repeated substitution.
The sign $\equiv$ will be used for definitions as well as for abbreviations in the trees.

We define $A \equiv T(a)$.

U^1

$$
\dfrac{a : U \qquad \dfrac{a : U \qquad \dfrac{a : U \qquad \dfrac{\dfrac{\dfrac{<\,: A \to A \to U \quad [x : A]^3}{\mathrm{ap}(<, x) : A \to U} \quad [y : A]^2}{\mathrm{ap}(\mathrm{ap}(<, x), y) : U} \qquad \dfrac{\dfrac{\dfrac{<\,: A \to A \to U \quad [y : A]^2}{\mathrm{ap}(<, y) : A \to U} \quad [z : A]^1}{\mathrm{ap}(\mathrm{ap}(<, y), z) : U} \qquad \dfrac{\dfrac{<\,: A \to A \to U \quad [x : A]^3}{\mathrm{ap}(<, x) : A \to U} \quad [z : A]^1}{\mathrm{ap}(\mathrm{ap}(<, x), z) : U}}{\pi(\mathrm{ap}(\mathrm{ap}(<, y), z), (q)\, \mathrm{ap}(\mathrm{ap}(<, x), z)) : U}}{\pi(\mathrm{ap}(\mathrm{ap}(<, x), y), (p)\pi(\mathrm{ap}(\mathrm{ap}(<, y), z), (q)\, \mathrm{ap}(\mathrm{ap}(<, x), z))) : U}\;(1)}{\pi(a, (z)\pi(\mathrm{ap}(\mathrm{ap}(<, x), y), (p)\pi(\mathrm{ap}(\mathrm{ap}(<, y), z), (q)\, \mathrm{ap}(\mathrm{ap}(<, x), z)))) : U}\;(2)}{\pi(a, (y)\pi(a, (z)\pi(\mathrm{ap}(\mathrm{ap}(<, x), y), (p)\pi(\mathrm{ap}(\mathrm{ap}(<, y), z), (q)\, \mathrm{ap}(\mathrm{ap}(<, x), z))))) : U}\;(3)}{\hat{P}(a, <) \equiv \pi(a, (x)\pi(a, (y)\pi(a, (z)\pi(\mathrm{ap}(\mathrm{ap}(<, x), y), (p)\pi(\mathrm{ap}(\mathrm{ap}(<, y), z), (q)\, \mathrm{ap}(\mathrm{ap}(<, x), z)))))) : U}
$$

We have $P(a, <) = T(\hat{P}(a, <)) = T(\pi(a, (x)\pi(a, (y)\pi(a, (z)\pi(\mathrm{ap}(\mathrm{ap}(<, x), y), (p)\pi(\mathrm{ap}(\mathrm{ap}(<, y), z), (q)\, \mathrm{ap}(\mathrm{ap}(<, x), z)))))))$
$= (\Pi x : A)(\Pi y : A)(\Pi z : A)[x < y \to (y < z \to x < z)] : \mathrm{set}$

U^2

$$
\dfrac{\dfrac{\dfrac{}{\hat{N} : U} \quad a : U}{\pi(\hat{N}, (q)a) : U} \qquad \dfrac{\dfrac{\dfrac{}{\hat{N} : U} \qquad \dfrac{\dfrac{<\,: A \to A \to U \quad \dfrac{[f : N \to A]^5 \quad \dfrac{[n : N]^4}{s(n) : N}}{\mathrm{ap}(f, s(n)) : A}}{\mathrm{ap}(<, \mathrm{ap}(f, s(n))) : A \to U} \qquad \dfrac{[f : N \to A]^5 \quad [n : N]^4}{\mathrm{ap}(f, n) : A}}{\mathrm{ap}(\mathrm{ap}(<, \mathrm{ap}(f, s(n))), \mathrm{ap}(f, n)) : U}}{\pi(\hat{N}, (n)\, \mathrm{ap}(\mathrm{ap}(<, \mathrm{ap}(f, s(n))), \mathrm{ap}(f, n))) : U}\;(4) \qquad \dfrac{}{\hat{\bot} : U}}{\pi(\pi(\hat{N}, (n)\, \mathrm{ap}(\mathrm{ap}(<, \mathrm{ap}(f, s(n))), \mathrm{ap}(f, n))), (p)\hat{\bot}) : U}}{\hat{Q}(a, <) \equiv \pi(\pi(\hat{N}, (q)a), (f)\pi(\pi(\hat{N}, (n)\, \mathrm{ap}(\mathrm{ap}(<, \mathrm{ap}(f, s(n))), \mathrm{ap}(f, n))), (p)\hat{\bot})) : U}\;(5)
$$

We have $Q(a, <) = T(\hat{Q}(a, <)) = T(\pi(\pi(\hat{N}, (q)a), (f)\pi(\pi(\hat{N}, (n)\, \mathrm{ap}(\mathrm{ap}(<, \mathrm{ap}(f, s(n))), \mathrm{ap}(f, n))), (p)\hat{\bot}))) = (\Pi f : N \to A)[(\Pi n : N)(\mathrm{ap}(f, s(n)) < \mathrm{ap}(f, n)) \to \bot] : \mathrm{set}$

In U^1 and U^2 we presuppose $A \to A \to U : \mathrm{set}$, which follows from

$$\dfrac{\dfrac{a : U \qquad \dfrac{a : U \qquad \overline{u : U}}{\pi(a,(x)u) : U}}{\pi(a,(y)\pi(a,(x)u)) : U}}{T(\pi(a,(y)\pi(a,(x)u))) = A \to A \to U : \mathrm{set}}$$

We will not explicitly write out premises like this in the following.

U^3

$$\dfrac{\overline{u : U} \qquad \dfrac{\dfrac{[a : U]^7 \qquad \dfrac{[a : U]^7 \qquad \overline{u : U}}{\pi(a,(x)u) : U}}{\pi(a,(y)\pi(a,(x)u)) : U} \qquad \dfrac{\begin{array}{c}[a : U]^7 \quad [<: A \to A \to U]^6 \\ U^1 \\ \hat{P}(a,<) : U\end{array} \qquad \begin{array}{c}[a : U]^7 \quad [<: A \to A \to U]^6 \\ U^2 \\ \hat{Q}(a,<) : U\end{array}}{\sigma(\hat{P}(a,<),(z)\hat{Q}(a,<)) : U}}{\sigma(\pi(a,(y)\pi(a,(x)u)),(<)\sigma(\hat{P}(a,<),(z)\hat{Q}(a,<))) : U}\,(6)}{\omega \equiv \sigma(u,(a)\sigma(\pi(a,(y)\pi(a,(x)u)),(<)\sigma(\hat{P}(a,<),(z)\hat{Q}(a,<)))) : U}\,(7)$$

Here we close the open assumptions $a : U$ and $<: A \to A \to U$ in U^1 and U^2.

We have $\Omega = T(\omega) = T(\sigma(u,(a)\sigma(\pi(a,(y)\pi(a,(x)u)),(<)\sigma(\hat{P}(a,<),(z)\hat{Q}(a,<)))))$
$= (\Sigma a : U)(\Sigma <: A \to A \to U)[P(a,<) \times Q(a,<)]$

We suppose that we have $A^* : \Omega$. We define $a \equiv p(A^*)$, $<_a \equiv p(q(A^*))$, $p_a \equiv p(q(q(A^*)))$ and $q_a \equiv q(q(q(A^*)))$. We also define $A \equiv T(a)$ and $B \equiv T(b)$.
The premises which show that the sets we use really are sets will as usual not be written out explicitly. They follow from U^1, U^2, U^3, and U : set.

D^1

$$\frac{[A^* : (\Sigma a : U)(\Sigma <_a : A \to A \to U)[P(a, <_a) \times Q(a, <_a)]]^{84}}{a \equiv p(A^*) : U}$$

D^{1b}

$$\frac{[B^* : (\Sigma a : U)(\Sigma <_a : A \to A \to U)[P(a, <_a) \times Q(a, <_a)]]^{83}}{b \equiv p(B^*) : U}$$

D^2

$$\dfrac{[A^* : (\Sigma a : U)(\Sigma <_a : A \to A \to U)[P(a, <_a) \times Q(a, <_a)]]^{84}}{\dfrac{q(A^*) : (\Sigma <_a : A \to A \to U)[P(a, <_a) \times Q(a, <_a)]}{<_a \equiv p(q(A^*)) : A \to A \to U}}$$

D^{2b}

$$\dfrac{[B^* : (\Sigma a : U)(\Sigma <_a : A \to A \to U)[P(a, <_a) \times Q(a, <_a)]]^{83}}{\dfrac{q(B^*) : (\Sigma <_b : B \to B \to U)[P(b, <_b) \times Q(b, <_b)]}{<_b \equiv p(q(B^*)) : B \to B \to U}}$$

D^3

$$\dfrac{[A^* : (\Sigma a : U)(\Sigma <_a : A \to A \to U)[P(a, <_a) \times Q(a, <_a)]]^{84}}{\dfrac{q(A^*) : (\Sigma <_a : A \to A \to U)[P(a, <_a) \times Q(a, <_a)]}{\dfrac{q(q(A^*)) : P(a, <_a) \times Q(a, <_a)}{p_a \equiv p(q(q(A^*))) : P(a, <_a)}}}$$

D^4

$$\dfrac{[A^* : (\Sigma a : U)(\Sigma <_a : A \to A \to U)[P(a, <_a) \times Q(a, <_a)]]^{84}}{\dfrac{q(A^*) : (\Sigma <_a : A \to A \to U)[P(a, <_a) \times Q(a, <_a)]}{\dfrac{q(q(A^*)) : P(a, <_a) \times Q(a, <_a)}{q_a \equiv q(q(q(A^*))) : Q(a, <_a)}}}$$

U^4

$$
\dfrac{
\dfrac{\stackrel{D^1}{a:U} \quad \stackrel{D^{1b}}{b:U}}{\pi(a,(q)b):U}
\quad
\dfrac{\stackrel{D^{1b}}{b:U} \quad
\dfrac{
\dfrac{\stackrel{D^1}{a:U} \quad
\dfrac{\stackrel{D^1}{a:U} \quad
\dfrac{
\dfrac{\dfrac{\stackrel{D^2}{<_a: A \to A \to U} \quad [x:A]^{10}}{\mathrm{ap}(<_a,x) : A \to U} \quad [y:A]^9}{\mathrm{ap}(\mathrm{ap}(<_a,x),y) : U}
\quad
\dfrac{\dfrac{\stackrel{D^{2b}}{<_b: B \to B \to U} \quad \dfrac{[f:A\to B]^{12} \quad [x:A]^{10}}{\mathrm{ap}(f,x):B}}{\mathrm{ap}(<_b,\mathrm{ap}(f,x)) : B \to U} \quad \dfrac{[f:A\to B]^{12} \quad [y:A]^9}{\mathrm{ap}(f,y):B}}{\mathrm{ap}(\mathrm{ap}(<_b,\mathrm{ap}(f,x)),\mathrm{ap}(f,y)) : U}
}{\pi(\mathrm{ap}(\mathrm{ap}(<_a,x),y),(p)\,\mathrm{ap}(\mathrm{ap}(<_b,\mathrm{ap}(f,x)),\mathrm{ap}(f,y))) : U}
}{\pi(a,(y)\pi(\mathrm{ap}(\mathrm{ap}(<_a,x),y),(p)\,\mathrm{ap}(\mathrm{ap}(<_b,\mathrm{ap}(f,x)),\mathrm{ap}(f,y)))) : U}\,(9)
}{\pi(a,(x)\pi(a,(y)\pi(\mathrm{ap}(\mathrm{ap}(<_a,x),y),(p)\,\mathrm{ap}(\mathrm{ap}(<_b,\mathrm{ap}(f,x)),\mathrm{ap}(f,y))))) : U}\,(10)
\quad
\stackrel{U^{4.1}}{\pi(a,(x)\,\mathrm{ap}(\mathrm{ap}(<_b,\mathrm{ap}(f,x)),z)) : U}
}{\sigma(\pi(a,(x)\pi(a,(y)\pi(\mathrm{ap}(\mathrm{ap}(<_a,x),y),(p)\,\mathrm{ap}(\mathrm{ap}(<_b,\mathrm{ap}(f,x)),\mathrm{ap}(f,y))))),(r)\pi(a,(x)\,\mathrm{ap}(\mathrm{ap}(<_b,\mathrm{ap}(f,x)),z))) : U}
}{\sigma(b,(z)\sigma(\pi(a,(x)\pi(a,(y)\pi(\mathrm{ap}(\mathrm{ap}(<_a,x),y),(p)\,\mathrm{ap}(\mathrm{ap}(<_b,\mathrm{ap}(f,x)),\mathrm{ap}(f,y))))),(r)\pi(a,(x)\,\mathrm{ap}(\mathrm{ap}(<_b,\mathrm{ap}(f,x)),z)))) : U}\,(11)
}{
\dfrac{
\dfrac{\sigma(\pi(a,(q)b),(f)\sigma(b,(z)\sigma(\pi(a,(x)\pi(a,(y)\pi(\mathrm{ap}(\mathrm{ap}(<_a,x),y),(p)\,\mathrm{ap}(\mathrm{ap}(<_b,\mathrm{ap}(f,x)),\mathrm{ap}(f,y))))),(r)\pi(a,(x)\,\mathrm{ap}(\mathrm{ap}(<_b,\mathrm{ap}(f,x)),z))))) : U}{(\lambda B^*)\sigma(\pi(a,(q)b),(f)\sigma(b,(z)\sigma(\pi(a,(x)\pi(a,(y)\pi(\mathrm{ap}(\mathrm{ap}(<_a,x),y),(p)\,\mathrm{ap}(\mathrm{ap}(<_b,\mathrm{ap}(f,x)),\mathrm{ap}(f,y))))),(r)\pi(a,(x)\,\mathrm{ap}(\mathrm{ap}(<_b,\mathrm{ap}(f,x)),z))))) : \Omega \to U}\,(83)
}{<_\omega \equiv (\lambda A^*)(\lambda B^*)\sigma(\pi(a,(q)b),(f)\sigma(b,(z)\sigma(\pi(a,(x)\pi(a,(y)\pi(\mathrm{ap}(\mathrm{ap}(<_a,x),y),(p)\,\mathrm{ap}(\mathrm{ap}(<_b,\mathrm{ap}(f,x)),\mathrm{ap}(f,y))))),(r)\pi(a,(x)\,\mathrm{ap}(\mathrm{ap}(<_b,\mathrm{ap}(f,x)),z))))) : \Omega \to \Omega \to U}\,(84)
}\,(12)
$$

$U^{4.1}$

$$
\dfrac{\stackrel{D^1}{a:U} \quad
\dfrac{\dfrac{\stackrel{D^{2b}}{<_b: B \to B \to U} \quad \dfrac{[f:A\to B]^{12} \quad [x:A]^8}{\mathrm{ap}(f,x):B}}{\mathrm{ap}(<_b,\mathrm{ap}(f,x)) : B \to U} \quad [z:B]^{11}}{\mathrm{ap}(\mathrm{ap}(<_b,\mathrm{ap}(f,x)),z) : U}
}{\pi(a,(x)\,\mathrm{ap}(\mathrm{ap}(<_b,\mathrm{ap}(f,x)),z)) : U}\,(8)
$$

When we apply $<_\omega : \Omega \to \Omega \to U$ to any two elements $A^* : \Omega$ and $B^* : \Omega$, we get the name
$ap(ap(<_\omega, A^*), B^*) = \sigma(\pi(a,(q)b),(f)\sigma(b,(z)\sigma(\pi(a,(x)\pi(a,(y)\pi(\mathrm{ap}(\mathrm{ap}(<_a,x),y),(p)\,\mathrm{ap}(\mathrm{ap}(<_b,\mathrm{ap}(f,x)),\mathrm{ap}(f,y))))),(r)\pi(a,(x)\,\mathrm{ap}(\mathrm{ap}(<_b,\mathrm{ap}(f,x)),z)))))$ in U.
We define the corresponding set as $A^* <_\omega B^* \equiv T(ap(ap(<_\omega, A^*), B^*))$
$= T(\sigma(\pi(a,(q)b),(f)\sigma(b,(z)\sigma(\pi(a,(x)\pi(a,(y)\pi(\mathrm{ap}(\mathrm{ap}(<_a,x),y),(p)\,\mathrm{ap}(\mathrm{ap}(<_b,\mathrm{ap}(f,x)),\mathrm{ap}(f,y))))),(r)\pi(a,(x)\,\mathrm{ap}(\mathrm{ap}(<_b,\mathrm{ap}(f,x)),z))))))$
$= (\Sigma f : A \to B)(\Sigma z : B)[(\Pi x : A)(\Pi y : A)(x <_a y \to \mathrm{ap}(f,x) <_b \mathrm{ap}(f,y)) \times (\Pi x : A)(\mathrm{ap}(f,x) <_b z)]$

In H we will now construct a proof t_ω of $\Omega^* <_\omega \Omega^*$ and then deduce a proof of absurdity.
These proofs consist of several parts:

Suppose we have $A^* : \Omega$ with $a \equiv p(A^*)$, $A \equiv T(a)$, $<_a \equiv p(q(A^*))$, $p_a \equiv p(q(q(A^*)))$ and $q_a \equiv q(q(q(A^*)))$.

