Universidad de La Laguna
Escuela Superior de Ingeniería y Tecnología
Sección de Ingeniería Informática

Construcción de un Compilador y un Intérprete de Scheme Usando Haskell

Building a Compiler and Interpreter for Scheme Using Haskell

Francisco Nebrera Perdomo

La Laguna, 6 de junio de 2015

D. **Casiano Rodríguez León**, con N.I.F. 42.020.072-S profesor Titular de Universidad adscrito al Departamento de Ingeniería Informática y de Sistemas de la Universidad de La Laguna, como tutor

C E R T I F I C A

Que la presente memoria titulada:

"*Construcción de un Compilador y un Intérprete de Scheme Usando Haskell.*"

ha sido realizada bajo su dirección por D. **Francisco Nebrera Perdomo**, con N.I.F. 79.064.507-Y.

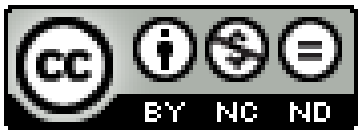
Y para que así conste, en cumplimiento de la legislación vigente y a los efectos oportunos firman la presente en La Laguna a 6 de junio de 2015

Agradecimientos

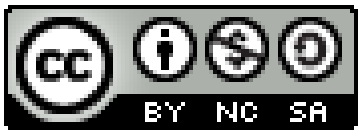
Casiano Rodríguez León
Jorge Riera Ledesma
Luz Marina Moreno de Antonio
Francisco de Sande González
Francisco Carmelo Almeida Rodríguez
Blas C. Ruiz
F. Gutiérrez
P. Guerrero
E. Gallardo
Daniel Díaz Casanueva

Licencia

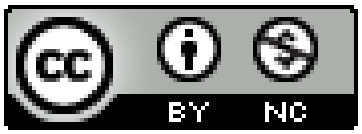
* Si NO quiere permitir que se compartan las adaptaciones de tu obra y NO quieres permitir usos comerciales de tu obra indica:



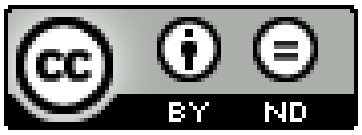
* Si quiere permitir que se compartan las adaptaciones de tu obra mientras se comparta de la misma manera y NO quieres permitir usos comerciales de tu obra indica:



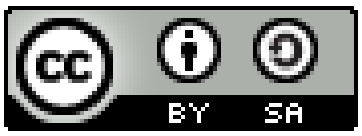
* Si quiere permitir que se compartan las adaptaciones de tu obra y NO quieres permitir usos comerciales de tu obra indica:



*Si NO quiere permitir que se compartan las adaptaciones de tu obra y quieres permitir usos comerciales de tu obra indica:



* Si quiere permitir que se compartan las adaptaciones de tu obra mientras se comparta de la misma manera y quieres permitir usos comerciales de tu obra (licencia de Cultura Libre) indica:



* Si quiere permitir que se compartan las adaptaciones de tu obra y quieres permitir usos comerciales de tu obra (licencia de Cultura Libre) indica:

Resumen

El objetivo de este trabajo ha sido crear un compilador e intérprete del lenguaje Scheme, usando Haskell. Para ello se ha hecho uso de un parser monádico como es Parsec.

La competencia [E6], que figura en la guía docente, indica que en la memoria del trabajo se ha de incluir: antecedentes, problemática o estado del arte, objetivos, fases y desarrollo del proyecto, conclusiones, y líneas futuras.

Se ha incluido el apartado de 'Licencia' con todas las posibles licencias abiertas (Creative Commons). En el caso en que se decida hacer público el contenido de la memoria, habrá que elegir una de ellas (y borrar las demás). La decisión de hacer pública o no la memoria se indica en el momento de subir la memoria a la Sede Electrónica de la ULL, paso necesario en el proceso de presentación del TFG.

El documento de memoria debe tener un máximo de 50 páginas.

No se deben dejar páginas en blanco al comenzar un capítulo, ya que el documento no está pensado para se impreso sino visionado con un lector de PDFs.

También es recomendable márgenes pequeños ya que, al firmar digitalmente por la Sede, se coloca un marco alrededor del texto original.

