Polynomial Functors in Agda:
Theory and Practice

A formalization and collection of applications of categories of polynomial functors

Master's thesis in Computer science and engineering

Marcus Jörgensson

André Muricy Santos

Polynomial Functors in Agda:
Theory and Practice

A formalization and collection of applications of categories of polynomial functors

ANDRÉ MURICY SANTOS   MARCUS JÖRGENSSON

UNIVERSITY OF
GOTHENBURG

CHALMERS
UNIVERSITY OF TECHNOLOGY

Abstract

The category of polynomial functors - **Poly** - has been studied for over two decades, and is known for its relevance to computer science and type theory **containers https://www.cs.le.ac.uk/people/ma139/docs/thesis.pdf**, as well as its deep mathematical structure **koch https://arxiv.org/abs/0906.4931**. However, in recent years, researchers in the growing field of Applied Category Theory have found wider interpretations of the content of **Poly** and related topics, leading to applications in dynamical systems theory **jaz http://davidjaz.com/Papers/DynamicalBook.pdf**, information theory **https://arxiv.org/pdf/2201.12878.pdf**, database theory **spivak db https://** and more. Many aspects of the well-behavedness and richness of interpretation of **Poly** are being collected in the textbook *Polynomial Functors - A Mathematical Theory of Interaction* **poly-book**.

However, the surrounding code ecosystem of **Poly** is lacking. It is difficult to find straightforward formalizations and/or programs that use **Poly**'s structure to their advantage. The present work seeks to contribute to this ecosystem by providing formalizations in Cubical Agda of many existing theorems and categorical constructs in **Poly** and closely related categories, along with an array of concrete examples of its application to real world, widely studied dynamical systems. Most of the code follows the aforementioned textbook as a general blueprint, but since the very authors admit to the youth of applied category theory as a field and specifically its perspective on **Poly**, we sometimes deviate and show new applications.

0.1 Acknowledgements

Firstly, we would like to acknowledge and thank Naïm Camille Favier for the help via the Agda Zulip in several proofs: solving a problem with characterizing lens equality, making the inclusion map in set coequalizers type-check, and for general guidance in Agda-related questions.

Secondly, we would like to acknowledge David Orion Girardo for his succinct and beautiful formulation of the exponential object we used and for inspiration in Agda.

Thirdly, we thank David Spivak, co-author of the textbook we follow for this thesis and many of the papers we cite. We could not have done it without his and Nelson Niu's wonderful book. Or without his warm encouragement to pursue this project.

Finally, we thank our supervisor Felix Cherubini, whose patience, guidance, and expertise were instrumental in bringing this thesis to life.

Contents

Contents

1

Introduction

Category theory is an area of mathematics concerned with abstraction and mathematical structure. The category of polynomial functors, named **Poly**, has been studied for over two decades. **Poly**[1] is known for its relevance to computer science, as well as for its deep mathematical structure. In recent years, researchers in the emerging field of applied category theory have found interpretations of **Poly** in several fields of science and engineering, including dynamical systems theory. The purpose of this thesis is to contribute to the code ecosystem of **Poly** and its use in dynamical systems.

1.1 Applied category theory and complex systems

In the study of complex systems, there is a concern of finding the right level of abstraction at which to seek commonalities between different systems. Sometimes, seemingly unrelated systems might reveal themselves to have some deep structural similarity, like critical transitions in ecosystems [1]. However, arriving at a fundamental mathematical theory of complex systems is profoundly difficult. The immense expressive power of category theory has been a staple of type theory and programming language theory for many decades, and this robust mathematical language is now finding its footing in expressing ideas from wider scientific fields as well, including fields under the umbrella of complex systems. Some of the applications include foundations of complex systems [2], game theory [3], probability [4], systems biology [5] and dynamical systems theory[6].

Polynomial functors and dynamical systems

In particular, **Poly** turns out to be useful in the modeling of so-called mode-dependent open dynamical systems. The "open" part of this class of systems refers to the fact that they not only have an internal dynamics and state space, but can interact with their environment by accepting external inputs and providing external outputs. The "mode-dependent" part refers to the fact that this paradigm of interaction *can change* depending on the internal state of the system, meaning they can, throughout time, accept inputs and provide outputs to different environments.

---

[1]**Poly** is also known as the category of containers.

This is a remarkably general form of system which **Poly** fully embeds in the categorical context, which means questions like "what does it mean to take the product of two systems?" often have a concrete and intuitive meaning for the dynamics of the resulting system.

1.2   The *poly-book*

Many aspects of the behavior and richness of **Poly** are collected in a newly written (still in draft at the time of writing) textbook, *Polynomial Functors: A Mathematical Theory of Interaction*, hereafter called the poly-book. The book is being written by David Spivak and Nelson Niu and consists of a deep dive into **Poly**, explaining its mathematical structure in detail as well as providing many practical consequences of this structure to modeling dynamical systems. The authors also gave a course based on the book, and the lectures are freely available on YouTube. In this lecture series, some categories closely related to **Poly** are explored, which hint at an even bigger potential for application.

1.3   Goal of this thesis

We believe that the surrounding ecosystem of **Poly** is lacking. It is difficult to find straightforward formalizations of the category, or programs/dynamical systems models that use its structure to their advantage. The goal of this thesis is to advance this ecosystem. We follow the poly-book and its associated lecture series as a blueprint, and rely on the explanations therein to contribute several categorical formalizations, as well as show how **Poly** can be used to model different examples of dynamical systems by providing implementations. Some of these examples are described informally in the poly-book, others we come up with. We don't aim to formalize all theorems or implement all the examples in the book, as this would be simply too much work for this thesis. Instead, we constrain ourselves to what we deem the most essential constructs and illustrative examples.

1.4   Outline

The outline of the thesis is as follows; first, some of the needed background is explained in chapter 2. Then, the theoretical part of the thesis is given, formalizing two categories of polynomial functors and their categorical constructs. The primary category, of lenses, is given in chapter 3 and charts in chapter 4. Chapter 5 consists of the practice part of the thesis, how **Poly** is used for dynamical systems, together with several examples. After, in chapter 6, a discussion of the formalization is given, as well as future work. Finally, in chapter 7, conclusions are drawn about how well the result of the thesis met its purpose.

2

Background

This chapter gives a short introduction to the technology used, as well as the different areas of knowledge needed to understand this thesis. For the technology, **Agda** and **Cubical Agda** are explained, as they provide the framework for the implementation. For the theory, a brief explanation of **Category Theory** is given, as well as some references to previous work on the main category of this thesis, **Poly**. Actually introducing the concepts of **Poly** is done in the theory section.

Some topics are not introduced in the background, but instead at the point where they are used. There is also a wealth of citations and references to deeper explorations of most concepts mentioned.

2.1 Agda

Agda is a dependently typed programming language [7] created at Chalmers. It is a functional language, with very similar syntax to Haskell, but with a more powerful type system which allows it to be used as a proof assistant. Dependent types means that types can depend on values. For example, it is possible for a function to have the type signature `natOrBool : (b : Bool)` $\rightarrow$ `if b then` $\mathbb{N}$ `else Bool`, which means it returns a natural number if the argument is true, and a boolean if the argument is false. In a non-dependent language, this would need to be expressed at the value level, so the function would have to return a value of type `Either` $\mathbb{N}$ `Bool`.

2.1.1 Unicode characters

Agda, compared to most other languages, allows, and makes heavy use of unicode characters in symbol names. This can look daunting, but allows for a concise syntax more akin to mathematical notation.

2.1.2 Propositions as Types

The Curry-Howard isomorphism [8] is the idea that there is a one-to-one correspondence between types in programming languages, propositions in logic, and expressions in algebra. Under this correspondence (isomorphism), a value of a type can be seen as a proof of the proposition the type corresponds to [9]. For example, the following data type in Agda corresponds to the logical operator `or` (written $A \vee B$):

Algebra	Logic	Types
$A + B$	$A \vee B$	`Either A B`
$A \times B$	$A \wedge B$	`Pair A B`
B^A	$A \implies B$	`A -> B`
$A = B$	$A \iff B$	`Iso A B`

```
data Either (A B : Type) : Set where
    inj₁ : A → Either A B
    inj₂ : B → Either A B
```

It takes two types as arguments, and produces another type as an argument - much like the proposition $A \vee B$ is constructed out of the two propositions A and B. This in turn also corresponds to summing in algebra - $A + B$. A summary of the isomorphism is given below:

Concretely, an element of type `Either A B` is either an element of `A`, contained in the first value constructor inj_1, or of the type `B`, contained in the second value constructor inj_2. This corresponds to the fact that that a proof of the logical proposition $A \vee B$ is either a proof of the proposition `A` or a proof of `B`.

The Curry-Howard isomorphism together with its powerful type system makes Agda a powerful proof checker.

2.1.3 Agda examples

The `Either` example above shows how (polymorphic) datatypes are defined, which should be similar to the GADTs [10] extension of Haskell.

It is also possible to create record types, for data types with only one constructor.

```
record Action : Type where
  constructor mkAction
  field
    write : Alphabet
    move : Movement
```

The unit type and unit function is defined as:

```
data ⊤ : Type where
    tt : ⊤
```

```
unit : {A : Type} → A → ⊤
unit _ = tt
```

The bottom (or False) type, and the absurd function from bottom:

```
data ⊥ : Type where
```

```
absurd : {A : Type} → ⊥ → A
absurd ()
```

The natOrBool function can be implemented as:

```
natOrBool : (b : Bool) → if b then ℕ else Bool
natOrBool true = 0
natOrBool false = true
```

More explanations and examples of Agda can be found in [7] and [9]. Otherwise, it should be possible to understand the more advanced concepts of Agda along the thesis.

2.1.4 Foreign function interface - Haskell

Agda has a fully-featured foreign function interface (FFI) with Haskell, providing access to the vast Haskell ecosystem. This is important to our application section, as it allows us to use well-established libraries for plotting graphs, writing versatile command-line interfaces, and fast matrix multiplication. For a more detailed description of the supporting Haskell code and what it's used for, see A.

2.2 Cubical Agda

Cubical **cubicalAgdaDocs** is an extension to Agda bringing ideas from Homotopy Type Theory (HoTT) [11]. The main contribution of Cubical Agda is how to deal with equality.

Equality in normal Agda is defined as a datatype. However, this (propositional) equality has certain problems. For example, it is not possible to prove function extensionality, that two functions are equal if they are equal for all arguments. Functional extensionality, which is very useful, is directly provable in Cubical.

Cubical discards this datatype and uses its own notion of equality, the path. An equality between two elements $a : A$ and $b : A$ is a path from a to b. In Cubical, this is achieved by a function $p : I \rightarrow A$ satisfying some conditions. Firstly, I is a special type representing an interval or a path, ranging from 0 to 1. Secondly, the beginning of the path must be a, that is $p(0) = a$. Finally, the end of the path must be b, $p(1) = b$. The important thing to note is that given `a b : A` a type `a ≡ b` is really a function `I → A` behind the scenes, satisfying the above conditions.

The equality `a ≡ b` is actually known as homogeneous equality, because a and b are of the same type. It is also possible to have an equality when a and b are of different types, named heterogeneous equality. In fact, homogeneous equality is a special case of heterogeneous equality, where a and b happen to be the same type. For this thesis, mostly equality between same types is used, so no further explanation of heterogeneous equality is needed.

It is also possible to combine Cubical Agda with propositional equality. Using the functions `pathToEq` and `eqToPath` it is possible to go back and forth between the two different notions of equality. Something that can be useful in some cases.

There exists a standard cubical library [12] which has plenty of constructs and proofs from HoTT. The Cubical library is extensively used in this project.

A useful property of Cubical Agda is that isomorphisms are the same as equalities. It is possible to go back and forth using `isoToPath` and `pathToIso`, from the standard Cubical library. This means that one way to prove an equality is to prove an isomorphism.

A final property, which is used heftily in this thesis is substition. Substitution makes it possible to transfer a property over an equality. The declaration of substitution, somewhat simplified, is

```
subst : {A : Type} {x y : A} {B : A → Type} → (x ≡ y) → B x → B y
```

The next subsection gives some more examples of Cubical Agda. Further information about Cubical Agda and HoTT can be found in **cubicalPaper** and [11].

2.2.1 Cubical examples

The first example is to show that the absurd function, define before, is actually unique. Given any other function from the empty type, it is the same function as absurd. To show equality between functions, function extensionality is used. Finally, the lemma is easily proved by pattern matching on the empty type.

```
absurdUnique : {A : Type} → (f : ⊥ → A) → f ≡ absurd
absurdUnique f = funExt lemma
    where
        lemma : (x : ⊥) → f x ≡ absurd x
        lemma ()
```

In propositional equality, `refl` is the way to prove that any element a, is equal to itself. The function `refl` can be recovered in Cubical as follows. Remember that this equality is a function from I which must be a at both endpoints, which it is, since it is constantly a along the whole interval.

```
refl : {A : Type} → {a : A} → a ≡ a
refl {a = a} i = a
```

2.3 Category Theory

We make use of CT in this project in three different regards: we sometimes need to think about category theory in general and what is true in a pure context, sometimes how to apply categorical concepts to concrete scientific and engineering contexts, and sometimes on the syntactic issue of how to translate statements about CT to statements in Agda and its type theories. We will briefly expand on these three regards now.

2.3.1 In general: pure and applied CT

We think of Category Theory as the formal study of composition - meaning connecting two relationships to form a larger one. It has some axioms that always hold for its central object of study, which is categories themselves, and definitions that build on top of these axioms that allow a highly rigorous but simple form of abstract reasoning about either structures that exist within categories or between them.

The definitions and associated axioms are extremely simple; a *category* $\mathcal{C}$ consists of the following data:

1. A collection of objects $\text{Ob}(\mathcal{C})$ (we're careful to not say "*set* of objects", to avoid size issues).

2. For every two objects A and B, a set of *arrows* (also called *morphisms* or simply *maps*) between them denoted by $\mathcal{C}(A , B)$. Individual arrows, when they need to be named, are represented like this: $f : A \to B$. This set is also called the *hom-set*.

3. For every object A, an identity arrow $id : A \to A$. Note that this means that the hom-set $\mathcal{C}(A , A)$ is never empty, for any category: it must contain at least the identity arrow.

4. A composition operator $\circ$, that gives rise to new arrows. For any three objects A, B and C and arrows $f : A \to B$ and $g : B \to C$, there exists an arrow $g \circ f : A \to C$. We read this as "g after f".

And the following laws:

1. $\forall f. f \circ id = id \circ f = f$ (identity)

2. $\forall f\ g\ h. f \circ (g \circ h) = (f \circ g) \circ h$ (associativity of composition)

The usual way of representing statements in category theory is through *commuting diagrams*, like so:

Here, the objects are A, B, C and D, and the arrows are $f : A \to B$, $g : B \to C$, $h : A \to D$ and $k : D \to C$. Composite arrows and identity arrows are almost always omitted, but are always there implicitly. *Commuting* means any path between two objects is the same, which in this case means $k \circ h = g \circ f$. In a sense, diagrams can be seen as category theory's analogue of abstract algebraic equations. Like equations handle symbols that can stand for anything, diagrams are a totally abstract representation of some structure in a category, containing objects and morphisms that are labeled and subject to constraints like being equal or unique, and should be interpreted as a statement.

In general settings of category theory, the meaning of the objects and arrows doesn't play any role - they are purposefully abstract. Concrete categories ought to have meaning to their arrows and what it means to compose them, like the category **Set**, where objects are sets and arrows are functions between them. They must also respect the laws to be categories of course. In our case, because the thesis has a large focus on applied category theory, we will also talk frequently about the meaning of

the arrows and objects in **Poly** and some closely related categories; it's what makes them interesting after all.

Defining the main point of study of category theory is barely scratching the surface. There are many constructions and structures within the theory that we cannot explain in any detail that would do them justice here; instead we will introduce other concepts in CT as they become relevant.

We'll end this section by highlighting the fact that, although the definition of a category may not seem like much, it in a sense reveals *the point* of category theory: it is a theory of *compositionality*. What it means to compose relationships between objects, what facts are preserved under different compositions, how composing can generalize: this and other deep structural questions are at the core of this theory.

It's in this spirit, of investigating compositionality, that *applied* category theory is done. The main introductory textbook on the subject, *Seven Sketches in Compositionality*, [13] does a good job of justifying this approach. The book usually starts with motivating examples in the real world, and what CT and its constructions and formalisms can bring to understanding them in better, or at least different ways. In our case this order is reversed; we present the theory first, but the goal of applying it is always present.

2.3.2 In code: Agda's `agda-categories` library

The poly-book demonstrates a large amount of category-theoretic constructs in the context of **Poly**, but we limit ourselves to the ones that are most essential. These tend to be quite ubiquitous, and as such, there are formulations of many of them in the Agda ecosystem: the three main ones we considered were the 1lab [14], the Cubical library's own **cubical-cat**, and the `agda-categories` [15] library. We choose the third one for a few reasons: firstly, it is very well documented, with a sensible user interface and active community. Secondly, it has more constructs: for example, it is the only one of the three in which exponential objects, a key feature of **Poly**, is present. Thirdly, it is interoperable with Cubical, since the definition of a category requires a user to provide an equivalence relation to express equality between arrows, where we can supply cubical equality.

The brunt of the work that `agda-categories` is doing for us is providing a general scaffolding for us to validate that our formalizations of the concepts in **Poly** are correct. Sometimes it also helps by providing pre-existing implementations of categories; for instance, polynomials are functors, and so we will often rely on the provided instance of the category of **Set** of sets and functions.

2.4 Polynomial Functors

The study of polynomial functors has an extensive literature in the pure category-theoretic setting, like the work of Joachim Kock [16][17], but we follow the **poly-book** closely in this thesis. It is an applied category theory textbook, which means it consistently provides interpretations of abstract mathematical structures that

correspond, mapping them to concrete concepts and applications. The most mature interpretations, as far as we can tell, are to database theory, decision systems, and dynamical systems. One of the authors of the book, David Spivak, has done work using **Poly** in the context of database theory [18]. The book spends some time developing intuitions on the decision-making perspective as well, but we could not find much work in this area. Finally, there's the interpretation of **Poly** in dynamical systems, which the book provides many examples of.

3

Theory: Lenses

This chapter covers the theory part of the formalization, consisting mainly of proofs of different categorical properties of the category **Poly**. All code is available open source at github [19], which is recommended to follow at the same time as reading the thesis.

3.1 The category Poly

The category **Poly** refers, in our context, to the category of polynomial endofunctors in **Set**, with morphisms as natural transformations. *Polynomial* is used here because these functors can be written down as high-school algebraic expressions, like:

$$p(y) = y^2 + 4y + 10$$

only instead of handling numbers, these polynomial expressions have *sets* plugged into the variable y. Sum, multiplication and exponentiation and numeric literals are all considered in the context of sets: a numeric literal like 7 represents the set $\underline{7}$, a set with 7 elements. A $+$ symbol will correspond to *the categorical coproduct*, multiplication (often written with no $*$ symbol, as in $7y$ represents $7*y$) corresponds to *the categorical product* and exponentiation to the exponential object: in our context, it means functions from the exponent to the base of a term. For example, $f : 7^2 \cong f : 2 \to 7$. This way of thinking algebraically is very convenient for our purposes, because plugging in "numbers" into our polynomial expressions to get "numbers" back will turn out to be a very useful kind of operation.

3.1.1 Objects are polynomial functors

The objects of **Poly** are polynomial functors, which can be written as expressions like the one shown above. Since **Poly** handles sets, though, we can create much more unorthodox looking expressions. For instance, the below polynomial is perfectly coherent:

$$p(y) = \mathbb{N}y^2 + 1 \tag{3.1}$$

and we can think intuitively about what it does: it takes a set as input, and sends it to the set of: either infinitely long tuples of functions from $\underline{2}$ to this set, or to the only element of the singleton set. Like the poly-book, we refer to each summand of a power of y in our polynomials as a *position*, and the exponent of that power as a

direction at that position. This suggests that a polynomial has a set of positions, and for each element in this set, a set of directions. Therefore a natural way to represent these polynomials in Agda is as a dependent set:

```
record Polynomial : Set₁ where
    constructor mkpoly
    field
        position : Set
        direction : position → Set
```

The way to write the example $p(y) = \mathbb{N}y^2 + 1$ in Agda using this type is then

```
p = mkpoly (ℕ ⊎ ⊤)  λ {(inj₁ x) → Bool ; (inj₂ y) → ⊥}
```

We remark that we might sometimes refer to a polynomial as having "a single position" when it is a *monomial*: what this really means is that, in the definition of such a polynomial, the set of directions does not depend on the set of positions. For instance, the monomial

$$p(y) = \mathbb{R}y^{\{a,b,c,d\}}$$

does not have any dependence between *the set of positions* (the real numbers) and *the sets of directions* (the set $\{a, b, c, d\}$). The sets of directions would normally be a family of sets, where each position could give a different one, but in this case, they're always the same set. This way of phrasing positions might be unsettling at first, but it makes more sense in everyday programming with polynomial functors, since what really matters when handling dependent sets is what distinction an indexing set performs. If two positions index the same set, we can almost always get away with thinking of them as the same thing.

