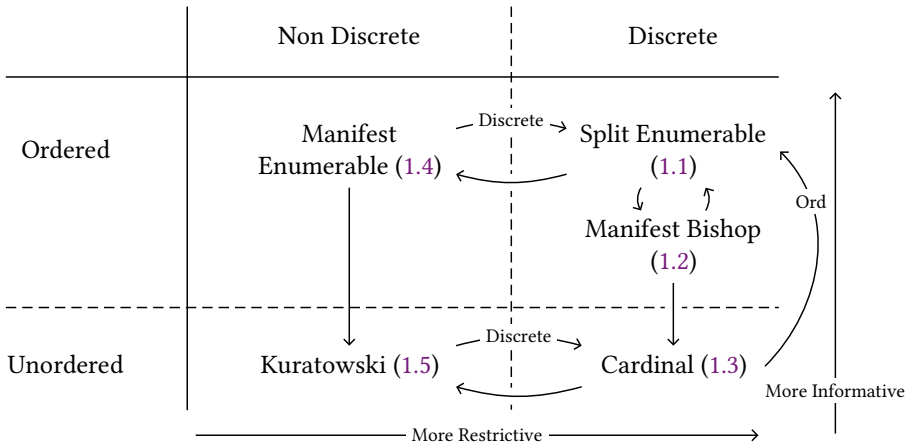

Fig. 1. Classification of finiteness predicates according to whether they are discrete (imply decidable equality) and whether they imply a total order.

1 FINITENESS PREDICATES

In this subsection, we will define and briefly describe each of the five predicates in Figure 1. We will also explain *why* there are five separate predicates: how can it be the case that so many different things describe "finiteness"? As we will see, some predicates are too informative (they tell us more about the underlying type other than it just being finite), or too restrictive (they don't allow certain finite types to be classified as finite). These diversions won't be dead-ends, however: the final predicate we will land on as the "correct" (or, more accurately, most useful) notion of finiteness will be built out of all of the others.

1.1 Split Enumerability

We will start with a simple notion of finiteness, called split enumerability. This predicate is perhaps the first definition of "finite" that someone might come up with (it's certainly the most common in dependently-typed programming): put simply, a split enumerable type is a type for which all of its elements can be listed.

Definition 1 (Split Enumerable Set). To say that some type A is split enumerable is to say that there is a list $\mathit{support} : \mathsf{List}\ A$ such that any value $x : A$ is in $\mathit{support}$.

$$\mathscr{E}!\ A = \Sigma[\ \mathit{support} : \mathsf{List}\ A\ ]\ ((x : A) \to x \in \mathit{support})$$

We call the first component of this pair the "support" list, and the second component the "cover" proof. An equivalent version of this predicate was called `Listable` in Firsov and Uustalu [2015].

This predicate is simple and useful, but we will see later on how it is perhaps a little imprecise. Before we dive in to exploring the predicate itself, though, we will need to explain some of the terms we used in its definition.

What is a List? In this paper we prefer a slightly unusual definition for the type of lists:

$$\mathsf{List} = [\![\ \mathbb{N}\ ,\ \mathsf{Fin}\ ]\!] \tag{1}$$

This is the definition for a *container* (Definition 2): effectively, the above definition says that "Lists

are a datatype whose shape is given by the natural numbers, and which can be indexed by numbers smaller than its shape".

If that seems needlessly complex, don't worry: this definition is precisely equivalent to the usual inductive one.

```
Fin zero     = ⊥
Fin (suc n)  = ⊤ ⊎ Fin n
```

Listing 2: Finite Prefixes of ℕ

```
data List (A : Type a) : Type a where
  []   : List A
  _::_ : A → List A → List A
```

And this isn't some kind of hand-waving equivalence, either: since we are working in HoTT, we can (and do) prove that the two types are equal, allowing us to use one or the other depending on whichever is more convenient, and subst in the other representation without loss of generality. That said, defining lists as containers will reveal several interesting connections and proofs about split enumerability and the other predicates, so for the remainder of the paper whenever we say List we will mean Equation 1.

We still must define containers themselves, of course. Containers are a well-studied topic in dependent type theory, with a rich theory: we won't dive in to that here.

Definition 2 (Containers). A container [Abbott et al. 2005] is a pair S, P where S is a type, the elements of which are called the *shapes* of the container, and P is a type family on S, where the elements of $P(s)$ are called the *positions* of a container. We "interpret" a container into a functor defined like so:

$$⟦\ S\ ,\ P\ ⟧\ X = Σ[\ s : S\ ]\ (P\ s \to X) \tag{3}$$

The definition of container is a little abstract: it is instructive to think of it more concretely for the case of lists. The container representing finite lists is a pair of a natural number n representing the length (or "shape") of the list, and a function $\mathsf{Fin}\ n \to A$, representing the indexing function into the list.

One of the nice things about containers is it gives us a generic way to define "membership":

$$x ∈ xs = \mathsf{fiber}\ (\mathsf{snd}\ xs)\ x \tag{4}$$

Here we're using the homotopy-theory notion of a fiber to define membership: a fiber for some function $B \to A$ and some point y in its codomain is a value x and a proof that $f\ x ≡ y$. Membership also makes more sense when described concretely in terms of lists: $x ∈ xs$ means "there is an index into xs such that the index points at an item equal to x".

```
fiber : (A → B) → B → Type _
fiber f y = Σ[ x ∈ _ ] (f x ≡ y)
```

Listing 5: A Fiber [Univalent Foundations Program 2013, definition 4.2.4]

Split Surjections. Now that we have our terms defined, let's look a little at how split enumerability relates to more traditional, classical notions of finiteness. In a classical setting we likely wouldn't mention "lists" or the like, and would instead define finiteness based on the existence of some injection or surjection, say a surjection from a finite prefix of the natural numbers. In HoTT, surjections (or, more precisely, *split* surjections [Univalent Foundations Program 2013, definition 4.6.1]), are defined like so:

$$\mathsf{SplitSurjective}\ f = ∀\ y \to \mathsf{fiber}\ f\ y \qquad\qquad A \twoheadrightarrow! B = Σ\ (A \to B)\ \mathsf{SplitSurjective} \tag{6}$$

As it turns out, our definition of finiteness here is precisely the same as a surjection-based one, in quite a deep way!

Lemma 1.1. *A proof of split enumerability is equivalent to a split surjection from a finite prefix of the natural numbers.*

$$\mathscr{E}!\ A \Leftrightarrow Σ[\ n : ℕ\ ]\ (\mathit{Fin}\ n \twoheadrightarrow!\ A)$$

Proof.

$$
\begin{array}{lll}
\mathscr{E}!\ A & \cong\langle\rangle & \text{Def. 1 } (\mathcal{E}!) \\
\Sigma[\ xs : \mathsf{List}\ A\ ]\ ((x : A) \to x \in xs) & \cong\langle\rangle & \text{Eqn. 4 } (\in) \\
\Sigma[\ xs : \mathsf{List}\ A\ ]\ ((x : A) \to \mathsf{fiber}\ (\mathsf{snd}\ xs)\ x) & \cong\langle\rangle & \text{Eqn. 6} \\
\Sigma[\ xs : \mathsf{List}\ A\ ]\ \mathsf{SplitSurjective}\ (\mathsf{snd}\ xs) & \cong\langle\rangle & \text{Eqn. 1 (List)} \\
\Sigma[\ xs : [\![\ \mathbb{N}\ , \mathsf{Fin}\ ]\!]\ A\ ]\ \mathsf{SplitSurjective}\ (\mathsf{snd}\ xs) & \cong\langle\rangle & \text{Eqn. 3} \\
\Sigma[\ xs : \Sigma[\ n : \mathbb{N}\ ]\ (\mathsf{Fin}\ n \to A)\ ]\ \mathsf{SplitSurjective}\ (\mathsf{snd}\ xs) & \cong\langle\ \mathsf{reassoc}\ \rangle & \text{Reassociation} \\
\Sigma[\ n : \mathbb{N}\ ]\ \Sigma[\ f : (\mathsf{Fin}\ n \to A)\ ]\ \mathsf{SplitSurjective}\ f & \cong\langle\rangle & \text{Eqn. 6} \\
\Sigma[\ n : \mathbb{N}\ ]\ (\mathsf{Fin}\ n \twoheadrightarrow! A) \ \blacksquare &&
\end{array}
$$

In the above proof syntax the $\equiv\langle\rangle$ connects lines which are definitionally equal, i.e. they are "obviously" equal from the type checker's perspective. Clearly, only one line isn't a definitional equality:

$$\mathsf{reassoc} : \Sigma\ (\Sigma\ A\ B)\ C \Leftrightarrow \Sigma[\ x : A\ ]\ \Sigma[\ y : B\ x\ ]\ C\ (x\ ,\ y)$$

This means that we could have in fact written the whole proof as follows:

```
split-enum-is-split-surj : ℰ! A ⇔ Σ[ n : ℕ ] (Fin n ↠! A)
split-enum-is-split-surj = reassoc
```

The simplicity of this proof, by the way, is why we preferred the container-based definition of lists over the traditional one.

In Firsov and Uustalu [2015], there is a proof that split enumerability and surjections from Fin are propositionally equivalent; i.e. a function from each to the other is provided. The fact that the two proofs are precisely equivalent is not proven (and, indeed, it is impossible to prove in MLTT).

Instances. To actually show that a type A is finite amounts to constructing a term of type $\mathcal{E}!\ A$. For simple types like Bool, that is simple: it just amounts to basically listing the constructors. As a slightly more complex example, consider the Fin type we've been using. Remember that split enumerability is in fact the same as a split surjection from Fin (Lemma 1.1): to show that Fin is split enumerable, then, we need only show that it has a split surjection from itself. We'll prove the following slightly more general statement:

```
↠!-ident : A ↠! A
↠!-ident .fst = id
↠!-ident .snd y .fst = y
↠!-ident .snd y .snd _ = y
```

```
ℰ!⟨2⟩ : ℰ
ℰ!⟨2⟩ .fs
ℰ!⟨2⟩ .sn
ℰ!⟨2⟩ .sn
```

Listing 7: Proof of $\mathcal{E}!$ Bool

Decidable Equality. One thing that characterises all split enumerable types is that they are all *discrete*, i.e. they have decidable equality.

```
Discrete A = (x y : A) → Dec (x ≡ y)
```

```
data Dec (A : Type a) : Type a where
  yes :    A → Dec A
  no  : ¬ A → Dec A
```

We will see later that this has implications for the space of types we're dealing with, but for now it simply provides a useful function on split enumerable types.

Lemma 1.2. *Split enumerability implies decidable equality.*

Proof. To prove that split enumerability implies decidable equality we'll take a quick detour through injections.

```
Injective f = ∀ x y → f x ≡ f y → x ≡ y          A ↣ B = Σ[ f ∶ (A → B) ] Injective f
```

These are useful because we know that any type which injects into a discrete type is itself discrete:

```
Discrete-pull-inj : A ↣ B → Discrete B → Discrete A
Discrete-pull-inj (f , inj) _≟_ x y =
  case (f x ≟ f y) of
    λ { (no ¬p) → no (¬p ∘ cong f)
      ; (yes p) → yes (inj x y p) }
```

And we can turn a split surjection from A to B into an injection from B to A:

```
surj-to-inj : (A ↠! B) → (B ↣ A)
surj-to-inj (f , surj) .fst x = surj x .fst
surj-to-inj (f , surj) .snd x y f⁻¹⟨x⟩≡f⁻¹⟨y⟩ =
  x                ≡˘⟨ surj x .snd ⟩
  f (surj x .fst) ≡⟨ cong f f⁻¹⟨x⟩≡f⁻¹⟨y⟩ ⟩
  f (surj y .fst) ≡⟨ surj y .snd ⟩
  y ■
```

Yielding a simple proof that any type with a split surjection from a discrete type is itself discrete:

```
Discrete-distrib-surj : (A ↠! B) → Discrete A → Discrete B
Discrete-distrib-surj = Discrete-pull-inj ∘ surj-to-inj
```

Since split enumerability is really just a split surjection from Fin, and since we know that Fin is discrete, the overall proof resolves quite simply:

```
ℰ!⇒Discrete : ℰ! A → Discrete A
ℰ!⇒Discrete = flip Discrete-distrib-surj discreteFin
                ∘ snd
                ∘ ℰ!⇔Fin↠! .fun
```

■

This lemma is also proven in Firsov and Uustalu [2015], although using a significantly different technique.

