Chesskell: Embedding a Two-Player Game in Haskell's type system

3rd Year Project Progress Report

Toby Bailey

November 20, 2020

Contents

1	Introduction				
2	Background Related Work				
3					
4	Project Functional Requirements				
5	Current Progress				
	5.1 Chess Types Overview	5			
	5.1.1 Team and PieceName				
	5.1.2 Position	5			
	5.1.3 The Pieces	6			
	5.1.4 The Board				
	5.2 Checking Chess Rules	7			
	5.2.1 Type families in Haskell	7			
	5.2.2 Move Lists	8			
	5.2.3 Moving the Pieces				
	5.3 EDSL	ç			
	5.3.1 Proxies and Singletons				
	5.3.2 EDSL Implementation				
	5.3.3 EDSL Board Creation				
	5.3.4 EDSL Issues				
	5.4 Testing				
6	Next Steps	13			
	6.1 Chess Rules	13			

	6.2	6.2 EDSL			
6.3 Performance					
	6.4 Revised Timetable/Plan				
		6.4.1	Weeks 9-10: 30th November to 13th December		
		6.4.2	Weeks 11-14: 14th December 2020 to 10th January 2021 14		
		6.4.3	Weeks 15-18: 11th January to 7th February		
		6.4.4	Weeks 19-22: 8th February to 7th March		
		6.4.5	Weeks 23-24: 8th March to 21st March		
		6.4.6	Weeks 25-29: 22nd March to 25th April		
7	Reso	ources,	Risks, and Ethical Considerations 15		
	7.1	Resou	rces		
	7.2 Risks				
			Professional, and Ethical Considerations		

1 Introduction

In 2020, video games are more popular than ever. In the US alone, an ESA report¹ estimates that there are more than 214 million individuals who play games. Considering this, it's surprising how many games are released with major bugs in their software—some of which end up being so notable that news and footage of them appear on mainstream media².

As programming languages have evolved, many have begun to address errors at compile time. Features like optional types are being introduced to languages such as Java and C#, and languages like Rust have pioneered ways of safely handling dynamic allocation through the use of owner types³. Many language compilers now force the developer to handle classes of errors that previously could only be encountered at runtime, such as null pointer exceptions. However, the most common programming language used for game development is C++[1], despite its' lack of automatic memory management and allowing unsafe pointer arithmetic. C++ is chosen for its' speed, but as computers (and gaming consoles) get more powerful, speed becomes less important than developer productivity. Research into bringing more type-safety into game programming environments is rare, and is the niche this project attempts to fill.

Recent versions of the *Glasgow Haskell Compiler* (GHC) support programming at the type level, allowing programmers to compute with types in the same way that languages like C or Python compute with values[2], using *type families* that emulate functions at the type-level. These computations run at compile time, before an executable of the source code is generated, allowing programmers to transform logic errors into type errors[3].

The aim of this project is to demonstrate a proof-of-concept; that it is possible to model Chess at the type-level, and that compiled programs comply with the rules.

2 Background

The project, nicknamed Chesskell, has the main aim of modelling the classic board game Chess in Haskell's type system. This type-level model will be interacted with via a Haskell-embedded Domain-Specific Language (DSL), for describing games of chess. This Embedded DSL (EDSL) will be modelled on Algebraic Notation, a method of writing down the moves associated with a particular match of chess.

https://www.theesa.com/wp-content/uploads/2020/07/Final-Edited-2020-ESA_Essential_ facts.pdf

²https://www.bbc.co.uk/news/technology-50156033

³https://doc.rust-lang.org/book/ch04-01-what-is-ownership.html

3 Related Work

There are many publicly available chess engines and chess-related programs written in Haskell⁴⁵, including a Haskell wiki article teaching the programming about Haskell using chess as an example⁶. Many of these are chess engines, which take in a board state and output the move (or set of moves) which are strongest, and so therefore performs move checking at the value-level to ensure that the moves that it outputs are valid. Chesskell differs from these in function, in that the end software will not output a list of strong moves; it will simply take in the moves performed, and state whether they are valid chess moves or not. We are not aware of any such type-level chess implementations in Haskell.