In H^1 we define a function $g \equiv (\lambda x)A^*_x : A \to \Omega$, so that $\mathrm{ap}(g, x) = \mathrm{ap}((\lambda x)A^*_x, x) = A^*_x$, where we have $A^*_x \equiv \langle a_x, \langle <_{a_x}, \langle p_{a_x}, q_{a_x} \rangle\rangle\rangle$ and a_x is the name of the set A_x that contains all $v : A$ such that $v <_a x$. We show in H^2 that this function is order-preserving and in H^3 that it is bounded by A^*. This combines to a proof of $A^* <_\omega \Omega^*$ and as A^* was an arbitrary element of Ω, $A^* <_\omega \Omega^*$ is true for all $A^* : \Omega$. But as $\Omega^* : \Omega$, which is shown in H^4, we get a proof t_ω of $\Omega^* <_\omega \Omega^*$.

In H^5 we prove that the binary relation $<_\omega$ defined on Ω is transitive and in H^6 that Ω contains no infinitely descending chains.

To deduce absurdity, we then define the function $(\lambda n)\Omega^* : N \to \Omega$, that, applied to any $n : N$, yields $\Omega^* : \Omega$, so that we have $\mathrm{ap}((\lambda n)\Omega^*, s(n)) = \Omega^*$ and $\mathrm{ap}((\lambda n)\Omega^*, n) = \Omega^*$ for any $n : N$ and thus $\mathrm{ap}((\lambda n)\Omega^*, s(n)) <_\omega \mathrm{ap}((\lambda n)\Omega^*, n)$ for all $n : N$, which is an infinitely descending chain in Ω. But as we have shown in H^6, Ω has no infinitely descending chains and we get a contradiction.

H

$$
\dfrac{
\dfrac{
\stackrel{H^1}{g \equiv (\lambda x)A^*_x : A \to \Omega}
\quad
\dfrac{
[A^* : \Omega]^{84}
\quad
\dfrac{
\stackrel{H^2}{x1 : (\Pi x : A)(\Pi y : A)(x <_a y \to A^*_x <_\omega A^*_y)}
\quad
\stackrel{H^3}{x2 : (\Pi x : A)(A^*_x <_\omega A^*)}
}{
\langle x1, x2\rangle : (\Pi x : A)(\Pi y : A)(x <_a y \to A^*_x <_\omega A^*_y) \times (\Pi x : A)(A^*_x <_\omega A^*)
}
}{
\langle A^*, \langle x1, x2\rangle\rangle : (\Sigma z : \Omega)[(\Pi x : A)(\Pi y : A)(x <_a y \to A^*_x <_\omega A^*_y) \times (\Pi x : A)(A^*_x <_\omega z)]
}
}{
\dfrac{
\langle g, \langle A^*, \langle x1, x2\rangle\rangle\rangle : (\Sigma g : A \to \Omega)(\Sigma z : \Omega)[(\Pi x : A)(\Pi y : A)(x <_a y \to \mathrm{ap}(g,x) <_\omega \mathrm{ap}(g,y)) \times (\Pi x : A)(\mathrm{ap}(g,x) <_\omega z)] = A^* <_\omega \Omega^*
}{
(\lambda A^*)\langle g, \langle A^*, \langle x1, x2\rangle\rangle\rangle : (\Pi A^* : \Omega)(A^* <_\omega \Omega^*)
}\ (84)
}
\quad
\stackrel{H^4}{\Omega^* : \Omega}
}{
t_\omega \equiv \mathrm{ap}((\lambda A^*)\langle g, \langle A^*, \langle x1, x2\rangle\rangle\rangle, \Omega^*) : \Omega^* <_\omega \Omega^*
}
$$

proof of $\perp$

$$
\dfrac{
\dfrac{
\stackrel{H^6}{q_\omega : (\Pi f : N \to \Omega)[(\Pi n : N)(\mathrm{ap}(f, s(n)) <_\omega \mathrm{ap}(f,n)) \to \perp]}
\quad
\dfrac{\stackrel{H^4}{\Omega^* : \Omega}}{(\lambda n)\Omega^* : N \to \Omega}
}{
\mathrm{ap}(q_\omega, (\lambda n)\Omega^*) : (\Pi n : N)(\mathrm{ap}((\lambda n)\Omega^*, s(n)) <_\omega \mathrm{ap}((\lambda n)\Omega^*, n)) \to \perp
}
\quad
\dfrac{
\dfrac{
\stackrel{H}{t_\omega : \Omega^* <_\omega \Omega^*}
\quad
\dfrac{
\dfrac{\stackrel{H^4}{\Omega^* : \Omega}}{(\lambda n)\Omega^* : N \to \Omega}
\quad
\dfrac{[n : N]^{85}}{s(n) : N}
}{
\mathrm{ap}((\lambda n)\Omega^*, s(n)) = \Omega^* : \Omega
}
\quad
\dfrac{
\dfrac{\stackrel{H^4}{\Omega^* : \Omega}}{(\lambda n)\Omega^* : N \to \Omega}
\quad
[n : N]^{85}
}{
\mathrm{ap}((\lambda n)\Omega^*, n) = \Omega^* : \Omega
}
}{
t_\omega : \mathrm{ap}((\lambda n)\Omega^*, s(n)) <_\omega \mathrm{ap}((\lambda n)\Omega^*, n)
}
}{
(\lambda n)t_\omega : (\Pi n : N)(\mathrm{ap}((\lambda n)\Omega^*, s(n)) <_\omega \mathrm{ap}((\lambda n)\Omega^*, n))
}\ (85)
}{
\mathrm{ap}(\mathrm{ap}(q_\omega, (\lambda n)\Omega^*), (\lambda n)t_\omega) : \perp
}
$$

H^1

As before we have $A^* : \Omega$ with $a \equiv p(A^*)$, $A \equiv T(a)$, $<_a \equiv p(q(A^*))$, $p_a \equiv p(q(q(A^*)))$ and $q_a \equiv q(q(q(A^*)))$.

We define the function $g : A \to \Omega$ that maps an element $x : A$ into $A^*_x \equiv \langle a_x, \langle <_{a_x}, \langle p_{a_x}, q_{a_x}\rangle\rangle\rangle : \Omega$, where a_x is defined as the name of the set A_x which contains all $v : A$ such that $v <_a x$. Formally this becomes

$$a_x \equiv \sigma(a, (v)\,\mathrm{ap}(\mathrm{ap}(<_a, v), x))$$

$$A_x = T(a_x) = (\Sigma v : A)(v <_a x))$$

To show that $A^*_x \equiv \langle a_x, \langle <_{a_x}, \langle p_{a_x}, q_{a_x}\rangle\rangle\rangle : \Omega$, we have to show that $a_x : U$, that $<_{a_x}$, the binary relation $<_a$ with restricton to A_x, is a binary relation on A_x, and that the properties $P(a_x, <_{a_x})$ and $Q(a_x, <_{a_x})$ are fulfilled, i.e. that $<_a$ is transitive and that there are no infinitely descending chains in A_x. This follows from the fact that $<_a$ is transitive and that A does not contain any infinitely descending chains. We give the formal proofs $p_{a_x} : P(a_x, <_{a_x})$ and $q_{a_x} : Q(a_x, <_{a_x})$ in $H^{1.1}$ and $H^{1.2}$.

H^1

$$
\dfrac{\dfrac{\dfrac{D^1}{a : U} \quad \dfrac{\dfrac{\overset{D^2}{<_a : A \to A \to U} \quad [v : A]^{13}}{\mathrm{ap}(<_a, v) : A \to U} \quad [x : A]^{33}}{\mathrm{ap}(\mathrm{ap}(<_a, v), x) : U}\,(13)}{a_x \equiv \sigma(a, (v)\,\mathrm{ap}(\mathrm{ap}(<_a, v), x)) : U} \quad \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\overset{D^2}{<_a : A \to A \to U} \quad \dfrac{[ax : A_x]^{15}}{p(ax) : A}}{\mathrm{ap}(<_a, p(ax)) : A \to U} \quad \dfrac{[bx : A_x]^{14}}{p(bx) : A}}{\mathrm{ap}(\mathrm{ap}(<_a, p(ax)), p(bx)) : U}\,(14)}{(\lambda bx)\,\mathrm{ap}(\mathrm{ap}(<_a, p(ax)), p(bx)) : A_x \to U}\,(15)}{<_{a_x} \equiv (\lambda ax)(\lambda bx)\,\mathrm{ap}(\mathrm{ap}(<_a, p(ax)), p(bx)) : A_x \to A_x \to U} \quad \dfrac{\overset{H^{1.1}}{p_{a_x} : P(a_x, <_{a_x})} \quad \overset{H^{1.2}}{q_{a_x} : Q(a_x, <_{a_x})}}{\langle p_{a_x}, q_{a_x} \rangle : P(a_x, <_{a_x}) \times Q(a_x, <_{a_x})}}{\langle <_{a_x}, \langle p_{a_x}, q_{a_x} \rangle \rangle : (\Sigma <_{a_x} : A_x \to A_x \to U)[P(a_x, <_{a_x}) \times Q(a_x, <_{a_x})]}}{\dfrac{\langle a_x, \langle <_{a_x}, \langle p_{a_x}, q_{a_x} \rangle \rangle \rangle : (\Sigma a : U)(\Sigma <_a : A \to A \to U)[P(a, <_a) \times Q(a, <_a)] = \Omega}{g \equiv (\lambda x)\langle a_x, \langle <_{a_x}, \langle p_{a_x}, q_{a_x} \rangle \rangle \rangle : A \to \Omega}\,(33)}
$$

$H^{1.1}$

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\overset{D^3}{p_a : (\Pi u : A)(\Pi v : A)(\Pi w : A)[u <_a v \to (v <_a w \to u <_a w)]} \quad \dfrac{[ax : A_x]^{26}}{p(ax) : A}}{\mathrm{ap}(p_a, p(ax)) : (\Pi v : A)(\Pi w : A)[p(ax) <_a v \to (v <_a w \to p(ax) <_a w)]} \quad \dfrac{[bx : A_x]^{25}}{p(bx) : A}}{\mathrm{ap}(\mathrm{ap}(p_a, p(ax)), p(bx)) : (\Pi w : A)[p(ax) <_a p(bx) \to (p(bx) <_a w \to p(ax) <_a w)]} \quad \dfrac{[cx : A_x]^{24}}{p(cx) : A}}{\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p_a, p(ax)), p(bx)), p(cx)) : p(ax) <_a p(bx) \to (p(bx) <_a p(cx) \to p(ax) <_a p(cx))} \quad \overset{H^{1.1.1}}{\diamond : p(ax) <_a p(bx)}}{\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p_a, p(ax)), p(bx)), p(cx)), \diamond) : p(bx) <_a p(cx) \to p(ax) <_a p(cx)} \quad \overset{H^{1.1.2}}{\nabla : p(bx) <_a p(cx)}}{\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p_a, p(ax)), p(bx)), p(cx)), \diamond), \nabla) : p(ax) <_a p(cx)} \quad \overset{H^{1.1.3}}{ax <_{a_x} cx = p(ax) <_a p(cx) : set}}{\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p_a, p(ax)), p(bx)), p(cx)), \diamond), \nabla) : ax <_{a_x} cx}\,(22)}{(\lambda \nabla)\,\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p_a, p(ax)), p(bx)), p(cx)), \diamond), \nabla) : bx <_{a_x} cx \to ax <_{a_x} cx}\,(23)}{(\lambda \diamond)(\lambda \nabla)\,\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p_a, p(ax)), p(bx)), p(cx)), \diamond), \nabla) : ax <_{a_x} bx \to (bx <_{a_x} cx \to ax <_{a_x} cx)}\,(24)}{(\lambda cx)(\lambda \diamond)(\lambda \nabla)\,\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p_a, p(ax)), p(bx)), p(cx)), \diamond), \nabla) : (\Pi w : A_x)[ax <_{a_x} bx \to (bx <_{a_x} w \to ax <_{a_x} w)]}\,(25)}{(\lambda bx)(\lambda cx)(\lambda \diamond)(\lambda \nabla)\,\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p_a, p(ax)), p(bx)), p(cx)), \diamond), \nabla) : (\Pi v : A_x)(\Pi w : A_x)[ax <_{a_x} v \to (v <_{a_x} w \to ax <_{a_x} w)]}\,(26)}{p_{a_x} \equiv (\lambda ax)(\lambda bx)(\lambda cx)(\lambda \diamond)(\lambda \nabla)\,\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p_a, p(ax)), p(bx)), p(cx)), \diamond), \nabla) : (\Pi u : A_x)(\Pi v : A_x)(\Pi w : A_x)[u <_{a_x} v \to (v <_{a_x} w \to u <_{a_x} w)] = P(a_x, <_{a_x})}
$$

$H^{1.1.1}$

$$
\dfrac{
\dfrac{
\dfrac{
\dfrac{
\dfrac{
\dfrac{
\dfrac{
\dfrac{
\dfrac{
\overset{D^2}{<_a : A \to A \to U} \quad \dfrac{[u : A_x]^{17}}{p(u) : A}
}{\mathrm{ap}(<_a, p(u)) : A \to U}
\quad \dfrac{[v : A_x]^{16}}{p(v) : A}
}{\mathrm{ap}(\mathrm{ap}(<_a, p(u)), p(v)) : U}
}{(\lambda v)\,\mathrm{ap}(\mathrm{ap}(<_a, p(u)), p(v)) : A_x \to U}\,(16)
}{<_{a_x} \equiv (\lambda u)(\lambda v)\,\mathrm{ap}(\mathrm{ap}(<_a, p(u)), p(v)) : A_x \to A_x \to U}\,(17)
\quad [ax : A_x]^{26}
}{\mathrm{ap}(<_{a_x}, ax) = \mathrm{ap}((\lambda u)(\lambda v)\,\mathrm{ap}(\mathrm{ap}(<_a, p(u)), p(v)), ax) = (\lambda v)\,\mathrm{ap}(\mathrm{ap}(<_a, p(ax)), p(v)) : A_x \to U}
\quad [bx : A_x]^{25}
}{\mathrm{ap}(\mathrm{ap}(<_{a_x}, ax), bx) = \mathrm{ap}((\lambda v)\,\mathrm{ap}(\mathrm{ap}(<_a, p(ax)), p(v)), bx) = \mathrm{ap}(\mathrm{ap}(<_a, p(ax)), p(bx)) : U}
}{ax <_{a_x} bx \equiv T(\mathrm{ap}(\mathrm{ap}(<_{a_x}, ax), bx)) = T(\mathrm{ap}(\mathrm{ap}(<_a, p(ax)), p(bx))) : \mathrm{set}}
\quad
\dfrac{
\dfrac{
\dfrac{
\overset{D^2}{<_a : A \to A \to U} \quad \dfrac{[ax : A_x]^{26}}{p(ax) : A}
}{\mathrm{ap}(<_a, p(ax)) : A \to U}
\quad \dfrac{[bx : A_x]^{25}}{p(bx) : A}
}{\mathrm{ap}(\mathrm{ap}(<_a, p(ax)), p(bx)) : U}
}{p(ax) <_a p(bx) \equiv T(\mathrm{ap}(\mathrm{ap}(<_a, p(ax)), p(bx))) : \mathrm{set}}
}{ax <_{a_x} bx = p(ax) <_a p(bx) : \mathrm{set}}
\quad [\diamond : ax <_{a_x} bx]^{23}
}{\diamond : p(ax) <_a p(bx)}
$$