1.2 Manifest Bishop Finiteness

We mentioned in the introduction that occasionally in constructive mathematics proofs will contain "too much" information. With split enumerability we can see an instance of this. Consider the following proof of the finiteness of bool:

```
ℰ!⟨2⟩ : ℰ! Bool
ℰ!⟨2⟩ .fst = [ false , true , false ]
ℰ!⟨2⟩ .snd false = 0 , refl
ℰ!⟨2⟩ .snd true  = 1 , refl
```

(8)

There is an extra false at the end of the support list. There's nothing terribly wrong with that: it is

still a valid proof of finiteness, after all, but it does mean that this proof has some extra information which we didn't necessarily intend to encode.

There is "slop" in the type of split enumerability: there are more distinct values than there are *usefully* distinct values. To reconcile this, we will disallow duplicates in the support list.

This is where manifest Bishop finiteness comes in: this is a definition of finiteness quite similar to split enumerability in other regards, except that it does not allows for duplicates in the support list.

How exactly to prohibit duplicates is the next question. One approach might be to change the definition of List, or introduce a new type NoDupeList, and use it in the predicate instead. However, this would mean we lose access to the functions we have defined on lists, and we have to change the definition of $\in$ as well.

There is a much simpler and more elegant solution: we insist that every *membership proof* must be unique. This would disallow a definition of $\mathcal{E}!$ Bool with duplicates, as there are multiple values which inhabit the type false $\in$ [false, true, false]. It also allows us to keep most of the split enumerability definition unchanged, just adding a condition to the returned membership proof in the cover proof.

To specify that a value must exist uniquely in HoTT we can use the concept of a *contraction* [Univalent Foundations Program 2013, definition 3.11.1].

$$\text{isContr } A = \Sigma[\ x : A\ ]\ \forall\ y \to x \equiv y \tag{9}$$

A contraction is a type with the least possible amount of information: it represents the tautologies. All contractions are isomorphic to $\top$.

By saying that a proof of membership is a contraction, we are saying that it must be *unique*.

$$x \in! xs = \text{isContr } (x \in xs) \tag{10}$$

Now a proof of $x \in! xs$ means that x is not just in xs, but it appears there *only once*.

With this we can define manifest Bishop finiteness:

Definition 3 (Manifest Bishop Finiteness). A type is manifest Bishop finite if there exists a list which contains each value in the type once.

$$\mathcal{B}\ A = \Sigma[\ \mathit{support} : \text{List } A\ ]\ ((x : A) \to x \in! \mathit{support})$$

The only difference between manifest Bishop finiteness and split enumerability is the membership term: here we require unique membership ($\in!$), rather than simple membership ($\in$). An equivalent version of this predicate was called `ListableNoDup` in Firsov and Uustalu [2015].

We use the word "manifest" here to distinguish from another common interpretation of Bishop finiteness, which we have called cardinal finiteness in this paper: this version of the proof is "manifest" because we have a concrete, non-truncated list of the elements in the proof.

The Relationship Between Manifest Bishop Finiteness and Split Enumerability. While manifest Bishop finiteness might seem stronger than split enumerability, it turns out this is not the case. Both predicates imply the other.

Going from manifest Bishop finiteness is relatively straightforward: to construct a proof of split enumerability from one of manifest Bishop finiteness, it suffices to convert a proof of $x \in! xs$ to one of $x \in xs$, for all x and xs. Since $\in!$ is defined as a contraction of $\in$, such a conversion is simply the fst function.

Going the other direction takes significantly more work.

Lemma 1.3. *Any split enumerable set is manifest Bishop finite.*

This lemma is proven in Firsov and Uustalu [2015]. We will only sketch the proof here: the "unique membership" condition in $\mathcal{B}$ means that we are not permitted duplicates in the support list. The first step in the proof, then, is to filter those duplicates out from the support list of the $\mathcal{E}!$ proof: we can do this using the decidable equality provided by $\mathcal{E}!$ (Lemma 1.2). From there, we need to show that the membership proof carries over appropriately.

We have now proved that every manifestly Bishop finite type is split enumerable, and vice versa. While the types are not *equivalent* (there are more split enumerable proofs than there are manifest Bishop finite proofs), they are of equal power.

From Manifest Bishop Finiteness to Equivalence. We have seen that split enumerability was in fact a split-surjection in disguise. We will now see that manifest Bishop finiteness is in fact an *equivalence* in disguise. We define equivalences as contractible maps [Univalent Foundations Program 2013, definition 4.4.1]:

$$\mathsf{isEquiv}\ f = \forall\ y \to \mathsf{isContr}\ (\mathsf{fiber}\ f\ y) \qquad A \simeq B = \Sigma[\ f : (A \to B)\ ]\ \mathsf{isEquiv}\ f \tag{11}$$

Lemma 1.4. *Manifest bishop finiteness is equivalent to an equivalence to a finite prefix of the natural numbers.*

$$\mathcal{B}\ A \Leftrightarrow \exists[\ n\ ]\ (\mathit{Fin}\ n \simeq A)$$

Proof.

$$\begin{array}{lll}
\mathcal{B}\ A & \cong\langle\rangle & \text{Def. 3 } (\mathcal{B}) \\
\Sigma[\ xs : \mathsf{List}\ A\ ]\ ((x : A) \to x \in! xs) & \cong\langle\rangle & \text{Eqn. 10 } (\in!) \\
\Sigma[\ xs : \mathsf{List}\ A\ ]\ ((x : A) \to \mathsf{isContr}\ (x \in xs)) & \cong\langle\rangle & \text{Eqn. 4 } (\in) \\
\Sigma[\ xs : \mathsf{List}\ A\ ]\ ((x : A) \to \mathsf{isContr}\ (\mathsf{fiber}\ (\mathsf{snd}\ xs)\ x)) & \cong\langle\rangle & \text{Eqn. 11} \\
\Sigma[\ xs : \mathsf{List}\ A\ ]\ \mathsf{isEquiv}\ (\mathsf{snd}\ xs) & \cong\langle\rangle & \text{Eqn. 1 (List)} \\
\Sigma[\ xs : [\![\ \mathbb{N}\ , \mathsf{Fin}\ ]\!]\ A\ ]\ \mathsf{isEquiv}\ (\mathsf{snd}\ xs) & \cong\langle\rangle & \text{Eqn. 3} \\
\Sigma[\ xs : \Sigma[\ n : \mathbb{N}\ ]\ (\mathsf{Fin}\ n \to A)\ ]\ \mathsf{isEquiv}\ (\mathsf{snd}\ xs) & \cong\langle\ \mathsf{reassoc}\ \rangle & \text{Reassociation} \\
\Sigma[\ n : \mathbb{N}\ ]\ \Sigma[\ f : (\mathsf{Fin}\ n \to A)\ ]\ \mathsf{isEquiv}\ f & \cong\langle\rangle & \text{Eqn. 11} \\
\exists[\ n\ ]\ (\mathsf{Fin}\ n \simeq A)\ \blacksquare &&
\end{array}$$

This proof is almost identical to the proof for Lemma 1.1: it reveals that enumeration-based finiteness predicates are simply another perspective on relation-based ones.

Firsov and Uustalu [2015] provides a proof of the related statement; that manifest Bishop finiteness implies an equivalence to a prefix of the natural number (and vice versa), but similarly to split enumerability they cannot show that the two are equivalent as we have done here.

As we are working in CuTT, a proof of equivalence between two types gives us the ability to *transport* proofs from one type to the other. This is extremely powerful, as we will see.

1.3 Cardinal Finiteness

While we have removed some of the unnecessary information from our finiteness predicates, one piece still remains. The two following proofs are both valid proofs of the finiteness of Bool, and both do not include any duplicates:

```
ℰ!⟨2⟩ : ℰ! Bool
ℰ!⟨2⟩ .fst = [ false , true ]
ℰ!⟨2⟩ .snd false = 0 , refl
ℰ!⟨2⟩ .snd true  = 1 , refl
```

```
ℰ!⟨2⟩′ : ℰ! Bool
ℰ!⟨2⟩′ .fst = [ true , false ]
ℰ!⟨2⟩′ .snd false = 1 , refl
ℰ!⟨2⟩′ .snd true  = 0 , refl
```

Clearly they're not the same though: the order of their support lists differs. Each finiteness predicate so far has contained an *ordering* of the underlying type. For our purposes, this is too much information: it means that when constructing the "category of finite sets" later on, instead of each type having one canonical representative, it will have $n!$, where n is the cardinality of the type[1].

What we want is a proof of finiteness that is a proposition.

(12)
$$\mathsf{isProp}\ A = (x\ y : A) \to x \equiv y$$

The mere propositions are one homotopy level higher than the contractions (Equation 9), the types for which all values are equal to some value. They represent the types for which all values are equal, or, the types isomorphic to $\bot$ or $\top$. You can also define propositions in terms of the contractions: propositions are the types whose paths are contractions. Soon (Equation 15) we will see the next homotopy level, which are defined in terms of the propositions.

Despite now knowing the precise property we want our finiteness predicate to have, we're not much closer to achieving it. To remedy the problem, we will use the following type:

(13)
```
data ||_|| (A : Type a) : Type a where
  |_|      : A → || A ||
  squash : (x y : || A ||) → x ≡ y
```

This is a *higher inductive type*. Normal inductive types have *point* constructors: constructors which construct values of the type. The first constructor here (|_|), or the constructor true for Bool, are both "point" constructors.

What makes this type higher inductive is that it also has *path* constructors: constructors which add new equalities to the type. The squash constructor here says that all elements of $\|A\|$ are equal, regardless of what A is. In this way it allows us to propositionally truncate types, turning information-containing proofs into mere propositions. Put another way, a proof of type $\|A\|$ is a proof that some A exists, without revealing *which* A.

To actually use values of this type we have the following eliminator:

(14)
$$\mathsf{rec} : \mathsf{isProp}\ B \to (A \to B) \to \|\ A\ \| \to B$$

This says that we can eliminate into any proposition: interestingly, this allows us to define a monad instance for ||_||, meaning we can use things like do-notation.

With this, we can define cardinal finiteness:

Definition 4 (Cardinal Finiteness). A type A is cardinally finite if there exists a propositionally truncated proof that A is manifest Bishop finite or equivalent to a finite prefix of the natural numbers.

$$\mathscr{C}\ A = \|\ \mathcal{B}\ A\ \|$$

This predicate is called Bishop finiteness in [Frumin et al. 2018].

[1]We actually do get a category (a groupoid, even) from manifest Bishop finiteness [Yorgey 2014]: it's the groupoid of finite sets equipped with a linear order, whose morphisms are order-preserving bijections. We do not explore this particular construction in any detail.

Deriving Uniquely-Determined Quantities. At first glance, it might seem that we lose any useful properties we could derive from $\mathcal{B}$. Luckily, this is not the case: we will show here how to derive decidable equality (Lemma 1.5) and cardinality (Lemma 1.6) out from under the truncation. Those two lemmas are proven in [Yorgey 2014] (Proposition 2.4.9 and 2.4.10, respectively), in much the same way as we have done here. Our contribution for this subsection is simply the formalisation.

First we'll show that decidable equality carries over from manifest Bishop finiteness. Before we do, note that the fact that we can do this says something interesting about propositional truncation: it has computational, or algorithmic, content. That is in contrast to other ways to "truncate" types: $\neg\neg P$, for instance, is a way to provide a "proof" of P without revealing anything about P in MLTT. No matter how much we prove that a function from P doesn't care about which P it got, though, we can never extract any kind of algorithm or computation from $\neg\neg P$.

