Lecture notes on the Semantics of Intuitionistic logic

Andrew W Swan

Abstract

These notes were originally written for the course Intuitionistic Logic 80818 at Carnegie Mellon University, Fall 2021. They are intended for students who have already seen some formal logic and aim to give an overview of the main ideas used in building and using models of intuitionistic theories, including Kripke models, topological models and realizability models.

Contents

1 Introduction

1.1 Why intuitionistic logic?

In intuitionistic logic we remove an axiom that many would view as fundamental in logic: the law of excluded middle, the statement $\varphi \vee \neg\varphi$ for all propositions φ. The original motivation for doing this was philosophical considerations, such as Brouwer's philosophy of intuitionism. He believed that mathematics is inherently a mental construction. Mathematical statements are true if they are known to be true and false if they are known to be false. If there is neither a proof of a statement, nor a proof of its negation (for example, famous open problems in mathematics such as the Riemann hypothesis and $P = NP$), then the statement is neither true nor false, according to intuitionism. Today intuitionism in its strictest form is not commonly believed by working mathematicians, but many still work without the law of excluded middle, in an approach to mathematics referred to as *constructivism.* Reasons for working constructively in mathematics include:

1. Constructive proofs are more appealing since they are more explicit, giving us a greater understanding of why a result is true.

2. Although many at first (including Brouwer himself) believed constructive mathematics to be impractical, large parts of mathematics are now known to work fine in constructive mathematics, as long definitions and statements of theorems are chosen correctly. This started with the work of Bishop, with later developments by many others.

3. Constructive proofs can be combined with powerful techniques in logic to get stronger versions of a theorem automatically. For example, given a constructive proof that a function exists, we can show there is a function with additional properties such as *continuity* and *computability*.

It is useful to know that large parts of mathematics can be done constructively, since otherwise studying constructive proofs would have little point. However, for this course we will mostly encounter the third point, as we study constructive proofs from the outside using formal logic.

Although Brouwer himself did not formalise his ideas using logic, intuitionism and constructive mathematics have been widely studied from a logical point of view, to further understand them and make them more precise. This started with Brouwer's student Heyting, who developed what we would now call intuitionistic logic, as well as some basic theories such as Heyting arithmetic (an intuitionistic version of Peano arithmetic) that we will see in this course. Today intuitionistic logic is a rich subject with many aspects. This course is mainly going to focus on the *semantics*, i.e. models, of intuitionistic logic. In classical logic, the term *model* usually only refers to one thing - the definition appearing in model theory. This definition would not get us very far in intuitionistic logic, and even for some quite basic results we need to consider other notions of model. In this course the notions of model will include

1. Kripke models
2. Heyting valued models
3. realizability models.

As mentioned above, some of these can be used to strengthen proofs, for example turning a proof of the existence of a function into a construction of a computable function. Another major theme is going to be independence and consistency proofs. Consider the axiom of countable choice. This states that if $\forall n \in \mathbb{N}\, \exists x\, \varphi(n, x)$ then there is a function f with domain $\mathbb{N}$ such that $\forall n \in \mathbb{N}\, \varphi(n, f(n))$. Cohen famously showed countable choice is independent of classical set theory **ZF**. However, there are weaker versions of countable choice that follow simply from the law of excluded middle. If $\forall n \in \mathbb{N}\, \exists m \in \mathbb{N}\, \varphi(n, m)$ is true, then we can non constructively prove the existence of a choice function f, simply by taking $f(n)$ to be the least natural number m such that $\varphi(n, m)$ is true. However, this is not provable in intuitionistic logic, and moreover we can show this using a natural example of a Heyting valued model. We can also show intuitionistic logic is consistent with many *anti-classical* axioms that contradict the law of excluded middle, such as Church's thesis, which states all functions $\mathbb{N} \to \mathbb{N}$ are computable.

1.2 Review of intuitionistic first order logic

This course will mostly be about the semantics (i.e. models) of intuitionistic logic and some simple theories based on intuitionistic logic, mostly variants of

Heyting arithmetic. Because of this we won't need to worry so much about the technicalities in the definition of proof that would arise in a course in proof theory. However, it is still useful to fix the definition of what we mean by formal proof, to refer back to later.

The system we will work with is multisorted intuitionistic natural deduction with sequent notation. To unpack this a bit more, we will consider a formal system with the following features,

1. It is intuitionistic - we will not assume the law of excluded middle (the axiom $\varphi \vee \neg\varphi$).
2. It is multisorted - we allow for theories where variables can range over different sorts. For example, we might have one sort for numbers and another sort for sets of numbers.
3. We will define proofs using natural deduction - most people find this the easiest and most intuitive form of formal logic.
4. We will use sequent notation to make natural deduction proofs easier to deal with formally.

Now for the formal definitions.

Definition 1.1. A *signature* consists of the following data:

1. A set $\mathcal{S}$ of *sorts*
2. A set $\mathcal{R}$ of *relation symbols*
3. For each relation symbol $R \in \mathcal{R}$, an *arity*, which is a finite list of sorts $S_1, \ldots, S_n \in \mathcal{S}$
4. A set $\mathcal{O}$ of *operator symbols*
5. For each operator symbol $O \in \mathcal{O}$, an *arity* of the operator, which is a finite list of sorts $S_1, \ldots, S_n \in \mathcal{S}$ together with 1 more sort, $T \in \mathcal{S}$. We will write the arity as $S_1, \ldots, S_n \to T$.

Next we define terms. We start with a countable supply of free variables of each sort. For now we will write the free variables of sort $S \in \mathcal{S}$ as $x_1^S, x_2^S, x_3^S, \ldots$. Later we will mostly drop the superscript, and indicate the sort of the variable other ways, for instance by the choice of letter or font.

Definition 1.2. Given a signature and the free variables, we inductively define terms, and simultaneously assign a sort to every term, as follows.

1. If x_i^S is a free variable of sort S, then it is also a term of sort S.
2. If $O \in \mathcal{O}$ is an operator symbol of arity $S_1, \ldots S_n \to T$ and $s_1, \ldots, s_n$ are terms of sort S_i for $i = 1, \ldots, n$, then $Os_1s_2 \ldots s_n$ is a term of sort T.

Next, we can define formulas.

Definition 1.3. We inductively define the set of *formulas* as follows.

1. If $R \in \mathcal{R}$ is a relation symbol, with arity $S_1, \dots, S_n \in \mathcal{S}$, and $s_1, \dots, s_n$ are terms where s_i has sort S_i for $i = 1, \dots, n$, then $Rs_1s_2 \dots s_n$ is a formula

2. $\perp$ is a formula

3. If φ and ψ are formulas, then the following are also formulas,

   (a) $\varphi \wedge \psi$

   (b) $\varphi \vee \psi$

   (c) $\varphi \to \psi$

4. If φ is a formula and x_i^S is a free variable, then the following are also formulas,

   (a) $\exists x_i^S \, \varphi$

   (b) $\forall x_i^S \, \varphi$

We will write substitution as $\varphi[x/t]$ where x is a free variable of sort $S \in \mathcal{S}$, and t is a term with the same sort S. We can read this as "x is replaced by t in φ." Formally, we define substitution as follows.

Definition 1.4. Let x be free variable of sort S and t a term of sort S. We first define substitution into terms $s[x/t]$ by induction on the definition of term. Namely,

1. If y is a free variable, we define $y[x/t]$ to be t if $x = y$, and otherwise $y[x/t]$ is defined to be y

2. For operator symbols $O \in \mathcal{O}$, we define $Os_1 \dots s_n[x/t]$ to be $O(s_0[x/t]) \dots (s_n[x/t])$.

We define $\varphi[x/t]$ for each formula φ again by induction

1. For a relation symbol $R \in \mathcal{R}$, we define $(Rs_1 \dots s_n)[x/t]$ to be $R(s_1[x/t]) \dots (s_n[x/t])$.

2. We define $(\varphi \Box \psi)[x/t]$ to be $\varphi[x/t] \Box \psi[x/t]$ where $\Box \in \{\wedge, \to, \vee\}$.

3. We define $(\forall y \, \varphi)[x/t]$ to be $\forall y \, (\varphi[x/t])$ when $y \neq x$ and to be $\forall y \, \varphi$ otherwise

4. We define $(\exists y \, \varphi)[x/t]$ to be $\exists y \, (\varphi[x/t])$ when $y \neq x$ and to be $\exists y \, \varphi$ otherwise

To help formulate proofs, we will first define sequents.

Definition 1.5. A *sequent* is a finite set of formulas $\varphi_1, \dots, \varphi_n$, and one more formula ψ. We will write this data as $\varphi_1, \dots, \varphi_n \vdash \psi$.

Finally, we can define proofs.

Definition 1.6. We define the set of *proofs* inductively by the following rules. Every proof proves a sequent, which is given simultaneously in the definition. We will refer to this as the *conclusion* of the proof.

To get off the ground, we first need the assumption rule. That is, whenever φ is an element of the set Γ, we have a proof

$$\frac{}{\Gamma \vdash \varphi}$$

We follow a general pattern that each logical connective has both introduction and elimination rules.

First the rules for conjunction:

$$\frac{\Gamma \vdash \varphi \qquad \Gamma \vdash \psi}{\Gamma \vdash \varphi \wedge \psi} \wedge I \qquad \frac{\Gamma \vdash \varphi \wedge \psi}{\Gamma \vdash \varphi} \wedge E_l \qquad \frac{\Gamma \vdash \varphi \wedge \psi}{\Gamma \vdash \psi} \wedge E_r$$

Disjunction:

$$\frac{\Gamma \vdash \varphi}{\Gamma \vdash \varphi \vee \psi} \vee I_l \quad \frac{\Gamma \vdash \psi}{\Gamma \vdash \varphi \vee \psi} \vee I_r \quad \frac{\Gamma \vdash \varphi \vee \psi \qquad \Gamma, \varphi \vdash \chi \qquad \Gamma, \psi \vdash \chi}{\Gamma \vdash \chi} \vee E$$

Implication:

$$\frac{\Gamma, \varphi \vdash \psi}{\Gamma \vdash \varphi \to \psi} \to I \qquad \frac{\Gamma \vdash \varphi \to \psi \qquad \Gamma \vdash \varphi}{\Gamma \vdash \psi} \to E$$

Now the rules for quantifiers, for all and exists. In both cases we assume x and y are free variables of the same sort, $S \in \mathcal{S}$, that t is a term of sort S. We also need to assume several technical conditions regarding free variables. Firstly, we need that the substitution $\varphi[x/t]$ "avoids free variable capture." That is, any occurrences of free variables in t do not become bound variables in $\varphi[x/t]$. For $\forall I$ we need that x does not occur free in any formula of Γ, and for $\exists E$ we need that x is not free in ψ. Finally, for both $\forall I$ and $\exists E$, we need that y is not free in φ unless $y = x$

$$\frac{\Gamma \vdash \varphi}{\Gamma \vdash \forall y\, \varphi[x/y]} \forall I \qquad \frac{\Gamma \vdash \forall x\, \varphi}{\Gamma \vdash \varphi[x/t]} \forall E$$

$$\frac{\Gamma \vdash \varphi[x/t]}{\Gamma \vdash \exists x\, \varphi} \exists I \qquad \frac{\Gamma \vdash \exists y\, \varphi[x/y] \qquad \Gamma, \varphi \vdash \psi}{\Gamma \vdash \psi} \exists E$$

Finally, we consider $\bot$, which only has an elimination rule, often referred to as *ex falso sequitur quodlibet* or just *ex falso*.

$$\frac{\Gamma \vdash \bot}{\Gamma \vdash \varphi} \bot E$$

We say a sequent $\Gamma \vdash \varphi$ is *provable* if it is the conclusion of some proof. We say a formula φ is provable if the sequent $\vdash \varphi$ is provable.

Definition 1.7. A *theory* over a given signature is a set T of formulas for that signature. We will refer to the elements of T as the *axioms* of the theory.

We write $T \vdash \varphi$ to mean that there is a finite set $\Gamma \subseteq T$ such that $\Gamma \vdash \varphi$, and say φ is a *theorem* of T.

2 Heyting Arithmetic

2.1 Equality

The theories we will consider will also have equality. For this we add a relation symbol $=^S$ of arity S, S for every sort S. We then add the following axioms whenever x is a free variable of sort S and s and t are terms of sort S. The axiom scheme is only for *atomic* formulas φ, i.e. those of the form $Rt_1 \ldots t_n$ for a relation symbol R, but holds without loss of generality for all formulas.

$$\Gamma \vdash x =^S x \qquad \frac{\Gamma \vdash s = t \qquad \Gamma \vdash \varphi[x/s]}{\Gamma \vdash \varphi[x/t]}$$

2.2 Review of first order Heyting arithmetic

The first formal theory we will see is (first order) Heyting arithmetic, **HA**. This is one of the most basic theories where we can formalise some real parts of mathematics. The signature for **HA** has a single sort, which we think of as the set of natural numbers. It has a nullary operator 0, a unary operator S, two binary operators $+$ and $\times$ and equality. We use *infix* notation for $+$ and $\times$. That is, we write $n + m$ and $n \times m$ rather than $+nm$ and $\times nm$ to match the usual notation in mathematics.

The axioms of **HA** are those for equality and as follows.

1. $\neg Sx = 0$
2. $Sx = Sy \to x = y$
3. $x + 0 = 0$
4. $x + (Sy) = S(x + y)$
5. $x \times 0 = x$
6. $x \times (Sy) = (x \times y) + x$
7. $(\varphi[x/0] \wedge \forall x\,(\varphi \to \varphi[x/S(x)])) \to \forall x\,\varphi$ for each formula φ (induction)

Although first order Heyting arithmetic is a solid base to work from, it is limited in the sense that we can only directly talk about numbers. However, many of the axioms that we will consider are most naturally stated in terms of functions or sets. We will therefore define two "higher order" extensions of **HA** where we can reason about functions or sets in addition to numbers.

2.3 Second order Heyting arithmetic

The first extension we will consider is *second order* Heyting arithmetic, which we abbreviate to **HAS**. The signature of **HAS** has two sorts: one for numbers, $\mathfrak{N}$ and one for sets of numbers, $\mathfrak{S}$. We write variables of sort $\mathfrak{N}$ using lower case letters, and variables of sort $\mathfrak{S}$ as upper case.

It has equality relations for both numbers and sets and a relation symbol $\in$. The arity of $\in$ is $\mathfrak{N}, \mathfrak{S}$. We write $\in nX$ using infix notation (i.e. as $n \in X$) to match the usual notation for set membership. Finally, **HAS** has the same operator symbols $0, S, +, \times$ as for **HA**.

The axioms of **HAS** include those of **HA** together with three more: *comprehension*, *extensionality* and *second order induction*.

Comprehension is an axiom scheme, which asserts for each formula φ the following:

$$\exists X \, \forall n \; n \in X \leftrightarrow \varphi$$

Extensionality is the following axiom:

$$\forall X, Y \, (\forall n \, n \in X \leftrightarrow n \in Y) \rightarrow (X = Y)$$

Finally, second order induction is the axiom below.

$$\forall X \, (0 \in X \, \wedge \, \forall n \, (n \in X \rightarrow Sn \in X)) \; \rightarrow \; \forall n \, n \in X$$

Note that second order induction and comprehension together imply the first order induction scheme, so we can drop first order induction from the definition without changing the set of theorems.

The richer language of **HAS** allows us to naturally state axioms that would be clumsy or even impossible to state in **HA**. For example, we can see our first example of *anti-classical* axioms, i.e. axioms that are provably false in classical logic. These are the *uniformity principle* (**UP**) and *unzerlegbarkeit* (**UZ**).

UP	$\forall X \, \exists n \, \varphi \; \rightarrow \; \exists n \, \forall X \, \varphi$
UZ	$\forall X \, (\varphi \vee \neg\varphi) \; \rightarrow \; (\forall X \, \varphi \, \vee \, \forall X \, \neg\varphi)$

We can see that **UP** implies **UZ**, and that **UZ** contradicts the law of excluded middle (for example by taking φ to be $\forall n \, n \notin X$). However, we will see later in the course that **UP** holds in *realizability* models of **HAS**, and so the theory is consistent.

2.4 Heyting arithmetic with finite types

The first extension of **HA** we are going to consider is Heyting arithmetic with *finite types*, which we abbreviate to $\mathbf{HA}_\omega$. Finite types allow us to also reason about functions. This includes functions that take a number as input and return a number as output. However, it goes much further - we can also have functions that take functions as input, and functions that take those as input, and so forth. We formalise this idea by first defining the set of *finite type symbols* as follows.

Definition 2.1. The set of *finite type symbols* is inductively generated by the following conditions:

1. N is a finite type symbol (the "type of numbers")

2. If σ and τ are type symbols, then so is $\sigma \times \tau$ (the "product type" of σ and τ)

3. If σ and τ are type symbols, then so is $\sigma \to \tau$ (the "type of functions from σ to τ")

We take the set of sorts of $\mathbf{HA}_\omega$ to be the finite type symbols.

Since we are thinking of $\sigma \to \tau$ as the sort of functions from σ to τ we should have some way to take an element of $\sigma \to \tau$ and an element of σ and return an element of τ by "function application." We do this by simply adding for all finite types σ and τ an operator symbol $\mathrm{Ap}^{\sigma,\tau}$ of arity $(\sigma \to \tau), \sigma \to \tau$. We will usually omit Ap when we write down terms of $\mathbf{HA}_\omega$. That is, for terms t, s of sorts $\sigma \to \tau$ and σ we will write $\mathrm{Ap}^{\sigma,\tau}\, ts$ simply as ts. When we have a series of function applications in a row, we follow a further notational convention of dropping parentheses according to "left associativity." Namely, for three terms r, s, t, we write rst to mean $(rs)t$, which is in turn notational shorthand for $\mathrm{Ap}(\mathrm{Ap}\, rs)t$.

Since we are thinking of $\sigma \times \tau$ as the product of σ and τ we should have some way to create new elements of $\sigma \times \tau$ by "pairing together an element of σ and an element of τ" and some way to take an element of $\sigma \times \tau$ and "project out" the component in σ and the component in τ. Again this is achieved by simply adding operator symbols. However, now we have Ap we can add them as constant symbols (i.e. 0-ary operator symbols) of the appropriate function type. Namely, we add a constant $\mathbf{p}^{\sigma,\tau}$ of sort $\sigma \to (\tau \to (\sigma \times \tau))$, a constant $\mathbf{p}_0^{\sigma,\tau}$ of sort $\sigma \times \tau \to \sigma$, and a constant $\mathbf{p}_1^{\sigma,\tau}$ of sort $\sigma \times \tau \to \tau$. In order for these constants to behave as described we will need to add the axioms below:

$$\begin{aligned} \mathbf{p}_0(\mathbf{p}xy) &= x \\ \mathbf{p}_1(\mathbf{p}xy) &= y \\ \mathbf{p}(\mathbf{p}_0 z)(\mathbf{p}_1 z) &= z \end{aligned}$$

We also add some constant symbols for generating new elements of a function sort. Firstly given an element of σ, we should get functions "constantly equal to that element." We implement this with a constant symbol $\mathbf{k}^{\sigma,\tau}$ of sort $\sigma \to (\tau \to \sigma)$. For this to behave as expected, we add the axiom below.

$$\mathbf{k}xy = x$$

We also want a way to compose two functions. It turns out that it's useful to have something a bit stronger that we can think of as "composition with a parameter." For this we add a constant symbol $\mathbf{s}^{\rho,\sigma,\tau}$ of sort $(\rho \to (\sigma \to \tau)) \to ((\rho \to \sigma) \to (\rho \to \tau))$, satisfying the axiom below:

$$\mathbf{s}xyz = xz(yz)$$

Finally, we have one more constant $\mathbf{r}^\sigma$, the *recursor*. One way to think about this is that we might want a way to "iterate" a function n times given a number

n. However, like with $\mathbf{s}$ it turns out that it's more useful to have a stronger version that takes a parameter as input. The sort of $\mathbf{r}^\sigma$ is $\sigma \to ((\sigma \to (N \to \sigma)) \to (N \to \sigma))$ and it satisfies the axioms below.

$$\mathbf{r}xy0 = x$$
$$\mathbf{r}xy(Sz) = y(\mathbf{r}xyz)z$$

We also add an equality relation for each sort, and the usual equality axioms. It also has the usual induction axiom scheme from **HA**.

Putting this all together we can formally define $\mathbf{HA}_\omega$ as follows.

Definition 2.2. *Heyting arithmetic with finite types*, $\mathbf{HA}_\omega$ is the theory defined as follows.

The set of sorts is the set of finite type symbols.

The set of operator symbols is an operator symbol $\mathrm{Ap}^{\sigma,\tau}$ of arity $(\sigma \to \tau), \sigma \to \tau$, for all finite types σ and τ together with the following constant symbols for all types σ, τ, ρ.

Constant symbol	Sort
0	N
S	$N \to N$
$\mathbf{p}^{\sigma,\tau}$	$\sigma \to (\tau \to (\sigma \times \tau))$
$\mathbf{p}_0^{\sigma,\tau}$	$\sigma \times \tau \to \sigma$
$\mathbf{p}_1^{\sigma,\tau}$	$\sigma \times \tau \to \tau$
$\mathbf{k}^{\sigma,\tau}$	$\sigma \to (\tau \to \sigma)$
$\mathbf{s}^{\rho,\sigma,\tau}$	$(\rho \to (\sigma \to \tau)) \to ((\rho \to \sigma) \to (\rho \to \tau))$
$\mathbf{r}^\sigma$	$\sigma \to ((\sigma \to (N \to \sigma)) \to (N \to \sigma))$

The only relation symbols are equality relations for each sort.

Firstly, we have the axioms for equality. In the language of $\mathbf{HA}_\omega$, this amounts to the following axioms.

$$x = x \qquad x = y \to y = x \qquad x = y \land y = z \to x = z$$
$$y = z \to \mathrm{Ap}\,xy = \mathrm{Ap}\,xz \qquad x = y \to \mathrm{Ap}\,xz = \mathrm{Ap}\,yz$$

Next, we have the formulas below, where we are following the conventions of writing $\mathrm{Ap}\,ts$ as ts for terms t and s, and taking application to be left associative.

$$\mathbf{p}_0(\mathbf{p}xy) = x \qquad \mathbf{p}_1(\mathbf{p}xy) = y \qquad \mathbf{p}(\mathbf{p}_0z)(\mathbf{p}_1z) = z$$
$$\mathbf{k}xy = x \qquad \mathbf{s}xyz = xz(yz)$$
$$\mathbf{r}xy0 = x \qquad \mathbf{r}xy(Sz) = y(\mathbf{r}xyz)z$$
$$Sx = Sy \to x = y \qquad \neg 0 = Sx$$

Finally, $\mathbf{HA}_\omega$ has the usual induction scheme from **HA**. That is, for each formula φ in the language of $\mathbf{HA}_\omega$, we have the following axiom.

$$(\varphi[x/0] \land \forall x\,(\varphi \to \varphi[x/S(x)])) \;\to\; \forall x\,\varphi$$

Although $\mathbf{HA}_\omega$ seems like a quite elaborate system, it is in some ways easier to deal with than **HAS**. We will see for example, that realizability models of $\mathbf{HA}_\omega$ are in a certain sense better behaved than those of of **HAS**. However, for this course, the main advantage of $\mathbf{HA}_\omega$ over **HAS** is that many of the axioms we are going to consider are stated in terms of functions. Although it is possible to implement functions as sets, using their graphs, it is more natural to just use a system that has a good notion of function already built in.

For example, we can now formulate the axiom of choice, a well known axiom usually viewed as uncontroversial today, but with an important place in the history and philosophy of mathematics. In $\mathbf{HA}_\omega$ it is most natural to not view it as a single axiom, but instead for all finite types σ and τ we have a separate axiom scheme $\mathbf{AC}^{\sigma,\tau}$ defined as the following, for each formula φ.

$$\forall x^\sigma \, \exists y^\tau \, \varphi \quad \to \quad \exists f^{\sigma \to \tau} \, \forall x^\sigma \, \varphi[y/fx]$$

We are going to see, for example,

1. using topological models, even the weakest version $\mathbf{AC}^{N,N}$ is independent of $\mathbf{HA}_\omega$
2. $\mathbf{AC}^{N,N}$ and $\mathbf{AC}^{N \to N,N}$ hold in certain realizability models and so are consistent with some anti-classical axioms.

We will also consider another axiom scheme, *function extensionality*, which is the following statement for all finite types σ and τ.

$$\forall f^{\sigma \to \tau} \, \forall g^{\sigma \to \tau} \, (\forall x^\sigma \, fx =^\tau gx) \to f =^{\sigma \to \tau} g$$

This is less often seen outside logic, since it usually viewed as obvious, and can be proved in set theory, when using the standard implementation of functions as graphs. However, it is often considered by philosophers of mathematics since it concerns the question of how we know when two objects (in this case functions) are equal to each other. One of the ideas we will see in this course is that although the axiom of choice is often seen as a non constructive axiom, this is in a certain sense only true when it is combined with function extensionality.

2.5 λ-terms in $\mathbf{HA}_\omega$

It might seem at first that $\mathbf{HA}_\omega$ is quite limited in what functions you can construct using the axioms. However, **s** and **k** turn out to be very powerful.

We first note that we can derive an "identity" term:

Definition 2.3. For each finite type, σ, we write $\mathbf{i}^\sigma$ for the closed term of $\mathbf{HA}_\omega$ defined as $\mathbf{s}^{\sigma,(\sigma\to\sigma),\sigma}\mathbf{k}^{\sigma,(\sigma\to\sigma)}\mathbf{k}^{\sigma,\sigma}$.

Proposition 2.4. *In* $\mathbf{HA}_\omega$ *we can prove* $\mathbf{i}x = x$.

Proof. By applying the axiom for $\mathbf{s}$ followed by the axiom for $\mathbf{k}$ we get the following equations:

$$\begin{aligned} \mathbf{i}x &:= \mathbf{skk}x \\ &= \mathbf{k}x(\mathbf{k}x) \\ &= x \end{aligned}$$

□

Using $\mathbf{s}$, $\mathbf{k}$ and $\mathbf{i}$ we can then show the powerful λ-abstraction lemma, that can be used to construct any function that can be written down as a term:

Lemma 2.5. *Let t be a term of $\mathbf{HA}_\omega$ of sort σ with a free variable x of sort τ. Then there is a term $\lambda x.t$ of sort $\tau \to \sigma$ satisfying the equation*

$$(\lambda x.t)y = t[x/y]$$

Proof. We construct $(\lambda x.t)$ by induction on the definition of terms.

If t is the free variable x, we take $(\lambda x.t)$ to be $\mathbf{i}^\sigma$, and it is clear this satisfies the lemma. (Note that in this case we have $\sigma = \tau$.)

If t is a constant c or free variable other than x, say y, (necessarily of sort σ) then we take $(\lambda x.t)$ to be $\mathbf{k}^{\tau,\sigma}c$ or $\mathbf{k}^{\tau,\sigma}y$ respectively. We again note that it is clear this satisfies the lemma, by the axiom for $\mathbf{k}$.

Finally, if t is of the form rs for a term r of sort $\rho \to \sigma$ and a term s of sort ρ, then we take $\lambda x.t$ to be $\mathbf{s}^{\tau,\rho,\sigma}(\lambda x.r)(\lambda x.s)$. We check that this works as follows.

$$\begin{aligned} (\lambda x.t)y &:= \mathbf{s}(\lambda x.r)(\lambda x.s)y \\ &= ((\lambda x.r)y)((\lambda x.s)y) && \text{axiom for } \mathbf{s} \\ &= r[x/y]\, s[x/y] && \text{inductive hypothesis} \\ &= (rs)[x/y] \end{aligned}$$

□

3 Omniscience Principles

3.1 Introducing the omniscience principles

In **HAS** and $\mathbf{HA}_\omega$ we now have formal languages where we can talk about mathematical statements that constructive mathematicians might be interested in. In this section we will see some important examples known as *omniscience principles*. These are used by constructive mathematicians to provide prototypical examples of nonconstructive statements. They are often used to illustrate to classical mathematicians the kind of statement that is not allowed when working constructively. They are also a useful tool for showing that something is not provable constructively: if we can use a statement φ to prove an omniscience principle, then φ must also be non constructive. However, for this argument to really work, we first need to show that the omniscience principles are not provable in the formal theories we are working in. By having a range of different principles, we can classify mathematical statements by "how non constructive" they are. This is one of the key ideas of a field known as *constructive reverse mathematics*.

Each omniscience principle has a universal quantifier ranging over all *binary sequences*, that is, functions from $\mathbb{N}$ to 2. When we are working in $\mathbf{HA}_\omega$ we implement this by viewing $\forall f \in 2^{\mathbb{N}}\ \varphi$ as notational shorthand for $\forall f^{N \to N}\ ((\forall n\, fn = 0 \vee fn = 1) \to \varphi)$. We first consider the strongest omniscience principle, the limited principle of omniscience.

Definition 3.1. The *limited principle of omniscience* (**LPO**) is the following statement:

$$\forall f \in 2^{\mathbb{N}}\ (\forall n\, f(n) = 0) \vee (\exists n\, f(n) = 1)$$

We first observe that **LPO** is easily provable in classical mathematics:

Proposition 3.2. *The law of excluded middle implies* **LPO**.

Proof. Let f be any binary sequence. By the law of excluded middle we know $\exists n\, f(n) = 1$ is either true or false. If it is true, then we are done. Suppose then that it is false. We need to show $\forall n\, f(n) = 0$. For each natural number n, we know that $f(n) = 0$ or $f(n) = 1$ (exercise!). However, if we had $f(n) = 1$, this would contradict $\neg(\exists n\, f(n) = 1)$. Hence we must have $f(n) = 0$. But we can now deduce $\forall n\, f(n) = 0$, as we needed. □

The intuition for why **LPO** is not constructively acceptable is that in order to know whether the sequence f contains a 1, we need to look at the entire sequence at once. We cannot tell whether or not there exists n such that $f(n) = 1$ by only looking at a finite portion of the sequence, whether that is the first five values, the first one million values, or first Graham's number values or higher. If we ever find a number n such that $f(n) = 1$, then we know for sure that $\exists n\, f(n) = 1$, but if $f(n) = 0$ for every value of n we have checked so far, then we have no way of knowing whether this will continue to be the case forever, or if we will find an n with $f(n) = 1$ sometime in the future.

