Dakka: A dependently typed Actor framework for Haskell

Philipp Dargel

2018-12-04

Contents

1 Introduction

The goal of this thesis is to explore how Haskell's type system can be leveraged to create an Actor frameworks, similar to Akka, that allows the user to better reason about the runtime behavior of the system. Haskell provides many tools in its type system that together with Haskell's purely functional nature enables us to formulate more strict constraints on Actor systems. To formulate these constraints I will leverage some of Haskell's dependent typing features. Another focus of the thesis is the testability of code written using the created framework.

I will show that leveraging Haskell's advantages can be used to create an Akka like Actor framework that enables the user to express many constraints inside the type system itself that have to be done through documentation in Akka. The implementation of the Actor framework and important design decisions will be discussed in detail. I will also show that excessive usage of the type system has some downsides that mostly relate to the maturity of Haskell's dependent typing features.

1.1 Motivation

Parallel programming plays an increasing role in the current tech environment and will become even more so. Single processor cores are not getting significantly faster, instead more cores are added to a single processor. As a result efficiently utilizing contemporary processors requires workloads to be processed in parallel. When dealing with parallelism the type system normally is not helpful in preventing bugs.

Distributed systems are also becoming more and more important with the advent of the Internet of Things (IoT). For IoT devices, it becomes even more important that software is bug free. Depending on the kind of deployment it may be hard to debug a device. Rolling out patches is also a hard task that would be nice to avoid if possible.

Actor systems provide a way of modeling programs that is particularly suited for parallel and distributed execution. There are many systems that use Actor systems for exactly that purpose. The Erlang language and the Akka framework for the Java Virtual Machine (JVM) are two of the most prominent examples.

Types may be used to check the behavior of a single actor internally, but are not often used to check properties about the Actor system as a whole. The lack of these global checks may make it possible for example to send messages to Actors that could never exist in the Actor system, or send messages to Actors that those can not handle.

Haskell provides a strong type system that can be used to express these kinds of invariants. Unfortunately the only major Actor framework for Haskell, *cloud-haskell* does not utilize it to do so.

1.2 Goals

I want to create an Actor framework for Haskell that leverages the type system to constrain Actor behaviors to minimize unexpected side effects. The main issue the type system can help in is ensuring that only messages can be sent that can be handled by the receiving Actor. It should ideally be possible for the user to add further constraints on messages and Actors or other parts of the system as they choose.

Runtime components of this Actor framework should be serializable. Serializability is desirable since it aids debugging, auditing, distribution and resilience. Debugging and auditing are aided by Serializability since we could store relevant parts of the system to further review them independent of the runtime environment. If we can store the state of the system we can also recover the whole system or parts of it by simply restoring a previous system state. These states could then also be sent to different processes or machines to migrate Actors from one node to another.

1.3 Result

I explored many aspects of Haskell's type system and dependent typing features and how to apply them to the domain of Actor frameworks. As a result I created an API that fulfills many of the target features. Actors implemented in the created API can be executed in a test environment and to a certain degree in a distributed environment. Since the main focus was the API and how to constrain Actors written with it, the runtime aspect is not yet fully implemented.

2 Fundamentals

2.1 Actor Model

The Actor Model is a way of modeling concurrent computation where the primitive of computation is called an Actor. A finite set of Actors that can communicate with each other is an Actor System. Actors can receive messages and are characterized by the way they respond to these Messages. In Response to a message an Actor may:

1. Send a finite number of messages to other Actors inside the same Actor System.
2. Add a finite number of new Actors to the Actor System.
3. Designate the behavior to be used for the next message it receives.

The Actor Model keeps these definitions very abstract. The high degree of abstraction makes Actor identification inside an Actor system and message ordering details of implementation.

2.2 Akka

Akka is an implementation of the Actor Model written in Scala for the Java Virtual Machine[Akka-docu]. Akka is not a straight implementation of the Actor Model. In some cases Akka deviates from the classic Actor Model, described above. All these change are made with care and make Akka more suited for real world use cases.

In Akka Actors are represented as classes that extend a common base class. When an Actor is created a new instance of the Actor's class is created. The Actor class' constructor may require additional arguments. Constructor arguments have to be supplied when an Actor is created. Actor classes have to provide an initial `receive` property which represents the Actors initial behavior. The type of the `receive` property is `PartialFunction[Any, Unit]` which means it's a possibly partial function that takes arguments of any type and returns a unit. An Actor class may have fields which represent internal state. In addition to fields they inherit the `become` method which provides a way to switch the behavior of the current Actor. Inside of its behavior the Actor has access to a reference to itself as well as to the sender of the currently handled message. Inside an Actor system messages of any type can be sent to any reference. There is a special message called `PoisonPill` which will terminate an Actor when received. When an Actor terminates it's designated supervisor is notified. Normally an Actors supervisor is the Actor that created it.[1]

In addition to these foundational Actors Akka provides more features for Actors like control over Actors mailboxes[2], message routing[3], clustering of Actor Systems[4] and more.

The way Akka is implemented distinguishes it from the traditional Actor Model in some cases and extends it:

- Actors have two kinds of state: The internal state of the Actor class instance and the current `Receive` behavior.
- A strict order on messages is enforced. For every pair of Actors in the Actorsystem it is ensured that messages from one of those Actors to the other are handled in the same order they were sent. A notable exception to this is the `Kill` message which terminates an Actor as soon as possible.
- Actors are named when they are created.
- Each Actor has access to the current Actor system via the `context` property. This gives any Actor access to every other Actor in the current Actor system. Actors can be enumerated or searched for by path.
- When an Actor terminates a message is sent to it's supervisor(s).
- Since Scala is not a pure language Actors can perform arbitrary operations in response to their behavior. As a result it is not possible to constrain the behavior of a piece of Scala code through its type. There is always a way to launch the missiles.
- Akka expects messages to be immutable.

There is also an alternative package to the described Actor base package which adds type information to Actors[5]. The main differences between those two packages are twofold. Actor references are parametrized by the type of message that the defining Actor may handle. Also Actors have to define what kind of message they may receive.

2.3 Cloud Haskell

Cloud Haskell is described by its authors as a platform for Erlang-style concurrent and distributed programming in Haskell.[6][7]

Since Erlang-style concurrency is implemented using the Actor model, Cloud Haskell already provides a fully fledged Actor framework for Haskell. In addition there are rich facilities to create distributed systems in Haskell.

Unfortunately Cloud Haskell has to be somewhat opinionated since some features it provides would not be possible otherwise. The biggest problem is the fact that Haskell does not provide a way to serialize functions at all[7]. Cloud Haskell solves the function serialization problem through the *distributed-static*[8] package, which requires some restrictions in the way functions are defined to work.