Lemma 1.5. *Any cardinal-finite set has decidable equality.*

$$\mathscr{C}\ A \to \mathit{Discrete}\ A$$

Proof. We already know that manifest Bishop finiteness implies decidable equality; to apply that proof to cardinal finiteness we'll use the eliminator in Equation 14. Our task, in other words, is to prove the following:

$$\mathsf{isProp}\ (\mathsf{Discrete}\ A)$$

To show that this type is a proposition we must show that any two given members of the type are equal, i.e. we are given two proofs of decidable equality on A and we must show that they are equal. Remember that Discrete A is a function of two arguments returning a Dec of whether those two arguments are equal or not. By function extensionality, to prove that that is a proposition we have to prove that Dec $(\equiv xy)$ is a proposition. This proof requires that we show that the payload of each of the constructors (yes and no) are propositions. no's payload is $\equiv xy \to \bot$, which is a proposition because $\bot$ is a proposition.

ref for wherever this definition was (15)

yes is a little more interesting: its payload is ≡xy. How can we prove that the path between x and y is a proposition? It turns out that there is a class of types for which all paths are propositions: the *sets*.

$$\mathsf{isSet}\ A = (x\ y : A) \to \mathsf{isProp}\ (x \equiv y)$$

This is the next homotopy level up from the propositions (Equation 12). More importantly, there is an important theorem relating to sets which *also* relates to decidable equality: Hedberg's theorem [Hedberg 1998]. This tells us that any type with decidable equality is a set.

$$\mathsf{Discrete}\ A \to \mathsf{isSet}\ A$$

And of course we know that A here has decidable equality: we were just given two proofs of that fact at the beginning of this proof!

This suffices to prove that decidable equality is itself a proposition, and therefore that we can apply Equation 14 and the proof that bishop finiteness implies decidable equality to cardinal finiteness, proving our goal. ■

The next thing we can derive from underneath the truncation in cardinal finiteness is a natural number representing the actual cardinality of the finite type. Of course $\mathbb{N}$ isn't a proposition, so the eliminator in equation 14 won't work for us here. Instead we will use the following:

(16)
$$\mathsf{rec{\to}set} : \mathsf{isSet}\ B \to (f : A \to B) \to (\forall\ x\ y \to f\,x \equiv f\,y) \to \| A \| \to B$$

This says that we can eliminate into a set as long as the function we use doesn't care about which value it's given: formally, f in this example has to be "coherently constant" [Kraus 2015].

With that, we can move on to the proof:

Lemma 1.6. *Given a cardinally finite type, we can derive the type's cardinality, as well as a propositionally truncated proof of equivalence with Fins of the same cardinality.*

$$\mathit{cardinality\text{-}is\text{-}unique} : \mathscr{C}\ A \to \exists[\ n\ ] \parallel \mathit{Fin}\ n \simeq A \parallel$$

Proof. The high-level overview of our proof is as follows:

$$\mathsf{cardinality\text{-}is\text{-}unique} = \mathsf{rec{\to}set}\ \mathsf{card\text{-}isSet}\ \mathsf{alg}\ \mathsf{const\text{-}alg} \circ \|\mathsf{map}\|\ \mathcal{B}{\Rightarrow}\mathsf{Fin}{\simeq}$$

It is the composition of two operations: first, with $\|\mathsf{map}\|$, we change the truncated proof of manifest bishop finiteness to a proof of equivalence with fin.

Then we use the eliminator from Equation 16 with three parameters. The first simply proves that that the output is a set:

$$\mathsf{card\text{-}isSet} : \mathsf{isSet}\ (\exists[\ n\ ] \parallel \mathsf{Fin}\ n \simeq A \parallel)$$

The second is the function we apply to the truncated value:

$$\begin{aligned}&\mathsf{alg} : \Sigma[\ n : \mathbb{N}\ ]\ (\mathsf{Fin}\ n \simeq A) \to \Sigma[\ n : \mathbb{N}\ ] \parallel \mathsf{Fin}\ n \simeq A \parallel\\ &\mathsf{alg}\ (n\ ,\ f{\simeq}A) = n\ ,\ |\ f{\simeq}A\ |\end{aligned}$$

And the third is a proof that that function is itself coherently constant:

$$\mathsf{const\text{-}alg} : (x\ y : \exists[\ n\ ]\ (\mathsf{Fin}\ n \simeq A)) \to \mathsf{alg}\ x \equiv \mathsf{alg}\ y$$

The tricky part of the proof is const-alg: here we need to show that alg returns the same value no matter its input. That output is a pair, the first component of which is the cardinality, and the second the truncated equivalence proof. The truncated proofs in the output are trivially equal by the truncation, so our obligation now has been reduced to:

$$\frac{(n :\ \mathbb{N})\quad (p : \mathsf{Fin}\ n \simeq A)\quad (m :\ \mathbb{N})\quad (q : \mathsf{Fin}\ m \simeq A)}{\equiv nm}$$

Given univalence we have $\mathsf{Fin}\ n \equiv \mathsf{Fin}\ m$, and the rest of our task is to prove:

$$\frac{\equiv \mathsf{Fin}\ n \mathsf{Fin}\ m}{\equiv nm}$$

This is a well-known puzzle in dependently-typed programming, and one that has a surprisingly tricky and complex proof. We do not include it here, since it has already been explored elsewhere, but it is present in our formalisation. ■

Going from Cardinal Finiteness to Manifest Bishop Finiteness. We know of course that we can convert any proof of manifest Bishop finiteness to a proof of Cardinal finiteness: it's just the truncation function $|\_|$. It's the other direction which presents a difficulty:

Theorem 1.7. *Any cardinal finite type with a total order is Bishop finite.*

Proof. The proof for this particular theorem is quite involved in the formalisation, so we only give its sketch here.

Our strategy will be to *sort* the support list of the proof for Bishop finiteness, and then prove that the sorting function is coherently constant, thereby satisfying the eliminator in Equation 16. We need to show, in other words, that sorting two support lists from proofs of manifest Bishop

finiteness on the same type with the same order always returns the same result. For simplicity's sake we will use insertion sort:

```
insert : E → List E → List E
insert x [] = x :: []
insert x (y :: xs) with x ⩽? y
... | inl x⩽y = x :: y :: xs
... | inr y⩽x = y :: insert x xs
```

```
sort : List E → List E
sort []       = []
sort (x :: xs) = insert x (sort xs)
```

And we prove that sort produces a list which is sorted, and a permutation of its input.

```
sort-sorts : ∀ xs → Sorted (sort xs)
```

```
sort-perm : ∀ xs → sort xs ↭ xs
```

We've introduced two new types here: Sorted is a predicate enforcing that the given list is sorted, and ↭ is a permutation relation between two lists. We take the definition of permutations from [Danielsson 2012]: two lists are permutations of each other if their membership proofs are all equivalent.

$$xs \leftrightsquigarrow ys = \forall\ x \to (x \in xs) \Leftrightarrow (x \in ys)$$

This definition fits particularly well for two reasons: first, it is defined on containers generically, which fits well with our finiteness predicates. Secondly, it is extremely straightforward to show that the support lists of any two proofs of manifest Bishop finiteness must be permutations of each other:

$$(xs\ ys : \mathcal{B}\ A) \to xs\ .\mathrm{fst} \leftrightsquigarrow ys\ .\mathrm{fst}$$

Almost all of the pieces are in place now: we know that the support lists of all proofs of $\mathcal{B}\ A$ are permutations of each other, and we know that sort returns a sorted permutation of its input. The final piece of the puzzle is the following:

```
sorted-perm-eq : ∀ xs ys → Sorted xs → Sorted ys → xs ↭ ys → xs ≡ ys
```

If two sorted lists are both permutations of each other they must be equal. Connecting up all the pieces we get the following:

```
perm-invar : ∀ xs ys → xs ↭ ys → sort xs ≡ sort ys
```

Because we know that all support lists of $\mathcal{B}\ A$ are permutations of each other this is enough to prove that sort is coherently constant, and therefore can eliminate from within a truncation. The second component of the output pair (the cover proof) follows quite naturally from the definition of permutations. ■

To the best of our knowledge, this is the first explicit proof of this theorem.

Restrictiveness. So far our explorations into finiteness predicates have pushed us in the direction of "less informative": however, as mentioned in the introduction, we can *also* ask how *restrictive* certain predicates are. Since split enumerability and manifest Bishop finiteness imply each other we know that there can be no type which satisfies one but not the other. We also know that manifest Bishop finiteness implies cardinal finiteness, but we do *not* have a function in the other direction:

(17)

$$\mathcal{C}\ A \to \mathcal{B}\ A$$

Solve the mystery of the mathcal

So the question arises naturally: is there a cardinally finite type which is *not* manifest Bishop finite? It turns out the answer is no! The proof of this fact is relatively short:

(18)
```
¬⟨𝒞 ⋂ ℬᶜ⟩ : ¬ Σ[ A : Type a ] 𝒞 A × ¬ ℬ A
¬⟨𝒞 ⋂ ℬᶜ⟩ (_ , c , ¬b) = rec isProp⊥ ¬b c
```

We can apply the function of type $\mathcal{B}\ A \to \bot$ (i.e. $\neg\ \mathcal{B}\ A$) to the value of type $\| \mathcal{B}\ A \|$ (i.e. $\mathcal{C}\ A$) using Equation 14, since $\bot$ is itself a proposition. This tells us that manifest bishop finiteness, cardinal finiteness, and split enumerability all refer to the same class of types.

Interestingly, while we cannot construct a function with the type in Equation 17, it does exist *classically*. In fact we can derive it from Equation 18 using the classical monad we developed in the introduction, since Equation 18 is actually equivalent to classical implication.

```
classical-impl : ¬ (A × ¬ B) → Classical (A → B)
classical-impl ¬A×¬B = do
  A? ← lem {A = A}
  B? ← lem {A = B}
  case (A? , B?) of
    λ { (inl a   , inl b   ) → pure (const b)
      ; (inl a   , inr ¬b ) → ⊥-elim (¬A×¬B (a , ¬b))
      ; (inr ¬a , inl b   ) → pure (const b)
      ; (inr ¬a , inr ¬b ) → pure (λ x → ⊥-elim (¬a x))
      }
```

1.4 Manifest Enumerability

Given that we have just proven that all of our finiteness predicates apply to the same types, the natural next step is to try find a predicate which applies to a different class of types. Let's first talk about what this new class of types might look like: what we're looking for is a type which is in some sense finite, but doesn't conform to any of the predicates we've seen so far. The *circle* (Listing 19) is such a type. The thing that this type has which precludes it from being, say, split enumerable, is its *higher homotopy structure*.

So far we have seen three levels of homotopy structure: the contractions (Equation 9), the propositions (Equation 12), and the sets (Equation 15). You may have noticed the pattern that each new level is generated by saying its paths are members of the previous level; if we apply that pattern again, we get to the next homotopy level: the groupoids.

```
data S¹ :
  base : S
  loop : base ≡ base
```

Listing 19: The Circle

(20)
```
isGroupoid A = (x y : A) → isSet (x ≡ y)
```

These types do not necessarily have unique identity proofs: there is more than one value which can inhabit the type ≡xy. The circle is one of the simplest examples of non-set groupoids: the constructor loop is the extra path in the type which isn't the identity path.

We now need to recall two facts: first, Hedberg's theorem tells us that every discrete type is a set. Second, every finiteness predicate we've seen thus far implies decidable equality. From this it's clear that all of the previous predicates are restricted to sets, and can't include types like the circle.

But the type certainly *seems* finite! It has finitely many points, for instance. In order to explore the "restrictiveness" axis in Figure 1, then, we'll need to construct a predicate which admits the circle. Manifest enumerability is one such predicate.