We can also motivate the idea by thinking about physical measurements. As we improve our equipment and carry out more experiments we can know physical quantities with greater and greater precision, but we will never reach absolute precision. For example, if we are given two platinum bars, we will eventually find out if they have different lengths, but if they have the same length we will never know for sure. Hence, there is no way in general to decide which of the two cases we are in.

Finally, we can understand this idea through *computability*. Given a computer program that outputs the numbers 0 and 1, we have no way to decide, in general whether it will output 0 forever, or whether it will eventually output 1 given enough time on an ideal computer. We will later make this last intuition precise using realizability.

The remaining omniscience principles are the weak limited principle of omniscience, the lesser limited principle of omniscience and Markov's principle.

Definition 3.3. The *weak limited principle of omniscience* (**WLPO**) is the following statement.

$$\forall f \in 2^{\mathbb{N}}\ (\forall n\, f(n) = 0) \lor \neg(\forall n\, f(n) = 0)$$

Definition 3.4. *Markov's principle* (**MP**) is the following statement.

$$\forall f \in 2^{\mathbb{N}}\ \neg(\forall n\, f(n) = 0) \to \exists n\, f(n) = 1$$

Markov's principle is not always included as an omniscience principle, and many constructive mathematicians view it as a perfectly reasonable axiom to use in constructive proofs. For example, it is viewed as acceptable according to the philosophy of recursive or "Russian" constructive mathematics. The idea is that existential quantifiers need to be justified by computable functions. In the case of Markov's principle, we imagine the binary sequence as a computer program outputting a sequence of 0's and 1's as it is given different inputs. If it is false that the program will always output 0, then we can write a program to find a number n such that the program outputs 1 given input n - we simply keep running the original program on higher and higher input values until it returns 1, and then return the input value where this happened.

Note however that Markov's principle is exactly what we need to get from **WLPO** to **LPO**. More precisely, we have the following proposition.

Proposition 3.5. **LPO** *is equivalent to the conjunction of* **WLPO** *and* **MP**.

Proof. There is a short proof in intuitionistic natural deduction that $\exists n\, f(n) = 1$ implies $\neg(\forall n\, f(n) = 0)$. From this it is clear that **LPO** implies **WLPO**. Similarly, given Markov's principle, we also have the converse statement, that $\neg(\forall n\, f(n) = 0)$ implies $\exists n\, f(n) = 1$, and so we also have that **WLPO** and **MP** together imply **LPO**.

We just need to check that **LPO** implies **MP**, but this is again straightforward: if $\exists n\, f(n) = 1$, then we are done, and if $\forall n\, f(n) = 0$, we can apply ex falso together with the assumption $\neg(\forall n\, f(n) = 0)$ to deduce $\exists n\, f(n) = 1$. ☐

Definition 3.6. The *lesser limited principle of omniscience* (**LLPO**) is the following statement.

$$\forall f \in 2^{\mathbb{N}}\ \forall n, m\, ((f(n) = 1 \land f(m) = 1) \to n = m) \longrightarrow (\forall n\, f(2n) = 0) \lor (\forall n\, f(2n+1) = 0)$$

Most people find lesser limited principle of omniscience to be the least intuitive of the omniscience principles. To explain it a bit more, the clause $\forall n, m\, ((f(n) = 1 \land f(m) = 1) \to n = m)$ says that f has "at most one 1." That is, $f(n)$ is equal to 0 for almost all n. It could be equal to 0 for all n, for example. However, we also allow for the possibility that $f(n)$ is equal to

1 for some n. In this case we know that $f(m)$ is equal to 0 whenever $m \neq n$. The statement $\forall n\, f(2n) = 0$ is telling us that if $f(n) = 1$, then n must be odd. Similarly, $\forall n\, f(2n+1) = 0$ tells us that if $f(n) = 1$, then n must be even. So **LLPO** is telling us that even if we don't know whether there is an n such that $f(n) = 1$, we can say either "if there is such an n it is odd" or "if there is such an n it is even."

Proposition 3.7. *The weak limited principle of omniscience implies the lesser limited principle of omniscience.*

Proof. Let f be a binary sequence with at most one 1. We apply **WLPO** to the sequence f' defined by $f'(n) := f(2n)$. If we have $\forall n\, f'(n) = 0$, then we are done, since $f(2n) = 0$ for all n.

Now suppose that we have $\neg(\forall n\, f'(n) = 0)$. We will show that for all n, we have $f(2n+1) = 0$. First recall that we know for each n that either $f(2n+1) = 0$ or $f(2n+1) = 1$. If we had $f(2n+1) = 1$, then it would imply that $f(2m) = 0$ for all numbers m, since we have $2m \neq 2n+1$ for all m. However, this would contradict $\neg(\forall n\, f'(n) = 0)$. Hence we have $f(2n+1) = 0$, and since this applies for all numbers n, we are done. □

Although **LLPO** is usually seen as not acceptable in constructive mathematics, in settings where we do not have countable choice it can be surprisingly harmless. A kind of realizability called *Lifschitz realizability* can be used to show that **LLPO** is consistent with *Church's thesis* (an anti classical axiom that says "all functions $\mathbb{N} \to \mathbb{N}$ are computable").

3.2 Review of the standard ordering on natural numbers

In order to help formalise some ideas we briefly review the standard ordering on natural numbers.

Definition 3.8. We write $x < y$ as shorthand for the formula $\exists z\, y = x + Sz$

The following standard properties of $<$ can be proved in **HA**. However, for this course we will omit the proofs.

Proposition 3.9. *The binary relation $x < y$ has the following properties.*

1. $\neg(x < x)$ *(irreflexivity)*
2. $x < y \wedge y < z \;\to\; x < z$ *(transitivity)*
3. $x < y \;\vee\; y < x \;\vee\; x = y$ *(trichotomy)*
4. $\neg(x < 0)$
5. $x < Sy \;\leftrightarrow\; (x = y \vee x < y)$

We also review some more standard notation. First of all, we write $x \leq y$ to mean $\exists z\, y = x + z$. This also satisfies a list of properties that we will assume without proof (but can be proved in **HA**).

Proposition 3.10. *The binary relation $x \leq y$ has the following properties.*

1. $x \leq x$ *(reflexivity)*
2. $x \leq y \wedge y \leq z \;\rightarrow\; x \leq z$ *(transitivity)*
3. $x \leq y \wedge y \leq x \;\rightarrow\; x = y$ *(anti-symmetry)*
4. $x \leq y \vee y \leq x$ *(linearity)*
5. $x \leq y \;\leftrightarrow\; x = y \vee x < y$

We will sometimes use the notational convention of "bounded quantifiers." Namely, we do the following:

1. We write $\forall x < y\, \varphi$ to mean $\forall x\, (x < y \rightarrow \varphi)$.
2. We write $\exists x < y\, \varphi$ to mean $\exists x\, (x < y \wedge \varphi)$
3. We write $\forall x \leq y\, \varphi$ to mean $\forall x\, (x \leq y \rightarrow \varphi)$.
4. We write $\exists x \leq y\, \varphi$ to mean $\exists x\, (x \leq y \wedge \varphi)$

Definition 3.11. Suppose we are given a formula $\varphi(x)$. We say n is the *least* number satisfying φ if the following holds:

$$\varphi(n) \,\wedge\, \forall x\, (\varphi(x) \rightarrow n \leq x)$$

Note that by the propositions above (in particular trichotomy), this is equivalent to the following:

$$\varphi(n) \,\wedge\, \forall x < n\, \neg\varphi(x)$$

We justify saying *the* least number by the fact that if n and m both have this property, then $m \leq n$ and $n \leq m$, and so $m = n$.

3.3 An explicit version of LPO and the axiom of unique choice

For each omniscience principle, we can also define an "explicit" version where we have a function that "witnesses" the truth of the omniscience principle.

For example, for **LPO**, we define this as follows.

Definition 3.12. The *explicit limited principle of omniscience* states that there is a function $F : 2^{\mathbb{N}} \rightarrow (2 \times \mathbb{N})$ with the following property. For every binary sequence f, exactly one of the following two conditions applies:

1. $\mathbf{p}_0\, F(f) \;=\; 0$ and $\forall n\, f(n) = 0$
2. $\mathbf{p}_0\, F(f) \;=\; 1$ and $f(\mathbf{p}_1\, F(f)) = 1$

Since we don't have a sort for 2 in $\mathbf{HA}_\omega$, we note that we can alternatively define F on *all* functions $\mathbb{N} \to \mathbb{N}$ and ignore those that are not binary sequences. Namely, we can formalise explicit **LPO** in $\mathbf{HA}_\omega$ as the following statement:

$$\exists F^{N \to N, N \times N} \, \forall f^{N \to N} \, (\forall n \, fn = 0 \, \vee \, fn = 1) \to \\ (\mathbf{p}_0(Ff) = 0 \wedge \forall n \, fn = 0) \vee (\mathbf{p}_0(Ff) = 1 \wedge f(\mathbf{p}_1(Ff)) = 1)$$

We can easily see that explicit **LPO** implies **LPO**. For the converse we sometimes (if working in $\mathbf{HA}_\omega$, for example) need an additional axiom, the axiom of unique choice.

We first introduce some notation.

Definition 3.13. We write $\exists! x \, \varphi(x)$ as notation for the following statement.

$$\exists x \, (\varphi(x) \, \wedge \, \forall y \, \varphi(y) \to x = y)$$

We say "there exists a *unique* x satisfying φ."

Definition 3.14. Given sorts σ and τ *axiom of unique choice*, $\mathbf{AC}_!^{\sigma,\tau}$ is the following statement for each formula φ.

$$\forall x^\sigma \, \exists! y^\tau \, \varphi(x, y) \; \to \; \exists f^{\sigma \to \tau} \, \forall x^\sigma \, \varphi(x, fx)$$

Theorem 3.15. *Assuming* **LPO** *and the axiom of unique choice, we can prove explicit* **LPO**.

Proof. We will just give an informal proof in constructive mathematics. Exercise: Think about how you would formalise this in $\mathbf{HA}_\omega$.

We are going to construct $F : (\mathbb{N} \to 2) \to (2 \times \mathbb{N})$ using unique choice.

We define $\varphi(f, x)$ to be following statement: Either $\mathbf{p}_0 x = 0$, $\mathbf{p}_1 x = 0$, and $\forall n \, f(n) = 0$ or $\mathbf{p}_0 x = 1$ and $\mathbf{p}_1 x$ is the least number n such that $f(n) = 1$.

We first check existence of x. By **LPO**, we know that either $\forall n \, f(n) = 0$ or $\exists n \, f(n) = 1$. In the former case we can take $x = \mathbf{p}00$. In the latter case, we can in fact deduce there is a *least* n such that $f(n) = 1$ (exercise!), and then take x to be $\mathbf{p}1n$.

We now need to check uniqueness. Suppose we have x and y such that $\varphi(f, x)$ and $\varphi(f, y)$ are both true. Note that we cannot have both $\forall n \, f(n) = 0$ and $\exists n \, f(n) = 1$ since these would contradict each other. It follows that $\mathbf{p}_0 x = \mathbf{p}_0 y$. To show the $\mathbf{p}_1 x = \mathbf{p}_1 y$, we split into the two cases $\forall n \, f(n) = 0$ and $\exists n \, f(n) = 1$. In the former case we have $\mathbf{p}_1 x = 0 = \mathbf{p}_1 y$. In the latter case we recall that the least number satisfying any condition is unique, and so we also have $\mathbf{p}_1 x = \mathbf{p}_1 y$ in that case, as we needed. ☐

4 Heyting Algebras and Topology

4.1 Posets and lattices

We first recall some basic theory about posets.

Definition 4.1. A *poset* is a set P, together with a binary relation $\leq$ satisfying the following axioms.

1. $x \leq x$ (reflexivity)
2. $x \leq y \wedge y \leq z \ \rightarrow\ x \leq z$ (transitivity)
3. $x \leq y \wedge y \leq x \ \rightarrow\ x = y$ (anti-symmetry)

We will sometimes write $x \leq y$ as $y \geq x$.

Definition 4.2. The *top* or *greatest* element of a poset P is $\top \in P$ such that for all $x \in P$, $x \leq \top$.

The *bottom* or *least* element of a poset P is $\bot \in P$ such that for all $x \in P$, $\bot \leq x$.

Definition 4.3. Let S be a set of elements of a poset P. We say z is the *least upper bound* or *join* of S if

1. for all $x \in S$, $x \leq z$ (z is an upper bound)
2. if $y \in P$ is such that for all x in S $x \leq y$, then $z \leq y$ (any other upper bound is greater)

Note that any set has at most one join. If it exists, we write it as $\bigvee S$. Given a two element set $\{x, y\}$, we write $\bigvee\{x, y\}$ as $x \vee y$.

We similarly define the *greatest lower bound* or *meet* of S as an element z such that

1. for all $x \in S$, $z \leq x$
2. if $y \in P$ is such that for all x in S $z \leq x$, then $y \leq z$.

The meet of a set is also unique, and when it exists we write it as $\bigwedge S$, and for two element sets $\{x, y\}$, we write $\bigwedge\{x, y\}$ as $x \wedge y$.

Definition 4.4. We say a poset P is *complete* if for every set $S \subseteq P$, $\bigvee S$ and $\bigwedge S$ both exist.

Proposition 4.5. *A poset P is complete if and only if it has all joins (or all meets).*

Proof. Let S be a set. We want to construct the greatest lower bound of S. Define $L := \{x \in P \mid \forall y \in S\, x \leq y\}$ the set of all lower bounds of S. We claim $\bigvee L$ is the greatest lower bound of S. For each $y \in S$, we know that for every x in L, $x \leq y$. It follows that $\bigvee L \leq y$. By applying this to each $y \in S$ we see $\bigvee L$ is a lower bound for S and it is clear it is the greatest one (since L contains every other lower bound). $\square$

Definition 4.6. A *lattice* is a poset P with least and greatest elements $\bot$ and $\top$, and any two elements $x, y \in P$ have a meet $x \wedge y$ and a join $x \vee y$.

Example 4.7. If P is any collection of sets, then it has a canonical ordering given by $x \leq y$ whenever $x \subseteq y$. If P is closed under binary intersection $x \cap y$ and union $x \cup y$, then $(P, \subseteq)$ is a lattice.

It is possible to recover the poset relation on a lattice from the operations $\wedge$ and $\vee$. Because of this, we can also think of lattices as *algebraic* structures (i.e. sets with operations satisfying equations).

4.2 Heyting algebras

To motivate Heyting algebras, we first consider an important example, the *Lindenbaum-Tarski algebra* of an intuitionistic theory.

Let T be a theory over some signature. We define an equivalence relation on formulas by $\varphi \sim \psi$ when $T \vdash \varphi \leftrightarrow \psi$. Let P be the quotient of the set of formulas by $\sim$. Note that we have a canonical ordering on P: we say $[\varphi] \leq [\psi]$ if $T \vdash \varphi \to \psi$, and note that this is preserved by the equivalence relation.

We see that P is a lattice, and moreover each part of the lattice structure corresponds naturally to logical connectives. For example $[\bot] \leq [\varphi]$ for all formulas φ, by ex falso. Similarly $[\top]$ is the greatest element of the lattice. We can show that $[\varphi \vee \psi]$ is greater than $[\varphi]$ and $[\psi]$ by the $\vee$ introduction rule, and the $\vee$ elimination rule precisely tells us that any other upper bound is greater than $[\varphi \vee \psi]$. Similarly we can use conjunction to construct meets in the Lindenbaum-Tarski algebra.

However, there is one logical connective that does not appear as part of the lattice structure, namely *implication*, $\to$. As before, we can translate the introduction and elimination rules into properties of the order structure. For all formulas φ, ψ and χ, we can use implication introduction to show that if $[\chi \wedge \varphi] \leq [\psi]$, then $[\chi] \leq [\varphi \to \psi]$. Using introduction elimination, we can also show the converse: if $[\chi] \leq [\varphi \to \psi]$, then $[\chi \wedge \varphi] \leq [\psi]$.

We can see Heyting algebras as posets that behave similar to Lindenbaum-Tarski algebras. We can think of the elements of a Heyting algebra as "truth values." We will use them to define certain models (*Heyting valued models*) of theories in intuitionistic logic. Formally, we define them as follows.

Definition 4.8. A *Heyting algebra* is a lattice $(P, \top, \bot, \wedge, \vee)$ together with a binary operation $\to$, *implication*, satisfying the condition below.

$$z \leq x \to y \qquad \text{if and only if} \qquad z \wedge x \leq y$$

4.3 Topological spaces

For this course, our main source of examples of Heyting algebras are going to be *topological spaces*. From the point of view of Heyting algebras, we can think of topological spaces as concrete examples of Heyting algebras, whose elements are all subsets of a fixed set.

The original motivation for topological spaces is to model "spaces" that appear in mathematics, such a spheres or 3-dimensional space.

Another idea that will be important in this course is the notion of morphism between topological spaces, *continuous* functions.

4.3.1 Some basic definitions and examples

Definition 4.9. Let X be a set. A *topology* on X is a collection $\mathcal{O}$ of subsets of X, satisfying the following conditions.

1. X and $\emptyset$ are elements of $\mathcal{O}$.
2. If U and V are elements of $\mathcal{O}$, their intersection $U \cap V$ also belongs to $\mathcal{O}$.
3. If $S \subseteq \mathcal{O}$ is a set of elements of $\mathcal{O}$, then their union $\bigcup S$ also belongs to $\mathcal{O}$.

We refer to the elements of $\mathcal{O}$ as *open sets*, and the elements of X as *points*. We say a set X together with a topology is a *topological space*.

Example 4.10. If X is any set, then we can define a topology by taking $\mathcal{O}$ to be the collection of *all* subsets of X. This is referred to as the *discrete* topology on X.

We can also define a topology by taking $\mathcal{O}$ to have just two elements, $\emptyset$ and X. This is referred to as the *indiscrete topology*.

Example 4.11. We define a topology on the set with two elements $2 := \{0, 1\}$ as the set of subsets $\{\emptyset, \{0\}, \{0, 1\}\}$. This topological space is referred to as *Sierpiński space*, S.

Example 4.12. We define a topology on the set with three elements $I = \{0, 1, 2\}$ as the set of subsets $\{\emptyset, \{0, 1\}, \{1\}, \{1, 2\}, \{0, 1, 2\}\}$. We will refer to this space as the *abstract interval*, or just the *interval*.

Example 4.13. If $(Q, \leq)$ is any poset, we can define a topology on Q as follows. A set $U \subseteq Q$ is *upwards closed* or an *upset* if whenever $x \in U$ and $x \leq y$, we have $y \in U$. The set of all upsets defines a topology on Q referred to as the *upset* or *Alexandrov* topology on the poset.

Example 4.14. If $(X, \mathcal{O})$ is a topological space, and Y is a subset of X, then we can also define a topology on Y as follows. We say a set $U \subseteq Y$ is open if for some open set V of X, we have $U = V \cap Y$. We refer to this as the *subspace topology* on Y.

Definition 4.15. Let $(X, \mathcal{O})$ be a topological space, and $Y \subset X$ any subset. We define the *interior* of Y, Y° to be the union of all open sets U such that $U \subseteq Y$.

Proposition 4.16. *We observe the following properties of interior.*

1. *For all Y, $Y^\circ \subseteq Y$.*
2. *For all Y, Y° is an open set.*

3. If Y is already an open set, then $Y^\circ = Y$.

Example 4.17. In Sierpiński space, the interior of $\{1\}$ is the empty set.

Proposition 4.18. *For any topological space $(X, \mathcal{O})$, the poset $\mathcal{O}$ of open sets ordered by inclusion is a complete Heyting algebra.*

Proof. We define joins using union and meets using (binary) intersection. Top and bottom element are given by X and $\emptyset$ respectively. It only remains to define the implication operator. We might try to define this using the canonical implication on subsets, i.e. for sets U and V, the set

$$\{x \in X \mid x \in U \to x \in V\}.$$

However, the above set is not necessarily open, even when U and V are. To get an open set, we use the interior operator. Namely, we define

$$U \to V := \{x \in X \mid x \in U \to x \in V\}^\circ.$$

It only remains to check that this satisfies the necessary condition to be the implication of a Heyting algebra (exercise!). □

NB: Using the law of excluded middle, we could alternatively define implication by

$$U \to V := (V \cup (X \setminus U))^\circ$$

Remark 4.19. *Since lattices of open sets have all joins, they must also have meets, by proposition 4.5. Whereas the join of a set $S \subseteq \mathcal{O}$ is just the union, $\bigvee S = \bigcup S$, the intersection $\bigcap S$ is not necessarily open. However, we can still explicitly describe the meet of S using the interior operator, $\bigwedge S = (\bigcap S)^\circ$.*

Definition 4.20. Let $(X, \mathcal{O}_X)$, $(Y, \mathcal{O}_Y)$ be topological spaces. We say a function $f : X \to Y$ is *continuous* if for every open set U of Y, the preimage, $f^{-1}(U)$ is an open set of X.

4.3.2 Neighbourhoods in a topological space

Definition 4.21. Let $(X, \mathcal{O})$ be a topological space. A *neighbourhood* of a point $x \in X$ is an open set $U \in \mathcal{O}$ such that $x \in U$.

We think of a neighbourhood of a point x as a set of points that are "nearby" x. For example, in a discrete topology, every point x has a neighbourhood $\{x\}$ that does not contain any other points. We can visualise this as a clear space around x that does not contain any elements. On the other hand in an indiscrete space, the only neighbourhood of a point x is the entire space. The points are so "close together" that you can't look at one without looking at the whole space. In Sierpiński space, every neighbourhood of 1 also contains 0, so 0 is "infinitely close" to 1, but this is not a symmetric relation: 0 has a neighbourhood that does not contain 1.

Definition 4.22. Let $(X, \mathcal{O})$ be a topological space. A *basis* of the topology, is a set of open sets $B \subseteq \mathcal{O}$ with the following property. For any open neighbourhood U of a point x, there exists an open neighbourhood V of x with $V \subseteq U$ and $V \in B$.

Equivalently, a basis is a set B such that for every open set U, there is a set $S \subseteq B$ such that $U = \bigcup S$.

Remark 4.23. *Note that the set of all open sets is a basis. This satisfies the definition, but is typically not useful.*

Note that if B is a basis of a topology, then a set is open if and only if it is equal to the union $\bigcup S$ for some $S \subseteq B$.

Example 4.24. If $(Q, \leq)$ is any poset, then we can define a basis of the upset topology as the set of upwards closed sets with least element. Note that there is a precise correspondence between the elements of the basis and elements of Q.

4.3.3 Product topologies

Definition 4.25. We define a topology on $\mathbb{N}^{\mathbb{N}}$ as follows. Given a finite sequence of numbers σ of length k, we define the set U_σ as follows.

$$U_\sigma := \{f : \mathbb{N} \to \mathbb{N} \mid \forall i < k\, f(i) = \sigma(i)\}$$

We define a set U to be open if it can be written as a union of sets of the form U_σ for finite sequences σ. We refer to the topological space with these open sets as *Baire space.*

We refer to the subspace topology on the set of binary sequences $2^{\mathbb{N}} \subseteq \mathbb{N}^{\mathbb{N}}$ as *Cantor space.*

We refer to the set of binary sequences with at most one 1, with the subspace topology as $\mathbb{N}_\infty$.

Proposition 4.26. *Assume* **LPO**. *Then $\mathbb{N}_\infty$ is isomorphic to the set $\mathbb{N} \amalg \{\infty\}$, where a set $U \subseteq \mathbb{N} \amalg \{\infty\}$ is open when it satisfies the following. If $\infty \in U$, then for some $n \in \mathbb{N}$, U contains every $m \in \mathbb{N}$ with $m \geq n$.*

Proposition 4.27. *Given $\mathbb{N}^{\mathbb{N}}$ with the Baire space topology, and $\mathbb{N}$ with the discrete topology, a function $F : \mathbb{N}^{\mathbb{N}} \to \mathbb{N}$ is continuous if and only if it satisfies the following condition. For every $f \in \mathbb{N}$, there is a natural number n such that for any $g \in \mathbb{N}$ satisfying the condition that $g(i) = f(i)$ for $i < n$, we have $F(g) = F(f)$.*

Proof. Since $\mathbb{N}$ has the discrete topology, every singleton set $\{m\}$ is open, and in any case we can write any set as a union of singletons. It follows that F is continuous if and only if $F^{-1}(\{m\})$ is open for every m. This says exactly that any element f of $F^{-1}(\{m\})$ has a basic open neighbourhood contained in $F^{-1}(\{m\})$. This precisely says that the condition described in the proposition holds whenever $F(f) = m$. However, F is continuous precisely when this condition holds for arbitrary m, so given a function f, we can just apply it with $m := F(f)$. □

The Baire space topology on $\mathbb{N}^{\mathbb{N}}$ is an instance of a more general construction known as the product topology.

Definition 4.28. Suppose we are given a set I, and a family of topological spaces $(X_i, \mathcal{O}_i)$ for each $i \in I$. We define a topology on the product $\prod_{i \in I} X_i$ as follows. Given a pair $\sigma = (F, (U_j)_{j \in F})$ consisting of a finite subset F of I, say $i_1, \ldots, i_k$, together with open sets $U_j \in \mathcal{O}_{i_j}$ for $j = 1, \ldots, k$, we define the set U_σ by the equation

$$U_\sigma := \{f \in \prod_{i \in I} X_i \mid \forall j \in F\; f(j) \in U_j\}$$

We refer to sets of the form U_σ as basic opens, and define a set to be open if it is a union of basic opens. We refer to the resulting topological space as the *product topology* of the family.

Proposition 4.29. *Baire space is the product of the (constant) family consisting of countably many copies of* $\mathbb{N}$ *with the discrete topology.*

4.3.4 Connectedness

An important concept when looking at the behaviour of topological models is that of connectedness, which informally is the idea that a space is "indecomposable" - it is impossible to break the space up cleanly into separate pieces.

Definition 4.30. A topological space X is *connected* if given open sets U and V such that $U \cup V = X$ and $U \cap V = \emptyset$ we have either $X = U$ or $X = V$.

Proposition 4.31. *A topological space X is connected if and only if for every discrete topological space Y, every map $X \to Y$ is constant.*

Example 4.32. Sierpiński space and the abstract interval are connected.

Example 4.33. Any discrete space with at least two distinct points is not connected.

Example 4.34. Baire space, Cantor space and $\mathbb{N}_\infty$ are not connected.

Definition 4.35. We say a topological space X is *locally connected* if for every point x of X and every open neighbourhood U of x, there is an open neighbourhood V of x such that $V \subseteq U$ and V is connected as a topological space with the subspace topology.

4.4 Formal topologies

It is sometimes convenient to take an alternative approach to topology, where we ignore the points of the topological space, and instead focus on the elements of a basis for the topology.

Definition 4.36. A *formal topology* is a poset $(B, \leq)$ together with a relation $\lhd$ of sort $B, \mathcal{P}(B)$ (i.e. a relation on B and sets of elements of B), satisfying the following axioms, for all $a, b \in B$ and $U, V \subseteq B$. We write $U^{\leq}$ to mean the *downwards closure of* U, i.e. $\{a \in B \mid \exists b \, b \in U \wedge a \leq b\}$.

1. $a \in U$ implies $a \lhd U$.
2. $a \lhd U$ and $a \lhd V$ implies $a \lhd U^{\leq} \cap V^{\leq}$.
3. If $a \lhd U$ and for all $b \in U$, $b \lhd V$, then $a \lhd V$.
4. $a \leq b$ implies $a \lhd \{b\}$

We refer to the relation $\lhd$ as the *covering relation* of the formal topology. When $b \lhd U$, we say b is *covered by* U.

Definition 4.37. We say an *open set* is a subset U of B satisfying the following conditions:

1. If $a \leq b$ and $b \in U$ then $a \leq b$. (downwards closure)
2. If $a \lhd U$, then $a \in U$.

Proposition 4.38. *The open sets of a formal topology ordered by inclusion form a complete Heyting algebra.*

Example 4.39. If $(B, \leq)$ is any poset, we can define a minimal covering relation by $b \lhd U$, whenever $b \leq a$ for some $a \in U$.

Example 4.40. Let $(X, \mathcal{O})$ be a topological space with a basis B. We define a formal topology as follows. We take the underlying poset to be B ordered by subset inclusion. If U is a set of basic open sets, We say $b \lhd U$ if $b \subseteq \bigcup U$. Then open sets of the formal topology correspond precisely to the open sets of the topological space. If $V \subset X$ is an open set of the topological space, we define an open set U of the formal topology, by taking U to be the set of basic opens b such that $b \subseteq V$. Given an open set U of the formal topology, we define an open set $V \subset X$ of the topological space by $V := \bigcup U$.

Example 4.41. As a special case of the previous example, we can define *formal Baire space* as follows. We take B to be the set of finite sequences of natural numbers, ordered by *reverse extension.* That is, for finite sequences σ and τ, we say $\sigma \leq \tau$ if the length of σ is greater than than of τ, and whenever i is less than the length of τ, we have $\tau(i) = \sigma(i)$. We define the covering relation $\lhd$ to be the smallest relation satisfying the axioms of a formal topology, and such that if $\sigma * \langle n \rangle \lhd U$ for all n, then $\sigma \lhd U$.

5 Heyting Valued Models

In this section we will see our first example of models of intuitionistic logic. The essential idea is that we think of the elements of a Heyting algebra as "truth values." We then assign each sentence a truth value in such a way that logical connectives $\wedge$, $\vee$ and $\rightarrow$ are sent to the corresponding operation in the Heyting algebra. This would work directly for propositional logic, but for first order we also need to deal with quantifiers. The way we will deal with this here is by making the additional assumption that we are given a *complete* Heyting algebra. We will then use this to interpret quantifiers; namely by sending universal quantifiers to meets, and existential quantifiers to joins.

5.1 Heyting valued models

Throughout this section we fix a signature with set of sorts $\mathfrak{S}$, set of operator symbols $\mathfrak{O}$ and set of relation symbols $\mathfrak{R}$.

We furthermore assume we are given a complete Heyting algebra P.