2.4 Dependent Typing

A dependent type is a type that depends on a value.[9] Dependent types are a way to express relationships between values inside of a type system. The canonic example for dependent types is a length indexed vector. A length indexed vector is a list which length is derivable from its type. This can be defined as a Haskell GADT[10]:

```
1  data Vec (l :: Nat) (a :: *) where
2      VNil  :: Vec 0 a
3      VCons :: a -> Vec l a -> Vec (l + 1) a
```

Where `Nat` is a kind that represents positive integers as types. A kind can be thought of as a type of types[11]. The kind of complete type in Haskell, like `Bool` is called *[12]. In contrast the type of a type constructor like `Maybe` has the kind * -> *. Where * -> * means that this type constructor will produce a type of kind * if it is provided with a type of kind *. The `DataKinds`[13] language extension provides a way to promote data types to kinds. The value constructors of the promoted datatype become types or type constructors. There are no values associated with the resulting types. The only use of these types is as shadow types or arguments to type families and typeclasses. Take for example the type `data Bool = True | False`. If we promote `Bool` to a kind we can create a version of `Maybe` that keeps the information whether or not it contains a value in its type:

```haskell
1  data CoolMaybe (b :: Bool) a where
2      CoolNothing :: CoolMaybe 'False a
3      CoolJust    :: a -> CoolMaybe 'True a
```

Notice that promoted values are preceded by a `'` to distinguish them from their value counterparts. For example, if we encounter the type `CoolMaybe 'True String` for example we know that a value of this type has to always contain a value of type `String`. We can use this information to create a safe version of `fromJust`:

```haskell
1  fromCoolJust :: CoolMaybe 'True a -> a
2  fromCoolJust (CoolJust a) = a
```

In order to call this function we first have to prove that the first type argument of `CoolMaybe` is `'True`. To define functions that operate on such a type we also have to reflect he transformations in the type.

```haskell
1  firstCool :: CoolMaybe a c -> CoolMaybe b c -> CoolMaybe (Or a b) c
2  firstCool (CoolJust a) (CoolJust _) = CoolJust a
3  firstCool (CoolJust a) CoolNothing  = CoolJust a
4  firstCool CoolNothing (CoolJust a)  = CoolJust a
5  firstCool CoolNothing CoolNothing   = CoolJust a
```

In this example `And` is a typefamily. That is, a function on types:

```haskell
1  type family Or (a :: Bool) (b :: Bool) :: Bool where
2      Or 'True b = 'True
3      Or a 'True = 'True
4      Or a    b = 'False
```

We have to write every possible pattern of value constructors in the definition of `firstCool` because Haskell can't prove that the resulting value has the expected type. For GADTs with more constructors than this becomes tedious. In those cases it might be beneficial to use typeclasses:

```haskell
1  data Proxy (a :: k) = Proxy
2
3  class ShowMEmpties (l :: [*]) where
4      showMEmpties :: Proxy l -> [String]
5  instance ShowMEmpties '[] where
6      showMEmpties _ = []
7  instance (ShowMEmpties as, Show a, Monoid a) => ShowMEmpties (a ': as)
       where
8      showMEmpties _ = show (mempty :: a) : showMEmpties (Proxy :: Proxy
           as)
```

These type classes shift the burden of proof onto the user. Each time you want to call `showMEmpties` you will also have to prove that an instance `ShowMEmpties` exists for the specific `l` you want to use. Nevertheless this is often the only way of working with data kinds.

2.4.1 singletons

Since types and values are fundamentally different in Haskell, there is no native way to demote a promoted data constructor back to a value. That means once we promote `True` to `'True` we can't retrieve a value of type `Bool` from the type `'True` even though the relationship is clear. In fact in the typesystem the kind `Bool` and the type `Bool` don't have any connection anymore. The *singletons* package tries to fix this shortcoming[14]. *singletons* provides a typefamily `Sing` that associates each promoted kind with its original type. For types in Haskell's *base* package *singletons* provides this mapping out of the box. For user defined types *singletons* provides facilities to derive `Sing`. Additionally *singletons* also provides a general way of promoting functions to type families[15].

Since *singletons* provides general ways to promote and demote types and functions the resulting code is quite opaque. As a result any library using the *singletons* library almost certainly will have unreadable compiler errors. Generating compile errors if a certain behavior violates some invariant is often the goal of using dependent types. In my implementation I chose not to use the *singletons* library for that reason. I still heavily relied on ideas the library is based on but performed the promotions by hand.

2.5 mtl Monad classes and Monad-transformers

The *mtl* library provides a suite of classes that generalizes different Monads[16]. A common way of creating Monads for specific use cases is by composing Monad-transformers[16] For example we can compose the `StateT` Monad-transformer with with `Writer` Monad to get a computation that has a state and allows do write some output:

```
1  type MyComp a = StateT Int (Writer [String]) a
```

If we now want to perform an operation on `Writer` we first have to enter the `StateT` Monad. Lifting an operation from the Monad-transformers argument Monad to the constructed outer Monad can be done through a lifting function specific to each Monad-transformer.

```
1  liftMTrans :: Monad m => m a -> MTrans m a
2  liftMTrans = ...
```

These lifting functions are generalized by the `MonadTrans` class provided by the *transformers* package, which is bundled with each GHC.

Explicitly lifting each operation to the appropriate level inside a chain of Monad-transformers is cumbersome. *mtl* Monad classes get rid of lifting altogether, by providing a class for common Monad capabilities. In the `MyComp` example instead of lifting a `Writer` operation inside `StateT` the `MonadWriter` class can be used. The `MonadWriter` class provides the same functions that are used to interact with the `Writer`Monad. The *mtl* package provides instances of `MonadWriter` and other Monad classes for common Monad-transformers in such a way that if a Monad inside the transformer chain provides the classes functionality the whole chain does too.

These Monad classes are not only useful for interacting with Monad-transformers, but also when creating Monads manually. If the created Monad shares some capabilities with a common Monad or Monad-transformer, these capabilities can be expressed by providing an instance of that Monad class for the created Monad.

Another use for these Monad classes is that they allow for a more general expression of monadic code. For example if you express a monadic computation in terms of `MonadWriter` instead of `Writer`, the computation can be used inside of any Monad that implements `MonadWriter`. This computation can now be executed in either the strict or the lazy variant of the `Writer` Monad or any Monad-transformer chain that contains a `Writer` or `WriterT`.

2.6 Haskell Language features

Modern Haskell development involves many language features that are not present in the base language of *Haskell2010*. These features have to explicitly be enabled by enabling language extensions. Especially working with dependent types and using more advanced features of Haskell's type system require many of these language extensions. Language extensions are enabled using `LANGUAGE` pragmas at the beginning of the file for which the extension should be enabled.

- `DataKinds`: Allows data types to be promoted to kinds and value constructors to types[13].
- `TypeFamilies`: Adds the ability to define type and data families. A type family can be thought of as a function on types[17]. Additionally this provides a way to associate types and type families with type classes[17]. When defining an instance of a class all associated types and type families have to be provided with a binding as well.
- `PolyKinds`: Allows mixing different kinds. For example `k` in `l :: [k]` could normally only be of kind `*` but with `PolyKinds` it may be any kind.

These extensions are the foundation for dependent typing in Haskell. This enables the definition of `not` on types of kind `Bool`:

```
1  type family Not (a :: Bool) :: Bool where
2      Not 'True  = 'False
3      Not 'False = 'True
```

Or even `elem`:

```
1  type family Elem (e :: k) (l :: [k]) :: Bool where
2      Elem e (e ': as) = 'True
3      Elem e (a ': as) = Elem e as
4      Elem e '[]       = 'False
```

2.6.1 Heterogeneous collections

Another example of the usage of some of these extensions are heterogeneous lists. That is lists that can hold values of different types at once. This can be achieved by defining a GADT `HList` that is parametrized by a list of types such that each element of `HList` has a corresponding entry in the list of types:

```
1  data HList (l :: [*]) where
2      HNil  :: HList '[]
3      HCons :: a -> HList as -> HList (a ': as)
4  infixr 5 `HCons`
```

With this we can now create Lists with where each element is of a different type:

```
1  l :: HList '[Int, String, Bool]
2  l = 42 `HCons` "Hello World" `HCons` False `HCons` HNil
```

It is also possible to create a lookup function for elements of a given type that is only defined if the list contains an element of that type:

```
1  class HElem e (l :: [*]) where
2      hElem :: HList l -> e
3
4  instance {-# Overlaps #-} HElem e (e ': as) where
5      hElem (HCons e _) = e
6  instance {-# Overlappable #-} HElem e as => HElem e (a ': as) where
7      hElem (HCons _ as) = hElem as
```

Unlike the previous example a type class is used instead of a type family. Matching rules differ between type families and type classes. Type families allow Non-Linear Patterns, that is the same variable may occur multiple times inside of the pattern, but type classes do not. Type classes are matched exclusively by structure. As a result both instance declarations of `HElem` look the same to compiler.