Definition 5 (Manifest Enumerability). Manifest enumerability is an enumeration predicate like Bishop finiteness or split enumerability with the only difference being a propositionally truncated membership proof.

$$\mathscr{E}\ A = \Sigma[\ \mathit{support} : \mathsf{List}\ A\ ]\ ((x : A) \to \|\ x \in \mathit{support}\ \|)$$

This predicate is novel, to the best of our knowledge.

It might not be immediately clear why this definition of enumerability allows the circle to conform while the others do not. The crux of the issue was that the cover proofs of the previous definitions didn't just tell us that some element was in the support list, they told us *where* it was in the support list. From the position we were able to derive decidable equality: that position is precisely what's hidden in manifest enumerability.

And indeed this means that the circle is manifestly enumerable.

$$\begin{aligned}
&\mathscr{E}\langle \mathsf{S}^1\rangle : \mathscr{E}\ \mathsf{S}^1\\
&\mathscr{E}\langle \mathsf{S}^1\rangle .\mathsf{fst}\ \ = [\ \mathsf{base}\ ]\\
&\mathscr{E}\langle \mathsf{S}^1\rangle .\mathsf{snd} = \|\mathsf{map}\|\ (0\ ,_) \circ \mathsf{isConnectedS}^1
\end{aligned}$$

We use a lemma here, proven in the Cubical Agda library, that S^1 is *connected*:

$$\mathsf{isConnectedS}^1 : (s : \mathsf{S}^1) \to \|\ \mathsf{base} \equiv s\ \|$$

Surjections. We already saw that split enumerability was the listed form of a split surjection: what we didn't explain was why the word "split" was placed before surjection. In the presence of higher homotopies than sets, split surjections are actually *not* a satisfactory definition of surjection. And we are most certainly in the presence of higher homotopies: just moments ago we were introduced to the circle. In these cases, the following definition of surjections is preferred [Univalent Foundations Program 2013, definition 4.6.1]:

$$\mathsf{Surjective}\ f = \forall\ y \to \|\ \mathsf{fiber}\ f\ y\ \| \qquad A \twoheadrightarrow B = \Sigma\ (A \to B)\ \mathsf{Surjective} \tag{21}$$

Much in the same way that split enumerability were split surjections, our new predicate of manifest enumerability corresponds to the proper surjections.

Lemma 1.8. *Manifest enumerability is equivalent to a surjection from a finite prefix of the natural numbers.*

$$\mathscr{E}\ A \Leftrightarrow \Sigma[\ n : \mathbb{N}\ ]\ (\mathit{Fin}\ n \twoheadrightarrow A)$$

Relation To Split Enumerability. It is trivially easy to construct a proof that any split enumerable type is manifest enumerable: we simply truncate the membership proof. Going the other way is more difficult, as we need to extract the membership proof from under a truncation. We do know what we need, however: the key difference between manifest enumerability and split enumerability is that the latter implied decidable equality. So that's the missing piece we should require in order to go from one to the other:

Lemma 1.9. *A manifestly enumerable type with decidable equality is split enumerable.*

Now that we know what extra bit of information we are allowed use in this proof, the path forward becomes a little more clear. In terms of the actual conversion function, the support list will stay the same, and only the return type of the cover proof needs to change: from $\|\ x \in xs\ \|$ to $x \in xs$.

That can be accomplished with the help of the following function:

```
recompute : Dec A → ‖ A ‖ → A
recompute (yes p)  _ = p
recompute (no ¬p) p = ⊥-elim (rec isProp⊥ ¬p p)
```
(22)

Given a decision procedure for some type, and a propositionally truncated value of that type, we can construct an element of the type.

In the case of $x \in xs$ we can construct a decision procedure for membership of a list, since we already have decidable equality on the elements of the list, proving our obligation.

1.5 Kuratowski Finiteness

We now finally arrive at the most important definition of finiteness: Kuratowski finiteness. As a definition, it is quite different from the predicates we've seen (it doesn't involve lists, for instance), but it plays a much larger role in the literature on finiteness predicates than, say, manifest enumerability.

We start with the definition of Kuratowski-finite subsets.

```
data 𝒦 (A : Type a) : Type a where
  []    : 𝒦 A
  _∷_ : A → 𝒦 A → 𝒦 A
  com : ∀ x y xs → x ∷ y ∷ xs ≡ y ∷ x ∷ xs
  dup  : ∀ x xs → x ∷ x ∷ xs ≡ x ∷ xs
  trunc : isSet (𝒦 A)
```
(23)

The first two constructors are point constructors, giving ways to create values of type $\mathcal{K}\ A$. They are also recognisable as the two constructors for finite lists, a type which represents the free monoid. The next two constructors add extra paths to the type: equations that usage of the type must obey. These extra paths turn the free monoid into the free *commutative* (com) *idempotent* (dup) monoid. The final constructor truncates the type $\mathcal{K}\ A$ to a set.

The Kuratowski finite subset is a free join semilattice (or, equivalently, a free commutative idempotent monoid). More prosaically, $\mathcal{K}$ is the abstract data type for finite sets, as defined in the Boom hierarchy [Boom 1981; Bunkenburg 1994]. However, rather than just being a specification, $\mathcal{K}$ is fully usable as a data type in its own right, thanks to HITs.

Other definitions of $\mathcal{K}$ exist (such as the one in [Frumin et al. 2018]) which make the fact that $\mathcal{K}$ is the free join semilattice more obvious. We have included such a definition in our formalisation, and proven it equivalent to the one above.

```
data 𝒦 (A : Type a) : Type a where
  η : A → 𝒦 A
  _∪_ : 𝒦 A → 𝒦 A → 𝒦 A
  ∅ : 𝒦 A
  ∪-assoc : ∀ xs ys zs → (xs ∪ ys) ∪ zs ≡ xs ∪ (ys ∪ zs)
  ∪-commutative : ∀ xs ys → xs ∪ ys ≡ ys ∪ xs
  ∪-idempotent : ∀ xs → xs ∪ xs ≡ xs
  ∪-identity : ∀ xs → xs ∪ ∅ ≡ xs
  trunc : isSet (𝒦 A)
```

Next, we need a way to say that an entire type is Kuratowski finite. For that, we will need to define membership of $\mathcal{K}$.

$$x \in [] \quad = \bot$$
$$x \in y :: ys = \| \equiv xy \uplus x \in ys \|$$

The com and dup constructors are handled by proving that the truncated form of $\uplus$ itself commutative and idempotent. The type of propositions is itself a set, satisfying the trunc constructor. This gives us enough to define Kuratowski finiteness.

Definition 6 (Kuratowski Finiteness). A type is Kuratowski finite if there exists a Kuratowski-finite subset of that type which contains every element of the type.

$$\mathcal{K}^{\mathrm{f}}\ A = \Sigma[\ xs : \mathcal{K}\ A\ ]\ ((x : A) \to x \in xs)$$

While Kuratowski finiteness is something of the standard formal definition of finiteness, it is quite separated from the enumeration-based definitions we have presented so far. It's difficult to relate to surjections and equivalences, and requires a different style of proof to reason about. As such, we want to get *away* from Kuratowski finiteness as quickly as possible. To do so we use the following lemma:

Lemma 1.10. *Kuratowski finiteness is equivalent to truncated manifest enumerability.*

$$\| \mathscr{E}\ A \| \Leftrightarrow \mathcal{K}^{f}\ A$$

Proof. This proof is constructed by providing a pair of functions, to and from each side of the equivalence. This pair implies an equivalence, because both source and target are propositions. This proof, as well as its auxiliary lemmas, are also provided in Frumin et al. [2018], although there the setting is HoTT rather than CuTT. ■

By relating Kuratowski finiteness—with a full equivalence, no less—to an enumerated predicate, we have made it possible to talk about Kuratowski finiteness without interacting with the type at all.

In the next subsection, we will explore the category of discrete Kuratowski finite sets. Under the hood, however, we will really be working with cardinal finite sets. We can do this in a fully rigorous way because Lemma 1.10 allows us to prove the following:

(24)
$$\mathscr{C}\ A \Leftrightarrow \mathcal{K}^{\mathrm{f}}\ A \times \mathrm{Discrete}\ A$$

2 TOPOS

In this subsection we will examine the categorical interpretation of finite sets. In particular, we will prove that discrete Kuratowski finite types form a Π-pretopos. A lot of the work for this proof has been done already: we have already proven that discrete Kuratowski finiteness is equivalent to cardinal finiteness (Theorem 24), meaning that we can work with the latter definition which is much simpler to prove things about.

There are two reasons we're interested in the categorical and topos-theoretic interpretation of finite sets: first, it's an important theoretical grounding for finite sets, which allows us to understand them in the context of other set-like constructions. Secondly, and more practically, the language of a topos is (or in our case the Π-pretopos) is a common standard framework for doing mathematics generally. This makes it a good basis for an API for building QuickCheck-like generators, for example.

2.1 Categories in HoTT

At first glance, HoTT seems like a perfect setting for category theory: the univalence axiom identifies isomorphisms with equality, a useful tool for category theory missing from MLTT. While this initial impression is broadly true, the construction of categories in HoTT is unfortunately quite complex and involved.

Much of this subsection is simply a summary of parts of Univalent Foundations Program [2013, section 9]. The formal proofs we provide are part translation of those proofs in that section, part from [Iversen 2018] [Hu and Carette 2020], and part our own.

First, we need to think about the type of objects and arrows. We cannot, unfortunately, leave them unrestricted: because of the potential for higher homotopy in HoTT types, we have to restrict the type of arrows to just the sets. This notion: that of a category with all the usual laws such that arrows are a set, is called a *precategory*.

(25)
```
record PreCategory ℓ₁ ℓ₂ : Type (ℓsuc (ℓ₁ ℓ⊔ ℓ₂)) where
  field
    Ob      : Type ℓ₁
    Hom     : Ob → Ob → Type ℓ₂
    Id      : ∀ {X} → Hom X X
    Comp    : ∀ {X Y Z} → Hom Y Z → Hom X Y → Hom X Z
    assoc-Comp : ∀ {W X Y Z}
                 (f : Hom Y Z)
                 (g : Hom X Y)
                 (h : Hom W X) →
                 Comp f (Comp g h) ≡ Comp (Comp f g) h
    Comp-Id : ∀ {X Y} (f : Hom X Y) → Comp f Id ≡ f
    Id-Comp : ∀ {X Y} (f : Hom X Y) → Comp Id f ≡ f
    Hom-Set : ∀ {X Y} → isSet (Hom X Y)
```

We will use long arrows to refer to morphisms within a category:

```
_⟶_ = Hom
```

From here, we can define a notion of isomorphisms.

(26)
```
Isomorphism : (X ⟶ Y) → Type ℓ₂
Isomorphism {X} {Y} f = Σ[ g : Y ⟶ X ] ((g · f ≡ Id) × (f · g ≡ Id))

X ≅ Y = Σ (X ⟶ Y) Isomorphism
```

It's a condition on this type which separates the precategories from the categories: if it satisfies a form of univalence, it the precategory is a full category.

(27)
```
univalent : {X Y : Ob} → (X ≡ Y) ≃ (X ≅ Y)
```

2.2 The Category of Sets

Next we'll look at how to construct the category of sets (in the HoTT sense). Much of this work comes directly from Rijke and Spitters [2015] and Univalent Foundations Program [2013, section 10] (the latter of which is in fact an updated and slightly less detailed version of the former). In particular, our treatment (and definition) of categories and topoi comes directly from those works.

We have provided in the formalisation a proof that sets in CuTT form a Π-pretopos: this proof (in HoTT) is in fact the main result of Rijke and Spitters [2015]; our contribution is simply the formalisation.

The objects are represented by a Σ:

$$\mathsf{Ob} = \Sigma[\ t : \mathsf{Type}_0\ ]\ \mathsf{isSet}\ t$$

This will be quite similar to our objects for finite sets.

Since sets in HoTT don't form a topos, there are quite a few smaller lemmas we need to prove to get as close as we can (a ΠW-pretopos): we won't include them here, other than the closure proofs in the following subsection.