Definition 5.1. A *Heyting valued model* consists of the following data:

1. For each sort $S \in \mathfrak{S}$ a set $\mathcal{M}_S$, together with a map $E_S : \mathcal{M}_S \rightarrow P$
2. For each operator symbol $O \in \mathfrak{O}$, of sort $S_1, \ldots, S_n \rightarrow T$, a function $[\![O]\!] : \mathcal{M}_{S_1} \times \ldots \times \mathcal{M}_{S_n} \rightarrow \mathcal{M}_T$, satisfying the following for any $a_1, \ldots, a_n$ with $a_i \in \mathcal{M}_{S_i}$:
$$E([\![O]\!](a_1, \ldots, a_n)) \;\geq\; E(a_1) \wedge \ldots \wedge E(a_n)$$
3. For each relation symbol $R \in \mathfrak{R}$ of sort $S_1, \ldots, S_n$, a function $[\![R]\!] : \mathcal{M}_{S_1} \times \ldots \times \mathcal{M}_{S_n} \rightarrow P$.

Finally, for technical reasons,[1] we will also assume the following non triviality condition. For every sort $S \in \mathfrak{S}$ there exists an element a of $\mathcal{M}_S$ with $E(a) = \top$.

Definition 5.2. A *topological model* is a Heyting valued model where the Heyting algebra is the lattice of open sets of a topological space.

For $a \in \mathcal{M}_S$, we refer to $E(a)$ as the *extent* of a. One way to think about this is that the object a does not exist absolutely, but only with truth value $E(a)$. We technically allow for the case $E(a) = \bot$ ("a does not exist"), but such objects will not affect any truth values in the Heyting valued model, so we can always just ignore them. When P is the lattice of open sets of a topological space, we can visualise $E(a)$ as the region of space where a is present.

Definition 5.3. A *variable assignment* is a function σ with domain the set of free variables such that for each free variable x^S of sort S, $\sigma(x^S) \in \mathcal{M}_S$.

Given a variable assignment σ, a free variable x^S and an element a of $\mathcal{M}_S$, we write for $\sigma[x^S \mapsto a]$ for the assigment that sends x^S to a, and sends y^T to $\sigma(y^T)$ when $y^T \neq x^S$.

[1] This relates to the formulation of first order logic we are using.

For each term t of sort $S \in \mathfrak{S}$, and each variable assignment σ, we define an element $[\![t]\!]_\sigma$ of $\mathcal{M}_S$ by induction on terms.

$$\begin{aligned} [\![x^S]\!]_\sigma &:= \sigma(x^S) \\ [\![Ot_1 \dots t_n]\!]_\sigma &:= [\![O]\!]([\![t_1]\!]_\sigma, \dots, [\![t_n]\!]_\sigma) \end{aligned}$$

Note that we can show the following lemmas by induction on terms.

Lemma 5.4. *Let s be a term of sort S and σ a variable assignment. Let $x_1, \dots, x_n$ be a list of variables including any occuring free in s. Then,*

$$E(\sigma(x_1)) \wedge \dots \wedge E(\sigma(x_n)) \leq [\![s]\!]_\sigma$$

Lemma 5.5. *Let t be a term and σ a variable assignment. Suppose x^S is a free variable, and s is a term of sort S. Then we have the following equality.*

$$[\![t[x^S/s]]\!]_\sigma = [\![t]\!]_{\sigma[x^S \mapsto [\![s]\!]_\sigma]}$$

For each formula φ and each variable assignment σ defined on all variables occuring free in φ, we define an element $[\![\varphi]\!]_\sigma$ of P, by induction on terms.

$$\begin{aligned} [\![Rt_1 \dots t_n]\!]_\sigma &:= [\![R]\!]([\![t_1]\!]_\sigma, \dots, [\![t_n]\!]_\sigma) \\ [\![\bot]\!]_\sigma &:= \bot \\ [\![\varphi \wedge \psi]\!]_\sigma &:= [\![\varphi]\!]_\sigma \wedge [\![\psi]\!]_\sigma \\ [\![\varphi \vee \psi]\!]_\sigma &:= [\![\varphi]\!]_\sigma \vee [\![\psi]\!]_\sigma \\ [\![\varphi \to \psi]\!]_\sigma &:= [\![\varphi]\!]_\sigma \to [\![\psi]\!]_\sigma \\ [\![\exists x^S \, \varphi(x)]\!]_\sigma &:= \bigvee_{a \in \mathcal{M}_S} E_S(a) \wedge [\![\varphi]\!]_{\sigma[x \mapsto a]} \\ [\![\forall x^S \, \varphi(x)]\!]_\sigma &:= \bigwedge_{a \in \mathcal{M}_S} E_S(a) \to [\![\varphi]\!]_{\sigma[x \mapsto a]} \end{aligned}$$

We again get a substitution lemma.

Lemma 5.6. *Let φ be a formula, x^S a free variable of sort $S \in \mathfrak{S}$ and s a term of sort S, such that the substitution $\varphi[x^S/s]$ avoids free variable capture. Then we have the following equality.*

$$[\![\varphi[x^S/s]]\!]_\sigma = [\![\varphi]\!]_{\sigma[x^S \mapsto [\![s]\!]_\sigma]}$$

Proof. By induction on formulas. □

We can similarly show the following lemma by induction on formulas.

Lemma 5.7. *Let φ be a formula, x^S a free variable of sort $S \in \mathfrak{S}$ that does not occur in φ and a any element of $\mathcal{M}_S$. Then we have*

$$[\![\varphi]\!]_\sigma = [\![\varphi]\!]_{\sigma[x^S \mapsto a]}$$

Given a finite list of formulas $\Gamma := \varphi_1, \ldots, \varphi_n$, we write $[\![\Gamma]\!]_\sigma$ for $[\![\varphi_1]\!]_\sigma \wedge \ldots \wedge [\![\varphi_n]\!]_\sigma$.

We can now prove the *soundness theorem* for Heyting valued models.

Theorem 5.8. *Suppose that $\Gamma \vdash \varphi$ is provable in intuitionistic first order logic, where $\Gamma = \psi_1, \ldots, \psi_n$, and $x_1, \ldots, x_m$ is a list of free variables including all those occuring in Γ or φ. Then for any free variable assignment σ, we have.*

$$[\![\Gamma]\!]_\sigma \;\wedge\; E(\sigma(x_1)) \wedge \ldots \wedge E(\sigma(x_m)) \;\leq\; [\![\varphi]\!]_\sigma$$

Proof. We show this by induction on proofs. We will just do some of the cases to illustrate the idea (the remainder are left as an exercise).

The case $\wedge I$ Suppose that we have deduced $\Gamma \vdash \varphi \wedge \psi$ by $\wedge I$, together with $\Gamma \vdash \varphi$ and $\Gamma \vdash \psi$. By the inductive hypothesis, we know that for any variable assignment σ and any list of free variables $x_1, \ldots, x_m$ including all those occuring free in Γ, φ or ψ, we have the following.

$$[\![\Gamma]\!] \wedge E(\sigma(x_1)) \wedge \ldots \wedge E(\sigma(x_n)) \leq [\![\varphi]\!]_\sigma$$
$$[\![\Gamma]\!] \wedge E(\sigma(x_1)) \wedge \ldots \wedge E(\sigma(x_n)) \leq [\![\psi]\!]_\sigma$$

It follows directly from the fact that $[\![\varphi \wedge \psi]\!]_\sigma = [\![\varphi]\!]_\sigma \wedge [\![\psi]\!]_\sigma$ by definition, and the definition of meet in a lattice that

$$[\![\Gamma]\!] \wedge E(\sigma(x_1)) \wedge \ldots \wedge E(\sigma(x_n)) \leq [\![\varphi \wedge \psi]\!]_\sigma$$

which is exactly what we needed to show.

The case $\wedge E$ Suppose we have deduce $\Gamma \vdash \varphi$ from $\Gamma \vdash \varphi \wedge \psi$. By the inductive hypothesis, we know that for any variable assignment σ and any list of free variables $x_1, \ldots, x_m$ including all those occuring free in Γ, φ or ψ, we have the following.

$$[\![\Gamma]\!]_\sigma \wedge E(\sigma(x_1)) \wedge \ldots \wedge E(\sigma(x_m)) \;\leq\; [\![\varphi \wedge \psi]\!]_\sigma$$

We can easily deduce the following from the definitions.

$$[\![\Gamma]\!]_\sigma \wedge E(\sigma(x_1)) \wedge \ldots \wedge E(\sigma(x_m)) \;\leq\; [\![\varphi]\!]_\sigma$$

It only remains to deal with the case where we are given a list of variables, say $x_1, \ldots, x_m$ including all those in Γ and φ, but missing out some free variables occuring in ψ but not Γ or φ. Let $y_1^{T_1}, \ldots, y_k^{T_k}$ be a list of all the free variables occuring in ψ, but not equal to x_i for any i. Let $a_1, \ldots, a_k$ be a list with $a_j \in \mathcal{M}_{T_j}$ and $E(a_j) = \top$ for $j = 1, \ldots, k$. Let τ be the variable assignment sending y_j to a_j and otherwise equal to σ. By lemma 5.7 we have $[\![\Gamma]\!]_\tau = [\![\Gamma]\!]_\sigma$ and $[\![\varphi]\!]_\tau = [\![\varphi]\!]_\sigma$, and by the basic properties of meet, we have

$$[\![\Gamma]\!]_\tau \wedge E(\sigma(x_1)) \wedge \ldots \wedge E(\sigma(x_m)) \wedge E(\tau(y_1)) \wedge \ldots \wedge E(\tau(y_k)) = [\![\Gamma]\!]_\tau \wedge E(\sigma(x_1)) \wedge \ldots \wedge E(\sigma(x_m))$$

We can thereby deduce the following, where this time the list $x_1, \ldots, x_m$ only needs to contain variables occuring free in Γ and φ.

$$[\![\Gamma]\!]_\sigma \wedge E(\sigma(x_1)) \wedge \ldots \wedge E(\sigma(x_m)) \;\leq\; [\![\varphi]\!]_\sigma$$

The remaining logical connectives are very similar and are left as an exercise. However, for quantifiers we need to be a bit careful and make use of the substitution lemmas.

The case $\exists E$ Suppose we have deduced $\Gamma \vdash \psi$ from $\Gamma \vdash \exists y\, \varphi[x/y]$ and $\Gamma, \varphi \vdash \psi$. We want to show the following, for all variable assignments σ and lists of variables $x_1, \ldots, x_m$ containing all the free variables in Γ and ψ:

$$[\![\Gamma]\!]_\sigma \wedge E(\sigma(x_1)) \wedge \ldots \wedge E(\sigma(x_m)) \leq [\![\psi]\!]_\sigma \tag{1}$$

For any $a \in \mathcal{M}_S$, we have the following by the inductive hypothesis,

$$[\![\Gamma]\!]_\sigma \wedge [\![\varphi]\!]_{\sigma[x \mapsto a]} \wedge E(\sigma(x_1)) \wedge \ldots \wedge E(\sigma(x_m)) \wedge E(a) \;\leq\; [\![\psi]\!]_\sigma$$

We deduce that we have the following:

$$\bigvee_{a \in \mathcal{M}_S} ([\![\Gamma]\!]_\sigma \wedge [\![\varphi]\!]_{\sigma[x \mapsto a]} \wedge E(\sigma(x_1)) \wedge \ldots \wedge E(\sigma(x_m)) \wedge E(a)) \;\leq\; [\![\psi]\!]_\sigma$$

By distributivity, we have

$$[\![\Gamma]\!]_\sigma \wedge E(\sigma(x_1)) \wedge \ldots \wedge E(\sigma(x_m)) \wedge \bigvee_{a \in \mathcal{M}_S} (E(a) \wedge [\![\varphi]\!]_{\sigma[x \mapsto a]}) \;\leq\; [\![\psi]\!]_\sigma$$

which is just

$$[\![\Gamma]\!]_\sigma \wedge E(\sigma(x_1)) \wedge \ldots \wedge E(\sigma(x_m)) \wedge [\![\exists x\, \varphi]\!]_\sigma \;\leq\; [\![\psi]\!]_\sigma$$

However, again by the inductive hypothesis we have

$$[\![\Gamma]\!]_\sigma \wedge E(\sigma(x_1)) \wedge \ldots E(\sigma(x_n)) \leq [\![\exists x\, \varphi]\!]_\sigma$$

It follows that

$$[\![\Gamma]\!]_\sigma \wedge E(\sigma(x_1)) \wedge \ldots \wedge E(\sigma(x_m)) \wedge [\![\exists x\, \varphi]\!]_\sigma = [\![\Gamma]\!]_\sigma \wedge E(\sigma(x_1)) \wedge \ldots \wedge E(\sigma(x_m))$$

However, we can now deduce (1).

The case $\exists I$ Suppose we have deduced $\Gamma \vdash \exists x^S\, \varphi(x)$ from $\Gamma \vdash \varphi(s)$, where s is a term of sort S, avoiding free variable capture. We want to show the following, for all variable assignments σ:

$$[\![\Gamma]\!]_\sigma \;\wedge\; E(\sigma(x_1)) \wedge \ldots \wedge E(\sigma(x_m)) \;\leq\; [\![\exists x^S\, \varphi]\!]_\sigma$$

Let $y_1^{T_1}, \ldots, y_k^{T_k}$ be a list of the free variables occuring in s but not equal to x_i for any i. Let $a_1, \ldots, a_k$ be a list with $a_j \in \mathcal{M}_{T_j}$ and $E(a_j) = \top$ for $j = 1, \ldots, k$. Let τ be the variable assignment obtained by setting the value at y_j to be a_j for each j, and otherwise agreeing with σ. By the inductive hypothesis, we have

$$[\![\Gamma]\!]_\tau \wedge E(\sigma(x_1)) \wedge \ldots \wedge E(\sigma(x_m)) \leq [\![\varphi[x^S/s]]\!]_\tau$$

By lemma 5.6 we have $[\![\varphi[x^S/s]]\!]_\tau = [\![\varphi]\!]_{\tau[x^S \mapsto [\![t]\!]_\tau]}$. We can therefore reason as follows,

$$\begin{aligned}
[\![\Gamma]\!]_\tau \wedge E(\sigma(x_1)) \wedge \ldots \wedge E(\sigma(x_m)) &\leq [\![\varphi[x^S/s]]\!]_\tau \\
&= [\![\varphi]\!]_{\tau[x \mapsto [\![t]\!]_\tau]} \\
&\leq \bigvee_{a \in \mathcal{M}_S} [\![\varphi]\!]_{\tau[x \mapsto a]} \\
&= [\![\exists x\, \varphi]\!]_\tau \\
&= [\![\exists x\, \varphi]\!]_\sigma
\end{aligned}$$

The case $\forall E$ Suppose that we have derived $\Gamma \vdash \varphi[x/s]$ from $\Gamma \vdash \forall x^S \varphi$. Suppose that we are given a list of variables $x_1, \ldots, x_m$ including all those occuring free in Γ and $\varphi[x/s]$, and σ is a variable assignment. By induction, we may assume we already have the following.

$$\begin{aligned}
[\![\Gamma]\!] \wedge E(\sigma(x_1)) \wedge \ldots \wedge E(\sigma(x_m)) &\leq [\![\forall x\, \varphi]\!]_\sigma \\
&= \bigwedge_{a \in \mathcal{M}_S} (E(a) \to [\![\varphi]\!]_{\sigma[x^S \mapsto a]}) \\
&\leq E([\![s]\!]_\sigma) \to [\![\varphi]\!]_{\sigma[x \mapsto [\![s]\!]_\sigma]} \\
&= E([\![s]\!]_\sigma) \to [\![\varphi[x/s]]\!]_\sigma
\end{aligned}$$

By the definition of Heyting implication, we can deduce

$$[\![\Gamma]\!] \wedge E(\sigma(x_1)) \wedge \ldots \wedge E(\sigma(x_m)) \wedge E([\![s]\!]_\sigma) \leq [\![\varphi[x/s]]\!]_\sigma$$

Note that the list of variables $x_1, \ldots, x_n$ includes any occuring free in s. Hence we have $E([\![s]\!]_\sigma) \geq E(x_1) \wedge \ldots E(x_m)$ by lemma 5.4. Hence, we have

$$\begin{aligned}
&[\![\Gamma]\!] \wedge E(\sigma(x_1)) \wedge \ldots \wedge E(\sigma(x_m)) \wedge E([\![s]\!]_\sigma) = \\
&\qquad\qquad\qquad [\![\Gamma]\!] \wedge E(\sigma(x_1)) \wedge \ldots \wedge E(\sigma(x_m))
\end{aligned}$$

But we are now done. □

Corollary 5.9. *Suppose that* $\vdash \varphi$. *Then for every variable assignment* σ *such that* $E(\sigma(x^S)) = \top$ *for all variables* x^S *occuring free in* φ, *we have* $[\![\varphi]\!]_\sigma = \top$. *(We also write this* $\mathcal{M} \vDash \varphi$*).*

5.2 Some simple examples of Heyting valued models

We now give some basic examples of Heyting valued models to illustrate how they work, and the use of the soundness theorem to get independence results.

5.2.1 Trivial Heyting valued models

We first consider the simplest Heyting algebra, 2, which just has two elements $\bot$ and $\top$. Assuming the law of excluded middle, this is a complete Heyting algebra.[2] We can also see this Heyting algebra as the lattice of open sets of the topological space with exactly one point.

In this case, we have a set $\mathcal{M}_S$ for each sort $S \in \mathfrak{S}$. As mentioned, the elements x with $E_S(x) = \bot$ don't play any role when assigning truth values, so by removing them, we may assume without loss of generality that $E_S(x) = \top$ for all $x \in \mathcal{M}_S$. We then see that the interpretation of an operator symbol O of sort $S_1, \ldots, S_n \to T$ is just a function $\mathcal{M}_{S_1} \times \ldots \times \mathcal{M}_{S_n} \to \mathcal{M}_T$. The interpretation of a relation symbol R of sort $S_1, \ldots, S_n$ is just a subset of $\mathcal{M}_{S_1} \times \ldots \times \mathcal{M}_{S_n}$.

We can see that this has recovered the usual notion of model for classical logic. Furthermore, note that the assigment of truth values is the same as usual for models. That is, we have the following equivalences:

$$
\begin{array}{lll}
[\![\varphi \wedge \psi]\!]_\sigma = \top & \text{iff} & [\![\varphi]\!]_\sigma = \top \text{ and } [\![\psi]\!]_\sigma = \top \\
[\![\varphi \vee \psi]\!]_\sigma = \top & \text{iff} & [\![\varphi]\!]_\sigma = \top \text{ or } [\![\psi]\!]_\sigma = \top \\
[\![\varphi \to \psi]\!]_\sigma = \top & \text{iff} & [\![\varphi]\!]_\sigma = \top \text{ implies } [\![\psi]\!]_\sigma = \top \\
[\![\exists x^S\, \varphi(x)]\!] = \top & \text{iff} & \text{there exists } a \in \mathcal{M}_S \text{ such that } [\![\varphi]\!]_{\sigma[x \mapsto a]} = \top \\
[\![\forall x^S\, \varphi(x)]\!] = \top & \text{iff} & \text{for all } a \in \mathcal{M}_S, [\![\varphi]\!]_{\sigma[x \mapsto a]} = \top
\end{array}
$$

Note that if the law of excluded middle is true in the metatheory where we are working, then it also holds in any such model.

5.2.2 Sierpiński space

In order to get results about intuitionistic logic that aren't already true for classical logic, we need to consider Heyting valued models for non trivial Heyting algebras.

We first consider the open sets of Sierpiński space. These consist of $\bot = \emptyset$, $\top = \{0, 1\}$ and an "intermediate truth value" $\{0\}$, with $\bot \leq \{0\} \leq \top$ corresponding to the open set containing just 0.

This is already enough to get some simple separation results, showing that some theorems easily provable in classical logic cannot be proved in intuitionistic logic. To illustrate the idea we will consider a very simple signature, which just one sort, no operator symbols and a single 0-ary relation symbol Q.

Theorem 5.10. *The formula $Q \vee \neg Q$ is not provable in intuitionistic logic.*

[2]When we are working in a constructive metatheory, we can instead use the set of all subsets of a singleton set.

Proof. We consider the Heyting valued model on Sierpiński space over the above signature defined as follows. We take $\mathcal{M}$ to be the one element set $\{*\}$, and define $E(*) := \top$. It remains to define the interpretation of the 0-ary relation symbol $[\![Q]\!]$, which just needs to be an element of the Heyting algebra. We take $[\![Q]\!]$ to be the intermediate truth value $\{0\}$. By definition this gives us the following for the unique variable assignment σ.

$$[\![Q]\!]_\sigma = \{0\}$$

We now consider the truth value of $\neg Q$. Recall that this is just notation for $Q \to \bot$.

$$\begin{aligned} [\![Q \to \bot]\!] &= [\![Q]\!] \to [\![\bot]\!] \\ &= \{0\} \to \emptyset \\ &= \{1\}^{\circ} \\ &= \emptyset \end{aligned}$$

Note that even though the truth value of Q was non trivial, the truth value of $\neg Q$ is still $\bot$. We can now calculate,

$$\begin{aligned} [\![Q \vee \neg Q]\!] &= [\![Q]\!] \cup [\![\neg Q]\!] \\ &= \{0\} \cup \emptyset \\ &= \{0\} \\ &\neq \top \end{aligned}$$

Now we observe that if $Q \vee \neg Q$ was provable in intuitionistic logic we would have by theorem 5.8 that $[\![Q \vee \neg Q]\!] = \top$. Since this is not the case, we can deduce that $Q \vee \neg Q$ is not provable. □

We give one more example using Sierpiński space, this time to illustrate the use of the extent predicate. We work over a signature with one sort, no operator symbols, and two relation symbols, a nulllary relation Q, and a unary relation R.

It's useful to note that we have the following lemmas. The first is for any Heyting algebra.

Lemma 5.11. *Let p, q be elements of a Heyting algebra. Then $p \to q = \top$ if and only if $p \leq q$.*

Proof. Suppose that $p \leq q$. Then $p \wedge r \leq q$ for any element r of the Heyting algebra. This implies that $p \wedge r \leq q$ if and only if $r \leq \top$. In other words $\top$ satisfies the property that uniquely characterises $p \to q$. It follows that $\top = p \to q$. Note that we can reverse each step of the above argument to show the converse. □

Lemma 5.12. *Let U be an open set of a topological space X. Then $U \to \bot = (X \setminus U)^{\circ}$ (the interior of the complement of U).*

Proof. By expanding out the definition of implication for the opens of a topological space. ☐

Theorem 5.13. *The following statement (sometimes referred to as* constant domain*) is not provable in intuitionistic logic.*

$$\forall x\,(Q \vee Rx) \;\rightarrow\; (Q \vee \forall x\, Rx)$$

Proof. We define a topological model on Sierpiński space as follows. We define $\mathcal{M}$ to have 2 elements, say $\mathcal{M} = \{a, b\}$. In order to satisfy the non triviality condition, we need $E(x) = \top$ for some x. Say $E(a) = \top$. We take the other extent to be the intermediate truth value $E(b) := \{0\}$. We again take Q to have the intermediate truth value $[\![Q]\!] = \{0\}$. We define the interpretation of R by $[\![R]\!](a) = \top$ and $[\![R]\!](b) = \bot$.

Note that $[\![\forall x\,(Q \vee Rx) \;\rightarrow\; (Q \vee \forall x\, Rx)]\!] = \top$ if and only if $[\![\forall x\,(Q \vee Rx)]\!] \subseteq [\![Q \vee \forall x\, Rx]\!]$. We explicitly compute both values as follows.

$$\begin{aligned}
[\![\forall x\,(Q \vee Rx)]\!] &= \bigwedge_{c \in \mathcal{M}} E(c) \to [\![Q \vee Rx]\!]_{x \mapsto c} \\
&= \bigwedge_{c \in \mathcal{M}} E(c) \to ([\![Q]\!] \cup [\![Rx]\!]_{x \mapsto c}) \\
&= (E(a) \to ([\![Q]\!] \cup [\![Rx]\!]_{x \mapsto a})) \wedge (E(b) \to ([\![Q]\!] \cup [\![Rx]\!]_{x \mapsto b})) \\
&= (\top \to (\{0\} \cup \top)) \cap (\{0\} \to (\{0\} \cup \bot)) \\
&= (\top \to \top) \cap (\{0\} \to \{0\}) \\
&= \top
\end{aligned}$$

$$\begin{aligned}
[\![Q \vee \forall x\, Rx]\!] &= \{0\} \cup \bigwedge_{c \in \mathcal{M}} E(c) \to [\![Rx]\!]_{x \mapsto c} \\
&= \{0\} \cup ((E(a) \to [\![Rx]\!]_{x \mapsto a}) \cap (E(b) \to [\![Rx]\!]_{x \mapsto b})) \\
&= \{0\} \cup ((\top \to \top) \cap (\{0\} \to \bot)) \\
&= \{0\} \cup (\top \cap \{1\}^{\circ}) \\
&= \{0\} \cup (\top \cap \emptyset) \\
&= \{0\} \\
&\neq [\![\forall x\,(Q \vee Rx)]\!]
\end{aligned}$$

☐

6 Kripke Models and Formal Topological Models

6.1 Forcing notation for Heyting valued models

It can be sometimes be useful to look at Heyting valued models from a different angle. Rather than looking at the truth value of a formula, which could be any element of the Heyting algebra, we fix a set of elements that are simple to describe or otherwise well behaved, and use this to give our alternative viewpoint.

For this we use the definition of basis that we have seen for topological spaces but generalises to all complete Heyting algebras (or in fact any complete poset).

Definition 6.1. A *basis* or *generating set* for a Heyting algebra P is a set $B \subseteq P$ with the following property. For any element p of P, we have

$$p = \bigvee \{q \in B \mid q \leq p\}$$

We can use this to define *forcing notation* or *Kripke-Joyal semantics* for Heyting valued models.

Definition 6.2. Let p be an element of a basis B of a Heyting algebra. We write $p \Vdash_\sigma \varphi$ for $p \leq [\![\varphi]\!]_\sigma$. If $p \Vdash_\sigma \varphi$, we say *p forces φ.*

Note that we can recover the truth value of φ with respect to a variable assignment σ from the forcing relation $\Vdash_\sigma$. Namely, we have the following equality.

$$[\![\varphi]\!]_\sigma = \bigvee \{p \in B \mid p \Vdash_\sigma \varphi\}$$

We can directly describe the forcing relation for logical formulas using the propositions below.

Proposition 6.3. *For all $p \in B$, and all formulas φ and ψ, we have the following.*

1. *$p \Vdash_\sigma \bot$ if and only if $p = \bot$.*
2. *$p \Vdash_\sigma \varphi \wedge \psi$ if and only if $p \Vdash_\sigma \varphi$ and $p \Vdash_\sigma \psi$.*
3. *$p \Vdash_\sigma \varphi \to \psi$ if and only if, for all $q \in B$ such that $q \leq p$ and $q \Vdash_\sigma \varphi$, we also have $q \Vdash_\sigma \psi$.*
4. *$p \Vdash_\sigma \forall x\, \varphi$ if and only if for all $a \in \mathcal{M}_S$ and for all $q \in B$ with $q \leq p$, if $q \Vdash E(a)$, then $q \Vdash_{\sigma[x \mapsto a]} \varphi$.*

Proof. 1 and 2 follow directly from the definitions of $\bot$ and $\wedge$ in a lattice.

We now show 3.

$$\begin{aligned} p \leq [\![\varphi]\!]_\sigma \to [\![\psi]\!]_\sigma \quad & \text{iff} \\ p \wedge [\![\varphi]\!]_\sigma \leq [\![\psi]\!]_\sigma \quad & \text{iff} \\ [\![\varphi]\!]_\sigma \leq p \to [\![\psi]\!]_\sigma \quad & \text{iff} \\ \bigvee \{r \mid r \Vdash_\sigma \varphi\} \leq p \to [\![\psi]\!]_\sigma \quad & \end{aligned}$$

However, the final inequality holds if and only if for every r such that $r \Vdash \varphi$, we have $r \wedge p \leq [\![\psi]\!]_\sigma$. However, this holds precisely when for all $q \in B$ such that $q \leq r$ and $q \leq p$ we have $q \Vdash_\sigma \psi$. Finally, observe that q is less than or equal to an element r such that $r \Vdash_\sigma \varphi$ if and only if $q \Vdash_\sigma \varphi$. Hence we now have the condition in 3.

We can show 4 by a very similar argument to 3. ☐

We say a basis B is *proper* if it does not contain $\bot$. Note that in this case we have $p \nVdash_\sigma \bot$ for all p.

We don't have such a different looking characterisation of the other logical connectives, but we can reformulate them a little, by using covering notation.

Definition 6.4. Let $S \subseteq P$[3] and $p \in B$. We write $p \lhd S$ and say p is *covered by* S when $p \leq \bigvee S$.

Proposition 6.5. *For all $p \in B$, and all formulas φ and ψ, we have the following.*

1. $p \Vdash_\sigma \varphi \vee \psi$ *if and only if* $p \lhd \{q \in B \mid q \Vdash_\sigma \varphi$ *or* $q \Vdash_\sigma \psi\}$.
2. $p \Vdash_\sigma \exists x\, \varphi$ *if and only if* $p \lhd \bigcup_{a \in \mathcal{M}_S} \{q \in B \mid q \leq E(a)$ *and* $q \Vdash_{\sigma[x \mapsto a]} \varphi\}$.

6.2 Kripke models

Let $(Q, \leq)$ be a poset with a bottom element. We will define Kripke models for the poset. To keep things simple, we will only consider signatures with no operator symbols.[4] Write $\mathfrak{S}$ for the set of sorts, and $\mathfrak{R}$ for the set of relation symbols.

Definition 6.6. A *Kripke model* consists of the following data for each $q \in Q$,

1. For each sort $S \in \mathfrak{S}$, an inhabited set $\mathcal{M}_{S,q}$
2. For each relation symbol $R \in \mathfrak{R}$ of sort $S_1, \ldots, S_n$ and $q \in Q$, a set $[R]_q \subseteq \mathcal{M}_{S_1,q} \times \ldots \times \mathcal{M}_{S_n,q}$

satisfying the following conditions for all $p, q \in Q$ such that $p \leq q$:

1. $\mathcal{M}_{S,p} \subseteq \mathcal{M}_{S,q}$ for all sorts S.
2. For each relation symbol $R \in \mathfrak{R}$ we have $[R]_p \subseteq [R]_q$.

Given a Kripke model, we define a topological model as follows. We first need to specify the topological space. We take it to be Q with the upset topology.