Constraints are only checked after the compiler already committed to a given declaration. In this context `HElem e (e ': as)` is equivalent to `(e ~ a)=> HElem e (a ': as)`. To prioritize which instance declaration will be chosen by the compiler the instances have to be annotated with overlapping instance pragmas.

2.6.2 Heterogeneous Maps

A heterogeneous map may hold values of different types at once. A type of a value is determined by the type of the key it is associated with. The easiest way to associate a value type with a key is to parametrize the key by the type of the value. The map itself is parametrized by the type of key used. A lookup function may then have the signature `lookup :: k v -> HMap k -> Maybe v` and `insert :: k v -> v -> HMap k -> HMap k`.

There are ways to implement a completely typesafe variant of `HMap`, but if there is no way of manipulating the map directly it is safe to use `unsafeCoerce` as long as the API is safe.

The base for this `HMap` will be a standard `Data.Map.Map`. To be able to use that map both keys and values have to be of a single type. This can be achieved by creating custom `Key` and `Elem` types that capture and hide the concrete value type.

```
1  data Key k where
2      Key :: k a -> Key k
3
4  data Elem where
5      Elem :: a -> Elem
6
7  newtype HMap k = HMap (Map (Key k) Elem)
8      deriving Eq
```

To be able to use `Key` and `Elem` as key and value of `Map Key` has to implement `Ord`. Additionally we need equality on `HMap` for which both `Key` and `Elem` have to implement `Eq`.

To implement either `Eq` or `Ord` it is necessary to have an instance `Ord (k a)` for all `a`. Unfortunately it is not possible to use the `forall` keyword in the context of instance declarations (yet [18]). A work around until `GHC 8.6` is to capture all commonly used classes inside of the `Key` and `Elem` constructors.

```
1  data Key k where
2      Key :: (Typeable (k a), Ord (k a), Show (k a))
3          => k a -> Key k
4
5  instance Show (Key k) where
```

```
 6        showsPrec d (Key k) = showsPrec d k
 7
 8  instance Eq (Key k) where
 9      Key a == Key b = Just a == typeRep b
10
11  instance Ord (Key k) where
12      -- first order by type then, if type are the same use Ord
13      Key a `compare` Key b = typeRep [a] `compare` typeRep [b]
14                           <> Just a      `compare` cast b
```

2.6.3 Typeable

In the process of compiling Haskell, all type information is removed since it is not needed at run-time[19]. Type information may be useful at runtime. If a type is hidden via existential quantification it may be useful to be able to get a `String` representation of the captured type for debug and/or `Show` purposes for example. Without some way of retrieving type information at runtime it would also be impossible to define an `Eq` instance for data types using existential quantification.

Runtime type information is provided by `Data.Typeable` in Haskell2010. The type class `Typeable` provides a single function `typeRep# :: TypeRep a` where `TypeRep a` is a representation of the type `a`. `typeRep#` and `TypeRep a` are only used internally. The module `Data.Typeable` exports ways to leverage this functionality. GHC will derive an instance of `Data.Typeable` for every data type, type class and promoted data constructors automatically[20]. Manually defining an instance of `Data.Typeable` will cause an error to ensure that the type representation is valid.

Showing a type

When dealing with complex types it is helpful to be able to have a way to print types at runtime. For example when capturing types with existential quantification it is helpful to include the captured type in the String representation of the data type. It is also quite useful to be able to print the type of an Actor, deep inside of an Actor system.

Since `TypeRep` implements `Show` we can print any type at runtime. The `Show` implementation of `TypeRep` does not produce output that is equivalent to the way types are represented in Haskell error messages. This mismatch is partly due to the fact that there is no way to represent type aliases using `TypeRep` and some issues with the `Show` implementation itself[21].

```
1  showsType :: forall a. Typeable a => ShowS
2  showsType = showString "<<"
3            . shows (typeRep (Proxy @a))
```

```
4                .  showString ">>"
```

Dynamic values and type casting

`Typeable` enables the creation of Dynamic values in Haskell. To represent a dynamic value, all we have to do is capture the `Typeable` instance of the given type. Dynamic values are implemented by `Data.Dynmaic` in `base`. To construct a dynamic value `toDyn :: Typeable a => a -> Dynamic` is used. To extract a value `fromDynamic :: Typeable a => Dynmaic -> Maybe a` which only returns a value if the expected type `a` is the same as the captured one. Data extraction is only possible because there are runtime type representations that can be compared.

In the same way values can be extracted from dynamic values, it is possible to define a way to conditionally cast a value of one type to another, as long as those two types are the same, where it is only known at runtime if that is the case:

```
1  cast :: forall a b. (Typeable a, Typeable b) => a -> Maybe b
```

`cast` does not provide a way to actually convert a value of one type to another. It only postpones the type equality check to the runtime.

Take the following function `appendIfString` for example:

```
1  appendIfString :: Typeable a => String -> a -> String
2  appendIfString str a = str ++ (fromMaybe "" (cast a))
```

If `appendIfString` is called with `appendIfString "Hello ""World"` it returns `"Hello World"`, but if it is called with `appendIfString "Hello "42` it returns `"Hello "`.

3 Implementation

3.1 Overview

The API is designed to be close to the API of Akka where appropriate. That means an Actor's behavior is modeled by a function from a message to an action. An Actors action is a Monad where all interactions with other Actors and the Actor system itself are functions that produce values in that Monad.

To be able to perform any type level computations on Actors and Actor systems there has to be some way of identifying specific kinds of Actors by type. Actors have to implement a typeclass `Actor a` where `a` is the type we can use do identify Actors by. The `Actor` class has a single function called

`behavior`, which describes the behavior of the Actor. What kind of messages an Actor can handle and what kind of Actors it may create in response has to be encoded in some way as well.

The Monad which models an Actors action is also a typeclass, that has roughly the form **`class`** `Monad m => ActorContext m`. In contrast `ActorContext` could also be defined as a concrete data type that implements `Monad`. This `ActorContext` is an *mtl* style Monad class, which makes it possible to have different implementations of Actor systems at once. This the creation of one implementation that is meant for testing Actors and another one that actually performs these actions inside of a distributed Actor system. Defining the Monad as an typeclass also makes it possible to use different backends without rewriting the actors themselves. One such backend may be *cloud-haskell*. Implementations for testing are also just different backends in this architecture.

3.2 Actor

Since Akka is not written in a pure functional language, each Actor can also invoke any other piece of code. The implicit capability to perform arbitrary actions is very useful for defining real world systems. So we have to provide a way to perform `IO` actions as well if we want to use this framework in a real world situation. Actors may also want to manage the Actor system itself in some capacities that exceed the Actor model axioms. For example stopping it all together, which also turns out to be very useful.

We need a way to identify specific Actors at compile time to be able to reason about them. The best way to do so is by defining types for Actors. Since Actors have a state this state type will be the type we will identify the Actor with. We could have chosen the message type but the state type is more characteristic.

```
1  data SomeActor = SomeActor
2    deriving (Eq, Show, Generic, Binary)
```

Note that we derive `Generic` and `Binary`. This allows the state of an Actor to be serialized.