$H^{1.1.2}$

$$
\dfrac{
\dfrac{
\dfrac{
\dfrac{
\dfrac{
\dfrac{
\dfrac{
\dfrac{
\dfrac{
\overset{D^2}{<_a : A \to A \to U} \quad \dfrac{[v : A_x]^{19}}{p(v) : A}
}{\mathrm{ap}(<_a, p(v)) : A \to U}
\quad \dfrac{[w : A_x]^{18}}{p(w) : A}
}{\mathrm{ap}(\mathrm{ap}(<_a, p(v)), p(w)) : U}
}{(\lambda w)\,\mathrm{ap}(\mathrm{ap}(<_a, p(v)), p(w)) : A_x \to U}\,(18)
}{<_{a_x} \equiv (\lambda v)(\lambda w)\,\mathrm{ap}(\mathrm{ap}(<_a, p(v)), p(w)) : A_x \to A_x \to U}\,(19)
\quad [bx : A_x]^{25}
}{\mathrm{ap}(<_{a_x}, bx) = \mathrm{ap}((\lambda v)(\lambda w)\,\mathrm{ap}(\mathrm{ap}(<_a, p(v)), p(w)), bx) = (\lambda w)\,\mathrm{ap}(\mathrm{ap}(<_a, p(bx)), p(w)) : A_x \to U}
\quad [cx : A_x]^{24}
}{\mathrm{ap}(\mathrm{ap}(<_{a_x}, bx), cx) = \mathrm{ap}((\lambda w)\,\mathrm{ap}(\mathrm{ap}(<_a, p(bx)), p(w)), cx) = \mathrm{ap}(\mathrm{ap}(<_a, p(bx)), p(cx)) : U}
}{bx <_{a_x} cx \equiv T(\mathrm{ap}(\mathrm{ap}(<_{a_x}, bx), cx)) = T(\mathrm{ap}(\mathrm{ap}(<_a, p(bx)), p(cx))) : \mathrm{set}}
\quad
\dfrac{
\dfrac{
\dfrac{
\overset{D^2}{<_a : A \to A \to U} \quad \dfrac{[bx : A_x]^{25}}{p(bx) : A}
}{\mathrm{ap}(<_a, p(bx)) : A \to U}
\quad \dfrac{[cx : A_x]^{24}}{p(cx) : A}
}{\mathrm{ap}(\mathrm{ap}(<_a, p(bx)), p(cx)) : U}
}{p(bx) <_a p(cx) \equiv T(\mathrm{ap}(\mathrm{ap}(<_a, p(bx)), p(cx))) : \mathrm{set}}
}{bx <_{a_x} cx = p(bx) <_a p(cx) : set}
\quad [\nabla : bx <_{a_x} cx]^{22}
}{\nabla : p(bx) <_a p(cx)}
$$

$H^{1.1.3}$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\overset{D^2}{<_a\colon A \to A \to U} \quad \dfrac{[u : A_x]^{21}}{p(u) : A}}{\mathrm{ap}(<_a, p(u)) : A \to U} \quad \dfrac{[w : A_x]^{20}}{p(w) : A}}{\mathrm{ap}(\mathrm{ap}(<_a, p(u)), p(w)) : U}}{(\lambda w)\,\mathrm{ap}(\mathrm{ap}(<_a, p(u)), p(w)) : A_x \to U}\,(20)}{<_{a_x} \equiv (\lambda u)(\lambda w)\,\mathrm{ap}(\mathrm{ap}(<_a, p(u)), p(w)) : A_x \to A_x \to U}\,(21) \quad [ax : A_x]^{26}}{\mathrm{ap}(<_{a_x}, ax) = \mathrm{ap}((\lambda u)(\lambda w)\,\mathrm{ap}(\mathrm{ap}(<_a, p(u)), p(w)), ax) = (\lambda w)\,\mathrm{ap}(\mathrm{ap}(<_a, p(ax)), p(w)) : A_x \to U} \quad [cx : A_x]^{24}}{\mathrm{ap}(\mathrm{ap}(<_{a_x}, ax), cx) = \mathrm{ap}((\lambda w)\,\mathrm{ap}(\mathrm{ap}(<_a, p(ax)), p(w)), cx) = \mathrm{ap}(\mathrm{ap}(<_a, p(ax)), p(cx)) : U}}{ax <_{a_x} cx \equiv T(\mathrm{ap}(\mathrm{ap}(<_{a_x}, ax), cx)) = T(\mathrm{ap}(\mathrm{ap}(<_a, p(ax)), p(cx))) : \mathrm{set}} \quad \dfrac{\dfrac{\dfrac{\overset{D^2}{<_a\colon A \to A \to U} \quad \dfrac{[ax : A_x]^{26}}{p(ax) : A}}{\mathrm{ap}(<_a, p(ax)) : A \to U} \quad \dfrac{[cx : A_x]^{24}}{p(cx) : A}}{\mathrm{ap}(\mathrm{ap}(<_a, p(ax)), p(cx)) : U}}{p(ax) <_a p(cx) \equiv T(\mathrm{ap}(\mathrm{ap}(<_a, p(ax)), p(cx))) : \mathrm{set}}}{ax <_{a_x} cx = p(ax) <_a p(cx) : \mathrm{set}}$$

$H^{1.2}$

$$\dfrac{\dfrac{\dfrac{\overset{D^4}{q_a : (\Pi f : N \to A)[(\Pi n : N)(\mathrm{ap}(f, s(n)) <_a \mathrm{ap}(f, n)) \to \bot]} \quad \dfrac{\dfrac{\dfrac{[f_q : N \to A_x]^{32} \quad [n : N]^{27}}{\mathrm{ap}(f_q, n) : A_x}}{p(\mathrm{ap}(f_q, n)) : A}}{(\lambda n)p(\mathrm{ap}(f_q, n)) : N \to A}\,(27)}{\mathrm{ap}(q_a, (\lambda n)p(\mathrm{ap}(f_q, n))) : (\Pi n : N)(p(\mathrm{ap}(f_q, s(n))) <_a p(\mathrm{ap}(f_q, n))) \to \bot} \quad \overset{H^{1.2.1}}{(\lambda n)\,\mathrm{ap}(r, n) : (\Pi n : N)(p(\mathrm{ap}(f_q, s(n))) <_a p(\mathrm{ap}(f_q, n)))}}{\dfrac{\mathrm{ap}(\mathrm{ap}(q_a, (\lambda n)p(\mathrm{ap}(f_q, n))), (\lambda n)\,\mathrm{ap}(r, n)) : \bot}{(\lambda r)\,\mathrm{ap}(\mathrm{ap}(q_a, (\lambda n)p(\mathrm{ap}(f_q, n))), (\lambda n)\,\mathrm{ap}(r, n)) : (\Pi n : N)(\mathrm{ap}(f_q, s(n)) <_{a_x} \mathrm{ap}(f_q, n)) \to \bot}\,(31)}}{q_{a_x} \equiv (\lambda f_q)(\lambda r)\,\mathrm{ap}(\mathrm{ap}(q_a, (\lambda n)p(\mathrm{ap}(f_q, n))), (\lambda n)\,\mathrm{ap}(r, n)) : (\Pi f : N \to A_x)[(\Pi n : N)(\mathrm{ap}(f, s(n)) <_{a_x} \mathrm{ap}(f, n)) \to \bot] = Q(a_x, <_{a_x})}\,(32)$$

$H^{1.2.1}$

$$\dfrac{\dfrac{\dfrac{[r : (\Pi n : N)(\mathrm{ap}(f_q, s(n)) <_{a_x} \mathrm{ap}(f_q, n))]^{31} \quad [n : N]^{30}}{\mathrm{ap}(r, n) : \mathrm{ap}(f_q, s(n)) <_{a_x} \mathrm{ap}(f_q, n)} \quad \dfrac{\overset{H^{1.2.1.1}}{\mathrm{ap}(f_q, s(n)) <_{a_x} \mathrm{ap}(f_q, n) = T(\mathrm{ap}(\mathrm{ap}(<_a, p(\mathrm{ap}(f_q, s(n)))), p(\mathrm{ap}(f_q, n)))) : \mathrm{set}} \quad \overset{H^{1.2.1.2}}{T(\mathrm{ap}(\mathrm{ap}(<_a, p(\mathrm{ap}(f_q, s(n)))), p(\mathrm{ap}(f_q, n)))) = p(\mathrm{ap}(f_q, s(n))) <_a p(\mathrm{ap}(f_q, n)) : \mathrm{set}}}{\mathrm{ap}(f_q, s(n)) <_{a_x} \mathrm{ap}(f_q, n) = p(\mathrm{ap}(f_q, s(n))) <_a p(\mathrm{ap}(f_q, n)) : \mathrm{set}}}{\mathrm{ap}(r, n) : p(\mathrm{ap}(f_q, s(n))) <_a p(\mathrm{ap}(f_q, n))}}{(\lambda n)\,\mathrm{ap}(r, n) : (\Pi n : N)(p(\mathrm{ap}(f_q, s(n))) <_a p(\mathrm{ap}(f_q, n)))}\,(30)$$

$H^{1.2.1.1}$

$$
\dfrac{
\dfrac{
\dfrac{
\dfrac{
\dfrac{
\dfrac{
\dfrac{
\overset{D^2}{<_a : A \to A \to U}
\qquad
\dfrac{[w : A_x]^{29}}{p(w) : A}
}{
\mathrm{ap}(<_a, p(w)) : A \to U
}
\qquad
\dfrac{[v : A_x]^{28}}{p(v) : A}
}{
\mathrm{ap}(\mathrm{ap}(<_a, p(w)), p(v)) : U
}
}{
(\lambda v)\,\mathrm{ap}(\mathrm{ap}(<_a, p(w)), p(v)) : A_x \to U
}\ (28)
}{
<_{a_x} \equiv (\lambda w)(\lambda v)\,\mathrm{ap}(\mathrm{ap}(<_a, p(w)), p(v)) : A_x \to A_x \to U
}\ (29)
\qquad
\dfrac{
[f_q : N \to A_x]^{32}
\qquad
\dfrac{[n : N]^{30}}{s(n) : N}
}{
\mathrm{ap}(f_q, s(n)) : A_x
}
}{
\mathrm{ap}(<_{a_x}, \mathrm{ap}(f_q, s(n))) = \mathrm{ap}((\lambda w)(\lambda v)\,\mathrm{ap}(\mathrm{ap}(<_a, p(w)), p(v)), \mathrm{ap}(f_q, s(n))) = (\lambda v)\,\mathrm{ap}(\mathrm{ap}(<_a, p(\mathrm{ap}(f_q, s(n)))), p(v)) : A_x \to U
}
\qquad
\dfrac{
[f_q : N \to A_x]^{32}
\qquad
[n : N]^{30}
}{
\mathrm{ap}(f_q, n) : A_x
}
}{
\mathrm{ap}(\mathrm{ap}(<_{a_x}, \mathrm{ap}(f_q, s(n))), \mathrm{ap}(f_q, n)) = \mathrm{ap}((\lambda v)\,\mathrm{ap}(\mathrm{ap}(<_a, p(\mathrm{ap}(f_q, s(n)))), p(v)), \mathrm{ap}(f_q, n)) = \mathrm{ap}(\mathrm{ap}(<_a, p(\mathrm{ap}(f_q, s(n)))), p(\mathrm{ap}(f_q, n))) : U
}
}{
\mathrm{ap}(f_q, s(n)) <_{a_x} \mathrm{ap}(f_q, n) \equiv T(\mathrm{ap}(\mathrm{ap}(<_{a_x}, \mathrm{ap}(f_q, s(n))), \mathrm{ap}(f_q, n))) = T(\mathrm{ap}(\mathrm{ap}(<_a, p(\mathrm{ap}(f_q, s(n)))), p(\mathrm{ap}(f_q, n)))) : \mathrm{set}
}
$$

$H^{1.2.1.2}$

$$
\dfrac{
\dfrac{
\dfrac{
\overset{D^2}{<_a : A \to A \to U}
\qquad
\dfrac{
\dfrac{
[f_q : N \to A_x]^{32}
\qquad
\dfrac{[n : N]^{30}}{s(n) : N}
}{
\mathrm{ap}(f_q, s(n)) : A_x
}
}{
p(\mathrm{ap}(f_q, s(n))) : A
}
}{
\mathrm{ap}(<_a, p(\mathrm{ap}(f_q, s(n)))) : A \to U
}
\qquad
\dfrac{
\dfrac{
[f_q : N \to A_x]^{32}
\qquad
[n : N]^{30}
}{
\mathrm{ap}(f_q, n) : A_x
}
}{
p(\mathrm{ap}(f_q, n)) : A
}
}{
\mathrm{ap}(\mathrm{ap}(<_a, p(\mathrm{ap}(f_q, s(n)))), p(\mathrm{ap}(f_q, n))) : U
}
}{
p(\mathrm{ap}(f_q, s(n))) <_a p(\mathrm{ap}(f_q, n)) \equiv T(\mathrm{ap}(\mathrm{ap}(<_a, p(\mathrm{ap}(f_q, s(n)))), p(\mathrm{ap}(f_q, n)))) : \mathrm{set}
}
$$

H^2

Suppose we have $x : A$ and $y : A$ such that $x <_a y$. Let A_x and A_y as before be the sets $A_x = (\Sigma v : A)(v <_a x)$ and $A_y = (\Sigma v : A)(v <_a y)$. An element of A_x consists thus of an element $w : A$ together with a proof of $w <_a x$.

In $H^{2.1}$ we construct an order-preserving function $f : A_x \to A_y$ in the following way: f maps a pair in A_x into the pair in A_y that contains the same $w : A$ as the pair in A_x and a proof of $w <_a y$, which follwos from $w <_a x$, $x <_a y$ and transitivity of $<_a$. In $H^{2.3}$ we give a formal proof of the order-preserving property of f.

In $H^{2.2}$ we define $\langle x, \triangle \rangle : A_y$ as the pair that consists of $x : A$ and the proof of $x <_a y$. For this $\langle x, \triangle \rangle$ we then have $\mathrm{ap}(f, ax) <_{a_y} \langle x, \triangle \rangle$ for all $ax : A_x$, which is proved in $H^{2.4}$.

Thus, for all $x : A$ and $y : A$, it follows from $x <_a y$ that $A_x^* <_\omega A_y^*$, or $\mathrm{ap}(g, x) <_\omega \mathrm{ap}(g, y)$, where g is the function $g \equiv (\lambda x)\langle a_x, \langle <_{a_x}, \langle p_{a_x}, q_{a_x} \rangle\rangle\rangle$.

H^2

$$
\dfrac{\dfrac{\dfrac{\overset{H^{2.1}}{f : A_x \to A_y} \quad \dfrac{\overset{H^{2.2}}{\langle x, \triangle\rangle : A_y} \quad \dfrac{\overset{H^{2.3}}{(\lambda ax)(\lambda bx)(\lambda\diamond)\diamond : (\Pi ax : A_x)(\Pi bx : A_x)(ax <_{a_x} bx \to fa <_{a_y} fb)} \quad \overset{H^{2.4}}{(\lambda ax)q(ax) : (\Pi ax : A_x)(fa <_{a_y} \langle x, \triangle\rangle)}}{\langle(\lambda ax)(\lambda bx)(\lambda\diamond)\diamond, (\lambda ax)q(ax)\rangle : (\Pi ax : A_x)(\Pi bx : A_x)(ax <_{a_x} bx \to fa <_{a_y} fb) \times (\Pi ax : A_x)(fa <_{a_y} \langle x, \triangle\rangle)}}{\langle\langle x, \triangle\rangle, \langle(\lambda ax)(\lambda bx)(\lambda\diamond)\diamond, (\lambda ax)q(ax)\rangle\rangle : (\Sigma z : A_y)[(\Pi ax : A_x)(\Pi bx : A_x)(ax <_{a_x} bx \to fa <_{a_y} fb) \times (\Pi ax : A_x)(fa <_{a_y} z)]}}{\langle f, \langle\langle x, \triangle\rangle, \langle(\lambda ax)(\lambda bx)(\lambda\diamond)\diamond, (\lambda ax)q(ax)\rangle\rangle\rangle : (\Sigma f : A_x \to A_y)(\Sigma z : A_y)[(\Pi ax : A_x)(\Pi bx : A_x)(ax <_{a_x} bx \to \mathrm{ap}(f, ax) <_{a_y} \mathrm{ap}(f, bx)) \times (\Pi ax : A_x)(\mathrm{ap}(f, ax) <_{a_y} z)] = A_x^* <_\omega A_y^*}\;(46)}{(\lambda\triangle)\langle f, \langle\langle x, \triangle\rangle, \langle(\lambda ax)(\lambda bx)(\lambda\diamond)\diamond, (\lambda ax)q(ax)\rangle\rangle\rangle : x <_a y \to A_x^* <_\omega A_y^*}
$$

$$
\dfrac{\dfrac{(\lambda\triangle)\langle f, \langle\langle x, \triangle\rangle, \langle(\lambda ax)(\lambda bx)(\lambda\diamond)\diamond, (\lambda ax)q(ax)\rangle\rangle\rangle : x <_a y \to A_x^* <_\omega A_y^*}{(\lambda y)(\lambda\triangle)\langle f, \langle\langle x, \triangle\rangle, \langle(\lambda ax)(\lambda bx)(\lambda\diamond)\diamond, (\lambda ax)q(ax)\rangle\rangle\rangle : (\Pi y : A)(x <_a y \to A_x^* <_\omega A_y^*)}\;(47)}{x1 \equiv (\lambda x)(\lambda y)(\lambda\triangle)\langle f, \langle\langle x, \triangle\rangle, \langle(\lambda ax)(\lambda bx)(\lambda\diamond)\diamond, (\lambda ax)q(ax)\rangle\rangle\rangle : (\Pi x : A)(\Pi y : A)(x <_a y \to A_x^* <_\omega A_y^*)}\;(48)
$$