2.3 Closure

The two most involved proofs for showing that discrete Kuratowski sets form a Π-pretopos are those proofs that show closure under Π and Σ. We will describe them here.

In [Frumin et al. 2018, Theorem 4.21], Kuratowski finite types are proven to be closed under surjections, products, and sums. Here we prove closure under products and sums, but also functions, Σ, and Π (and furthermore our closure proofs are given on all of the finiteness predicates that they apply to).

Closure of the Ordered Predicates. First, we will show that split enumerability (and, by extension, manifest enumerability) are closed under Π and Σ. This is the first stepping stone on our way to prove that cardinal finiteness is closed under the same.

Practically speaking, these proofs also open up a wide number of other closure proofs to us. By proving that dependent products and sums are finite, we get the non-dependent cases for free.

Lemma 2.1. *Split enumerability is closed under* Σ.

$$\_/\Sigma/\_ : \mathscr{E}!\ A \to (\forall\ x \to \mathscr{E}!\ (U\ x)) \to \mathscr{E}!\ (\Sigma\ A\ U)$$

Proof. Our task is to construct the two components of the output pair: the support list, and the cover proof. We'll start with the support list: this is constructed by taking the Cartesian product of the input support lists.

```
sup-Σ : List A →
        ((x : A) → List (U x)) →
        List (Σ A U)
sup-Σ xs ys = do x ← xs
                 y ← ys x
                 [ x , y ]
```

We use do notation here because we're working the list monad: this applies the latter function (ys) to every element of the list xs, and concatenates the results.

To show that this does indeed cover every element of the target type is a little intricate, but not necessarily difficult. ■

Next we'll look at closure under Π. In MLTT, this is of course not provable: since all of the finiteness predicates we have seen so far imply decidable equality, and since we don't have any kind of decidable equality on functions in MLTT, we know that we won't be able to show that any kind of function is finite; even one like Bool → Bool.

CuTT is not so restricted. Since we have things like function extensionality and transport, we can indeed prove the finiteness of function types. Our proof here makes use directly of the univalence axiom, and makes use furthermore of all the previous closure proofs.

Theorem 2.2. *Split enumerability is closed under dependent functions (Π-types).*

$$\_|\Pi|\_ : \mathcal{E}!\ A \to ((x : A) \to \mathcal{E}!\ (U\ x)) \to \mathcal{E}!\ ((x : A) \to U\ x)$$

Proof. Let A be a split enumerable type, and U be a type family from A, which is split enumerable over all points of A.

As A is split enumerable, we know that it is also manifestly Bishop finite (Lemma 1.3), and consequently we know $A \simeq \mathsf{Fin}\ n$, for some n (Lemma 1.4). We can therefore replace all occurrences of A with $\mathsf{Fin}\ n$, changing our goal to:

$$\frac{\mathcal{E}!\ (\mathsf{Fin}\ n) \quad ((x : \mathsf{Fin}\ n) \to \mathcal{E}!\ \ (U\ x))}{\mathcal{E}!\ ((x : \mathsf{Fin}\ n) \to U\ x)}$$

We then define the type of n-tuples over some type family.

$$\begin{aligned}
&\mathsf{Tuple} : \forall\ n \to (\mathsf{Fin}\ n \to \mathsf{Type}_0) \to \mathsf{Type}_0\\
&\mathsf{Tuple}\ \mathsf{zero}\quad f = \top\\
&\mathsf{Tuple}\ (\mathsf{suc}\ n)\ f = f\,\mathsf{f0} \times \mathsf{Tuple}\ n\ (f \circ \mathsf{fs})
\end{aligned}$$

We can show that this type is equivalent to functions (proven in our formalisation):

$$\mathsf{Tuple}\ n\ U \Leftrightarrow ((i : \mathsf{Fin}\ n) \to U\ i)$$

And therefore we can simplify again our goal to the following:

$$\frac{\mathcal{E}!\ (\mathsf{Fin}\ n) \quad ((x : \mathsf{Fin}\ n) \to \mathcal{E}!\ (U\ x))}{\mathcal{E}!\ \ (\mathsf{Tuple}\ n\ U)}$$

We can prove this goal by showing that $\mathsf{Tuple}\ n\ U$ is split enumerable: it is made up of finitely many products of points of U, which are themselves split enumerable, and $\top$, which is also split enumerable. Lemma 2.1 shows us that the product of finitely many split enumerable types is itself split enumerable, proving our goal. ■

Closure on Cardinal Finiteness. Since we don't have a function of type $\mathcal{C}\ A \to \mathcal{B}\ A$, closure proofs on $\mathcal{B}$ do not transfer over to $\mathcal{C}$ trivially (unlike with $\mathcal{E}!$ and $\mathcal{B}$). The cases for $\bot$, $\top$, and Bool are simple to adapt: we can just propositionally truncate their Bishop finiteness proof.

Non-dependent operators like $\times$, $\uplus$, and $\to$ are also relatively straightforward: since $\|\_\|$ forms a monad, we can apply n-ary functions to values inside it, combining them together.

$$\begin{aligned}
|\times| : &\ \mathcal{B}\ A \to\\
&\ \mathcal{B}\ B \to\\
&\ \mathcal{B}\ (A \times B)
\end{aligned}$$

Into a truncated context:

```
_||×||_ : 𝒞 A →
          𝒞 B →
          𝒞 (A × B)
xs ||×|| ys = do
  x ← xs
  y ← ys
  | x |×| y |
```

Unfortunately, for the dependent type formers like Σ and Π, the same trick does not work. We have closure proofs like:

$$\frac{\mathcal{B}\ A \quad ((x : A) \to \mathcal{B}\ (U\ x))}{\mathcal{B}\ ((x : A) \to U\ x)}$$

If we apply the monadic truncation trick we can derive closure proofs like the following:

$$\frac{\| \mathcal{B}\ A \| \quad \| ((x : A) \to \mathcal{B}\ (U\ x)) \|}{\| \mathcal{B}\ ((x : A) \to U\ x) \|}$$

However our *desired* closure proof is the following:

$$\frac{\| \mathcal{B}\ A \| \quad ((x : A) \to \| \mathcal{B}\ (U\ x) \|)}{\| \mathcal{B}\ ((x : A) \to U\ x) \|}$$

They don't match!

The solution would be to find a function of the following type:

$$((x : A) \to \| \mathcal{B}\ (U\ x) \|) \to \| (x : A) \to \mathcal{B}\ (U\ x) \|$$

However we might be disheartened at realising that this is a required goal: the above equation is *extremely* similar to the axiom of choice!

Definition 7 (Axiom of Choice). In HoTT, the axiom of choice is commonly defined as follows [Univalent Foundations Program 2013, lemma 3.8.2]. For any set A, and a type family U which is a set at all the points of A, the following function exists:

$$((x : A) \to \| U(x) \|) \to \| (x : A) \to U(x) \|$$

Luckily the axiom of choice *does* hold for cardinally finite types, allowing us to prove the following:

Lemma 2.3. *The axiom of choice holds for finite sets.*

$$C\ A \to ((x : A) \to \| U(x) \|) \to \| (x : A) \to U(x) \|$$

Proof. Let A be a cardinally finite type, U be a type family on A, and f be a dependent function of type $\Pi(x : A), \| U(x) \|$.

First, since our goal is itself propositionally truncated, we have access to values under truncations: put another way, in the context of proving our goal, we can rely on the fact that A is manifestly Bishop finite. Using the same technique as we did in Lemma 2.2, we can switch from working with dependent functions from A to n-tuples, where n is the cardinality of A. This changes our goal to the following:

$$\mathsf{Tuple}\ n\ (\|\_\| \circ U) \to \| \mathsf{Tuple}\ n\ U \| \tag{28}$$

Since $\|\_\|$ is closed under finite products, this function exists (in fact, using the fact that $\|\_\|$ forms a monad, we can recognise this function as `sequenceA` from the `Traversable` class in Haskell). ■

This lemma is a well-known folklore theorem.

This gets us all of the necessary closure proofs on C, and as a result we know the following theorem.

Theorem 2.4. *Decidable Kuratowski finite sets form a* Π*-pretopos.*

2.4 The Absence of the Subobject Classifier

It's a little unsatisfying that our topos construction has so many caveats: we have to prove a lot of small, uninteresting lemmas just to get to a Π-pretopos, all because we can't prove the one or two larger, simple lemmas which would show that sets form a topos. So what exactly are we missing?

Well, one of the characteristic features of topos theory is that there are a wide variety of equivalent ways to show that something is a topos (a natural consequence of their being a wide variety of things which qualify as toposes). For the direction we have been going, though, the big missing feature is the *subobject classifier*.

A subobject in this context refers to a subset. In set theory, we can often describe a subset of some set A with the following notation:

$$\{x \mid x \in A; P(x)\}$$

This is the subset of elements in A which satisfy some predicate P.

Type theoretically, the way to express the same would be $\Sigma\ A\ P$: if we wanted to describe the subset of $\mathbb{N}$smaller than 10 we would write $\Sigma[\ n : \mathbb{N}\ ]\ n < 10$. In general, however, this type holds too many elements to properly classify the subsets of the larger set: there may be more than one inhabitant of $P\ x$ for any given x. For *propositions*, however, (i.e. where P is a proposition), Σ represents a perfectly valid encoding of subsets.

The subobject classifier is an object within the topos (which must be a contraction) which classifies monomorphisms (injections). We can actually show that the "subset" notion we just defined does in fact classify monomorphisms in sets in HoTT (in fact directly through univalence), but at this point we run into our one and only size problem in this thesis. The actual object corresponding to the subobject classifier is the following:

$$\begin{aligned}&\text{Prop-univ} : \text{Type}_1\\ &\text{Prop-univ} = \Sigma[\ t : \text{Type}_0\ ]\ \text{isProp}\ t\end{aligned}$$

The problem here, crucially, is that the universe level of this type is one higher than the universe level of the types it bounds. In other words, this is *not* an object in our Π-pretopos of sets, where the types are all of universe level 0.

Remember that the purpose of universe levels was to prevent Girard's paradox. However, there is an axiom which removes universe levels to a certain extent which does *not* imply the paradox: propositional resizing.

Definition 8 (Propositional Resizing). The axiom of propositional resizing states that the following two types, for any universe level u, are equivalent:

$$\Sigma[\ t : \text{Type}\ u\ ]\ \text{isProp}\ t \simeq \Sigma[\ t : \text{Type}\ (\ell\text{suc}\ u)\ ]\ \text{isProp}\ t$$

If propositional resizing holds, then we *can* in fact construct a subobject classifier, for both sets and finite sets.

Mention lem?

3 SEARCH

A common theme in dependently-typed programming is that proofs of interesting theoretical things often correspond to useful algorithms in some way related to that thing. Finiteness is one

such case: if we have a proof that a type A is finite, we should be able to search through all the elements of that type in a systematic, automated way.

As it happens, this kind of search is a very common method of proof automation in dependently-typed languages like Agda. Proofs of statements like "the following function is associative"

```
_∧_ : Bool → Bool → Bool
false ∧ false = false
false ∧ true  = false
true  ∧ false = false
true  ∧ true  = true
```

can be tedious: the associativity proof in particular would take $2^3 = 8$ cases. This is unacceptable! There are only finitely many cases to examine, after all, and we're *already* on a computer: why not automate it? A proof that Bool is finite can get us much of the way to a library to do just that.

Similar automation machinery can be leveraged to provide search algorithms for certain "logic programming"-esque problems. Using the machinery we will describe in this subsection, though, when the program says it finds a solution to some problem that solution will be accompanied by a formal *proof* of its correctness.

In this subsection, we will describe the theoretical underpinning and implementation of a library for proof search over finite domains, based on the finiteness predicates we have introduced already. The library will be able to prove statements like the proof of associativity above, as well as more complex statements. As a running example for a "more complex statement" we will use the countdown problem, which we have been using throughout: we will demonstrate how to construct a prover for the existence of, or absence of, a solution to a given countdown puzzle.