An Actor now has to implement the `Actors` type class. On this typeclass we can ensure that the Actor state is serializable and can be printed in human-readable form to be included in error messages and log entries.

```
1  class (Show a, Binary a) => Actor a where
```

The first member of this class will be a type family that maps a given Actor state type (Actor type for short) to a message type this Actor can handle. If the message type is not specified, it is assumed that the Actor only understands `()` as a message.

```
1    type Message a
```

```
2    type Message a = ()
```

Notice that the default definition for `Messag a` is `()`. This default definition is meant for actors that do not actually expect any data inside of their messages, but rather use messages as triggers for their behavior. Another sensible default definition for `Message a` would be `Void`, which would indicate that the actor can not receive any messages at all. If the message type of an Actor is `Void` the only way for that actor to act is on receiving signals.

To be able message sending in distributed systems we have to ensure that they are serializable. They have to fulfill the same constraints as the Actor type itself. For this we create a constraint type alias (through the language extension `ConstraintKinds`):

```
1    type RichData a = (Show a, Binary a)
```

Now the class header can be changed to:

```
1    class (RichData a, RichData (Message a)) => Actor a where
```

Instead of a constraint type alias we could also have used a new class and provided a single `instance (Show a, Binary a) => RichData a`. This would have allowed `RichData` to be partially applied. There is currently no need to do this, since the `RichData` constraint doesn't have to be sent around by itself.

Next we have to define a way for Actors to handle Messages.

```
1    behavior :: Message a -> ActorContext ()
```

`ActorContext` will be a class that provides the Actor with a way to perform its actions.

Additionally we have to provide a start state the Actor has when it is first created:

```
1    startState :: a
2    default startState :: Monoid a => a
3    startState = mempty
```

3.3 ActorContext

We need a way for Actors to perform the Actor operations. The most straightforward way to implement these actions would be to use a Monad transformer for each action. Creating and sending could be modeled with `WriterT [SystemMessage]` where `SystemMessage` encapsulates each both the intent to create an actor as well as sending a message to a specific actor. Changing the internal state

of the actor could be achieved through `StateT s` where `s` is the internal state of the actor or the actor type to be more specific.

But here we encounter several issues:

1. To change the state we must know which Actors behavior we are currently describing. Since the actor type has to be deductible from the context type. That means the type has to be more complex then simply `ActorContext a`.

2. To send a message we must ensure that the target Actor can handle the message. If we allow values of `SystemMessage` to be created freely we can not ensure that the receiving actor can handle the sent message.

3. To create an Actor we have to pass around some reference to the Actor type of the Actor to create.

The first issue can be solved by adding the Actor type to `ActorContext` as a type parameter.

To be able to send a message in a type safe way, we need to retain the Actor type. If we would make the Actor type explicit in the `WriterT` type though, we would only be able to send messages to Actors of that exact type. Luckily there is a way to get both. Using the language extension `ExistentialQuantification` we can capture the Actor type with a constructor without exposing it. To retrieve the captured type you have to pattern match on the constructor. We can also use `ExistentialQuantification` to close over the Actor type in the create case. With this technique we can create a wrapper for send and create actions:

```
1  data SystemMessage
2      = forall a. Actor a => Send (ActorRef a) (Message a)
3      | forall a. Actor a => Create (Proxy a)
4    deriving (Eq, Show)
```

`ActorRef` provides some way to identify an Actor inside an Actor system we will define later. `Proxy a` is just a datatype with a unit constructor and a phantom type, that provides a way to pass references to types around. `SystemMessage` could also have been defined in GADT-notation, which would have been semantically equivalent.

Unfortunately we cannot derive `Generic` for data types that use existential quantification and thus can not get a `Binary` instance for free. But as I will show later we do not need to serialize values of `SystemMessage` so this is fine for now.

With all this information we can define `ActorContext` as follows:

```
1  newtype ActorContext a v
2      = ActorContext (StateT a (Writer [SystemMessage]) v)
3    deriving
4      ( Functor
```

```
5        , Applicative
6        , Monad
7        , MonadWriter [SystemMessage]
8        , MonadState a
9        )
```

Notice that we only need one `Writer` since we combined create and send actions into a single type. Since `ActorContext` is nothing more than the composition of several Monad transformers it is itself a Monad. Using `GeneralizedNewtypeDeriving` we can derive several useful Monad instances. The classes `MonadWriter` and `MonadState` are provided by the `mtl` package.

Since we added the Actor type to the signature of `ActorContext` we need to change definition of `behavior` to reflect this:

```
1    behavior :: Message a -> ActorContext a ()
```

By deriving `MonadState` we get a variety of functions to change the Actors state. The other Actor actions can now be defined as functions:

3.3.1 send

Since there is an instance for `MonadWriter [SystemMessage]` for `ActorContext`, we can use `tell` from `MonadWriter` to emit `SystemMessage`s.

```
1  send :: Actor a => ActorRef a -> Message a -> ActorContext b ()
2  send ref msg = tell [Send ref msg]
```

Notice that the resulting `ActorContext` does not have `a` as its Actor type but rather some other type `b`. `a` is the type of Actor the message is sent to and `b` is the type of Actor of which the behavior is being described. The `send` function does not have an `Actor b` constraint since this would needlessly restrict the use of the function itself. When defining an Actor it is already ensured that whatever `b` is, it will be an `Actor`.

We can also provide an akka-style send operator as a convenient alias for `send`:

```
1  (!) = send
```

3.3.2 create

As with `send` we can use `MonadWriter` to emit `SystemMessage`s.

```
1  create' :: Actor b => Proxy b -> ActorContext a ()
2  create' b = tell [Create b]
```

As indicated by the `'` this version of create is not intended to be the main one. For that we define:

```
1  create :: forall b a. Actor b => ActorContext a ()
2  create = create' (Proxy @b)
```

In combination with `TypeApplications` this enables us to create Actors by just writing `create @TheActor` wich shortens the ordinary `create'(Proxy :: Proxy TheActor)`.

3.3.3 ActorRef

We need a way to reference Actors inside an Actor system. The most straightforward way to do this is by creating a data type to represent these references. This type also has to hold the Actor type of the Actor it is referring to. But how should we encode the Actor reference? The simplest way would be to give each Actor some kind of identifier and just store the identifier:

```
1  newtype ActorRef a = ActorRef ActorId
```

References of this kind cannot be be created by the user since you should not be able to associate any `ActorId` with any Actor type, since there is no way of verifying that a given id is associated a given Actor type at compile time. The best way to achieve this is to modify the signature of `create` to return a reference to the just created Actor.

```
1  create :: forall a. Actor a => ActorContext b (ActorRef a)
```

Additionally it would be useful for Actors to have a way to get a reference to themselves. We can give Actors a way to refer to themselves by adding:

```
1  self :: ActorContext a (ActorRef a)
```

To `ActorContext`.

Composing references

If we assume that a reference to an Actor is represented by the Actors path, relative to the Actor system root, we could in theory compose Actor references or even create our own. To allow for Actor reference composition in a typesafe manner we need to know what Actors an Actor may create. To expose which Actors an Actor may create, we add a new type family to the `Actor` class.

```
1      type Creates a :: [*]
2      type Creates a = '[]
```

The type family `Create` has the kind `[*]`, which represents a list of all Actor types Actor `a` can create. We additionally provide a default that is the empty list. So if we do not override the `Creates` type family for a specific Actor, we assume that this Actor does not create any other Actors.