$H^{2.1}$

$$
\dfrac{\dfrac{\dfrac{[aw : A_x]^{34}}{p(aw) : A} \quad \dfrac{[\triangle : x <_a y]^{46} \quad \dfrac{\dfrac{[aw : A_x]^{34}}{q(aw) : p(aw) <_a x} \quad \dfrac{\dfrac{\dfrac{\overset{D^3}{p_a : (\Pi u : A)(\Pi v : A)(\Pi w : A)[u <_a v \to (v <_a w \to u <_a w)]} \quad \dfrac{[aw : A_x]^{34}}{p(aw) : A}}{\mathrm{ap}(p_a, p(aw)) : (\Pi v : A)(\Pi w : A)[p(aw) <_a v \to (v <_a w \to p(aw) <_a w)]} \quad [x : A]^{48}}{\mathrm{ap}(\mathrm{ap}(p_a, p(aw)), x) : (\Pi w : A)[p(aw) <_a x \to (x <_a w \to p(aw) <_a w)]} \quad [y : A]^{47}}{\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p_a, p(aw)), x), y) : p(aw) <_a x \to (x <_a y \to p(aw) <_a y)}}{\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p_a, p(aw)), x), y), q(aw)) : x <_a y \to p(aw) <_a y}}{\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p_a, p(aw)), x), y), q(aw)), \triangle) : p(aw) <_a y}}{\langle p(aw), \mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p_a, p(aw)), x), y), q(aw)), \triangle)\rangle : (\Sigma v : A)(v <_a y) = A_y}}{f \equiv (\lambda aw)\langle p(aw), \mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p_a, p(aw)), x), y), q(aw)), \triangle)\rangle : A_x \to A_y}\;(34)
$$

$H^{2.2}$

$$\frac{[x : A]^{48} \quad [\triangle : x <_a y]^{46}}{\langle x, \triangle \rangle : (\Sigma v : A)(v <_a y) = A_y}$$

$H^{2.3}$

$$\cfrac{\cfrac{\cfrac{\overset{H^{2.3.1}}{\diamond : fa <_{a_y} fb}}{(\lambda\diamond)\diamond : ax <_{a_x} bx \to fa <_{a_y} fb}\ (40)}{(\lambda bx)(\lambda\diamond)\diamond : (\Pi bx : A_x)(ax <_{a_x} bx \to fa <_{a_y} fb)}\ (41)}{(\lambda ax)(\lambda bx)(\lambda\diamond)\diamond : (\Pi ax : A_x)(\Pi bx : A_x)(ax <_{a_x} bx \to fa <_{a_y} fb)}\ (42)$$

$H^{2.3.1}$

$$\cfrac{\overset{H^{2.3.1.1}}{\diamond : p(fa) <_a p(fb)} \quad \cfrac{\cfrac{\cfrac{\cfrac{\overset{D^2}{<_a : A \to A \to U} \quad \overset{H^{2.3.1.1.1}}{p(fa) : A}}{\mathrm{ap}(<_a, p(fa)) : A \to U}}{\mathrm{ap}(\mathrm{ap}(<_a, p(fa)), p(fb)) : U} \quad \overset{H^{2.3.1.1.2}}{p(fb) : A}}{p(fa) <_a p(fb) \equiv T(\mathrm{ap}(\mathrm{ap}(<_a, p(fa)), p(fb))) : \mathrm{set}} \quad \overset{H^{2.3.1.2}}{T(\mathrm{ap}(\mathrm{ap}(<_a, p(fa)), p(fb))) = fa <_{a_y} fb : \mathrm{set}}}{p(fa) <_a p(fb) = fa <_{a_y} fb : \mathrm{set}}}{\diamond : fa <_{a_y} fb}$$

$H^{2.3.1.1}$

$$\dfrac{\dfrac{\stackrel{H^{2.3.1.1.1}}{p(ax) = p(fa) : A} \quad \stackrel{H^{2.3.1.1.2}}{p(bx) = p(fb) : A} \quad \dfrac{\dfrac{\dfrac{\stackrel{D^2}{<_a : A \to A \to U} \quad [x : A]^{35}}{\mathrm{ap}(<_a, x) : A \to U} \quad [y : A]^{35}}{\mathrm{ap}(\mathrm{ap}(<_a, x), y) : U}}{x <_a y \equiv T(\mathrm{ap}(\mathrm{ap}(<_a, x), y)) : \mathrm{set}}\,(35)}{p(ax) <_a p(bx) = p(fa) <_a p(fb) : \mathrm{set}} \quad \dfrac{[\diamond : ax <_{a_x} bx]^{40} \quad \stackrel{H^{2.3.1.1.3}}{ax <_{a_x} bx = p(ax) <_a p(bx) : \mathrm{set}}}{\diamond : p(ax) <_a p(bx)}}{\diamond : p(fa) <_a p(fb)}$$

$H^{2.3.1.1.1}$

$$\dfrac{\dfrac{\stackrel{H^{2.1}}{(\lambda aw)\langle p(aw), \mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p_a, p(aw)), x), y), q(aw)), \triangle)\rangle : A_x \to A_y} \quad [ax : A_x]^{42}}{fa \equiv \mathrm{ap}((\lambda aw)\langle p(aw), \mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p_a, p(aw)), x), y), q(aw)), \triangle)\rangle, ax) = \langle p(ax), \mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p_a, p(ax)), x), y), q(ax)), \triangle)\rangle : A_y}}{p(fa) = p(\langle p(ax), \mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p_a, p(ax)), x), y), q(ax)), \triangle)\rangle) = p(ax) : A}$$

$H^{2.3.1.1.2}$

$$\dfrac{\dfrac{\stackrel{H^{2.1}}{(\lambda aw)\langle p(aw), \mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p_a, p(aw)), x), y), q(aw)), \triangle)\rangle : A_x \to A_y} \quad [bx : A_x]^{41}}{fb \equiv \mathrm{ap}((\lambda aw)\langle p(aw), \mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p_a, p(aw)), x), y), q(aw)), \triangle)\rangle, bx) = \langle p(bx), \mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p_a, p(bx)), x), y), q(bx)), \triangle)\rangle : A_y}}{p(fb) = p(\langle p(bx), \mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p_a, p(bx)), x), y), q(bx)), \triangle)\rangle) = p(bx) : A}$$

$H^{2.3.1.1.3}$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\overset{D^2}{<_a : A \to A \to U} \quad \dfrac{[v : A_x]^{37}}{p(v) : A}}{\mathrm{ap}(<_a, p(v)) : A \to U} \quad \dfrac{[w : A_x]^{36}}{p(w) : A}}{\mathrm{ap}(\mathrm{ap}(<_a, p(v)), p(w)) : U}}{(\lambda w)\,\mathrm{ap}(\mathrm{ap}(<_a, p(v)), p(w)) : A_x \to U}\,(36)}{<_{a_x} \equiv (\lambda v)(\lambda w)\,\mathrm{ap}(\mathrm{ap}(<_a, p(v)), p(w)) : A_x \to A_x \to U}\,(37) \quad [ax : A_x]^{42}}{\mathrm{ap}(<_{a_x}, ax) = \mathrm{ap}((\lambda v)(\lambda w)\,\mathrm{ap}(\mathrm{ap}(<_a, p(v)), p(w)), ax) = (\lambda w)\,\mathrm{ap}(\mathrm{ap}(<_a, p(ax)), p(w)) : A_x \to U} \quad [bx : A_x]^{41}}{\mathrm{ap}(\mathrm{ap}(<_{a_x}, ax), bx) = \mathrm{ap}((\lambda w)\,\mathrm{ap}(\mathrm{ap}(<_a, p(ax)), p(w)), bx) = \mathrm{ap}(\mathrm{ap}(<_a, p(ax)), p(bx)) : U}}{ax <_{a_x} bx \equiv T(\mathrm{ap}(\mathrm{ap}(<_{a_x}, ax), bx)) = T(\mathrm{ap}(\mathrm{ap}(<_a, p(ax)), p(bx))) : \mathrm{set}} \quad \dfrac{\dfrac{\dfrac{\overset{D^2}{<_a : A \to A \to U} \quad \dfrac{[ax : A_x]^{42}}{p(ax) : A}}{\mathrm{ap}(<_a, p(ax)) : A \to U} \quad \dfrac{[bx : A_x]^{41}}{p(bx) : A}}{\mathrm{ap}(\mathrm{ap}(<_a, p(ax)), p(bx)) : U}}{p(ax) <_a p(bx) \equiv T(\mathrm{ap}(\mathrm{ap}(<_a, p(ax)), p(bx))) : \mathrm{set}}}{ax <_{a_x} bx = p(ax) <_a p(bx) : \mathrm{set}}$$

$H^{2.3.1.2}$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\overset{D^2}{<_a : A \to A \to U} \quad \dfrac{[u : A_y]^{39}}{p(u) : A}}{\mathrm{ap}(<_a, p(u)) : A \to U} \quad \dfrac{[v : A_y]^{38}}{p(v) : A}}{\mathrm{ap}(\mathrm{ap}(<_a, p(u)), p(v)) : U}}{(\lambda v)\,\mathrm{ap}(\mathrm{ap}(<_a, p(u)), p(v)) : A_y \to U}\,(38)}{<_{a_y} \equiv (\lambda u)(\lambda v)\,\mathrm{ap}(\mathrm{ap}(<_a, p(u)), p(v)) : A_y \to A_y \to U}\,(39) \quad \overset{H^{2.3.1.2.1}}{fa : A_y}}{\mathrm{ap}(<_{a_y}, fa) = \mathrm{ap}((\lambda u)(\lambda v)\,\mathrm{ap}(\mathrm{ap}(<_a, p(u)), p(v)), fa) = (\lambda v)\,\mathrm{ap}(\mathrm{ap}(<_a, p(fa)), p(v)) : A_y \to U} \quad \overset{H^{2.3.1.2.2}}{fb : A_y}}{\mathrm{ap}(\mathrm{ap}(<_{a_y}, fa), fb) = \mathrm{ap}((\lambda v)\,\mathrm{ap}(\mathrm{ap}(<_a, p(fa)), p(v)), fb) = \mathrm{ap}(\mathrm{ap}(<_a, p(fa)), p(fb)) : U}}{fa <_{a_y} fb \equiv T(\mathrm{ap}(\mathrm{ap}(<_{a_y}, fa), fb)) = T(\mathrm{ap}(\mathrm{ap}(<_a, p(fa)), p(fb))) : \mathrm{set}}$$

$H^{2.3.1.2.1}$

$$\dfrac{\overset{H^{2.1}}{(\lambda aw)\langle p(aw), \mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p_a, p(aw)), x), y), q(aw)), \triangle)\rangle : A_x \to A_y} \quad [ax : A_x]^{42}}{fa \equiv \mathrm{ap}((\lambda aw)\langle p(aw), \mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p_a, p(aw)), x), y), q(aw)), \triangle)\rangle, ax) = \langle p(ax), \mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p_a, p(ax)), x), y), q(ax)), \triangle)\rangle : A_y}$$

$H^{2.3.1.2.2}$

$$\dfrac{\overset{H^{2.1}}{(\lambda aw)\langle p(aw), \mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p_a, p(aw)), x), y), q(aw)), \triangle)\rangle : A_x \to A_y} \quad [bx : A_x]^{41}}{fb \equiv \mathrm{ap}((\lambda aw)\langle p(aw), \mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p_a, p(aw)), x), y), q(aw)), \triangle)\rangle, bx) = \langle p(bx), \mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p_a, p(bx)), x), y), q(bx)), \triangle)\rangle : A_y}$$

$H^{2.4}$

$$\dfrac{\dfrac{\overset{H^{2.4.1}}{q(ax) : p(fa) <_a p(\langle x, \triangle\rangle)} \quad \dfrac{\dfrac{\dfrac{\dfrac{\overset{D^2}{<_a : A \to A \to U} \quad \overset{H^{2.3.1.1.1}}{p(fa) : A}}{\mathrm{ap}(<_a, p(fa)) : A \to U} \quad \dfrac{\overset{H^{2.2}}{\langle x, \triangle\rangle : A_y}}{p(\langle x, \triangle\rangle) : A}}{\mathrm{ap}(\mathrm{ap}(<_a, p(fa)), p(\langle x, \triangle\rangle)) : U}}{p(fa) <_a p(\langle x, \triangle\rangle) \equiv T(\mathrm{ap}(\mathrm{ap}(<_a, p(fa)), p(\langle x, \triangle\rangle))) : \mathrm{set}} \quad \overset{H^{2.4.2}}{T(\mathrm{ap}(\mathrm{ap}(<_a, p(fa)), p(\langle x, \triangle\rangle))) = fa <_{a_y} \langle x, \triangle\rangle : \mathrm{set}}}{p(fa) <_a p(\langle x, \triangle\rangle) = fa <_{a_y} \langle x, \triangle\rangle : \mathrm{set}}}{\dfrac{q(ax) : fa <_{a_y} \langle x, \triangle\rangle}{(\lambda ax)q(ax) : (\Pi ax : A_x)(fa <_{a_y} \langle x, \triangle\rangle)}\ (45)}$$

$H^{2.4.1}$

$$\dfrac{\dfrac{\overset{H^{2.4.1.1}}{p(ax) = p(fa) : A} \quad \dfrac{[ax : A_x]^{45}}{q(ax) : p(ax) <_a x}}{q(ax) : p(fa) <_a x} \quad \dfrac{\overset{H^{2.2}}{\langle x, \triangle\rangle : A_y}}{p(\langle x, \triangle\rangle) = x : A}}{q(ax) : p(fa) <_a p(\langle x, \triangle\rangle)}$$

$H^{2.4.1.1}$

$$\dfrac{\dfrac{\overset{H^{2.1}}{(\lambda aw)\langle p(aw), \mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p_a, p(aw)), x), y), q(aw)), \triangle)\rangle : A_x \to A_y} \quad [ax : A_x]^{45}}{fa \equiv \mathrm{ap}((\lambda aw)\langle p(aw), \mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p_a, p(aw)), x), y), q(aw)), \triangle)\rangle, ax) = \langle p(ax), \mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p_a, p(ax)), x), y), q(ax)), \triangle)\rangle : A_y}}{p(fa) = p(\langle p(ax), \mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p_a, p(ax)), x), y), q(ax)), \triangle)\rangle) = p(ax) : A}$$

$H^{2.4.2}$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\overset{D^2}{<_a : A \to A \to U} \quad \dfrac{[u : A_y]^{44}}{p(u) : A}}{\mathrm{ap}(<_a, p(u)) : A \to U} \quad \dfrac{[v : A_y]^{43}}{p(v) : A}}{\mathrm{ap}(\mathrm{ap}(<_a, p(u)), p(v)) : U}}{(\lambda v)\,\mathrm{ap}(\mathrm{ap}(<_a, p(u)), p(v)) : A_y \to U}\ (43)}{<_{a_y} \equiv (\lambda u)(\lambda v)\,\mathrm{ap}(\mathrm{ap}(<_a, p(u)), p(v)) : A_y \to A_y \to U}\ (44) \quad \overset{H^{2.4.2.1}}{fa : A_y}}{\mathrm{ap}(<_{a_y}, fa) = \mathrm{ap}((\lambda u)(\lambda v)\,\mathrm{ap}(\mathrm{ap}(<_a, p(u)), p(v)), fa) = \mathrm{ap}((\lambda v)\,\mathrm{ap}(\mathrm{ap}(<_a, p(fa)), p(v))) : A_y \to U} \quad \overset{H^{2.2}}{\langle x, \triangle\rangle : A_y}}{\mathrm{ap}(\mathrm{ap}(<_{a_y}, fa), \langle x, \triangle\rangle) = \mathrm{ap}((\lambda v)\,\mathrm{ap}(\mathrm{ap}(<_a, p(fa)), p(v)), \langle x, \triangle\rangle) = \mathrm{ap}(\mathrm{ap}(<_a, p(fa)), p(\langle x, \triangle\rangle)) : U}}{fa <_{a_y} \langle x, \triangle\rangle \equiv T(\mathrm{ap}(\mathrm{ap}(<_{a_y}, fa), \langle x, \triangle\rangle)) = T(\mathrm{ap}(\mathrm{ap}(<_a, p(fa)), p(\langle x, \triangle\rangle))) : \mathrm{set}}$$

$H^{2.4.2.1}$

$$\dfrac{\overset{H^{2.1}}{(\lambda ax)\langle p(ax), \mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p_a, p(ax)), x), y), q(ax)), \triangle)\rangle : A_x \to A_y} \quad [ax : A_x]^{45}}{fa \equiv \mathrm{ap}((\lambda ax)\langle p(ax), \mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p_a, p(ax)), x), y), q(ax)), \triangle)\rangle, ax) = \langle p(ax), \mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p_a, p(ax)), x), y), q(ax)), \triangle)\rangle : A_y}$$

H^3

We show that $A_x^* <_\omega A^*$, or $\mathrm{ap}(g, x) <_\omega A^*$, for any $x : A$.