More perspectives on polynomials

Understanding polynomial functors is crucial, and there are many ways to think about them. The book showcases other standard ways to visualize polynomials, which are helpful both to interpretations of the category in different applications and to build intuition about the purely categorical perspective, so we'll introduce them now. The first is of *corolla forests*:

$$(3.2)$$

The term *corolla* refers to a tree that has a depth of 1, and *forest* refers to the fact that there are many such small trees. The set of positions is the set of roots of the forest (in this case $\underline{4}$) and the set of directions at each position is the set of arrows at each root (in this case $\underline{5}$, $\underline{2}$, $\underline{2}$ and $\underline{0}$). This then corresponds to $p(y) = y^5 + y^2 + y^2 + 1 \approx y^5 + 2y^3 + 1$. This perspective lends intuition to the decision-making interpretation of polynomials. Very loosely speaking, the set of directions is "directions in which one can go" given that one is at that set's associated position. Since polynomials can handle infinite sets, even uncountably infinite ones,

one can imagine such infinite corolla trees. For example, the polynomial 3.1 could be visualized as:

$$\bullet \quad \bigvee \bigvee \bigvee \quad \cdots \tag{3.3}$$

Another perspective on polynomials, emphasizing the container datatype view, is as Haskell algebraic data types. The polynomial above for instance can be written, in Haskell, as this:

```haskell
data P y = Fst y y y y y | Snd y y | Trd y y | Frth
  deriving Functor
```

And this type can trivially implement the `Functor` typeclass, so much so that it can be derived automatically as above (given that the language extension `DeriveFunctor` is enabled of course).

Since polynomials act on objects and on functions, we can define these applications as such:

```
_⦅_⦆ : Polynomial → Set → Set
_⦅_⦆ (mkpoly position direction) Y = Σ position λ x → (direction x → Y)
infixl 30 _⦅_⦆

applyFn : {A B : Set} → (p : Polynomial) → (A → B) → p ⦅ A ⦆ → p ⦅ B ⦆
applyFn (mkpoly position direction) f (fst , snd) = fst , λ x → f (snd x)
```

And we can use these definitions to quickly prove that **Poly**'s objects really are functors, we formulate of a record in `agda-categories` corresponding to an endofunctor in the library's provided instance of **Set**:

```
F-resp : {p : Polynomial} {A B : Set} {f g : A → B} {x : p ⦅ A ⦆ } →
    f ≡ g → applyFn p f x ≡ applyFn p g x
F-resp {x = posApp , dirApp} pr = λ i → posApp , (pr i) ∘ dirApp

conv : {A B : Set} {f g : A → B} → ({x : A} → f x Eq.≡ g x) → f ≡ g
conv p = funExt λ _ → eqToPath p

asEndo : (p : Polynomial) → Functor (Sets zero) (Sets zero)
asEndo p = record
    { F₀ = λ x → p ⦅ x ⦆
    ; F₁ = λ f → applyFn p f
    ; identity = Eq.refl
    ; homomorphism = Eq.refl
    ; F-resp-≈ = λ {_} {_} {f} {g} proof →
        pathToEq (F-resp {f = f} {g = g} (conv proof))
    }
```

note: The book acknowledges the unfortunate naming clash with the equivalent category of containers, where in that context, positions are called directions and

directions are called shapes, but we will stick to the poly-book terminology from here on out.

3.1.2 Arrows are natural transformations - *dependent lenses*

Arrows on **Poly** are so-called *dependent lenses*. The reason for the naming is that, for arrows between monomials, we recover the usual definition of lenses in Haskell and other languages, as a pair of non-dependent "getter and setter" functions. A dependent lens has the additional property that "the types we are allowed to set depend on the value we get". We will explore this notion in depth, and in much more intuitive and concrete contexts, in the applications part of the thesis. Since non-dependent lenses (lenses between monomials) are a very special case of dependent lenses, we will refer to the arrows in **Poly** as simply *lenses* from here on out, like the Agda definition.

Perspectives on lenses

Another use of the corolla forest perspective is that it makes visualizing maps between polynomials quite intuitive. Consider the polynomials $p(y) = y^5 + 2y^2 + y$ and $q(y) = y^3 + y^2 + 2y + 1$. First, their visualization as corollas is as such:

Now a map between them can be visualized in the following way. Each position in p gets sent to a position in q, and for each of these sendings, its directions are sent back to the directions at the original position:
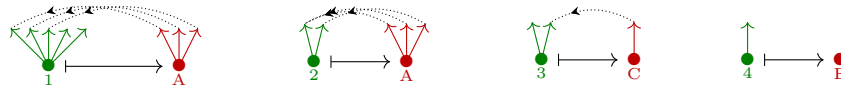
In the Haskell view of polynomials, lenses can be given as the following type synonym (requires `RankNTypes` language extension):

```haskell
type Lens f g = forall a. f a -> g a
```

An example of a lens between the two polynomials above would look like the following (requires `LambdaCase` language extension):

```haskell
data P y = Fst y y y y y | Snd y y | Trd y y | Frth
  deriving Functor

data Q y = A y y y | B y y | C y | D y | E
  deriving Functor

exampleLens :: Lens P Q
exampleLens = \case
  Fst y1 y2 y3 y4 y5 -> _
  Snd y1 y2 -> _
  Trd y1 y2 -> _
  Frth -> _
```

Here we can see that the map on positions translates to a map on value constructors; for each of the summands of p, first a value constructor in q must be chosen. Once we fill that in we get:

```haskell
exampleLens :: Lens P Q
exampleLens = \case
  Fst y1 y2 y3 y4 y5 -> A y1 y2 y3
  Snd y1 y2 -> A y1 y1 y2
  Trd y1 y2 -> C y2
  Frth -> E
```

And now we recover the idea of a function "back", that is a map on directions *at that position*: for each of the targeted value constructors in q, a lens must fill its slots with the source constructor's y values. For example, in this case, the three slots of the A constructor in the first branch must be filled by the five available choices in the source constructor Fst. In an abstract sense, we can see this as a function of type $\underline{3} \to \underline{5}$ - a choice of source out of 5 possible sources for each of the 3 slots.

Expressing more general kinds of lenses in Haskell using this scheme becomes hard quickly however. The dependency between the value of the map on positions and the type of the map of directions is hidden and baked into the syntax of the language. There are attempts to do this in Haskell [20] [21], but we only give this Haskell perspective for the sake of intuition. Polynomials and lenses are much more natural in Agda.

Lenses in Agda

Here is their definition:

```agda
record Lens (from to : Polynomial) : Set where
    constructor _⇆_
    open Polynomial
    field
        mapPosition : position from → position to
        mapDirection : (fromPos : position from) →
```

```
                        direction to (mapPosition fromPos) →
                        direction from fromPos
```

The constructor _⇆_ is an attempt at representing the two maps: on positions below and going forward (because they are the coefficients of the polynomials) and on directions on top going backwards (because they are exponents of y at each position).

As they are maps between functors, one can expect that lenses are natural transformations. Indeed, they are, and here's how to formulate them as such in the context of `agda-categories`:

```
asNatTransLens : {p q : Polynomial} →
          Lens p q →
          NaturalTransformation (asEndo p) (asEndo q)
asNatTransLens (f ⇆ f♯) = record {
    η = λ { X (posP , dirP) → f posP , dirP ∘ f♯ posP } ;
    commute = λ f₁ → Eq.refl ;
    sym-commute = λ f₁ → Eq.refl
    }
```

Now in order to have something that we can even test forms a valid category, we need two more pieces of data: the identity arrow, and a description of how arrows compose. Here are these definitions in Agda.

3.1.2.1 The identity lens

It is what one would expect: A polynomial remains untouched in both its set of positions and the direction at each position, by having the `fromPos` argument ignored by the map on directions:

```
idLens : {A : Polynomial} → Lens A A
idLens = id ⇆ λ _ → id
```

3.1.2.2 Composing lenses

Again it is done in the "obvious" way:

```
_∘ₚ_ : {A B C : Polynomial} → Lens B C → Lens A B → Lens A C
_∘ₚ_ (f ⇆ f♯) (g ⇆ g♯) = (f ∘ g) ⇆ (λ i → g♯ i ∘ f♯ (g i))
```

The map on positions of the resulting lens is straightforward function composition of the maps on positions of the composing lenses, and the map on directions is also a composition with the care that maps on directions are dependent.

3.1.3 Categorical axioms

That the associativity and identity laws hold for **Poly** follows directly from the definitions of these types, which means that they can be proved via Cubical Agda's `refl`. We take this chance to show just how **Poly** is represented as a category in `agda-categories`' formulation:

```
-- Categorical/Instance/Poly.agda
○-resp-≈ : {A B C : Polynomial}
           {f h : Lens B C}
           {g i : Lens A B} →
           f ≡ h → g ≡ i → (f ○ₚ g) ≡ (h ○ₚ i)
○-resp-≈  p q ii = (p ii) ○ₚ (q ii)

Poly : Category (lsuc lzero) lzero lzero
Poly = record
    { Obj = Polynomial
    ; _⇒_ = Lens
    ; _≈_ = _≡_
    ; id = idLens
    ; _○_ = _○ₚ_
    ; assoc = refl
    ; sym-assoc = refl
    ; identityˡ = refl
    ; identityʳ = refl
    ; identity² = refl
    ; equiv = record { refl = refl ; sym = sym ; trans = _·_ }
    ; ○-resp-≈ = ○-resp-≈
    }
```

This record definition should not be surprising: **Poly**'s objects are `Polynomial`s, its arrows are dependent `Lens`es, our notion of equality between arrows is the cubical equality `_≡_`, the identity lens and composition are the previously defined values in Agda. The rest of the data provided to the record correspond to the categorical laws plus some convenience that `agda-categories` provides: for instance, the proof `sym-assoc` is not normally needed, but it makes it so that the opposite of opposite categories are equal "on the nose" to the original categories, and not just provably equal. We will not spend time going over design decisions of `agda-categories`, but for a more detailed account of the reasoning behind the library, see the authors' paper [15] on it.

3.2   Polynomial equality

There are some occasions where polynomials need to be compared for equality. For example, a convenient way of characterizing monoidal structures in **Poly** is by simply expressing the laws on types, which requires polynomials to be compared; the right unit law for instance says can be written as

```
rightUnit : (p q : Polynomial) → p ⊗ Y ≡ p
```

(in the case of monoidal structures, we stick to the `agda-categories` framework and don't rely on equality of objects at all, but polynomial equality is both useful elsewhere and an illuminating exercise).

Since a polynomial is a record consisting of the `position : Set` and

`direction : position → `**`Set`** a characterization of equality for this record is needed, to make it easy to use and prove equalities between polynomials. It is very directly expressible as a Σ-type, on which there are already many properties and lemmas in the Cubical library. Here's that representation, along with proofs that it's equal to the record one:

```
PolyAsSigma : Set₁
PolyAsSigma = Σ[ position ∈ Set ] (position → Set)

polyToSigma : Polynomial → PolyAsSigma
polyToSigma (mkpoly position direction) = position , direction

polyFromSigma : PolyAsSigma → Polynomial
polyFromSigma (position , direction) = mkpoly position direction

poly≡polySigma : Polynomial ≡ PolyAsSigma
poly≡polySigma = isoToPath (iso polyToSigma
                                polyFromSigma
                                (λ _ → refl)
                                (λ _ → refl))
```

add code we actually use for poly equality

3.3   Lens equality

In category theory, we are usually interested instead in the algebra of morphisms, that is, we want to know when morphisms are equal; very often we want to ascertain that a composite of two arrows is equal to some induced morphism, for instance. We often want to prove that a morphism is unique. For this we need a good way of characterizing equality of lenses, and this characterization of equality needs to be convenient enough to be usable in practical proofs about **Poly**.

A lens consists of two components, with the second component depending on the first. This suggests a Σ-type structure, which inclines us to use the same strategy as with polynomial equality, but in lenses things are slightly more complicated, since the second field depends on the first field *applied* to something. Still, representing lenses as Σ-types is relatively straightforward:

```
LensAsSigma : Polynomial → Polynomial → Type
LensAsSigma (mkpoly posP dirP) (mkpoly posQ dirQ)
    = Σ[ mapPos ∈ (posP → posQ) ]
    ((fromPos : posP) → dirQ (mapPos fromPos) → dirP fromPos)

sigmaToLens : LensAsSigma p q → Lens p q
sigmaToLens (mapPos , mapDir) = mapPos ⇆ mapDir

lensToSigma : Lens p q → LensAsSigma p q
lensToSigma  (mapPos ⇆ mapDir) = mapPos , mapDir
```

```
lens≡lensSigma : (Lens p q) ≡ (LensAsSigma p q)
lens≡lensSigma = isoToPath (iso lensToSigma
                                sigmaToLens
                                (λ _ → refl)
                                (λ _ → refl))
```

As for the equality itself, we must have some dependency expressed in the types of the two arrows provided. This is unlike polynomial equality, since polynomials require only equality of types, whereas lenses requires both equality of types and of values.

If lenses and polynomials are both Σ-types in disguise, why not just use Σ-types?

The main reason is *usability.* This is a convenience-driven decision: we want our implementation of **Poly** to not only be amenable to formalization and having theorems proven about it, but also to be *nice to program in.* Working with Σ-types causes readability to suffer too much, because the names of the constructor and fields are lost. We then make the compromise of converting between the record Σ-type representations only in the equality modules, since the proofs that the representations are equal allows use to carry equality over between the types, a very useful feature of Cubical Agda.

3.4   Initial object

A category has an initial object if there exists an (initial) object that has an unique arrow from every other object. In **Poly**, the initial object is the polynomial $p(y) = 0$ with no positions (and therefore also no directions). It is the functor that sends any set to the empty set.

```
𝟘 : Polynomial
𝟘 = mkpoly ⊥ λ ()
```

The steps to prove that 0 is an initial object is to firstly construct a lens to every other object, and to secondly show that this each of these lenses are unique.

Since 0 has no positions (nor directions) both the map on positions and the map on directions to any other polynomial p is given by the function $λ$ () (often named absurd) from the empty type.

```
lensFromZero : {p : Polynomial} → Lens 𝟘 p
lensFromZero = (λ ()) ⇆ (λ ())
```

The final step is to show that this lensFromZero from 0 to p is unique. This is done by assuming that there exists another lens f from 0 to p and then prove that f actually is the same lens as lensFromZero.

```
lensFromZeroUnique : {p : Polynomial} (f : Lens 𝟘 p) → lensFromZero ≡ f
lensFromZeroUnique _ = lens≡ (funExt λ ()) (funExt λ ())
```

This is implemented by using the lens equality function and the fact that the absurd function (from the empty type) is unique utilizing function extensionality. We plug these proofs in to the `agda-categories` characterization of the initial object:

```
open import Categories.Object.Initial Poly

zeroIsInitial : IsInitial 𝟘
zeroIsInitial = record { ! = lensFromZero ; !-unique = lensFromZeroUnique }

initialZero : Initial
initialZero = record { ⊥ = 𝟘 ; ⊥-is-initial = zeroIsInitial }
```

3.5   Terminal object

Dually to the initial object, the terminal object is an object such that there is an arrow to every other object in the category. In **Poly**, the terminal object is defined as the polynomial with a single position, but no directions for that position. It is the functor that sends all sets to the singleton set.

```
𝟙 : Polynomial
𝟙 = mkpoly ⊤ (λ _ → ⊥)
```

Similarly to the initial object, to show that 1 is a terminal object is to first construct a lens **from** every other object, and to show that each of these lenses is unique.

The map on positions from any other polynomial to 1 is the function always returning unit (often also named unit). For this position, the map on directions goes from ⊥ since 1 has no directions, thus the map on directions is again the absurd function.

```
lensToOne : {p : Polynomial} → Lens p 𝟙
lensToOne = (λ _ → tt) ⇆ λ _ ()
```

Finally, for uniqueness, any other lens f from p to 1 is the same as lensToOne. For map on positions equality, refl is used to show that any two functions (with same domain) to ⊤ is the same. For map on directions equality, uniqueness of the absurd function is used, as in the initial object. The more advanced version of lens equality is used to do the function extensionality in the background, allowing for more concise syntax.

```
lensToOneUnique : {p : Polynomial} (f : Lens p 𝟙) →  lensToOne ≡ f
lensToOneUnique _ = lens≡∀∀ refl (λ _ ())
```

And again we plug these proofs in to the `agda-categories` characterization of the initial object:

```
open import Categories.Object.Terminal Poly

oneIsTerminal : IsTerminal
```

```
oneIsTerminal = record { ! = lensToOne ; !-unique = lensToOneUnique }

terminalOne : Terminal
terminalOne = record { ⊤ = 𝟙 ; ⊤-is-terminal = oneIsTerminal }
```

3.6   Product

The product of two objects A and B is an object $A \times B$ together with arrows π_1 and π_2 such that, given any other object T and arrows from T to A and B, the following diagram commutes and the arrow ! exists and is unique:
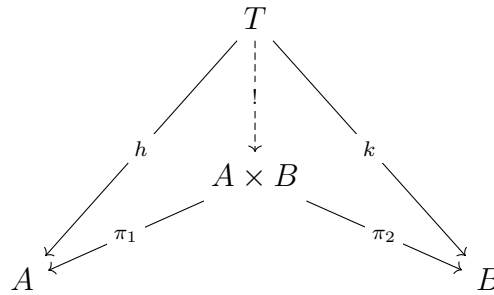
The arbitrary object T can be intuitively thought of as a "candidate" object to be the product. This is an example of a *universal property* in category theory and the arrows h and k are said to be *factorized* through the compositions $\pi_1 \circ$! and $\pi_2 \circ$! respectively.

Poly has all products, which means that for any two polynomials p and q we can construct their product as the polynomial $p * q$.

Firstly, the product polynomial of two polynomials is defined. The position set is the product of p's and q's position sets, while the direction at a position is either a direction of p or a direction of q. This is another instance of the fact that these polynomials behave just like the ones in high school algebra: for example, given $p(y) = y^2$ and $q(y) = 5y^3 + 2y$, their product is given by $p * q(y) = 5y^5 + 2y^3$. In Agda, the product is given as:

```
_*_ : Polynomial → Polynomial → Polynomial
(mkpoly posP dirP) * (mkpoly posQ dirQ) =
    mkpoly (posP × posQ) (λ {(posP , posQ) → (dirP posP) ⊎ (dirQ posQ)})
```

A simple example of the product is provided in figure 3.1.

Continuing with figure **??**, it is natural to ask how π_1 and π_2 are defined. They are defined to take the projections on the positions, and injections on the map on the directions back. Figure 3.1 makes this clear with the black dotted line projecting one of the positions $p * q$ to p and then injecting the direction back to $p * q$.

```
π₁ : {p q : Polynomial} → Lens (p * q) p
π₁ = proj₁ ⇆ λ _ → inj₁
```
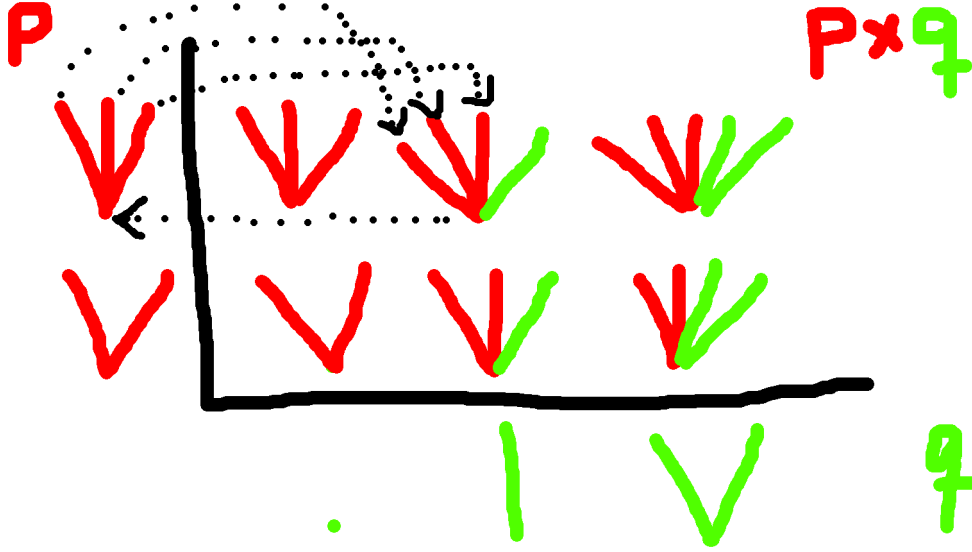
Figure 3.1: The polynomial p has 2 positions, and q has 3 positions. Their product $p * q$ has both p and q's positions, totaling $2 * 3 = 6$ positions. The directions at each position is the disjoint union of p' and q's directions at that position.

```
π₂ : {p q : Polynomial} → Lens (p * q) q
π₂ = proj₂ ⇆ λ _ → inj₂
```

Given an object c as the domain of two lenses f and g, $\langle f, g \rangle$ should be the factorization lens of f and g. This factorization must be unique, which is proved later as the last step. The factorization lens is defined by taking the pair of functions for map on positions. For the map on directions either f or g is used depending on the direction.

```
⟨_,_⟩ : {p q r : Polynomial} → Lens p q → Lens p r → Lens p (q * r)
⟨ f ⇆ f♯ , g ⇆ g♯ ⟩ = < f , g > ⇆ λ posP → [ f♯ posP , g♯ posP ]
```