There have been allusions to chess at the type-level through solving the N-queens problem in dependently typed languages, such as Idris[4]⁷. The N-queens problem makes use of some chess rules, including the Queen's attack positions (a straight line in any direction); but as the end goal is not to successfully model a game of chess, it is not a full type-level chess implementation.

Despite the apparent lack of work on Chess at the type level specifically, there has been work on Haskell-Embedded DSLs in other domains to enforce certain behaviour at compile time. [5] is an EDSL for music composition, which checks if the created piece of music conforms to some musical ruleset before compilation of the program. WASH/CGI[6] is an EDSL for type-safe server-side web scripting, which ensures thread-safe server operations and valid XHTML generation.

4 Project Functional Requirements

- 1. The system will contain a type-level model of chess, which will:
 - a) Support all rules within the FIDE official chess rules.
 - b) Have an internal representation of the board, the pieces, and other information relevant to the state of play.
 - c) Be able to detect when a rule of chess has been broken, and give out a compile-time error to explain which rule was broken and how.
- 2. This type-level system will run on the Glasgow Haskell Compiler.
- 3. This type-level system will be accessible via an Embedded Domain-Specific Language, which will:
 - a) Describe a game of chess, move by move.

⁴https://github.com/mlang/chessIO

⁵https://github.com/nionita/Barbarossa

⁶https://wiki.haskell.org/Learning_Haskell_with_Chess

⁷https://github.com/ExNexu/nqueens-idris

- b) Be written in Haskell source files, and be valid Haskell according to the Haskell grammar.
- c) Be type-checked by the type-level model, and will cause errors if the game described breaks the established rules of chess.

5 Current Progress

5.1 Chess Types Overview

The software has a full representation of a game of chess at the type-level, explained in more depth below. As the game is checked move by move, the only representation required is the board at the time of checking; containing all pieces on that board.

The most important types are the board, and the individual pieces on it. Both of these types are composed of other types, making them *composite types*; the board is comprised of pieces, and the pieces contain that piece's team, name, and position.

5.1.1 Team and PieceName

Both of these types are simple algebraic data types, with all inhabitants defined in code. The team defines whether a piece is black or white, and the piece's name states which chess piece it is; a pawn, a rook, and so on.

The declarations below, combined with the -XDataKinds extension, create two things. The first is a value-level representation of Team and Piecename, where Black is a value of type Team. The second is a *type-level* representation, where 'Black is a type and Team is a kind⁸.

5.1.2 Position

The type for holding a piece's position makes use of two more types, one for columns and the other for rows. In chess, columns are labelled with letters and rows are labelled with numbers; "a1" is the top-left of the board, and "h8" is the bottom-right. The column type is another simple algebraic data type with inhabitants listed. The row type is a type-level implementation of Peano natural numbers.

⁸Note that 'Black has no term-level values, since it does not have kind *

```
data Column = A | B | C | D | E | F | G | H
data Nat where
    Z :: Nat
    S :: Nat -> Nat
```

The Position type itself is a *product type*, a type whose cardinality⁹ is equal to the product of the types it takes in. Since Position takes in a type of cardinality 8 (Column) and a type of infinite cardinality (Nat), it has infinite cardinality, even though only 64 of them are valid chess positions.

```
data Position where
At :: Column -> Nat -> Position
```

5.1.3 The Pieces

Each piece contains information relevant for the rule check: that piece's team, name, and an information type. The information type contains a Nat and a Position, to represent the number of moves that piece has taken, and its current position on the board. Recording the number of moves the piece has taken is important for several rules in chess, including castling and *En Passant*. As such, it's included in the PieceInfo type.

```
-- Pieceinfo type, containing general information
data PieceInfo where
    Info :: Nat -> Position -> Bool -> PieceInfo

-- Piece type, containing the most important information for a piece
data Piece where
    MkPiece :: Team -> PieceName -> PieceInfo -> Piece
```

5.1.4 The Board

At the type level, the most common way of storing n elements is with the built-in list type (similar to the value-level). However, a singly-linked list has issues; how can we ensure that the chess board is the appropriate size without a length check each move? This would take at least 64 operations, since the length must be computed recursively; as well as 7 addition operations to put together the list lengths.