A_x is the set that consists of all $v : A$ such that $v <_a x$, or formally, $A_x = (\Sigma v : A)(v <_a x)$.
An element $ax : A_x$ is thus evaluated into a pair that consists of an element $w : A$ together with a proof of $w <_a x$.

We show that $(\lambda a)p(a)$ is an order-preserving function $A_x \to A$, where p is the left projection; because $ax <_{a_x} bx$ means the same as $p(ax) <_a p(bx)$ for any $ax : A_x$ and $bx : A_x$ (which was formally shown in $H^{2.3.1.1.3}$), this follows immediately.

If q is the right projection, then $(\lambda a)q(a)$ applied to an element $ax : A_x$ yields a proof of $p(ax) <_a x$ for all $ax : A_x$. Thus $(\lambda a)p(a)$ is bounded by x.

H^3

$$
\dfrac{
\dfrac{
\dfrac{\dfrac{[ax : A_x]^{49}}{p(ax) : A}}{(\lambda ax)p(ax) : A_x \to A}\;(49)
\qquad
\dfrac{[x : A]^{52} \quad
\dfrac{
\dfrac{
\dfrac{
\dfrac{
\dfrac{[\diamond : ax <_{a_x} bx]^{50} \quad \overset{H^{2.3.1.1.3}}{ax <_{a_x} bx = p(ax) <_a p(bx) : \mathrm{set}}}{\diamond : p(ax) <_a p(bx)}
}{(\lambda\diamond)\diamond : (ax <_{a_x} bx \to p(ax) <_a p(bx))}\;(50)
}{(\lambda bx)(\lambda\diamond)\diamond : (\Pi bx : A_x)(ax <_{a_x} bx \to p(ax) <_a p(bx))}\;(41)
}{(\lambda ax)(\lambda bx)(\lambda\diamond)\diamond : (\Pi ax : A_x)(\Pi bx : A_x)(ax <_{a_x} bx \to p(ax) <_a p(bx))}\;(42)
\quad
\dfrac{\dfrac{[ax : A_x]^{51}}{q(ax) : p(ax) <_a x}}{(\lambda ax)q(ax) : (\Pi ax : A_x)(p(ax) <_a x)}\;(51)
}{\langle(\lambda ax)(\lambda bx)(\lambda\diamond)\diamond, (\lambda ax)q(ax)\rangle : (\Pi ax : A_x)(\Pi bx : A_x)(ax <_{a_x} bx \to p(ax) <_a p(bx)) \times (\Pi ax : A_x)(p(ax) <_a x)}
}{\langle x, \langle(\lambda ax)(\lambda bx)(\lambda\diamond)\diamond, (\lambda ax)q(ax)\rangle\rangle : (\Sigma z : A)[(\Pi ax : A_x)(\Pi bx : A_x)(ax <_{a_x} bx \to p(ax) <_a p(bx)) \times (\Pi ax : A_x)(p(ax) <_a z)]}
}{\langle(\lambda a)p(a), \langle x, \langle(\lambda ax)(\lambda bx)(\lambda\diamond)\diamond, (\lambda ax)q(ax)\rangle\rangle\rangle : (\Sigma f : A_x \to A)(\Sigma z : A)[(\Pi ax : A_x)(\Pi bx : A_x)(ax <_{a_x} bx \to \mathrm{ap}(f, ax) <_a \mathrm{ap}(f, bx)) \times (\Pi ax : A_x)(\mathrm{ap}(f, ax) <_a z)] = A_x^* <_\omega A^*}
}{x2 \equiv (\lambda x)\langle(\lambda a)p(a), \langle x, \langle(\lambda ax)(\lambda bx)(\lambda\diamond)\diamond, (\lambda ax)q(ax)\rangle\rangle\rangle : (\Pi x : A)(A_x^* <_\omega A^*)}\;(52)
$$

H^4

We now show that $\Omega^* \equiv \langle \omega, \langle <_\omega, \langle p_\omega, q_\omega \rangle \rangle \rangle$ is an element of Ω.
We have shown in U^3 that $\omega : U$, where we essentially needed the assumption $u : U$ with $T(u) = U$. In H^5 we show that the binary relation $<_\omega : \Omega \to \Omega \to U$, which we have defined in U^4, is transitive and in H^6 that Ω does not contain any infinitely descending chains.

$$\dfrac{\overset{U^3}{\omega : U} \quad \dfrac{\overset{U^4}{<_\omega : \Omega \to \Omega \to U} \quad \dfrac{\overset{H^5}{p_\omega : P(\omega, <_\omega)} \quad \overset{H^6}{q_\omega : Q(\omega, <_\omega)}}{\langle p_\omega, q_\omega \rangle : P(\omega, <_\omega) \times Q(\omega, <_\omega)}}{\langle <_\omega, \langle p_\omega, q_\omega \rangle \rangle : (\Sigma <_\omega : \Omega \to \Omega \to U)[P(\omega, <_\omega) \times Q(\omega, <_\omega)]}}{\Omega^* \equiv \langle \omega, \langle <_\omega, \langle p_\omega, q_\omega \rangle \rangle \rangle : (\Sigma a : U)(\Sigma <_a : A \to A \to U)[P(a, <_a) \times Q(a, <_a)] = \Omega}$$

H^5

We prove the property of transitivity of the binary relation $<_\omega$ on Ω, i.e. that for $A^* : \Omega$, $B^* : \Omega$ and $C^* : \Omega$, $A^* <_\omega B^*$ and $B^* <_\omega C^*$ implies $A^* <_\omega C^*$.

This is shown in the following way:

Suppose we have $A^* <_\omega B^*$ and $B^* <_\omega C^*$ for $A^* : \Omega$, $B^* : \Omega$ and $C^* : \Omega$. Then we have order-preserving functions $f : A \to B$ and $g : B \to C$, and bounds $z : B$ and $z' : C$. The function f is the element $p(ab)$, where ab is a proof of $A^* <_\omega B^*$, as we show in $H^{5.1}$. The function g is the element $p(bc)$, where bc is a proof of $B^* <_\omega C^*$, as we show in a first step in $H^{5.2}$. In $H^{5.4.1}$ and $H^{5.4.2}$ we prove that f is order-preserving and that g is order-preserving on the range of f.

In $H^{5.2}$ we then define in a second step the function $gf : A \to C$ as a composition of the functions f and g. This function has the form $(\lambda x)\,\mathrm{ap}(p(bc), \mathrm{ap}(p(ab), x))$. In $H^{5.4}$ we show that this function is order-preserving and in $H^{5.5}$ that it is bounded by the element $p(q(bc))$, which we get in $H^{5.3}$.

This combines to a proof of $A^* <_\omega C^*$. As $A^* <_\omega C^*$ follows from the assumptions $A^* <_\omega B^*$ and $B^* <_\omega C^*$ and as A^*, B^* and C^* were arbitrary elements in Ω, we get a proof p_ω of $P(\omega, <_\omega)$.

H^5

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\overset{H^{5.2}}{(\lambda x)\,\mathrm{ap}(p(bc),\mathrm{ap}(p(ab),x)) : A \to C} \quad \dfrac{\overset{H^{5.3}}{p(q(bc)) : C} \quad \dfrac{\overset{H^{5.4}}{x3 : (\Pi x : A)(\Pi y : A)(x <_a y \to \mathrm{ap}(p(bc),\mathrm{ap}(p(ab),x)) <_c \mathrm{ap}(p(bc),\mathrm{ap}(p(ab),y)))} \quad \overset{H^{5.5}}{x4 : (\Pi x : A)(\mathrm{ap}(p(bc),\mathrm{ap}(p(ab),x)) <_c p(q(bc)))}}{\langle x3, x4\rangle : (\Pi x : A)(\Pi y : A)(x <_a y \to \mathrm{ap}(p(bc),\mathrm{ap}(p(ab),x)) <_c \mathrm{ap}(p(bc),\mathrm{ap}(p(ab),y))) \times (\Pi x : A)(\mathrm{ap}(p(bc),\mathrm{ap}(p(ab),x)) <_c p(q(bc)))}}{\langle p(q(bc)), \langle x3, x4\rangle\rangle : (\Sigma z' : C)[(\Pi x : A)(\Pi y : A)(x <_a y \to \mathrm{ap}(p(bc),\mathrm{ap}(p(ab),x)) <_c \mathrm{ap}(p(bc),\mathrm{ap}(p(ab),y))) \times (\Pi x : A)(\mathrm{ap}(p(bc),\mathrm{ap}(p(ab),x)) <_c z')]}}{\langle (\lambda x)\,\mathrm{ap}(p(bc),\mathrm{ap}(p(ab),x)), \langle p(q(bc)), \langle x3, x4\rangle\rangle\rangle : (\Sigma gf : A \to C)(\Sigma z' : C)[(\Pi x : A)(\Pi y : A)(x <_a y \to \mathrm{ap}(gf,x) <_c \mathrm{ap}(gf,y)) \times (\Pi x : A)(\mathrm{ap}(gf,x) <_c z')] = A^* <_\omega C^*} \ (58)}{(\lambda bc)\langle (\lambda x)\,\mathrm{ap}(p(bc),\mathrm{ap}(p(ab),x)), \langle p(q(bc)), \langle x3, x4\rangle\rangle\rangle : B^* <_\omega C^* \to A^* <_\omega C^*} \ (59)}{(\lambda ab)(\lambda bc)\langle (\lambda x)\,\mathrm{ap}(p(bc),\mathrm{ap}(p(ab),x)), \langle p(q(bc)), \langle x3, x4\rangle\rangle\rangle : A^* <_\omega B^* \to (B^* <_\omega C^* \to A^* <_\omega C^*)}}{(\lambda C^*)(\lambda ab)(\lambda bc)\langle (\lambda x)\,\mathrm{ap}(p(bc),\mathrm{ap}(p(ab),x)), \langle p(q(bc)), \langle x3, x4\rangle\rangle\rangle : (\Pi C^* : \Omega)[A^* <_\omega B^* \to (B^* <_\omega C^* \to A^* <_\omega C^*)]}}{(\lambda B^*)(\lambda C^*)(\lambda ab)(\lambda bc)\langle (\lambda x)\,\mathrm{ap}(p(bc),\mathrm{ap}(p(ab),x)), \langle p(q(bc)), \langle x3, x4\rangle\rangle\rangle : (\Pi B^* : \Omega)(\Pi C^* : \Omega)[A^* <_\omega B^* \to (B^* <_\omega C^* \to A^* <_\omega C^*)]}}{p_\omega \equiv (\lambda A^*)(\lambda B^*)(\lambda C^*)(\lambda ab)(\lambda bc)\langle (\lambda x)\,\mathrm{ap}(p(bc),\mathrm{ap}(p(ab),x)), \langle p(q(bc)), \langle x3, x4\rangle\rangle\rangle : (\Pi A^* : \Omega)(\Pi B^* : \Omega)(\Pi C^* : \Omega)[A^* <_\omega B^* \to (B^* <_\omega C^* \to A^* <_\omega C^*)] = P(\omega, <_\omega)}$$

$H^{5.1}$

$$\dfrac{[ab : (\Sigma f : A \to B)(\Sigma z : B)[(\Pi x' : A)(\Pi y' : A)(x' <_a y' \to \mathrm{ap}(f,x') <_b \mathrm{ap}(f,y')) \times (\Pi x' : A)(\mathrm{ap}(f,x') <_b z)]]^{59}}{p(ab) : A \to B}$$

$H^{5.2}$

$$\dfrac{\dfrac{\dfrac{[bc : (\Sigma g : B \to C)(\Sigma z' : C)[(\Pi x' : B)(\Pi y' : B)(x' <_b y' \to \mathrm{ap}(g,x') <_c \mathrm{ap}(g,y')) \times (\Pi x' : B)(\mathrm{ap}(g,x') <_c z')]]^{58}}{p(bc) : B \to C} \quad \dfrac{\overset{H^{5.1}}{p(ab) : A \to B} \quad [x : A]^{53}}{\mathrm{ap}(p(ab),x) : B}}{\mathrm{ap}(p(bc),\mathrm{ap}(p(ab),x)) : C}}{(\lambda x)\,\mathrm{ap}(p(bc),\mathrm{ap}(p(ab),x)) : A \to C} \ (53)$$

$H^{5.3}$

$$\dfrac{\dfrac{[bc : (\Sigma g : B \to C)(\Sigma z' : C)[(\Pi x' : B)(\Pi y' : B)(x' <_b y' \to \mathrm{ap}(g, x') <_c \mathrm{ap}(g, y')) \times (\Pi x' : B)(\mathrm{ap}(g, x') <_c z')]]^{58}}{q(bc) : (\Sigma z' : C)[(\Pi x' : B)(\Pi y' : B)(x' <_b y' \to \mathrm{ap}(p(bc), x') <_c \mathrm{ap}(p(bc), y')) \times (\Pi x' : B)(\mathrm{ap}(p(bc), x') <_c z')]}}{p(q(bc)) : C}$$

$H^{5.4}$

$$\dfrac{\dfrac{\dfrac{\dfrac{\overset{H^{5.4.1}}{\mathrm{ap}(\mathrm{ap}(p(q(q(bc))), \mathrm{ap}(p(ab), x)), \mathrm{ap}(p(ab), y)) : \mathrm{ap}(p(ab), x) <_b \mathrm{ap}(p(ab), y) \to \mathrm{ap}(p(bc), \mathrm{ap}(p(ab), x)) <_c \mathrm{ap}(p(bc), \mathrm{ap}(p(ab), y))} \quad \dfrac{\overset{H^{5.4.2}}{\mathrm{ap}(\mathrm{ap}(p(q(q(ab))), x), y) : x <_a y \to \mathrm{ap}(p(ab), x) <_b \mathrm{ap}(p(ab), y)} \quad [\triangle : x <_a y]^{54}}{\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p(q(q(ab))), x), y), \triangle) : \mathrm{ap}(p(ab), x) <_b \mathrm{ap}(p(ab), y)}}{\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p(q(q(bc))), \mathrm{ap}(p(ab), x)), \mathrm{ap}(p(ab), y)), \mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p(q(q(ab))), x), y), \triangle)) : \mathrm{ap}(p(bc), \mathrm{ap}(p(ab), x)) <_c \mathrm{ap}(p(bc), \mathrm{ap}(p(ab), y))}}{(\lambda\triangle)\, \mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p(q(q(bc))), \mathrm{ap}(p(ab), x)), \mathrm{ap}(p(ab), y)), \mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p(q(q(ab))), x), y), \triangle)) : x <_a y \to \mathrm{ap}(p(bc), \mathrm{ap}(p(ab), x)) <_c \mathrm{ap}(p(bc), \mathrm{ap}(p(ab), y))}\,(54)}{(\lambda y)(\lambda\triangle)\, \mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p(q(q(bc))), \mathrm{ap}(p(ab), x)), \mathrm{ap}(p(ab), y)), \mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p(q(q(ab))), x), y), \triangle)) : (\Pi y : A)(x <_a y \to \mathrm{ap}(p(bc), \mathrm{ap}(p(ab), x)) <_c \mathrm{ap}(p(bc), \mathrm{ap}(p(ab), y)))}\,(55)}{x3 \equiv (\lambda x)(\lambda y)(\lambda\triangle)\, \mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p(q(q(bc))), \mathrm{ap}(p(ab), x)), \mathrm{ap}(p(ab), y)), \mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p(q(q(ab))), x), y), \triangle)) : (\Pi x : A)(\Pi y : A)(x <_a y \to \mathrm{ap}(p(bc), \mathrm{ap}(p(ab), x)) <_c \mathrm{ap}(p(bc), \mathrm{ap}(p(ab), y)))}\,(56)$$