Finally, we need to supply the proofs associated with the product. There are two proofs needed to show that the diagram **??** commutes, and one for showing the uniqueness of $\langle f, g \rangle$. The proofs for commuting is shown with refl, but the uniqueness is left out for space reasons - we will do this relatively frequently in the rest of the thesis, since many of these proofs get quite long. Interested readers can of course see the code for the formalization**code**. But the idea is that if we have another factorization function h it must be the same as $\langle f, g \rangle$.

```
project₁ : {p q r : Polynomial} {f : Lens r p} {g : Lens r q} →
           π₁ ∘ₚ ⟨ f , g ⟩ ≡ f
project₁ = refl

project₂ : {p q r : Polynomial} {f : Lens r p} {g : Lens r q} →
           π₂ ∘ₚ ⟨ f , g ⟩ ≡ g
```

```
project₂ = refl

unique : {p q r : Polynomial} {h : Lens p (q * r)} {f : Lens p q}
    {g : Lens p r → (π₁ ∘ₚ h) ≡ f} →
    (π₂ ∘ₚ h) ≡ g →
    ⟨ f , g ⟩ ≡ h
unique = ...
    where
        lemma : ⟨ π₁ ∘ₚ h , π₂ ∘ₚ h ⟩ ≡ h
        lemma = ...
```

3.6.1   Productized lens

Two lenses between different pairs of polynomials can be represented as a single lens between two products, as follows.

```
⟨_×_⟩ : {a b c d : Polynomial} → (f : Lens a c) (g : Lens b d)
    → Lens (a * b) (c * d)
⟨ (f ⇆ f♯) × (g ⇆ g♯) ⟩
    = (λ {(a , b) → f a , g b}) ⇆
        λ {(a , b) (inj₁ dirC) → inj₁ (f♯ a dirC)
         ; (a , b) (inj₂ dirD) → inj₂ (g♯ b dirD)}
```

3.6.2   Generalized product

Similarly to how the pi type is a generalization of the product type in **Type**, we can create the product of all polynomials indexed by any type. Think of it as a "fold", in the sense that each element of a set produces a new polynomial that is productized with the "accumulator" of this fold (all the products taken so far). This construction will be particularly important for the exponential object.

```
ΠPoly : Σ[ indexType ∈ Set ] (indexType → Polynomial) → Polynomial
ΠPoly (indexType , generatePoly) = mkpoly pos dir
  where
    -- Embedding all polynomial positions into one position
    pos : Set
    pos = (index : indexType) → position (generatePoly index)

    -- Direction is exactly one of the polynomials' directions
    dir : pos → Set
    dir pos = Σ[ index ∈ indexType ] direction (generatePoly index) (pos index)
```

Using this generalized product type with *Bool* as the index type is equivalent to having a normal product - which makes sense, since it means taking a product of two polynomials. This is proved in `Various.agda`.

```
productIsΠPoly : {p q : Polynomial}
    → ΠPoly (Bool , tupleToFunFromBool (p , q)) ≡ (p * q)
```

```
productIsΠPoly = ...

tupleToFunFromBool : {ℓ : Level} {A : Set ℓ} → (A × A) → Bool → A
tupleToFunFromBool (a , b) true = a
tupleToFunFromBool (a , b) false = b
```

With the generalized product it is also useful to have the factorizer $\langle f, g \rangle$ generalized, now factorizing a generalized amount of lenses instead of just f and g.

```
-- A function A→(B*C) is the same as two functions A→B and A→C
universalPropertyProduct : {p : Polynomial} {Index : Type}
    {generate : Index → Polynomial}
    → Lens p (ΠPoly (Index , generate)) ≡ ((i : Index) → Lens p (generate i))
universalPropertyProduct = ...
```

3.6.3 Monoid

The product forms a monoid with 1, which we show by constructing a monoidal category (**Poly**, ×, 1). Any In fact, the product is a *symmetric* monoidal structure, since $A \times B$ is isomorphic to $B \times A$. This is one place where the existing structure of `agda-categories` is extremely useful, as we only need to call some of its helpers to prove this and rely on the fact that all cartesian categories form monoidal categories with respect to the product. The code for it is:

```
-- Categorical/Poly/Monoidal/Product.agda

open import Categories.Category.Monoidal
import Categories.Category.Cartesian as Cartesian

binaryProducts : Cartesian.BinaryProducts Poly
binaryProducts = record { product = prod }

cartesian : Cartesian.Cartesian Poly
cartesian = record { terminal = terminalOne ;
                     products = binaryProducts }

productMonoidal : Monoidal Poly
productMonoidal
    = Cartesian.CartesianMonoidal.monoidal Poly cartesian

open import Categories.Category.Monoidal.Symmetric productMonoidal
productSymmetricMonoidal : Symmetric
productSymmetricMonoidal
    = Cartesian.CartesianSymmetricMonoidal.symmetric Poly cartesian
```

3.7 Coproduct

The coproduct is the dual construction to the product, obtained from reversing all arrows. If we do this to the product diagram shown before, this is obtained:
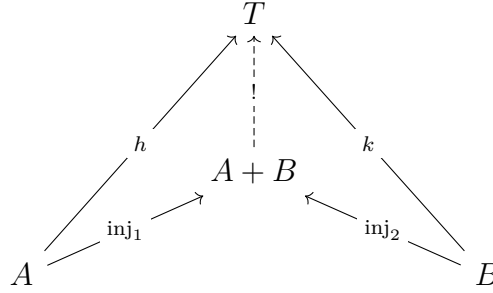
Poly also has all coproducts. The coproduct $p+q$ of any p and q is defined by taking the coproduct on just the positions and the *projection* on the directions. Again, consider the analogy with high-school algebra. Given two polynomials $p(y) = 4y^2$ and $q(y) = y^2 + 3y$, their sum is given by $(p+q)(y) = 5y^2 + 3y$. In this example, four of the positions with 2 as a direction came from p and one of them came from q, hence we pattern match on the resulting polynomials's position. As a shorthand, we use Agda's [_,_] function.

```
_+_  : Polynomial → Polynomial → Polynomial
(mkpoly posA dirA) + (mkpoly posB dirB) = mkpoly (posA ⊎ posB) [ dirA , dirB ]
```

The example of coproduct is provided in figure 3.2.

> reverse the black dotted arrow

The injections i_1 and i_2 are defined by taking the injections on the positions and keeping the directions on that position using id. The example **??** shows how i_1 behave for one position.

```
i₁ : {p q : Polynomial} → Lens p (p + q)
i₁ = inj₁ ⇆ λ _ → id

i₂ : {p q : Polynomial} → Lens q (p + q)
i₂ = inj₂ ⇆ λ _ → id
```

Given an object c as the codomain of two functions f and g, $[f, g]$ should be the (unique) factorization lens of f and g. This function is defined by using f or g' behavior depending what the position corresponds to.

```
[_,_]ₚ : {p q r : Polynomial} → Lens p r → Lens q r → Lens (p + q) r
[ f ⇆ f♯ , g ⇆ g♯ ]ₚ = [ f , g ] ⇆ [ f♯ , g♯ ]
```

The proofs needed correspond to the ones from the product.

```
inject₁ : {f : Lens p a} {g : Lens q a} → [ f , g ]ₚ ∘ₚ i₁ ≡ f
inject₁ = refl

inject₂ : {f : Lens p a} {g : Lens q a} → [ f , g ]ₚ ∘ₚ i₂ ≡ g
```
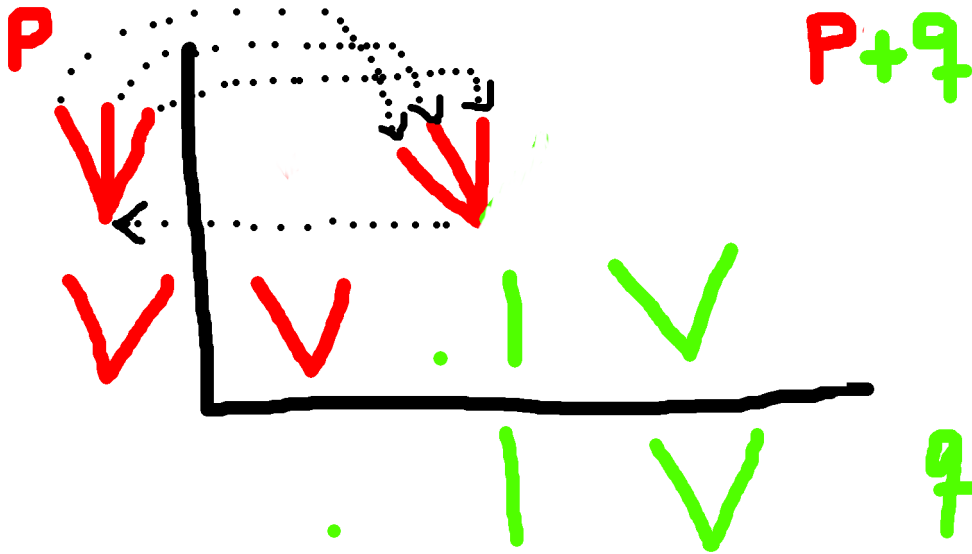
Figure 3.2: The polynomial p has 2 positions, and q has 3 positions. Their coproduct $p+q$ has the union of p and q's positions, totaling $2+3=5$ positions. The directions at each position is the corresponding direction at that position.

```
inject₂ = refl

unique : {F : Polynomial} {h : Lens (A + B) F} {f₁ : Lens A F} {f₂ : Lens B F}
    → (h ∘ₚ i₁) ≡ f₁
    → (h ∘ₚ i₂) ≡ f₂
    → [ f₁ , f₂ ]ₚ ≡ h
unique p₁ p₂ = ...
    where
        lemma : [ h ∘ₚ i₁ , h ∘ₚ i₂ ]ₚ ≡ h
        lemma = ...
```

3.7.1 Generalized coproduct

The coproduct can also be generalized to work with any amount of polynomials indexed by an index-type. This corresponds to the sigma type, or exists operator.

```
ΣPoly : Σ[ indexType ∈ Set ] (indexType → Polynomial) → Polynomial
ΣPoly (indexType , generatePoly) = mkpoly pos dir
  where
    -- It is the positions of one of the polynomials
    pos : Set
    pos = Σ[ index ∈ indexType ] (position (generatePoly index))
```

```
    -- It is the direction of the polynomial for the position
    dir : pos → Set
    dir (index , positionAtIndex) = direction (generatePoly index)
                                              positionAtIndex
```

Using the generalized coproduct with bool as index type is the same as the normal coproduct. This is proved in Various.agda.

```
coproductIsΣPoly : {p q : Polynomial} → ΣPoly (Bool , tupleToFunFromBool (p , q))
coproductIsΣPoly = ...
```

The generalized factorizer for the generalized coproduct can now be defined.

```
-- A function (A+B)→C is the same as two functions A→C and B→C
universalPropertyCoproduct : {Index : Type} {generate : Index → Polynomial}
    → Lens (ΣPoly (Index , generate)) p ≡ ((i : Index) → Lens (generate i) p)
universalPropertyCoproduct = ...
```

3.7.2 Monoid

To show the coproduct is a monoid, we do almost the same as for the product. Again we rely on `agda-categories`' helpers to define the monoidal category (**Poly**, +, 0) - which is also *symmetric* monoidal. This time we rely on the fact that the category is *co*-cartesian, but the code is almost identical:

```
-- Categorical/Poly/Monoidal/Coproduct.agda

open import Categories.Category.Monoidal
import Categories.Category.Cocartesian as Cocartesian

binaryCoproducts : Cocartesian.BinaryCoproducts Poly
binaryCoproducts = record { coproduct = coprod }

coproductCocartesian  : Cocartesian.Cocartesian Poly
coproductCocartesian = record { initial = initialZero ;
                                coproducts = binaryCoproducts }

coproductMonoidal : Monoidal Poly
coproductMonoidal =
  Cocartesian.CocartesianMonoidal.+-monoidal Poly coproductCocartesian

open import Categories.Category.Monoidal.Symmetric coproductMonoidal
productSymmetricMonoidal : Symmetric
productSymmetricMonoidal =
  Cocartesian.CocartesianSymmetricMonoidal.+-symmetric Poly coproductCocartesian
```

3.7.3 Coproductized lens

It is also easy to define a parallel coproduct lens. Like for the product, two lenses can be made to represent a single lens between coproducts.

```
⟨_⊎_⟩ : {p q r w : Polynomial} → (f : Lens p r) (g : Lens q w)
        → Lens (p + q) (r + w)
⟨_⊎_⟩ {p} {q} {r} {w} (f ⇆ f♯) (g ⇆ g♯) = mp ⇆ md
    where mp : position (p + q) → position (r + w)
          mp = map f g
          md : (fromPos : position (p + q)) →
               direction (r + w) (mp fromPos) →
               direction (p + q) fromPos
          md (inj₁ x) d = f♯ x d
          md (inj₂ y) d = g♯ y d
infixl 30 ⟨_⊎_⟩
```

3.8 Parallel product

The parallel product (also called Dirichlet product) is another useful binary operator on polynomials. It is defined as the pair of positions and pair of directions. This is the first operator that does not have a direct high school algebra correspondence, as it multiplies on both coefficients and exponents: given $p(y) = 2y^4 + 3y2$ and $q(y) = 2y^3$, $(p \otimes q)(y) = 4y^12 + 6y^6$.

```
_⊗_ : Polynomial → Polynomial → Polynomial
(mkpoly posA dirA) ⊗ (mkpoly posB dirB)
    = mkpoly (posA × posB) (λ (posA , posB) → (dirA posA) × (dirB posB))
```

The parallel product will turn out to be extremely important in the context of dynamical systems. More on this in the applications chapter.

3.8.1 Monoid

The parallel product forms a monoid with y. Since it is not the canonical product (or coproduct), as far as we know, we can't rely on the previous tricks to construct a monoidal category out of it. However, we can define the monoidal category construction manually, which is a lot of code in "volume". This is done, for consistency in `Categorical/Poly/Monoidal/ParallelProduct.agda`, but an alternative proof of the monoidality and symmetry of $\otimes$ is offered here, by using polynomial equality.

3.9 Composition product

Since polynomials are functors, and they are in fact *endo*functors, they can be composed to give rise to new polynomial functors. The composition operator is given by the symbol ◁. Here we can recover high-school algebra intuition. Composing functors corresponds to plugging in entire expressions, instead of numbers, into a

polynomial. For instance, given $p(y) = y^2$ and $q(y) = y^2 + 3y$, the composition product is given as $(p \lhd q)(y) = (y^2 + 3y)^2$.

```
-- Proposition 5.2, page 158. Note: not same definition used. We here treat posit
-- as inhabitants of the same set, which makes a lot of proofs easier down the li
_◁_ : Polynomial → Polynomial → Polynomial
p ◁ q = mkpoly pos dir
  where
    module p = Polynomial p
    module q = Polynomial q

    pos : Set
    pos = (Σ[ i ∈ p.position ] (p.direction i → q.position))

    dir : pos → Set
    dir (i , j) = Σ[ a ∈ p.direction i ] q.direction (j a)
infixl 27 _◁_
```

This definition is heavily inspired by David Orion Girardo's implementation here **daig**. The key insight from his implementation is that the positions in the resulting polynomial are a pair, the first element of which is the same set as the postcomposed polynomial. This sticks closer to the idea that polynomials are in fact being composed. In terms of corolla forests, the composition product stacks trees on top of one another:

> Give an example as a drawing.

3.9.1 Monoid

The composition product is an important monoidal structure in **Poly**. It is not symmetric, like the previous three structures, but it has massive consequences from both a theoretical as well as application point of view.

3.9.2 Composite power

3.10 Cartesian closure

Poly has exponential objects, that is, the family of lenses from any polynomial p to any polynomial q is found as an object q^p in **Poly**, and it is given by the following polynomial:

$$\Pi$$

We define this polynomial as follows in Agda:

```
-- CategoryData/Exponential

-- Exponential object.
-- Theroem 4.27, page 130 in Poly book.
_^_ : (r : Polynomial) → (q : Polynomial) → Polynomial
r ^ (mkpoly posQ dirQ) = ΠPoly (posQ , λ j → r ◁ (Y + Constant (dirQ j)))
infixl 30 _^_
```

29

where `Constant` is a helper function that takes a set to the constant polynomial sending every set to that set.

We do not prove that this object is the exponential object via the universal property of exponential objects, which would employ `agda-categories`' `Exponential` construction. Instead, we prove an equivalent statement, related to currying: that the following natural isomorphism exists

> **write isomorphism given by currying**

This statement is the same as the poly-book proves. In fact, we follow the same chain of isomorphisms as the book to conclude this. In Agda, this looks like:

> **write finished exponential**

The existence of the exponential object implies the existence of the canonical evaluation lens. Its implementation is as follows:

> **write eval arrow**

The cartesian closure of **Poly** has interesting implications, which we will consider when we talk about future work possibilities from this thesis.

3.11   Quadruple adjunction

Poly has a deep relationship with **Set**. One place where this is reflected is with the connection between the following functors:

- Constant functor $C : \textbf{Set} \to \textbf{Poly}$, which sends a set A to the constant polynomial $p(y) = A$; that is, the polynomial that sends all sets to A. This is a fully faithful functor, which means **Set** is a full subcategory of **Poly**.

- Linear functor $L : \textbf{Set} \to \textbf{Poly}$, which sends a set A to the linear polynomial $p(y) = Ay$; that is, the polynomial that sends a set y to the set of A-tuples of y. This is also a fully faithful functor.

- Plug in $\underline{0}$ functor $p(0) : \textbf{Poly} \to \textbf{Set}$, which sends a polynomial to a subset of its positions, namely the constant ones. For example, its action on $p(y) = 5y^{\mathbb{R}} + \mathbb{N}$ is $p(y) = 5 * 0^{\mathbb{R}} + \mathbb{N} = \mathbb{N}$. Again, $\underline{0}$ here stands for the set with 0 elements.

- Plug in $\underline{1}$ functor $p(1) : \textbf{Poly} \to \textbf{Set}$, which sends a polynomial to the set of its positions. Its action on the example polynomial above is $p(y) = 5 * 1^{\mathbb{R}} + \mathbb{N} = 5 + \mathbb{N}$.

Note that the notation $p(1)$ here is being used to represent a functor from **Poly** to **Set**, but most of the time it is used to talk about the set of positions directly.

These functors turn out to form a chain of adjunctions: $L \dashv p(1) \dashv C \dashv p(0)$. Adjunctions are forms of weak equivalences between categories, given by pairs of functors: the left and right adjoints. The intuition is that the left adjoint "adds as much information" as possible to the objects and morphisms it acts on, whereas the right adjoint "removes as little information" as possible in order to make the objects and morphisms it acts on "fit" in its target category. In the case of the relationship between **Poly** and **Set**, there are two functors putting different copies

of **Set** inside **Poly** (the fully faithful ones), and two functors that "forget" that they act on something *contains* sets.

For an illustration of what is meant by "forgetting", take the the first adjoint pair $L \dashv p(1)$ as an example. It acts the following way: $p(1)$ makes a polynomial functor "forget" it's a functor, while L "reminds" a set of "how to be" a functor.

The proof that this adjunction exists is relatively trivial for these functors, so we won't include it in the main text, but it is of course present in the accompanying formalization, under `Categorical/Adjunction`. It is done via the `agda-categories` scaffolding and via the unit-counit definition of adjunctions.

3.12   Equalizer

Equalizers are an important categorical structure, because when they exist in a category that also has products, it can be shown that the category is complete - in other words, it has all finite limits. A limit is an object, together with several morphisms from that object, in a category that "represent" or somehow summarize diagrams in that category.

An equalizer in general is an object E and a morphism eq, such that the following diagram commutes:

The objects A, B and T and the arrows f and g are arbitrary, and the morphism h is such that $f \circ h = g \circ h$. The equalizer object and its arrow are named this way for the intuitive reason that they select the part of A under which both arrows f and g agree, i.e. they "give the same output" in B. The notion of "giving output" is perhaps exploiting a **Set**-based intuition, but we believe **Poly** is close enough to **Set** to justify this. The object T relates to the universal property of the equalizer, in that it can be thought of as another "candidate" for the equalizer object along with its arrow h. What makes the actual equalizer E special is that, for any such candidate, there exists a unique arrow ! from that object to the equalizer, such that $eq \circ ! = h$.

The dual construction to the equalizer is the coequalizer, and as usual, is obtained by reversing all the arrows:

$$A \underset{g}{\overset{f}{\rightrightarrows}} B \xrightarrow{inc} C$$
$$B \xrightarrow{h} T, \quad C \dashrightarrow T$$
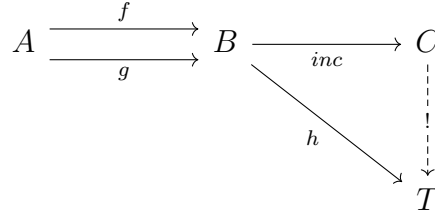
The intuition behind the coequalizer, in a **Set** context, is that it inspects the functions f and g, and any time these functions disagree on an input $a : A$, the two targets $b_1 : B$ and $b_2 : B$ get collapsed together by the arrow *inc*. This is done for all inputs. So, for example, if the functions disagree on all inputs, the set C becomes the singleton set, because *inc* will have to collapse every disagreeing output, eventually collecting all of them.

In **Poly**, the equalizer has two parts, as most constructions. For two given lenses $f : p \to q$ and $g : p \to q$, the equalizer is a polynomial that:

- Has the equalizer set of the map on positions of these lenses. This easily expressible with Σ-types:

  ```
  EqualizedPosition : Set
  EqualizedPosition = Σ (position p)
                        (λ z → mapPosition f z ≡ mapPosition g z)
  ```

- Has the *co*equalizer set of the map on directions of these lenses *at their equalized positions*. Normally, the partially applied functions `mapDirection f posp` and `mapDirection g posp` would have different types, since the second argument (a direction in q) depends on on the values of `mapPosition f posp` and `mapPosition g posp`. But in this case, we know these values to be the same, since the map on positions is the equalizer map, so the partially applied maps on directions have the same type, which means they can be coequalized. Cubical has a construction for this, `SetCoequalizer`.

