MASARYK UNIVERSITY
FACULTY OF INFORMATICS

Game development in Haskell

BACHELOR'S THESIS

Jan Rychlý

Brno, Spring 2021

MASARYK UNIVERSITY
FACULTY OF INFORMATICS

Game development in Haskell

BACHELOR'S THESIS

Jan Rychlý

Brno, Spring 2021

Declaration

Hereby I declare that this paper is my original authorial work, which I have worked out on my own. All sources, references, and literature used or excerpted during elaboration of this work are properly cited and listed in complete reference to the due source.

Jan Rychlý

Advisor: doc. Mgr. Jan Obdržálek, PhD.

Acknowledgements

These are the acknowledgements for my thesis, which can span multiple paragraphs.

Abstract

This is the abstract of my thesis, which can
span multiple paragraphs.

Keywords

Haskell, functional paradigm, game development, Apecs

Contents

List of Tables

List of Figures

Introduction

// introduction draft written for the VB000 assignment

Video games are a special kind of application that many consider an art form and rewarding to develop. However, they generally involve a complex system with a non-trivial state, a certain amount of pseudo-randomness, and user/player input handling. This makes for non-deterministic programs that are usually incredibly difficult to test efficiently.

Conversely, functional programming strives to eliminate mutable state and make code more deterministic, which allows for programs to be safer and easier to test. These and other benefits have naturally led to people trying out game development in functional languages, but it remains mostly a matter of passion projects. That said, even though the vast majority of the video game industry still uses imperative languages like C++, the communities *are* very active, and there are hundreds of games, blog posts, and libraries that help with game programming in functional languages.

The focus of this thesis narrows down to exploring game development in Haskell in the context of small-scale 2D games. The goal is to give an overview of the process, then compare this approach to a more conventional and imperative one and ultimately highlight the features of Haskell that are beneficial and those that become hurdles in the context of programming a video game.

This is done through reimplementing a single game with an already existing imperative implementation in Haskell, first using the Apecs[1] library and for a second time without it. After a further discussion about chosen technologies in the following chapter, said three implementations are described and analyzed in chapters 2, 3, and 4. Then they are more closely compared and the pros and cons of Haskell in game development are evaluated and demonstrated in chapter 5.

We find that the Apecs library makes developing games in Haskell much more approachable. On the other hand it goes against the functional philosophy, and using it will generally result in very imperative code wrapped in monads that lacks the expressiveness and apparent

---

1. CARPAY, Jonas. *Apecs: Fast Entity-Component-System library for game programming*. Available also from: `https://hackage.haskell.org/package/apecs`.

safeness of regular Haskell. Yet, from the second reimplementation, we learn that some use of monads is beneficial, and it makes the code cleaner and more elegant. In both cases, the development was mostly a smooth experience without a single major hick-up, unlike what often happens when dealing with a C++ compiler.

1 Motivation and used methods

1.1 Why functional programming matters

Functional languages are a subset of declarative languages, where the programmer states "what" instead of "how". Unlike in imperative languages, we *declare* what we want a program to return by combining functions (other declarations) instead of giving the computer serialized instructions (*imperatives*). That is manifested in the lack of assignment statements and lesser control structures. What sounds like a bug is actually a feature — once variables are assigned their value, it can't be changed, and the burden of prescribing the flow of control is removed[2]. Moreover, a pure function has no side effects and its return value depends solely on its arguments. This makes for deterministic programs that are easier to test, debug and argue about their correctness.

John Hughes describes the key benefits of functional paradigm in his paper *Why Functional Programming Matters*[2]. He first explains how modularity of code is clearly very important, since separate modules are easier to write and test and then proceeds to show how functional programming increases modularity through higher-order functions and lazy evaluation, demonstrating their importance on several examples.

Additionally, Haskell is a purely functional programming language that is also *strongly typed*. This means that one doesn't need to worry about memory errors causing crashes because everything is caught by the type-checker during compilation. The types also serve as documentation and can help greatly with writing and understanding code. On the other hand types can also be inferred by the compiler so it is not necessary to explicitly declare the type of everything. Furthermore the type system allows extensive user-defined data types, which makes the code even more expressive. All of this potentially increases productivity of a functional programmer even further.