1. For each sort $S \in \mathfrak{S}$, we define $\mathcal{M}_S := \bigcup_{q \in Q} \mathcal{M}_{S,q}$.
2. For each sort $S \in \mathfrak{S}$, and $a \in \mathcal{M}_S$, we define $E_S(a) := \{q \in Q \mid a \in \mathcal{M}_{S,q}\}$.
3. For each relation symbol $R \in \mathfrak{R}$ of sort $S_1, \ldots, S_n$ and for elements $a_1, \ldots, a_n$ with each a_i belonging to $\mathcal{M}_{S_i}$, we define $[\![R]\!](a_1, \ldots, a_n)$ to be $\{q \in Q \mid (a_1, \ldots, a_n) \in [R]_q\}$.

[3]Typically we will only consider the case $S \subseteq B$.

[4]In order to deal with operator symbols it is necessary to use a construction called *sheafification* that makes the definition more complicated. We will see a version of this later for topological models of $\mathbf{HA}_\omega$. Exercise: Think about what goes wrong in the definition of Heyting valued model here when we have a binary operation symbol.

Note that by the definition of Kripke model, we can see that for each a, $E_S(a)$ is an open set in the topology, i.e. an upwards closed subset of Q. Similarly, the condition on relations in the definition of Kripke model tells us that $[\![R]\!](a_1, \ldots, a_n)$ is always an open set. Hence we do indeed have a well defined topological model on the upset topology.

We can now explicitly describe the forcing relation for topological models derived from Kripke models in this way. First recall that we can view the elements of Q itself as a basis for the upset topology, where $q \in Q$ corresponds to the upset $\{r \in Q \mid q \leq r\}$. Under this correspondence, we have for any upwards closed set U that $p \lhd U$ if and only if $p \in U$. Also recall that the correspondence between elements of Q ordered by $\leq$ and upwards closed sets with least element ordered by $\subseteq$, the order is reversed. Hence if we want to describe the forcing relation in terms of the order on Q, we need to reverse all the inequalities in proposition 6.3. This gives us the following definition:

$p \nVdash_\sigma \bot$		always
$p \Vdash_\sigma \varphi \wedge \psi$	iff	$p \Vdash_\sigma \varphi$ and $p \Vdash_\sigma \psi$
$p \Vdash_\sigma \varphi \vee \psi$	iff	$p \Vdash_\sigma \varphi$ or $p \Vdash_\sigma \psi$
$p \Vdash_\sigma \varphi \to \psi$	iff	for all $q \geq p$ if $q \Vdash_\sigma \varphi$ then $q \Vdash_\sigma \psi$
$p \Vdash_\sigma \exists x\, \varphi$	iff	there exists $a \in \mathcal{M}_p$ such that $p \Vdash_{\sigma[x \mapsto a]} \varphi$
$p \Vdash_\sigma \forall x\, \varphi$	iff	for all $q \geq p$ and $a \in \mathcal{M}_q$, $q \Vdash_{\sigma[x \mapsto a]} \varphi$

6.3 Formal topological models

Kripke models turn out to be useful in some situations, but are limited in some ways. For this reason, we consider a generalisation that includes some more examples of Heyting valued models, even some that are not topological models, but behaves very in a similar way to Kripke models, with a similar explicit description of their forcing relation. Instead of a poset, we consider the more general definition of formal topology. Fix a formal topology B with ordering relation $\leq$ and covering relation $\lhd$. We again restrict to signatures without operator symbols for simplicity.

Definition 6.7. A *formal topological model* consists of the following data for each $p \in B$,

1. For each sort $S \in \mathfrak{S}$, a set $\mathcal{M}_{S,p}$
2. For each relation symbol $R \in \mathfrak{R}$ of sort $S_1, \ldots, S_n$ and $p \in B$, a set $[R]_p \subseteq \mathcal{M}_{S_1,p} \times \ldots \times \mathcal{M}_{S_n,p}$

satisfying the following conditions for all $q, p \in B$:

1. If $q \leq p$, then $\mathcal{M}_{S,p} \subseteq \mathcal{M}_{S,q}$. If $q \lhd U$ and a is an element of $\bigcap_{p \in U} \mathcal{M}_{S,p}$ then $a \in \mathcal{M}_{S,q}$.
2. For each relation symbol $R \in \mathfrak{R}$, if $q \leq p$, then $[R]_p \subseteq [R]_q$. If $q \lhd U$ and $a \in \bigcap_{p \in U} [R]_p$ then $a \in [R]_q$.

Similarly to Kripke models, we can define a Heyting valued model over the open sets of the formal topology as follows.

1. For each sort $S \in \mathfrak{S}$, we define $\mathcal{M}_S := \bigcup_{q \in B} \mathcal{M}_{S,q}$.
2. For each sort $S \in \mathfrak{S}$, and $a \in \mathcal{M}_S$, we define $E_S(a) := \{q \in B \mid a \in \mathcal{M}_{S,q}\}$.
3. For each relation symbol $R \in \mathfrak{R}$ of sort $S_1, \ldots, S_n$ and for elements $a_1, \ldots, a_n$ with each a_i belonging to $\mathcal{M}_{S_i}$, we define $[\![R]\!](a_1, \ldots, a_n)$ to be $\{q \in B \mid (a_1, \ldots, a_n) \in [R]_q\}$.

Note that we have chosen the definition of formal topological model so that extent and relation symbols are interpreted as open sets with respect to the formal topology, and so we do get a Heyting valued model over the complete Heyting algebra of open sets.

Note furthermore that for each element p of B, we can define U_p to be the least open set containing p to get a basis for the Heyting algebra of open sets. We also have $U_p \lhd V$ if and only if $p \lhd V$. We can then use propositions 6.3 and 6.5 to explicitly describe the forcing relation with respect to elements p of B as follows.

$p \Vdash_\sigma \bot$	iff	$p \lhd \emptyset$
$p \Vdash_\sigma \varphi \wedge \psi$	iff	$p \Vdash_\sigma \varphi$ and $p \Vdash_\sigma \psi$
$p \Vdash_\sigma \varphi \vee \psi$	iff	$p \lhd \{q \in B \mid q \Vdash_\sigma \varphi$ or $q \Vdash_\sigma \psi\}$
$p \Vdash_\sigma \varphi \rightarrow \psi$	iff	for all $q \leq p$ if $q \Vdash_\sigma \varphi$ then $q \Vdash_\sigma \psi$
$p \Vdash_\sigma \exists x\, \varphi$	iff	$p \lhd \{q \in B \mid \exists a \in \mathcal{M}_q\, q \Vdash_{\sigma[x \mapsto a]} \varphi\}$
$p \Vdash_\sigma \forall x\, \varphi$	iff	for all $q \leq p$ and $a \in \mathcal{M}_q$, $q \Vdash_{\sigma[x \mapsto a]} \varphi$

We say a formal topology is *proper* if it is never the case that $p \lhd \emptyset$. In this case we can see that $p \nVdash_\sigma \bot$ for all $p \in B$.

7 Heyting Valued Models of Second Order Arithmetic

7.1 Standard models

The definition of Heyting valued model is a very general one. This has advantages - we will see later the completeness theorem for complete Heyting algebras, which says that if a formula holds in every Heyting valued model, then it is provable in intuitionistic logic. However, it can also be too flexible. If we are interested in a specific theory, then each time we want a model with a particular property, we need to specify the model completely, carefully ensuring each part is chosen to ensure the axioms of our theory of interest hold. For this reason, it is often useful to have some notion of *standard model* of a theory we are interested in. This is a set recipe for generating Heyting valued models of the theory, from any complete Heyting algebra. Once the complete Heyting algebra has been chosen, the remaining details of the model are already specified in the general definition, ensuring that the axioms of our theory of interest hold. For this course the theory of interest will be either **HAS** or $\mathbf{HA}_\omega$, but the same idea of standard model also appears for many other theories in logic, such as subtheories of classical second order arithmetic, set theory (both constructive and classical), higher order logic and type theory.

7.2 Some new notation regarding variable assignments

It is sometimes notationally inconvenient to always keep track of variable assignments. Hence we introduce some notation to avoid having to write them out every time.

Suppose we are given a signature $(\mathfrak{S}, \mathfrak{O}, \mathfrak{R})$ and a model over the signature $((\mathcal{M}_S)_{S \in \mathfrak{S}}, (\llbracket O \rrbracket)_{O \in \mathfrak{O}}, (\llbracket R \rrbracket)_{R \in \mathfrak{R}})$ and complete Heyting algebra $(P, \bigvee, \wedge, \to)$. We define a new signature by extending the set of operator symbols $\mathfrak{O}$ with a new constant of sort S for each $a \in \mathcal{M}_S$ for each sort $S \in \mathfrak{S}$. Given a variable assignment σ and a formula φ whose free variables are $x_1, \ldots, x_n$, note that $\varphi[x_1, \ldots, x_n / \sigma(x_1), \ldots, \sigma(x_n)]$ is a closed formula over the extended signature.

We use the notational convention

$$\llbracket \varphi[x_1, \ldots, x_n / \sigma(x_1), \ldots, \sigma(x_n)] \rrbracket := \llbracket \varphi \rrbracket_\sigma$$

Note that this defines a unique element $\llbracket \psi \rrbracket$ of P for any closed formula ψ of the extended language.

7.3 Standard models of second order Heyting arithmetic

We first see a standard way to construct Heyting valued models of **HAS** for any complete Heyting algebra $(P, \bigvee, \wedge, \to)$. We have two sorts to deal with: numbers and sets. We will define our Heyting valued models to be *global*. That is the extent for each sort will be defined to be constantly equal to $\top$. Because of this, the interpretation of quantifiers the models can be simplified as follows.

$$\llbracket \exists x^S \, \varphi(x) \rrbracket_\sigma := \bigvee_{a \in \mathcal{M}_S} \llbracket \varphi \rrbracket_{\sigma[x \mapsto a]}$$
$$\llbracket \forall x^S \, \varphi(x) \rrbracket_\sigma := \bigwedge_{a \in \mathcal{M}_S} \llbracket \varphi \rrbracket_{\sigma[x \mapsto a]}$$

Since we have two sorts we just need to define two sets: $\mathcal{M}_N$ for numbers and $\mathcal{M}_S$ for sets. We simply define $\mathcal{M}_N$ to be the set of (external) natural numbers $\mathbb{N}$. We define zero and successor to simply be the same as the external ones.

This leaves the question of what to use for sets $\mathcal{M}_S$ and how to interpret the relations $\in$ and $=$. In classical logic, we think of sets of numbers as either containing a number or not. However, in models of intuitionistic logic it is useful to allow for more possibilities. When we define a set X we can take the truth value of $n \in X$ to be anything, i.e. any element of the Heyting algebra P of truth values. Thinking topologically, we allow X to contain n within some regions of space, while not containing it within others. To formalise this we take $\mathcal{M}_S$ to be the set of functions assigning a truth value to each natural number:

$$\mathcal{M}_S := P^{\mathbb{N}}$$

This idea also suggests to us an interpretation of the membership relation. For $n \in \mathbb{N}$ and $A : \mathbb{N} \to P$, we define

$$[\![n \in A]\!] := A(n)$$

Finally, we need to define the equality relations on each sort. Since we can prove decidable equality for numbers in **HA**, we are forced to just take the simplest definition:

$$[\![n = m]\!] = \begin{cases} \top & n = m \\ \bot & n \neq m \end{cases}$$

The equality on sets is also now fixed. Firstly by extensionality we need to have $\bigwedge_{n \in N} [\![n \in X \leftrightarrow n \in Y]\!] \leq [\![X = Y]\!]$. However, in order to satisfy the axioms for equality we also need to have $[\![X = Y]\!] \leq \bigwedge_{n \in N} [\![n \in X \leftrightarrow n \in Y]\!]$. Hence we are forced to define equality of sets this way:

$$[\![X = Y]\!] \ := \ \bigwedge_{n \in N} [\![n \in X \leftrightarrow n \in Y]\!]$$

Theorem 7.1. *For any complete Heyting algebra $(P, \bigvee, \wedge, \to)$, the Heyting valued model above satisfies all of the theorems of* **HAS** *(i.e. $[\![\varphi]\!]_\sigma = \top$ for any formula φ provable in* **HAS** *and any variable assignment σ).*

Proof. By the soundness theorem for intuitionistic logic, it suffices to show that the model satisfies the axioms of **HAS**. In fact we can use any axiomatisation that ends up giving the same theorems. We do not need to prove first order induction, for example, because we can derive it from second order induction and comprehension.

We can also derive many axioms from the following statement.

$$X = Y \ \leftrightarrow \ (\forall n \, n \in X \leftrightarrow n \in Y) \tag{2}$$

We can easily show that the above axiom holds in our models, since we chose to define equality of sets in a way that forces this to be the case.

Clearly (2) implies extensionality, since it is the right to left direction of the bi-implication. However, it also implies axioms of equality, including all of the equality axioms for sets:

$$\begin{array}{ll} X = X & X = Y \to Y = X \\ X = Y \to (Y = Z \to X = Z) & X = Y \to (n \in X \to n \in Y) \end{array}$$

It still remains to check the equality axioms for the sort of numbers, namely, we need to show

$$\begin{array}{ll} n = n & n = m \to m = n \\ n = m \to (m = k \to n = k) & n = m \to (n \in X \to m \in X) \end{array}$$

However, all are straightforward to check for our chosen definition of equality of numbers, which matches up with the external "true" equality of numbers.

We now check comprehension. We are given a formula $\varphi(x)$ and need to verify that for all variable assignments σ we have

$$[\![\exists X\, \forall n\, (n \in X \leftrightarrow \varphi(n))]\!] = \top$$

Note that it is sufficient to find a specific element $A \in \mathcal{M}_S$ such that $[\![n \in A]\!] = [\![\varphi(n)]\!]$ for all n. However, we can do this by simply defining $A(n) := [\![\varphi(n)]\!]$.

We finally check second order induction. We need to show that for each $A \in \mathcal{M}_S$ we have

$$[\![(0 \in A \wedge \forall n\, (n \in A \to Sn \in A)) \ \to\ \forall n\, n \in A]\!] = \top$$

By unfolding the definitions and using the basic properties of implication and meet in a complete Heyting algebra, this amounts to showing the following inequality for each n:

$$A(0) \wedge \bigwedge_{m \in \mathbb{N}} (A(m) \to A(S(m))) \ \leq\ A(n) \tag{3}$$

We show this by induction on n in the metatheory where we are working. The case $n = 0$ is easy from the definition of meet. Suppose we have already shown (3) for n; we will prove it for Sn. It follows from the inductive hypothesis and the basic properties of meet that we have

$$A(0) \wedge \bigwedge_{m \in \mathbb{N}} (A(m) \to A(S(m))) = A(0) \wedge \bigwedge_{m \in \mathbb{N}} (A(m) \to A(S(m))) \wedge A(n)$$

We can then reason as follows

$$\begin{aligned} A(0) \wedge \bigwedge_{m \in \mathbb{N}} (A(m) \to A(S(m))) \wedge A(n) &\leq A(n) \wedge (A(n) \to A(S(n))) \\ &\leq A(S(n)) \end{aligned}$$

□

It is a characteristic of the standard model approach that we do not have a general completeness theorem. There are examples of formulas that hold in every standard model that are not provable within the theory we are interested in, in this case **HAS**. In fact, for standard models of **HAS**, this is the case for all true first order formulas. In other words, if a first order formula is true externally, in the metatheory where we are working, then it is also true in the model. Results of this kind are known as *absoluteness* theorems.

Theorem 7.2. *Assume the law of excluded middle and that in the Heyting algebra* $(P, \bigvee, \wedge, \to)$ *we have* $\bot \neq \top$*. Let* φ *be a formula over the signature of* **HA** *(i.e. no second order variables free or bound, in particular no second order quantifiers, no* $\in$ *and no set equality relations). Given a function* σ *from free variables to* $\mathbb{N}$ *write* $\varphi[\sigma]$ *for the result of replacing each free variable* x *by* $\sigma(x)$*. Then for any function* σ *from variables to* $\mathbb{N}$*,* $[\![\varphi]\!]_\sigma = \top$ *if and only if* $\varphi[\sigma]$ *is true, and otherwise* $[\![\varphi]\!]_\sigma = \bot$*.*

Proof. This is proved by induction on the set of formulas. We give some of the cases of the induction, with the remainder left as an exercise.

Disjunction Suppose that $\varphi \vee \psi[\sigma]$ is true. If $\varphi[\sigma]$ is true, then $[\![\varphi]\!]_\sigma = \top$ by the inductive hypothesis, and so also $[\![\varphi \vee \psi]\!]_\sigma = \top$. Similarly if $\psi[\sigma]$ is true. Hence in either case we have $[\![\varphi \vee \psi]\!]_\sigma = \top$.

Suppose on the other hand that $\varphi \vee \psi[\sigma]$ is false. Then $\varphi[\sigma]$ and $\psi[\sigma]$ are both false. Hence $[\![\varphi]\!]_\sigma = [\![\psi]\!]_\sigma = \bot$. Hence $[\![\varphi \vee \psi]\!]_\sigma = \bot$. By the law of excluded middle, it follows that if $[\![\varphi \vee \psi]\!]_\sigma = \top$ then $\varphi \vee \psi[\sigma]$ is true.

Implication Suppose that $\varphi \to \psi[\sigma]$ is true. By the law of excluded middle, either $\psi[\sigma]$ is true or $\varphi[\sigma]$ is false. In the former case we have $[\![\varphi]\!]_\sigma = \top$ and in the latter case $[\![\psi]\!]_\sigma = \bot$. In either case we can deduce $[\![\varphi \to \psi]\!]_\sigma = \top$. On the other hand, suppose that $\varphi \to \psi[\sigma]$ is false. Then $\psi[\sigma]$ is false, and by the law of excluded middle, $\varphi[\sigma]$ is true. Hence $[\![\varphi]\!]_\sigma = \bot$ and $[\![\psi]\!]_\sigma = \top$. It follows that $[\![\varphi \to \psi]\!]_\sigma = \bot$. Again using the law of excluded middle, we can also deduce that if $[\![\varphi \to \psi]\!]_\sigma = \top$, then $\varphi[\sigma]$ is true.

Existential quantification Suppose that $(\exists x\, \varphi(x))[\sigma]$ is true. Then there exists a number n such that $\varphi[\sigma[x \mapsto n]]$ is true. Hence $[\![\varphi]\!]_{\sigma[x \mapsto n]} = \top$. It follows that $[\![\exists x\, \varphi]\!]_\sigma = \top$.

Suppose that $(\exists x\, \varphi)[\sigma]$ is false. In this case, for every n in $\mathbb{N}$, $\varphi[\sigma[x \mapsto n]]$ is false. Hence for each n, $[\![\varphi]\!]_{\sigma[x \mapsto n]} = \bot$. Hence $[\![\exists x\, \varphi]\!]_\sigma = \bot$. By the law of excluded middle, it follows that if $[\![\exists x\, \varphi]\!]_\sigma = \top$, then $(\exists x\, \varphi)[\sigma]$ is true. □

7.4 Some examples of standard models of HAS

7.4.1 The trivial Heyting algebra

We now have a way to construct a model of **HAS** given any complete Heyting algebra. We first consider what this looks like in the simplest case: the trivial Heyting algebra $2 = \{\bot, \top\}$.[5]

In this case the sort of sets is modelled by functions $\mathbb{N} \to 2$. However, these correspond precisely to actual subsets of $\mathbb{N}$. As always the sort of numbers is just the external set of numbers $\mathbb{N}$. Combining this with our previous general discussion of trivial Heyting valued models we get the following theorem.

Theorem 7.3. *Let φ be a formula over the signature of* **HAS**. *Then φ holds in the trivial standard model of* **HAS** *if and only if it is true (in the metatheory where we are working).*

[5] Once again we assume the law of excluded middle in order to show this is complete. For a constructive version, we can instead use the power set of a singleton.

7.4.2 Sierpiński space

We now turn to our simplest non trivial example of a complete Heyting algebra: Sierpiński space. We can use this to show that **HAS** is different from second order Peano arithmetic. That is, we show that the law of excluded middle is not provable in **HAS**. We specifically show that the following instance of excluded middle is not provable.

$$\forall X\, 0 \in X \ \vee\ 0 \notin X$$

In order to do this, it suffices to find an element A of $\mathcal{M}_S$ in the standard model on Sierpiński space such that

$$[\![0 \in A \ \vee\ 0 \notin A]\!] \neq \top$$

We can define such an A as follows. The value of $A(n)$ for $n \neq 0$ does not matter, so we can take it to be $\emptyset$, for example. We take $A(0)$ to be the intermediate truth value $\{0\}$.

We can then calculate similarly to before,

$$\begin{aligned}
[\![0 \in A \ \vee\ 0 \notin A]\!] &= A(0) \cup (\{0,1\} \setminus A(0))^{\mathrm{o}} \\
&= \{0\} \cup \{1\}^{\mathrm{o}} \\
&= \{0\} \cup \emptyset \\
&= \{0\} \\
&\neq \{0,1\}
\end{aligned}$$

7.5 Cantor space

In the previous example, we only showed that an instance of excluded middle is not provable in **HAS**. It is a more difficult problem to give an example of a formula that is consistent with **HAS** that contradicts the law of excluded middle. For example, for any formula ψ, we can prove $\neg\neg(\psi \vee \neg\psi)$ in intuitionistic logic. Hence $\neg(\psi \vee \neg\psi)$ is never consistent with **HAS**. However, we have previously seen for example that $A \vee \neg A$ is not provable in intuitionistic logic with a constant relation symbol A.

To see our first non trivial consistency statement with **HAS**, we will use the standard topological model on Cantor space.

Theorem 7.4. *The following formula is consistent with* **HAS***, i.e. if we add it as an axiom it is still impossible to derive* $\perp$*.*

$$\neg(\forall X\, 0 \in X \ \vee\ 0 \notin X)$$

Proof. We will show the formula holds in the standard topological model on Cantor space. It will follow from the soundness theorem that it is consistent. That is, we show

$$[\![\neg(\forall X\, 0 \in X \ \vee\ 0 \notin X)]\!] = \top$$

It suffices to show

$$[\![\forall X\, 0 \in X \ \vee\ 0 \notin X]\!] = \perp$$

Since we are working in a topological model, we can explicitly describe $\bot$ as $\emptyset$. Hence we will derive a contradiction from the assumption that the open set $[\![\forall X\, 0 \in X \vee 0 \notin X]\!]$ contains an element f.

By expanding out the interpretation of universal quantification and using the explicit description of meet in a topological space, we have

$$f \in \left(\bigcap_{A \in \mathcal{M}_S} [\![0 \in A \vee 0 \notin A]\!] \right)^{\circ} \subseteq \bigcap_{A \in \mathcal{M}_S} [\![0 \in A \vee 0 \notin A]\!]$$

We will show this leads to a contradiction by finding A such that f does not belong to $[\![0 \in A \vee \neg 0 \in A]\!]$. We define this A as follows.

$$A(n) = \begin{cases} 2^{\mathbb{N}} \setminus \{f\} & n = 0 \\ \emptyset & \text{otherwise} \end{cases}$$

We clearly have $f \notin [\![0 \in A]\!]$ by construction.

However, we can calculate

$$\begin{aligned} [\![0 \notin A]\!] &= (2^{\mathbb{N}} \setminus A(n))^{\circ} \\ &= \{f\}^{\circ} \\ &= \emptyset \end{aligned}$$

Hence we have $f \notin [\![0 \in A \vee 0 \notin A]\!]$, as we needed. □

7.6 The standard model of HAS on $\mathbb{N}_\infty$

We can formulate **LLPO** in **HAS** as the following formula:

$$\forall x\, (x \in X \vee x \notin X) \wedge \forall x, y\, (x \in X \wedge y \in X \to x = y) \to \\ (\forall x\, 2x \notin X) \vee (\forall x\, 2x + 1 \notin X)$$

Suppose we want to verify that **LLPO** is not provable in **HAS**. Since **HAS** says something is true for each binary sequence with at most one 1, we might guess that it is a good idea to try a topological model where points are exactly the set of binary sequences with at most one 1. It turns out that this works, and we can indeed use the standard topological model on $\mathbb{N}_\infty$ to show the independence of **LLPO**.

Theorem 7.5. LLPO *does not hold in the standard model of* **HAS** *on* $\mathbb{N}_\infty$.

Proof. We define $A : \mathbb{N} \to \mathcal{O}_{\mathbb{N}_\infty}$ as follows:

$$A(n) := \{f \in \mathbb{N}_\infty \mid f(n) = 1\}$$

We can verify that A satisfies the hypotheses of **LLPO** as follows. For each n we have

$$[\![n \in A]\!] = \{f \in \mathbb{N}_\infty \mid f(n) = 1\} \cup \{f \in \mathbb{N}_\infty \mid f(n) \neq 1\}^{\circ}$$

However, $\{f \in \mathbb{N}_\infty \mid f(n) \neq 1\}$ is already open, so we can continue.

$$\begin{aligned}\llbracket n \in A \rrbracket &= \{f \in \mathbb{N}_\infty \mid f(n) = 1\} \cup \{f \in \mathbb{N}_\infty \mid f(n) \neq 1\} \\ &= \mathbb{N}_\infty\end{aligned}$$

We also have for all $n, m \in \mathbb{N}$ such that $n \neq m$

$$\begin{aligned}\llbracket n \in X \wedge m \in X \rrbracket &= \{f \in \mathbb{N}_\infty \mid f(n) = 1\} \cap \{f \in \mathbb{N}_\infty \mid f(m) = 1\} \\ &= \emptyset\end{aligned}$$

It only remains to check that

$$\llbracket (\forall x\, 2x \notin X) \vee (\forall x\, 2x + 1 \notin X) \rrbracket \neq \top$$

We have

$$\begin{aligned}\llbracket \forall x\, 2x \notin X \rrbracket &= \cap_{n \in \mathbb{N}} \{f \in \mathbb{N}_\infty \mid f(2n) = 0\}^{\mathrm{o}} \\ &= \{f \in \mathbb{N}_\infty \mid \forall n\, f(2n) = 0\}^{\mathrm{o}}\end{aligned}$$

Note, however, that if $f(n) = 0$ for all n, then every neighbourhood of f contains a function g such that $g(2n) = 1$ for some n. Hence in this case $f \notin \llbracket \forall x\, 2x \notin X \rrbracket$. But we similarly have $f \notin \llbracket \forall x\, 2x + 1 \notin X \rrbracket$, so we are done. □

8 Heyting Valued Models of Arithmetic with Finite Types

8.1 Partial equivalence relations and H-sets

The theories that we are interested in typically have equality relations for each sort. For such theories it is often conventient to merge the equality relation and extent predicate together into a single binary relation by the following method.

Suppose that we are given signature $(\mathfrak{S}, \mathfrak{O}, \mathfrak{R})$ such that $\mathfrak{R}$ includes a binary relation symbol $=$ on a sort S. Suppose further we have a Heyting valued model for that signature over a complete Heyting algebra $(H, \bigvee, \wedge, \rightarrow)$ and that $=$ satisfies the axioms for equality.

We extend the signature with a new binary relation symbol $\approx$ and interpret it in the Heyting valued model as follows

$$\llbracket a \approx b \rrbracket := E(a) \wedge E(b) \wedge \llbracket a = b \rrbracket$$

Note that the extended Heyting valued model satisfies $\forall x\, \forall y\, x \approx y \leftrightarrow x = y$, which follows directly from the way that universal quantifiers are interpreted in the model. In particular $\approx$ satisfies all the axioms for equality, since $=$ does.

Futhermore, we can recover the extent predicate from $\approx$ by the following equation.

$$E(a) = \llbracket a \approx a \rrbracket$$

We can also see that $[\![\approx]\!]$ has the following properties:

$$[\![a \approx b]\!] = [\![b \approx a]\!]$$
$$[\![a \approx b]\!] \wedge [\![b \approx c]\!] \leq [\![a \approx c]\!]$$

We can think of this as a "Heyting valued" version of the following definition.

Definition 8.1. Let X be any set. A *partial equivalence relation* on X is a binary relation $E \subseteq X \times X$ satisfying the following conditions:

1. For all $x, y \in X$, $E(x, y)$ if and only if $E(y, x)$ (E is symmetric).
2. For all $x, y, z \in X$, if $E(x, y)$ and $E(y, z)$, then $E(x, z)$ (E is transitive).

We refer sets with H-valued relations satisfying the above as H-sets:

Definition 8.2. Let $(H, \bigvee, \wedge, \rightarrow)$ be a complete Heyting algebra. An H-set is a set X, together with a function $\approx : X \times X \rightarrow H$ satisfying the following for all $x, y, z \in X$:

$$x \approx y = y \approx x$$
$$x \approx y \wedge y \approx z \leq x \approx z$$

Given an H-set $(X, \approx)$ we will write $E(x)$ for $x \in X$ as notation for $x \approx x$.

8.2 Singletons in Heyting valued models of HAS and H-sets

In order to motivate an important aspect of the standard model on $\mathbf{HA}_\omega$, we will first take a closer look at how singleton sets work in **HAS**.

Recall that a singleton set is one with exactly one element. We formalise this in **HAS** as follows:

Definition 8.3. We say that X is a *singleton* if there exists a number x such that $x \in X$ and for all numbers x, y such that $x \in X$ and $y \in X$ we have $x = y$.

If we apply this to the standard Heyting valued model of **HAS** on a complete Heyting algebra $(H, \bigvee, \wedge, \rightarrow)$, we obtain the following definition.

Definition 8.4. We say $A : \mathbb{N} \rightarrow H$ is a *singleton* if

1. $\bigvee_{n \in \mathbb{N}} A(n) = \top$
2. for all n, m such that $m \neq n$ we have $A(n) \wedge A(m) = \bot$

Note that given any $n \in \mathbb{N}$, we can define a singleton $\underline{n}$ as follows:

$$\underline{n}(m) := \begin{cases} \top & m = n \\ \bot & m \neq n \end{cases}$$

However, it is important to note that in general these are not the only examples of singletons. Suppose that we have $p, q \in H$ such that $p \wedge q = \bot$ and $p \vee q = \top$. In that case we can define a singleton that can take one of two different values:

$$A(k) := \begin{cases} p & k = n \\ q & k = m \\ \bot & k \neq n \text{ and } k \neq m \end{cases}$$

To give a more concrete example of this we can define the following on Cantor space:

$$A(k) := \begin{cases} \{f : \mathbb{N} \to 2 \mid f(0) = 0\} & k = 0 \\ \{f : \mathbb{N} \to 2 \mid f(0) = 1\} & k = 1 \\ \emptyset & k > 1 \end{cases}$$

We note however, that these examples only work because Cantor space is not connected. We can use connectedness to get some more control over what singletons can look like.