Before we can introduce the equalizer polynomial and lens, we must address an issue: in the Cubical setting, SetCoequalizers can only be constructed out of functions that target types that are sets in the HoTT sense, so types of truncation level 0. In our case, this means that the directions of our polynomials must be sets, which is not a guarantee we have in the **Poly** category we've been working with; we only guarantee that these polynomials' targeted sets have *universe* level 0. To get around this, we work in a category that is identical to **Poly**, except the directions of the polynomials are guaranteed to be sets. This is the category **SetPoly**, where the objects are:

```
record SetPolynomial : Set₁ where
    constructor mksetpoly
    field
        poly : Polynomial
        isDirSet : ∀ {p : position poly} → isSet (direction poly p)
```

and the morphisms are lenses as usual, just wrapped in a new type constructor:

```
record SetLens (from to : SetPolynomial) : Set where
    constructor ⇆ˢ
    field
        lens : Lens (poly from) (poly to)
```

With that, we can introduce the equalizer polynomial, which is then given by:

```
eqPoly : Polynomial
eqPoly = mkpoly EqualizedPosition $
        λ ( posp , equal ) →
            SetCoequalizer (mdf posp)
                          (λ x → mdg posp
                                    (subst (direction q) equal x))
```

and the equalizer lens is:

```
mpe : position (eqObj) → position p
mpe = fst
mde : (fromPos : position (poly eqObj)) →
      direction p (mpe fromPos) →
      direction (poly eqObj) fromPos
mde _ dir = inc dir
eqLens : SetLens eqObj pˢ
eqLens = ⇆ˢ (mpe ⇆ mde)
```

Proving that this is in fact the equalizer of the **SetPoly** category is quite involved, so we will omit it from the text. It is of course included in the code. The idea behind the proof, however, is simple: since an equalizing lens of two lenses $f, g : p \rightarrow q$ must make sure following it with either f or g is equal, then the maps comprising f and g must be guaranteed to land in the same results after it. That is, the map on positions must be equalized on the inputs, and the map on directions must be equalized on the outputs. Equalizing two functions on outputs is, loosely speaking, what a coequalizer map does: it identifies potentially differing outputs of two functions.

3.13   Coequalizer

move coequalizer to future work. It's going to be really hard

3.14   Various properties of polynomials

We went over the intuition for polynomial functors as algebraic polynomials the beginning of this chapter, and many of these properties have been formalized. For example, here's a proof that the sum of two constant polynomials is still constant, exercise 4.1 in the poly-book:

```
isConstant : Polynomial → Type₁
isConstant (mkpoly pos dir) = (p : pos) → dir p ≡ ⊥

constantClosedUnderPlus : {p q : Polynomial} →
    isConstant p →
```

```
    isConstant q →
    isConstant (p + q)
constantClosedUnderPlus isConstantP isConstantQ (inj₁ x) = isConstantP x
constantClosedUnderPlus isConstantP isConstantQ (inj₂ y) = isConstantQ y
```

Other proofs of this kind are provided in the files `Cubical/Proofs.agda` and `Cubical/Various.agda`. Here's a list of them:

- The coproduct is the generalized coproduct indexed by $\underline{2}$.

- The product is the generalized product indexed by $\underline{2}$.

- A sum of two linear polynomials is still linear.

- A multiplication of two monomials is still a monomial.

- A lens to y is the same thing as a function from positions to directions. The idea is that the map on positions has no choice (a single position), and the map on directions is then a choice of position in the source polynomial.

- Given $p(y) = y^2 + y$, $p(\underline{2}) = \underline{6}$.

- $p(1)$ is equal to the set of positions of p.

- Given two constant polynomials $p(y) = \underline{3}$ and $q(y) = \underline{2}$, $p^q(y) = \underline{9}$.

4

Theory: Charts

The previous chapter covered the category **Poly**, of polynomials and lenses. This chapter covers a closely related category, the category of polynomials and charts, **Chart**. First, the category **Chart** is defined, then some universal constructions are proved, and finally the relation between **Poly** and **Chart** as commuting squares is shown.

4.1 The category Chart

The objects of **Chart** are exactly the same as **Poly**: polynomial functors. What differs are the arrows.

4.1.1 Chart

Arrows in **Chart** are called charts. The difference between charts and lenses is that while lenses has the map on directions going backwards, charts has the map on directions going forwards. This makes charts easier to reason about, since the map on directions and the map on positions goes the same way. An example chart can be seen in figure 4.1. The definition of `Chart` in Agda is as follows:

```
record Chart (p q : Polynomial) : Set where
    constructor _⇉_
    field
        mapPos : position p → position q
        mapDir : (i : position p) → direction p i → direction q (mapPos i)
```

The constructor name is $\rightrightarrows$ to indicate that both the map on positions and the map on directions goes forward. Notably, despite being maps between functors, charts are *not* natural transformations, and just what categorical structure they correspond to is not known. If we try to formulate it as an agda-categories natural transformation, like below, we get an unfillable hole:

```
asNatTransChart : {p q : Polynomial} →
    Chart p q →
    NaturalTransformation (asEndo p) (asEndo q)
asNatTransChart (f ⇉ f♯) = record {
    η = λ { X (posP , dirP) → (f posP) , (λ x → dirP {!   !}) } ;
```
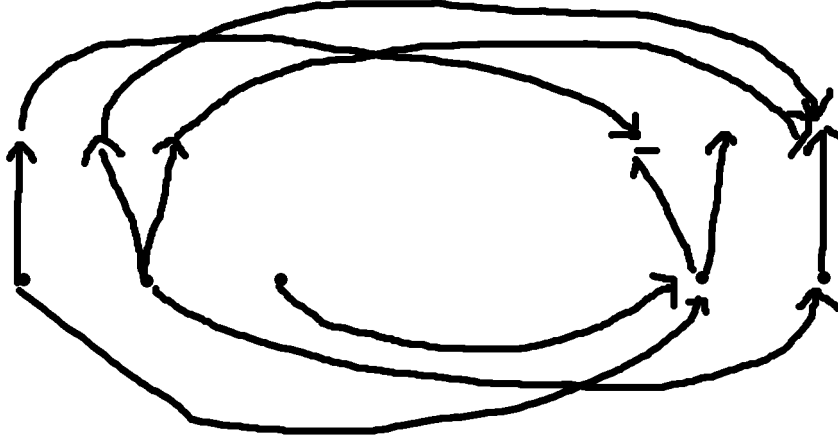
Figure 4.1: The polynomial p has 3 positions, and q has 2 positions. The map on position is a map between p's and q's positions. The map on direction is for each position of p, a map forward on the directions..

```
    ...
  }
```

In the hole, `x` is of type `direction q (f posP)`, but there is no such data to provide.

4.1.2 Identity chart

The identity chart, in similarity to the identity lens, maps positions to itself and directions to itself. In code:

```
idChart : {p : Polynomial} → Chart p p
idChart = id ⇉ λ _ → id
```

4.1.3 Composing charts

Composition of charts is naturally defined as following all the arrows. For the map on position function composition is used. For the map on directions, function composition is used as well, but with some extra care in dealing with the dependency of the map on positions.

```
_∘c_ : {p q r : Polynomial} → Chart q r → Chart p q → Chart p r
(f ⇉ f♭) ∘c (g ⇉ g♭) = (f ∘ g) ⇉ (λ i → f♭ (g i) ∘ g♭ i)
```

4.1.4 Associativity and identity laws

The associativity, left identity, and right identity laws are proved directly by `refl`. This fulfills all the requirements of a category and thus **Chart** is fully defined.

4.2 Chart equality

In similarity to polynomials and lenses, charts also needs a characterization of equality. Thus the same procedure is used, to define charts as Σ-type, and show that the Σ-type is equal to the record definition.

```
ChartAsΣ : (p q : Polynomial) → Type
ChartAsΣ p q = Σ[ mapPos ∈ (position p → position q) ]
                    ((i : position p) → direction p i → direction q (mapPos i))

chartAsΣToChart : {p q : Polynomial} → ChartAsΣ p q → Chart p q
chartAsΣToChart (mapPos , mapDir) = mapPos ⇉ mapDir

chartToChartAsΣ : {p q : Polynomial} → Chart p q → ChartAsΣ p q
chartToChartAsΣ (mapPos ⇉ mapDir) = mapPos , mapDir

chartAsΣ≡Chart : {p q : Polynomial} → ChartAsΣ p q ≡ Chart p q
chartAsΣ≡Chart {p} {q} = isoToPath
    (iso chartAsΣToChart chartToChartAsΣ (λ b → refl) λ a → refl)
```

Equality of charts is now defined for the Σ-type and then transferred to the chart defined as a record. There are two variants of chart equality, the direct, and slightly weaker variant, `chart≡`. As well as the more powerful variant, `chart≡∀`.

The weak variant of chart equality, is directly defined by using, Σ`PathTransport→Path`Σ. This results in the following type.

```
chart≡ : {p q : Polynomial} {f g : Chart p q}
    → (mapPos≡ : mapPos f ≡ mapPos g)
    → subst (λ x → (i : position p) → direction p i → direction q (x i))
        mapPos≡ (mapDir f) ≡ mapDir g
    → f ≡ g
```

The subst for the map on direction equality proof can easily be simplified. This results in the more powerful variant of chart equality, which is defined by using Σ`PathP` and function extensionality twice. This result in the following type.

```
chart≡∀ : {p q : Polynomial} {f g : Chart p q}
    → (mapPos≡ : mapPos f ≡ mapPos g)
    → ((i : position p) → (x : direction p i)
        → subst (λ h → direction q (h i)) mapPos≡
            (mapDir f i x) ≡ mapDir g i x)
    → f ≡ g
```

With chart equality available, some universal constructions in **Chart** will be proved next.

4.3 Initial object

The initial object of **Chart** is the same object 0 as for **Poly**. Of course, the chart from 0 to any other polynomial is not the same as the lens from 0, since the types are different. But the implementation is the same, since the position is $\bot$, the absurd function is used.

```
chartFromZero : {p : Polynomial} → Chart 0 p
chartFromZero = (λ ()) ⇉ (λ ())
```

The uniqueness proof is identical to the uniqueness proof for the initial object in **Poly**.

```
unique : {p : Polynomial} (f : Chart 0 p) → chartFromZero ≡ f
unique _ = chart≡ (funExt λ ()) (funExt λ ())
```

4.4 Terminal object

The terminal object of **Chart** is not the same as for **Poly**! The terminal object of **Chart** is Y, having one position, and exactly one direction at that position.

```
Y : Polynomial
Y = mkpoly ⊤ (λ _ → ⊤)
```

The chart from any polynomial to Y is implemented as unit for both the map on positions and map on directions.

```
chartToY : {p : Polynomial} → Chart p Y
chartToY = (λ _ → tt) ⇉ (λ _ _ → tt)
```

Uniqueness of `chartToY` follows directly from the uniqueness of the unit function.

```
unique : {p : Polynomial} (f : Chart p Y) → chartToY ≡ f
unique _ = chart≡ refl refl
```

4.5 Product

The product of **Chart** is the parallel product $\otimes$ defined in 3.8. Since the parallel product forms a monoid with Y, it makes sense that that the terminal object is Y.

The projections from `p ⊗ q` is defined by projecting both the position and directions.

```
π₁ : {p q : Polynomial} → Chart (p ⊗ q) p
π₁ = fst ⇉ λ i → fst
```

```
π₂ : {p q : Polynomial} → Chart (p ⊗ q) q
π₂ = snd ⇉ λ i → snd
```

The factorization chart is given by running the two functions in parallel on both the map on position and map on directions.

```
⟨_,_⟩ : {p q r : Polynomial} → Chart p q → Chart p r → Chart p (q ⊗ r)
⟨ f ⇉ f♭ , g ⇉ g♭ ⟩ = (λ i → (f i , g i)) ⇉ (λ i d → f♭ i d , g♭ i d)
```

The proofs for commutation is given as `refl`, and uniqueness is given by slightly longer proof using a lemma. Full code can be found in **Categorical/Chart/Product.agda**.

4.6 Coproduct

The coproduct of charts is the same as for lenses. The injections and factorization are defined the same as well.

```
i₁ : {p q : Polynomial} → Chart p (p + q)
i₁ = inj₁ ⇉ λ _ → id
```

```
i₂ : {p q : Polynomial} → Chart q (p + q)
i₂ = inj₂ ⇉ λ _ → id
```

```
[_,_]c : {p q r : Polynomial} → Chart p r → Chart q r → Chart (p + q) r
[ f ⇉ f♭ , g ⇉ g♭ ]c = [ f , g ] ⇉ [ f♭ , g♭ ]
```

The proofs are also similar, and can be found in **Categorical/Chart/Coproduct.agda**.

4.7 Commuting square between lenses and charts

One interesting thing about charts is how they behave together with lenses. This behavior can be condensed into a square of four polynomials, with lenses and charts going between them, such that the square commutes. Such a square is shown in figure 4.2.
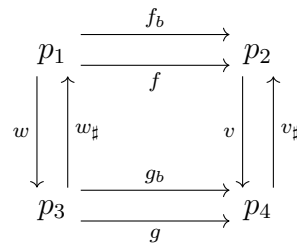

Figure 4.2: A square consists of four polynomials. Vertically there is a lens from p_1 to p_3 and from p_2 to p_4. Horizontally there is a chart from p_1 to p_2 and from p_3 to p_4.

For a square to commute there are two requirements. The first requirement is that the map on positions commute. Starting at a position in p_1, following w and then g, ending up at a position in p_4, should be the same as following f and then v.

The second requirement is that the map on directions commute. That is, $w_\sharp$ followed by f_b should be the same as g_b followed by $v_\sharp$. Since these functions depend on the different map on positions there are some plumbing to make this type correct, which can be seen in the Agda definition down below. Furthermore, the type that the map on directions commute depends on that the map on positions the commute. This explains the use of a Σ-type as well as the use of subst to make the types match.

In Agda, the conditions of a square commuting is defined as:

```
LensChartCommute : {p₁ p₂ p₃ p₄ : Polynomial} (w : Lens p₁ p₃) (v : Lens p₂ p₄)
    (f : Chart p₁ p₂) (g : Chart p₃ p₄) → Type
LensChartCommute {p₁} {p₂} {p₃} {p₄} (w ⇆ w♯) (v ⇆ v♯) (f ⇉ f♭) (g ⇉ g♭)
    = Σ mapPos≡ mapDir≡
    where
        mapPos≡ : Type
        mapPos≡ = (i : position p₁) → v (f i) ≡ g (w i)

        mapDir≡ : mapPos≡ → Type
        mapDir≡ p≡ = (i : position p₁) → (x : direction p₃ (w i))
            → f♭ i (w♯ i x) ≡
                v♯ (f i) (subst (direction p₄) (sym (p≡ i)) (g♭ (w i) x))
```

The use of commuting squares will be of importance in reducing dynamical systems to simpler ones, which is covered the applications section.

4.8 Composing squares between lenses and charts

Commuting squares are also composable in two different ways. They are composable both vertically and horizontally. This can be seen in figure **??**.

In Agda, horizontal and vertical composition is given as follows. However, the part of proof that the commutation of map on directions compose is unfinished, due to the complexity of an explosion of subst's, and left for future work, discussed further in **??**.

```
horizontialComposition : {p₁ p₂ p₃ p₄ p₅ p₆ : Polynomial}
    (f : Chart p₁ p₂) (g : Chart p₃ p₄) (h : Chart p₂ p₅) (r : Chart p₄ p₆)
    (w : Lens p₁ p₃) (v : Lens p₂ p₄) (m : Lens p₅ p₆)
    → LensChartCommute w v f g → LensChartCommute v m h r
    → LensChartCommute w m (h ∘c f) (r ∘c g)


verticalComposition : {p₁ p₂ p₃ p₄ p₅ p₆ : Polynomial}
    (f : Chart p₁ p₂) (g : Chart p₃ p₄) (r : Chart p₅ p₆) (h : Lens p₃ p₅)
    (w : Lens p₁ p₃) (v : Lens p₂ p₄) (m : Lens p₄ p₆)
```

Do diagram of squares composition - extracted this todo because otherwise latex gives a Float lost error

$\rightarrow$ `LensChartCommute w v f g` $\rightarrow$ `LensChartCommute h m g r`
$\rightarrow$ `LensChartCommute (h ∘`$_p$` w) (m ∘`$_p$` v) f r`

Add section about
stationary systems
thingy

5

Applications

As we mentioned in the background section, polynomials have applications to many different areas of study, of which the main three highlighted in the poly-book are dynamical systems, decision systems and databases. Our implemented applications are all instances of the first: we pay no attention to the other two perspectives. Several of our examples are listed in the book and its associated YouTube lecture series, but a few are our own contribution.

It is worth mentioning that many of the ideas in this chapter come from the Categorical Systems Theory book, even though that book does not even mention polynomial functors. What it does do is explain how dynamical systems are modeled by lenses in extreme, welcome detail. It also pays a lot of attention to the distinction between different theories of dynamical systems: stochastic, discrete, continuous etc. This matters to us because several of our examples are continuous systems that are discretized, so strictly speaking we are working within a theory of discrete dynamical systems. David Jaz Myers, the author, is also responsible for the chart-related part of the applications, through his guest lectures on the YouTube series as well as his book.

We will start by talking about the class of systems that **Poly** mainly talks about - ones that can be expressed through lenses. We'll then explain what our categorical constructs mean in the context of application to dynamical systems, and we'll end by showcasing implemented examples of these dynamical systems, some of which clearly display the advantages of this categorical framework.

5.1   Mode-dependent open dynamical systems

First, let's define what a dynamical system is, starting with the simplest case: a closed dynamical system. It can be thought of as a machine with some functionality that "runs" through time. It has a notion of a state space S (in our case a set of values) and a function $f : S \rightarrow S$, called the *dynamics* of the system. The purpose of the dynamics is to reveal what the system's state is at each timestep. Given an initial state S_0, one can then run several steps of a system by simply composing f with itself that many times. This sort of system is called *closed* because it does not take any inputs apart from the initial state, and does not provide any outputs. The only thing one might be interested in is the state of the system. An example of a closed dynamical system is the logistic map, defined by the following equation:

$$x_{t+1} = rx_t(1 - x_t)$$

The state space S here is $\mathbb{R}$, and the body of the function f is the expression on the right. The next state is defined in terms of the last state and a parameter r, so given an initial state x_0, one can generate an infinite sequence of states by running this function repeatedly. *Open* dynamical systems on the other hand are systems that, in addition to possessing a state space, also provide output to and receive input from their environment. "Environment" here is meant to be interpreted loosely: it is the collection of surrounding systems or user inputs, or some other source of data. Open dynamical systems are then composed of five pieces of data: an input set A, an output set B, a state space set S, and two functions: $\texttt{update} : S \times A \rightarrow S$ and $\texttt{readout} : S \rightarrow B$. This naturally suggests composition: one systems's output can be another system's input, if the types match.

Finally, *mode-dependent* open dynamical systems are ones in which the sets A and B can vary, and as such, which system is providing output to which other system's input can also change depending on the states of the system. This will become clearer in the next section, where we'll provide visualizations of what is being described here.

5.1.1   Wiring diagrams

Wiring diagrams are a way of expressing morphisms of monoidal categories that is much more visually intuitive and applies to a wide array of contexts, including dynamical systems[22]. In our case, we will use them to express open dynamical systems and their connections.
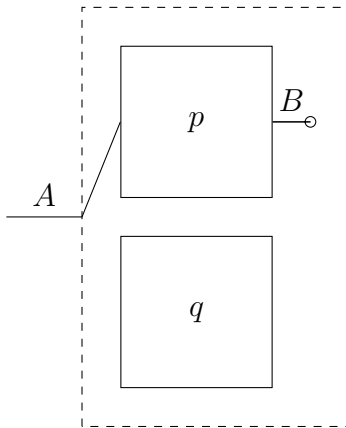
An important fact about a system like the one above shows is that they give morphisms in **Poly**. In this case, this system gives a composition of several arrows, which we will break down exactly how.

5.2 Polynomial functors in dynamical systems

The three main objects of our study - polynomials, lenses and charts - have important roles in the study of dynamical systems, and the categorical structures we've formalized do as well, especially the parallel product $\otimes$.

5.2.1 Polynomials as interfaces

A polynomial can be seen as the interface or "API" of a dynamical system, by describing what its outputs and inputs are. In particular, the set of positions of a polynomial gives the possible outputs of a system, while the directions at a given position p are the possible inputs given that the state is such that the output from that state is p. In this way, the system is mode-dependent. Take the polynomial below for instance.

$$p(y) = \mathbb{R}y^{\texttt{string}} + \top y^{\mathbb{N} \times \mathbb{N}}$$

As an interface to an abstract system, it applies to any system that, when its state is such that it outputs real numbers, the possible inputs are strings. When its state is such that it's outputting $\top$ (the singleton set), then the possible inputs to give it are pairs of natural numbers. Notice that this says nothing about what the state space of a system should concretely be, only that its state can specify its interface. Mode-dependence makes it so that writing a polynomial as an abstract box in a wiring diagram is hard, because the input and output wires change. Consider the polynomial above. As boxes, its interface would be both boxes below, depending again on the state:

draw

A good way of integrating this mode dependence inherent to polynomials into regular wiring diagrams is an open problem.