$H^{5.4.1}$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[bc : (\Sigma g : B \to C)(\Sigma z' : C)[(\Pi x' : B)(\Pi y' : B)(x' <_b y' \to \mathrm{ap}(g, x') <_c \mathrm{ap}(g, y')) \times (\Pi x' : B)(\mathrm{ap}(g, x') <_c z')]]^{58}}{q(bc) : (\Sigma z' : C)[(\Pi x' : B)(\Pi y' : B)(x' <_b y' \to \mathrm{ap}(p(bc), x') <_c \mathrm{ap}(p(bc), y')) \times (\Pi x' : B)(\mathrm{ap}(p(bc), x') <_c z')]}}{q(q(bc)) : (\Pi x' : B)(\Pi y' : B)(x' <_b y' \to \mathrm{ap}(p(bc), x') <_c \mathrm{ap}(p(bc), y')) \times (\Pi x' : B)(\mathrm{ap}(p(bc), x') <_c p(q(bc)))}}{p(q(q(bc))) : (\Pi x' : B)(\Pi y' : B)(x' <_b y' \to \mathrm{ap}(p(bc), x') <_c \mathrm{ap}(p(bc), y'))} \quad \dfrac{\overset{H^{5.1}}{p(ab) : A \to B} \quad [x : A]^{56}}{\mathrm{ap}(p(ab), x) : B}}{\mathrm{ap}(p(q(q(bc))), \mathrm{ap}(p(ab), x)) : (\Pi y' : B)(\mathrm{ap}(p(ab), x) <_b y' \to \mathrm{ap}(p(bc), \mathrm{ap}(p(ab), x)) <_c \mathrm{ap}(p(bc), y'))} \quad \dfrac{\overset{H^{5.1}}{p(ab) : A \to B} \quad [y : A]^{55}}{\mathrm{ap}(p(ab), y) : B}}{\mathrm{ap}(\mathrm{ap}(p(q(q(bc))), \mathrm{ap}(p(ab), x)), \mathrm{ap}(p(ab), y)) : \mathrm{ap}(p(ab), x) <_b \mathrm{ap}(p(ab), y) \to \mathrm{ap}(p(bc), \mathrm{ap}(p(ab), x)) <_c \mathrm{ap}(p(bc), \mathrm{ap}(p(ab), y))}$$

$H^{5.4.2}$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[ab : (\Sigma f : A \to B)(\Sigma z : B)[(\Pi x' : A)(\Pi y' : A)(x' <_a y' \to \mathrm{ap}(f, x') <_b \mathrm{ap}(f, y')) \times (\Pi x' : A)(\mathrm{ap}(f, x') <_b z)]]^{59}}{q(ab) : (\Sigma z : B)[(\Pi x' : A)(\Pi y' : A)(x' <_a y' \to \mathrm{ap}(p(ab), x') <_b \mathrm{ap}(p(ab), y')) \times (\Pi x' : A)(\mathrm{ap}(p(ab), x') <_b z)]}}{q(q(ab)) : (\Pi x' : A)(\Pi y' : A)(x' <_a y' \to \mathrm{ap}(p(ab), x') <_b \mathrm{ap}(p(ab), y')) \times (\Pi x' : A)(\mathrm{ap}(p(ab), x') <_b p(q(ab)))}}{p(q(q(ab))) : (\Pi x' : A)(\Pi y' : A)(x' <_a y' \to \mathrm{ap}(p(ab), x') <_b \mathrm{ap}(p(ab), y'))} \quad [x : A]^{56}}{\mathrm{ap}(p(q(q(ab))), x) : (\Pi y' : A)(x <_a y' \to \mathrm{ap}(p(ab), x) <_b \mathrm{ap}(p(ab), y'))} \quad [y : A]^{55}}{\mathrm{ap}(\mathrm{ap}(p(q(q(ab))), x), y) : x <_a y \to \mathrm{ap}(p(ab), x) <_b \mathrm{ap}(p(ab), y)}$$

$H^{5.5}$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[bc : (\Sigma g : B \to C)(\Sigma z' : C)[(\Pi x' : B)(\Pi y' : B)(x' <_b y' \to \mathrm{ap}(g, x') <_c \mathrm{ap}(g, y')) \times (\Pi x' : B)(\mathrm{ap}(g, x') <_c z')]]^{58}}{q(bc) : (\Sigma z' : C)[(\Pi x' : B)(\Pi y' : B)(x' <_b y' \to \mathrm{ap}(p(bc), x') <_c \mathrm{ap}(p(bc), y')) \times (\Pi x' : B)(\mathrm{ap}(p(bc), x') <_c z')]}}{q(q(bc)) : (\Pi x' : B)(\Pi y' : B)(x' <_b y' \to \mathrm{ap}(p(bc), x') <_c \mathrm{ap}(p(bc), y')) \times (\Pi x' : B)(\mathrm{ap}(p(bc), x') <_c p(q(bc)))}}{q(q(q(bc))) : (\Pi x' : B)(\mathrm{ap}(p(bc), x') <_c p(q(bc)))} \quad \dfrac{\overset{H^{5.1}}{p(ab) : A \to B} \quad [x : A]^{57}}{\mathrm{ap}(p(ab), x) : B}}{\mathrm{ap}(q(q(q(bc))), \mathrm{ap}(p(ab), x)) : \mathrm{ap}(p(bc), \mathrm{ap}(p(ab), x)) <_c p(q(bc))}}{x4 \equiv (\lambda x)\, \mathrm{ap}(q(q(q(bc))), \mathrm{ap}(p(ab), x)) : (\Pi x : A)(\mathrm{ap}(p(bc), \mathrm{ap}(p(ab), x)) <_c p(q(bc)))}\ (57)$$

H^6

We show that Ω cannot contain infinitely descending chains:

Suppose that there is an infinitely descending chain ... $<_\omega A^*_{s(n)} <_\omega A^*_n <_\omega$ ... in Ω, and let $a_n = p(A^*_n)$, $<_{a_n} = p(q(A^*_n))$, and $A_n = T(a_n)$.

$A^*_{s(n)} <_\omega A^*_n$ expresses that we have an order-preserving function $f_n : A_{s(n)} \to A_n$, which we define in $H^{6.5}$, and a range-dominating element $z_n : A_n$, which we define in $H^{6.6}$.

We define in $H^{6.2}$ by recursion functions $R_n : A_n \to A_0$ such that

$$\begin{array}{lcll} \mathrm{ap}(R_0, x) & = & x & \text{for } x : A_0 \\ \mathrm{ap}(R_{s(n)}, x) & = & \mathrm{ap}(R_n, \mathrm{ap}(f_n, x)) & \text{for } x : A_{s(n)} \end{array}$$

where for each $n : N$, $f_n : A_{s(n)} \to A_n$ is the order-preserving function defined in $H^{6.5}$. Thus, for each $n : N$, R_n is order-preserving, which is shown by induction in $H^{6.4.1}$.

In $H^{6.3}$ we then define $F \equiv N \to A_0$ as the composition of R_n and $(\lambda n)z_n$, where $(\lambda n)z_n$ maps each $n : N$ into the range-dominating element $z_n : A_n$. We then get an infinitely descending chain

$$\ldots <_{a_0} \mathrm{ap}(R_{s(n)}, z_{s(n)}) <_{a_0} \mathrm{ap}(R_n, z_n) <_{a_0} \ldots <_{a_0} \mathrm{ap}(R_1, z_1) <_{a_0} z_0$$

or

$$\mathrm{ap}(F, s(n)) <_{a_0} \mathrm{ap}(F, n) <_{a_0} \ldots <_{a_0} \mathrm{ap}(F, 1) <_{a_0} \mathrm{ap}(F, 0)$$

in A_0, which is formalized in $H^{6.4}$. But A_0 does not have infinitely descending chains, as we show in $H^{6.1}$. Thus Ω cannot contain infinitely descending chains.

We define $A_n^* \equiv \mathrm{ap}(g,n)$, where $g : N \to \Omega$, $a_n \equiv p(\mathrm{ap}(g,n)) \equiv p(A_n^*)$, $<_{a_n} \equiv p(q(\mathrm{ap}(g,n))) \equiv p(q(A_n^*))$ and $A_n \equiv T(a_n)$ for all $n : N$.

H^6

$$\dfrac{\dfrac{\dfrac{\dfrac{\overset{H^{6.1}}{q_{a_0} : Q(a_0,<_{a_0})} \quad \overset{H^{6.3}}{F \equiv (\lambda n)\,\mathrm{ap}(R_n, z_n) : N \to A_0}}{\mathrm{ap}(q_{a_0},F) : (\Pi n : N)(\mathrm{ap}(F,s(n)) <_{a_0} \mathrm{ap}(F,n)) \to \bot} \quad \overset{H^{6.4}}{(\lambda n)\,\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(R'_n, \mathrm{ap}(f_n, z_{s(n)})), z_n), \mathrm{ap}(q(q(q(\mathrm{ap}(*,n)))), z_{s(n)})) : (\Pi n : N)(\mathrm{ap}(F,s(n)) <_{a_0} \mathrm{ap}(F,n))}}{\mathrm{ap}(\mathrm{ap}(q_{a_0},F), (\lambda n)\,\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(R'_n, \mathrm{ap}(f_n, z_{s(n)})), z_n), \mathrm{ap}(q(q(q(\mathrm{ap}(*,n)))), z_{s(n)}))) : \bot}}{(\lambda *)\,\mathrm{ap}(\mathrm{ap}(q_{a_0},F), (\lambda n)\,\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(R'_n, \mathrm{ap}(f_n, z_{s(n)})), z_n), \mathrm{ap}(q(q(q(\mathrm{ap}(*,n)))), z_{s(n)}))) : (\Pi n : N)(\mathrm{ap}(g,s(n)) <_\omega \mathrm{ap}(g,n)) \to \bot}\,(78)}{q_\omega \equiv (\lambda g)(\lambda *)\,\mathrm{ap}(\mathrm{ap}(q_{a_0},F), (\lambda n)\,\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(R'_n, \mathrm{ap}(f_n, z_{s(n)})), z_n), \mathrm{ap}(q(q(q(\mathrm{ap}(*,n)))), z_{s(n)}))) : (\Pi f : N \to \Omega)[(\Pi n : N)(\mathrm{ap}(f,s(n)) <_\omega \mathrm{ap}(f,n)) \to \bot] = Q(\omega,<_\omega)}\,(79)$$

$H^{6.1}$

$$\dfrac{\dfrac{\dfrac{\dfrac{[g : N \to \Omega]^{79} \qquad \overline{0 : N}}{A_0^* \equiv \mathrm{ap}(g,0) : (\Sigma a : U)(\Sigma <_a : A \to A \to U)[P(a,<_a) \times Q(a,<_a)]}}{q(A_0^*) : (\Sigma <_{a_0} : A_0 \to A_0 \to U)[P(a_0,<_{a_0}) \times Q(a_0,<_{a_0})]}}{q(q(A_0^*)) : P(a_0,<_{a_0}) \times Q(a_0,<_{a_0})}}{q_{a_0} \equiv q(q(q(A_0^*))) : Q(a_0,<_{a_0})}$$

$H^{6.2}$

$$\dfrac{[n : N]^{69} \quad \dfrac{[z : A_0]^{66}}{(\lambda z)z : A_0 \to A_0}\,(66) \quad \dfrac{\dfrac{[y : A_m \to A_0]^{68} \quad \dfrac{\dfrac{\overset{H^{6.5}}{(\lambda n)f_n : (\Pi n : N)(A_{s(n)} \to A_n)} \quad [m : N]^{68}}{\mathrm{ap}((\lambda n)f_n, m) = f_m : A_{s(m)} \to A_m} \quad [x : A_{s(m)}]^{67}}{\mathrm{ap}(f_m, x) : A_m}}{\mathrm{ap}(y, \mathrm{ap}(f_m,x)) : A_0}}{e(m,y) \equiv (\lambda x)\,\mathrm{ap}(y, \mathrm{ap}(f_m,x)) : A_{s(m)} \to A_0}\,(67)}{R_n \equiv R(n, (\lambda z)z, (m,y)e(m,y)) : A_n \to A_0}\,(68)$$

$H^{6.3}$

$$\dfrac{\dfrac{\stackrel{H^{6.2}}{R_n \equiv R(n,(\lambda z)z,(m,y)e(m,y)) : A_n \to A_0} \qquad \dfrac{\stackrel{H^{6.6}}{(\lambda n)z_n : (\Pi n : N)A_n} \quad [n:N]^{69}}{\mathrm{ap}((\lambda n)z_n, n) = z_n : A_n}}{\mathrm{ap}(R_n, z_n) : A_0}}{F \equiv (\lambda n)\,\mathrm{ap}(R_n, z_n) : N \to A_0}\ (69)$$

$H^{6.4}$

$$\dfrac{\dfrac{\dfrac{\dfrac{H^{6.4.1} \qquad \dfrac{\dfrac{H^{6.5} \quad [n:N]^{77}}{\mathrm{ap}((\lambda n)f_n, n) = f_n : A_{s(n)} \to A_n} \quad \dfrac{H^{6.6} \quad \dfrac{[n:N]^{77}}{s(n):N}}{\mathrm{ap}((\lambda n)z_n, s(n)) = z_{s(n)} : A_{s(n)}}}{\mathrm{ap}(f_n, z_{s(n)}) : A_n}}{\mathrm{ap}(R'_n, \mathrm{ap}(f_n, z_{s(n)})) : (\Pi y : A_n)[\mathrm{ap}(f_n, z_{s(n)}) <_{a_n} y \to \mathrm{ap}(R_n, \mathrm{ap}(f_n, z_{s(n)})) <_{a_0} \mathrm{ap}(R_n, y)]} \quad \dfrac{H^{6.6} \quad [n:N]^{77}}{\mathrm{ap}((\lambda n)z_n, n) = z_n : A_n}}{\mathrm{ap}(\mathrm{ap}(R'_n, \mathrm{ap}(f_n, z_{s(n)})), z_n) : \mathrm{ap}(f_n, z_{s(n)}) <_{a_n} z_n \to \mathrm{ap}(R_n, \mathrm{ap}(f_n, z_{s(n)})) <_{a_0} \mathrm{ap}(R_n, z_n)} \quad H^{6.4.2}}{\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(R'_n, \mathrm{ap}(f_n, z_{s(n)})), z_n), \mathrm{ap}(q(q(q(\mathrm{ap}(*, n)))), z_{s(n)})) : \mathrm{ap}(R_n, \mathrm{ap}(f_n, z_{s(n)})) <_{a_0} \mathrm{ap}(R_n, z_n)} \quad \stackrel{H^{6.4.3}}{\mathrm{ap}(R_n, \mathrm{ap}(f_n, z_{s(n)})) = \mathrm{ap}(F, s(n)) : A_0} \quad \stackrel{H^{6.4.4}}{\mathrm{ap}(R_n, z_n) = \mathrm{ap}(F, n) : A_0}}{\dfrac{\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(R'_n, \mathrm{ap}(f_n, z_{s(n)})), z_n), \mathrm{ap}(q(q(q(\mathrm{ap}(*, n)))), z_{s(n)})) : \mathrm{ap}(F, s(n)) <_{a_0} \mathrm{ap}(F, n)}{(\lambda n)\,\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(R'_n, \mathrm{ap}(f_n, z_{s(n)})), z_n), \mathrm{ap}(q(q(q(\mathrm{ap}(*, n)))), z_{s(n)})) : (\Pi n : N)(\mathrm{ap}(F, s(n)) <_{a_0} \mathrm{ap}(F, n))}\ (77)}$$

$H^{6.4.1}$

$$\dfrac{[n:N]^{77} \quad \stackrel{H^{6.4.1.1}}{(\lambda x)(\lambda y)(\lambda o)o : (\Pi x : A_0)(\Pi y : A_0)[x <_{a_0} y \to \mathrm{ap}(R_0, x) <_{a_0} \mathrm{ap}(R_0, y)]} \quad \stackrel{H^{6.4.1.2}}{e'(m, \mathrm{ind}) : (\Pi x : A_{s(m)})(\Pi y : A_{s(m)})[x <_{a_{s(m)}} y \to \mathrm{ap}(R_{s(m)}, x) <_{a_0} \mathrm{ap}(R_{s(m)}, y)]}}{R'_n \equiv R'(n, (\lambda x)(\lambda y)(\lambda o)o, (m, \mathrm{ind})e'(m, \mathrm{ind})) : (\Pi x : A_n)(\Pi y : A_n)[x <_{a_n} y \to \mathrm{ap}(R_n, x) <_{a_0} \mathrm{ap}(R_n, y)]}\ (76)$$