1.1.1 Game development specifics

Functional languages are great tools but we know that not every tool is fit for every job. One of the consequences of the functional purity is that the state of the program has to be modeled explicitly as an argument and is therefore immutable. There are monads that help us abstract from this but the monads themselves are in a way still an explicit work around.

Conversely, games are real-time, interactive applications simulating often very complex systems and hold a non-trivial state that is updated many times a second. Such state generally involves a representation of the game world with all the objects existing in it, their properties and flags, the current state of the input devices and many other variables. Since its beginning, the video game industry has been dominated by imperative languages, that make it easy to model a game world and alter it globally through references and side effects inside of decomposed functions. And because of their established position, there is a plethora of libraries and game engines with supporting documentation and tutorials. Besides, a company will most likely have no problem finding skilled C++ programmers with interest in the game industry, where as finding their functional counterparts might be much harder. Additionally, in the case of C++ the performance also fits the requirements of large games.

However, it does come with a price — modules of such programs may be more dependent and entangled, which makes the whole less flexible and with implicit state more prone to bugs and harder to test and debug. That is, while testing is already a large issue due to the nature of video games. Automated testing is not sufficient and companies have to hire teams of game testers to test games manually. And in terms of high performance, which is generally connected to lower-level languages, developers must wrestle with the lower-level nature, producing problems as well.

We can see that video games, like any other software, could benefit from purely functional design, provided that we are able to model the game state efficiently enough. Another consideration in the real world are the available frameworks and whether the cost of potential pioneering is worth to us.

1.1.2 Existing work

Indeed, people have tried developing games in Haskell and a decent progress has been made over the years. There are libraries like ...
Also games that were published commercially and non-commercially ...
... blog posts like the one by John Vargas (Dino Rush), ?reddit activity?, talks given in the past, Yampa by Keera Studios, elerea, Chucklefish's new game engine ?Cove?, Indigo in Scala

1.2 Asteroids by Atari as an example

simple example with enough interesting world but not too complex for the scope of a ba thesis

1.3 Rendering and interfacing with the OS

SDL2 and other existing Haskell libraries, why I chose SDL2

2 Using the Apecs library — hAsteroids

2.1 About Apecs

Many libraries for game programming in Haskell have come out over the years. One of the more recent ones is Apecs — "a fast, type-driven Entity–Component–System library for game programming."[1] Entity–Component–System (or ECS) is a data-oriented architectural pattern often used in video game engines. It provides better performance by increasing data locality. Instead of the object-oriented pattern where data is grouped by their related **entity**/object, in ECS , we group pieces of data by their character and call them **components**. Simply put, we exchange an array of structures for a structure of arrays. That way, when running functions — which define a **system** — on entities every frame of the game, we iterate only over the arrays of components we need for the system rather than iterating over a much larger data set of whole objects as it would be in the object-oriented paradigm. Typical example of a system would be updating positions of all entities based on their velocity regardless of whether that entity is a player, projectile or a falling anvil.

[https://medium.com/ingeniouslysimple/entities-components-and-systems-89c31464240d]

And since both Unity and Unreal engine use Entity-Component design, we chose Apecs as the current state-of-the-art Haskell library for the traction it has received in the community despite it not being the only ECS library in existence[1].

---

1. The making of Ecstasy was actually inspired by author's issues with Apecs.
`https://reasonablypolymorphic.com/blog/why-take-ecstasy/`

Listing 2.1: Defining instance for Component

```haskell
newtype Position = Position (V2 Double)
instance Component Position where
    type Storage Position = Map Position
```

Listing 2.2: Simplified world state type example

```haskell
data World =
    World
    { record1 :: !(Unique Player)
    , record2 :: !(Map Enemy)
    , record3 :: !(Map Bullet)
    , record4 :: !(Map Position)
    , record5 :: !(Global Time)
    }
```

To define a component in Apecs means to define an instance of the class `Component`, as we see in the 2.1. The `Component` class requires us to state how we want to store the given component by assigning a type alias to the specific storage type. We can define our `Stores` or use one of those provided with the library: `Map`, `Unique`, `Global`. With `Map`, there can be multiple components of that type, each belonging to a particular entity. With `Unique`, at most one component may exist belonging to a particular entity. Furthermore, with `Global`, at most one component instance can exist, and it belongs to the special *global* entity together with every other entity. Finally, we call `makeWorld`, which uses Template Haskell to generate `World` product type along with `initWorld` function and instances of the `Has` class needed for altering contents of `World` through the other functions in Apecs. The resulting `World` may look close to something as shown in listing 2.1.