5.2.2 The parallel product $\otimes$

We come to a great motivating factor for the naming of this structure: it corresponds to putting systems in parallel. Given two polynomials p and q with their own interfaces, taking their parallel product $p \otimes q$ corresponds to creating a polynomial that has *both interfaces simultaneously*. Intuitively, one can think that any system that has this interface can only run one timestep upon being given both inputs, and it will output both outputs. The structure, in some sense, doesn't "touch" the behavior of the systems at all. In the below picture:

The resulting polynomial functor is the dashed line around the two inner boxes, representing p and q.

5.2.3 Lenses as behaviors

A particular kind of lens specifies behaviours of systems, namely any lens of shape $f : Sy^S \to p$, where p is a polynomial like the one described above and S is the state space set. Since lenses are maps on positions and on directions, in this case we get $f : S \to p(1)$; and $f\# : (\texttt{curState} : S) \to A(\texttt{f curState}) \to S$. This corresponds to our definition of open systems, with the added mode dependence: the map on positions is exactly the `readout` function, and the map on directions is the `update` function with the added constraint that not any A is valid - A is now a type family, so only the fibrations of A at the position (output) at the current state are valid.

If we draw a behavior-specifying lens on our previous example, not much changes visually, but we denote the state inside the box space as below:

This can then be seen as a lens from $S^S \to p$.

5.2.4 Lenses as wiring patterns

The other role that lenses play in systems of this kind is that they correspond to wiring patterns. For example, consider the abstract drawing of boxes below:

This corresponds to a lens from an abstract polynomial placed in the inner box to the abstract polynomial corresponding to the outermost box. We say "abstract" because we need to know nothing about the dynamics or concrete behavior of the systems here, all we need is the types of the interfaces to match. Now there are other wiring patterns possible here:

And they all correspond to different lenses. The intuition for the lens structure is that the map on positions selects which of the outer outputs will be served by the inner outputs, and the map on directions will fill in the needed inputs to the inner boxes.

5.2.5 Charts as system transformations (??)

Charts also have a deep interpretation in this context: it turns out that they correspond to . We give an implementation of the example given by David Jaz in his YouTube guest lecture for the polynomial functors course.

5.3   Implementing dynamical systems

Although dynamical systems are "just" lenses of a certain form, we choose a convenient representation for them so as to align with our thinking, analogously to how we choose to represent lenses, charts and polynomials as records and not Σ-types. A lot of this code is directly translated from the Idris-based implementation in the poly-book's github repository. Our characterization is the following record:

```
record DynamicalSystem : Set₁ where
    constructor mkdyn
    field
        state : Set -- S
        interface : Polynomial -- p
        dynamics : Lens (selfMonomial state) interface -- Sy^S → p
```

When it comes to implementing a framework for coding up dynamical systems in this context, some patterns appear over and over. Some of these patterns are:

- Transforming a pure function $f : A \to B$ into a memoryless dynamical system, i.e. one that has A as an input, B as output, A as a state space and whose dynamics are given by running the function f on `readout` and just replace the state with the input in `update`.

  ```
  functionToDynamicalSystem : (A B : Set) → (A → B) → DynamicalSystem
  functionToDynamicalSystem A B f =
    mkdyn B (monomial B A) (id ⇆ (\_ → f))
  ```

- Create an `emitter` polynomial that represents an interface of no inputs (i.e. takes the singleton set) and always outputs the same type.

  ```
  emitter : Set → Polynomial
  emitter = linear
  ```

  It should not be surprising that this is just the linear polynomial; we give it a special name to bind it to the current interpretation, in the spirit of *concepts with an attitude*[23].

- Taking the parallel product of two systems. Recall that the parallel product is an operation on objects (polynomials) but it also it is also possible to construct a canonical lens from two lenses $f : A \to C$ and $g : B \to D$ of type $\langle f \otimes g \rangle : A \otimes B \to C \otimes D$. In the context of systems *being* morphisms, this corresponds to putting them "in parallel" and giving rise to a new system, with both state spaces, interfaces and behaviors simultaneously.

  ```
  _&&&_ : DynamicalSystem → DynamicalSystem → DynamicalSystem
  mkdyn stateA interfaceA dynamicsA &&& mkdyn stateB interfaceB dynamicsB
      = mkdyn (stateA × stateB)
              (interfaceA ⊗ interfaceB)
              ⟨ dynamicsA ⊗ dynamicsB ⟩
  ```

- Plugging dynamics and wiring patterns together. If lenses account both for dynamics *and* wiring patterns, then it seems natural that a dynamical system (really a lens) that just contains several "free-floating" boxes that are simply sitting in parallel could be composed with a lens that represents a wiring pattern to give rise to a system that is actually usable. Indeed, this is the case, and the function that does this is called `install` by both us and the poly example code:

```
install : (d : DynamicalSystem) →
          (a : Polynomial) →
          Lens (DynamicalSystem.interface d) a →
          DynamicalSystem
install d a l = mkdyn (DynamicalSystem.state d)
                      a
                      (l ∘ₚ (DynamicalSystem.dynamics d))
```

- Outer box lens representations. An outer box that represents an enclosure, as well as enclosures of a couple different possible inputs are given: we use `auto` for most systems, since we're interested in just running them as closed systems and seeing what happens, but sometimes we're interested in giving constant inputs that come from outside the environment, and for that we use `constI`:

```
encloseFunction : {t u : Set} → (t → u) → Lens (monomial t u) Y
encloseFunction f = (λ _ → tt) ⇆ (λ fromPos _ → f fromPos)

auto : {m : Set} → enclose (emitter m)
auto = encloseFunction λ _ → tt

constI : {m : Set} → (i : m) → enclose (selfMonomial m)
constI i = encloseFunction λ _ → i
```

One can easily imagining extending `constI` to taking a function that operates on the number of timesteps so as to vary the input over time as well. This is something that is done in dynamical systems theory in a very informal capacity, under the intuition that the parameters of the system vary "slowly" with respect to the timescale of the state updates. With this framework, we gain the ability of making it precise for free. Of course, if we wanted to make the parameter variation respond to the system's dynamics, we'd need to make it a system in itself; this is a design consideration that we hope future developers take into account.

- Finally, there's a function that does what one accustomed to dynamical systems simulations would expect: given an abstract representation of a dynamical system (a value of type `DynamicalSystem` in our case), and an initial condition, it produces an infinite stream of outputs.

```
{-# TERMINATING #-}
run : (d : DynamicalSystem) →
      enclose (DynamicalSystem.interface d) →
```

```
        DynamicalSystem.state d →
        Stream (Polynomial.position (DynamicalSystem.interface d)) _
  run d e initialState =  [ output ] ++ (run d e next)
      where
          output : Polynomial.position (DynamicalSystem.interface d)
          output = Lens.mapPosition (DynamicalSystem.dynamics d)
                                  initialState
          next : DynamicalSystem.state d
          next = Lens.mapDirection (DynamicalSystem.dynamics d)
                                  initialState
                                  (Lens.mapDirection e output tt)
```

We need the Agda pragma `TERMINATING` here because the termination checker can't be sure that this code terminates, since `run` calls itself on input that is not guaranteed to be smaller.. We can then `take` on the resulting stream to produce regular Agda lists or vectors, and use them to render plots or print to the screen, and as long as we always run this in the context of a finite `take` call, the program should always terminate.

There are more helpers and utilities peppered throughout the code, but this is the basic glue needed to work with **Poly** in the context of dynamical systems, as the examples will demonstrate.

5.3.1   Examples: Discrete systems

We split the example implementations in the class of discrete- and continuous-time. As mentioned previously, all examples are in the strict sense discrete; the continuous ones are simply discretized with some *dt*. However, the split is still valid if one considers the spirit of the implementations.

5.3.1.1   Fibonacci

We start with the most basic example of all: a Fibonacci sequence generator implemented as a dynamical system in **Poly**. The wiring diagram of this system is as follows:

fibwd

To break it down: the + function takes two natural numbers and produces one output, and the *id* function takes a natural number and outputs a natural number. In terms of dynamical systems, however, we have to think about what the `readout` and `update` functions are doing for each system: `readout` can only read what the current state is, and `update` can only update the state given an input. At every timestep, all systems produce their outputs based on the current state, then all the inputs are read, and the states of all systems are updated. Therefore, the effect of the `id` function is only felt one timestep later. Hence it plays the subtle role of storing the Fibonacci value computed in the previous timestep, which is why its output is fed as an input to the + function. The code for this system is this:

```
delay : (A : Set) → DynamicalSystem
delay A = functionToDynamicalSystem A A id
```

```
plus : DynamicalSystem
plus = functionToDynamicalSystem (Nat × Nat) Nat (uncurry _+ℕ_)

prefib : DynamicalSystem
prefib = plus &&& delay Nat

fibWiringDiagram : Lens (interface prefib) (emitter Nat)
fibWiringDiagram = (λ {(sumOutput , idOutput) → idOutput})
                   ⇆
                   (λ {(sumOutput , idOutput) _ →
                          (idOutput , sumOutput) , sumOutput })

fibonacci : DynamicalSystem
fibonacci = install prefib (emitter Nat) fibWiringDiagram
```

This example provides a good opportunity to understand dynamical systems in the context of polynomials, so let's take the chance to dig in further. In keeping with the previously mentioned intuition, the `fibWiringDiagram` lens is just one of many possible wiring patterns for an emitter-enclosed system that contains two "floating" dynamical systems that have the interfaces of `plus` and `delay Nat`. The outer outputs are the target of the lens, so the polynomial `emitter Nat`, which is a polynomial that takes the singleton set as input and outputs natural numbers. The inner inputs are given by the wiring pattern in the lens, specifically in the second function (the map on directions). This part of the code:

```
(λ {(sumOutput , idOutput) _ →
    (idOutput , sumOutput) , sumOutput })
```

is, on the right hand side, saying that the two inputs to `plus` are given by the output of `delay Nat` and the its own output, named `idOutput` and `sumOutput` respectively in the lambda. The input to `delay Nat`, on the other hand, is given by the output of the sum: in the next time step, this value will appear in the variable `idOutput`.

Two final remarks about the Fibonacci sequence generator: the first is that it is actually an efficient way to express Fibonacci dynamics, since there is no recursive calls involved - and the second is that we went into as much detail as possible explaining it so that we can rely on some of the intuition built here to explain the more complicated systems that follow.

5.3.1.2 Flip Flop

Flip Flop is a simple example with two dynamical systems showing how one system can simulate the behavior of another system. This simulation is exactly a special kind of commuting square between charts and lenses.

Both systems needs to have the same interface, which in this case is a linear polynomial {on, off}y, that can output either on or off, and always takes unit as input. However, both the state and the behavior of the systems can differ. One of the systems is a flip

flop system, with states on or off, with behaviour that outputs this state, and flips the state in the update. The other system has natural numbers as states, outputs the two different states depending on the natural number modulus 2, and updates the state by always taking the successor of the state. Running these systems has the exact same output. This can be expressed concisely by a chart transforming the state space of the natural numbers to either on or off. Then by putting this chart in a commuting square such as shown in diagram **??**.

```
data Switch : Set where
    on : Switch
    off : Switch


toggle : Switch → Switch
toggle on = off
toggle off = on


-- | Commonly used where input to enclosed dynamical system where updateState only
ignoreUnitInput : {A B : Set} → (A → B) → A → ⊤ → B
ignoreUnitInput f a tt = f a


-- | Note: linear interface is used to accept only 1 possible input.
--    Readout defined as id to expose state.
flipFlop : DynamicalSystem
flipFlop = mkdyn Switch (linear Switch) (id ⇆ ignoreUnitInput toggle)


-- | Result is: on, off, on, off...
flipFlopRan : Vec Switch 10
flipFlopRan = take 10 $ run flipFlop auto on


modNat : Nat → Switch
modNat n = if n % 2 == 0 then on else off


-- | To compare flipFlop and counter they need to have the same interface.
counter : DynamicalSystem
counter = mkdyn Nat (linear Switch) (modNat ⇆ ignoreUnitInput suc)


-- | Result is: on, off, on, off...
counterRan : Vec Switch 10
counterRan = take 10 $ run counter auto 0


-- | Morphism between p dynamicalSystems with states Nat and Switch.
morphSystem : Nat → Switch
morphSystem = modNat


-- | The square expressing the simulation.
square : LensChartCommute (dynamics counter) (dynamics flipFlop) (morphSystem ⇉ λ
square = law₁ , law₂
```

$$Ny^1 \xrightarrow[\quad \%2 \quad]{\quad \%2 \quad} 2y^1$$

$$\%2 \downarrow \uparrow suc \qquad id \downarrow \uparrow toggle$$

$$2y^1 \xrightarrow{\quad id_{chart} \quad} 2y^1$$
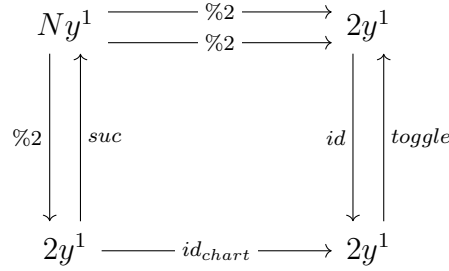
This showcased how a simple chart in a commuting square can compare behaviour of two different dynamical systems. By generalizing, and using more advanced charts and commuting squares, more advanced comparisons can be made.

5.3.1.3 Simulating state machines

A more advanced example, in comparison to flip flop, is how one state machine can simulate another. This is often expressed as that one state machine is more minimal than another, that is, it has less states but achieves the same behaviour. This is exactly what the special commuting square can express. Because the same idea is followed as for the flip flop example, this example will not be discussed further, but can be found in **Dynamical/Chart/SimulateStateMachine.agda**.

5.3.1.4 Moore machine

A Moore machine is the same as a lens to a monomial. A Moore machine can be defined as:

```
record MooreMachine {State Input Output : Set} : Set where
    constructor mkMooreMachine
    field
        readout : State → Output
        update : State → Input → State
```

Since a moore machine is a polynomial, it is possible to define the dynamical system as a moore machine and then transport it into a lens, or the other way around.

5.3.1.5 Deterministic finite automaton

A deterministic finite state automaton (DFA) is also a special kind of lens, the lens to the monomial $2y^{alphabet}$. A DFA can be defined as:

```
record DFS {State Alphabet : Set} : Set where
    constructor mkDFS
    field
        -- Transition function
        update : State → Alphabet → State
        -- Partitions all states into recognized and unrecognized states.
        recognized : State → Bool
```

The equality can be defined as a simple isomorphism.

The same trick as for the Moore machine can be used to provide an initial state to the DFA.

5.3.1.6 Turing machines

5.3.2 Examples: Continuous systems

Now we move on to discretized continuous-time systems. They may be treated in much more rigorous mathematical detail than we do here, by representing the state space as topological spaces or manifolds [24], but as we've mentioned, to showcase their implementation in **Poly** we discretize them simply according to a differential time dt supplied by the user.

Supporting Haskell code

Unfortunately, not all the code in this thesis could be kept in Agda, since it would require too much reimplementation of basic user functionality, or would not be efficient. Therefore, there are the following accompanying tiny Haskell libraries:

- A command-line interface library for running the continuous-time systems, since each of them requires different parameters, and using the `optparse-applicative` **optparse** library.

- A plotting library, which renders each system as a collection of values of the system at a sequence of timsteps, written using the `Chart` **chart-lib** library (not to be confused, of course, with the Chart arrow between polynomials).

- A matrix-multiplication library with a single function which takes a list-of-lists representation of a matrix and returns its pseudoinverse. This is needed for the last example. Taking the pseudoinverse of a matrix is a costly operation, and there was an attempt to do it in Agda, but the associated proofs were not trivial and the performance was prohibitively slow. Therefore, we outsource matrix multiplication to the `hmatrix`[25] Haskell library, which relies on underlying BLAS **blas** algorithms for speed, while passing `trustMe` proofs to guarantee matrix sizes at the Agda level.

For more information on this supporting code, see the Appendix A. With this context, we move on the the systems in question.

5.3.2.1 Lotka-Volterra

The Lotka-Volterra predator-prey model is an essential system [26] that portrays how the population of predators interacts with the population of prey in a given environment. The two main quantities in the model are two real numbers, abstractly representing the amount of the two types of animals - these are its states, in our context. The model is given by the two equations below, where r and f represent

our populations of rabbits and foxes - our chosen species of prey and predator:

$$\dot{r} = \alpha r - \beta f r \tag{5.1}$$

$$\dot{f} = \delta f r - \gamma f \tag{5.2}$$

The logic of it is that, for rabbits, they will reproduce naturally at a rate α, while being hunted at a rate βf. In other words, their encounters with and subsequent predation by foxes is modeled by multiplying the number of foxes by a "fox appetite" parameter β. For foxes, their population can be thought of as flourishing or growing at the rate that they can encounter and eat rabbits. This aspect is captured by the term $\delta f r$. They naturally die at a rate of γ. Note that δ doesn't have to match β - the rate at which foxes eat rabbits is not necessarily the same as the rate at which fox populations grow as a result of being well-fed, though the intuitive relationship between these parameters could be captured by a more sophisticated model.

This system can be represented in terms of our open dynamical systems by considering the rabbit and fox populations as each inhabited by an individual system. They output their population (so `readout` is the identity function) and that gets fed as an input to the other system, along with any other parameters - in this case α, β, δ and γ.
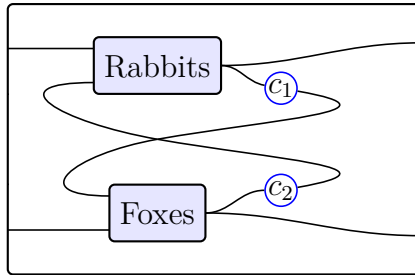
Thinking of the systems individually allows us a "separation of concerns" in specifying these systems' behavior that isn't there in environments like Matlab or Python. The code for the rabbit system is:

```
rabbits : DynamicalSystem
rabbits = mkdyn ℝ (mkpoly ℝ λ _ → ℝ × ℝ) (readout ⇆ update)
  where readout : ℝ → ℝ
        readout state = state
        update : ℝ → ℝ × ℝ → ℝ
        update state (birthRabbits , deathRabbits) =
            state + dt * (state  * (birthRabbits - deathRabbits))
```

The differential equation is directly discretized, so the differential amount of time dt goes to the right hand side of 5.1 and multiplies the entire thing; this is then added to the current state to produce a new state.

Similarly, the fox system is given as:

```
foxes : DynamicalSystem
foxes = mkdyn ℝ (mkpoly ℝ λ _ → ℝ × ℝ) (readout ⇆ update)
  where readout : ℝ → ℝ
```
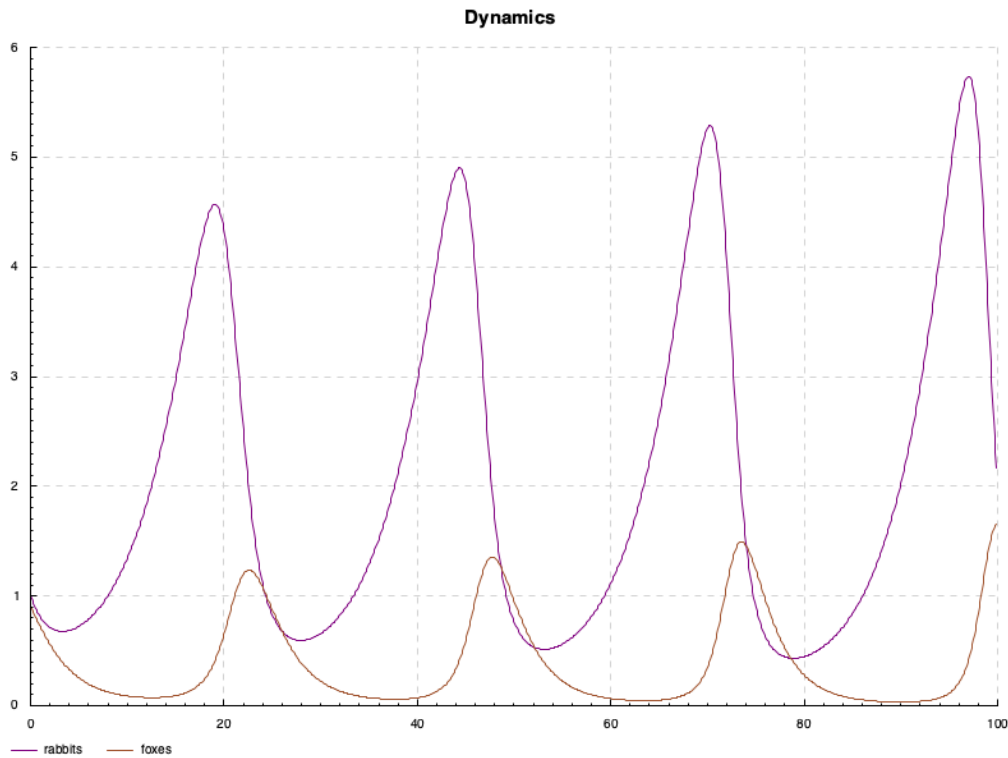
Figure 5.1: Lotka-Volterra oscillatory dynamics. As the population of rabbits grows, the population of foxes catches up, causing it to go down.

```
readout state = state
update : ℝ → ℝ × ℝ → ℝ
update state (birthFoxes , deathFoxes) =
    state + dt * (state * (birthFoxes - deathFoxes))
```

Note that each of these systems have only two parameters, whereas by the equations 5.1 and 5.2, they ought to have three: 5.1 should get α and β from "outside" the system, whereas 5.2 should get δ and γ. It is a somewhat arbitrary choice that we make to make the β and γ parameters part of the larger system, that is the one given rise to by the composition of these systems tensored with the wiring diagram, and not constants within the system themselves, for instance.