Performing such a check each move would be wasteful. Instead, length-indexed vectors are used. Nat values are used to specify the length of a vector, and it would generate a type error if the vector was any other length. As with most things in Haskell, they are defined recursively; an empty vector has length 0, and you express a vector of length (n + 1) by pushing an element to the front of a vector of length n.

⁹A type's cardinality is the number of inhabitants it has.

For example, a vector of 3 boolean types has the type Vec (S (S Z))) Bool. The board can be expressed using these vectors—a vector of 8 vectors of 8 pieces. (Maybe Piece is used instead, since a board square does not necessary contain a piece.)

```
type Eight = (S (S (S (S (S (S (S Z)))))))

type Row = Vec Eight (Maybe Piece) -- Row of eight Maybe
    Pieces

type Board = Vec Eight Row -- Grid of eight rows
```

Although this is the main board type, it is augmented with a BoardDecorator, so named because the intention is to be similar to the decorator design pattern[7] (with the exception that subclassing and superclassing are not features of Haskell). BoardDecorator should be used instead of Board, since it provides additional information:

- The last team to move;
- The last position moved to;
- The White and Black king positions, stored as a tuple;
- The number of moves in the game thus far.

This information prevents repeated passes over the Board which existed in previous versions of the program, to find the last piece to move as well as king positions. While there is the overhead that the BoardDecorator must be updated at each move, the code is much conceptually clearer with the use of the additional information.

```
data BoardDecorator where
   Dec :: Board -> Team -> Position -> (Position, Position) ->
   Nat -> BoardDecorator
```

5.2 Checking Chess Rules

All of the rule-checking is done via type families that take in the BoardDecorator and output either a correct result, or a type-level error.

5.2.1 Type families in Haskell

Although Haskell allows type-level computation via type families, these type families must be fully saturated. That is, they can't be partially applied like normal Haskell functions can. However, Li-yao Xia's work on First Class Families¹⁰ defines a kind of type-level interpreter, allowing similar behaviour to partial application.

¹⁰https://github.com/Lysxia/first-class-families

This means that the board can be mapped over, performing computation on the pieces therein without writing a new function for each map. Incrementing a piece's move count, or changing the name of a piece, is much easier with First Class Families than it would otherwise be with regular type families.

5.2.2 Move Lists

Each piece, depending on its team and name, can move to a certain number of spaces. For instance, a King can move a single space in any direction. The PieceMoveList type family formalises this, returning a list of spaces that a piece can move to (given a piece and board decorator).

```
data PieceMoveList :: Piece -> BoardDecorator -> Exp [Position]
-- If it's the pawn's first move, move two spaces, otherwise
  move one
type instance Eval (PieceMoveList (MkPiece team Pawn info)
  boardDec) = Eval (If (Eval ((IsZero . GetMoveCount) info))
   (PawnStartMove (MkPiece team Pawn info) boardDec)
   (PawnPostStart (MkPiece team Pawn info) boardDec))
-- Bishops move diagonally
type instance Eval (PieceMoveList (MkPiece team Bishop info)
  boardDec) = Eval (AllReachableDiag team boardDec (Eval
   (GetPosition info)))
-- Knights move in an L-shape
type instance Eval (PieceMoveList (MkPiece team Knight info)
  boardDec) = Eval (AllReachableGivenList team boardDec (Eval
   (GetAllKnightPositions (Eval (GetPosition info)))))
-- Rooks move in straight lines
type instance Eval (PieceMoveList (MkPiece team Rook info)
  boardDec) = Eval (AllReachableStraightLine team boardDec
   (Eval (GetPosition info)))
-- Queens can move in any direction
type instance Eval (PieceMoveList (MkPiece team Queen info)
  boardDec) = Eval (AllReachableLineAndDiag team boardDec (Eval
   (GetPosition info)))
-- Kings can move a single space in any direction
type instance Eval (PieceMoveList (MkPiece team King info)
  boardDec) = KingMoveList (MkPiece team King info) boardDec
```

There is another function, PieceAttackList, which is similar to PieceMoveList but gives the list of spaces that a piece can attack. This is different from a piece's move list in a handful of cases (for instance, pieces can attack the opposite team's King, but

cannot move to its position and take it).

5.2.3 Moving the Pieces

The core movement function, aptly named Move, performs a series of checks on the board state, moves the piece, and then performs further checks. The actual movement code is omitted here for brevity; but all moves except for pawn promotion are complete.