$H^{6.4.1.1}$

$$
\dfrac{[o : x <_{a_0} y]^{70} \quad \dfrac{\dfrac{\dfrac{\dfrac{\overset{H^{6.2}}{R(n, (\lambda z)z, (m, y)e(m, y)) : A_n \to A_0}}{(\lambda n)R(n, (\lambda z)z, (m, y)e(m, y)) : (\Pi n : N)(A_n \to A_0)}\,(69) \quad \dfrac{}{0 : N}}{\mathrm{ap}((\lambda n)R(n, (\lambda z)z, (m, y)e(m, y)), 0) = R(0, (\lambda z)z, (m, y)e(m, y)) = (\lambda z)z : A_0 \to A_0} \quad [x : A_0]^{71}}{\mathrm{ap}(R(0, (\lambda z)z, (m, y)e(m, y)), x) = x : A_0} \quad \dfrac{\dfrac{\dfrac{\overset{H^{6.2}}{R(n, (\lambda z)z, (m, y)e(m, y)) : A_n \to A_0}}{(\lambda n)R(n, (\lambda z)z, (m, y)e(m, y)) : (\Pi n : N)(A_n \to A_0)}\,(69) \quad \dfrac{}{0 : N}}{\mathrm{ap}((\lambda n)R(n, (\lambda z)z, (m, y)e(m, y)), 0) = R(0, (\lambda z)z, (m, y)e(m, y)) = (\lambda z)z : A_0 \to A_0} \quad [y : A_0]^{72}}{\mathrm{ap}(R(0, (\lambda z)z, (m, y)e(m, y)), y) = y : A_0}}{\dfrac{o : \mathrm{ap}(R(0, (\lambda z)z, (m, y)e(m, y)), x) <_{a_0} \mathrm{ap}(R(0, (\lambda z)z, (m, y)e(m, y)), y)}{\dfrac{(\lambda o)o : x <_{a_0} y \to \mathrm{ap}(R(0, (\lambda z)z, (m, y)e(m, y)), x) <_{a_0} \mathrm{ap}(R(0, (\lambda z)z, (m, y)e(m, y)), y)}{\dfrac{(\lambda y)(\lambda o)o : (\Pi y : A_0)[x <_{a_0} y \to \mathrm{ap}(R(0, (\lambda z)z, (m, y)e(m, y)), x) <_{a_0} \mathrm{ap}(R(0, (\lambda z)z, (m, y)e(m, y)), y)]}{(\lambda x)(\lambda y)(\lambda o)o : (\Pi x : A_0)(\Pi y : A_0)[x <_{a_0} y \to \mathrm{ap}(R(0, (\lambda z)z, (m, y)e(m, y)), x) <_{a_0} \mathrm{ap}(R(0, (\lambda z)z, (m, y)e(m, y)), y)]}\,(72)}\,(71)}\,(70)}
$$

$H^{6.4.1.2}$

$$
\dfrac{\overset{H^{6.4.1.2.1}}{x5 : \mathrm{ap}(R_m, \mathrm{ap}(f_m, w)) <_{a_0} \mathrm{ap}(R_m, \mathrm{ap}(f_m, v))} \quad \overset{H^{6.4.5}}{\mathrm{ap}(R_{s(m)}, w) = \mathrm{ap}(R_m, \mathrm{ap}(f_m, w)) : A_0} \quad \overset{H^{6.4.6}}{\mathrm{ap}(R_{s(m)}, v) = \mathrm{ap}(R_m, \mathrm{ap}(f_m, v)) : A_0}}{\dfrac{x5 : \mathrm{ap}(R_{s(m)}, w) <_{a_0} \mathrm{ap}(R_{s(m)}, v)}{\dfrac{(\lambda \mathrm{ov})x5 : w <_{a_{s(m)}} v \to \mathrm{ap}(R_{s(m)}, w) <_{a_0} \mathrm{ap}(R_{s(m)}, v)}{\dfrac{(\lambda v)(\lambda \mathrm{ov})x5 : (\Pi y : A_{s(m)})[w <_{a_{s(m)}} y \to \mathrm{ap}(R_{s(m)}, w) <_{a_0} \mathrm{ap}(R_{s(m)}, y)]}{e'(m, \mathrm{ind}) \equiv (\lambda w)(\lambda v)(\lambda \mathrm{ov})x5 : (\Pi x : A_{s(m)})(\Pi y : A_{s(m)})[x <_{a_{s(m)}} y \to \mathrm{ap}(R_{s(m)}, x) <_{a_0} \mathrm{ap}(R_{s(m)}, y)]}\,(75)}\,(74)}\,(73)}
$$

$H^{6.4.1.2.1}$

$$
\dfrac{\dfrac{\dfrac{[\mathrm{ind} : (\Pi x : A_m)(\Pi y : A_m)[x <_{a_m} y \to \mathrm{ap}(R_m, x) <_{a_0} \mathrm{ap}(R_m, y)]]^{76} \quad \dfrac{\dfrac{H^{6.5} \quad [m : N]^{76}}{\mathrm{ap}((\lambda n)f_n, m) = f_m : A_{s(m)} \to A_m} \quad [w : A_{s(m)}]^{75}}{\mathrm{ap}(f_m, w) : A_m}}{\mathrm{ap}(\mathrm{ind}, \mathrm{ap}(f_m, w)) : (\Pi y : A_m)[\mathrm{ap}(f_m, w) <_{a_m} y) \to \mathrm{ap}(R_m, \mathrm{ap}(f_m, w)) <_{a_0} \mathrm{ap}(R_m, y)]} \quad \dfrac{\dfrac{H^{6.5} \quad [m : N]^{76}}{\mathrm{ap}((\lambda n)f_n, m) = f_m : A_{s(m)} \to A_m} \quad [v : A_{s(m)}]^{74}}{\mathrm{ap}(f_m, v) : A_m}}{\mathrm{ap}(\mathrm{ap}(\mathrm{ind}, \mathrm{ap}(f_m, w)), \mathrm{ap}(f_m, v)) : \mathrm{ap}(f_m, w) <_{a_m} \mathrm{ap}(f_m, v) \to \mathrm{ap}(R_m, \mathrm{ap}(f_m, w)) <_{a_0} \mathrm{ap}(R_m, \mathrm{ap}(f_m, v))} \quad H^{6.4.1.2.1.1}}{x5 \equiv \mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{ind}, \mathrm{ap}(f_m, w)), \mathrm{ap}(f_m, v)), \mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p(q(q(\mathrm{ap}(*, m)))), w), v), \mathrm{ov})) : \mathrm{ap}(R_m, \mathrm{ap}(f_m, w)) <_{a_0} \mathrm{ap}(R_m, \mathrm{ap}(f_m, v))}
$$

$H^{6.4.1.2.1.1}$

$$\dfrac{\dfrac{\dfrac{\dfrac{H^{6.4.1.2.1.1.2}}{p(q(q(\mathrm{ap}(*,m)))) : (\Pi x : A_{s(m)})(\Pi y : A_{s(m)})(x <_{a_{s(m)}} y \to \mathrm{ap}(f_m, x) <_{a_m} \mathrm{ap}(f_m, y))} \quad [w : A_{s(m)}]^{75}}{\mathrm{ap}(p(q(q(\mathrm{ap}(*,m)))), w) : (\Pi y : A_{s(m)})(w <_{a_{s(m)}} y \to \mathrm{ap}(f_m, w) <_{a_m} \mathrm{ap}(f_m, y))} \quad [v : A_{s(m)}]^{74}}{\mathrm{ap}(\mathrm{ap}(p(q(q(\mathrm{ap}(*,m)))), w), v) : w <_{a_{s(m)}} v \to \mathrm{ap}(f_m, w) <_{a_m} \mathrm{ap}(f_m, v)} \quad [\mathrm{ov} : w <_{a_{s(m)}} v]^{73}}{\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(p(q(q(\mathrm{ap}(*,m)))), w), v), \mathrm{ov}) : \mathrm{ap}(f_m, w) <_{a_m} \mathrm{ap}(f_m, v)}$$

$H^{6.4.1.2.1.1.2}$

$$\dfrac{\dfrac{\dfrac{[* : (\Pi n : N)(A^*_{s(n)} <_\omega A^*_n)]^{78} \qquad [m : N]^{76}}{\mathrm{ap}(*, m) : (\Sigma f : A_{s(m)} \to A_m)(\Sigma z : A_m)[(\Pi x : A_{s(m)})(\Pi y : A_{s(m)})(x <_{a_{s(m)}} y \to \mathrm{ap}(f, x) <_{a_m} \mathrm{ap}(f, y)) \times (\Pi x : A_{s(m)})(\mathrm{ap}(f, x) <_{a_m} z)]}}{q(\mathrm{ap}(*, m)) : (\Sigma z : A_m)[(\Pi x : A_{s(m)})(\Pi y : A_{s(m)})(x <_{a_{s(m)}} y \to \mathrm{ap}(f_m, x) <_{a_m} \mathrm{ap}(f_m, y)) \times (\Pi x : A_{s(m)})(\mathrm{ap}(f_m, x) <_{a_m} z)]}}{q(q(\mathrm{ap}(*, m))) : (\Pi x : A_{s(m)})(\Pi y : A_{s(m)})(x <_{a_{s(m)}} y \to \mathrm{ap}(f_m, x) <_{a_m} \mathrm{ap}(f_m, y)) \times (\Pi x : A_{s(m)})(\mathrm{ap}(f_m, x) <_{a_m} z_m)}$$

$H^{6.4.2}$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[* : (\Pi n : N)(A^*_{s(n)} <_\omega A^*_n)]^{78} \qquad [n : N]^{77}}{\mathrm{ap}(*, n) : (\Sigma f : A_{s(n)} \to A_n)(\Sigma z : A_n)[(\Pi x : A_{s(n)})(\Pi y : A_{s(n)})(x <_{a_{s(n)}} y \to \mathrm{ap}(f, x) <_{a_n} \mathrm{ap}(f, y)) \times (\Pi x : A_{s(n)})(\mathrm{ap}(f, x) <_{a_n} z)]}}{q(\mathrm{ap}(*, n)) : (\Sigma z : A_n)[(\Pi x : A_{s(n)})(\Pi y : A_{s(n)})(x <_{a_{s(n)}} y \to \mathrm{ap}(f_n, x) <_{a_n} \mathrm{ap}(f_n, y)) \times (\Pi x : A_{s(n)})(\mathrm{ap}(f_n, x) <_{a_n} z)]}}{q(q(\mathrm{ap}(*, n))) : (\Pi x : A_{s(n)})(\Pi y : A_{s(n)})(x <_{a_{s(n)}} y \to \mathrm{ap}(f_n, x) <_{a_n} \mathrm{ap}(f_n, y)) \times (\Pi x : A_{s(n)})(\mathrm{ap}(f_n, x) <_{a_n} z_n)}}{q(q(q(\mathrm{ap}(*, n)))) : (\Pi x : A_{s(n)})(\mathrm{ap}(f_n, x) <_{a_n} z_n)} \quad \dfrac{\dfrac{H^{6.6}}{(\lambda n)z_n : (\Pi n : N)A_n} \quad \dfrac{[n : N]^{77}}{s(n) : N}}{\mathrm{ap}((\lambda n)z_n, s(n)) = z_{s(n)} : A_{s(n)}}}{\mathrm{ap}(q(q(q(\mathrm{ap}(*, n)))), z_{s(n)}) : \mathrm{ap}(f_n, z_{s(n)}) <_{a_n} z_n}$$

In $H^{6.2}$ we have $e(m,y) \equiv (\lambda x)\,\mathrm{ap}(y, \mathrm{ap}(f_m, x))$.

According to N-equality we have $R_{s(n)} \equiv R(s(n), (\lambda z)z, (m,y)e(m,y)) = e(n, R(n, (\lambda z)z, (m,y)e(m,y))) = e(n, R_n) = (\lambda x)\,\mathrm{ap}(R_n, \mathrm{ap}(f_n, x))$

because we have $F \equiv (\lambda n)\,\mathrm{ap}(R_n, z_n)$, we get:

$\mathrm{ap}(F, n) = \mathrm{ap}((\lambda n)\,\mathrm{ap}(R_n, z_n), n) = \mathrm{ap}(R_n, z_n)$

$\mathrm{ap}(F, s(n)) = \mathrm{ap}((\lambda n)\,\mathrm{ap}(R_n, z_n), s(n)) = \mathrm{ap}(R_{s(n)}, z_{s(n)}) = \mathrm{ap}((\lambda x)\,\mathrm{ap}(R_n, \mathrm{ap}(f_n, x)), z_{s(n)}) = \mathrm{ap}(R_n, \mathrm{ap}(f_n, z_{s(n)}))$

Formally, we get:

$H^{6.4.3}$

$$\dfrac{\overset{H^{6.3}}{F \equiv (\lambda n)\,\mathrm{ap}(R_n, z_n) : N \to A_0} \quad \dfrac{[n : N]^{77}}{s(n) : N}}{\mathrm{ap}(F, s(n)) = \mathrm{ap}(R_n, \mathrm{ap}(f_n, z_{s(n)})) : A_0}$$

$H^{6.4.4}$

$$\dfrac{\overset{H^{6.3}}{F \equiv (\lambda n)\,\mathrm{ap}(R_n, z_n) : N \to A_0} \quad [n : N]^{77}}{\mathrm{ap}(F, n) = \mathrm{ap}(R_n, z_n) : A_0}$$

For $v : A_{s(m)}$ and $w : A_{s(m)}$ it also follows
$\mathrm{ap}(R_{s(m)}, v) = \mathrm{ap}((\lambda x)\,\mathrm{ap}(R_m, \mathrm{ap}(f_m, x)), v) = \mathrm{ap}(R_m, \mathrm{ap}(f_m, v))$
and
$\mathrm{ap}(R_{s(m)}, w) = \mathrm{ap}((\lambda x)\,\mathrm{ap}(R_m, \mathrm{ap}(f_m, x)), w) = \mathrm{ap}(R_m, \mathrm{ap}(f_m, w))$

Formally, we get:

$H^{6.4.5}$

$$\dfrac{\dfrac{\dfrac{\overset{H^{6.2}}{R_n : A_n \to A_0}}{(\lambda n)R_n : (\Pi n : N)(A_n \to A_0)}\,(69) \quad \dfrac{[m : N]^{76}}{s(m) : N}}{\mathrm{ap}((\lambda n)R_n, s(m)) = R_{s(m)} : A_{s(m)} \to A_0} \quad [w : A_{s(m)}]^{75}}{\mathrm{ap}(R_{s(m)}, w) = \mathrm{ap}(R_m, \mathrm{ap}(f_m, w)) : A_0}$$

$H^{6.4.6}$

$$\dfrac{\dfrac{\dfrac{\overset{H^{6.2}}{R_n : A_n \to A_0}}{(\lambda n)R_n : (\Pi n : N)(A_n \to A_0)}\,(69) \quad \dfrac{[m : N]^{76}}{s(m) : N}}{\mathrm{ap}((\lambda n)R_n, s(m)) = R_{s(m)} : A_{s(m)} \to A_0} \quad [v : A_{s(m)}]^{74}}{\mathrm{ap}(R_{s(m)}, v) = \mathrm{ap}(R_m, \mathrm{ap}(f_m, v)) : A_0}$$

$H^{6.5}$

$$\cfrac{\cfrac{[* : (\Pi n : N)(A^*_{s(n)} <_\omega A^*_n)]^{78} \qquad [n : N]^{63}}{\mathrm{ap}(*, n) : (\Sigma f : A_{s(n)} \to A_n)(\Sigma z : A_n)[(\Pi x : A_{s(n)})(\Pi y : A_{s(n)})(x <_{a_{s(n)}} y \to \mathrm{ap}(f, x) <_{a_n} \mathrm{ap}(f, y)) \times (\Pi x : A_{s(n)})(\mathrm{ap}(f, x) <_{a_n} z)]}}{\cfrac{f_n \equiv p(\mathrm{ap}(*, n)) : A_{s(n)} \to A_n}{(\lambda n) f_n : (\Pi n : N)(A_{s(n)} \to A_n)}\ (63)}$$

$H^{6.6}$

$$\cfrac{\cfrac{\cfrac{[* : (\Pi n : N)(A^*_{s(n)} <_\omega A^*_n)]^{78} \qquad [n : N]^{64}}{\mathrm{ap}(*, n) : (\Sigma f : A_{s(n)} \to A_n)(\Sigma z : A_n)[(\Pi x : A_{s(n)})(\Pi y : A_{s(n)})(x <_{a_{s(n)}} y \to \mathrm{ap}(f, x) <_{a_n} \mathrm{ap}(f, y)) \times (\Pi x : A_{s(n)})(\mathrm{ap}(f, x) <_{a_n} z)]}}{q(\mathrm{ap}(*, n)) : (\Sigma z : A_n)[(\Pi x : A_{s(n)})(\Pi y : A_{s(n)})(x <_{a_{s(n)}} y \to \mathrm{ap}(f_n, x) <_{a_n} \mathrm{ap}(f_n, y)) \times (\Pi x : A_{s(n)})(\mathrm{ap}(f_n, x) <_{a_n} z)]}}{\cfrac{z_n \equiv p(q(\mathrm{ap}(*, n))) : A_n}{(\lambda n) z_n : (\Pi n : N) A_n}\ (64)}$$

5 Coquand's paradox of trees

The paradox of trees can be seen as a version of Russell's paradox, where sets are replaced by trees and sets that do not contain themselves by normal trees.