El tipo de letra base ha de ser de 14ptos.

Palabras clave: Palabra reservada1, Palabra reservada2, ...

Abstract

Here should be the abstract in a foreing language...

Keywords: *Keyword1, Keyword2, Keyword2, ...*

Índice general

Índice de figuras

Índice de tablas

Capítulo 1

Introducción

Todo empezó con un hilo en el foro de internet forocoches.com (`https://www.forocoches.com`), en él hablaban de que la programación funcional iba a tener cada día más relevancia porque cuenta con ventajas de las cuales la imperativa carece.

A raíz de ello, me interesé por este paradigma y empecé a (e incluso terminé de) leer numerosos libros sobre el lenguaje y la programación funcional en general, y a crear pequeños programas en Haskell. Haskell es muy interesante debido a que es el lenguaje con mayor nivel de abstracción en el que he programado hasta hoy.

Haskell es idóneo para crear lenguajes de dominio específico. En otras palabras, antes de escribir un compilador se captura el lenguaje a compilar (el lenguaje fuente) en un tipo. Las expresiones de ese tipo representarán términos en el lenguaje fuente y normalmente son bastante similares al mismo, a pesar de ser, realmente, tipos de Haskell.

Luego se representa el lenguaje objetivo como otro tipo más. Finalmente, el compilador es realmente una función del tipo fuente al tipo objetivo y las traducciones son fáciles de escribir y leer. Las optimizaciones también son funciones como cualquier otra (ya que realmente en Haskell todo es una función, y además, currificada) que mapean del dominio del lenguaje fuente al codominio del lenguaje objetivo.

Por ello los lenguajes funcionales con sintaxis ligera y un fuerte sistema de tipos se consideran muy adecuados para crear compiladores y muchas otras cosas cuya finalidad es la traducción.

Además, Haskell cuenta con mecanismos de abstracción muy fuertes que permiten escribir códigos escuetos que se comportan muy bien, como por ejemplo:

- Reconocimiento de patrones
- Tipos de datos algebraicos (generalizados o no)
- Lambdas (y por ello, mónadas)
- Plegados de listas

A continuación se describen los mencionados constructos en su sección correspondiente.

1.1. Reconocimiento de patrones y tipos de datos algebraicos

Para entender qué es el reconocimiento de patrones primero debemos saber qué es **casar**. Para ello usaremos las acepciones pertinentes del diccionario de la Real Academia:

- Dicho de dos o más cosas: Corresponder, conformarse, cuadrar.
- Unir, juntar o hacer coincidir algo con otra cosa. Casar la oferta con la demanda.

- Disponer y ordenar algo de suerte que haga juego con otra cosa o tengan correspondencia entre sí.

Es un término que se usa bastante en las expresiones regulares, para ver si una expresión casa con un texto dado, y en qué lugar. Veamos un ejemplo de reconocimiento de patrones:

```
dime :: Int -> String
dime 1 = "¡Uno!"
dime 2 = "¡Dos!"
dime 3 = "¡Tres!"
dime 4 = "¡Cuatro!"
dime 5 = "¡Cinco!"
dime x = "No está entre 1 y 5"
```

La función **dime** hace reconocimiento de patrones con su primer argumento, de tipo **Int**, y va de arriba a abajo intentando encontrar una coincidencia. Cuando recibe un número entre 1 y 5, lo canta con ahínco, si no lo encuentra, nos devolverá un mensaje diciéndonoslo. Notar además que si hubiéramos puesto la línea **dime x = "No está entre 1 y 5"** al principio, nuestra función siempre devolvería "No está entre 1 y 5", aun siendo cierto. Por tanto, debemos ordenar los patrones por probabilidad; de los menos probables a los más probables.

Cuando hablamos de reconocimiento de patrones hablamos, en realidad, de reconocimiento de constructores. En concreto en Haskell existen dos tipos de constructores, los constructores de tipos (los tipos que aparecen en las declaraciones de las funciones) y los constructores de valor (aquellos que se suelen poner entre paréntesis, y son funciones que recibiendo un valor crean un tipo que encapsula dicho valor).