5.3 EDSL

5.3.1 Proxies and Singletons

Value-level information is not naturally available to Haskell's type-level facilities, due to type erasure. As such, there are various methods to pass such information up to the type level.

Proxy types are parameterised with kind-polymorphic type variable[2]. A Proxy value can have type Proxy Int or Proxy 'Black, for example. Proxies are used to make runtime board decorator information available at the type level in Chesskell.

Singleton types mimic dependent types[8]; each type has a single, unambiguous value. This value can be used to pass around type information at the value level (and the corresponding type can be used to pass around value information at the type level). Chesskell uses singletons to perform computation with positions and piece names.

5.3.2 EDSL Implementation

The EDSL is implemented through a Continuation Passing Style[9] scheme, with inspiration taken from the Flat Builders pattern developed by Dmitrij Szamozvancev[10]. The core idea is that a value is transformed via a series of continuation functions, until the final continuation function returns a value.

```
-- Starts with a value of type t, gets a continuation of type
(t -> m), and applies it to the value

type Spec t = forall m. (t -> m) -> m

-- Converts a value of type s into a value of type t, passing that t-value into a continuation

type Conv s t = s -> Spec t

-- Ends the continuation stream

type Term t r = t -> r
```

The chess game starts with a Proxy value, with its type containing a BoardDecorator type. Continuations are passed, transforming that value, until the chess game ends (or until a rule is broken). The core continuations defined are named after the pieces, such as pawn and king. Each of them takes in an SPosition, a singleton version of Position. A new datatype, MoveArgs, is defined in order to simplify the process of passing information between the continuations.

```
-- StartDec is a board decorator holding all the information
for the start of a chess game

chess :: Spec (Proxy StartDec)
chess cont = cont (Proxy @StartDec)

data MoveArgs where
    MA :: BoardDecorator -> Position -> PieceName -> Position
    -> MoveArgs

pawn :: Proxy (b :: BoardDecorator) -> SPosition fromPos ->
    Spec (Proxy (MA b fromPos 'Pawn))

pawn (dec :: Proxy b) (from :: SPosition fromPos) cont = cont
    (Proxy @(MA b fromPos Pawn))
```

The next continuation, to, takes in another SPosition, performs the move computation, puts the resulting board decorator into a Proxy type, and passes that Proxy into the continuation given. The final relevant definition is of end, which ends the chess game (and the continuation stream).

Using the above continuations, we can lay out a chess game, move by move. Consider a short game, involving a 3-move checkmate by White. The game can be laid out as below in the EDSL, and will compile successfully:

```
-- Describes a game with a checkmate by White in 3 moves
game = chess

pawn _e2 to _e4 -- White begins
pawn _f7 to _f5 -- Black move
queen _d1 to _f3
pawn _g7 to _g5
queen _f3 to _h5 -- Checkmate!
end
```

However, should Black attempt to move after checkmate, an error will be logged, since no moves by Blac will be valid:

```
-- Below results in a type error:

-- * The Black King is in check after a Black move. This is not allowed.

-- * When checking the inferred type
-- game :: Data.Proxy.Proxy (TypeError ...)

game = chess

pawn _e2 to _e4   -- White begins
pawn _f7 to _f5   -- Black move
queen _d1 to _f3
pawn _g7 to _g5
queen _f3 to _h5  -- Checkmate!
pawn _g5 to _g4  -- Invalid move by black - Black

King is in checkmate
end
```

Or, should White attempt an impossible move in the middle of the game, such as moving a Queen through another piece, a different type error will occur:

```
-- Below results in a type error:

-- * There is no valid move from D1 to D3.

-- The Queen at D1 can move to: E2, F3, G4, H5

-- * When checking the inferred type

-- game :: Data.Proxy.Proxy (...)

game = chess

pawn _e2 to _e4   -- White begins

pawn _f7 to _f5   -- Black move

queen _d1 to _d3  -- White Queen moves through a

Pawn - not allowed

pawn _g7 to _g5

queen _f3 to _h5  -- Checkmate!

end
```