A tree can be formed by linking together a collection of leaves or already existing trees under a common root. Each member of the collection is an immediate subtree of this new tree. We define a normal tree to be a tree that is not equal to any of its immediate subtrees.

The paradox is constructed in the following way:[18]

Suppose that we have a collection of all normal trees. Let r be the tree that is formed by linking together all normal trees under a common root. Is r a normal tree?
If r is a normal tree, then r is a member of the collection of all normal trees and thus an immediate subtree of r. This contradicts the assumption of r being normal.
Suppose r is not normal, i.e. r is equal to one of its immediate subtrees. Then r must be normal, because all immediate subtrees of r are normal, which again leads to a contradiction.

In the formalization of the paradox of trees we essentially need the rules

$$u : U, \qquad T(u) = U$$

to be able to form the tree of all normal trees.

We let $\perp$ denote the empty set, which has the name $\hat{\perp}$ in U. Inconsistency means that $\perp$ contains an element.

Let a be the name of a set A. We define negation by

$$\neg a \equiv \pi(a, (x)\hat{\perp})[a : U]$$

so that

$$T(\neg a) = T(\pi(a, (x)\hat{\perp})) = (\Pi x : T(a))T(\hat{\perp}) = T(a) \to T(\hat{\perp}) = \neg T(a)$$

We define Tree as the well-ordering in which U represents the different ways to form a tree and the decoding function T represents the parts of a tree formed by an element in U.

$$\text{tree} \equiv w(u, (t)t) : U, \qquad \text{Tree} = T(\text{tree}) = (W\ t : U)T(t)$$

In order to express the property of being normal we define

$$\text{normal}(t) \equiv \text{wrec}(t, (x, f, c)\pi(x, (y)\pi(\widehat{\text{Id}}(\text{tree}, f(y), \text{sup}(x, f)), (z)\hat{\perp})))$$

so that

$$T(\text{normal}(\text{sup}(a, b))) = T(\text{wrec}(\text{sup}(a, b), (x, f, c)\pi(x, (y)\pi(\widehat{\text{Id}}(\text{tree}, f(y), \text{sup}(x, f)), (z)\hat{\perp}))))$$

$$= T(\pi(a, (y)\pi(\widehat{\text{Id}}(\text{tree}, b(y), \text{sup}(a, b)), (z)\hat{\perp}))) = (\Pi y : T(a))\neg\,\text{Id}(\text{Tree}, b(y), \text{sup}(a, b))$$

For better readability we here use the notations b and f instead of $(y)b(y)$ and $(y)f(y)$.

In D^3 we formally show $\text{normal}(t) : U[t : \text{Tree}]$.
Then we define the set of normal trees as

$$nt \equiv \sigma(\text{tree}, (x)\,\text{normal}(x)) : U$$

so that

$$T(nt) = T(\sigma(\text{tree}, (x)\,\text{normal}(x))) = (\Sigma x : \text{Tree})T(\text{normal}(x))$$

In D^2 we then define the tree of all normal trees as $\text{sup}(nt, p) : \text{Tree}$, where p is the left projection.
In D^1 we prove that $\text{sup}(nt, p)$ is normal by constructing an element of the set $T(\text{normal}(\text{sup}(nt, p)))$.
But because $\text{sup}(nt, p)$ itself is a member of the set of normal trees, which contains all the parts $\text{sup}(nt, p)$ consists of, $\text{sup}(nt, p)$ must be equal to one of its subtrees, as we show in D, and we get the contradiction.

[18] We follow the idea presented by Coquand 1992

D

$$\dfrac{\dfrac{\overset{D^1}{\mathrm{lemma} : (\Pi y : T(nt))(\mathrm{Id}(\mathrm{Tree}, p(y), \mathrm{sup}(nt,p)) \to \bot)} \quad \dfrac{\overset{D^2}{\mathrm{sup}(nt,p) : \mathrm{Tree}} \quad \overset{D^1}{\mathrm{lemma} : T(\mathrm{normal}(\mathrm{sup}(nt,p)))}}{\langle \mathrm{sup}(nt,p), \mathrm{lemma}\rangle : T(nt)}}{\mathrm{ap}(\mathrm{lemma}, \langle \mathrm{sup}(nt,p), \mathrm{lemma}\rangle) : \mathrm{Id}(\mathrm{Tree}, p(\langle \mathrm{sup}(nt,p), \mathrm{lemma}\rangle), \mathrm{sup}(nt,p)) \to \bot} \quad \dfrac{\dfrac{\dfrac{\overset{D^2}{\mathrm{sup}(nt,p) : \mathrm{Tree}} \quad \overset{D^1}{\mathrm{lemma} : T(\mathrm{normal}(\mathrm{sup}(nt,p)))}}{\langle \mathrm{sup}(nt,p), \mathrm{lemma}\rangle : T(nt)}}{p(\langle \mathrm{sup}(nt,p), \mathrm{lemma}\rangle) = \mathrm{sup}(nt,p) : \mathrm{Tree}}}{\mathrm{ref}(\mathrm{sup}(nt,p)) : \mathrm{Id}(\mathrm{Tree}, p(\langle \mathrm{sup}(nt,p), \mathrm{lemma}\rangle), \mathrm{sup}(nt,p))}}{\mathrm{russell} \equiv \mathrm{ap}(\mathrm{ap}(\mathrm{lemma}, \langle \mathrm{sup}(nt,p), \mathrm{lemma}\rangle), \mathrm{ref}(\mathrm{sup}(nt,p))) : \bot}$$

D^1

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\overset{D^{1.1}}{\mathrm{idpeel}(z,(x)(\lambda v)v) : T(\mathrm{normal}(p(y)) \to T(\mathrm{normal}(\mathrm{sup}(nt,p)))} \quad \dfrac{[y : T(nt)]^{10}}{q(y) : T(\mathrm{normal}(p(y)))}}{\mathrm{ap}(\mathrm{idpeel}(z,(x)(\lambda v)v), q(y)) : (\Pi y : T(nt))(\mathrm{Id}(\mathrm{Tree}, p(y), \mathrm{sup}(nt,p)) \to \bot)} \quad [y : T(nt)]^{10}}{\mathrm{ap}(\mathrm{ap}(\mathrm{idpeel}(z,(x)(\lambda v)v), q(y)), y) : \mathrm{Id}(\mathrm{Tree}, p(y), \mathrm{sup}(nt,p)) \to \bot} \quad [z : \mathrm{Id}(\mathrm{Tree}, p(y), \mathrm{sup}(nt,p))]^9}{\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{idpeel}(z,(x)(\lambda v)v), q(y)), y), z) : \bot}}{(\lambda z)\,\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{idpeel}(z,(x)(\lambda v)v), q(y)), y), z) : \mathrm{Id}(\mathrm{Tree}, p(y), \mathrm{sup}(nt,p)) \to \bot}\,(9)}{\mathrm{lemma} \equiv (\lambda y)(\lambda z)\,\mathrm{ap}(\mathrm{ap}(\mathrm{ap}(\mathrm{idpeel}(z,(x)(\lambda v)v), q(y)), y), z) : (\Pi y : T(nt))(\mathrm{Id}(\mathrm{Tree}, p(y), \mathrm{sup}(nt,p)) \to \bot)}\,(10)$$

$D^{1.1}$

$$\dfrac{\dfrac{[y : T(nt)]^{10}}{p(y) : \mathrm{Tree}} \quad \overset{D^2}{\mathrm{sup}(nt,p) : \mathrm{Tree}} \quad [z : \mathrm{Id}(\mathrm{Tree}, p(y), \mathrm{sup}(nt,p))]^9 \quad \overset{D^{1.1.1}}{T(\mathrm{normal}(x)) \to T(\mathrm{normal}(y)) : \mathrm{set}} \quad \dfrac{[v : T(\mathrm{normal}(x))]^7 \quad \dfrac{\dfrac{\overset{D^3}{(\lambda a)\,\mathrm{normal}(a) : \mathrm{Tree} \to U} \quad [x : \mathrm{Tree}]^8}{\mathrm{ap}((\lambda a)\,\mathrm{normal}(a), x) = \mathrm{normal}(x) : U}}{T(\mathrm{normal}(x)) : \mathrm{set}}}{(\lambda v)v : T(\mathrm{normal}(x)) \to T(\mathrm{normal}(x))}\,(7)}{\mathrm{idpeel}(z,(x)(\lambda v)v) : T(\mathrm{normal}(p(y))) \to T(\mathrm{normal}(\mathrm{sup}(nt,p)))}\,(8)$$

$D^{1.1.1}$

$$\dfrac{\dfrac{\dfrac{\overset{D^3}{(\lambda a)\,\mathrm{normal}(a) : \mathrm{Tree} \to U} \quad [x : \mathrm{Tree}]^8}{\mathrm{ap}((\lambda a)\,\mathrm{normal}(a), x) = \mathrm{normal}(x) : U}}{T(\mathrm{normal}(x)) : \mathrm{set}} \quad \dfrac{\dfrac{\overset{D^3}{(\lambda a)\,\mathrm{normal}(a) : \mathrm{Tree} \to U} \quad [y : \mathrm{Tree}]^8}{\mathrm{ap}((\lambda a)\,\mathrm{normal}(a), y) = \mathrm{normal}(y) : U}}{T(\mathrm{normal}(y)) : \mathrm{set}}}{T(\mathrm{normal}(x)) \to T(\mathrm{normal}(y)) : \mathrm{set}}$$

D^2

$$\dfrac{\dfrac{\dfrac{\dfrac{}{u : U} \quad \dfrac{[t : T(u)]^1 \quad \dfrac{}{T(u) = U : \mathrm{set}}}{t : U}}{\mathrm{tree} \equiv w(u, (t)t) : U}\,(1) \quad \dfrac{\overset{D^3}{(\lambda a)\,\mathrm{normal}(a) : \mathrm{Tree} \to U} \quad [x : \mathrm{Tree}]^5}{\mathrm{ap}((\lambda a)\,\mathrm{normal}(a), x) = \mathrm{normal}(x) : U}}{nt \equiv \sigma(\mathrm{tree}, (x)\,\mathrm{normal}(x)) : U}\,(5) \quad \dfrac{[y : T(nt)]^6}{p(y) : \mathrm{Tree}}}{\sup(nt, p) : \mathrm{Tree}}\,(6)$$

D^3

$$\dfrac{[a : (W\ t : U)T(t)]^4 \quad \dfrac{}{U : \mathrm{set}} \quad \dfrac{[x : U]^3 \quad \dfrac{\dfrac{\dfrac{\dfrac{}{u : U} \quad \dfrac{[t : T(u)]^1 \quad \dfrac{}{T(u) = U : \mathrm{set}}}{t : U}}{\mathrm{tree} \equiv w(u, (t)t) : U}\,(1) \quad \dfrac{[y : T(x)]^2 \quad [f(y) : (W\ t : U)T(t)[y : T(x)]]^3}{[f(y) : (W\ t : U)T(t)]^3} \quad \dfrac{[x : U]^3 \quad [f(y) : (W\ t : U)T(t)[y : T(x)]]^3}{\sup(x, f) : (W\ t : U)T(t)}}{\widehat{\mathrm{Id}}(\mathrm{tree}, f(y), \sup(x, f)) : U} \quad \dfrac{}{\hat{\bot} : U}}{\pi(\widehat{\mathrm{Id}}(\mathrm{tree}, f(y), \sup(x, f)), (z)\hat{\bot}) : U}}{\pi(x, (y)\pi(\widehat{\mathrm{Id}}(\mathrm{tree}, f(y), \sup(x, f)), (z)\hat{\bot})) : U}\,(2)}{\dfrac{\mathrm{normal}(a) \equiv \mathrm{wrec}(a, (x, f, c)\pi(x, (y)\pi(\widehat{\mathrm{Id}}(\mathrm{tree}, f(y), \sup(x, f)), (z)\hat{\bot}))) : U}{(\lambda a)\,\mathrm{normal}(a) : \mathrm{Tree} \to U}\,(4)}\,(3)$$

6 Conclusion

The later, presumably consistent version of Martin-Löf's type theory does not contain the rules $u : U$, $T(u) = U$ anymore. Instead, a second universe U' is introduced by

$$\frac{}{u : U'} \qquad \frac{}{T'(u) = U : \text{set}}$$

$$\frac{a : U}{t(a) : U'} \qquad \frac{a : U}{T'(t(a)) = T(a) : \text{set}}$$

where t gives every element in U a name in U' and T' decodes the elements of U' into the sets they represent.
In the same way we can introduce a third universe U'' and iterate the process.[19]
Now the name of each universe is no longer contained in the universe itself but in the universe of one level higher. Thus the constructions we presented cannot be performed anymore:

Girard's paradox is based on the assumption that Ω has a name in U, where Ω is defined as the set of all sets that have a name in U and the properties P and Q, that is, as the set of all ordinal numbers. Without the rules $u : U$, $T(u) = U$, Ω does not have a name in U and can therefore not contain itself. $\Omega^* : \Omega$ can no longer be proved and the proof of absurdity that is based on $\Omega^* : \Omega$ can no longer be constructed.

Coquand's paradox of trees cannot be formalized without the rules $u : U$, $T(u) = U$, because with our definition of $\text{Tree} = (W\ t : U)T(t)$ we cannot construct $\text{sup}(nt, p)$ as the tree of all normal trees. The rules $u : U$, $T(u) = U$ are essentially needed in forming the name for the set of normal trees $nt = \sigma(\text{tree}, (x)\,\text{normal}(x))$, because only with these rules Tree has a name $\text{tree} = w(u, (t)t)$ in U.

On the background of the presumably consistent version of type theory the paradoxes we presented are prohibited:
A tree of all normal trees does not exist, neither does a greatest ordinal number, i.e. a set of all ordinal numbers, where an ordinal number is defined as the set of all preceding ordinals.
The same solution can be found for the village-barber paradox: There does not exist such a barber.
Finally, for Russell's paradox we draw the conclusion that there cannot be a set of all sets that do not contain themselves.

[19] Martin-Löf 1984, p. 89

References

Per Martin-Löf, *Intuitionistic Type Theory*, Bibliopolis, Napoli, 1984.

Aarne Ranta, *Type-Theoretical Grammar*, Clarendon Press, Oxford, 1994.

Bengt Nordström, Kent Petersson, Jan M. Smith, *Programming in Martin-Löf's Type Theory. An Introduction*, Oxford University Press, 1990.

Per Martin-Löf, An intuitionistic theory of types. In Giovanni Sambin, Jan Smith, eds, *Twenty-Five Years of Constructive Type Theory*, Claredon Press, Oxford, 1998.

Erik Palmgren, On universes in type theory. In Giovanni Sambin, Jan Smith, eds, *Twenty-Five Years of Constructive Type Theory*, Claredon Press, Oxford, 1998.

A.S. Troelstra, D. van Dalen, *Constructivism in Mathematics, Volume II*, North-Holland, Amsterdam, 1988.

Encyclopedia of Mathematics, Kluwer Academic Publishers, 1993.

v. Mangoldt/Knopp, *Einführung in die Höhere Mathematik, Vierter Band*, S. Hirzel Verlag, Stuttgart, 1973.

Jean van Heijenoort, *From Frege to Gödel. A Source Book in Mathematical Logic, 1879-1931*, Harvard University Press, Cambridge, 1967.

Charles C. Pinter, *Set Theory*, Addison-Wesley Publishing Company, Bucknell University, 1971.

Paul R. Halmos, *Naive Set Theory*, Springer-Verlag, New York, 1974.

Thierry Coquand, *The paradox of trees in type theory*, BIT 32, no. 1, 10–14, 1992

Folke Tersman, *Fem filosofiska frågor*, Wahlström&Wikstrand, 2001.