Having components, we need a way to run systems. In Apecs, the `SystemT` monad transformer is where changes to the world happen. Therefore any "system" in the ECS sense must be a function returning `SystemT w m a`. One such "micro-system" is the `newEntity` function. It accepts a tuple of components and adds them into their records — the new entity itself is defined by the components stored under its ID.

```haskell
newEntity (Ship $ pi / 2 * 3, Position $ V2 x y, Velocity $ V2 0 0)
```

Another noteworthy functions to build systems are the component map functions shown in the listing 2.1. They are the means of altering the world state.

Listing 2.3: Component maps documentation

```haskell
-- | Maps a function over all entities with a cx, and writes their cy.
cmap :: forall w m cx cy. (Get w m cx, Members w m cx, Set w m cy) => (cx ->

-- | Monadically iterates over all entities with a cx, and writes their cy.
cmapM :: forall w m cx cy. (Get w m cx, Set w m cy, Members w m cx) => (cx ->

-- | Monadically iterates over all entities with a cx
cmapM_ :: forall w m c. (Get w m c, Members w m c) => (c -> SystemT w m ())
```
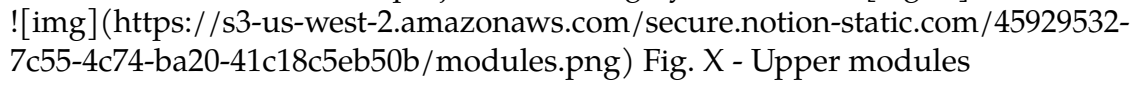
[[https://hackage.haskell.org/package/apecs-0.9.2/docs/Apecs.html](https://hackage.
0.9.2/docs/Apecs.html)] - probably simplify the types with an asterisk

cmap accepts a function that takes a tuple of components and returns some other tuple of components. It internally iterates over entities with at least those components matching the mapped function's input tuple and writes the output tuple components to those entities. cmapM works similarly only as its name suggests the mapped function returns the component tuple wrapped in the system monad, which allows it to execute side effects. And with cmaM_ there is no direct writing, only side effects.

2.2 Writing of hAsteroids

To familiarize our selves further with the library and with the process of using it for its main purpose — creating games in Haskell — we did exactly so.

Module structure of the project looks roughly as shown in [Fig. X] ![img](https://s3-us-west-2.amazonaws.com/secure.notion-static.com/45929532-7c55-4c74-ba20-41c18c5eb50b/modules.png) Fig. X - Upper modules may have direct dependencies on the lower modules, arrows show most of the important ones.

The core of the game's design is the components, as they form the primary data structure acting as the state. There are more approaches to designing them, but in hAsteroids, we have three categories of components: marker components, shared components, and control

components. Marker components serve two purposes: they contain information that is unique for a given type of game object, and that way, they also mark an entity as that object. Shared components include characteristics that are shared by more types of game objects like position. Lastly, control components are all global and are used in one way or another to control the run of the game. hAsteroids has one `Unique` marker component called `Ship`, which marks an entity representing the player's ship and stores the angle of the direction the ship is facing. The three other marker components are `Map` stored. They are `Asteroid` — holding asteroid size — `Ufo` — holding saucer size and a countdown to the next UFO's shot being fired — and `Bullet` — storing whether the player or a UFO shot it. Next, there are the shared components `Position`, `Velocity` and `TimeToLive` and several `Global` control ones like `ShipLives`, `ShipState`, `GameLoopState`, `WaveTime` and few others. []Fig. X] shows which entity types have which components.