Definition 8.5. We say a complete Heyting algebra $(H, \bigvee, \wedge \to)$ (or more generally any poset with finite meets and least element) is *connected* if for all $p, q \in H$ such that $p \vee q = \top$ and $p \wedge q = \bot$ we have $p = \top$ or $q = \top$.

In particular the lattice of open sets of a topological space is connected as a Heyting algebra if and only if the topological space is connected.

Proposition 8.6. *Suppose that $(H, \bigvee, \wedge \to)$ is a connected complete Heyting algebra. Then for every singleton $A : \mathbb{N} \to H$ in the standard model of* **HAS** *such that $A(n) = \bot$ for $n > 1$, there exists $n \in \{0, 1\}$ such that $[\![A = \underline{n}]\!] = \top$.*

Proof. Define $p := A(0)$ and $q := A(1)$. We then have $p \vee q = \top$ and $p \wedge q = \bot$. By connectedness, we have either $p = \top$ or $q = \top$. In the former case we have $[\![A = \underline{0}]\!] = \top$ and in the latter case we have $[\![A = \underline{1}]\!] = \top$. ☐

More generally, we define singletons for H-sets as follows:

Definition 8.7. Suppose $(X, \approx)$ is an H-set for a complete Heyting algebra $(H, \bigvee, \wedge, \to)$. We define a new H-set $S(X)$ of *singletons* as follows. The underlying set of $S(X)$ consists of functions $A : X \to H$. We define a function $E : S(X) \to H$ by

$$E(A) := \bigvee_{x \in X} (x \approx x \wedge \bigwedge_{y \in X} A(y) \leftrightarrow x \approx y)$$

We then define the H-valued partial equivalence relation on $S(X)$ by

$$A \approx B \ := \ E(A) \wedge E(B) \wedge \bigwedge_{x \in X} A(x) \leftrightarrow B(x)$$

Note that given an H-set $(X, \approx)$ and a map $A : X \to H$, we can define a Heyting valued model for the signature with one sort, and a unary relation symbol A, and a binary relation $=$ satisfying the axioms of reflexivity, symmetry and transitivity. In this case, $E(A)$ is the interpretation of the formula

$$\exists x \, \forall y \, A(y) \leftrightarrow x = y$$

We can show, e.g. by giving a proof in intuitionisitic logic, that this is equivalent to the conjunction of the formulas

$$\exists x \, A(x) \qquad \forall x, y \, A(x) \wedge A(y) \;\to\; x = y \qquad \forall x, y \; x = y \wedge A(x) \;\to\; A(y)$$

The first two state that there is exactly one x such that $A(x)$, while the last is one of the axioms for equality.

For an H-set X, we can define a canonical inclusion $i : X \to S(X)$ such that $x \approx y \leq i(x) \approx i(y)$ for all $x, y \in X$ by

$$i(x)(y) := x \approx y$$

Note that we can define a Heyting valued model for a signature with two sorts, an equality relation on each sort and a unary operation symbol, where we interpret the two sorts as X and $S(X)$, and i as the the unary operation symbol. In this case we can show that the following formulas all evaluate to $\top$ in the interpretation of intuitionistic logic. Firstly, the axiom of equality

$$\forall x^X \, \forall y^X \; x = y \to ix = iy$$

Secondly the following two axioms, stating that i is a bijection

$$\forall z^{S(X)} \, \exists x^X \; z = i(x) \qquad \forall x^X \, \forall y^X \; ix = iy \to x = y$$

Also note that the above Heyting valued model also admits an interpretation for a relation $\in$ of arity $Y, S(Y)$, and we can show that the following axioms are satisfied.

$$\forall z \, \forall z' \, (\forall y \, y \in z \leftrightarrow y \in z') \leftrightarrow z = z'$$

$$\forall z \, \exists! y \, y \in z$$

Definition 8.8. Let $(X, \approx_X)$ and $(Y, \approx_Y)$ be H-sets. A *functional relation from X to Y* is a function $F : X \times Y \to H$ such that for all $x, x' \in X$ and $y, y' \in Y$ we have

$$x \approx x' \;\wedge\; y \approx y' \;\wedge\; F(x, y) \leq F(x', y')$$

and for all $x \in X$, we have

$$x \approx x \leq \bigvee_{y \in Y} y \approx y \wedge F(x, y)$$

and for all $x \in X$ and $y, y' \in Y$, we have

$$x \approx x \wedge y \approx y \wedge y' \approx y' \wedge F(x, y) \wedge F(x, y') \leq y \approx y'$$

We can again view these conditions as certain logical formulas evaluating to $\top$ in a Heyting valued model. Namely, we have two axioms of equality

$$\forall x \, \forall x', x = x' \wedge F(x, y) \to F(x', y) \qquad \forall x \, \forall y \, \forall y' \, y = y' \wedge F(x, y) \to F(x, y')$$

and the following two formulas stating that F is total and single valued.

$$\forall x \, \exists y \, F(x, y) \qquad \forall x \, \forall y \, \forall y' \, F(x, y) \wedge F(x, y') \to y = y'$$

Definition 8.9. We say an H-set $(Y, \approx_Y)$ is *weakly complete* or *Higgs complete* if for every H-set $(X, \approx_X)$ and every functional relation F from X to Y, there is a function $f : X \to Y$ such that for all $x \in X$ we have

$$x \approx x \leq F(x, f(x))$$

Lemma 8.10. *For any H-set $(Y, \approx_Y)$, $S(Y)$ is weakly complete.*

Proof. Suppose we have a functional relation F from $(X, \approx_X)$ to $S(Y)$. We define the function $f : X \to S(Y)$ by taking $f(x) := \lambda y.F(x, i(y))$. We need to check that $x \approx x \leq f(x) \approx f(x)$ and that $x \approx x \leq F(x, \lambda y.F(x, i(y)))$ for all x. In both cases we will use the description of the definitions in terms of formulas evaluating to $\top$ in a Heyting valued model. First, note that to show $x \approx x \leq f(x) \approx f(x)$, it suffices to prove in intuitionistic logic that for all x, $f(x)$ is a singleton. However, we may assume that for every x there exists a unique z such that z is a singleton and we have $F(x, z)$. Since z is a singleton, there is a unique y such that $z = i(y)$. But we have now shown $f(x)$ must be the singleton $i(y)$.

The same argument also shows that $x \approx x \leq F(x, f(x))$. We have shown in intuitionistic logic that for all x, there is a unique z such that $F(x, z)$ and a unique y such that $z = i(y) = f(x)$. However, it follows by the axioms of equality that we have $F(x, f(x))$. ☐

Definition 8.11. Suppose we are given H-sets $(X, \approx_X)$ and $(Y, \approx_Y)$, we define an H-valued partial equivalence relation $\approx_{X \times Y}$ on $X \times Y$ as follows:

$$(x, y) \approx_{X \times Y} (x', y') := x \approx_X x' \wedge y \approx_Y y'$$

Definition 8.12. Suppose we are given H-sets $(X, \approx_X)$ and $(Y, \approx_Y)$, we define an H-valued partial equivalence relation $\approx_{Y^X}$ on Y^X as follows:

$$f \approx g := \bigwedge_{x, x' \in X} x \approx x' \to f(x) \approx g(x')$$

Note that we can define a Heyting valued model on a signature with two three sorts X, Y, Y^X and a term Ap of sort $Y^X, X \to Y$, where we interpret each sort as the corresponding set of the same name, and Ap as function application. In this case the following formula evaluates to $\top$ in the Heyting valued model:

$$\forall f, g \, (\forall x \, fx = gx) \leftrightarrow f = g$$

Lemma 8.13. *If $(Y, \approx_Y)$ is a weakly complete H-set and $(X, \approx_X)$ is any H-set, then $(Y^X, \approx)$ is weakly complete.*

Proof. We need to check that given a functional relation F from an H-set $(Z, \approx_Z$ to $(Y^X, \approx)$ we can find $f : Z \to Y^X$ such that for all z, $z \approx z \leq F(z, f(z))$.

We define a new functional relation G from $Z \times X$ to Y by $G((z,x),y) := \bigvee_{h:X\to Y} F(z,h) \wedge h(x) \approx_Y y$. We will first check that G is a functional relation. We can then deduce that there exists $g : Z \times X \to Y$ such that for all $x \in X$ and $z \in Z$ $(z,x) \approx (z,x) \leq G((z,x), g(z,x))$. We then define $f : Z \to Y^X$ defined by $f(z)(x) := g(z,x)$ and check that for all $x \in X$ we have $x \approx x \leq F(x, f(x))$.

We will check both of the statements above again using the descriptions in terms of formulas evaluating to $\top$ in a Heyting valued model. We first need to prove that $\forall (z,x) \, \exists! y \, G((x,z),y)$. However, we can make the assumption $\forall z \, \exists! f \, F(x,f)$ by the definition of functional relation, and we can make the assumption $\forall f \, \forall x \, \exists! y \, f(x) = y$ by the definition of the equality relation for Y^X. From these assumptions it follows that $\forall (z,x) \, \exists! y \, G((x,z),y)$.

For the second statement we need to check, it suffices to prove in intuitionisitic logic that for all z we have $F(z, f(z))$. We may assume that there exists h of sort Y^X such that $F(z,h)$. We can then show that for all x of sort X we have $g(z,x) = h(x)$. Then since the equality relation on Y^X was chosen to satisfy extensionality, we can deduce $f(z) = h$, and thereby deduce $F(z, f(z))$, as required. □

8.3 Standard Heyting valued models of $\mathbf{HA}_\omega$

Fix a complete Heyting algebra $(H, \bigvee, \wedge, \to)$. We will define the standard Heyting valued model of $\mathbf{HA}_\omega$ on H in such a way that each function sort $\sigma \to \tau$ behaves similarly to functional relations in the standard model of **HAS**. This will ensure that the models are non trivial. It will also have the side effect that the models satisfy both function extensionality and the axiom of unique choice.

We will ensure that each sort of $\mathbf{HA}_\omega$ is interpreted as a weakly complete H-set. We will then use the H-valued partial equivalence relation on the H-set to define both the extent and the equality relation in the Heyting valued model.

8.3.1 The domains of the model, extent and equality

We define $\mathcal{M}_N$ to be $S(\mathbb{N})$, with partial equivalence relation defined as in definition 8.7.

Given finite types σ and τ, we define $\mathcal{M}_{\sigma\times\tau}$ to be $\mathcal{M}_\sigma \times \mathcal{M}_\tau$, with partial equivalence relation defined as in definition 8.11.

We define $\mathcal{M}_{\sigma\to\tau}$ to be the set of functions from $\mathcal{M}_\sigma$ to $\mathcal{M}_\tau$, with partial equivalence relation defined as in definition 8.12. Note that from the definition of the partial equivalence relation, we can see that function extensionality has to hold in the model.

We can also see that unique choice holds, as follows. Let φ be any formula. We want to show $\forall x^\sigma \, \exists! y^\tau \, \varphi(x,y)$ holds in the model. Suppose that $z_1^{\rho_1}, \ldots, z_n^{\rho_n}$

is a list of the free variables occuring in φ not equal to x and y.

$$(c_1,\ldots,c_n) \approx (c'_1,\ldots,c'_n) := c_1 \approx c'_1 \wedge \ldots \wedge c_n \approx c'_n \wedge \\ [\![\forall x^\sigma\, \exists! y^\tau\, \varphi(x,y,c_1,\ldots,c_n)]\!]$$

Finally, note that we have a functional relation from $Z \times \mathcal{M}_\sigma$ to $\mathcal{M}_\tau$, which is just $[\![\varphi]\!]$, viewed as a map $Z \times \mathcal{M}_\sigma \times \mathcal{M}_\tau \to H$. Since we ensured that $\mathcal{M}_\tau$ is weakly complete, this gives us a function from $Z \times \mathcal{M}_\sigma$ to $\mathcal{M}_\tau$. Given any element $(c_1,\ldots,c_n)$ of Z, we can apply the function above to them, to get a function $\mathcal{M}_\sigma \to \mathcal{M}_\tau$, which is what we needed to find an element f of $\mathcal{M}_{\sigma\to\tau}$. By the way that we defined this function, we do indeed have the inequality below, confirming that f does witness unique choice holding in the model.

$$E(c_1) \wedge \ldots \wedge E(c_n) \wedge [\![\forall x\, \exists! y \varphi]\!] \le E(f) \wedge [\![\forall x\, \varphi(x, f(x))]\!]$$

8.3.2 The application operation

We now need to show now to interpret the operator symbols. We define $[\![\mathrm{Ap}]\!] : \mathcal{M}_{\sigma\to\tau} \times \mathcal{M}_\sigma \to \mathcal{M}_\tau$ to be the external function application map. It is important to note that we know this is a well defined function, since we implemented $\mathcal{M}_{\sigma\to\tau}$ so that its underlying set is just the set of functions from $\mathcal{M}_\sigma$ to $\mathcal{M}_\tau$. We also need to check that for all $f : \mathcal{M}_\sigma \to \mathcal{M}_\tau$ and all $x \in \mathcal{M}_\sigma$ we have $f \approx f \wedge x \approx x \le f(x) \approx f(x)$ to satisfy the definition of Heyting valued model. However, this is ensured by definition 8.12. Moreover, given $f, f' : \mathcal{M}_\sigma \to \mathcal{M}_\tau$ and $x, x' \in \mathcal{M}_\sigma$ we have $f \approx f' \wedge x \approx x' \le f(x) \approx f'(x')$, which implies the above, but also shows that the Heyting valued model satisfies the axioms of equality $f = f' \to \mathrm{Ap}(f,x) = \mathrm{Ap}(f',x)$ and $x = x' \to \mathrm{Ap}(f,x) = \mathrm{Ap}(f,x')$.

8.3.3 Constant symbols

We interpret each constant symbol $\mathbf{p}, \mathbf{p}_i, \mathbf{k}, \mathbf{s}$ as the corresponding "external" version. For example, given sorts σ and τ, we need to take $\mathbf{k}$ to be a global element of $\mathcal{M}_{\sigma\to(\tau\to\sigma)}$. This means it needs to be a function $\mathcal{M}_\sigma \to \mathcal{M}_{\tau\to\sigma}$. But $\mathcal{M}_{\tau\to\sigma}$ is itself the set of functions from $\mathcal{M}_\tau$ to $\mathcal{M}_\sigma$. Hence we take $[\![\mathbf{k}]\!]$ to be the function that takes an element a of $\mathcal{M}_\sigma$ and returns the function constantly equal to x. This needs to be a global element, i.e. such that $E([\![\mathbf{k}]\!]) = \top$. This amounts to showing that $a \approx a' \le [\![\mathbf{k}]\!](a) \approx [\![\mathbf{k}]\!](a')$. However, this is the same as showing $a \approx a' \le \bigwedge_{b\in\mathcal{M}_\sigma} [\![\mathbf{k}]\!](a)(b) \approx [\![\mathbf{k}]\!](a')(b)$. Since $[\![\mathbf{k}]\!](a)(b) = a$ for all a and b, this is easy to show. We also need to check that this does satisfy the axiom for $\mathbf{k}$, namely $\forall x\, \forall y\, \mathbf{k}xy = x$. This amounts to showing $a \approx a \wedge b \approx b \le [\![\mathbf{k}]\!]ab \approx a$, which again follows directly from the definition of $[\![\mathbf{k}]\!]$.

We can argue similarly for the other constants $\mathbf{p}, \mathbf{p}_i, \mathbf{s}$.

This only leaves the constant symbols relating directly to numbers, namely $0, S$ and the recursor $\mathbf{r}$. We take $[\![0]\!]$ to be $i(0)$. We need to take $[\![S]\!]$ to be a function $S(\mathbb{N}) \to S(\mathbb{N})$. One way to do this is to note that it suffices to find a functional relation from $S(\mathbb{N}) \to S(\mathbb{N})$. However, by composing with the function $i : \mathbb{N} \to S(\mathbb{N})$ and its inverse, which is a functional relation from $S(\mathbb{N})$ to $\mathbb{N}$, it suffices to find a function from $\mathbb{N}$ to $\mathbb{N}$, which we can take to just be the external successor function. However, we can also explicitly describe $[\![S]\!](A)$ for $A \in S(\mathbb{N})$. It is simply the function $\mathbb{N} \to H$ defined by $\lambda n.A(S(n))$. The induction axioms are proved using external induction in a similar way to the standard Heyting valued model for **HAS**.

Defining $[\![\mathbf{r}]\!]$ amounts to finding a function $S(\mathbb{N}) \to \mathcal{M}_\sigma$, given an element a of $\mathcal{M}_\sigma$ and a function $f : \mathcal{M}_\sigma \times S(\mathbb{N}) \to \mathcal{M}_\sigma$. As before, we observe that it suffices to define a function g from $\mathbb{N} \to \mathcal{M}_\sigma$. We define g by (external) recursion as $g(0) = a$, and $g(Sn) = f(g(n), i(n))$. It is clear that this satisfies the relavant equations by definition.

By the above reasoning we can deduce the soundness theorem for standard Heyting valued models of $\mathbf{HA}_\omega$:

Theorem 8.14. *The above model satisfies the axioms of* $\mathbf{HA}_\omega$*, as well as function extensionality and unique choice.*

9 Independence of choice and omiscience principles over $\mathbf{HA}_\omega$

9.1 The topological model over Cantor space

Theorem 9.1. *Markov's principle does not hold in the standard topological model of* $\mathbf{HA}_\omega$ *on Cantor space.*

Proof. We first construct a global element of the sort $N \to N$. It suffices to find a function $g : \mathbb{N} \to S(\mathbb{N})$. So for each n, we need $g(n)$ to be a function from $\mathbb{N}$ to the open subsets of $2^\mathbb{N}$. We take this to be the "generic" function

$$g(n)(m) := \{f \in 2^\mathbb{N} \mid f(n) = m\}$$

We can then calculate

$$\begin{aligned}
[\![\neg\forall n\, g(n) = 0]\!] &= (2^\mathbb{N} \setminus [\![\forall n\, g(n) = 0]\!])^\circ \\
&= (2^\mathbb{N} \setminus (\bigcap_{n\in\mathbb{N}} [\![g(n) = 0]\!])^\circ)^\circ \\
&= (2^\mathbb{N} \setminus \{\lambda n.0\}^\circ)^\circ \\
&= (2^\mathbb{N} \setminus \emptyset)^\circ \\
&= 2^\mathbb{N}
\end{aligned}$$

However, we also have

$$\begin{aligned}
[\![\exists n\, g(n) = 1]\!] &= \{f \in 2^\mathbb{N} \mid \exists n\, f(n) = 1\} \\
&= 2^\mathbb{N} \setminus \{\lambda n.0\}
\end{aligned}$$

Hence

$$\begin{aligned}[\![\neg\forall n\, g(n) = 0 \;\to\; \exists n\, g(n) = 1]\!] &= (2^{\mathbb{N}} \setminus \{\lambda n.0\})^{\circ} \\ &= 2^{\mathbb{N}} \setminus \{\lambda n.0\} \\ &\neq 2^{\mathbb{N}}\end{aligned}$$

□

Corollary 9.2. LPO *does not hold in the standard topological model of* $\mathbf{HA}_\omega$ *on Cantor space.*

9.2 Independence of countable choice

To show that the axiom of countable choice does not hold in $\mathbf{HA}_\omega$, we first consider a weaker version of the result, that has a simpler proof. We extend the signature of $\mathbf{HA}_\omega$ in include a binary relation symbol A of sort N, N. Write $\mathbf{HA}^+_\omega$ for the theory with the same axioms as $\mathbf{HA}_\omega$ over the larger signature.

We note that the axiom scheme of countable choice now includes some extra formulas, that do not occur for countable choice over $\mathbf{HA}_\omega$, namely those formulas where A occurs somewhere. In particular $\mathbf{AC}^{N,N}$ now includes the following formula.

$$\forall x^N\, \exists y^N\, Axy \;\to\; \exists f^{N\to N}\, \forall x^N\, Axf(x)$$

Theorem 9.3. $\mathbf{AC}^{N,N}$ *is not provable in* $\mathbf{HA}^+_\omega$.

Proof. We work in the standard topological model of $\mathbf{HA}_\omega$ over $I^{\mathbb{N}}$.

In order to make this a Heyting valued model over the extended signature, we need to show how to interpret the binary relation symbol A. We define it as follows:

$$[\![Anm]\!] := \begin{cases} \{f : \mathbb{N} \to I \mid f(n) = 0 \text{ or } f(n) = 1\} & m = 0 \\ \{f : \mathbb{N} \to I \mid f(n) = 2 \text{ or } f(n) = 1\} & m = 1 \\ \emptyset & \text{otherwise} \end{cases}$$

We clearly have

$$[\![\forall x\, \exists y\, Axy]\!] = \top$$

Hence, if $\mathbf{AC}^{N,N}$ held, we would have an element f of $N \to N$ such that

$$\lambda n.1 \in E(f) \wedge [\![\forall x\, Axf(x)]\!]$$

Hence for some connected open neighbourhood U of $\lambda n.1$, we would have $U \subseteq E(f) \cap [\![\forall x\, Axf(x)]\!]$.

It follows that for all n, we have $[\![f(n) = 0 \vee f(n) = 1]\!] \subseteq U$, and $[\![f(n) = 0 \wedge f(n) = 1]\!] \cap U = \bot$. Hence $U \subseteq [\![f(n) = 0]\!] \cup [\![f(n) = 1]\!]$ and $[\![f(n) = 0]\!] \cap [\![f(n) = 1]\!] = \emptyset$. Since U is connected we can deduce that for each n, either

$U \subseteq [\![f(n) = 0]\!]$ or $U \subseteq [\![f(n) = 1]\!]$. That is, f has to correspond to an actual function $\mathbb{N} \to 2$ in the metatheory where we are working.

To get a contradiction from the assumption, we need to show

$$U \not\subseteq [\![\forall x\, Axf(x)]\!]$$

We will show in fact that

$$\lambda n.1 \notin [\![\forall x\, Axf(x)]\!]$$

Suppose that $\lambda n.1 \in [\![\forall x\, Axf(x)]\!]$. In this case it would have a basic open neighbourhood $U_\sigma \subseteq [\![\forall x\, Axf(x)]\!]$ for some finite sequence σ of elements of I. Let n be any number greater than the length of σ.

We assume that $[\![f(n) = 0]\!] = \top$, with a similar proof applying for the case $[\![f(n) = 1]\!] = \top$.

Note that we can easily define a function $g : \mathbb{N} \to I$ such that $g(i) = \sigma(i)$ for $i < |\sigma|$ and such that $g(n) = 2$. Since $[\![f(n) = 0]\!]$, we have $g \in [\![f(n) = 0]\!]$. But since $[\![f(n) = 0]\!] \leq [\![An0]\!]$, this implies $g \in [\![An0]\!]$, which contradicts the definition of $[\![A]\!]$. □

We will now show that $\mathbf{AC}^{N,N}$ also fails for $\mathbf{HA}_\omega$ itself. The rough idea is to combine the above proof for $\mathbf{HA}_\omega^+$ with an idea based on the omniscience principle **LLPO**. **LLPO** says that given a binary sequence f with at most one 1, either $f(2n) = 0$ for all n or $f(2n+1) = 0$. However, if $f(n) = 0$ for all n, then both cases hold, and there is no canonical way to choose one. This leads us to consider the following instance of countable choice. Suppose we have a countable family of binary sequences $f_m : \mathbb{N} \to 2$ for $m \in \mathbb{N}$ and for each m there exists $i \in \{0, 1\}$ such that for all n, $f_m(2n + i) = 0$. Countable choice would imply there is a function $g : \mathbb{N} \to 2$ such that for all m and for all n, $f_m(2n + g(m)) = 0$. We will show that this is not provable in $\mathbf{HA}_\omega$.

The proof also uses some less precise general rules of thumb:

1. If we want to find a topological model where an implication does not hold, it is often helpful to consider a topological space that "looks similar" to the antecedent of the implication.

2. If we want to combine the ideas of two constructions together, it can be useful to combine the topological spaces together in a very simple way, such as binary product.

Theorem 9.4. *The following instance of countable choice is not provable in* $\mathbf{HA}_\omega$*.*

$$\forall f^{N \times N \to N}\, \forall m^N\, \exists i^N\, (i = 0 \vee i = 1) \wedge \forall n^N\, f(m, 2n + i) = 0 \;\to$$
$$\exists g^{N \to N}\, \forall m^N\, (g(m) = 0 \vee g(m) = 1) \wedge \forall n^N\, f(m, 2n + g(m)) = 0$$

Proof. We take X to be the topological space defined as the following subspace of $2^{\mathbb{N}\times\mathbb{N}} \times I^{\mathbb{N}}$.

$$X := \{(h,k) \in 2^{\mathbb{N}\times\mathbb{N}} \times I^{\mathbb{N}} \mid \forall m\, k(m) \geq 0 \to \forall n\, h(m,2n) = 0 \wedge \\ k(m) \geq 2 \to \forall n\, h(m, 2n+1) = 0\}$$

We work over the standard topological model of $\mathbf{HA}_\omega$ on X and define a global element f of $\mathcal{M}_{N\times N\to N}$ as follows.

$$[\![f(m,n) = i]\!] := \{(h,k) \in X \mid h(n,m) = i\}$$

Note that the above does give a functional relation from $\mathcal{M}_N \times \mathcal{M}_N$ to $\mathcal{M}_N$. In particular we can see that each $[\![f(m,n) = i]\!]$ is an open set, since it is the intersection of X with an open set of $2^{\mathbb{N}\times\mathbb{N}} \times I^{\mathbb{N}}$. Hence there is a global element f of $\mathcal{M}_{N\times N\to N}$ satisfying it.

Furthermore, we have the following equalities for all m

$$\begin{aligned} [\![\forall n\, f(m,2n) = 0]\!] &= \{(h,k) \in X \mid k(m) \geq 0\} \\ &= X \,\cap\, 2^{\mathbb{N}\times\mathbb{N}} \times \{k \in I^{\mathbb{N}} \mid k(m) \geq 0\} \\ [\![\forall n\, f(m,2n+1) = 0]\!] &= \{(h,k) \in X \mid k(m) \geq 2\} \\ &= X \,\cap\, 2^{\mathbb{N}\times\mathbb{N}} \times \{k \in I^{\mathbb{N}} \mid k(m) \geq 2\} \end{aligned}$$

However, we can now show there is no element g such that

$$(\lambda n.\lambda m.0, \lambda n.1) \in [\![\forall m^N\, (g(m) = 0 \vee g(m) = 1) \wedge \forall n^N\, f(m, 2n + g(m)) = 0]\!]$$

by a similar argument to theorem 9.3. □

10 Completeness and existence properties

10.1 Completeness theorem for Heyting valued models

In logic, it is common to consider not just soundness theorems (as we have seen so far) but converse versions, known as *completeness* theorems, where we show that if something holds in every model, then it is provable. For the kinds of model we commonly study in intuitionistic logic, we can often do something stronger: for a given theory T, we can find a single *canonical* model such that a closed formula φ is provable in T if and only if it holds in the model. In this section we will see how this works for Heyting valued models on a complete Heyting algebra.

The essential idea is to take the Heyting algebra to be the Lindenbaum-Tarski algebra of a theory T, take the domains $\mathcal{M}_S$ to be terms of sort S and then show $[\![\varphi]\!] = [\varphi]$. However, there are a few issues to deal with to make this precise.

The first problem is that we have only defined Heyting valued model for complete Heyting algebras and the Lindenbaum-Tarski algebra is typically not complete. To deal with this we use a construction called *Dedekind-MacNeille completion*.

Lemma 10.1. *Let $(P, \wedge, \vee, \top, \bot, \rightarrow)$ be a Heyting algebra (not necessarily complete). Then we can construct a complete Heyting algebra $(\overline{P}, \bigwedge, \rightarrow)$ and a function $\iota : P \rightarrow \overline{P}$ such that ι preserves Heyting implication, and any meets and joins that already exist in P.*

Proof. We define an *c-ideal* to be a set $U \subseteq P$ satisfying the following conditions:

1. $\bot \in U$
2. If $p \in U$ and $q \leq p$, then $q \in U$ (i.e. U is downwards closed)
3. If $S \subseteq P$ and $\bigvee S$ already exists in P, then $\bigvee S \in U$

We take $\overline{P}$ to be the set of all c-ideals, ordered by inclusion. We define $\iota(p)$ to be the downwards closure of $\{p\}$.

Note that $\overline{P}$ is a complete Heyting algebra. For example, given c-ideals U and V, we define $U \rightarrow V$ to be

$$U \rightarrow V := \{p \in P \mid \forall q \leq p \, q \in U \rightarrow q \in V\}$$

We can see in particular that $\{p\}^{\leq} \rightarrow \{q\}^{\leq} = \{p \rightarrow q\}^{\leq}$.

Given $S \subseteq \overline{P}$, we can define the join of S by

$$\bigvee S := \left\{ \bigvee X \mid X \subseteq \bigcup S \text{ and } \bigvee X \text{ exists} \right\}$$

One again note that the definition of c-ideal was chosen precisely to ensure that given $S \subseteq P$ such that $\bigvee S$ exists, we have

$$\bigvee \iota(S) = \left\{ \bigvee S \right\}^{\leq}$$

This precisely ensures that any joins that exist in P are preserved by ι. □

Now fix a theory T over a signature $(\mathfrak{S}, \mathfrak{O}, \mathfrak{R})$. If L is the Lindenbaum-Tarski algebra on T, we will define a Heyting valued model on $\overline{L}$.

Given a sort $S \in \mathfrak{S}$, we define $\mathcal{M}_S$ to be the set of all terms of sort S. It is an important point that this really means all terms, not just closed terms. In particular, for the completeness proof to work correctly at quantifiers we will need to use the fact that free variables of sort S are included in $\mathcal{M}_S$. We define $E(t) := \top$, i.e. we construct a global model.

From this definition, it is clear that for each operator symbol $O \in \mathfrak{O}$, we can define $[\![O]\!](t_1, \ldots, t_n)$ simply to be $Ot_1 \ldots t_n$. Similarly, for each relation symbol $R \in \mathfrak{R}$, we can define $[\![R]\!](t_1, \ldots, t_n)$ to be $\iota([Rt_1 \ldots t_n])$ (that is, we send the formula $Rt_1 \ldots t_n$ to the correspond equivalence class $[Rt_1 \ldots t_n]$ in the Lindenbaum-Tarski algebra, and then include it into the Dedekind-MacNeille completion with ι.