5.3.3 EDSL Board Creation

There are additional statements than the above, to help create an initial chess board for easy testing. A short example is below:

```
-- A simple example board created,
example
          = create
                                    -- create chess board
                                    -- Set a white rook at a8
                put _Wh _R at _a8
                put _Wh _R at _h1
                put _Bl _P at _h7
                                    -- Set a black rook at h7
                                    -- Set black as last team
                lastTeam _Bl
  moved
            startMoves
                                    -- Begin chess game
                rook _a8 to _a2
                pawn _h7 to _h6
                rook _h1 to _h2
            end
```

5.3.4 EDSL Issues

Early versions of the EDSL crashed GHC, due to the initial definition of chess, whereby StartDec was not manually created by the programmer, but pieced together through a series of type family applications:

```
chess :: Spec (Proxy StartDec)
chess cont = cont (Proxy :: Proxy StartDec)
```

GHC would run out of memory to allocate, and would usually crash. This behaviour is unusual, and so we filed a GHC bug report¹¹. As a work-around until the bug is fixed, StartDec was manually written out in full in the source Haskell file.

Additionally, adding the move number to the type error¹² causes a spike in memory usage and compile time; not as excessive as above, as GHC does not crash, but still double or triple what it should be. The move number will only be re-added to the error if this problem can be fixed or worked around, since it would hamper the testing of a large number of games.

5.4 Testing

The test suite is a mixture of HSpec tests[11] and module declarations. The original plan was to have all tests be HSpec behaviour tests; many of these HSpec tests make use of the type-level testing facilities laid out in other works[10]. However, running the test suite causes excessive memory usage, exceeding 27 GB in some cases.

Work on a solution is ongoing, but for now, many tests for the EDSL's new features are expressed as high-level declarations in a Haskell module, and seeing whether they

¹¹https://gitlab.haskell.org/ghc/ghc/-/issues/18902

¹²e.g. Error at move 1: Pawn cannot move to...

compile or not (and if not, what error they throw). This will hopefully only be temporary, but it should suffice as a rudimentary method of integration testing.

6 Next Steps

As the project's main feature set is nearly complete, the remaining time will be spent testing, evaluating, and improving the project. Full-length chess games will be expressed in the EDSL, to determine if the program compiles correct chess games; and the subtle errors will be introduced to those games, to test that incorrect games fail to compile. A comprehensive integration testing plan/suite will be drawn up before the end of the year.

6.1 Chess Rules

At the moment, the only chess rule missed out is *mandatory promotion*. Promotion is currently optional, meaning a Pawn can reach the opposite end of the board and remain a pawn. However, according to the FIDE 2018 Laws of Chess¹³ promotion is not optional, and must occur.

The current implementation of promotion does not involve the Move function, but since the pawn must be promoted at the end of its' move, the Move function needs to include this functionality.

6.2 EDSL

The only major addition to the EDSL is mandatory promotion, as we explain above. However, a few additional helper functions or syntax changes may be helpful during development, and so the EDSL is subject to change.

6.3 Performance

The current memory usage of the EDSL is acceptable, but only short (i10 moves) chess games have been described and tested thus far. Memory usage and compile time are likely to rise with the length of the chess game, and so memory usage optimisations may become necessary as testing continues.

6.4 Revised Timetable/Plan

6.4.1 Weeks 9-10: 30th November to 13th December

These two weeks will be spent implementing mandatory promotion, and performing integration/system testing. A curated data set of chess games will be expressed in

¹³https://handbook.fide.com/chapter/E012018

the EDSL and put through the GHC compiler. Some of these games will be valid, and some will be invalid. Only valid games should compile, and invalid games should fail to compile with an appropriate type error.

6.4.2 Weeks 11-14: 14th December 2020 to 10th January 2021

This is the first section of allocated empty space. If the test suite is posing any issues (with regards to compile time and memory usage) then this additional month should help with solving them. But should the testing be going well, with promotion complete, then this extra time will be spent either starting early on the dissertation, or adding extensions to the project to further explore the modelling of games at the type-level.

6.4.3 Weeks 15-18: 11th January to 7th February

Writing the dissertation is planned for this stretch of time; it will be planned out, section by section, with any relevant graphs, figures, and citations gathered.

The module CS324 Computer Graphics has a coursework due on the 20th January; as such, planning the dissertation could take longer than expected. However, since a month is set aside for just planning, delays are unlikely.