We can now also use the `Create` typefamily to enforce the assumption on the `create'` and `create` functions that the type of any Actor created by an Actor has to be present in the `Create` a list.

```
1  create' :: (Actor b, Elem b (Creates a)) => Proxy b -> ActorContext a
       ()
```

Where `Elem` is a type family of kind `k -> [k] -> Constraint` that works the same as `elem` only on the type level.

```
1  type family Elem (e :: k) (l :: [k]) :: Constraint where
2      Elem e (e ': as) = ()
3      Elem e (a ': as) = Elem e as
```

There are three things to note with The `Elem` type family:

1. It is partial. It has no pattern for the empty list. Since it's kind is `Constraint` this means, the constraint is not met if we would enter that case either explicitly or through recursion.
2. The first pattern of `Elem` is non-linear. That means that a variable appears twice. `e` appears as the first parameter and as the first element in the list. This is only permitted on type families in Haskell. Without this feature it would be quite hard to define this type family at all.
3. We leverage that n-tuples of `Constraints` are `Constraints` themselves. In this case `()` can be seen as an 0-tuple and thus equates to `Constraint` that always holds.

The `Creates` typefamily is useful for defining assumptions that concern the hierarchy of the Actorsystem. For example we can formulate an assumption that states that all Actors in a given Actor system fulfill a certain constraint.

```
1  type family AllActorsImplement
2    (c :: * -> Constraint) (a :: *) :: Constraint where
3      AllActorsImplement c a
4          = (c a, AllActorsImplementHelper c (Creates a))
5  type family AllActorsImplementHelper
6    (c :: * -> Constraint) (as :: [*]) :: Constraint where
7      AllActorsImplementHelper c '[]       = ()
8      AllActorsImplementHelper c (a ': as)
```

```
9              = (AllActorsImplement c a, AllActorsImplementHelper c as)
```

We can also enumerate all Actor types in a given Actor system.

What we can't do unfortunately is to create a type of kind `Data.Tree` that represents the whole Actor system since it may be infinite. The tree representation of the following example would be infinite.

```
1  data A = A
2  instance Actor A where
3      type Creates A = '[B]
4      ...
5
6  data B = B
7  instance Actor B where
8      type Creates B = '[A]
9      ...
```

The type for an Actor system that starts with `A` would have to be `'Node A '[Node B '[Node A '[...]]]`. We can represent any finite path inside this tree as a type.

Since any running Actor system has to be finite we can use the fact that we can represent finite paths inside an Actor system for our Actor references. We can parametrize our Actor references by the path of the Actor that it refers to.

The Actor type is not sufficient to refer to a given Actor. Since an Actor may create multiple Actors of the same type you also need a way to differentiate between them in order to reference them directly. The easiest way would be to order created Actors by creation time and use an index inside the resulting list. There are two problems with this approach. Firstly we lose some type safety since we can now construct Actor references to Actors for which we can not confirm that they exist at compile time. Secondly this index would not be unambiguous since an older Actor may die and thus an index inside the list of child Actors would point to the wrong Actor. We could take the possibility of Actors dying into account by giving each immediate child Actor an unique identifier. Composing an Actor reference would require the knowledge of the exact identifier in that case. Having to know the unique identifier for an Actor to create an Actor reference to it would make composition unfeasible.

I decided to remove the ability to compose Actor references since it would impose to many restrictions onto the form that Actor references could take. Furthermore the usability would be potentially limited.

Type families created in the process of implementing composition are still useful for other purposes. These type families allow type level computation on specific groups of Actors deep inside of an Actor system.

Implementation specific references

Different implementations of `ActorContext` might want to use different data types to refer to Actors. Since we do not provide a way for the user to create references themselves we do not have to expose the implementation of these references.

The most obvious way to create implementation specific data types is to associate a given `ActorContext` implementation with a specific reference type. This can be done using an additional type variable on the class, a type family or a data family. Here the data family seems the best choice to represent implementation specific reference data types since it's injective. The injectivity allows us to not only know the reference type from from an `ActorContext` implementation but also the other way round.

```
1        data CtxRef m :: * -> *
```

Additionally we have to add some constraints to `CtxRef`. Since we need to be able to serialize `CtxRef`, equality and a way to show them would also be nice. To ensure that `CtxRef` is serializable we can reuse the `RichData` constraint.

```
1    class (RichData (CtxRef m)), ...) => ActorContext ... where
```

In our simple implementation I'm using a single `Word` as a unique identifier but we can't assume that every implementation wants to use it.

Now we have another problem though. Messages should be able to include Actor references. If the type of these references now depends on the `ActorContext` implementation we need a way for messages to know this reference type. We can achieve this by adding the Actor context as a parameter to the `Message` type family.

```
1    type Message a :: (* -> *) -> *
```

Here we come in a bind because of the way we chose to define `ActorContext`. The functional dependency in `ActorContext a m | m -> a` forces us to create unwieldy typesignatures in this case. It states that we know `a` if we know `m`. This means that if we expose `m` to `Message` the message is now bound to a specific `a`. This is problematic since we only want to expose the type of reference, not the Actor type of the current context to the `Message`. Doing so would bloat every signature that wants to move a message from one context to another with equivalence constraints like:

```
1    forall a b m n. (ActorContext a m, ActorContext b n, Message a m ~
        Message b n) => ...
```

Instead we add the reference type itself as a parameter to `Message`. This alleviates the problem only a little bit, since we need the actual `ActorContext` type to retrieve the concrete reference type. So we would only delay the constraint dance and move it a little bit. These constraints would mean many

additional type parameters to types and functions that do not actually care about them. Compile errors would also come more cluttered, without adding useful information to the user.

Due to all of the stated concerns, I decided against the idea of `ActorContext` implementation specific reference types.

Instead of trying to create different representations for Actor references I chose to represent them using a `ByteString`. Since Actor references have to be serializable anyway we can represent them by a `ByteString`.

```
1  newtype ActorRef a = ActorRef ByteString
```

This might go a little against our ideal, to keep the code as typesafe as possible, but in this case the trade off should be considered acceptable. Firstly other data types that might have taken the place of `ByteString` would not be any safer. We can still keep the user from being able to create references by themselves by not exporting the `ActorRef` constructor. We could expose it to `ActorContext` implementers through an internal package.

Sending references

A core feature that is necessary for an Actor system to effectively communicate is the ability to send Actor references as messages to other Actors.

The most trivial case would be that the message to the Actor is an Actor reference itself.

```
1  instance Actor Sender where
2      type Message Sender = ActorRef Receiver
3      ...
```

This way limits the Actor type of the receiver to be a single concrete type. In particular we have to know the type of the Actor (`Receiver` in the following) when defining the Actor that handles the reference (`Sender` in the following). So we would like this reference type to be more generic. A simple way to do this is to add a type parameter to the Sender that represents the `Receiver`.

```
1  instance (Actor a, c a) => Actor (Sender a) where
2      type Message (Sender a) = ActorRef a
3      ...
```

`c` may take any constraint that the `Receiver` Actor has to fulfill as well. This is more generic but `a` still represents a concrete type at runtime. The way this is normally done in Haskell is by extracting the commonalities of all `Receiver` types into a typeclass and ensure that all referenced Actors implement that typeclass.

```
1  class Actor a => Receiver a
2  instance Receiver a => Actor (Sender a) where
3      type Message (Sender a) = ActorRef a
4      ...
```

This is a variation on the previous implementation, since we only consolidated `c` into the `Receiver` class. Unfortunately we can not use `forall` in constraint contexts (yet; see `QuantifiedConstraints`). To get around this restriction we can create a new message type that encapsulates the constraint like this:

```
1  data AnswerableMessage c = forall a. (Actor a, c a) =>
       AnswerableMessage (ActorRef a)
```

With this we can define the Sender like this:

```
1  class Actor a => Receiver a
2  instance Actor Sender where
3      type Message Sender = AnswerableMessage Receiver
4      ...
```

`Receiver` should not perform long running tasks, since that would provide a way to circumvent the Actor model somewhat. Since any functions defined on `Receiver` are executed the context of the `Sender`, the message implicitly contains instructions for the `Sender` to run. Ideally the class should only provide a way to construct a message the `Receiver` understands from a more generic type. We can express this with a typeclass like this:

```
1  class Actor a => Understands m a where
2      convert :: m -> Message a
```

A Sender may use this class like this:

```
1  instance Actor Sender
2      type Message Sender = AnswerableMessage (Understands SomeType)
3      onMessage (AnswerableMessage ref) = do
4          ref ! convert someType
```

Solving the problem of sending generic Actor references presents a huge problem though. Using existential quantification prevents `AnswerableMessage` from being serialized. Serializability is a core requirement for messages though.