Theorem 10.2. *For the Heyting valued model defined above, we have for all formulas φ and all variable assigments σ,*

$$[\![\varphi]\!]_\sigma = \iota([\varphi[\sigma]])$$

Proof. This is proved by induction on formulas. □

10.2 Existence properties in logic

It is a key characteristic of constructive mathematics that proving statements of the form $\exists x\, \varphi(x)$ should require something more than the same statement in classical logic. Namely, we should only be able to prove this if we can explicitly find a witness. We formalise this idea through *existence properties*.

Definition 10.3. A theory T satisfies the *disjunction property* if whenever $T \vdash \varphi \vee \psi$, either $T \vdash \varphi$ or $T \vdash \psi$.

Definition 10.4. A theory T satisfies the *term existence property* if whenever $T \vdash \exists x\, \varphi(x)$, there is a closed term t such that $T \vdash \varphi(t)$.

Definition 10.5. A theory T satisfies the *definable existence property* if whenever φ is a formula whose only free variable is x and $T \vdash \exists x\, \varphi(x)$ there is a formula ψ, whose only free variable is x such that $T \vdash \exists! x\, (\varphi(x) \wedge \psi(x))$.

Note that the term existence property implies the definable existence property, and if a theory T satisfies the definable existence property, we can extend it to a theory with the term existence property by adding a constant symbol c_φ and axiom $\varphi(c_\varphi)$ whenever φ is a formula whose only free variable is x such that $T \vdash \exists! x\, \varphi(x)$.

Definition 10.6. Suppose T has a sort N together with a constant symbol 0 of sort N and a unary operation symbol S of sort $N \to N$. For each natural number $n \in \mathbb{N}$, we define a term $\underline{n}$ of sort N by induction:

$$\underline{0} := 0$$
$$\underline{n+1} := S\underline{n}$$

We say T satisfies the *numerical existence property* if whenever $T \vdash \exists x^N\, \varphi(x)$, there exists $n \in \mathbb{N}$ such that $T \vdash \varphi(\underline{n})$.

Although we only need 0 and S in the signature for the above definition to make sense, in practice we only consider theories that are extensions of **HA**, i.e. every theorem of **HA** is also provable in T.

The disjunction and numerical existence property hold for all theories widely viewed as foundations for constructive mathematics. On the other hand, they can never hold for theories used as foundations of mathematics based on classical logic. We can see this using an argument based on Gödel's incompleteness theorem. If T is a theory in classical logic where we can interpret Peano arithmetic, then we can formalise the statement "T is either consistent or not consistent" inside T, and prove it as a direct instance of the law of excluded middle. However, by Gödel's result we can show that T does not prove either of the disjuncts (as long as T is consistent and recursively axiomatisable).

The status of the definable existence property is less clear. There are examples of theories in classical logic that satisfy it, including Peano arithmetic and any set theory extending **ZF** + $V =$ **OD**. There are also examples of theories used in constructive mathematics that do not satisfy the definable existence property, such as the set theories **IZF** and **CZF**

10.3 Numerical existence property for Heyting arithmetic

We now show a simple technique for proving that a theory has the numerical existence property. To illustrate the idea, we will just show the result for **HA**. However, this is a robust argument that can be generalised and adapted to diverse theories, including for example **HAS** and $\mathbf{HA}_\omega$.

Definition 10.7. Given a complete Heyting algebra $(P, \bigvee, \wedge, \rightarrow)$, we define the *connectification* P^* to have underlying set $P \amalg \{\top_*\}$ with the ordering $\leq$ defined so that $\top_*$ is the top element of the poset and otherwise the ordering agrees with that of P.

Note that P^* is itself a complete Heyting algebra, and it is connected in the following very strong sense.

Lemma 10.8. *Suppose that $S \subseteq P^*$ is such that $\bigvee S = \top_*$. Then for some $p \in S$ we have $p = \top_*$.*

We write $\top_P$ for the "old" top element, which is now strictly below the new top. We make use of the following key lemma, whose proof is left as an exercise.

Lemma 10.9. *The map $\pi : P^* \to P$ sending p to $p \wedge \top_P$ preserves Heyting implication and all meets and joins.*

Theorem 10.10. HA *satisfies the disjunction property and numerical existence property.*

Proof. We take P to be the Dedekind MacNeille completion of the Lindenbaum-Tarski algebra for **HA**. We then define a Heyting valued model over the connectification P^* as follows.

We define $\mathcal{M}$ to be the set of all terms, including terms with free variables, as we did for the completeness theorem. However, we define a non trivial existence predicate as follows:

$$E(t) := \begin{cases} \top_* & \text{if } t = \underline{n} \text{ for some } n \in \mathbb{N} \\ \top_P & \text{otherwise} \end{cases}$$

If we write $[\![\varphi]\!]^P_\sigma$ for the interpretation of φ in the canontical model and $[\![\varphi]\!]^*_\sigma$ in the new model on P^* defined above, then by lemma 10.9 and induction on formulas, we have

$$[\![\varphi]\!]^P_\sigma = [\![\varphi]\!]^*_\sigma \wedge \top_P$$

Now assume that $\mathbf{HA} \vdash \exists n\, \varphi(n)$. One can show that the axioms of **HA** hold in the model (exercise). Hence by the soundness theorem for intuitionistic logic, we have (for any variable assignment σ)

$$[\![\exists n\, \varphi]\!]^*_\sigma = \top_*$$

By lemma 10.8 and the interpretation of existential quantifiers it follows that there exists a term t such that

$$E(t) \wedge [\![\varphi(t)]\!]^*_\sigma = \top_*$$

However, it follows that $E(t) = \top_*$, and therefore that $t = \underline{n}$ for some $n \in \mathbb{N}$.

Now by the observation above, we can deduce

$$[\![\varphi(\underline{n})]\!]^P_\sigma = \top_P$$

However, finally we can apply the completeness theorem to show $\mathbf{HA} \vdash \varphi(\underline{n})$, as required. $\square$

11 Partial Combinatory Algebras

So far we have considered models of intuitionistic logic based on topology and more generally on complete Heyting algebras. The idea behind realizability is to instead work with an abstract notion of computation, called *partial combinatory algebra* (pca), which is the topic of this section.

11.1 Some notation and terminology for partial functions

The "computable functions" of a pca will be in particular partial functions, so we will first define some notation and terminology that will help us to deal with partial functions better.

Definition 11.1. Let X and Y be sets. A *partial function* from X to Y is a subset f of $X \times Y$ such that whenever $(x, y) \in f$ and $(x, y') \in f$ we have $y = y'$. We write $f : X \rightharpoondown Y$.

We write $f(x)\downarrow$ to mean that there exists $y \in Y$ (necessarily unique) such that $(x, y) \in f$. We say f *is defined at* x. We write $f(x) = y$ to mean $(x, y) \in f$. Note in particular that writing $f(x) = y$ implicitly implies $f(x)\downarrow$. However, we often write out $f(x)\downarrow$ explicitly anyway to draw attention to the fact that f is defined at x.

Suppose that $f : X \rightharpoondown Y$ and $g : Z \rightharpoondown Y$. For $x \in X$ and $z \in Z$, we write $f(x) \simeq g(z)$ to mean that $f(x)\downarrow$ if and only if $g(z)\downarrow$, and that if they are defined, then $f(x) = g(z)$.

In particular, if X is a set with one element, say $*$, then partial functions $X \rightharpoondown Y$ are just subsets of singletons of Y. In this case $f(*)\downarrow$ means the subsingleton is inhabited, and $f(*) = y$ means the subsingleton contains the element y (and so is an actual singleton).

Also note that we have the following propositions.

Proposition 11.2. *Any function $X \to Y$ is in particular a partial function.*

Proposition 11.3. *Given partial functions $f : X \rightharpoondown Y$ and $g : Y \rightharpoondown Z$, we define the composition $g \circ f$ to be the partial function $X \rightharpoondown Z$, which is defined at x if and only if $f(x)\downarrow$ and $g(f(x))\downarrow$, and in this case $(g \circ f)(x) = g(f(x))$.*

11.2 Partial applicative structures

Before defining partial combinatory algebras, we will first define a weaker notion, partial applicative structure.

Definition 11.4. A *partial applicative structure* (pas) is a set $\mathcal{A}$, together with a partial binary operation $\cdot : \mathcal{A} \times \mathcal{A} \rightharpoonup \mathcal{A}$.

Definition 11.5. Let $\mathcal{A}$ be a pas. We say a partial function $f : \mathcal{A} \rightharpoonup \mathcal{A}$ is *representable* if there exists $a \in \mathcal{A}$ such that for all $b \in \mathcal{A}$ we have $f(b) \simeq a \cdot b$.

The intuitive idea behind representable partial functions is *programs-as-data*. The elements of $\mathcal{A}$ are mathematical objects that we wish to discuss (e.g. we can take $\mathcal{A} = \mathbb{N}$ or $\mathcal{A} = \mathbb{N}^{\mathbb{N}}$). We think of representable functions as programs that operate on this data. By definition, every representable function can be "coded" or "represented" as an element of $\mathcal{A}$ itself.

Example 11.6. Any total binary operation is in particular a partial binary operation. Applying this to $\mathbb{N}$ with the binary operation $+$, we get that the representable partial function $\mathbb{N} \rightharpoonup \mathbb{N}$ are precisely the total functions of the form $\lambda m.m + n$ for each $n \in \mathbb{N}$.

Example 11.7. Note that we can define an injective function i from closed terms of $\mathbf{HA}_\omega$ of sort $N \to N$ to $\mathbb{N}$. Given a closed term t of sort $N \to N$, we write $\ulcorner t \urcorner$ for the corresponding element of $\mathbb{N}$. Define the partial operation by setting $n \cdot m = k$ when $n = \ulcorner t \urcorner$ and $\mathbf{HA}_\omega \vdash tm = k$ (and otherwise it is undefined). An encoding of terms as natural numbers in this way is often referred to as a *Gödelnumbering*

Definition 11.8. Let $\mathcal{A}$ be a partial applicative structure and assume that we have a supply of free variables $x_1, x_2, \ldots$. The set of $\mathcal{A}$*-terms* is defined inductively as follows.

1. Every free variable x_i is an $\mathcal{A}$-term
2. Every element of $\mathcal{A}$ is an $\mathcal{A}$-term
3. If s and t are $\mathcal{A}$-terms, then so is $s \cdot t$

Definition 11.9. Given an $\mathcal{A}$-term t and a list of free variables $x_1, \ldots, x_n$ including all those that occur free in t, we define a partial function $|t| : \mathcal{A}^n \rightharpoonup \mathcal{A}$ by induction on terms.

1. $|x_i|$ is a total function, defined as $|x_i|(a_1, \ldots, a_n) := a_i$
2. $|a|$ is a total function, defined as $|a|(a_1, \ldots, a_n) := a$
3. We define $|s \cdot t|(a_1, \ldots, a_n)$ to be defined only when $|s|(a_1, \ldots, a_n)$ and $|t|(a_1, \ldots, a_n)$ are defined, and in this case

$$|s \cdot t|(a_1, \ldots, a_n) \simeq |s|(a_1, \ldots, a_n) \cdot |t|(a_1, \ldots, a_n)$$

In particular if an $\mathcal{A}$-term is closed (contains no free variables), then it defines a partial function on a set with one element.

We will deal with $\mathcal{A}$-terms quite a lot, so we will use some notational conventions to make them easier to write:

1. We will often omit the binary operator, writing $s \cdot t$ simply as st.
2. Write $\cdot$ left associatively. That is, we write rst to mean $(rs)t$.
3. We will often write $|t|$ simply as t, omitting the bars.
4. We write substitution into $\mathcal{A}$-terms the same way as substitution of terms in intuitionistic logic (i.e. $t[x/s]$ where s is an $\mathcal{A}$-term).
5. If t is a closed term and $t \downarrow$, then there is a unique element a of $\mathcal{A}$ such that $t \simeq a$. We say t *denotes* a.

11.3 Partial combinatory algebras

Definition 11.10. A *partial combinatory algebra* (pca) is a pas $(\mathcal{A}, \cdot)$ such that there are $\mathbf{k}, \mathbf{s} \in \mathcal{A}$ such that for all $a, b, c \in \mathcal{A}$,

1. $\mathbf{k}a \downarrow$ and $\mathbf{k}ab = a$ (in particular $\mathbf{k}ab$ is always defined).
2. $\mathbf{s}a \downarrow$, $\mathbf{s}ab \downarrow$ and

$$\mathbf{s}abc \simeq ac(bc)$$

We say a partial combinatory algebra is *non-trivial* if $\mathbf{s} \neq \mathbf{k}$.

Similarly to the definition of $\mathbf{HA}_\omega$, we can use $\mathbf{k}$ to generate constant functions and $\mathbf{s}$ to define the "dependent composition" of two functions. In particular, we have the following proposition.

Proposition 11.11. *Let $\mathcal{A}$ be a pca. Then every constant function $\mathcal{A} \to \mathcal{A}$ is representable. If partial functions $f : \mathcal{A} \rightharpoonup \mathcal{A}$ and $g : \mathcal{A} \rightharpoonup \mathcal{A}$ are representable, then so is $g \circ f$. The identity function $\mathcal{A} \to \mathcal{A}$ is also representable (by $\mathbf{i} := \mathbf{skk}$).*

Similarly to in $\mathbf{HA}_\omega$ we can prove a λ-abstraction lemma. Since we are now working with partial functions, we phrase the lemma a bit differently. It is particularly important that $\lambda x.t$ is always a defined term, even if (for example) t is already closed and not defined.

Lemma 11.12. *Let $\mathcal{A}$ be a pca, and t an $\mathcal{A}$-term whose only free variable are included in the list $x, y_1, \ldots, y_n$. Then there is an $\mathcal{A}$-term $\lambda x.t$ such that all the free variables of $\lambda x.t$ are included in the list $y_1, \ldots, y_n$, for all $b_1, \ldots, b_n \in \mathcal{A}$, we have $(\lambda x.t)(b_1, \ldots, b_n) \downarrow$ and for all $a, b_1, \ldots, b_n \in \mathcal{A}$ we have*

$$(\lambda x.t)(b_1, \ldots, b_n)a \simeq t(a, b_1, \ldots, b_n)$$

Proof. We define $\lambda x.t$ and check that it works by induction on terms.

1. If $t = x$, we define $\lambda x.t$ to be $\mathbf{i}$
2. If $t = y_i$, we define $\lambda x.t$ to be $\mathbf{k}y_i$
3. If $t = a$ for $a \in \mathcal{A}$, we define $\lambda x.t$ to be $\mathbf{k}a$
4. If $t = r \cdot s$, we define $\lambda x.t$ to be $\mathbf{s}(\lambda x.r)(\lambda x.s)$

□

We can use λ-abstraction to define $\mathbf{y}$ combinators. This construction can only be carried out in "untyped" settings, such as in pcas. It does not work for $\mathbf{HA}_\omega$, for example. This is not the most common definition of $\mathbf{y}$-combinator. It is however the most appropriate definition when working with a partially defined application operator, and therefore the only definition we will consider in this course.

Theorem 11.13. *For any pca, $\mathcal{A}$, there is an element $\mathbf{y} \in \mathcal{A}$ with the following properties. For all $a \in \mathcal{A}$, we have $\mathbf{y}a\downarrow$, and for all $b \in \mathcal{A}$, we have*

$$\mathbf{y}ab \simeq a(\mathbf{y}a)b$$

Proof. We first define

$$t := \lambda x.\lambda y.\lambda z.y(xx)z$$

We then define $\mathbf{y}$ to be the closed term tt.

First note that we can calculate

$$\begin{aligned}\mathbf{y} &= tt \\ &= (\lambda x.\lambda y.\lambda z.y(xxy)z)t \\ &\simeq \lambda y.\lambda z.y(tty)z\end{aligned}$$

In particular, we can see that $\mathbf{y}$ is defined, since λ-abstraction is always defined and so it does denote an element of $\mathcal{A}$.

Now given $a \in \mathcal{A}$, we have

$$\mathbf{y}a \simeq \lambda z.a(tty)z$$

Again using the fact that λ-abstraction is always defined, we can see in particular that $\mathbf{y}a\downarrow$.

Finally for any $b \in \mathcal{A}$, we have

$$\begin{aligned}\mathbf{y}ab &\simeq a(tta)b \\ &= a(\mathbf{y}a)b\end{aligned}$$

□

Neither of the two examples of partial applicative structures we have seen are pcas. For example 11.6 it is clear that constant functions are not representable, for example. It is harder to see why example 11.7 is not a pca. For now we just remark that the main difficultly is that the $\mathbf{s}$ combinator of $\mathbf{HA}_\omega$ is a typed operation of sort $(N \to N \to N) \to (N \to N) \to N \to N$. However, to get a pca we would need a term of sort $N \to N$ that takes Gödelnumbers as input, which turns out to be more difficult.

11.4 Two examples of term pcas

To get some non trivial examples of pcas, we will use *term models*. We first define what we mean by term.

Definition 11.14. We define the set $\mathcal{T}$ of *pca-terms* to be inductively defined as follows:

1. $\mathbf{k}$ is a pca-term
2. $\mathbf{s}$ is a pca-term
3. If s and t are pca-terms, then so is $s \cdot t$

We define our first term model as follows. Let $\sim$ be the smallest equivalence relation on $\mathcal{T}$ satisfying the following conditions for all $r, s, t \in \mathcal{T}$:

1. $\mathbf{k}rs \sim r$
2. $\mathbf{s}rst \sim rt(st)$
3. If $r \sim s$, then $r \cdot t \sim s \cdot t$
4. If $s \sim t$, then $r \cdot s \sim r \cdot t$

Example 11.15. We can give $\mathcal{T}/{\sim}$ the structure of a pca, where we define $[r] \cdot [s] := [r \cdot s]$, $\mathbf{k} := [\mathbf{k}]$ and $\mathbf{s} := [\mathbf{s}]$.

The preceding example has the advantage of being simple. However, it is difficult to say anything concrete about it, even to show it is non trivial.

We therefore consider a second term model, which has a bit more complicated definition, but is easier to describe concretely. It also has the advantage of being easy to implement on electronic computers, especially so with the help of modern programming languages with a built in notion of inductively defined type, such as Haskell.

We first define a subset of terms that we refer to as normal.

Definition 11.16. We say a term t is *reducible* if it satisfies any of the conditions below:

1. $t = \mathbf{k}rs$ for some terms r and s
2. $t = \mathbf{s}rsu$ for some terms r, s and t
3. $t = rs$, for some terms r and s where either r or s is reducible

We say a term t is *normal* if it is not reducible. We write the set of all normal terms as $\mathcal{T}_0$.

Definition 11.17. We define a ternary relation on terms, natural numbers and normal terms. Given a term t, number n and normal form s, we will write the relation as $t \to_n s$ and say *t reduces to s at stage n*. We define the relation as the smallest one satisfying the conditions below.

1. If t is normal, then $t \to_n t$.
2. If either t or r is reducible and $t \to_n t'$ and $r \to_n r'$ and $t'r' \to_n v$, then $tr \to_n v$ (note that $t'r'$ is either normal or one of the two remaining cases below).
3. If t and r are normal, then $\mathbf{k}tr \to_n t$.
4. If t, r, u and v are normal, then $\mathbf{s}tru \to_{n+1} v$ whenever $tu(ru) \to_n v$.

Note that we have chosen the definition to ensure we have the following lemma.

Lemma 11.18. *Reducibility at n satisfies the following statements.*

1. *If $t \to_n t'$, then $t \to_{n+1} t'$.*
2. *If $t \to_n t'$ and $t \to_n t''$, then $t' = t''$.*
3. *If t is normal, then $t \to_n t$ for all n.*

We can now give $\mathcal{T}_0$ the structure of a pas by defining $t \cdot s$ to be r if there exists n such that $t \cdot s \to_n r$ (and $t \cdot s$ is undefined if there is no such n).

Theorem 11.19. *$\mathcal{T}_0$ with the above application operator is a pca.*

Proof. By lemma 11.18 we can see that the above definition does give a well defined partial binary operator $\cdot$. It remains to check the axioms for $\mathbf{k}$ and $\mathbf{s}$. First note that if r and t are normal, then so are $\mathbf{k}r$, and $\mathbf{s}rt$, so we do have $\mathbf{k}r\downarrow$ and $\mathbf{s}rt\downarrow$. We clearly have $\mathbf{k}rt = r$. It only remains to check $\mathbf{s}rtu \simeq ru(tu)$ for normal forms r, t, u. However, by definition $\mathbf{s}rtu \to_{n+1} v$ for some normal form v if and only if $ru(tu) \to_n v$. We can see from this that $\mathbf{s}rtu\downarrow$ if and only if $ru(tu)\downarrow$, and when they are defined we clearly have $\mathbf{s}rtu = ru(tu)$, as required. □

11.5 Extended pcas and computable functions

We now have a non trivial example of a pca, giving us some kind of notion of computation. However, at the moment the only thing we can "compute" is normal terms. For this reason, we also consider a variant of the definition of partial combinatory algebra that allows us to encode numbers as a subset of the pca.

Definition 11.20. An extended pca, or pca$^+$ is a partial combinatory algebra $(\mathcal{A}, \cdot)$ with the additional constants $\mathbf{p}$, $\mathbf{p}_0$, $\mathbf{p}_1$, 0, S, P, and $\mathbf{d}$, satisfying the axioms below. Given $n \in \mathbb{N}$ we will write $\underline{n}$ for the element of $\mathcal{A}$ defined recursively by $\underline{0} := 0$ and $\underline{n+1} := S\underline{n}$.

1. For all $a, b \in \mathcal{A}$, $\mathbf{p}ab\downarrow$, $\mathbf{p}_0(\mathbf{p}ab) \simeq a$ and $\mathbf{p}_1(\mathbf{p}ab) \simeq b$
2. $S\underline{n}\downarrow$ and $S\underline{n} \neq \underline{0}$ for any $n \in \mathbb{N}$

3. For all $n \in \mathbb{N}$, $P(S\underline{n}) = \underline{n}$

4. For all $n, m \in \mathbb{N}$ and all $a, b \in \mathcal{A}$, $\mathbf{d}\underline{n}\underline{m}ab = a$ if $n = m$ and $\mathbf{d}\underline{n}\underline{m}ab = b$ if $n \neq m$

In fact, any non trivial pca can be made into an extended pca. We won't give full details here, but we give some definitions to illustrate how the proof works. First note that we can define $\mathbf{p}$, $\mathbf{p}_0$ and $\mathbf{p}_1$ with the correct properties, using the λ-abstraction lemma.

$$\mathbf{p} := \lambda x.\lambda y.\lambda z.zxy \qquad \mathbf{p}_0 := \lambda w.w(\lambda x.\lambda y.x) \qquad \mathbf{p}_1 := \lambda w.w(\lambda x.\lambda y.y)$$

We then implement the boolean values $\bot$ and $\top$ with the properties that $\bot ab = a$ and $\top ab = b$:

$$\top := \mathbf{k} \qquad \bot := \lambda x.\lambda y.y \; (= \mathbf{ki})$$

We use the above to define an encoding of the natural numbers:

$$0 := \mathbf{p}\bot\bot \qquad S := \lambda x.\mathbf{p}\top x$$

With these definitions it is clear we can define $P := \mathbf{p}_1$, and we can define a simpler version of $\mathbf{d}$, that decides whether or not a number is equal to 0, as $\mathbf{d}_0 := \mathbf{p}_0$. Note that by repeatedly using $\mathbf{d}_0$ we can find $a \in \mathcal{A}$ such that for all $b \in \mathcal{A}$, and all $n, m \in \mathbb{N}$,

$$ab\underline{n}\underline{m} = \begin{cases} \top & n = m = 0 \\ \bot & (n = 0 \wedge m \neq 0) \vee (n \neq 0 \wedge m = 0) \\ \lambda x.\lambda y.b(Pn)(Pm)xy & \text{otherwise} \end{cases}$$

We can then define $\mathbf{d} := \mathbf{y}a$.

However, there are often many ways to choose the pca$^+$ structure, and it can be useful to ensure the structure is simple to describe explicitly. To get a term model with a simple description of the pca$^+$ structure we adjust the definition of $\mathcal{T}_0$ from the previous section by explicitly adding new constants. Namely we define terms as follows:

Definition 11.21. We inductively define the set $\mathcal{T}^+$ of *pca$^+$-terms* by

1. $\mathbf{k}$ is a pca$^+$-term
2. $\mathbf{s}$ is a pca$^+$-term
3. If s and t are pca$^+$-terms, then so is $s \cdot t$
4. 0 is a pca$^+$-term
5. S is a pca$^+$-term
6. P is a pca$^+$-term
7. $\mathbf{d}$ is pca$^+$-term

Definition 11.22. We say a term t is *reducible* if it satisfies any of the conditions below:

1. $t = \mathbf{k}rs$ for some terms r and s
2. $t = \mathbf{s}rsu$ for some terms r, s and t
3. $t = rs$, for some terms r and s where either r or s is reducible
4. $t = \mathbf{p}_0(\mathbf{p}rs)$ or $t = \mathbf{p}_1(\mathbf{p}rs)$ for some terms r and s
5. $t = P(S\underline{n})$ for some $n \in \mathbb{N}$
6. $t = \mathbf{d}\underline{n}\underline{m}rs$ for some $n, m \in \mathbb{N}$ and terms r and s

We say a term t is *normal* if it is not reducible. We write the set of all normal terms as $\mathcal{T}_0^+$.

We adjust the definition of reducibility at stage n (definition 11.17) by adding the following clauses:

1. If t and s are normal terms, then $\mathbf{p}_0(\mathbf{p}ts) \to_n t$ and $\mathbf{p}_1(\mathbf{p}ts) \to_n s$
2. If $m \in \mathbb{N}$, then $P(S\underline{m}) \to_n \underline{m}$
3. If $m, l \in \mathbb{N}$ and t, s are normal terms, then $\mathbf{d}\underline{m}\underline{l}ts \to_n t$ if $m = l$ and $\mathbf{d}\underline{m}\underline{l}ts \to_n s$ if $m \neq l$

As before, for normal terms t and s, we define $t \cdot s$ to be r if there exists $n \in \mathbb{N}$ such that $t \cdot s \to_n r$ (and $t \cdot s$ is undefined if there is no such n). As before, this gives us a pca, and we can make this into an extended pca by interpreting each constant as "itself."

Definition 11.23. We say a partial function $\mathbb{N} \rightharpoonup \mathbb{N}$ is *computable* relative to a pca$^+$ $(\mathcal{A}, \cdot, \mathbf{s}, \mathbf{k}, \mathbf{p}, \mathbf{p}_0, \mathbf{p}_1, 0, S, P, \mathbf{d})$ if there exists $a \in \mathcal{A}$ such that the following holds. For all $n \in \mathbb{N}$, if $f(n)\downarrow$, then $t\underline{n}\downarrow$ and $t\underline{n} = \overline{f(n)}$, and whenever there exists $m \in \mathbb{N}$ such that $t\underline{n} = \underline{m}$, we have $f(n)\downarrow$ with $\overline{f(n)} = m$.

In fact if a partial function $\mathbb{N} \rightharpoonup \mathbb{N}$ is computable relative to $\mathcal{T}_0^+$, then it is computable relative to *any* extended pca, justifying the following definition.

Definition 11.24. We refer to partial functions $\mathbb{N} \rightharpoonup \mathbb{N}$ that are computable relative to $\mathcal{T}_0^+$ simply as *computable.*

The class of computable partial functions can be defined many different ways. This is the main definition we will see in this course, but in other courses you may have seen definitions in terms of abstract versions of mechanical or electronic computers, such as Turing machines or register machines, definitions in terms of definability in the λ-calculus or combinatory logic, or the definition of recursive functions. In general we will follow the rule of thumb referred to as the Church-Turing thesis,[6] which states that any time we can write down a precise but possibly informally stated procedure for finding a number, we could implement it as a computable function in the formal sense.

[6]Not to be confused with the axiom Church's thesis that we will also study in this course.

12 Realizability

12.1 Realizability models for intuitionistic logic

A classic way to motivate realizability is formalise an intuitive idea known as the Brouwer-Heyting-Kolmogorov, or BHK, interpretation of intuitionistic logic. The BHK interpretation asserts for each logical connective what it means to constructively justify the truth of formulas built from that connective. The most important cases to consider are disjunction, existential quantifiers and implication. To assert a disjunction $\varphi \vee \psi$ it is essential, according to BHK, to either assert φ or to assert ψ. In particular, we need to have a choice of φ or ψ, that tells us which we going to show. Similarly, to assert $\exists x\, \varphi(x)$ we need to have both a witness t and then we must also assert $\varphi(t)$ for that particular t. To assert an implication $\varphi \to \psi$, we must have an effective rule that tells us how to produce witnesses of ψ from witnesses of φ. In realizability we make the vague term "rule" precise by defining to be a representable function in a pca.

Another viewpoint of realizability is that the powerset of a pca $\mathcal{A}$, $\mathcal{P}(\mathcal{A})$ can play a similar role to the complete Heyting algebra in Heyting valued models. In fact the definition of realizability model for intuitionistic logic with only relation symbols is identical to that of Heyting valued model.

Fix a signature $(\mathfrak{S}, \mathfrak{O}, \mathfrak{R})$ and a partial combinatory algebra $\mathcal{A}$. Will will also assume we have pairing and projection constants $\mathbf{p}$, $\mathbf{p}_0$ and $\mathbf{p}_1$. These could be either constructed in a general pca, or part of the structure of an extended pca. We will write $\top$ for $\lambda x.\lambda y.x$ and $\bot$ for $\lambda x.\lambda y.y$.

Definition 12.1. A *realizability model* for $(\mathfrak{S}, \mathfrak{O}, \mathfrak{R})$ with pca $\mathcal{A}$ consists of the following data:

1. For each sort $S \in \mathfrak{S}$ a set $\mathcal{M}_S$
2. For each sort S a function $E_S : \mathcal{M}_S \to \mathcal{P}(\mathcal{A})$
3. For each relation symbol $R \in \mathfrak{R}$ of sort $S_1, \ldots, S_n$ a function $[\![R]\!] : \mathcal{M}_{S_1} \times \ldots \times \mathcal{M}_{S_n} \to \mathcal{P}(\mathcal{A})$
4. For each operator symbol $O \in \mathfrak{O}$ of sort $S_1, \ldots, S_n \to T$ a function $[\![O]\!] : \mathcal{M}_{S_1} \times \ldots \times \mathcal{M}_{S_n} \to \mathcal{M}_T$ such that there exists $e \in \mathcal{A}$ with the following property. For all $x_1, \ldots, x_n$ with $x_i \in \mathcal{M}_{S_i}$ and all $a_1, \ldots, a_n$ with $a_i \in E_{S_i}(x_i)$ we have $ea_1 \ldots a_n \downarrow$ and $ea_1 \ldots a_n \in E_T([\![O]\!](x_1, \ldots, x_n))$. We say that e *realizes* or *tracks* $[\![O]\!]$.