Importante: los dos guiones al principio de cada línea representan comentarios en Haskell, y su presencia aquí es de carácter didáctico.

```
data Persona = CrearPersona String Int
                            |      |
                            |      |
                            |      La edad de la persona
                            El nombre de la persona
```

A la izquierda del igual está el constructor de tipos. A la derecha del igual están los constructores de datos. El constructor de tipos es el nombre del tipo y usado en las declaraciones de tipos. Los constructores de datos son funciones que producen valores del tipo dado. Si solo hay un constructor de datos, podemos llamarlo igual que el de tipo, ya que es imposible sustituirlos sintácticamente (recuerda, los constructores de tipos van en las declaraciones, los constructores de valor en las ecuaciones).

```
data Persona = Persona String Int
     |         |
     |         |
     |         Constructor de valor (o datos)
     El nombre de la persona
```

El tipo del último ejemplo se conoce como **tipo de dato algebraico**; tipos de datos construidos mediante la combinación de otros tipos. El reconocimiento de patrones es una manera de desestructurar un tipo de dato algebraico, seleccionar una ecuación basada en su constructor y luego enlazar los componentes a variables. Cualquier constructor puede aparecer en un patrón; ese patrón casa con un valor si la etiqueta del patrón es la misma que la etiqueta del valor y todos los subpatrones casan con sus correspondientes componentes.

Importante: el reconocimiento de patrones es en realidad reconocimiento de constructores.

1.2. Variables de tipo

Las variables de tipo son aquellas que se declaran en **data** después del nombre del tipo que vamos a crear. Su finalidad principal es hacer saber qué puede formar parte del tipo, y además permitir a cualquier tipo formar parte de nuestro tipo personalizado. Veámoslo con un ejemplo:

```
data Persona a = PersonaConCosa String a | PersonaSinCosa String
               |                         |
               |                         |
               |                        Podemos usarla aquí
              Introduciendo una variable de tipo aquí
```

En los siguientes ejemplos se ilustra el deber de informar al compilador qué tipo queremos que nuestra función devuelva, y así producir un tipo **Persona Int**, **Persona String**,...,etc.

```
franConEdad :: Persona Int
franConEdad = PersonaConCosa "fran" 25

franSinEdad :: Persona Int
franSinEdad = PersonaSinCosa "fran"
```

Ahora llega el reconocimiento de patrones propiamente dicho; según se encuentre el constructor **PersonaConCosa String a** ó **PersonaSinCosa String**, nuestra función debe ser programada para actuar en consecuencia:

```
getNombre :: Persona Int -> String
getNombre (PersonaConCosa nombre _) = nombre
getNombre (PersonaSinCosa nombre)   = nombre

getEdad :: Persona Int -> Maybe Int
getEdad (PersonaConCosa _ edad) = Just edad
getEdad (PersonaSinCosa _)      = Nothing
```

Como vemos, a las variables **nombre** y **edad** respectivamente se le han enlazado sus valores reales, que son los que nuestra función devuel-

ve. Como el constructor **PersonaSinCosa** sólo contiene el nombre y no la edad, utilizamos el tipo **Maybe** para devolver **Nothing** en caso de que ese patrón (constructor) sea reconocido. En el otro caso, devolvemos **Just edad** ya que en este caso la tenemos.

1.3. Lambdas

En un lenguaje funcional, poner nombre a todas las funciones que usemos podría resultar tedioso. Además, es cómodo definir funciones al vuelo, es decir, rápidamente y en el punto del programa en el que realmente sea necesario. Las lambdas sirven para este propósito.

Las lambdas se suelen declarar entre paréntesis para que el compilador sepa que se tratan de un "todo". No obstante, en el caso de las mónadas hay veces en que no son demasiado necesarios y se va resolviendo todo mediante la indentación.

La sintaxis de las lambdas es la siguiente:

```
\arg1 arg2 ... argn -> cuerpo_función
```

Es decir, para que Haskell sepa que estamos trabajando con una lambda, se usa la backslash \ y a continuación se encuentran dos partes bien diferenciadas, separadas por una flecha →:

- A la izquierda de la flecha → la lista de nombres de argumentos, separados por espacios.
- A la derecha de la flecha → el cuerpo de la función. El tipo de esa expresión será el tipo retorno de la lambda.