6.4.4 Weeks 19-22: 8th February to 7th March

Once the previous detailed planning stage is complete, writing shall begin; with a detailed enough plan, this section should not take longer than a month. An initial draft will be completed by the 7th of March.

The project itself will be evaluated during this period, examined for any subtle bugs, and the code will be cleaned up and completed. Work on the project presentation will also start in earnest, writing out a script.

6.4.5 Weeks 23-24: 8th March to 21st March

These two weeks are more empty space; set aside to act as a buffer for delays in dissertation writing. The project presentation is expected on the 19th March; so this empty space will ideally be spent completing the presentation and rehearsing it.

6.4.6 Weeks 25-29: 22nd March to 25th April

Drafting, re-drafting, and refining the dissertation (with the help of the supervisor) will take place during this time. This is the final stretch, and will be spent ensuring that the final piece of writing is as good as it can be.

Revision for examinations is key during this period; however, since the dissertation should be complete and this period is for evaluation and not the main bulk of writing, there should be ample time.

7 Resources, Risks, and Ethical Considerations

7.1 Resources

Developing this program requires a computer (and Operating System) which can run the latest version of GHC (8.10.2, at the time of writing). GHC runs on most Linux distributions, as well as Windows and MacOS. In other words, it runs on almost all consumer-grade PCs, and so development is unlikely to be a problem. The development environment will be the same as the "production" environment; it is a tool designed for those involved in the compilation process.

There are many Chess rulesets, casual or otherwise. The project will use the FIDE official ruleset, because it gives comprehensive, detailed rules that are widely used¹⁴.

7.2 Risks

The biggest risk going into the project was that it was unachievable; however, that risk has not materialised—the implementation is almost complete. The largest project-related risks left are the memory usage and compile time issues.

There are, of course, other risks not directly associated with the project; in 2020, there is a large-scale pandemic going on, which could pose considerable personal risk. However, by following Government and University guidelines, the chances of contracting COVID-19 will be kept to a minimum, and the project should not be interrupted by sudden illness.

7.3 Legal, Professional, and Ethical Considerations

As this project is undertaken alone, with no involvement from third parties (aside from the project supervisor), there are no additional legal, professional, or ethical considerations to consider. There will be no personal data processing, and no pre-existing Haskell type-level Chess implementation to plagiarise.

References

- [1] R. Nystrom, Game Programming Patterns, p. 6. Genever Benning, 2014.
- [2] B. A. Yorgey, S. Weirich, J. Cretin, S. Peyton Jones, D. Vytiniotis, and J. P. Magalhães, "Giving haskell a promotion," in *Proceedings of the 8th ACM SIGPLAN workshop on Types in language design and implementation*, pp. 53–66, 2012.
- [3] S. Maguire, Thinking With Types: Type-Level Programming In Haskell. 2018.

¹⁴It also includes rules for the players themselves, but these meta-rules are beyond the scope of the project.

- [4] E. Brady, "Embedded domain specific languages in idris," 2015. https://www.cs.ox.ac.uk/projects/utgp/school/idris-tutorial.pdf, pg. 51.
- [5] D. Szamozvancev and M. B. Gale, "Well-typed music does not sound wrong (experience report)," in *Proceedings of the 10th ACM SIGPLAN International Symposium on Haskell*, pp. 99–104, 2017.
- [6] P. Thiemann, "An embedded domain-specific language for type-safe server-side web scripting," *ACM Trans. Internet Technol.*, vol. 5, p. 1–46, Feb. 2005.
- [7] H. Mu and S. Jiang, "Design patterns in software development," in 2011 IEEE 2nd International Conference on Software Engineering and Service Science, pp. 322–325, 2011.
- [8] R. A. Eisenberg and S. Weirich, "Dependently typed programming with singletons," *ACM SIGPLAN Notices*, vol. 47, no. 12, pp. 117–130, 2012.
- [9] G. J. Sussman and G. L. Steele, "Scheme: A interpreter for extended lambda calculus," *Higher-Order and Symbolic Computation*, vol. 11, no. 4, pp. 405–439, 1998.
- [10] D. Szamozvancev, "Well-typed music does not sound wrong," 2017.
- [11] S. Hengel, "Behavior-driven development in haskell,"