To serialize arbitrary types we would need some kind of sum-type where each constructor corresponds with one concrete type. Since we can enumerate every Actor type of Actors inside a given

Actor system from the root Actor, we could use this to create a dynamic union type. An example of a dynamic union type would be `Data.OpenUnion` from the `freer-simple` package. To construct this type we need a reference to the root Actor, so that type has to be exposed to the Actor type in some way, either as an additional type parameter to the `Actor` class or to the `Message` typefamily. Adding a type parameter to `Actor` or `Message` would require rewriting ab big chunk of the codebase. Sending `ActorRef` values directly is the only possible way for now.

3.3.4 Flexibility and Effects

By defining `ActorContext` as a datatype, we force any environment to use exactly this data type. This is problematic since Actors can only perform their three Actor actions in the implementation as discussed so far. `ActorContext` is not flexible enough to express anything else. We could change the definition of `ActorContext` to be a Monad transformer over `IO` and provide a `MonadIO` instance. This would defeat our goal to be able to reason about Actors, since we could now perform any `IO` we wanted.

Luckily Haskell's type system is expressive enough to solve this problem. Due to this expressiveness there is a myriad of different solutions for this problem. We will take a look at two approaches that integrate well into existing programming paradigms used in Haskell and other functional languages.

Both approaches involve associating what additional action an Actor can take with the `Actor` instance definition. This is done by creating another associated typefamily in `Actor`. The value of this type-family will be a list of types, that identify what additional actions can be performed. What this type will be depends on the chosen approach. The list in this case will be an actual Haskell list but promoted to a kind. This is possible through the `DataKinds` extension.

mtl style Monad classes

In this approach we use mtl style Monad classes to communicate additional capabilities of the Actor. This is done by turning `ActorContext` into a class itself where `create` and `send` are class members and `MonadState a` is a superclass.

The associated typefamily will look like this:

```
1    type Capabilities a :: [(* -> *) -> Constraint]
2    type Capabilities a = '[]
```

With this the signature of `behavior` will change to:

```
1    behavior :: (ActorContext ctx, ImplementsAll (ctx a) (Capabilities a)
         ) => Message a -> ctx a ()
```

Where `ImplementsAll` is a typefamily of kind `Constraint` that checks that the concrete context class fulfills all constraints in the given list:

```
1  type family ImplementsAll
2    (a :: k) (c :: [k -> Constraint]) :: Constraint where
3      ImplementsAll a (c ': cs) = (c a, ImplementsAll a cs)
4      ImplementsAll a '[]       = ()
```

To be able to run the behavior of a specific Actor the chosen `ActorContext` implementation has to also implement all Monad classes listed in `Capabilities`.

```
1  newtype SomeActor = SomeActor ()
2    deriving (Eq, Show, Generic, Binary, Monoid)
3  instance Actor SomeActor where
4      type Capabilities SomeActor = '[MonadIO]
5      behavior () = do
6          liftIO $ putStrLn "we can do IO action now"
```

Since `MonadIO` is in the list of capabilities, we can use its `liftIO` function to perform arbitrary `IO` actions inside the `ActorContext`.

`MonadIO` may be a bad example though since it exposes too much power to the user. What we would need here is a set of more fine grain Monad classes, that each only provide access to a limited set of IO operations. Examples would be: a network access Monad class, file system class, logging class, etc. These would be useful even outside of this Actor framework.

the Eff Monad

The `Eff` Monad as described in the `freer`, `freer-effects` and `freer-simple` packages is a free Monad[22] that provides an alternative way to Monad classes and Monad transformers to combine different effects into a single Monad.

In category theory a free Monad is the simplest way to turn a functor into a Monad[22]. In other words it's the most basic construct for that the Monad laws hold given a functor. The definition of a free Monad involves a hefty portion of category theory. We will only focus on the aspect that a free Monad provides a way to describe monadic operations, without providing interpretations immediatel. Instead there can be multiple ways to interpret these operations.

When using the `Eff` Monad there is only one monadic operation:

```
1  send :: Member eff effs => eff a -> Eff effs a
```

`effs` has the kind `[* -> *]` and `Member` checks that `eff` is an element of `effs`. Every `eff` describes a set of effects. We can describe the Actor operations with a GADT that can be used as effects in `Eff`:

```
1  data ActorEff a v where
2      Send   :: Actor b => ActorRef b -> Message b -> ActorEff a ()
3      Create :: Actor b => Proxy b -> ActorEff a ()
4      Become :: a -> ActorEff a ()
```

With this we can define the functions:

```
1  send :: (Member (ActorEff a) effs, Actor b)
2       => ActorRef b -> Message b -> Eff effs ()
3  send ref msg = Freer.send (Send ref msg)
4
5  create :: forall b a effs.
6            (Member (ActorEff a), Actor b)
7         => Eff effs ()
8  create = Freer.send $ Create (Proxy @b)
9
10 become :: Member (ActorEff a) effs => a -> Eff effs ()
11 become = Freer.send . Become
```

We can also define these operations without a new data type using the predefined effects for `State` and `Writer`:

```
1  send :: (Member (Writer [SystemMessage]) effs, Actor b)
2       => ActorRef b -> Message b -> Eff effs ()
3  send ref msg = tell (Send ref msg)
4
5  create :: forall b a effs.
6            (Member (Writer [SystemMessage]), Actor b)
7         => Eff effs ()
8  create = tell $ Create (Proxy @b)
```

`become` does not need a corresponding function in this case since `State` already defines everything we need.

3.4 Testing

One of the goals of the Actor framework is testability of Actors written in the framework. The main way that testability is achieved, is by implementing a special `ActorContext` that provides a way to execute an Actors behavior in a controlled environment. The name of this

ActorContext is MockActorContext. MockActorContext has to provide implementations for create, send and MonadState. Additionally we need a way to execute a MockActorContext. One way to define MockActorContext is using Monad transformers in conjunction with GeneralizedNewtypeDeriving.

```
 1  newtype MockActorContext a v = MockActorContext
 2      ( ReaderT (ActorRef a)
 3          ( StateT CtxState
 4              (Writer [SystemMessage])
 5          ) v
 6      )
 7    deriving
 8      ( Functor
 9      , Applicative
10      , Monad
11      , MonadWriter [SystemMessage]
12      , MonadReader (ActorRef a)
13      )
```

Where CtxState is used to keep track of Actor instances, that currently are known to the context.

```
 1  data CtxState = CtxState
 2      { nextId :: Word
 3      , states :: HMap ActorRef
 4      }
 5    deriving
 6      ( Show
 7      , Eq
 8      )
```

MonadState is a prerequisite for ActorContext so an instance of that has to be provided.

```
 1  instance Actor a => MonadState a (MockActorContext a) where
 2      get = do
 3          ref <- ask
 4          MockActorContext . gets $ ctxLookup ref
 5      put a = do
 6          ref <- ask
 7          MockActorContext $ do
 8              CtxState i m <- get
 9              let m' = HMap.hInsert ref a m
10              put $ CtxState i m'
```