Definamos la lambda más sencilla que existe, lo único que hace es devolver su argumento:

```
\x -> x
```

Definamos una lambda que eleve al cubo un número:

```
\x -> x*x*x
```

Veamos ahora una lambda que sume sus dos argumentos:

```
\x y -> x + y
```

Y por último, veamos una que ignora su primer argumento y devuelva el segundo:

```
\_ x -> x
```

Como vemos, se puede usar el patrón subrayado (barra baja) para expresar que no nos importa el valor del primer parámetro, ya que sólo usamos el segundo. Las lambdas tienen mucha importancia en Haskell, y son muy útiles.

1.4. Plegados de listas

Pondremos un ejemplo real codificado por mí, un DFA hecho mediante un plegado de listas por la izquierda.

```
probarDFA :: DFA -> [Char] -> Bool
probarDFA (DFA i a t) = a . foldl' t i
```

Un DFA se podría implementar en programación imperativa con un bucle for que fuera sobreescribiendo el estado en cada iteración, haciendo un lookup en su tabla de estados dependiendo de su estado actual y el símbolo leído.

Esto, en Haskell, se puede hacer usando la función **foldl** (aunque aquí por razones de rendimiento y uso de memoria se ha optado por **foldl'**).

Se trata de empezar con un acumulador (en este caso, el estado inicial), y nos vamos moviendo por la lista (cadena de entrada) de izquierda a derecha, haciendo un lookup con el acumulador y el carácter leído en ese instante, y luego el resultado (estado siguiente) se convertirá en el nuevo acumulador y se repetirá el proceso.

La función **scanl** nos permite ver la lista de todos los valores que ha ido tomando el acumulador durante la ejecución del programa, y se comporta como **foldl** pero devolviendo la lista completa:

```
*DFA> leerDFA "dfa1.txt"
Cadena:AAABABABA
["Q1","Q2","Q1","Q2","Q2","Q2","Q1"]
```

Luego de la aplicación de **foldl'** vemos un punto, que significa composición, es decir, aplicará la función **a** al resultado de **foldl'**, donde **a** es una función que comprueba si ese estado pertenece a la lista de estados finales, devolviendo un booleano que indicará la aceptación o rechazo de la cadena por el autómata.

Como vemos, en Haskell con una sóla línea se pueden hacer virguerías, el programa completo que simula un DFA, leyendo desde fichero y pidiendo continuamente entrada tras computar la anterior, ocupa 35 líneas.

Capítulo 2

Librería Parsec

2.1. Introducción a Parsec

En el capítulo anterior se ha realizado una introducción a la programación funcional con Haskell, y ahora se describirá un módulo muy útil en la creación del intérprete, Parsec.

Parsec es un módulo de Haskell, un conjunto de funciones exportables que suelen tener una finalidad común y se pueden importar en otros programas. En nuestro intérprete hemos importado Parsec, pues es el módulo utilizado en el tutorial que seguimos, Write Yourself a Scheme in 48 hours.

Parsec se diseñó desde cero como una librería de parsers con capacidades industriales. Es simple, segura, está bien documentada, provee de buenos mensajes de error y es rápida. Se define como un transformador de mónadas que puede ser apilado sobre mónadas arbitrarias, y también es paramétrico en el tipo de flujo de entrada. La documentación de la versión usada en el presente Trabajo Fin de Grado se puede consultar online en `https://hackage.haskell.org/package/parsec-3.1.9`

Parsec se puede leer en "inglés plano"(siempre que nuestros parsers

tengan los nombres adecuados). Se pueden hacer analogías entre las funciones de Parsec y las expresiones regulares, como veremos en el ejemplo de código de este capítulo.

La mejor manera de describir el módulo es mediante un ejemplo, un parser del conocido formato JSON.