The final system is then a composition with a certain wiring diagram lens:

```
preLV : DynamicalSystem
preLV = rabbits &&& foxes

-- Wiring diagram is an lens between monomials
lotkaVolterraWiringDiagram : ℝ → ℝ →
    Lens (DynamicalSystem.interface preLV) (selfMonomial (ℝ × ℝ))
lotkaVolterraWiringDiagram foxPerCapDeath foxHunger =
  outerOutput ⇆ innerInput
    where outerOutput : ℝ × ℝ → ℝ × ℝ
          outerOutput (rabbitOutput , foxOutput) =
```

```
          rabbitOutput , foxOutput
      innerInput : (outputs : ℝ × ℝ) →
                direction (selfMonomial (ℝ × ℝ))
                          (outerOutput outputs) →
                direction (interface preLV) outputs
      innerInput (r , f) (rabMaxPerCapGrowth , howNutritiousRabbitsAre) =
        (rabMaxPerCapGrowth , foxHunger * f) ,
        (foxPerCapDeath * r , howNutritiousRabbitsAre)
-- Final system is composition of wiring diagram and dynamics
lotkaVolterra : ℝ → ℝ → DynamicalSystem
lotkaVolterra β γ = install preLV
                          (selfMonomial (ℝ × ℝ))
                          (lotkaVolterraWiringDiagram β γ)
```

Running this system through the command line interface produces the expected oscillating dynamics of the Lotka-Volterra model, albeit with some numerical instability causing the oscillations to grow slightly. Figure 5.1 shows this dynamics, obtained by running the system through the command line interface with the following parameters.

```
./Plot LotkaVolterra --alpha 0.2 --beta 0.2 --delta 0.5 \
                     --gamma 0.4 --r0 1.0 --f0 0.9 --dt 0.1
```

make this pretty. perhaps improve the plotting code as well? it's a bit cringe

5.3.2.2 Lorenz system

The Lorenz system[27] is what is known as a chaotic system, in an elementary form. It was introduced by Edward Lorenz as a simplified model of atmospheric convection. The three dimensionless (as in, pure proportions with no physical units) variables x, y and z correspond to rate of convection, temperature difference between ascending and descending air currents, and how temperature changes with height. It's given by the following differential equations:

$$\dot{x} = \sigma(y - x) \tag{5.3}$$

$$\dot{y} = x(\rho - z) - y \tag{5.4}$$

$$\dot{z} = xy - \beta z \tag{5.5}$$

The system has three numeric parameters, σ, ρ and β and for specific values of these parameters ($\sigma = 10.0$, $\rho = 28.0$, $\beta = 8/3$), it displays chaotic dynamics - dynamics that have the property of becoming hard to predict as time goes on. What this means is that for two arbitrarily close initial conditions, the distance between the trajectories followed by running the system with these initial conditions grows exponentially, until the trajectories are maximally distant. Figure 5.2 shows an example trajectory of the system with the parameters mentioned above.

This is known as a "strange attractor": the system's trajectories eternally oscillate between two attracting states, but it never converges to a single orbit, jumping back and forth. The system is a greatly simplified model of the atmosphere, but provides
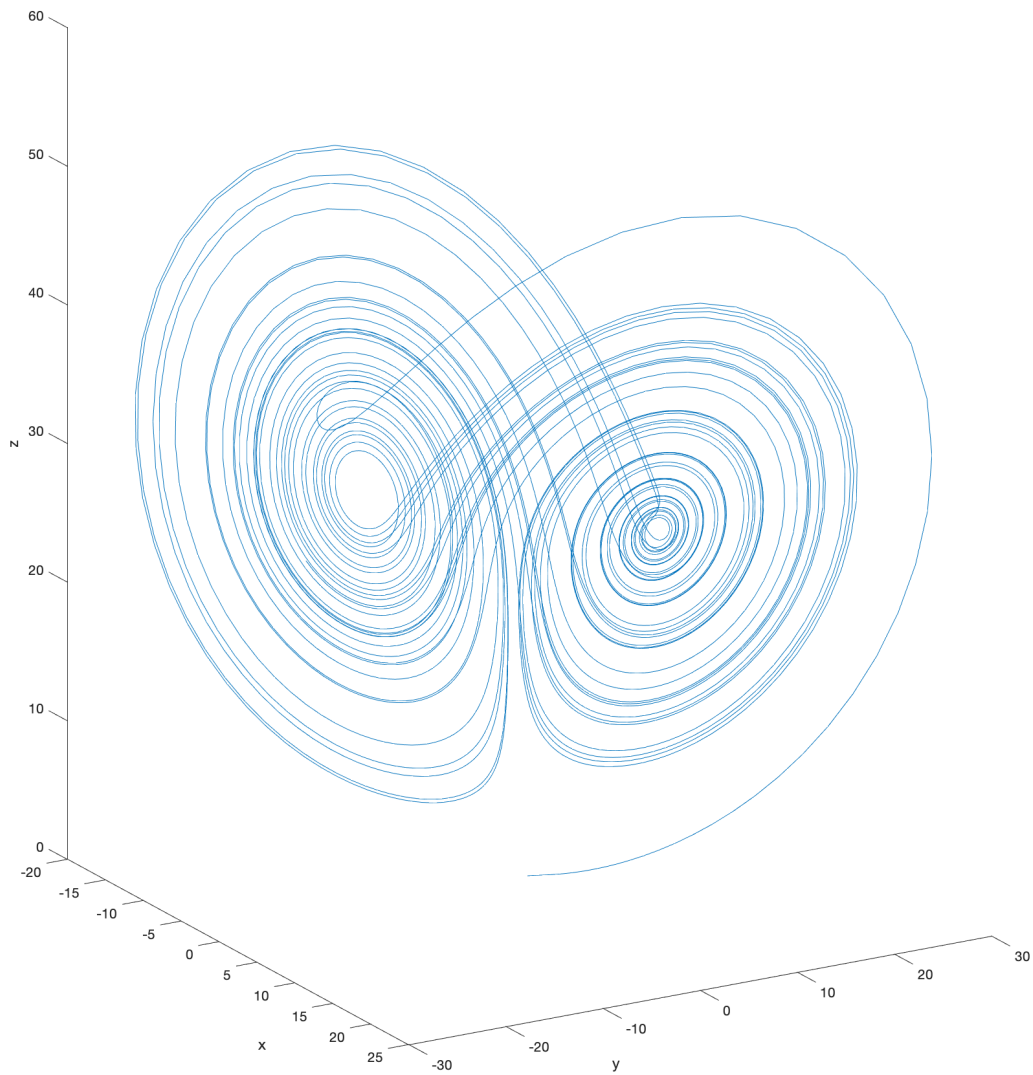
Figure 5.2: "Lorenz butterfly", a visualization of a trajectory of the Lorenz system. dt is set to 0.01.

a good "essence" of chaotic dynamics that are hard to predict, even in just three variables.

Implementing the Lorenz system in terms of **Poly** follows the same pattern as the Lotka-Volterra model, by capturing each equation in its own box with its own internal state, with the difference that the external parameters are now fixed, so the inputs are now only the external variables upon which each box should depend. In Agda, the systems are given as:

```
-- First order differential equations
x : ℝ → DynamicalSystem
x dt = mkdyn X (mkpoly X λ _ → Y) (readout ⇆ update)
  where readout : X → X
        readout state = state
        update : X → Y → X
        update (xnt state) (ynt y) =
            xnt (state + dt * (σ * (y - state)))

y : ℝ → DynamicalSystem
y dt = mkdyn Y (mkpoly Y λ _ → X × Z) (readout ⇆ update)
  where readout : Y → Y
        readout state = state
        update : Y → X × Z → Y
        update (ynt state) ( xnt x , znt z ) =
            ynt (state + dt * (x * (ρ - z) - state))

z : ℝ → DynamicalSystem
z dt = mkdyn Z (mkpoly Z λ _ → X × Y) (readout ⇆ update)
  where readout : Z → Z
        readout state = state
        update : Z → X × Y → Z
        update (znt state) (xnt x , ynt y) =
            znt (state + dt * (x * y - β * state))
```

And they're again tensored, producing a system with 3 inputs and 3 outputs, then wired in the right way (x is given as an input to y and z, y to x and z, and z to y, as per the system equations)

```
preLorenz : ℝ → DynamicalSystem
preLorenz dt = x dt &&& y dt &&& z dt

-- Wiring diagram is an lens between monomials
lorenzWiringDiagram : Lens (interface (preLorenz _))
                           (emitter (X × Y × Z))
lorenzWiringDiagram = mp ⇆ md
  where mp : X × Y × Z → X × Y × Z
        mp (x , y , z) = x , y , z
        md : X × Y × Z → ⊤ → Y × (X × Z) × (X × Y)
```
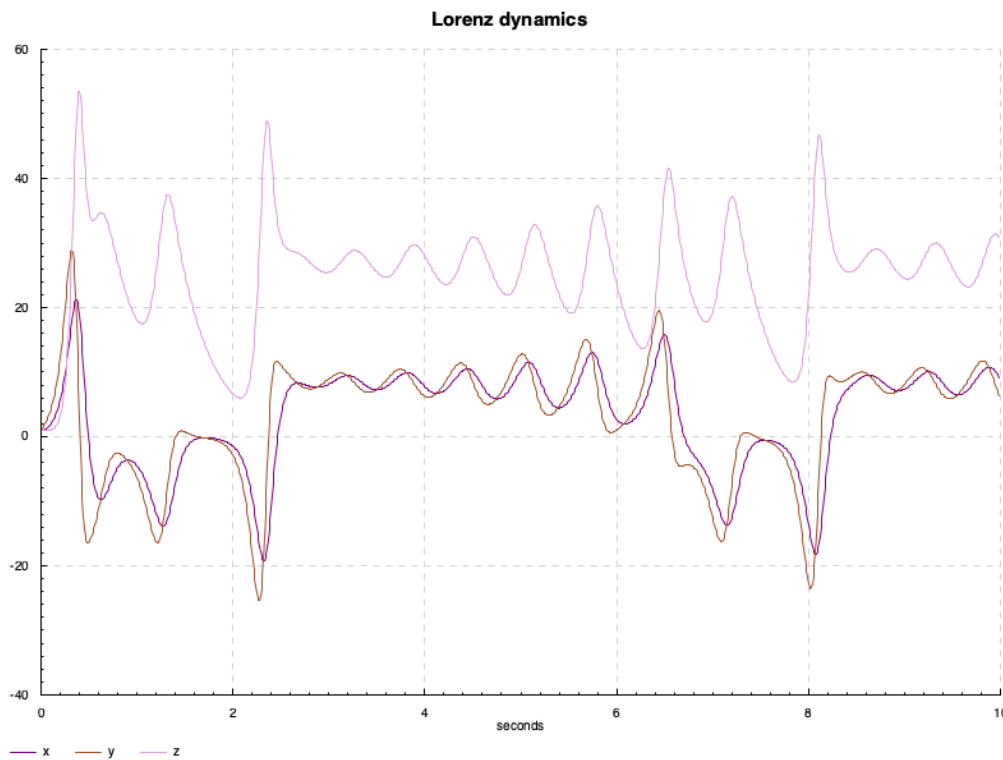
Figure 5.3: Lorenz system chaotic dynamics. The three variables keep oscillating in ever-changing patterns, never settling to a stable cycle.

```
md (x , y , z) _ = y , (x , z) , (x , y)

-- Final system is composition of wiring diagram and dynamics
lorenz : ℝ → DynamicalSystem
lorenz dt = install (preLorenz dt)
                    (emitter (X × Y × Z))
                    lorenzWiringDiagram
```

An example run of this system, obtained by `take` of 1000 timesteps at $dt = 0.01$, is given in Figure 5.3.

5.3.2.3 Hodgkin-Huxley model

The two examples given in this section are closed systems, meaning they take no input - what this really means is that they take the singleton set as input, so each timestep the element `tt` is given as a value and the system produces the next output. Now we come to a slightly bigger and more realistic system with 4 equations: the Hodgkin-Huxley [28] model. It is a system that models action potentials in the giant squid neuron, by thinking of the neuron membranes as circuits, so that the relevant quantities are all electrical (charge, current, capacitance etc). This model captures the flow of sodium and potassium ions through ion channels in the membrane, like in Figure 5.4, obtained from [29].

The interesting thing about the Hodgkin-Huxley model is that it actually lends itself to
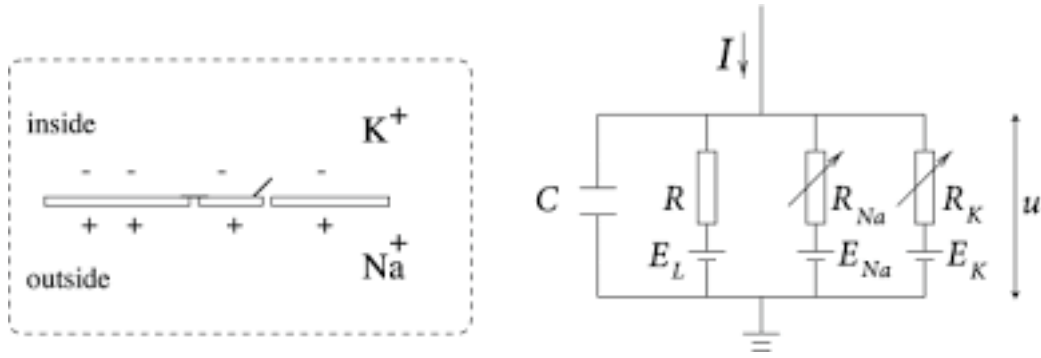
Figure 5.4: Analogy schematic of a neuronal membrane modeled as a circuit, from the book Neuronal Dynamics.

being implemented as an open system. There are several equivalent ways of formulating HH as a 4-dimensional dynamical system, but the equations we use are according to this online tutorial **https://mark-kramer.github.io/Case-Studies-Python/HH.html** and given by:

$$\dot{V} = G_{na}m^3h(E_{na} - V_s) + G_kn^4(E_k - V_s) + G_L * (E_L - V_s) + I_e \qquad (5.6)$$

$$\dot{m} = \alpha_m(V)(1.0 - m) - \beta_m(V)m \qquad (5.7)$$

$$\dot{h} = \alpha_h(V)(1.0 - h) - \beta_h(V)h \qquad (5.8)$$

$$\dot{n} = \alpha_n(V)(1.0 - n) - \beta_n(V)n \qquad (5.9)$$

The variables here mean the following: V is the resulting voltage between the intra and extracellular mediums, m encodes potassium channel activation probability, h encodes sodium channel activation probability and n encodes sodium channel *inactivation* probability, and I_e is the driven input voltage. All other symbols are constants and helper functions that have their own deep theoretical justification which we will not get into, we'll just explain the basic logic of the system. The idea is that potassium and sodium channels are positively charged, so their flow across their membrane corresponds to current. However, their flow is not only affected by voltage, but also by concentration: in a chemical setting, substances can saturate a solution regardless of their charge. The interacting probabilities of channels activating (opening, thus allowing particles to flow from more to less concentrated environments) or inactivating give rise to interesting voltage dynamics - the voltage is in fact the output we're interested in for this particular application. Given a strong enough input current, the system can be made to display oscillating behavior, which corresponds to the neuron firing.

The input current can be a constant value, in which case it might as well be treated as a parameter, or an input to an open dynamical system as we've been describing in the framework we work in. This allows us to model "changing a parameter" of a dynamical system, but in a way that is totally internal to our modeling language and does not rely on ad-hoc formulations or programming. For a description of the informal process of analyzing changing a parameter as "turning

a knob" at a timescale that is much bigger than the system's dynamics, see this lecture by Shane Ross on the book Nonlinear Dynamics and Chaos by Strogatz: https://www.youtube.com/watch?v=BBd68_q3Dgg.

The code for the equations themselves, represented as systems, follows the same pattern as the two previous examples. It also involves many fixed constants, discovered experimentally, and so it is quite big, so for the sake of clarity we omit most of it for this system. Instead we focus on the wiring pattern, as well use of `constI` as a constant input to the system to showcase running a system:

```
preHH : ℝ → DynamicalSystem
preHH dt = voltage dt &&&
           potassiumActivation dt &&&
           sodiumActivation dt &&&
           sodiumInactivation dt


hodgkinHuxleyWiringDiagram : Lens (interface (preHH _)) (selfMonomial ℝ)
hodgkinHuxleyWiringDiagram =
    (λ {(v , m , h , n) → v }) ⇆
    (λ {((v , m , h , n)) Ie → (Ie , m , h , n) , v , v , v })


hodgkinHuxley : ℝ → DynamicalSystem
hodgkinHuxley dt = install (preHH dt)
                           (selfMonomial ℝ)
                           hodgkinHuxleyWiringDiagram


hhSeq : ℝ → Stream ℝ _
hhSeq dt = run (hodgkinHuxley dt) (constI Ie) (V₀ , m∞ , n∞ , h∞)
  where V₀ : ℝ
        V₀ = -70.0
        m∞ : ℝ
        m∞ = 0.05
        n∞ : ℝ
        n∞ = 0.54
        h∞ : ℝ
        h∞ = 0.34
        Ie : ℝ
        Ie = 10.0
```

Notice the constant driven input current of 10nA. An example run of the system under this setup, with $dt = 0.02$ and run for 5000 steps, is seen in Figure 5.5.

Since this is a system with input, however, we can exploit this in order to do parameter variation: if we want to see how the system's dynamics vary with a change in parameters that is comparatively slow, we can create a new system whose responsibility is to output the changing current.

Adding this is simple: we take our finished Hodgkin-Huxley system and tensor it with a new system whose only responsibility is to grow its own output at a rate
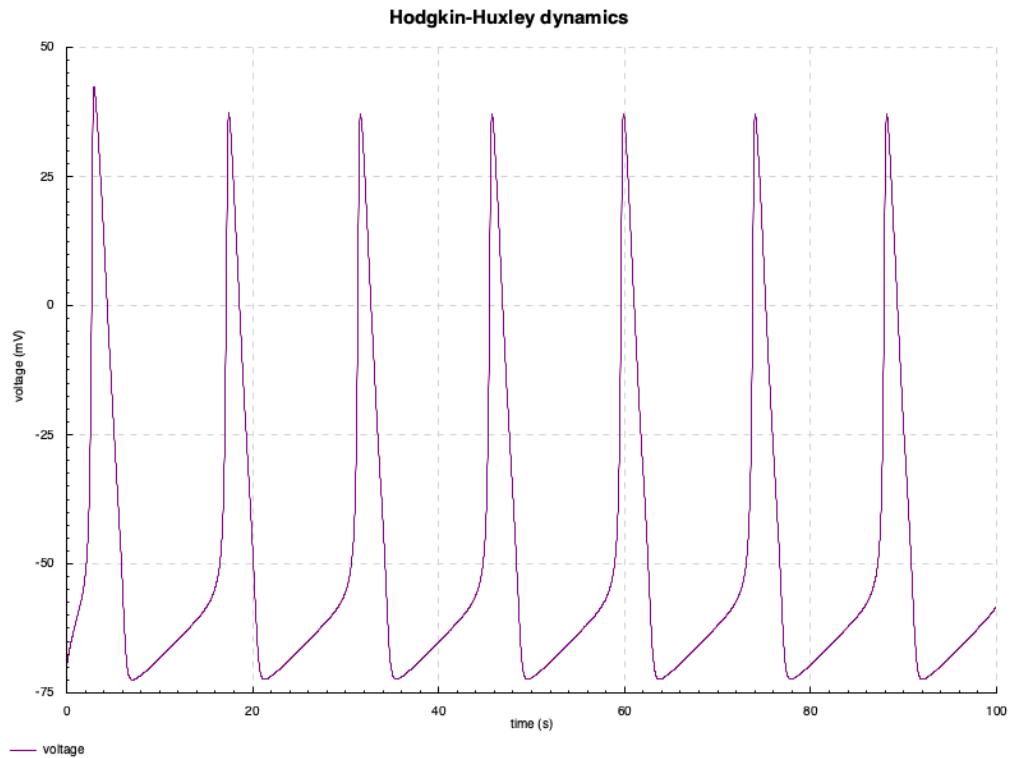
Figure 5.5: Hodgkin-Huxley model dynamics with input current $I_e = 10$nA and $dt = 0.02$. The simulated neuron fires repeatedly.

slower than HH's. This "current grower" system is so simple we inline it, using the helper `functionToDynamicalSystem`. It could have been implemented in other ways - and in fact, this could be its own investigation, like how changing parameters quickly/slowly/exponentially or whatever, could landing systems in different dynamics landscapes. We choose a simple ticking, linear increase at each timestep here.

```
preHHWithInput : ℝ → DynamicalSystem
preHHWithInput dt = hodgkinHuxley dt &&&
                    functionToDynamicalSystem ℝ ℝ
                                              λ x → x + (dt * 0.02)
```