With this the actual definition of `ActorContext` for `MockActorContext` is pretty short.

```
 1  instance (Actor a, MockActorContext a `CanRunAll` a)
 2           => ActorContext a (MockActorContext a) where
 3
 4      self = ask
 5
 6      create' _ = do
 7        ref <- MockActorContext ctxCreateActor
 8        tell [Left $ Creates ref]
 9        pure ref
10
11      p ! m = tell [Right $ Send p m]
```

To execute a single `MockActorContext` action, all Monad transformer actions have to be executed. It does not make sense though, to export this capability directly, since `CtxState` should not be visible to the user. So exported variants on the core running function construct `CtxState` values themselves.

```
 1  runMockInternal :: forall a v. Actor a
 2                  => MockActorContext a v -> ActorRef a -> CtxState
 3                  -> ((v, CtxState), [SystemMessage])
 4  runMockInternal (MockActorContext ctx) ref
 5      = runWriter
 6      . runStateT (runReaderT ctx ref)
```

3.5 executing in a distributed environment

When executing an Actor inside a distributed environment, one has to take care of message passing and the actual concurrent execution. cloud-haskell already provides a solution for this problem, in form of an Erlang-style Actor framework. Problematically it's messages are untyped and every Actor has access to `IO`. This enables us to execute the previously defined typed Actors on top of it.

As with the testing case the central entry point will be the `ActorContext`. All actions in cloud-haskell are inside of the `Process` Monad. So we need to keep track of the Actors state and access to the `Process` Monad. This can be achieved using a `newtype` wrapper around `StateT` transformer of `Process`:

```
 1  newtype DistributedActorContext a v
 2      = DistributedActorContext
 3          { runDAC :: StateT a Process v
 4          }
```

```
5        deriving (Functor, Applicative, Monad, MonadState a, MonadIO)
```

GeneralizedNewtypeDeriving is used to derive the normally not derivable instances like MonadIO
. ActorContext has to be implemented manually.

```
1   instance (DistributedActorContext a `A.CanRunAll` a)
2            => A.ActorContext a (DistributedActorContext a) where
3       self = A.ActorRef . encode <$> liftProcess getSelfPid
4       (A.ActorRef pid) ! m = liftProcess $ send (decode pid) m
5
6       create' a = liftProcess $ do
7           nid <- processNodeId <$> getSelfPid
8           pid <- spawn nid (staticRunActor a)
9           return . A.ActorRef . encode $ pid
```

All we have to do to implement self and (!) is to wrap/unwrap the process id in an ActorRef
and use the functions that Process gives us. Notice that create' uses staticRunActor instead
of runActor. More on this in the next section

Executing an Actor in this context, now means dispatching a Created signal and continuously
polling for messages.

```
1   runActor :: forall a proxy.
2            (A.Actor a, DistributedActorContext a `A.CanRunAll` a)
3          => proxy a -> Process ()
4   runActor _
5     = void
6     . runStateT (runDAC runActor') $ A.startState @a
7   where
8     runActor' = initActor *> forever awaitMessage
9     initActor = A.behavior . Left $ A.Created
10    awaitMessage = A.behavior
11              . Right =<< liftProcess (expect @(A.Message a))
```

3.5.1 Creating Actors

Creating an Actor means spawning a new Process that executes runActor for that specific Actor
type. The problem here is that the instruction on what the Process should do has to be serializ-
able. Since functions are not Serializability in Haskell cloud-haskell provides a workaround with the
distributed-static package.

It provides a way to serialize references to values and functions that are known at compile time and compose these. This is done using a heterogeneous map that has to be manually populated with all static values that the program may encounter while running. Unfortunately there is no way to register polymorphic functions like `runActor` this way. Luckily there is a way to enumerate all Actor types that exist in a given Actor system and could register a version of `runActor` for each one. As the hierarchy of the Actor system, described by the `Creates` typefamily is potentially infinite, the registration would have to perform cycle detection on the type level.

Alternatively we can also parametrize the actor context by a type list of kind `[*]`. This list represents a flattened version of the Actor system's hierarchy. Each actor type inside of the actor system has to be a an element of that list. All ActorContext operations have to check that the relevant Actors are in fact elements of that list. Although this solution may appear trivial, I discovered it at a very late stage of the project. As a result this solution is not incorporated into the main library code. A proof of concept exists in the source repository at `./cloud-fix/Main.hs`. This proof of concept also includes a running example of a solution for the dining philosophers problem as well as a ping pong example.

4 Results

I demonstrated that it is possible to create an Actor framework in Haskell that is capable of expressing many constraints about it's hierarchy and the capabilities of the Actors in it, using the type system. The created framework allows a wide range of properties of actors to be expressed and reasoned about at compile time. Furthermore the `Capabilities` mechanism and the the ability to run Actors defined in Dakka to be against multiple backends makes it extensible, too. To test Actors they just have to be run against a testing backend. If the provided testing backend isn't sufficient it can be augmented with additional capabilities by implementing appropriate type classes or using newtypes.

4.1 Dependent types in Haskell

Dependent types are a powerful tool in Haskell. Unfortunately their usability is somewhat limited since the language support for them is also limited. The lack of native support does not make its use impossible but prevalent usage cumbersome. Promotion of values to types and demotion from types to values has to be done manually. The *singletons*, that aims to aleviate this problem, is easy enough to use but produces hard to debug type expressions. As a result of this I decided to reduce the usage of dependent types in my code and do without the *singletons* library altogether. Even though they are not dependent types many of the more advanced type-level-computation features Haskell provides were useful. Dependent typing in Haskell definetly requires better native support before it can be widely be adopted.

4.2 Cloud Haskell

Since most of the effort went into creating a typed interface rather then its actual execution, I can't comment on how cloud haskell actually performs as a backend for the created interface.

The implementation was very straightforward for the most part since cloud haskell provides similar primitives to the API itself. Parts that required the serialization of polymorphic values were not that easily implemented. As statet in[7], GHC native support for static values was planned originally. Without native support for static values serializing polymorphic values requires that all possible type parameters can be enumerated. Functions with the resulting monomorphic types can then be registered as static values.

4.3 Future Work

Although Dakka can be used to create working Actor systems some parts can be improved. These range from minor concers about the codebase to future research topics. The codebase could use some cleanup and improvements in usability.

4.3.1 General cleanup

The codebase has been growing organically in the course of the project. As a result, the code can be streamlined and cleaned up. Since the creation process was also a learning experience it contains a few remnants of experiments that are no longer needed. For example I would like to rename `ActorContext` to better to `ActorAction`. For this framework to be usable it also requires better documentation. Both Haddock comments inside of the code and basic tutorials would be needed.

4.3.2 Automatically flattening the Actor System type hierarchy

Currently the user has to provide a flat representation of all Actor types inside of an Actor system if they want to run it manually. This representation could be derived from the root Actor of the Actor system. For this all entries inside of the `Creates` type family have to be recursively aggregated. Since the graph of Actor types inside of an Actor system may have cycles, flattening this graph requires cycle detection inside a type computation.

4.3.3 Polymorphic Actors

Currently the API is not designed with polymorphic Actor types in mind. The `RootActor` is an exaple of a polymorphic Actor. It is parametrized by the Actors it should create on startup. In this case the

`Creates` type is trivially defined as `Creates (RootActor l) = l`. It would be interesting to explore more complex polymorphic Actor types.

4.3.4 Support more Akka features

Although the API is inspired by Akka, it only provides a very small subset of its operation. For real world use cases this has to be expanded. It would be useful for example to provide an Actor with an initial state on creation. This features was explored briefly in the `cloud-fix`/`Main.hs` proof of concept.