Ship - Ship, Position, Velocity

UFO - Ufo, TimeToLive, Position, Velocity

Bullet - Bullet, TimeToLive, Position, Velocity

Asteroid - Asteroid, Position, Velocity

Now, the `main` function is the entry point of the program. It first initializes the SDL libraries and then creates a window and a renderer. Next, using `loadResources`, it loads object textures into a hash map, prerenders fonts saving them into another hash map and wraps it all together with the renderer and a few stateful random generators into one product type called `Resources`. Then the game world with our components is created by calling `initWorld`, and together with resources, they are passed to `gameLoop` through a stack of Reader monad transformers (`SystemWithResources`).

`gameLoop` is one compound system responsible for updating and drawing the world and the menus, and it loops until the player quits the game. It also measures the time every frame and calls `SDL.Delay` if it was updated and drawn too quickly for the targeted 60 FPS. World updating is split into two functions: `reactToInput` and `stepScene`. The scene (a menu or the world) is then drawn by `drawScene`.

`reactToInput` manages the state of the input, and as its name suggests, it reacts to it. Depending on the global `GameLoopState` component, it either transitions between the states (`InMenu`, `Playing`, `Paused`, `GameOver`, `Quit`) or when in the `Playing` state, it also allows the player

to control the ship. That is done by a `cmapM` call with a lambda that changes the angle of the ship, increases its velocity or creates a new bullet entity — all conditioned by the input state.

   `stepScene` takes care of simulating physics and game rules over time when the loop is in the `Playing` state. This is divided into multiple function calls:

- `cmap \$ stepKinetics dT`

  iterates over all entities and adds their velocity vector multiplied by time `dT` to their position vector and also takes care of wrapping the space — if an entity flies out of the screen on one side, it comes back in from the other side;

- `cmap \$ decelerateShip dT`

  simply applies deceleration to the ship by scaling down its velocity vector slightly;

- `cmapM \$ stepShipState dT`

  is responsible for transitioning between ship states (`Alive`, `Exploding Int`, `Respawning Int` where the integers serve as countdown timers for the state transition) and the "explosion animation";

- `cmapM_ \$ ufosShoot dT`

  iterates over all `Ufo` components decrementing the time to shoot and when it reaches 0, it creates a new `Bullet`. The algorithm for finding a shooting direction is different for the two UFO sizes. Small UFOs are more accurate because the algorithm uses the law of sines to calculate the bullet trajectory based on the ship's current position and velocity. Large UFOs shoot in quarter of π increments towards the ship's current location.

- `awardLifeIf10000`

  For reaching every 10000 score points ship lives are incremented by 1.

- a lambda incrementing `WaveTime`

  Simply adds the frame delta time `dT` to the wave time counter.

- `spawnUfos`

  Randomly creates new UFO entities on the left side of the screen, with the chances increasing as the time spent in one wave (`WaveTime`) passes.

- `spawnNewAsteroidWaveIfCleared`

  uses `cfold` from Apecs to count all asteroids and when there are none it starts counting up using the `WavePauseTimer`. When the timer reaches 1500 it calls `spawnNewAsteroidWave` from the `Initialize` module, which uses the random stateful generators from the `WithResources` reader monad to create new asteroids.

- a lambda decrementing all `TimeToLive` components

  Simply subtracts the frame delta time `dT` from all the time to live components.

- `destroyDeadBullets` and `destroyDeadUfos`

  use `cmapM_` to destroy all the components for entities that run out of "time to live."

- `detectAndHandleCollisions`

  Collision handling has its own module — `Collisions`. There the collisions are detected and handled individually between each entity group. The general idea is that we use `cmapM_` inside of `cmapM_` as an equivalent of nested for loops. This way, for every asteroid, we iterate (or map) over all the other entities, checking for collision. We do the similar for the rest of the combination pairs, using in total $\binom{6}{2} = 6$ algorithms. All the collisions are detected simply as a question of "is a point or any of the points inside of a rectangle, an ellipse or a circle." A detected collision always results in some effect — the colliding entities are removed, except an asteroid may break into two smaller ones if it is not already the smallest size, and in the case of the ship, one life is subtracted.