We then wire it with HH, taking care to provide the current increaser this time using the `auto` helper to always give `tt` as input to the resulting system; we no longer need to provide it an external input of any kind, since this is now built into the system. We also take care with the wiring pattern to give our "current grower" system's output to both itself and to HH.

```
hhWithInputWiring : Lens (interface (preHHWithInput 0.0)) (emitter ℝ)
hhWithInputWiring =
    (λ { (hhOut , _) → hhOut}) ⇆
    λ { (hhOut , fnOut) _ → fnOut , fnOut }


hhWithInput : ℝ → DynamicalSystem
hhWithInput dt = install (preHHWithInput dt) (emitter ℝ) hhWithInputWiring
```
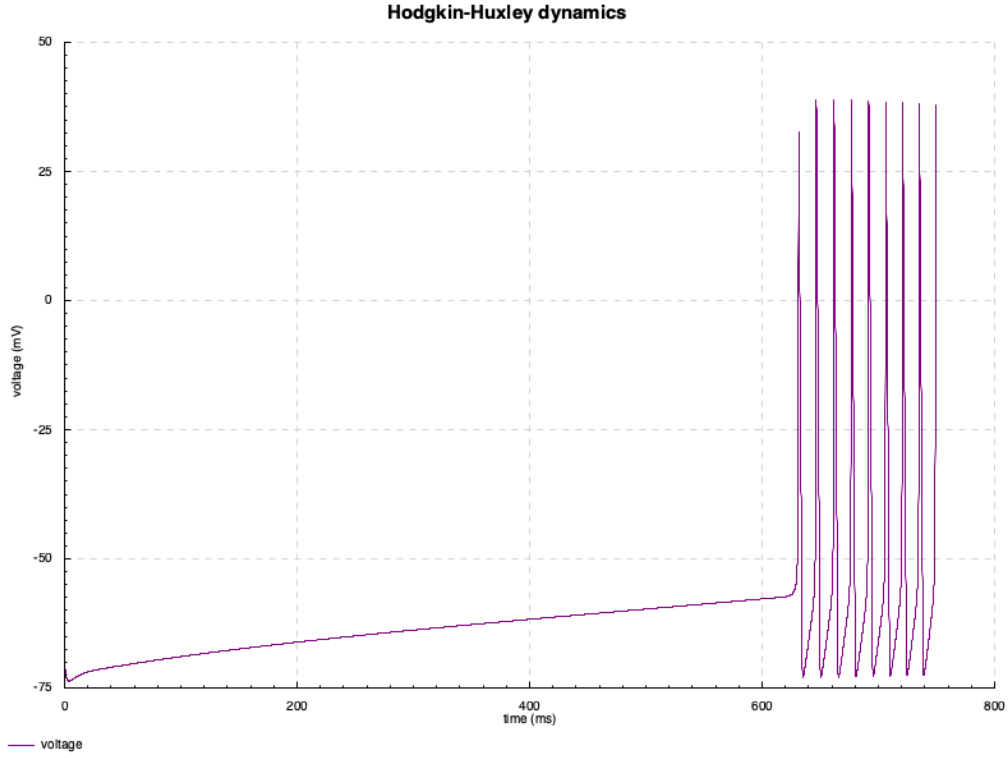
Figure 5.6: Hodgkin-Huxley model dynamics with $dt = 0.05$ and input current growing linearly from -5nA at a rate of $0.02dt$ per timestep. The simulated neuron starts by not firing, with the resting state slowly growing as the input current grows, then fires repeatedly.

```
hhSeqWithInput : ℝ → Stream ℝ _
hhSeqWithInput dt = run (hhWithInput dt) auto ((V₀ , m∞ , n∞ , h∞) , -5.0)
  where V₀ : ℝ
        V₀ = -70.0
        m∞ : ℝ
        m∞ = 0.05
        n∞ : ℝ
        n∞ = 0.54
        h∞ : ℝ
        h∞ = 0.34
```

The resulting dynamics is what one would expect from Hodgkin-Huxley at the input current levels explored by the current grower, and can be seen in 5.6.

5.3.2.4 Reservoir computer

Reservoir computing (RC) refers to a broad range of machine learning techniques [30] in the general paradigm of recurrent neural networks. Generally speaking, they involve exploiting the intrisic computations that happen in random, fixed recurrent neural networks (RNNs) to learn how to reproduce complex dynamics i.e. nonlinear

Figure 5.7: Reservoir setup when compared to neural networks. The dashed box represents an input/output boundary around the reservoir. Figure with modification from Cucchi et al. (2022) [33]. Licensed under CC BY.

timeseries, and have found many applications [31][32]. What makes the RC approach attractive is that, since the source of the dynamics is fixed, training is simplified when compared to regular neural networks, which require backpropagation: RCs can be tr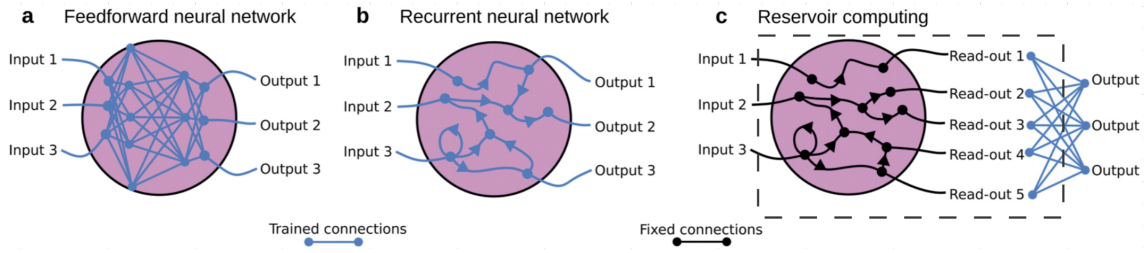ained by simple linear regression on the output layer. Another advantage is that again, since the dynamics are fixed, they do not need to be implemented in software, and can instead be implemented in physical systems [33]. Figure 5.7 contains a schematic showcasing the RC approach compared to neural networks.

The most representative type of reservoir computer is the *Echo State Network* (ESN) [34]. In this setup, the reservoir is composed of a large collection of stateful nodes. All nodes are connected to all other nodes, and these connections are weighted. Nodes also carry their own state. The reservoir has an input layer, through which data points of the timeseries to be learned is fed, and an output layer, through which the reservoir's output dynamics are exposed.

To actually use a reservoir to predict a timeseries, the process is organized in three stages, as follows:

- **Training / data collection**: In this stage, the reservoir accumulates training data from an instance of the nonlinear timeseries it should learn to predict. Both the states and inputs to the system are stored, and this stage ends with the output layer being trained with some form of linear regression.

- **Warm-up / touching**: The reservoir states are reset, and another instance of the nonlinear timeseries (the test data) is provided as input to the reservoir for a limited number of timesteps. This is done so that the system's states have time to reflect an input series and not its own initial values.

- **Prediction / going**: The output of the reservoir is now looped back on itself and given as input. Now it is running essentially as a closed dynamical system, and its outputs should be considered its predictions of how it "thinks" the test timeseries will continue throughout time.

The *touch and go* terminology is to invoke an intuition of gradually letting go of some system and letting it run on its own after guiding it for a while. Concretely, we use the following update rule for the reservoir states:

$$res(t+1) = tanh(W_{res} * res(t) + W_{in} * input) \tag{5.10}$$

where res is a vector containing the reservoir states, and the following training method:

$$W_{out} = (H_{states}^T * H_{states} + rI)^{-1} * (H_{states}^T * H_{inputs}) \tag{5.11}$$

To explain the variables in order: H_{states} is a matrix representation of the history of states of the reservoir, accumulated in stage 1, and given by 5.10. r is the ridge parameter, which multiplies the identity matrix I. Ridge regression is helpful to avoid overfitting. H_{inputs} is a history of system inputs.

Reservoir computer in terms of Poly

We will show an implementation of a reservoir computer that learns the dynamics of the Lorenz system, whose implementation we have already shown. The dashed box in Figure 5.7 is reminiscent of a polynomial as an interface, and the system taking in inputs from different sources depending on its outputs suggests a mode-dependent dynamical system. Mode-dependence on the type level is not needed, since the systems in question always receive inputs and provide outputs of the same types, but the wiring pattern changes through time. Our situation is with the three wiring patterns below, each corresponding to one of the stages mentioned:

`draw`

Not all of the Agda code is included, since it is quite big and again, readers can consult the implementation in `Dynamical/Reservoir/ModeDependent.agda`. But some key areas to highlight are the following definitions.

Firstly, the state of the reservoir system:

```
data ReservoirState (numNodes : ℕ) (systemDim : ℕ) : Set where
  Coll : (nodeStates : Vec ℝ numNodes)
         (counter : ℕ)
         (statesHistory : Vec (Vec ℝ numNodes) counter)
         (systemHistory : Vec (Vec ℝ systemDim) counter) →
         ReservoirState numNodes systemDim
  Touch : (nodeStates : Vec ℝ numNodes)
          (counter : ℕ)
          (outputWeights : OutputWeights numNodes systemDim)
          →
          ReservoirState numNodes systemDim
  Go : (nodeStates : Vec ℝ numNodes)
       (outputWeights : OutputWeights numNodes systemDim)
       →
       ReservoirState numNodes systemDim
```

The reservoir ought to start in the *collection* state, where it will accumulate some amount of states AND system history, as well as keep a current set of states for the

nodes. It then goes into the *touching* state, where it has a separate counter keeping track of how many state updates to perform before starting prediction. Although the touching state doesn't use the trained output weights obtained from the collection state, it needs to carry it around in order to pass it to the *going* state, where the reservoir is now only concerned with its dynamics.

Second, the output of the reservoir:

```
data ReservoirOutput (systemDim : ℕ) : Set where
    stillColl : ReservoirOutput systemDim
    stillTouch : ReservoirOutput systemDim
    predicting : Vec ℝ systemDim → ReservoirOutput systemDim
```

A regular sum type, since we don't care about what the output of the system is until it's actually ready to predict.

Thirdly, the parallelizing of the systems in question. The interesting thing about this is the `reservoir` function, which takes some parameters to construct a `DynamicalSystem`. Here it is made clear that the input and reservoir weights are fixed:

```
preLorRes : (numNodes trainingSteps touchSteps : ℕ) →
            (dt : ℝ) →
            InputWeights numNodes 3 →
            ReservoirWeights numNodes →
            DynamicalSystem
preLorRes numNodes trainingSteps touchSteps dt inputWeights reservoirWeights =
  -- Training system
  lorenz dt &&&
  -- Test system
  lorenz dt &&&
  -- Reservoir of dynamics + readout layer
  reservoir numNodes 3 trainingSteps touchSteps inputWeights reservoirWeights
```

The types `InputWeights` and `ReservoirWeights` are just synonyms for matrices. Finally, the wiring pattern lens, which captures the changing inputs depending on system outputs:

```
lrWiringDiagram : (numNodes trainingSteps touchSteps : ℕ) →
                  (dt : ℝ) →
                  (iw : InputWeights numNodes 3) →
                  (rw : ReservoirWeights numNodes) →
                  Lens (interface (preLorRes numNodes
                                            trainingSteps
                                            touchSteps
                                            dt
                                            iw
                                            rw))
                  (emitter (ReservoirOutput 3 × (X × Y × Z)))\
lrWiringDiagram numNodes trainingSteps touchSteps dt iw rw =
    outerOutputsFrom ⇆ innerInputsFrom
```

```
where outerOutputsFrom : (X × Y × Z) ×
                           (X × Y × Z) ×
                           ReservoirOutput 3 →
                           ReservoirOutput 3 × (X × Y × Z)
      outerOutputsFrom (_ , test , ro) = ro , test
      -- Provide inputs from different sources
      -- depending on reservoir's output
      innerInputsFrom : (X × Y × Z) ×
                          (X × Y × Z) ×
                          ReservoirOutput 3 →
                          ⊤ →
                          (⊤ × ⊤ × Vec ℝ 3)
      innerInputsFrom (lorOut , lorTestOut , stillColl) tt =
        tt , tt , Lorenz.outToVec lorOut
      innerInputsFrom (lorOut , lorTestOut , stillTouch) tt =
        tt , tt , Lorenz.outToVec lorTestOut
      innerInputsFrom (lorOut , lorTestOut , predicting x) tt =
        tt , tt , x
```

The `outerOutputsFrom` function just always provides the reservoir output along with the test sequence, so as to facilitate exposing the data to a list we can plot. The `innerInputsFrom` function, which takes care of filling inputs, is key to representing the three stages: the input to the reservoir is *its own output* when it's started predicting.

For deep mathematical reasons, reservoirs (or any other model of chaotic systems) cannot perfectly reproduce the learned time series. For an in-depth explanation of this, see [35]. However, they can track the system's dynamics quite well for a while before diverging. Figure 5.8 shows the result of our run over 200 predictions, with 8000 training steps and 600 reservoir nodes - the system, although small, has clearly learned some of the intrinsic properties of the input sequence.
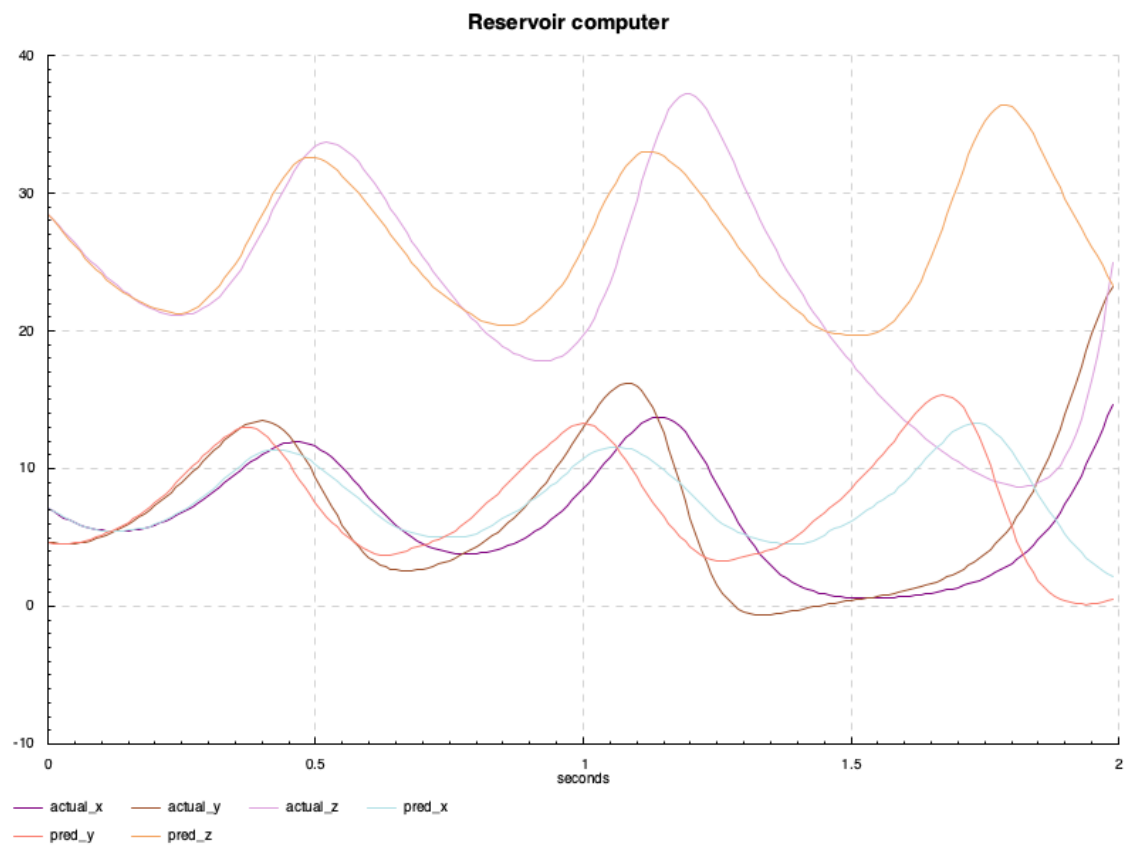
Figure 5.8: Reservoir dynamics: the reservoir's output along with the test sequence it's trying to predict along all dimensions.

6

Discussion

This chapter covers a discussion on certain topics about the formalization of **Poly**. The characterization of equality between lenses is discussed, as well as the ease of use of **Poly** in the implementation of dynamical systems. Further, the consequence of mixing Cubical Agda with agda-categories is discussed. Finally, some areas of future work is given.

6.1   Equality of lenses

The fact that lenses (as well as polynomials and charts) are defined as (dependent) sigma types makes equality of lenses (as well as polynomials and charts) a major pain point. The first component of the sigma type is directly comparable, but for the second component, `subst` is needed to make the types comparable. This problem shows up every time an equality between lenses needs to be proved. If the proof of the first component being equal is `refl`, the `subst` is easily removed by using `substRefl`. But, anytime the proof of the first component is non-trivial, the complexity of the equality proof quickly increases. Therefore, the more powerful variant of lens equality were developed, `lensEquals3`, where the `subst` is pushed inside the term as much as possible. This variant of lens equality showed to be very useful and is used in most of the places in the code. Hopefully, the efforts put in to characterize lens equality is useful to any future formalizations of **Poly**.

> Rename lensE-quals3

6.2   Ease of implementation

The implementation of a couple of dynamical systems has led us, the authors of the thesis, to an impression of how it is to use **Poly** in practical implementation. There are some pain points, such as dealing with time, getting lost in data types, and debugging. There are also some nice points, such as the use of interfaces.

When designing and implementing a dynamical system, it is important to consider the time and flow of information. For example, given an input to a system, the output for that input is not received until the next time step. This creates a delay, which in fact was used by the Fibonacci example, where the identity system simply delayed the input by 1 frame. However, taking into account into these delays in bigger systems creates a lot of complexity to keep track of.

When creating dynamical systems a lot of data types needs to be created, which ends up making it easy to get lost in the details. There needs to be data types for the state of each system, for the positions and output of each system, as well as for each input at each output. On paper, this is easier to handle, since it is possible to abuse notation and leave details out, but in Agda everything is explicit making it a bit verbose. Also, using polynomials created by the operators defined such as addition or parallel product is annoying. In these cases the data types used behind the scenes are coproduct and product, which have general constructor names that is terrible to work with when implementing the dynamical system, or wiring the wiring diagram.

A problem, most likely due to the early state of the ecosystem of **Poly**, is how to debug systems. When the systems didn't work as expected a lot of effort was put into figuring out what was going wrong. A tool more similar to debuggers for imperative languages would have been useful, to step through the dynamical systems and look into their internal state. To make it possible to do some debugging, functionality similar to Haskell's trace was added, as a somewhat better, but still far from perfect way to debug systems.

Explain how we debug systems, better than I do in the last sentence above.

The importance put on interfaces provides a nice way of reasoning about and building systems. Each system is given a polynomial that represent the interface, then wiring systems together in a wiring diagram is done entirely in terms of the interfaces of the involved systems. Different implementations with the same interface is can then be installed as a last step to get the full system. This provides a good level of abstraction when developing programs.

6.3   Mixing Cubical Agda with agda-categories

A big consequence of mixing Cubical Agda with agda-categories is that it hinders the formalization to be merged into either of the repositories. The choice were to use Cubical Agda for its usefulness in proofs, and agda-categories for its richness of categorical constructs. However, Cubical Agda has its own category theory library, and would not want any code using agda-categories to be merged into it. At the same time, agda-categories would not want any code that depends on Cubical Agda. This hinders the code to be merged into any of these repositories.

The code could be modified to be mergable into the Cubical Agda library. All the categorical constructs used from agda-categories could be rewritten to use the Cubical Agda's constructs. The constructs only available in agda-categories could be defined and added to the Cubical library. At the core, all the proofs should remain the same, although they might need some slight modifications. An important detail that is needed to be considered when adapting the code to only use Cubical Agda is the notion of wild polynomials.

6.4 Wild polynomials

There are some requirements differing between providing an instance of a category in agda-categories versus in the Cubical Agda library. One important difference is that Cubical Agda requires the arrows to be sets, using the cubical notion `isSet`. In order to provide a proof that arrows are sets, it is required that the positions and all directions of polynomials are sets. This gives rise to two different variants of polynomials. The normal polynomials, also called wild polynomials, without any requirement of the positions or directinos being sets, used to define a category in agda-categories. As well as the more strict variant of polynomials, called `SetPolynomial`, which is needed to define an instance of a category in Cubical Agda library.

In the code, both variants of polynomials are provided, but wild polynomials are mostly used as they are more general. The use of wild polynomials also avoids the use of `isSet`, which would add noise to the code. However, set polynomials together with an instance of a category for Cubical agda library can be found in the code. In fact, some constructs and proofs require polynomials to be sets, such as the proof showing that **Poly** has equalizers.

Is the terminology of Wild polynomial correct here, and is the name of Set Polynomial correct here as well?

6.5 Future work

There are many areas for future work in formalizing **Poly**.

6.5.1 Comonoids are small categories

The poly-book explains how comonoids in **Poly** are isomorphic to small categories. An interesting thing to do is to show this isomorphism in Agda. An initial attempt exists in the code, with the correspondance of objects with positions, as well as arrows with directions, also dealing with the codomain mismatch. But the correspondence between the comonoid laws and the category laws remains unfinished.

6.5.2 Bicomodules are parametric right adjoints

A construct covered in the poly-book is bicomodules, which relates to comonoids. Bicomodules has a practical use as data migration for databases [36] as well as effect handlers. Therefore, implementation and use of bicomodules should be of special interest for software developers.