The API for writing searches over finite domains comes from the language of the Π-pretopos: with it we will show how to compose QuickCheck-like generators for proof search, with the addition of some automation machinery that allows us to prove things like the associativity in a couple of lines:

(29)
```
∧-assoc : ∀ x y z → (x ∧ y) ∧ z ≡ x ∧ (y ∧ z)
∧-assoc = ∀ℵ₀ⁿ 3 λ x y z → (x ∧ y) ∧ z ≟ x ∧ (y ∧ z)
```

We have already, in previous sections, explored the theoretical implications of Cubical Type Theory on our formalisation. With this library for proof search, however, we will see two distinct practical applications which would simply not be possible without computational univalence. First and foremost: our proofs of finiteness, constructed with the API we will describe, have all the power of full equalities. Put another way any proof over a finite type A can be lifted to any other type with the same cardinality. Secondly our proof search can range over functions: we could, for instance, have asked the prover to find if *any* function over Bool is associative, and if so return it to us.

```
some-assoc : Σ[ f ∶ (Bool → Bool → Bool) ] ∀ x y z → f (f x y) z ≡ f x (f y z)
some-assoc = ∃ℵ₀ⁿ 1 λ f → ∀?ⁿ 3 λ x y z → f (f x y) z ≟ f x (f y z)
```

The usefulness of which is dubious, but we will see a more interesting application soon.

3.1 How to make the Typechecker do Automation

For this prover we will not resort to reflection or similar techniques: instead, we will trick the type checker to do our automation for us. This is a relatively common technique, although not so much outside of Agda, so we will briefly explain it here.

To understand the technique we should first notice that some proof automation *already* happens in Agda, like the following:

```
obvious : true ∧ false ≡ false
obvious = refl
```

The type checker does not require us to manually explain each step of evaluation of true ∧ false. While it's not a particularly impressive example of automation, it does nonetheless demonstrate a principle we will exploit: closed terms will compute to a normal form if they're needed to type check. The type checker will perform β-reduction as much as it can.

So our task is to rewrite proof obligations like the one in Equation 29 into ones which can reduce completely. As it turns out, we have already described the type of proofs which can "reduce completely": *decidable* proofs. If we have a decision procedure over some proposition P we can run that decision during type checking, because the decision procedure itself is a proof that the decision will terminate. In code, we capture this idea with the following pair of functions:

```
True : Dec A → Type₀
True (yes _) = ⊤
True (no  _) = ⊥
```

```
toWitness : (decision : Dec A) →
            { _ : True decision } → A
toWitness (yes x) = x
```

The first is a function which derives a type from whether a decision is successful or not. This function is important because if we use the output of this type at any point we will effectively force the unifier to run the decision computation. The second takes—as an implicit argument—an inhabitant of the type generated from the first, and uses it to prove that the decision can only be true, and the extracts the resulting proof from that decision. All in all, we can use it like this:

```
extremely-obvious : true ≢ false
extremely-obvious = from-true (! (true ≟ false))
```

This technique will allow us to automatically compute any decidable predicate.

3.2 Omniscience

So we now know what is needed of us for proof automation: we need to take our proofs and make them decidable. In particular, we need to be able to "lift" decidability back over a function arrow. For instance, given x, y, and z we already have Dec $((x \wedge y) \wedge z \equiv x \wedge (y \wedge z))$ (because equality over booleans is decidable). In order to turn this into a proof that $\wedge$ is associative we need Dec $(\forall\ x\ y\ z \to (x \wedge y) \wedge z \equiv x \wedge (y \wedge z))$. The ability to do this is described formally by the notion of "Exhaustibility".

```
Exhaustible p A = ∀ {P : A → Type p} → (∀ x → Dec (P x)) → Dec (∀ x → P x)
```

We say a type A is exhaustible if, for any decidable predicate P on A, the universal quantification of the predicate is decidable.

This property of Bool would allow us to automate the proof of associativity, but it is in fact not strong enough to find individual representatives of a type which support some property. For that we need the more well-known related property of *omniscience*.

```
Omniscient p A = ∀ {P : A → Type p} → (∀ x → Dec (P x)) → Dec (∃[ x ] P x)
```

The "limited principle of omniscience" [Bishop 1967] is a classical principle which says that omniscience holds for all sets. It doesn't hold constructively, of course: it lies a little bit below LEM

in terms of its non-constructiveness, given that it can be derived from LEM but LEM cannot be derived from it.

Omniscience implies exhaustibility: we can use the usual rule of $\neg\exists x.P(x) \iff \forall x.\neg P(x)$ to turn omniscience for some predicate P into exhaustibility for some predicate $\neg\neg P$. Usually we don't have double negation elimination constructively, but since P is decidable it's actually present in this case:

```
Dec→DoubleNegElim : (A : Type a) → Dec A → ¬ ¬ A → A
Dec→DoubleNegElim A (yes p) _       = p
Dec→DoubleNegElim A (no ¬p) contra = ⊥-elim (contra ¬p)
```

All together, this gives us the following proof:

```
Omniscient→Exhaustible : Omniscient p A → Exhaustible p A
Omniscient→Exhaustible omn P? =
  map-dec
    (λ ¬∃P x → Dec→DoubleNegElim _ (P? x) (¬∃P ∘ (x ,_)))
    (λ ¬∃P ∀P → ¬∃P λ p → p .snd (∀P (p .fst)))
    (! (omn (! ∘ P?)))
```

Our focus here is on those types for which omniscience *does* hold, which includes the (ordered) finite types. Perhaps surprisingly, it is not *only* finite types which are exhaustible. Certain infinite types can be exhaustible [Escardo 2007], but an exploration of that is beyond the scope of this work.

All of the finiteness predicates imply exhaustibility. To prove that fact we'll just show that the Kuratowski finite types are exhaustible: since it's the weakest predicate, and can be derived from all the others.

Lemma 3.1. *Kuratowski finiteness implies exhaustibility.*

Manifest enumerability is similarly the weakest of the ordered predicates:

Lemma 3.2. *Manifest enumerability implies omniscience.*

We won't provide these full proofs here, since they are rather tedious and don't provide much insight.

Finally, there is a form of omniscience which works with Kuratowski finiteness:

```
Prop-Omniscient p A = ∀ {P : A → Type p} → (∀ x → Dec (P x)) → Dec ‖ ∃[ x ] P x ‖
```

By truncating the returned Σ we don't reveal which A we've chosen which satisfies the predicate: this means that it can be pulled out of the Kuratowski finite subset without issue.

```
𝒦ᶠ⇒Prop-Omniscient : 𝒦ᶠ A → Prop-Omniscient p A
𝒦ᶠ⇒Prop-Omniscient K P? =
  PropTrunc.rec
    (isPropDec squash)
    (map-dec |_| refute-trunc ∘ λ xs → ℰ⇒Omniscient xs P?)
    (𝒦ᶠ⇒‖ℰ‖ K)
```

With the knowledge that any Kuratowski finite type implies exhaustibility we know that we can do proof search over all of the types we have proven to be Kuratowski finite: the 0, 1, and 2 types; (dependent) sums and products; and any type proven to be equivalent to these. It's still not entirely

clear how to actually *use* this automation without incurring so much boilerplate as to defeat the point, though.

3.3 An Interface for Proof Automation

In this subsection we will present the more user-friendly interface to the library, designed to be used to automate away tedious proofs in an easy way.

The Design of the Interface. The central idea of the interface to the proof search library are the following two functions:

```
∀? : ℰ! A →
     (∀ x → Dec (P x)) →
     Dec (∀ x → P x)
∀? ℰ!⟨A⟩ = ℰ!⇒Exhaustible ℰ!⟨A⟩

∃? : ℰ! A →
     (∀ x → Dec (P x)) →
     Dec (∃[ x ] P x)
∃? ℰ!⟨A⟩ = ℰ!⇒Omniscient ℰ!⟨A⟩
```

Clearly they're just restatements of exhaustibility and omniscience. However, we can combine these functions with the automation technique from above to create the following:

```
∀↯ : (ℰ!⟨A⟩ : ℰ! A) →
     (P? : ∀ x → Dec (P x)) →
     ⦃ _ : True (∀? ℰ!⟨A⟩ P?) ⦄ →
     ∀ x → P x
∀↯ _ _ ⦃ t ⦄ = toWitness t

∃↯ : (ℰ!⟨A⟩ : ℰ! A) →
     (P? : ∀ x → Dec (P x)) →
     ⦃ _ : True (∃? ℰ!⟨A⟩ P?) ⦄ →
     ∃[ x ] P x
∃↯ _ _ ⦃ t ⦄ = toWitness t
```

This automation procedure allows us to state the property succinctly, and have the type checker go and run the decision procedure to solve it for us. Here's an example of its use:

```
∧-idem : ∀ x → x ∧ x ≡ x
∧-idem = ∀↯ ℰ!⟨2⟩ λ x → x ∧ x ≟ x
```

Instances. One bit of cruft in the above proof is the need to specify the particular finiteness proof for bools. While this isn't any great burden in this case, it of course becomes more difficult in more complex circumstances.

To solve this we can use Agda's instance search. This changes the definitions of our automation functions to the following:

```
∀↯ : ⦃ ℰ!⟨A⟩ : ℰ! A ⦄ →
     (P? : ∀ x → Dec (P x)) →
     ⦃ _ : True (∀? ℰ!⟨A⟩ P?) ⦄ →
     ∀ x → P x
∀↯ _ ⦃ t ⦄ = toWitness t

∃↯ : ⦃ ℰ!⟨A⟩ : ℰ! A ⦄ →
     (P? : ∀ x → Dec (P x)) →
     ⦃ _ : True (∃? ℰ!⟨A⟩ P?) ⦄ →
     ∃[ x ] P x
∃↯ _ ⦃ t ⦄ = toWitness t
```

And this also changes the idempotency proof to the following:

(30)

```
∧-idem : ∀ x → x ∧ x ≡ x
∧-idem = ∀↯ λ x → x ∧ x ≟ x
```

Again, there's not any great revelation in ease of use here, but more complex examples really benefit. Especially when we build the full set of instances: any expression built out of products and sums will automatically have an instance. This will allow us, for instance, to perform proof search over tuples, which gives us some degree of automation for proof search in tuples.

```
∧-comm : ∀ x y → x ∧ y ≡ y ∧ x
∧-comm = curry (∀ⁱ (uncurry (λ x y → x ∧ y ≟ y ∧ x)))
```

These instances aren't limited to non-dependent sums and products, either: for Σ, for instance, we already have a proof that $\mathcal{E}!\ A \to (\forall x \to \mathcal{E}!\ (B\ x)) \to \mathcal{E}!\ (\Sigma\ A\ B)$. Since A is finite, we can construct a finite constraint that "B is finite at all points of A", and use that to statically build our instance.

```
_ : ℰ! (Σ[ s ∶ Bool ] (if s then Fin 3 else Fin 4))
_ = it
```

The it function here is a clever helper function. It's defined like so:

```
it : ⦃ _ : A ⦄ → A
it ⦃ x ⦄ = x
```

Basically it searches for an instance for the type in the hole that it's put into: it's a way of asking Agda to "find an instance which fits here".

Generic Currying and Uncurrying. While we have arguably removed the bulk of the boilerplate from the automated proofs, there is still the case of the ugly noise of currying and uncurrying. In this subsection, we take inspiration from Allais [2019] to develop a small interface to generic n-ary functions and properties. We will describe it briefly here.