Another big feature of Akka is Actor discovery. It's possible to search for specific Actors inside of an Actor system. Together with the ability to name Actors and search them by name this is a powerful feature. It would be interesting to examine whether or not this feature in particular can be added in a typesafe fashion.

4.3.5 Better type error messages

When using type level computation in Haskell compiler error messages currently are not very refined in general. A common error is that two type expressions to not unify to the same type or that no definition for a type family exists for some given types. These errors may occur deep inside of complex type expressions and may be dealing with types that have little connection to the types in the context that the error was caused by. This makes debugging hard for the author of these type expressions and basically impossible for anyone else without consulting the source code of these expressions themselves. When creating a framework it can not be expected of a user to consult the source code of the framework each time a type expression from that framework causes a compiler error.

In Dakka these type expressions are used to prevent the user from using the framework wrongly. The error messages thus should tell the user what they did wrong and ideally how to fix this.

There are techniques to aid the user here though.

4.3.6 Make Actor creation easier

Currently you have to perform many steps to create an Actor in Dakka. You have to create a data type for your Actor and implement the `Actor` class for it.

The data type has to also have `Show`, `Eq` and `Binary` instances. `Show` and `Eq` instances can be derived. Deriving `Binary` is not directly possible. To obtain a `Binary` instance without implementing it manually you have to derive `Generic` and than create an empty `Binary` implementation. To be able to derive `Generic` instance though you have to enable the `DeriveGeneric` language extension. With the `DeriveAnyClass` language extension you can remove the empty `Binary` implementation with

an entry in the `deriving` clause of the data type. `DerivingAnyClass` can not always determine what derving strategy should be used for a specific class or the user does not want to use a certain deriving strategy. This comes up most often when using `DerivingAnyClass` in conjunction with `GeneralizedNewtypeDeriving`. When creating an Actor it is very likely that the internal state of that Actor can be modeled by an existing data type. Defining the Actor as a `newtype` is thus a reasonable thing to do. For these cases the `DerivingStrategies` language extension exists, which enables the user to specify the desired deriving strategy manually.

Implementing the `Actor` class consists of providing an implementation of `behavior`. If the Actor wants to create other actors the `Creates` associated type family also has to be overridden. If the Actor wants to do anything but the basic Actor operations the `Capabilities` associated type family has to be overridden.

All of these potential steps add up to a substantial amount of boilerplate code. It would be nice if the Amount of Boilerplate could be reduced without weakening the constraints of the API. The most promising way to achieve this is seems to be TemplateHaskell.

Bibliography

[1] Actors - akka documentation. Available from: https://doc.akka.io/docs/akka/2.5.16/actors.html, [31/08/2018].

[2] Mailboxes - akka documentation. Available from: https://doc.akka.io/docs/akka/2.5.16/mailboxes.html, [31/08/2018].

[3] Routing - akka documentation. Available from: https://doc.akka.io/docs/akka/2.5.16/routing.html, [31/08/2018].

[4] Cluster specification - akka documentation. Available from: https://doc.akka.io/docs/akka/2.5.16/common/cluster.html, [31/08/2018].

[5] Actors - akka documentation. Available from: https://doc.akka.io/docs/akka/2.5.16/typed/actors.html, [31/08/2018].

[6] Cloud-haskell documentation. Available from: https://haskell-distributed.github.io/documentation.html, [13/11/2018].

[7] Jeff Epstein SP-J Andrew P. Black. Towards haskell in the cloud. Available from: https://www.microsoft.com/en-us/research/wp-content/uploads/2016/07/remote.pdf, [13/11/2018].

[8] Distributed-static. Available from: https://hackage.haskell.org/package/distributed-static-0.3.8/docs/Control-Distributed-Static.html, [18/11/2018].

[9] Thorsten Altenkirch JM Conor McBride. Why dependent types matter. Available from: http://www.cs.nott.ac.uk/~psztxa/publ/ydtm.pdf, [13/11/2018].

[10] Simon Peyton Jones GWSW. Wobbly types: Type inference for generalised algebraic data types. Available from: https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/old-wobbly.pdf?from=https%3A%2F%2Fresearch.microsoft.com%2Fen-us%2Fum%2Fpeople%2Fsimonpj%2Fpapers%2Fgadt%2FMS-CIS-05-26.pdf, [03/12/2018].

[11] Pierce BC. 2002. Types and programming languages. Cambridge, Massachusetts 02142: The MIT Press.

[12] Kind polymorphism. Available from: https://downloads.haskell.org/~ghc/8.6.2/docs/html/users_guide/glasgow_exts.html#kind-polymorphism, [28/11/2018].

[13] Brent A. Yorgey JC Stephanie Weirich. Giving haskell a promotion. Available from: http://dreixel.net/research/pdf/ghp.pdf, [13/11/2018].

[14] Richard A. Eisenberg SW. Dependently typed programming with singletons. Available from: https://cs.brynmawr.edu/~rae/papers/2012/singletons/paper.pdf, [13/11/2018].

[15] Richard A. Eisenberg JS. Promoting functions to type families in haskell. Available from: https://cs.brynmawr.edu/~rae/papers/2014/promotion/promotion.pdf, [13/11/2018].

[16] Jones MP. Functional programming with overloading and higher-order polymorphism. Available from: https://web.cecs.pdx.edu/~mpj/pubs/springschool95.pdf, [28/11/2018].

[17] Oleg Kiselyov C-cS Simon Peyton Jones. 2010. Fun with type functions. Available from: https://www.microsoft.com/en-us/research/wp-content/uploads/2016/07/typefun.pdf?from=https%3A%2F%2Fresearch.microsoft.com%2F%7Esimonpj%2Fpapers%2Fassoc-types%2Ffun-with-type-funs%2Ftypefun.pdf, [04/12/2018].

[18] Gert-Jan Bottu TS Georgios Karachalias. Quantified class constraints. Available from: https://i.cs.hku.hk/~bruno/papers/hs2017.pdf, [06/08/2018].

[19] Ralf Lämmel SPJ. Scrap your boilerplate with class: Extensible generic functions. Available from: https://www.microsoft.com/en-us/research/wp-content/uploads/2016/07/gmap3.pdf, [03/12/2018].

[20] 10.1. Language options — Glasgow Haskell Compiler 8.4.3 User's Guide. Available from: https://downloads.haskell.org/~ghc/8.4.3/docs/html/users_guide/glasgow_exts.html#deriving-typeable-instances, [31/08/2018].

[21] #14341 (Show instance for TypeReps is a bit broken) – GHC. Available from: https://ghc.haskell.org/trac/ghc/ticket/14341, [31/08/2018].

[22] Swierstra W. 2008. Functional pearl, data types à la carte. Available from: http://www.cs.ru.nl/~W.Swierstra/Publications/DataTypesALaCarte.pd, [03/12/2018].

5 Appendix

All code produced including this thesis itself can be found on github at https://github.com/chisui/dakka. This document is based on the following commit.

Commit: `cc37e94c8e33c8e94e36241bcdfa90d3b51e9ddc`

sha256: `0z72haiy15djf3mpbvxmlhs9apsx9g4dq9yf02dd3sq2ijp993b8`

To browse the repository at this commit visit

https://github.com/chisui/dakka/tree/cc37e94c8e33c8e94e36241bcdfa90d3b51e9ddc.

To verify the hash download the tarball, unpack it and hash its contents by running:

```
nix-prefetch-url --unpack \
  https://github.com/chisui/dakka/archive/cc37e94c8e33c8e94e36241bcdfa90d3b51e9ddc.tar.gz \
  0z72haiy15djf3mpbvxmlhs9apsx9g4dq9yf02dd3sq2ijp993b8
```