6.5.3 Composition of squares between charts and lenses

The commuting squares between charts and lenses have been defined, as `LensChartCommute`. Showing that these squares compose both horizontally and vertically is an important step in finalizing the formalization of the double category between charts and lenses. The proof that the map on positions of the squares compose is done, but the proof for map on directions remains undone, due to an explosion of subst's.

6.5.4 Consistency of distributed data types

There is some speculation that polynomial functors can be a good model of consistent distributed data storage, like Conflict-Free Replicated Data Types, so the categorical setting could be fruitful for proving consistency of CRDTs.

6.5.5 Other categorical properties

The categorical constructs and properties formalized in this thesis comprise only a handful of all the properties that **Poly** exhibits. A natural extension would therefore be to formalize more constructs. For instance the coequalizer, which is described informally in natural language in the poly-book, is an important construct as it shows that **Poly** has all small colimits. The poly-book contains plenty of proofs and examples that could be formalized.

6.5.6 Dynamical systems

This thesis implements several dynamical systems, but it is a large field. Systems come in all sorts of flavors: time delayed, open, closed, chaotic, stable, unstable, discrete, continuous, stochastic, deterministic, hybrid etc. It would be interesting to continue to explore this landscape while leveraging the properties of **Poly** in the dynamical systems. For instance, our use of mode-dependence, in which we believe lies the true potential of **Poly**, is minor. Implementing more types of dynamical systems could also lead to common patterns or problems arising, whose solutions could lead to a better ecosystem in using **Poly** for modeling dynamics.

6.5.7 Interaction between different fields that Poly models

Poly is good at modeling databases, dynamical systems and information theory, three fields that are connected in loose ways. **Poly** could be a setting in which to deepen these connections.

7
Conclusion

This thesis has successfully succeeded in both formalizing a significant amount of theory of **Poly** as well as implementing dynamical systems using the category while relying heavily on the theory. The theory chapters have formalized the categories themselves, initial objects, terminal objects, products, coproducts, composition and parallel product as monoidal structures, exponential object, charts, quadruple adjunction with sets, as well as many more various proofs, examples, and lemmas needed. The implementation part has shown how arrows can be seen as dynamics in dynamical systems, how polynomials act as interfaces, how to wire systems together, how to implement systems, how to install behaviour into wiring diagrams, and how to run systems arriving at concrete programs. How arrows between correspond to different concepts such as DFA's or Moore machines depending on the polynomials of the arrow shows the power polynomials and lenses as an abstract generalization posses. Some of the programs implemented are Fibonacci sequence generators, Turing machines, the Lorenz system, the Hodgkin-Huxley model and reservoir computers. Combining theory and applications, commuting squares between lenses and charts have been used to show how one dynamical system can simulate another for two different examples.

The work of this thesis has succeeded in contributing to the polynomial functors ecosystem, showing the feasibility of formalizing **Poly** and associated categories, as well providing a solid ground to build up more advanced concepts, both theoretical and practical.

Bibliography

[1] M. Scheffer, S. Carpenter, J. A. Foley, *et al.*, "Catastrophic shifts in ecosystems," *Nature*, vol. 413, pp. 591–596, 2001. DOI: 10.1038/35098000. [Online]. Available: https://doi.org/10.1038/35098000.

[2] E. Zafiris, "Complex systems from the perspective of category theory: I. functioning of the adjunction concept," *Axiomathes*, vol. 15, pp. 147–158, 2005. DOI: 10.1007/s10516-005-3790-0. [Online]. Available: https://doi.org/10.1007/s10516-005-3790-0.

[3] N. Ghani, J. Hedges, V. Winschel, and P. Zahn, *Compositional game theory*, 2018. arXiv: 1603.04641 [cs.GT].

[4] P. Perrone, *Markov categories and entropy*, 2022. arXiv: 2212.11719 [cs.IT].

[5] J. C. Baez and B. S. Pollard, "A compositional framework for reaction networks," *Reviews in Mathematical Physics*, vol. 29, no. 09, p. 1 750 028, 2017.

[6] S. Libkind, A. Baas, E. Patterson, and J. Fairbanks, "Operadic modeling of dynamical systems: Mathematics and computation," *Electronic Proceedings in Theoretical Computer Science*, vol. 372, pp. 192–206, Nov. 2022. DOI: 10.4204/eptcs.372.14. [Online]. Available: https://doi.org/10.4204/eptcs.372.14.

[7] [Online]. Available: https://agda.readthedocs.io/en/v2.6.3/getting-started/what-is-agda.html.

[8] W. A. Howard, "The formulae-as-types notion of construction," *To H. B. Curry: Essays on Combinatory Logic, Lambda Calculus and Formalism*, vol. 2, pp. 479–490, 1980.

[9] A. Bove and P. Dybjer, "Dependent types at work," in *Language Engineering and Rigorous Software Development: International LerNet ALFA Summer School 2008, Piriapolis, Uruguay, February 24 - March 1, 2008, Revised Tutorial Lectures*, A. Bove, L. S. Barbosa, A. Pardo, and J. S. Pinto, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 57–99, ISBN: 978-3-642-03153-3. DOI: 10.1007/978-3-642-03153-3_2. [Online]. Available: https://doi.org/10.1007/978-3-642-03153-3_2.

[10] [Online]. Available: https://en.wikibooks.org/wiki/Haskell/GADT#GADTs.

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

[12] Agda, *Agda/cubical: An experimental library for cubical agda*. [Online]. Available: https://github.com/agda/cubical.

[13]   B. Fong and D. I. Spivak, *Seven sketches in compositionality: An invitation to applied category theory*, 2018. arXiv: `1803.05316 [math.CT]`.

[14]   *1lab - a formalised, cross-linked reference resource for cubical methods in homotopy type theory.* Accessed: 2023-05-17, 2023. [Online]. Available: `https://1lab.dev`.

[15]   J. Z. S. Hu and J. Carette, "Formalizing category theory in agda," in *Proceedings of the 10th ACM SIGPLAN International Conference on Certified Programs and Proofs*, ACM, Jan. 2021. DOI: `10.1145/3437992.3439922`. [Online]. Available: `https://doi.org/10.1145/3437992.3439922`.

[16]   N. GAMBINO and J. KOCK, "Polynomial functors and polynomial monads," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 154, no. 1, pp. 153–192, Sep. 2012. DOI: `10.1017/s0305004112000394`. [Online]. Available: `https://doi.org/10.1017/s0305004112000394`.

[17]   J. Kock, *Polynomial functors and trees*, 2009. arXiv: `0807.2874 [math.CT]`.

[18]   D. I. Spivak, *Functorial aggregation*, 2023. arXiv: `2111.10968 [math.CT]`.

[19]   [Online]. Available: `https://github.com/amuricys/polynomial-functors`.

[20]   Iceland_jack, *Iceland_jack's attempt at polynomial functors in haskell*, `https://web.archive.org/web/20230523131853/https://twitter.com/Iceland_jack/status/1646315831402323970`, Accessed: 2023-05-17, 2023.

[21]   S. Visscher, *Sjoerd visscher's attempt at polynomial functors in haskell*, `https://web.archive.org/web/20230523131533/https://gist.github.com/sjoerdvisscher/0fd871b1c2df4f4750945eb08378fb4a`, Accessed: 2023-05-17, 2023.

[22]   D. Vagner, D. I. Spivak, and E. Lerman, *Algebras of open dynamical systems on the operad of wiring diagrams*, 2015. arXiv: `1408.1598 [math.CT]`.

[23]   e. a. Urs Schreiber Mike Shulman, *Concept with an attitude*, Online; accessed 17-May-2023, 2023. [Online]. Available: `https://ncatlab.org/nlab/show/concept+with+an+attitude`.

[24]   D. J. Myers, "Categorical systems theory," Manuscript in preparation, 2022.

[25]   *Hmatrix: A haskell library of linear algebra, numerical computation, and general utilities for scientific computation*, Accessed: 2023-05-17. [Online]. Available: `https://hackage.haskell.org/package/hmatrix`.

[26]   J. D. Murray, *Mathematical Biology: I. An Introduction*, 3rd. Springer, 2002.

[27]   E. N. Lorenz, "Deterministic nonperiodic flow," *Journal of the Atmospheric Sciences*, vol. 20, no. 2, pp. 130–141, 1963.

[28]   A. L. Hodgkin and A. F. Huxley, "A quantitative description of membrane current and its application to conduction and excitation in nerve," *The Journal of physiology*, vol. 117, no. 4, p. 500, 1952.

[29]   W. Gerstner, W. M. Kistler, R. Naud, and L. Paninski, *Neuronal Dynamics: From single neurons to networks and models of cognition.* Cambridge University Press, 2014. [Online]. Available: `https://neuronaldynamics.epfl.ch/index.html`.

[30]   B. Schrauwen, D. Verstraeten, and J. Campenhout, "An overview of reservoir computing: Theory, applications and implementations," Jan. 2007, pp. 471–482.

[31] A. Ferreira, T. Ludermir, R. Aquino, M. Lira, and O. Neto, "Investigating the use of reservoir computing for forecasting the hourly wind speed in short-term," Jun. 2008, pp. 1649–1656. DOI: 10.1109/IJCNN.2008.4634019.

[32] H. Jaeger, "Adaptive nonlinear system identification with echo state networks," in *Proceedings of the 2002 45th IEEE Conference on Decision and Control*, IEEE, vol. 3, 2002, pp. 2696–2701.

[33] M. Cucchi, S. Abreu, G. Ciccone, D. Brunner, and H. Kleemann, "Hands-on reservoir computing: A tutorial for practical implementation," *Neuromorphic Computing and Engineering*, vol. 2, no. 3, p. 032 002, Aug. 2022. DOI: 10.1088/2634-4386/ac7db7. [Online]. Available: https://dx.doi.org/10.1088/2634-4386/ac7db7.

[34] H. Jaeger, "The "echo state" approach to analysing and training recurrent neural networks-with an erratum note," Bonn, Germany: German National Research Center for Information Technology GMD Technical Report, Tech. Rep. 148, 2001.

[35] S. H. Strogatz, *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering*. Westview press, 2001.

[36] [Online]. Available: https://github.com/agda/cubical.

[37] *Nix: The purely functional package manager*, Accessed: 2023-05-17. [Online]. Available: https://nixos.org/nix/.

[38] *Chart: A haskell library for producing 2d charts and plots*, Accessed: 2023-05-17. [Online]. Available: https://hackage.haskell.org/package/Chart.

[39] *Optparse-applicative: Utilities and combinators for parsing command line options*, Accessed: 2023-05-17. [Online]. Available: https://hackage.haskell.org/package/optparse-applicative.

[40] *Glasgow haskell compiler: Gadts*, Accessed: 2023-05-17. [Online]. Available: https://downloads.haskell.org/ghc/latest/docs/users_guide/exts/gadt.html.

[41] *Glasgow haskell compiler: Datakinds*, Accessed: 2023-05-17. [Online]. Available: https://downloads.haskell.org/ghc/latest/docs/users_guide/exts/data_kinds.html#extension-DataKinds.

Bibliography

A

Appendix 1: Haskell supporting code

We use the Agda FFI to interact with Haskell and its ecosystem for three main purposes: fast matrix inversion, plotting graphs, and a convenient command-line interface for running the continuous dynamical systems and plotting them. The way we plug the Haskell code and its ecosystem with the Agda code is by using the Nix package manager[37] and its tools.

A.1 Matrix inversion

In the reservoir example 5.3.2.4, many matrix operations are used. Many of these are implemented directly in Agda - see the `Dynamical/Reservoir/Matrix` folder - using facilities from the standard library's `Data.Vec` type, but inverting a matrix is an operation that requires a sophisticated algorithm that is quite performance-intensive. For this, we decided to rely on an existing implementation in the Haskell ecosystem, using the HMatrix [25] library. This is done by declaring a postulate and using the `FOREIGN` pragma in the Agda module, and accessing the Haskell module by importing it in the generated Haskell code:

```
...
{-# FOREIGN GHC
import HsMatrix
#-}

postulate
  invertMatrixAsListTrusted : List (List ℝ) → List (List ℝ)
{-# COMPILE GHC invertMatrixAsListTrusted = invertMatrixAsList #-}
...
```

The matrix is temporarily represented as a list of lists, where each list is a row, to simplify translating the matrix dimensions to the Agda level - this way, we only need to care about type conversions directly in Agda. The code in the Haskell module is simply this:

```
...
import Numeric.LinearAlgebra qualified as HMat
```

```
...
invertMatrixAsList :: [[Double]] -> [[Double]]
invertMatrixAsList = HMat.toLists . HMat.pinv . HMat.fromLists
```

We then convert the matrix back to our Agda representation - a length-indexed vector of length-indexed vectors, by using the `trustMe` construct from the standard library to convince Agda that the matrix has the right dimensions.:

```
...
open import Relation.Binary.PropositionalEquality.TrustMe
...
fl : ∀ {A : Set} {n} → (l : List A) → {l : L.length l ≡ n} → Vec A n
fl l {refl} = fromList l

fromListOfLists : ∀ {n m} →
    (l : List (List ℝ)) →
    {p₁ : L.length l ≡ n} →
    {p₂ : M.map L.length (L.head l) ≡ M.just m}
    → Matrix ℝ n m
fromListOfLists [] {refl} = 𝕄 []
fromListOfLists (x :: xs) {refl} {p} =
    𝕄 (fl x {maybepr p} ::
            (V.map (\x → fl x {trustMe}) $ fl xs {trustMe}))
    where maybepr : ∀ {m n} → M.just m ≡ M.just n → m ≡ n
          maybepr refl = refl

_⁻¹ : ∀ {n : ℕ} → Matrix ℝ n n → Matrix ℝ n n
_⁻¹ {n} (𝕄 m) =
  let asList = toList $ V.map toList m
      inverted = invertMatrixAsListTrusted asList
  in fromListOfLists inverted {trustMe} {trustMe}
infixl 40 _⁻¹
```

In this way, the `_⁻¹` operation can be made to take in square matrices and return square matrices of the same dimension.

A.2   Graph plotting

The figures in the continuous systems are all generated via the Haskell library Chart [38]. This case is similar to the above, with the simplification that there is no need to use `trustMe`, instead the Agda postulate for the plotting function runs in IO. There is the small issue of implementing our own product type to guarantee the FFI that there is no dependent typing involved. The postulate code is as follows:

```
open import Data.Product as P hiding (_×_) renaming (_,_ to _,p_)
record _×_ (A B : Set) : Set where
  constructor _,_
  field
```

```
    fst : A
    snd : B

fromSigma : {A B : Set} → A P.× B → A × B
fromSigma ( a ,p b ) = a , b

postulate
  plotDynamics : String →
                 String →
                 String →
                 Float →
                 List (String × List Float) → IO ⊤

{-# FOREIGN GHC import HsPlot #-}
{-# COMPILE GHC plotDynamics = plotToFile #-}
{-# COMPILE GHC _×_ = data (,) ((,)) #-}
```

The corresponding Haskell code is small enough that we include the entire module:

```haskell
-- Dynamical/Plot/src/HsPlot.hs
{-# LANGUAGE BlockArguments #-}
{-# LANGUAGE ImportQualifiedPost #-}
{-# LANGUAGE OverloadedStrings #-}

module HsPlot where

import Graphics.Rendering.Chart.Easy
import Graphics.Rendering.Chart.Backend.Cairo
import Control.Monad (forM_)
import Data.Text qualified as T

plotToFile :: T.Text ->
              T.Text ->
              T.Text ->
              Double ->
              [(T.Text, [Double])] ->
              IO ()
plotToFile xaxisTitle yaxisTitle title dt lines =
    toFile def (T.unpack . T.replace " " "_" $ T.toLower title <> ".png") $
    do
        layout_title .= T.unpack title
        layout_x_axis . laxis_title .= T.unpack xaxisTitle
        layout_y_axis . laxis_title .= T.unpack yaxisTitle
        setColors . fmap opaque $
            [purple, sienna, plum,
             powderblue, salmon, sandybrown,
             cornflowerblue, blanchedalmond,
```

```
                firebrick, gainsboro, honeydew]
    forM_ lines \(name, l) ->
        plot (line (T.unpack name) [zip [0, dt..] l ])
```

The string arguments are given as `Data.Text.Text`, even though `Chart` handles `String`s, because that is what Agda's `Data.String.String` corresponds to; see the table in the official documentation.

A.3   Command-line interface

Finally, there is the issue being able to conveniently run a system of choice, with a given selection of parameters. This uses optparse-applicative [39], a well-established Haskell library. The FFI code looks much the same as above, but the main interest now is to convert between types that are pattern-matched on in the main Agda program. The Agda code, with the type containing the paramaters of the different systems is this

```
{-# FOREIGN GHC
import OptparseHs
#-}

data SystemParams : Set where
  ReservoirParams : (rdim : ℕ) →
                    (trainSteps touchSteps outputLength : ℕ) →
                    (lorinitx lorinity lorinitz dt variance : Float) →
                    SystemParams
  LorenzParams       : (lorinitx lorinity lorinitz dt : Float) → SystemParams
  HodgkinHuxleyParams : (dt : Float) → SystemParams
  LotkaVolterraParams : (α β δ γ r0 f0 dt : Float) → SystemParams
{-# COMPILE GHC SystemParams = data SystemParams
    (ReservoirParams |
     LorenzParams |
     HodgkinHuxleyParams |
     LotkaVolterraParams) #-}

record Options : Set where
  constructor mkopt
  field
    system  : SystemParams

{-# COMPILE GHC Options = data Options (Options) #-}

postulate
  parseOptions : IO Options
{-# COMPILE GHC parseOptions = parseOptions #-}
```

And its corresponding Haskell type is:

```haskell
data SystemParams where
  ReservoirParams ::
    { rdimf         :: Integer
    , trainStepsf   :: Integer
    , touchStepsf   :: Integer
    , outputLengthf :: Integer
    , lorinitx      :: Double
    , lorinity      :: Double
    , lorinitz      :: Double
    , dt            :: Double
    , variance      :: Double
    } -> SystemParams
  LorenzParams ::
    { lorinitx' :: Double
    , lorinity' :: Double
    , lorinitz' :: Double
    , dt'       :: Double
    } -> SystemParams
  HodgkinHuxleyParams ::
    { dt'' :: Double }
    -> SystemParams
  LotkaVolterraParams ::
    { alpha :: Double
    , beta :: Double
    , delta :: Double
    , gamma :: Double
    , r0 :: Double
    , f0 :: Double
    , dt''' :: Double }
    -> SystemParams
deriving instance Show SystemParams
```

The code for parsing this type looks like this:

```haskell
systemParser :: Parser SystemParams
systemParser = hsubparser $
    command "Reservoir"
       (info reservoirParamsParser (progDesc "Reservoir system"))
  <> command "Lorenz"
       (info lorenzParamsParser (progDesc "Lorenz system"))
  <> command "HodgkinHuxley"
       (info hodgkinHuxleyParamsParser (progDesc "Hodgkin-Huxley system"))
  <> command "LotkaVolterra"
       (info lotkaVolterraParamsParser (progDesc "Lotka-Volterra system"))

reservoirParamsParser :: Parser SystemParams
```

```haskell
reservoirParamsParser = ReservoirParams
  <$> option auto
    (long "numNodes" <> short 'n' <> metavar "NUMNODES" <>
        help "Number of nodes in the reservoir")
  <*> option auto
    (long "trainingSteps" <> short 't' <> metavar "TRAINSTEPS" <>
        help "Number of data points to train on")
  -- ... other params

lorenzParamsParser :: Parser SystemParams
lorenzParamsParser = LorenzParams <$> -- ... params
-- other systems with a similar type signature
```

It is worth remarking here that Haskell and optparse-applicative allow for a concise expression of more type safety than a simple sum type, through the use of the language extensions for GADTs [40] and DataKinds [41], by having the `SystemParams` type be parametrized by the kind of a separate type called `ParamType`:

```haskell
data ParamType
    = ReservoirParamsType
    | LorenzParamsType
    | HodgkinHuxleyParamsType
    | LotkaVolterraParamsType
  deriving Show
```

Which we can then be given as an argument to `SystemParams`:

```haskell
data SystemParams (p :: ParamType) where
  ReservoirParams ::
    { rdimf         :: Integer
    , trainStepsf   :: Integer
    , touchStepsf   :: Integer
    , outputLengthf :: Integer
    , lorinitx      :: Double
    , lorinity      :: Double
    , lorinitz      :: Double
    , dt            :: Double
    , variance      :: Double
    } -> SystemParams ReservoirParamsType
    ... other constructors
```

This way, the individual parsers can be made to only parse specific sets of parameters. Forgetting an argument would give a compilation error:

```haskell
lorenzParamsParser :: Parser (SystemParams LorenzParamsType)
lorenzParamsParser = LorenzParams
  <$> option auto (long "x0" <> short 'x' <> metavar "X0" <>
    help "Initial x value")
  <*> option auto (long "y0" <> short 'y' <> metavar "Y0" <>
```

```
  help "Initial y value")
-- <*> option auto (long "z0" <> short 'z' <> metavar "Z0" <>
--  help "Initial z value") error! (SystemParams LorenzParamsType)
-- expects this float
<*> option auto (long "dt" <> short 't' <> metavar "DT" <>
  help "Time step")
```

This is a preferable solution in the Haskell level, but converting the types given as arguments to a higher-kinded type in Agda to the DataKinds-produced types proved tricky, so we opted for the simpler solution of a pure sum type.