The basic idea of currying and uncurrying generically is to allow ourselves to work with a generic and flexible representation of function arguments which can be manipulated more easily than a simple function itself. Our first task, then, is to define that representation of function arguments. As in Allais [2019], our representation is a tuple which is in some sense a "second order" indexed type. By second order here we mean that it is an indexed type indexed by another indexed type. The reason for this complexity is that our solution is to be fully level-polymorphic. To start, we define a type representing a vector of universe levels:

```
Levels : ℕ → Type₀
Levels zero = ⊤
Levels (suc n) = Level × Levels n
```

```
max-level : Levels n → Level
max-level {zero} _ = ℓzero
max-level {suc n} (x , xs) =
  x ℓ⊔ max-level xs
```

This will be used to assign our tuple the correct universe level generically. Next, we define the list of types (this type is indexed by the list of universe levels of each type):

```
Types : ∀ n → (ls : Levels n) → Type (ℓsuc (max-level ls))
Types zero ls = ⊤
Types (suc n) (l , ls) = Type l × Types n ls
```

And finally, the tuple, indexed by its list of types:

```
(|_|)⁺ : Types (suc n) ls →
         Type (max-level ls)
(|_|)⁺ {n = zero } (X , Xs) = X
(|_|)⁺ {n = suc n} (X , Xs) = X × (| Xs |)⁺
```

```
(|_|) : Types n ls →
        Type (max-level ls)
(|_|) {n = zero} _ = ⊤
(|_|) {n = suc n} = (|_|)⁺ {n = n}
```

The reason for two separate functions here is to avoid the ⊤-terminated tuples we would need if

we just had one. This means that, for instance, to represent a tuple of a Bool and $\mathbb{N}$we can write (true , 2) instead of (true , 2 , tt).

Next we turn to how we will represent functions. In Agda there are three ways to pass function arguments: explicitly, implicitly, and as an instance. We will represent these three different versions with a data type:

```
data ArgForm : Type₀ where expl impl inst : ArgForm
```

And then we can make a type for functions in the general sense: a type which has this sum type as a parameter.

```
_[_]→_ : Type a → ArgForm → Type b → Type (a ℓ⊔ b)
A [ expl ]→ B =      A   → B
A [ impl ]→ B = { _ : A } → B
A [ inst ]→ B = ⦃ _ : A ⦄ → B
```

And we can show that this is isomorphic to a normal function:

```
[_$] : ∀ form → (A [ form ]→ B) ⇔ (A → B)
```

This of course is only a representation of *non*-dependent functions. Dependent functions are defined in a similar way:

```
Π[_$] : ∀ {B : A → Type b} fr → (x : A Π[ fr ]→ B x) ⇔ ((x : A) → B x)
```

Using both of these things, we can now define a generic type for multi-argument functions:

```
(|_|)[_]→_ : Types n ls → ArgForm → Type ℓ → Type (max-level ls ℓ⊔ ℓ)
(|_|)[_]→_ {n = zero} Xs fr Y = Y
(|_|)[_]→_ {n = suc n} (X , Xs) fr Y = X [ fr ]→ (| Xs |)[ fr ]→ Y
```

We can also define multi-argument dependent functions in a similar way. Similarly to how we had to define two tuple types in order to avoid the ⊤-terminated tuples, we have two definitions for multi-argument dependent functions. We only include the nonempty version here for brevity.

```
pi-arrs-plus :
  (Xs : Types (suc n) ls) →
  ArgForm →
  (y : (| Xs |)⁺ → Type ℓ) →
  Type (max-level ls ℓ⊔ ℓ)
pi-arrs-plus {n = zero  } (X , Xs) fr Y = x : X Π[ fr ]→ Y x
pi-arrs-plus {n = suc n } (X , Xs) fr Y =
  x : X Π[ fr ]→ xs :(| Xs |)⁺ Π[ fr ]→ Y (x , xs)
```

Finally, this all allows us to define an isomorphism between generic multi-argument dependent functions and their uncurried forms.

```
Π[_^_$] : ∀ n {ls ℓ} fr {Xs : Types n ls} {Y : (| Xs |) → Type ℓ} →
          (xs :(| Xs |)Π[ fr ]→ Y xs) ⇔ ((xs : (| Xs |)) → Y xs)
```

The use of all of this is that we can take the user-supplied curried version of a function and transform it into a version which takes instance arguments for each of the types.

```
∃?ⁿ : ⦇ map-types ℰ! Xs ⦈[ inst ]⟶
      xs :⦇ Xs ⦈Π[ expl ]⟶
      Dec (P xs) [ expl ]⟶
      Dec (Σ ⦇ Xs ⦈ P)
∃?ⁿ =  [ n ^ inst $] .inv λ fs
    ⟶ ℰ!⇒Omniscient (tup-inst n fs)
    ∘ Π[ n ^ expl $] .fun
```

```
∃↯ⁿ :
  insts :⦇ map-types ℰ! Xs ⦈Π[ inst ]⟶
  ((P? : xs :⦇ Xs ⦈Π[ expl ]⟶ Dec (P xs))
    ⟶ {⦃ _ : True
        (ℰ!⇒Omniscient
          (tup-inst n insts)
          (Π[ n ^ expl $] .fun P?)) ⦄}
    ⟶ Σ ⦇ Xs ⦈ P)
∃↯ⁿ =
  Π[ n ^ inst $] .inv
    λ fs P? ⦃ p ⦄ ⟶ toWitness p
```

While the type signatures involved are complex, the usage is not. Finally, here is how we can automate the proof of commutativity fully:

```
∧-comm : ∀ x y ⟶ x ∧ y ≡ y ∧ x
∧-comm = ∀↯ⁿ 2 λ x y ⟶ x ∧ y ≟ y ∧ x
```

With that, we now have a simple interface to a proof search library, which can be used to automate away certain tedious proofs.

Automation of simple proofs like the associativity of conjunction is all well and good, but tasks like that are more tedious than they are difficult. What about more difficult problems? In the next subsection we will look at a problem which is too complex to be solved by the simple instance-search solver we have constructed here. Instead, we will have to combine instance search with manual construction of finiteness proofs, optimisation of representation, and some other tricks. At the end of it, we will have a solver for countdown.

3.4 Countdown

The Countdown problem [Hutton 2002] is a well-known puzzle in functional programming (which was apparently turned into a TV show). As a running example in this paper, we will produce a verified program which lists all solutions to a given countdown puzzle: here we will briefly explain the game and our strategy for solving it.

The idea behind countdown is simple: given a list of numbers, contestants must construct an arithmetic expression (using a small set of functions) using some or all of the numbers, to reach some target. Here's an example puzzle:

> Using some or all of the numbers 1, 3, 7, 10, 25, and 50 (using each at most once), construct an expression which equals 765.

We'll allow the use of +, −, ×, and ÷. The answer is at the bottom of this page[2].

Our strategy for finding solutions to a given puzzle is to describe precisely the type of solutions to a puzzle, and then show that that type is finite. So what is a "solution" to a countdown puzzle? Broadly, it has two parts:

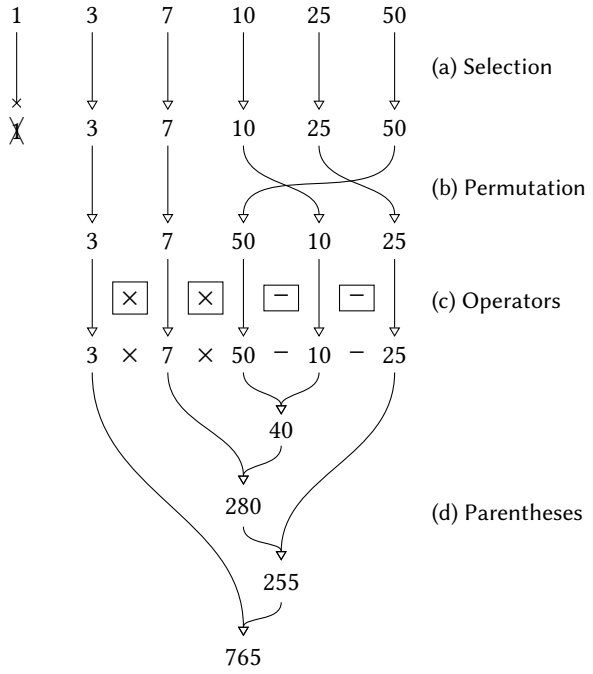

Fig. 2. The components of a transformation which makes up a Countdown candidate solution

A Transformation from a list of numbers to an expression.

A Predicate showing that the expression is valid and evaluates to the target.

The first part is described in Figure 2.

This transformation has four steps. First (Fig. 2a) we have to pick which numbers we include in our solution. We will need to show there are finitely many ways to filter n numbers.

Secondly (Fig. 2b) we have to permute the chosen numbers. The representation for a permutation is a little trickier to envision: proving that it's finite is trickier still. We will need to rely on some of the more involved lemmas later on for this problem.

The third step (Fig. 2c) is a vector of length n of finite objects (in this case operators chosen from +, ×, −, and ÷). Although it is complicated slightly by the fact that the n in this n-tuple is dependent on the amount of numbers we let through in the filter in step one. (in terms of types, that means we'll need a Σ rather than a $\times$, explanations of which are forthcoming).

Finally (Fig. 2d), we have to parenthesise the expression in a certain way. This can be encapsulated by a binary tree with a certain number of leaves: proving that that is finite is tricky again.

Once we have proven that there are finitely many transformations for a list of numbers, we will then have to filter them down to those transformations which are valid, and evaluate to the target. This amounts to proving that the decidable subset of a finite set is also finite.

Finally, we will also want to optimise our solutions and solver: for this we will remove equivalent expressions, which can be accomplished with quotients. We have already introduced and described countdown: in this subsection, we will fill in the remaining parts of the solver, glue the pieces together, and show how the finiteness proofs can assist us to write the solver.

[2] Answer: $3 \times (7 \times (50 - 10) - 25)$

Finite Vectors. We'll start with a simple example: for both the selection (Fig. 2a) and operators (Fig. 2c) subsection, all we need to show is that a vector of some finite type is itself finite. To describe which elements to keep from an n-element list, so instance, we only need a vector of Booleans of length n. Similarly, to pick n operators requires us only to provide a vector of n operators. And we can prove in a straightforward way that a vector of finite things is itself finite.

```
ℰ!⟨Vec⟩ : ℰ! A → ℰ! (Vec A n)
ℰ!⟨Vec⟩ {n = zero  } ℰ!⟨A⟩ = ℰ!⟨Poly⊤⟩
ℰ!⟨Vec⟩ {n = suc n } ℰ!⟨A⟩ = ℰ!⟨A⟩ |×| ℰ!⟨Vec⟩ ℰ!⟨A⟩
```

We've already shown that there are finitely many booleans, the fact that there are finitely many operators is similarly simple to prove:

```
ℰ!⟨Op⟩ : ℰ! Op
ℰ!⟨Op⟩ .fst = +′ ∷ ×′ ∷ -′ ∷ ÷′ ∷ []
ℰ!⟨Op⟩ .snd +′ = 0 , refl
ℰ!⟨Op⟩ .snd ×′ = 1 , refl
ℰ!⟨Op⟩ .snd -′ = 2 , refl
ℰ!⟨Op⟩ .snd ÷′ = 3 , refl
```

Finite Permutations. A more complex, and interesting, step of the transformation is the first step (Fig. 2b), where we need to specify the permutation to apply to the chosen numbers.

Our first attempt at representing permutations might look something like this:

```
Perm : ℕ → Type₀
Perm n = Fin n → Fin n
```

the idea is that Perm n represents a permutation of n things, as a function from positions to positions. Unfortunately such a simple answer won't work: there are no restrictions on the operation of the function, so it could (for instance), send more than one input position into the same output.

What we actually need is not just a function between positions, but an *isomorphism* between them. In types:

```
Perm : ℕ → Type₀
Perm n = Isomorphism (Fin n) (Fin n)
```

Where an isomorphism is defined as follows:

```
Isomorphism : Type a → Type b → Type (a ℓ⊔ b)
Isomorphism A B = Σ[ f ∶ (A → B) ] Σ[ g ∶ (B → A) ] (f ∘ g ≡ id) × (g ∘ f ≡ id)
```

Should this isomorphism definition be put earlier in the intro with the

While it may look complex, this term is actually composed of individual components we've already proven finite. First we have Fin n → Fin n: functions between finite types are, as we know, finite (Theorem 2.2). We take a pair of them: pairs of finite things are *also* finite (Lemma 2.1). To get the next two components we can filter to the subobject: this requires these predicates to be decidable. We will construct a term of the following type:

```
Dec (f ∘ g ≡ id)
```

So can we construct such a term? Yes!