Finally, we have a non triviality condition, that for every sort $S \in \mathfrak{S}$ there exists an element a of $\mathcal{M}_S$ where $E_S(a)$ is non empty.

We now describe the interpretation of intuitionistic logic in realizability models.

As for Heyting valued models we assign for each term t of sort T and each variable assignment α, an element $[\![t]\!]_\alpha$ of $\mathcal{M}_T$. For realizability models, we further ensure that these assignments are realized, in the following sense. Suppose that $x_1^{S_1}, \ldots, x_n^{S_n}$ is a list of variables including all those that occur free in t. Then there is $e \in \mathcal{A}$ with the following property. For every free variable assignment α and for $f_1, \ldots, f_n$ with each $f_i \in E_{S_i}(\alpha(x_i))$, we will ensure that $ef_1 \ldots f_n \downarrow$, with $ef_1 \ldots f_n \in [\![t]\!]_\alpha$. We define the value of $[\![t]\!]_\alpha$ exactly the same as for Heyting valued models, namely, by induction with

$$[\![x]\!]_\alpha := \alpha(x_i)$$
$$[\![Ot_1 \ldots t_n]\!]_\alpha := [\![O]\!]([\![t_1]\!]_\alpha, \ldots, [\![t_n]\!]_\alpha)$$

Lemma 12.2. *There exists a realizer $e \in \mathcal{A}$ for each term t with free variables amongst $x_1, \ldots, x_n$, as described above.*

Proof. Fix a list of variables $x_1, \ldots, x_n$. We will show how to construct realizers for terms containing only free variables in the list $x_1, \ldots, x_n$. $[\![x_i]\!]_\alpha$ is realized by $\lambda x_1, \ldots, x_n.x_i$. Suppose we are already given realizers $f_1, \ldots, f_m$ for $t_1, \ldots, t_m$, and that $[\![O]\!]$ is realized by $e \in \mathcal{A}$. Then $[\![Ot_1 \ldots t_m]\!]_\alpha$ is realized by

$$\lambda x_1, \ldots, x_n.e(f_1 x_1 \ldots x_n) \ldots (f_m x_1 \ldots x_n)$$

□

We now show how to define truth values of formulas. For each formula φ and variable assignment, α, we will define $[\![\varphi]\!]_\alpha \subseteq \mathcal{A}$.

$$\begin{aligned}
[\![Rt_1 \ldots t_n]\!]_\alpha &:= [\![R]\!]([\![t_1]\!]_\alpha, \ldots, [\![t_n]\!]_\alpha) \\
[\![\bot]\!]_\alpha &:= \emptyset \\
[\![\varphi \wedge \psi]\!]_\alpha &:= \{e \mid \mathbf{p}_0 e \in [\![\varphi]\!]_\alpha \text{ and } \mathbf{p}_1 e \in [\![\psi]\!]_\alpha\} \\
[\![\varphi \vee \psi]\!]_\alpha &:= \{e \mid (\mathbf{p}_0 e = \top \text{ and } \mathbf{p}_1 e \in [\![\varphi]\!]_\alpha) \text{ or } (\mathbf{p}_0 e = \bot \text{ and } \mathbf{p}_1 e \in [\![\psi]\!]_\alpha)\} \\
[\![\varphi \to \psi]\!]_\alpha &:= \{e \mid \text{for all } f \in [\![\varphi]\!]_\alpha,\ ef \downarrow \text{ and } ef \in [\![\psi]\!]_\alpha\} \\
[\![\exists x^S\, \varphi]\!]_\alpha &:= \bigcup_{a \in \mathcal{M}_S} \{e \mid \mathbf{p}_0 e \in E_S(a) \text{ and } \mathbf{p}_1 e \in [\![\varphi]\!]_{\alpha[x \mapsto a]}\} \\
[\![\forall x^S\, \varphi]\!]_\alpha &:= \bigcap_{a \in \mathcal{M}_S} \{e \mid \text{for all } f \in E_S(a),\ ef \downarrow \text{ and } ef \in [\![\varphi]\!]_{\alpha[x \mapsto a]}\}
\end{aligned}$$

Even more so than for Heyting valued models, it can be useful to instead phrase this definition using forcing notation. For a given variable assignment α, we write $e \Vdash_\alpha \varphi$ to mean $e \in [\![\varphi]\!]_\alpha$. We can then describe $\Vdash_\alpha$ explicitly as follows.

$e \nVdash_\alpha \bot$		always
$e \Vdash_\alpha \varphi \wedge \psi$	iff	$\mathbf{p}_0 e \Vdash_\alpha \varphi$ and $\mathbf{p}_1 e \Vdash_\alpha \psi$
$e \Vdash_\alpha \varphi \vee \psi$	iff	either $\mathbf{p}_0 e = \top$ and $\mathbf{p}_1 e \Vdash_\alpha \varphi$, or $\mathbf{p}_0 e = \bot$ and $\mathbf{p}_1 e \Vdash_\alpha \psi$
$e \Vdash_\alpha \varphi \to \psi$	iff	if $f \Vdash_\alpha \varphi$, then $ef \downarrow$ and $ef \Vdash_\alpha \psi$
$e \Vdash_\alpha \exists x^S\, \varphi$	iff	there exists $a \in \mathcal{M}_S$ such that $\mathbf{p_0} e \in E_S(a)$ and $\mathbf{p}_1 e \Vdash_{\alpha[x \mapsto a]} \varphi$
$e \Vdash_\alpha \forall x^S\, \varphi$	iff	for all $a \in \mathcal{M}_S$ and for all $f \in E_S(a), ef \downarrow$ and $ef \Vdash_{\alpha[x \mapsto a]} \varphi$

Before showing the soundness theorem, we first introduce some notation. For any pca $\mathcal{A}$, we can encode finite lists of elements of $\mathcal{A}$ as single elements of $\mathcal{A}$, in a similar manner to how natural numbers are implemented in arbitrary pcas. We write the encoding of lists as $[e_1, \ldots, e_n] \in \mathcal{A}$, which is defined by induction on the length of the list by

$$[] := \mathbf{p}\bot\bot$$
$$[e_1, \ldots, e_{n+1}] := \mathbf{p}\top(\mathbf{p}e_{n+1}[e_1, \ldots, e_n])$$

If Γ is a finite set of formulas that can be written as $\{\varphi_1, \ldots, \varphi_n\}$ and α a variable assignment, then we write $e \Vdash_\alpha \Gamma$ to mean $e = [e_1, \ldots, e_n]$ and for each i, $e \Vdash_\alpha \varphi_i$.

Theorem 12.3. *If $\Gamma \vdash \varphi$ is provable in intuitionistic logic, and $x_1^{S_1}, \ldots, x_n^{S_n}$ is a list of variables including all of those that occur free in Γ and φ, then we can find $e \in \mathcal{A}$ such that for all variable assignments α, all $f_1, \ldots, f_n$ with $f_i \in E_{S_i}(\alpha(x_i))$, and all g such that $g \Vdash_\alpha \Gamma$, we have $ef_1 \ldots f_n g \downarrow$ and*

$$ef_1 \ldots f_n g \Vdash_\alpha \varphi$$

Proof. To a large extent, this is very similar to the soundness theorem for Heyting valued models. We again work by induction on the definition of provability, proving some cases as an example, and leaving the rest as exercises.

The case $\vee E$ Suppose that we have deduced $\Gamma \vdash \chi$ from the hypotheses $\Gamma \vdash \varphi \vee \psi$, $\Gamma, \varphi \vdash \chi$ and $\Gamma, \psi \vdash \chi$. By the inductive hypothesis, we may assume we already have realizers for $\Gamma \vdash \varphi \vee \psi$, $\Gamma, \varphi \vdash \chi$ and $\Gamma, \psi \vdash \chi$. Suppose that $x_1, \ldots, x_n$ is a list of free variables including any occurring in Γ or ψ. We will assume that the list $x_1, \ldots, x_n$ also includes any free variables occurring in φ and ψ: we can do this using the non triviality condition in a very similar manner to Heyting valued models. Hence we have $e, f, g \in \mathcal{A}$ such that for all variable assignments α and all $h_1, \ldots, h_n$ such that $h_i \in E_{S_i}(\alpha(x_i))$ and all k, l, m such that $k \Vdash_\alpha \Gamma$, $m \Vdash_\alpha \varphi$ and $l \Vdash_\alpha \psi$, we have

$$eh_1 \ldots h_n k \Vdash_\alpha \varphi \vee \psi \tag{4}$$
$$fh_1 \ldots h_n km \Vdash_\alpha \chi \tag{5}$$
$$gh_1 \ldots h_n kl \Vdash_\alpha \chi \tag{6}$$

It follows from (4) that either $\mathbf{p}_0(eh_1 \ldots h_n k) = \top$ and $\mathbf{p}_1(eh_1 \ldots h_n k) \Vdash_\alpha \varphi$ or $\mathbf{p}_0 eh_1 \ldots h_n k = \bot$ and $\mathbf{p}_1(eh_1 \ldots h_n k) \Vdash_\alpha \psi$. In the former case we have $\mathbf{p}_0(eh_1 \ldots h_n k)fg = f$ and in the latter $\mathbf{p}_0(eh_1 \ldots h_n k)fg = g$. In the former case, we have by (5) that $fh_1 \ldots h_n k\mathbf{p}_1(eh_1 \ldots h_n k) \Vdash_\alpha \chi$, and in the latter case, by (6) we have $gh_1 \ldots h_n k\mathbf{p}_1(eh_1 \ldots h_n k) \Vdash_\alpha \chi$. Hence, in either case we have

$$\mathbf{p}_0(eh_1 \ldots h_n k)fgh_1 \ldots h_n k(\mathbf{p}_1(eh_1 \ldots h_n k)) \Vdash_\alpha \chi$$

So we can take our required realizer to be

$$\lambda x_1, \ldots, x_n.\lambda y.\mathbf{p}_0(ex_1 \ldots x_n y)fgx_1 \ldots x_n y(\mathbf{p}_1(ex_1 \ldots x_n y))$$

The case $\exists I$ Suppose we have deduced $\Gamma \vdash \exists x\, \varphi$ from $\Gamma \vdash \varphi[x/t]$. Let $x, y_1, \ldots, y_n$ be a list of free variables including all those occurring in Γ and φ. By the inductive hypothesis, we have e such that for all variable assignments α, and all $f_1, \ldots, f_n$ with $f_i \in E(\alpha(x_i))$, and all $k\mathcal{A}$ such that $k \Vdash_\alpha \Gamma$, we have

$$e f_1 \ldots f_n k \Vdash_\alpha \varphi[x/t]$$

In the interpretation of terms we ensured that there is $g \in \mathcal{A}$ such that for all such α and f_i we have $g f_1 \ldots f_n \in [\![t]\!]_\alpha$. Hence we can deduce

$$\mathbf{p}(g f_1 \ldots f_n k)(e f_1 \ldots f_n k) \Vdash_\alpha \exists x\, \varphi$$

We can therefore take our realizer to be

$$\lambda x_1, \ldots, x_n.\lambda y.\mathbf{p}(g x_1 \ldots x_n y)(e x_1 \ldots x_n y)$$

□

12.2 Realizability models for $\mathbf{HA}_\omega$

We now show how to construct standard realizability models for $\mathbf{HA}_\omega$. Instead of giving one definition, we will give two different models for each pca$^+$, $\mathcal{A}$, that we will refer to as the *intensional* model and the *extensional* model.

12.2.1 The intensional model

We define $\mathcal{M}_S$ and E_S for each sort S of $\mathbf{HA}_\omega$ by induction on finite types. We will ensure that $\mathcal{M}_S \subseteq \mathcal{A}$, and then define $E_S(e) := \{e\}$.

We define $\mathcal{M}_N$ to be the copy of the standard natural numbers in $\mathcal{A}$, i.e. $\{\underline{n} \mid n \in \mathbb{N}\}$.

Suppose we have already defined $\mathcal{M}_\sigma$ and $\mathcal{M}_\tau$. We define $\mathcal{M}_{\sigma\times\tau}$ and $\mathcal{M}_{\sigma\to\tau}$ as follows.

$$\mathcal{M}_{\sigma\times\tau} := \{\mathbf{p}ef \in \mathcal{A} \mid e \in \mathcal{M}_\sigma \text{ and } f \in \mathcal{M}_\tau\}$$

$$\mathcal{M}_{\sigma\to\tau} := \{e \in \mathcal{A} \mid \text{for all } f \in \mathcal{M}_\sigma,\ ef\downarrow \text{ and } ef \in \mathcal{M}_\tau\}$$

We interpret equality to be the standard one, namely

$$[\![x = y]\!] := \begin{cases} \mathcal{A} & \text{if } x = y \\ \emptyset & \text{otherwise} \end{cases}$$

We interpret application by

$$[\![\mathrm{Ap}]\!](e, f) := e \cdot f$$

We interpret each constant symbol as the corresponding pca$^+$ constant in $\mathcal{A}$. Namely,

$$\begin{array}{lll} [\![\mathbf{s}^{\sigma,\tau,\rho}]\!] := \mathbf{s} & [\![\mathbf{k}^{\sigma,\tau}]\!] := \mathbf{k} & \\ [\![\mathbf{p}^{\sigma,\tau}]\!] := \mathbf{p} & [\![\mathbf{p}_0^{\sigma,\tau}]\!] := \mathbf{p}_0 & [\![\mathbf{p}_1^{\sigma,\tau}]\!] := \mathbf{p}_1 \\ [\![0]\!] := 0 & [\![S]\!] := S & \end{array}$$

This only leaves the recursor combinator, $\mathbf{r}$, which we define by combining the fixpoint combinator $\mathbf{y}$ with $\mathbf{d}$ from the pca$^+$ structure. Our first (incorrect) attempt would be to define $[\![\mathbf{r}^{\sigma,\tau}]\!]$ as,

$$\mathbf{r} := \lambda x.\lambda y.\mathbf{y}(\lambda u.\lambda z.\mathbf{d}0zx(y(u(Pz))(Pz)))$$

This would certainly give us that for all x, y and z we have

$$\mathbf{r}xyz \simeq \mathbf{d}0zx(y(\mathbf{r}xy(Pz))(Pz))$$

and in particular

$$\begin{aligned}\mathbf{r}xy(Sz) &\simeq \mathbf{d}0(Sz)x(y(\mathbf{r}xy(P(Sz)))(Pz)) \\ &\simeq y(\mathbf{r}xyz)z\end{aligned}$$

Note, however, that there is no way to show that for the above definition $\mathbf{r}xy0\downarrow$, since in order to show this, we would need that $\mathbf{d}00x(y(\mathbf{r}xy(P0))(P0))\downarrow$, which can only be defined once the subterm $\mathbf{r}xy(P0)$ is defined. Hence, we adjust the above definition to get the one below, exploiting the fact that λ-terms are always defined.

$$[\![\mathbf{r}^{\sigma}]\!] := \lambda x.\lambda y.\mathbf{y}(\lambda u.\lambda z.\mathbf{d}0z(\mathbf{k}x)(\lambda w.y(u(Pz))(Pz))\top)$$

Theorem 12.4. *The above model satisfies all the axioms (and hence all the theorems) of* $\mathbf{HA}_{\omega}$.

Proof. Note that we have interpreted equality in way that makes it straightforward to check the axioms of identity. By the above reasoning, each of the equations associated with the constants also holds in the model. It only remains to check that induction holds. Similarly to the recursor, we can do this using the $\mathbf{y}$ combinator. This time it is slightly simpler. Suppose we are given $e \in \mathcal{A}$ such that

$$e \Vdash_{\alpha} \varphi(0) \wedge \forall n\, \varphi(n) \to \varphi(Sn)$$

We can then show by induction, that for all $n \in \mathbb{N}$ we have

$$\mathbf{y}(\lambda u.\lambda z.(\mathbf{d}0z(\mathbf{k}(\mathbf{p}_0 e))(\lambda w.\mathbf{p}_1 e(u(Pz)))\top))\underline{n} \Vdash_{\alpha} \varphi(\underline{n})$$

It follows that

$$\lambda x.\mathbf{y}(\lambda u.\lambda z.(\mathbf{d}0z(\mathbf{k}(\mathbf{p}_0 x))(\lambda w.\mathbf{p}_1 x(u(Pz)))\top)) \Vdash \varphi(0) \wedge \forall n\, \varphi(n) \to \varphi(Sn) \;\to\; \forall n\, \varphi(n)$$

□

Theorem 12.5. *Markov's principle holds in the intensional models of* $\mathbf{HA}_{\omega}$.

Proof. First, note that we can show the following in general for a pca$^+$, $\mathcal{A}$. Suppose that e is such that for all $n \in \mathbb{N}$ $e\underline{n}$ and either $e\underline{n} = \top$ or $e\underline{n} = \bot$. If there is an n such that $e\underline{n} = \top$, then we can find the first such instance computably and uniformly in e. We deduce this as a special case of the stronger statement that we can compute f such that for all n, if there is $k \in \mathbb{N}$ such that $e\underline{n+k} = \top$, then $f\underline{n}\downarrow$ and $f\underline{n} = \underline{k}$ where k is the least such value.

$$f := \mathbf{y}(\lambda u.\lambda z.((ez)(\mathbf{k}\underline{0})\lambda w.(S(u(Sz))))\top)$$

However, we can now see that Markov's principle is realized by

$$\lambda x.\mathbf{y}(\lambda u.\lambda z.((xz)(\mathbf{k}\underline{0})\lambda w.(S(u(Sz))))\top)\underline{0}$$

□

Theorem 12.6. *The axiom of choice* $\mathbf{AC}^{\sigma,\tau}$ *holds in intensional models of* $\mathbf{HA}_\omega$ *for all sorts* σ *and* τ.

Proof. Suppose that we have

$$e \Vdash_\alpha \forall x^\sigma\, \exists y^\tau\, \varphi(x, y)$$

Note that this gives us an element e' of $\mathcal{M}_{\sigma\to\tau}$ defined by $\lambda z.\mathbf{p}_0(ez)$. Meanwhile, we also have

$$\lambda z.\mathbf{p}_1(ez) \Vdash_\alpha \forall x^\sigma\, \varphi(x, e'x)$$

We can hence see that $\mathbf{AC}^{\sigma,\tau}$ is realized by $\lambda x.\mathbf{p}(\lambda z.\mathbf{p}_0(ez))(\lambda z.\mathbf{p}_1(ez))$. □

12.2.2 The extensional model

We now give a second way to construct realizability models of $\mathbf{HA}_\omega$, this time ensuring that we also get extensionality and the axiom of unique choice.

We again define $\mathcal{M}_S$ and E_S for each sort S by induction on finite types. At each stage we will ensure that for each $x \in \mathcal{M}_S$, $E_S(x)$ is inhabited. It is possible to construct these as quotients of the ones in the intensional model, but to get a more concrete description of the model will will define them directly.

As before, we define $\mathcal{M}_N := \mathbb{N}$ and $E_N(n) := \{n\}$.

For products, we define

$$\mathcal{M}_{\sigma\times\tau} := \mathcal{M}_\sigma \times \mathcal{M}_\tau$$
$$E_{\sigma\times\tau}((x, y)) := \{e \mid \mathbf{p}_0 e \in E_\sigma(x) \text{ and } \mathbf{p}_1 e \in E_\tau(y)\}$$

Finally, for function types, we recall that $e \in \mathcal{A}$ *tracks* $F : \mathcal{M}_\sigma \to \mathcal{M}_\tau$ if for all $x \in \mathcal{M}_\sigma$ and all $a \in E_\sigma(x)$ we have $ea\downarrow$ with $ea \in E_\tau(F(x))$. We then define

$$\mathcal{M}_{\sigma\to\tau} := \{F \in \mathcal{M}_\tau^{\mathcal{M}_\sigma} \mid \exists e \in \mathcal{A}\ e \text{ tracks } F\}$$
$$E_{\sigma\to\tau}(F) := \{e \in \mathcal{A} \mid e \text{ tracks } F\}$$

Note that every element of $\mathcal{M}_{\sigma\to\tau}$ is in particular a function from $\mathcal{M}_\sigma$ to $\mathcal{M}_\tau$, so we can define application simply by

$$[\![\mathrm{Ap}]\!](F, x) := F(x)$$

We again define equality to be the standard one, namely,

$$[\![x = y]\!] := \begin{cases} \mathcal{A} & \text{if } x = y \\ \emptyset & \text{otherwise} \end{cases}$$

Theorem 12.7. *The above model satisfies all the axioms of* $\mathbf{HA}_\omega$*, as well as the axioms of unique choice, function extensionality and Markov's principle.*

Proof. The proof that the axioms of $\mathbf{HA}_\omega$ and Markov's principle hold is essentially the same as for the intensional model.

For function extensionality, suppose that t and s are both terms of sort $\sigma \to \tau$, that α is any variable assignment and that for some $e \in \mathcal{A}$ we have $e \Vdash_\alpha \forall x^\sigma\, tx = sx$. Then, by the definition of the model, $[\![s]\!]_\alpha$ and $[\![t]\!]_\alpha$ are both functions from $\mathcal{M}_\sigma$ to $\mathcal{M}_\tau$, and for any $a \in \mathcal{M}_\sigma$, and $f \in E_\sigma(a)$, we have $ef \Vdash_{\alpha[x\mapsto a]} tx = sx$, which by the interpretation of equality and application implies $[\![t]\!]_\alpha(a) = [\![s]\!]_\alpha(a)$. Since this applies for arbitrary a we have, by function extensionality in our meta theory, that $[\![t]\!]_\alpha = [\![s]\!]_\alpha$. However, by the interpretation of equality, this implies that for any $g \in \mathcal{A}$, we have $g \Vdash_\alpha s = t$. Hence, we have a realizer for function extensionality given simply by $\mathbf{i}$.

We can show unique choice holds by similar arguments. □

12.3 Standard realizability models for HAS

We now show how to construct realizability models of **HAS**. As before, we take the natural number sort $\mathcal{M}_N$ to simply be the usual one. This leaves the problem of how to define the sort of sets. One way to motivate the definition, is to view definition as similar to the one for Heyting valued models of **HAS**. Instead of defining a set to be a function from $\mathbb{N}$ to the Heyting algebra, we take sets to be functions $A : \mathbb{N} \to \mathcal{P}(\mathcal{A})$. We recall that the standard Heyting valued models were defined to be global. It turns out that the way to do this in realizability models is to take E_S to be constantly equal to the same, non empty set. We will take this to be $E_S(A) := \{\top\}$.

We interpret membership and equality, again similarly to Heyting valued models.

$$[\![n \in A]\!] := A(n)$$

$$[\![A = B]\!] := \bigcap_{n\in\mathbb{N}} \{e \in \mathcal{A} \mid \forall f \in A(n)\, \mathbf{p}_0 e\underline{n} f \in B(n) \;\wedge\; \forall f \in B(n)\, \mathbf{p}_1 e\underline{n} f \in A(n)\}$$

Theorem 12.8. *The above model satisfies all the axioms (and hence all the theorems) of* **HAS**.

Proof. Just as for Heyting valued models of **HAS**, extensionality follows directly from the interpretation of equality in the model.

We check comprehension. Let φ be any formula. We define $A : \mathbb{N} \to \mathcal{P}(\mathcal{A})$ by $A(n) := [\![\varphi(\underline{n})]\!]_\alpha$. We then have the following, immediately from our interpretation of $\in$.

$$\mathbf{k}(\mathbf{pii}) \Vdash_\alpha \forall n\, n \in A \leftrightarrow \varphi(n)$$

It only remains to check second order induction. As before, we will use the **y**-combinator. Suppose that

$$e \Vdash_\alpha 0 \in A \;\wedge\; \forall n\, n \in A \to Sn \in A$$

Similarly to before, we have

$$\lambda x.\mathbf{d}0x(\mathbf{k}(\mathbf{p}_0 e))(\lambda y.(\mathbf{p}_1 e)x)\top \Vdash_\alpha \forall x\, x \in A$$

□

Theorem 12.9. *Every standard realizabilty of* **HAS** *satisfies the uniformity principle.*

Proof. Suppose that $e \Vdash \forall X\, \exists n\, \varphi(X, n)$. Let $A \in \mathcal{M}_S$. We define the model so that $E_S(A) = \{\top\}$. It follows that $e\top\downarrow$ and $e\top \Vdash \exists n\, \varphi(X, n)$. Hence there exists $n \in \mathcal{M}_N = \mathbb{N}$ such that $\mathbf{p}_0(e\top) \in E_N(n)$ and $\mathbf{p}_1(e\top) \Vdash \varphi(A, n)$. We hence have $\mathbf{p}_0(e\top) = \underline{n}$. However, $\mathbf{p}_0(e\top)$ can only be equal to at most one numeral, and this value of independent of A, i.e. will get the same numeral $\mathbf{p}_0(e\top)$ for any $A \in \mathcal{M}_S$. It follows that we have

$$\lambda x.\mathbf{p}(\mathbf{p}_0(x\top))(\lambda y.\mathbf{p}_1(x\top)) \Vdash \forall X\, \exists n\, \varphi(X, n) \;\to\; \exists n\, \forall X\, \varphi(X, n)$$

□

13 Kleene Realizability

13.1 Encoding $\mathcal{T}_0$ in arithmetic

So far the main non trivial examples of pca that we've seen are $\mathcal{T}_0$ and $\mathcal{T}_0^+$, the pca of normal forms and inside first reduction and the extended version. As it stands, this is not something we can formalise in **HA**, or even in $\mathbf{HA}_\omega$; we cannot even define the set of normal forms in this setting. Hence it is useful to have way to view normal terms as numbers. That is, we need a Gödelnumbering of terms. To define this, first note that we can define an bijection from $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$. There are various ways to do this. For example, note that the function $\langle -, - \rangle : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ given by the following definition is definable and provably a bijection already in **HA**.

$$\langle n, m \rangle := \frac{1}{2}(n + m)(n + m + 1) + m$$

We can then define an injective function $\ulcorner - \urcorner : \mathcal{T}^+ \to \mathbb{N}$ as follows:

$$\begin{array}{llll} \ulcorner 0 \urcorner := 0 = \langle 0, 0\rangle & \ulcorner S \urcorner := \langle 0,1\rangle & \ulcorner P \urcorner := \langle 0,2\rangle & \ulcorner \mathbf{d} \urcorner := \langle 0,3\rangle \\ \ulcorner \mathbf{k} \urcorner := \langle 0,4\rangle & \ulcorner \mathbf{s} \urcorner := \langle 0,5\rangle & & \\ \ulcorner \mathbf{p} \urcorner := \langle 0,6\rangle & \ulcorner \mathbf{p}_0 \urcorner := \langle 0,7\rangle & \ulcorner \mathbf{p}_1 \urcorner := \langle 0,8\rangle & \\ & s \cdot t := \langle 1, \langle \ulcorner s \urcorner, \ulcorner t \urcorner \rangle\rangle & & \end{array}$$

Theorem 13.1. *There is a total computable function $f : \mathbb{N} \to \mathbb{N}$ such that*

$$f(\langle\langle n, m\rangle, k\rangle) = \begin{cases} \langle 1, \ulcorner r \urcorner\rangle & n = \ulcorner t \urcorner \text{ for } t \in \mathcal{T}_0^+ \text{ and } t\underline{m} \to_k r \\ \langle 0, 0\rangle & \text{otherwise} \end{cases}$$

As it turns out, we can now define application in **HA**, in the following sense.

Theorem 13.2. *There is a formula $\varphi(l, m, n)$ in the language of arithmetic such that $\varphi(l, m, n)$ is true if and only if $l = \ulcorner s \urcorner$ for a term s, $m = \ulcorner t \urcorner$ for a normal term t, and s reduces to t at stage n. Furthermore, we may assume φ is a* negative *formula, i.e. it does not contain disjunction or existential quantifiers and that* **HA** *proves $\forall l, m, n\, \varphi(l, m, n) \vee \neg\varphi(l, m, n)$.*

Definition 13.3. The axiom of Church's thesis $\mathbf{CT}_0!$ is the following sentence of $\mathbf{HA}_\omega$.

$$\forall f^{N \to N}\, \exists e^N\, \forall n^N\; e \cdot n \downarrow \wedge\, e \cdot n = fn$$

Lemma 13.4. *There is an element $\mathbf{e}$ of $\mathcal{T}_0^+$ with the property that for all $t \in \mathcal{T}^+$, we have $\mathbf{e}\ulcorner t \urcorner = t'$ if t evaluates to t' at some stage n, and otherwise is undefined. In particular, if t is normal, then $\mathbf{e}\ulcorner t \urcorner = t$.*

Proof. First note that using the decidability combinator for numbers, $\mathbf{d}$, and the fact that projection is computable, we can ensure that $\mathbf{e}\langle 0, 0\rangle = 0$, that $\mathbf{e}\langle 0, 1\rangle = S$, that $\mathbf{e}\langle 0, 2\rangle = P$, and similarly for all the other constants. Using the $\mathbf{y}$ combinator we can also ensure $\mathbf{e}\langle 1, \langle n, m\rangle\rangle \simeq \mathbf{e}n(\mathbf{e}m)$. In particular, we can see that $\mathbf{e}n$ is defined whenever n is the Gödelnumber for a normal term. □

13.2 The first Kleene algebra

The first Kleene algebra, $\mathcal{K}_1$ is often seen as the key example of pca$^+$ and is the one that was originally used for realizability, the general theory of pca's being a later generalisation of this example. $\mathcal{K}_1$ is the first example we will see of an ω-pca, so we can assume that the underlying set is equal to $\mathbb{N}$, and that in the pca$^+$ structure 0 is the actual zero of $\mathbb{N}$ and that S represents the actual successor function. Furthermore, it is defined so that the representable partial functions are exactly the computable partial functions. In fact these features characterise $\mathcal{K}_1$ uniquely up to isomorphism, using a non trivial argument due to Blum. However, for this course we will just show how to define an extended pca with these properties.