  Once the scene is stepped, it is drawn by the already mentioned `drawScene` function from the `Draw` module. If the loop state is `InMenu` or `GameOver` only text is drawn on a cleared black screen. That is done by `drawCenteredTexts`, which only calls a monadic `zipWith`

with `drawCenteredText` on a list of y coordinates and a list of text keys. `drawCenteredText` itself then looks up the text texture in the hash map that is part of the `WithResources` environment, queries its width and finally calls `drawText` with the coordinates for the texture to be drawn centered.

When the loop is in the `Paused` state, there is text being drawn as well as the world. Moreover, when the state is `Playing`, only the world is drawn. This is taken care of by `drawWorld`, which calls functions to draw the background, entities and the UI (number of lives and the score). Entities are drawn using `cmapM_` with a lambda and a wrapper function around the `SDL.copy` and `SDL.copyEx`, which copy the texture to the rendering target.

maybe a footnote - `https://wiki.libsdl.org/SDL_RenderCopy]` (`https://wiki.libsdl.org/SDL_RenderCopy`) is what SDL.Copy is binding to

2.3 Reflection

From this example alone, we learn several valuable lessons right away. First, Apecs makes game programming in Haskell relatively accessible. Once we understand the ECS principle, every world modification is "intuitively imperative" thanks to the `System` monad and the component map functions. Any system side effects can be added to an existing `cmap f` call with a simple change of the return types.

```haskell
-- steer and thrust
handleInput input =
    cmap \$ \(Ship a, Velocity vel) ->
                ( Ship \$ a + steering input
                , Velocity \$ vel + thrust input a
                )
-- we realized that we also want to be able to pause the game
handleInput' input =
    cmapM \$ \(Ship a, Velocity vel) -> do
                when (wasPressed input escapeKeycode) (set global Paused)
                pure ( Ship \$ a + steering input
                     , Velocity \$ vel + thrust input a
                     )
```

Moreover, the `WithResources` reader monad provides easy access to resources without having to pass them along everywhere as a function argument.

Not less important is the fact that the game works. The development experience was smooth, with no major hick-ups. Because Haskell is a statically-typed high-level language, there is no reason to worry about random invalid memory access making our game crash, and everything is type-safe.

However, this approach is far from perfect. Handling collisions on an individual basis would scale very poorly, so some universal interface would be better. This could be achieved by defining a class and its instances for the marker components. More polymorphism could also differentiate functions that require the `System` monad, the `WithResources` monad or both. In the current state, there are many functions, such as `reactToInput`, which do not use the resources but still have access to them since everything returns the transformed `SystemWithResources` monad that has `WithResources` inside of it. Furthermore, said monad transformer stack[2] includes `IO`, so any function with this type can perform input or output effects, which goes against functional purity and nullifies many of the reasons why one would choose Haskell as a language in the first place.

Another issue we observe is partially tied to the nature of ECS — there is no way to destroy all entity's components automatically in Apecs. One has to do destroy them explicitly. That way, nothing protects us from accidentally adding a component to an entity that will never be destroyed. This could be mitigated by creating helper functions for entity creation and destruction.

Overall, because Apecs is so powerful, it makes it easy for the programmer to rely on it too much and produce code that does not reach all the potential benefits purely functional programming has to offer.

---

2. `type SystemWithResources = SystemT World (ReaderT Resources IO)` where `SystemT` is only a `newtype` around `ReaderT` defined in Apecs

3 Focusing on functional purity - pure-asteroids

3.1 Game engines in purely functional style

// refferencing John Carmack's keynote at Quakecon 2013, discussing its benefits, maybe mentioning `https://indigoengine.io/`

3.2 Writing of pure-asteroids

3.3 Reflection

// sort of a immediate reflection

4 Analyzing an existing C++ implementation

// explaining a bit about the choice of the particular implementation and analyzing it, again with a "reflection" part

5 Comparing the approaches

Conclusion

Bibliography

1. CARPAY, Jonas. *Apecs: Fast Entity-Component-System library for game programming*. Available also from: `https://hackage.haskell.org/package/apecs`.

2. HUGHES, John. Why Functional Programming Matters. *The Computer Journal*. 1989, vol. 32, no. 2, pp. 98–107. ISSN 0010-4620. Available from DOI: `10.1093/comjnl/32.2.98`.

Index

A An appendix

Here you can insert the appendices of your thesis.