We basically need to construct decidable equality for functions between Fin ns: of course, this decidable equality is provided by the fact that such functions are themselves finite.

All in all we can now prove that the isomorphism, and by extension the permutation, is finite:

```
iso-finite : ℬ A →
             ℬ B →
             ℬ (Σ[ f,g : (A → B) × (B → A) ]
                 ( (f,g .fst ∘ f,g .snd ≡ id) ×
                   (f,g .snd ∘ f,g .fst ≡ id)))
iso-finite ℬ⟨A⟩ ℬ⟨B⟩ =
  filter
    (λ _ → isPropEqs)
    (λ { (f , g) → (f ∘ g) ≟² B id && (g ∘ f) ≟² A id})
    ((ℬ⟨A⟩ |→| ℬ⟨B⟩) |×| (ℬ⟨B⟩ |→| ℬ⟨A⟩))
```

Unfortunately this implementation is too slow to be useful. As nice and declarative as it is, fundamentally it builds a list of all possible pairs of functions between Fin n and itself (an operation which takes in the neighbourhood of $O(n^n)$ time), and then tests each for equality (which is likely worse than $O(n^2)$ time). We will instead use a factoriadic encoding: this is a relatively simple encoding of permutations which will reduce our time to a blazing fast $O(n!)$. It is expressed in Agda as follows:

```
Perm : ℕ → Type₀
Perm zero    = ⊤
Perm (suc n) = Fin (suc n) × Perm n
```

It is a vector of positions, each represented with a Fin. Each position can only refer to the length of the tail of the list at that point: this prevents two input positions mapping to the same output point, which was the problem with the naive encoding we had previously. And it also has a relatively simple proof of finiteness:

```
ℰ!⟨Perm⟩ : ℰ! (Perm n)
ℰ!⟨Perm⟩ {n = zero  } = ℰ!⟨⊤⟩
ℰ!⟨Perm⟩ {n = suc n } = ℰ!⟨Fin⟩ |×| ℰ!⟨Perm⟩
```

Parenthesising. Our next step is figuring out a way to encode the parenthesisation of the expression (Fig. 2d). At this point of the transformation, we already have our numbers picked out, we have ordered them a certain way, and we have also selected the operators we're interested in. We have, in other words, the following:

$$3 \times 7 \times 50 - 10 - 25 \tag{31}$$

Without parentheses, however, (or a religious adherence to BOMDAS) this expression is still ambiguous.

$$3 \times ((7 \times (50 - 10)) - 25) = 765 \tag{32}$$

$$(((3 \times 7) \times 50) - 10) - 25 = 1015 \tag{33}$$

The different ways to parenthesise the expression result in different outputs of evaluation.

So what data type encapsulates the "different ways to parenthesise" a given expression? That's what we will figure out in this subsection, and what we will prove finite.

One way to approach the problem is with a binary tree. A binary tree with n leaves corresponds in a straightforward way to a parenthesisation of n numbers (Fig. 2d). This doesn't get us much closer to a finiteness proof, however: for that we will need to rely on *Dyck* words.

Definition 9 (Dyck words). A Dyck word is a string of balanced parentheses. In Agda, we can express it as the following:

```
data Dyck : ℕ → ℕ → Type₀ where
  done : Dyck zero zero
  ⟨_ : Dyck (suc n) m → Dyck n (suc m)
  ⟩_ : Dyck n m → Dyck (suc n) m
```

A fully balanced string of n pairs of parentheses has the type Dyck zero n. Here are some example strings:

```
_ : Dyck 0 2
_ = ⟨ ⟩ ⟨ ⟩ done
```

```
_ : Dyck 0 3
_ = ⟨ ⟩ ⟨ ⟨ ⟩ ⟩ done
```

The first parameter on the type represents the amount of unbalanced closing parens, for instance:

```
_ : Dyck 1 2
_ = ⟩ ⟨ ⟩ ⟨ ⟩ done
```

Already Dyck words look easier to prove finite than straight binary trees, but for that proof to be useful we'll have to relate Dyck words and binary trees somehow. As it happens, Dyck words of length $2n$ are isomorphic to binary trees with $n - 1$ leaves, but we only need to show this relation in one direction: from Dyck to binary tree. To demonstrate the algorithm we'll use a simple tree definition:

```
data Tree : Type₀ where
  leaf : Tree
  _*_ : Tree → Tree → Tree
```

The algorithm itself is quite similar to stack-based parsing algorithms.

```
dyck→tree : Dyck zero n → Tree
dyck→tree d = go d (leaf , _)
  where
  go : Dyck n m → Vec Tree (suc n) → Tree
  go (⟨ d)  ts           = go d (leaf , ts)
  go (⟩ d)  (t₁ , t₂ , ts) = go d (t₂ * t₁ , ts)
  go done (t , _)      = t
```

Putting It All Together. At this point we have each of the four components of the transformation defined. From this we can define what an expression is:

```
ExprTree : ℕ → Type₀
ExprTree zero    = ⊥
ExprTree (suc n) = Dyck 0 n × Vec Op n

Transformation : List ℕ → Type₀
Transformation ns =
  Σ[ s : Subseq (length ns) ]
    let n = count s
    in Perm n × ExprTree n
```

Notice that we don't allow expressions with no numbers.

The proof that this type is finite mirrors its definition closely:

```
ℰ!⟨ExprTree⟩ : ℰ! (ExprTree n)
ℰ!⟨ExprTree⟩ {n = zero } = ℰ!⟨⊥⟩
ℰ!⟨ExprTree⟩ {n = suc n} = ℰ!⟨Dyck⟩ |×| ℰ!⟨Vec⟩ ℰ!⟨Op⟩

ℰ!⟨Transformation⟩ : ℰ! (Transformation ns)
ℰ!⟨Transformation⟩ = ℰ!⟨Subseq⟩ |Σ| λ _ → ℰ!⟨Perm⟩ |×| ℰ!⟨ExprTree⟩
```

Filtering to Correct Expressions. We now have a way to construct, formally, every expression we can generate from a given list of numbers. This is incomplete in two ways, however. Firstly, some expressions are invalid: we should not, for instance, be able to divide two numbers which do not divide evenly. Secondly, we are only interested in those expressions which actually represent solutions: those which evaluate to the target, in other words. We can write a function which tells us if both of these things hold for a given expression like so:

```
eval : Tree Op ℕ → Maybe ℕ
eval (leaf x) = just x
eval (xs ⟨ op ⟩ ys) = do
  x ← eval xs
  y ← eval ys
  x !⟨ op ⟩! y
```

```
_!⟨_⟩!_ : ℕ → Op → ℕ → Maybe ℕ
x !⟨ +′ ⟩! y = just $! (x + y)
x !⟨ ×′ ⟩! y = just $! (x * y)
x !⟨ -′ ⟩! y =
  if x <ᴮ y
    then nothing
    else just $! (x − y)
x !⟨ ÷′ ⟩! zero = nothing
x !⟨ ÷′ ⟩! suc y =
  if rem x (suc y) ≡ᴮ 0
    then just $! (x ÷ suc y)
    else nothing
```

With this all together, we can finally write down the type of all solutions to a given countdown problem.

```
Solution ns n = Σ[ e : Transformation ns ] (eval (transform ns e) ≡ just n)
```

And, because the predicate here is decidable and a mere proposition, we can prove that there are finitely many solutions:

```
𝒞! (Solution ns n)
```

And we can apply this to a particular problem like so:

```
exampleSolutions : 𝒞! (Solution [ 1 , 3 , 7 , 10 , 25 , 50 ] 765)
exampleSolutions = 𝒞!⟨Solutions⟩
```

Typecheck this in Agda and it will evaluate to a list of the valid answers for that problem.

REFERENCES

Michael Abbott, Thorsten Altenkirch, and Neil Ghani. Containers: Constructing strictly positive types. *Theoretical Computer Science*, 342(1):3–27, September 2005. ISSN 0304-3975. doi: 10.1016/j.tcs.2005.06.002.

Guillaume Allais. Generic level polymorphic n-ary functions. In *Proceedings of the 4th ACM SIGPLAN International Workshop on Type-Driven Development - TyDe 2019*, pages 14–26, Berlin, Germany, 2019. ACM Press. ISBN 978-1-4503-6815-5. doi: 10.1145/3331554.3342604.

Errett Bishop. *Foundations of Constructive Analysis*. McGraw-Hill Series in Higher Mathematics. McGraw-Hill, New York, 1967.

H. J. Boom. Further thoughts on Abstracto. *Working Paper ELC-9, IFIP WG 2.1*, 1981.

Alexander Bunkenburg. The Boom Hierarchy. In John T. O'Donnell and Kevin Hammond, editors, *Functional Programming, Glasgow 1993*, Workshops in Computing, pages 1–8. Springer London, 1994. ISBN 978-3-540-19879-6 978-1-4471-3236-3. doi: 10.1007/978-1-4471-3236-3_1.

Nils Anders Danielsson. Bag Equivalence via a Proof-Relevant Membership Relation. In *Interactive Theorem Proving*, Lecture Notes in Computer Science, pages 149–165. Springer, Berlin, Heidelberg, August 2012. ISBN 978-3-642-32346-1 978-3-642-32347-8. doi: 10.1007/978-3-642-32347-8_11.

Martin Escardo. Infinite sets that admit fast exhaustive search. In *22nd Annual IEEE Symposium on Logic in Computer Science (LICS 2007)*, pages 443–452, Wroclaw, Poland, 2007. IEEE. ISBN 978-0-7695-2908-0. doi: 10.1109/LICS.2007.25.

Denis Firsov and Tarmo Uustalu. Dependently typed programming with finite sets. In *Proceedings of the 11th ACM SIGPLAN Workshop on Generic Programming - WGP 2015*, pages 33–44, Vancouver, BC, Canada, 2015. ACM Press. ISBN 978-1-4503-3810-3. doi: 10.1145/2808098.2808102.

Dan Frumin, Herman Geuvers, Léon Gondelman, and Niels van der Weide. Finite Sets in Homotopy Type Theory. In *Proceedings of the 7th ACM SIGPLAN International Conference on Certified Programs and Proofs*, CPP 2018, pages 201–214, New York, NY, USA, 2018. ACM. ISBN 978-1-4503-5586-5. doi: 10.1145/3167085.

Michael Hedberg. A coherence theorem for Martin-Löf's type theory. *Journal of Functional Programming*, 8(4):413–436, July 1998. ISSN 0956-7968, 1469-7653. doi: 10.1017/S0956796898003153.

Jason Z. S. Hu and Jacques Carette. Proof-relevant Category Theory in Agda. *arXiv:2005.07059 [cs]*, May 2020.

Graham Hutton. The Countdown Problem. *J. Funct. Program.*, 12(6):609–616, November 2002. ISSN 0956-7968. doi: 10.1017/S0956796801004300.

Frederik Hanghøj Iversen. Fredefox/cat, May 2018.

Nicolai Kraus. The General Universal Property of the Propositional Truncation. *arXiv:1411.2682 [math]*, page 35 pages, September 2015. doi: 10.4230/LIPIcs.TYPES.2014.111.

Egbert Rijke and Bas Spitters. Sets in homotopy type theory. *Mathematical Structures in Computer Science*, 25(5):1172–1202, June 2015. ISSN 0960-1295, 1469-8072. doi: 10.1017/S0960129514000553.

The Univalent Foundations Program. *Homotopy Type Theory: Univalent Foundations of Mathematics*. https://homotopytypetheory.org/book, Institute for Advanced Study, 2013.

Brent Abraham Yorgey. *Combinatorial Species and Labelled Structures*. PhD thesis, University of Pennsylvania, Pennsylvania, January 2014.