Definition 13.5. We define $\mathcal{K}_1$ to be the pca with underlying set $\mathbb{N}$, and application defined as follows:

$$n \cdot m := \begin{cases} l & \text{if } \mathbf{e}\underline{n}\underline{m} = \underline{l} \\ \text{undefined} & \text{otherwise} \end{cases}$$

Note that we cannot use $\mathbf{k}$ and $\mathbf{s}$ from $\mathcal{T}_0^+$ directly, but we can still define them as follows. For $\mathbf{k}$ we use the fact that pairing is computable.

$$\mathbf{k} := \ulcorner \lambda x.\langle 1, \langle \ulcorner \mathbf{k} \urcorner, x \rangle \rangle \urcorner$$

We define $\mathbf{s}$ using $\mathbf{e}$ from lemma 13.4. Our first attempt would be $\mathbf{s}_0$, as defined below.

$$\mathbf{s}_0 := \ulcorner \lambda x.\lambda y.\lambda z.\mathbf{e}(\mathbf{e}xz)(\mathbf{e}yz) \urcorner$$

As an element of $\mathcal{T}_0^+$, this takes the Gödelnumbers of two normal terms as input, and then evaluates them. However, we need to ensure that $\mathbf{s}x$ and $\mathbf{s}xy$ are Gödelnumbers for terms, rather than the terms themselves. Hence, we again use the computability of the pairing operator, and define

$$\mathbf{s}_1 := \ulcorner \lambda x.\lambda y.\langle 1, \langle \langle 1, \langle \mathbf{s}_0, x \rangle \rangle, y \rangle \rangle \urcorner$$
$$\mathbf{s} := \ulcorner \lambda x.\langle 1, \langle \mathbf{s}_1, x \rangle \rangle \urcorner$$

Theorem 13.6. *Church's thesis holds in both standard realizability models of* $\mathbf{HA}_\omega$ *on* $\mathcal{K}_1$.

Proof. For the intensional model, $\mathcal{M}_{N \to N}$ is precisely the set of $n \in \mathbb{N}$ representing a total function $\mathbb{N} \to \mathbb{N}$. So given such an n, it's clear that if we want a realizer for $\exists e^N, \forall m^N \, e \cdot m\downarrow \wedge e \cdot m = fm$, then the first component should just be n. We still need to show how to find the second component, which should be a realizer for $\forall m^N \, n \cdot m\downarrow \wedge n \cdot m = fm$. The key point is that from theorem 13.2 the statement $\forall m^N \, n \cdot m\downarrow \wedge e \cdot m = fm$ is equivalent to one of the form $\forall m \, \exists k \, \varphi(n, m, k)$ where φ is negative. Furthermore, in the presence of Markov's principle (which always holds in the realizability models we consider in this course), this is equivalent to $\forall m \, \neg \forall k \, \neg \varphi(n, m, k)$, which is entirely negative. However, negative formulas ψ are always *self-realizing*. That is, we can find f such that f realizes ψ whenever ψ is true. We can apply this here to get a realizer for $\forall m^N \, n \cdot m\downarrow \wedge \, n \cdot m = fm$. □

Definition 13.7. We say an extended pca satisfies the *computability axiom* if there exists $\mathbf{c} \in \mathcal{K}_1$ with the following property. For all $a, b, c \in \mathcal{K}_1$, $\mathbf{c}abc\downarrow$ and $\mathbf{c}abc = 0$ or $\mathbf{c}abc = 1$, and for all a, b, $ab\downarrow$ if and only if there exists c such that $\mathbf{c}abc = 1$.

Lemma 13.8. $\mathcal{K}_1$ *satisfies the computability axiom.*

Proof. By theorem 13.1. □

Theorem 13.9. *There is no $e \in \mathcal{K}_1$ with the following property: For all a such that a is total, $ea \downarrow$ and $ea = 0$ if $a\underline{n} = 0$ for all n, and $ea = 1$ if $a\underline{n} \neq 0$ for some n.*

Proof. If there was such a term e, then by theorem 13.1 we could use it to construct a term e' such that $e'ab = \underline{1}$ if $ab \downarrow$ and $e'ab = \underline{0}$ if $ab \uparrow$. However, this is not possible for any strictly partial pca (exercise). □

Corollary 13.10. **WLPO** *does not hold in either of the standard realizability models of* $\mathbf{HA}_\omega$ *on* $\mathcal{K}_1$.

Proof. Suppose there was a realizer e of

$$e \Vdash \forall f^{N \to N} (\forall n\, fn = 0) \vee \neg(\forall n\, fn = 0)$$

We will use e to contradict theorem 13.9. Suppose a is total. Then this gives us an element of $\mathcal{M}_{N \to N}$, directly for the intensional model, and as the function that a represents for the extensional model. In either case we have $ea \Vdash (\forall n\, an = 0) \vee \neg(\forall n\, an = 0)$. Hence either $\mathbf{p}_0(ea) = \top$ and $\mathbf{p}_1(ea) \Vdash \forall n\, an = 0$, or $\mathbf{p}_0(ea) = \bot$ and $\mathbf{p}_1(ea) \Vdash \neg\forall n\, an = 0$. In the former case, we have that for all n, $\mathbf{p}_1(ea)n \Vdash an = 0$, and so $an = 0$, and in the latter case it is false that $an = 0$ for all n, because otherwise $\lambda n.\top$ would be a realizer of $\forall n\, an = 0$. Hence $e' := \lambda a.\mathbf{p}_0(ea)01$ has the required property to contradict theorem 13.9. □

Theorem 13.11. *There is no $e \in \mathcal{K}_1$ with the following property: For all a such that a is total and $an \neq 0$ at most once, $ea \downarrow$ with $ea \in \{0, 1\}$ and for all n, $a(2n + (ea)) = 0$.*

Proof. Assume there is such an e. Using the $\mathbf{y}$ combinator, we can define a as follows. For each n, we will define $a(2n)$ and $a(2n+1)$. We first check if n is least such that $\mathbf{c}ean = 1$. If not, we take both $a(2n)$ and $a(2n+1)$ to be 0. If it is, then we know that $ea \downarrow$. We then check the value of ea. If it is 0, we take $a(2n) = 1$ and $a(2n+1) = 0$. If it is 1, we take $a(2n) = 0$ and $a(2n+1) = 1$. If it is anything else, then we take $a(2n) = a(2n+1) = 0$. Note that we have ensured that whatever happens a is a total binary sequence with at most one 1. Hence $ea \downarrow$, and so $\mathbf{c}ean = 1$ for some n. By considering the least such n, we get a contradiction. □

Corollary 13.12. **LLPO** *does not hold in either of the standard realizability models of* $\mathbf{HA}_\omega$ *on* $\mathcal{K}_1$.

13.3 The Kleene Tree

We now work in **HAS**.

Definition 13.13. A set T of numbers is a *tree* if it is inhabited, every element of T is the encoding of a finite sequence $[a_0, \ldots, a_{n-1}]$ where $a_i \in \{0, 1\}$, and whenever $[a_0, \ldots, a_{n-1}, a_n] \in T$ we have also $[a_1, \ldots, a_n] \in T$ (we say T is *prefix closed*).

An *infinite path* through the tree is a binary sequence $f : \mathbb{N} \to 2$ such that for all n we have $[f(0), f(1), \ldots, f(n)] \in T$.

We say T is *infinite* if for all n, T contains a binary sequence of length greater than n.

We say T is *decidable* if we have $n \in T \vee n \notin T$ for all n.

Theorem 13.14. *In the standard realizability model of* **HAS** *on* $\mathcal{K}_1$ *there is an infinite tree with no infinite branch (the* Kleene tree*).*

Proof. We again use the fact that $\mathcal{K}_1$ satisfies the computability axiom. We first define a set $T \subseteq \mathbb{N}$ externally to the model. We take T to be the set of codes for sequences $[a_0, \ldots, a_{n-1}]$ such that for all $e < n$ with $\mathbf{c}eem = 1$ for some $m < n$, we have $a_i \neq ei$. Note that T cannot contain any computable infinite branches. Suppose $e \in \mathcal{K}_1$ is a total binary sequence. Then $ee \downarrow$, and so $\mathbf{c}ee = 1$ for some m. However, we can see that for all $n > \max(e, m)$ and all finite sequences $[a_0, \ldots, a_{n-1}] \in T$, we have $a_e \neq ee$, and so e cannot be a branch of the tree. However, T is infinite, and moreover for each n we can find $[a_0, \ldots, a_{n-1}] \in T$ computably: we evaluate $\mathbf{c}eem$ for each $m, e < n$, and whenever $\mathbf{c}eem = 1$, we can evaluate ee, and then ensure $a_e \neq ee$. If $\mathbf{c}eem = 0$ for all $m < n$, we can take a_e to be either 0 or 1.

We can then view T as an element $\bar{T}$ of $\mathcal{M}_S$ by defining

$$\bar{T}(n) = \begin{cases} \{\top\} & n \in T \\ \emptyset & n \notin T \end{cases}$$

By the way that we defined T, we can computably decide whether or not a binary sequence belongs to T, and so we have a realizer that $\bar{T}$ is decidable. Furthermore, using the fact that T has no computable infinite branch, we can find a realizer witnessing that $\bar{T}$ has no infinite branch, and similarly, we can find a realizer witnessing that $\bar{T}$ is infinite using the fact that T is "computably infinite." □

Corollary 13.15. *In the standard realizability model of* **HAS** *on* $\mathcal{K}_1$ *there is an infinite open cover of Cantor space with no finite subcover.*

Proof. We take the open cover to consist of U_σ for each finite binary sequence σ such that $\sigma \notin \bar{T}$. Now for every infinite binary sequence $f \in 2^{\mathbb{N}}$, there is n such that $[f(0), \ldots, f(n)] \notin \bar{T}$. Hence we can take $\sigma := [f(0), \ldots, f(n)]$ to get $f \in U_\sigma$ and $\sigma \notin \bar{T}$. However, given any finite set $\sigma_1, \ldots, \sigma_k$ of finite binary sequences in the complement of $\bar{T}$, we can take n to be the maximum length, and then find $\tau \in \bar{T}$ of length n. However, we can now define f to be the infinite binary sequence such that $f(j) = \tau(j)$ for $j \leq n$, and $f(j) = 0$ for $j > n$. We then have $f \notin U_{\sigma_i}$ for $1 \leq i \leq k$. □

14 The Second Kleene Algebra and Function Realizability

14.1 The second Kleene algebra

Definition 14.1. A partial function $F : \mathbb{N}^{\mathbb{N}} \rightharpoonup \mathbb{N}$ is *continuous* if for all $f \in \mathbb{N}^{\mathbb{N}}$ such that $F(f)\downarrow$ there is $n \in \mathbb{N}$ such that for all $g \in \mathbb{N}^{\mathbb{N}}$, if $g(i) = f(i)$ for $i < n$, then $F(g)\downarrow$ and $F(g) = F(f)$.

A partial function $F : \mathbb{N}^{\mathbb{N}} \rightharpoonup \mathbb{N}^{\mathbb{N}}$ is *continuous* if for all n the partial function sending f to $F(f)(n)$ is continuous.

Note that every continuous function $\mathbb{N}^{\mathbb{N}} \to \mathbb{N}^{\mathbb{N}}$ is in particular a continuous function.

The key idea behind the second Kleene algebra is that we can encode partial continuous functions $\mathbb{N}^{\mathbb{N}} \rightharpoonup \mathbb{N}^{\mathbb{N}}$ as elements of $\mathbb{N}^{\mathbb{N}}$. We first show how to encode partial continuous functions $\mathbb{N}^{\mathbb{N}} \rightharpoonup \mathbb{N}$. Write $\mathbb{N}^{<\omega}$ for the set of finite sequences of natural numbers. Note that we can view any function $f : \mathbb{N}^{<\omega} \to \mathbb{N} + \{\bot\}$ as a continuous partial function function $F : \mathbb{N}^{\mathbb{N}} \rightharpoonup \mathbb{N}$ defined by

$$F(g) := \begin{cases} f(g(0), \ldots, g(n-1)) & f(g(0), \ldots, g(n-1)) \in \mathbb{N} \text{ and } n \text{ is least such} \\ \text{undefined} & \text{otherwise} \end{cases}$$

This in fact defines a surjective function from $(\mathbb{N} + \{\bot\})^{\mathbb{N}^{<\omega}}$ to continuous partial functions $\mathbb{N}^{\mathbb{N}} \rightharpoonup \mathbb{N}$. Given continuous $F : \mathbb{N}^{\mathbb{N}} \rightharpoonup \mathbb{N}$, we can define $f : \mathbb{N}^{<\omega} \to \mathbb{N} + \{\bot\}$ on $(a_0, \ldots, a_{n-1})$ as follows. If $F(g) = F(h)$ whenever $g(i) = h(i) = a_i$ for $i < n$, then we take $f(a_0, \ldots, a_{n-1}) := F(g)$, and otherwise we take $f(a_0, \ldots, a_{n-1})$ to be $\bot$. We say f is an *associate* of the function F.

However, we have a canonical bijection between $(\mathbb{N} + \{\bot\})^{\mathbb{N}^{<\omega}}$ and $\mathbb{N}^{\mathbb{N}}$ by composing with bijections $\mathbb{N}^{<\omega} \cong \mathbb{N}$ and $\mathbb{N} + \{\bot\} \cong \mathbb{N}$. This is one way of understanding the explicit definition below.

Definition 14.2. We define a function $|$ from $\mathbb{N}^{\mathbb{N}} \times \mathbb{N}^{\mathbb{N}}$ to partial functions $\mathbb{N} \rightharpoonup \mathbb{N}$. We define $f|g(n)$ to be $f(\langle n, [g(0), \ldots, g(m-1)]\rangle) - 1$ if m is the least such number with $f(\langle n, [g(0), \ldots, g(m-1)]\rangle) > 0$. If there is no such m, then $f|g(n)$ is undefined. We then convert this into a partial binary operator giving a partial applicative structure on $\mathbb{N}^{\mathbb{N}}$ by

$$f \cdot g(n) := \begin{cases} f|g & f|g \text{ is total} \\ \text{undefined} & \text{otherwise} \end{cases}$$

The partial applicative structure has elements $\mathbf{s}$ and $\mathbf{k}$ making it a partial combinatory algebra, that we call the *second Kleene algebra*, $\mathcal{K}_2$.

We have a canonical way to make $\mathcal{K}_2$ into an extended pca.

We define 0 to be the function constantly equal to 0. Note that the function sending $f : \mathbb{N} \to \mathbb{N}$ to the function $\lambda n. f(n) + 1$ is evidently continuous, and so has an associate, S that we use for the successor combinator. Note that for each n, the numeral $\underline{n}$ is precisely the constant function $\lambda x.n$.

14.2 Function realizability

We refer to realizability over $\mathcal{K}_2$ as *function realizability.* We will show two key properties of function realizability: that every function $\mathbb{N}^\mathbb{N} \to \mathbb{N}$ is continuous, and that we have the axiom of choice $\mathbf{AC}^{N \to N, N}$. These two axioms are sometimes combined together into a single axiom called *continuous choice*, which states that whenever $\forall f^{N \to N} \, \exists x^N \, \varphi(f, x)$ there exists a continous function $F : (N \to N) \to N$ such that for all $f \in \mathbb{N}^\mathbb{N}$ we have $\varphi(f, F(f))$. However, we will consider them separately. We first look at the axiom of choice.

Note that we have a continuous way to take a function $f : \mathbb{N} \to \mathbb{N}$ and evaluate it: i.e. return the numeral $\underline{f(n)}$ given f and $\underline{n}$ as input. We can also go the other way, and given an associate f for a continuous function F such that $F(\underline{n})$ is a numeral for all n, we can find, continuously in f, a function $g : \mathbb{N} \to \mathbb{N}$ such that $F(\underline{N}) = \underline{g(n)}$.

Using this, we can show the following for realizability models on $\mathcal{K}_2$.

Theorem 14.3. *$\mathbf{AC}^{N \to N, N \to N}$ holds in the standard extensional realizability model of $\mathbf{HA}_\omega$ on $\mathcal{K}_2$.*

Definition 14.4. Let $F : \mathbb{N}^\mathbb{N} \rightharpoonup \mathbb{N}^\mathbb{N}$. A *modulus of convergence function* is a function $M : \mathbb{N}^\mathbb{N} \rightharpoonup \mathbb{N}$ such that if $F(g)\downarrow$, then also $M(g)\downarrow$ and for all $h : \mathbb{N} \to \mathbb{N}$ if $h(i) = g(i)$ for $i < M(g)$ then $F(h)\downarrow$ and $F(h) = F(g)$.

Note that assuming $\mathbf{AC}^{N \to N, N}$, F is continuous if and only if it admits a modulus of convergence function. In fact we have the following theorem.

Theorem 14.5. *We can find $m \in \mathcal{K}_2$ with the following property. For all f, if $F : \mathbb{N}^\mathbb{N} \rightharpoonup \mathbb{N}^\mathbb{N}$ is the partial continuous function that f represents, then mf represents the modulus of convergence function for F.*

Using this result we can show the following for realizability models.

Theorem 14.6. *In both standard realizabliity models of $\mathbf{HA}_\omega$ on $\mathcal{K}_2$, every function $\mathbb{N}^\mathbb{N} \to \mathbb{N}$ is continuous.*

Moreover, in the intensional model, there is a function $m : \mathbb{N}^\mathbb{N} \to \mathbb{N}^\mathbb{N}$ such that for all f, $m(f)$ is a modulus of convergence function for f.

15 Realizability with Truth

One the key features of realizability is that it gives us models that satisfy axioms that do not necessarily hold externally in the metatheory where we are working (such as Church's thesis). However, it is sometimes preferable to consider a variation with the same idea that realizers capture the computational information implicit in intuitionisitic proofs, but where only true sentences can be realized. We will refer to this variation as *realizability with truth.*

This allows us to show properties such as *Church's rule*: Given a constructive proof (e.g. in $\mathbf{HA}_\omega$) that $\forall x^N \, \exists y^N \, \varphi(x, y)$, there is an algorithm e such that $\forall x^N \, \exists y^N \, \varphi(x, e \cdot y)$, provably in $\mathbf{HA}_\omega$. Moreover, in contrast to the anticlassical axioms of Kleene realizability, this algorithm remains valid when we move to stronger theories, even with classical logic, and with careful phrasing, we can even deduce the statement is "true" in the metatheory where we are working.

15.1 An external description

We will first give an "external" description of realizability with truth along similar lines to the other models with have considered in the course so far.

We fix a signature $(\mathfrak{S}, \mathfrak{O}, \mathfrak{R})$ of sorts, operator symbols and relation symbols, and a pca$^+$, $\mathcal{A}$.

Definition 15.1. A *realizability model with truth* consists of the following:

1. For each sort $S \in \mathfrak{S}$ a set $\mathcal{M}_S$
2. For each sort S a function $E_S : \mathcal{M}_S \to \mathcal{P}(\mathcal{A})$
3. For each relation symbol $R \in \mathfrak{R}$ of sort $S_1, \ldots, S_n$, a set $[\![R]\!]^\top \subseteq \mathcal{M}_{S_1} \times \ldots \times \mathcal{M}_{S_n}$ and a function $[\![R]\!] : \mathcal{M}_{S_1} \times \ldots \times \mathcal{M}_{S_n} \to \mathcal{P}(\mathcal{A})$. We require that if $[\![R]\!](a_1, \ldots, a_n)$ is a non empty subset of $\mathcal{A}$, then $(a_1, \ldots, a_n) \in [\![R]\!]^\top$.
4. For each operator symbol $O \in \mathfrak{O}$ of sort $S_1, \ldots, S_n \to T$ a function $[\![O]\!] : \mathcal{M}_{S_1} \times \ldots \times \mathcal{M}_{S_n} \to \mathcal{M}_T$ such that there exists $e \in \mathcal{A}$ with the following property. For all $x_1, \ldots, x_n$ with $x_i \in \mathcal{M}_{S_i}$ and all $a_1, \ldots, a_n$ with $a_i \in E_{S_i}(x_i)$ we have $ea_1 \ldots a_n \downarrow$ and $ea_1 \ldots a_n \in E_T([\![O]\!](x_1, \ldots, x_n))$.

We think of elements of $[\![R]\!]^\top$ as being those tuples $(a_1, \ldots, a_n)$ for which R is *true*. Hence the condition states that if R is realized, then it is true. However, we do not require the converse statement, so there can be true statements that are not realized. We think of the realizers as providing evidence that a formula is true, so if we have evidence that something is true, then it is true, but there are also true statements that we don't have enough evidence to know for certain.

For each variable assigment α, we define the interpretation of each term $[\![t]\!]_\alpha$ the same as for realizability models.

For formulas, we first define what it means for a formula to be *true* in the model.

$R(t_1, \ldots, t_n)$	true wrt α iff	$([\![t_1]\!]_\alpha, \ldots, [\![t_n]\!]_\alpha) \in [\![R]\!]$
$\bot$	true wrt α	never
$\varphi \wedge \psi$	true wrt α iff	φ is true and ψ is true
$\varphi \vee \psi$	true wrt α iff	φ is true or ψ is true
$\varphi \to \psi$	true wrt α iff	φ is true implies ψ is true
$\exists x^S\, \varphi$	true wrt α iff	there exists $a \in \mathcal{M}_S$ such that φ is true wrt $\alpha[x \mapsto a]$
$\forall x^S\, \varphi$	true wrt α iff	for all $a \in \mathcal{M}_S, \varphi$ is true wrt $\alpha[x \mapsto a]$

We extend realizability with truth from atomic formulas to all formulas by induction, as follows. Note that this is almost the same as the realizability interpretation we saw before, except that we adjust the definition for implication and universal quantifiers.

$e \nVdash_\alpha \bot$		always
$e \Vdash_\alpha \varphi \wedge \psi$	iff	$\mathbf{p}_0 e \Vdash_\alpha \varphi$ and $\mathbf{p}_1 e \Vdash_\alpha \psi$
$e \Vdash_\alpha \varphi \vee \psi$	iff	either $\mathbf{p}_0 e = \top$ and $\mathbf{p}_1 e \Vdash_\alpha \varphi$, or $\mathbf{p}_0 e = \bot$ and $\mathbf{p}_1 e \Vdash_\alpha \psi$
$e \Vdash_\alpha \varphi \to \psi$	iff	if $f \Vdash_\alpha \varphi$, then $ef\downarrow$ and $ef \Vdash_\alpha \psi$, and $\varphi \to \psi$ is true
$e \Vdash_\alpha \exists x^S \varphi$	iff	there exists $a \in \mathcal{M}_S$ such that $\mathbf{p_0} e \in E_S(a)$ and $\mathbf{p}_1 e \Vdash_{\alpha[x\mapsto a]} \varphi$
$e \Vdash_\alpha \forall x^S \varphi$	iff	for all $a \in \mathcal{M}_S$ and for all $f \in E_S(a), ef\downarrow$ and $ef \Vdash_{\alpha[x\mapsto a]} \varphi$ and $\forall x^S \varphi$ is true

As usual, we have a soundness theorem for realizability with truth:

Theorem 15.2. *If $\Gamma \vdash \varphi$ is provable in intuitionistic logic, and $x_1^{S_1}, \ldots, x_n^{S_n}$ is a list of variables including all of those that occur free in Γ and φ, then we can find $e \in \mathcal{A}$ such that for all variable assignments α, all $f_1, \ldots, f_n$ with $f_i \in E_{S_i}(\alpha(x_i))$, and all g such that $g \Vdash_\alpha \Gamma$, we have $ef_1 \ldots f_n g\downarrow$ and*

$$ef_1 \ldots f_n g \Vdash_\alpha \varphi$$

The key new result for realizability with truth is that if a formula is realized, then it is true, in the following sense.

Theorem 15.3. *Let α be a variable assignment. Suppose there is $e \in \mathcal{A}$ such that $e \Vdash_\alpha \varphi$. Then φ is true wrt α.*

Proof. We prove this by induction on formulas. Note that the definition of realizability with truth was adjusted precisely to give us the cases of implication and universal quantifiers in the inductive argument.

We still need to check the other cases, but these are straightforward.

We show the case of disjunction as an example. Suppose that $e \Vdash_\alpha \varphi \vee \psi$. Then either $\mathbf{p}_0 e = \top$ and $\mathbf{p}_1 e \Vdash_\alpha \varphi$ or $\mathbf{p}_0 e = \bot$ and $\mathbf{p}_1 e \Vdash_\alpha \psi$. In the former case we have φ is true by the inductive hypothesis, and so $\varphi \vee \psi$ is true, and similarly in the latter case, ψ and so also $\varphi \vee \psi$ is true. □

We can define realizability with truth models of $\mathbf{HA}_\omega$ by the following definition:

We take $\mathcal{M}_N$ to be $\mathbb{N}$, and for sorts σ and τ, we take $\mathcal{M}_{\sigma\times\tau}$ to be $\mathcal{M}_\sigma \times \mathcal{M}_\tau$ just as in the extensional model of $\mathbf{HA}_\omega$. However, we adjust the definition of $\mathcal{M}_{\sigma\to\tau}$ by taking it to be the set of *all* functions from $\mathcal{M}_\sigma$ to $\mathcal{M}_\tau$. As usual we take $E_{\sigma\to\tau}(f)$ to be the set of all $e \in \mathcal{A}$ that track f. We again define equality to be absolute.

We can use the realizability with truth model to extract algorithms from proofs in $\mathbf{HA}_\omega$ that give us information about true formuals of arithmetic. For example, we can show the following:

Theorem 15.4. *Suppose that $\varphi(x, y)$ is a formula of $\mathbf{HA}_\omega$ whose only free variables are x^N and y^N, and that $\mathbf{HA}_\omega \vdash \forall x\, \exists y\, \varphi(x, y)$. Then we can find a computable function, say $e \in \mathcal{T}_0$ such that for all $n \in \mathbb{N}$, we have $e\underline{n}\downarrow$ and $\varphi(\underline{n}, \underline{e \cdot n})$ is true.*

Proof. Suppose that $\mathbf{HA}_\omega \vdash \forall x\, \exists y\, \varphi(x,y)$. Then using the soundness theorem for the realizability with truth model of $\mathbf{HA}_\omega$ for the pca $\mathcal{T}_0^+$, we get $a \in \mathcal{T}_0^+$ such that $a \Vdash \forall x\, \exists y\, \varphi(x,y)$. Hence we can take $e := \lambda x.\mathbf{p}_0(ax)$. For each n, we have $\mathbf{p}_1(a\underline{n}) \Vdash \varphi(\underline{n}, \underline{en})$, and so $\varphi(\underline{n}, \underline{en})$ is true. □

15.2 The internal version

In theorem 15.4 we saw that given a proof in $\mathbf{HA}_\omega$, we could extract an algorithm witnessing $\forall x\, \exists y\, \varphi(x,y)$. However, we were only able to show that for each n, $\varphi(\underline{n}, \underline{e \cdot n})$ is true, not that the statement is provable in $\mathbf{HA}_\omega$. It is also possible to give this stronger result using a similar technique. However, to do this, it is necessary to consider a variation where we carry out parts of the definition of realizability with truth inside $\mathbf{HA}_\omega$, and use this to prove parts of the soundness theorem, again inside $\mathbf{HA}_\omega$.

We won't cover this in complete detail, but to give a rough idea, one can follow this outline:

1. Instead of working with the general theory of pcas, we only look at specific pcas that we can formalise in $\mathbf{HA}_\omega$, with the application operator in the pca appearing as a formula with 3 free variables, which is provably functional. It is possible to do this for $\mathcal{T}_0^+$ by working with terms via their Gödelnumbering. However, it turns out to be more useful to consider $\mathcal{K}_1$.

2. Instead of defining realizability models in general, we work directly with ones that we can define in $\mathbf{HA}_\omega$. We only consider the realizability with truth model of $\mathbf{HA}_\omega$ above. We define $\mathcal{M}_\sigma$ and E_σ by external induction on the sorts σ. We take $\mathcal{M}_\sigma$ to be the sort σ itself. We define E_σ to be a formula with at most two free variables, x^σ and e^N. E.g. we define $E_{\sigma\to\tau}(x,e)$ to be the internal statement that e tracks (as an element of $\mathcal{K}_1$) the function x.

3. We define the realizability interpretation by external induction on formulas. Namely, for each (external) formula φ, we define a formula of $\mathbf{HA}_\omega$ with an additional free variable, e, which we denote $e \Vdash \varphi$.

4. We also prove by induction on formulas, that it is provable in $\mathbf{HA}_\omega$ that $\forall e^N\, e \Vdash \varphi \ \to\ \varphi$.

5. We prove the soundness theorem again by an external inductive argument, this time on proofs. Namely, given a proof $\Gamma \vdash \varphi$, we can find $e \in \mathbb{N}$ such that $\mathbf{HA}_\omega \vdash\ e \Vdash \varphi$.

By following this outline, we can sketch a proof of a stronger version of theorem 15.4.

Theorem 15.5. *Suppose that $\varphi(x,y)$ is a formula of $\mathbf{HA}_\omega$ whose only free variables are x^N and y^N, and that $\mathbf{HA}_\omega \vdash \forall x\, \exists y\, \varphi(x,y)$. Then we can find $n \in \mathbb{N}$ such that $\mathbf{HA}_\omega \vdash \forall x^N\, \varphi(x, \underline{n} \cdot x)$. We say $\mathbf{HA}_\omega$ satisfies* Church's rule.

Proof. Suppose that $\mathbf{HA}_\omega \vdash \forall x^N \, \varphi(x, \underline{n} \cdot x)$. Then we can find $e \in \mathbb{N}$ such that $\mathbf{HA}_\omega \vdash \underline{e} \Vdash \forall x^N \, \varphi(x, y)$. Hence $\mathbf{HA}_\omega \vdash \forall x^N \, (\mathbf{p}_1 \underline{e} x) \Vdash \varphi(x, \mathbf{p}_0 \underline{e})$, and so we have $\mathbf{HA}_\omega \vdash \forall x^N \, \varphi(x, \mathbf{p}_0 \underline{e})$. It turns out that one can show $\mathbf{HA}_\omega \vdash \underline{\lambda x.\mathbf{p}_0(ex)} = \lambda x.\mathbf{p}_0(\underline{e}x)$, and so taking $e' := \lambda x.\mathbf{p}_0(ex)$, we have $\mathbf{HA}_\omega \vdash \forall x^N \, \varphi(x, \underline{e'}x)$. □