2.2. Librerías necesarias

```
import Text.ParserCombinators.Parsec hiding ((<|>), many)
import Text.Parsec.Numbers (parseFloat)
import Control.Applicative
import Control.Monad
import Prelude hiding (Null, null)
```

2.3. Formato a parsear

El formato JSON (JavaScript Object Notation) es de los más comunes hoy en día para el intercambio de información a través de la red. Es un formato sencillo y fácil de parsear, y por ello está ganando terreno frente a su competidor principal, XML. Sus principales elementos son:

- Number
- String
- Boolean
- Array
- Object
- Null

La principal aplicación de Haskell siempre han sido los parsers, tanto para compiladores como para propósito general.

Empezaremos definiendo los parsers más sencillos, cuyo fin es devolver argumentos que entrarán en constructores de valor para tipos de JSON que siempre valgan lo mismo. Estos 3 tipos son: **true**, **false** y **null**.

```
alwaysTrue :: Parser Bool
alwaysTrue = pure True

alwaysFalse :: Parser Bool
alwaysFalse = pure False

alwaysNull :: Parser String
alwaysNull = pure "null"
```

La misión de *pure* :: $a\ f\ a$ (donde en este caso **f** es la mónada **Parser**) no es otra que envolver los dos **Bool** y la **String** en un valor monádico, devolviendo de este modo un **Parser Bool** o un **Parser String**. Por tanto, estas funciones devuelven un **Parser**, que cuando se ejecuta (mediante la función **parse**) devuelve un **Bool** o una **String**.

Ahora lo que debemos hacer es usar el parser **string**, que intenta casar con una cadena dada, devolviéndola en caso de que consiga casar:

```
matchTrue :: Parser String
matchTrue = string "true"

matchFalse :: Parser String
matchFalse = string "false"

matchNull :: Parser String
matchNull = string "null"
```

Por último, no devolveremos la cadena propiamente dicha, sino un valor puro (por ello antes definimos funciones que usan **pure**):

```
boolTrue :: Parser Bool
boolTrue = matchTrue *> alwaysTrue

boolFalse :: Parser Bool
boolFalse = matchFalse *> alwaysFalse

null :: Parser String
null = matchNull *> alwaysNull
```

Aquí usamos un operador de la clase de tipos **Applicative**, que en inglés se suele llamar "star arrow". Este operador ejecuta primero el parser de la izquierda, luego el de la derecha, y devuelve sólo lo que parsee el de la derecha (el sitio al que apunta la flecha).

Ahora veamos qué pasa si un token puede pertenecer a un tipo aún más general:

```
bool :: Parser Bool
bool = boolTrue <|> boolFalse
```

$<|>$ es el operador de elección, y se parece mucho a la barra vertical | de las expresiones regulares. Pueden ser encadenados tantos parsers como queramos. Este operador lo que hace es:

- 1. intenta el parser de la izquierda, que no debería consumir entrada...(ver **try**)
- 2. intenta el parser de la derecha.

Si el parser de la izquierda consume entrada, podríamos usar **try**, el cual intenta ejecutar ese Parser, y, si falla, vuelve al estado anterior, es decir, deja la entrada sin consumir. Sólo funciona a la izquierda de

$< | >$, es decir, si queremos encadenar varios **try**, deben estar a la izquierda de la cadena de $< | >$.

try es como un lookahead, y se puede ver como algo para procesar cosas de manera atómica, **try** es realmente backtracking, y por ello no es demasiado eficiente.

Como en este caso las string que vamos a parsear no tienen prefijos coincidentes, no hace falta usar **try** por si hay que volver a empezar.

Ahora empezaremos a ver algo que se parece aún más a las expresiones regulares:

```
stringLiteral :: Parser String
stringLiteral = char '"' >> (many (noneOf ['"'])) << char '"'
```

Aquí vemos que Parsec puede leerse casi en "inglés plano", ya que esta línea casi se autodescribe. Primero, debe encontrarse un carácter ¨, luego vemos la función **many**, que equivale a la estrella * de las expresiones regulares, es decir, podría haber muchos, uno o ninguna ocurrencia del parser que reciba **many**.

Luego vemos una función noneOf, que es un parser que acepta todo menos los carácteres que pertenezcan a una lista determinada, en este caso acepta todo menos las comillas dobles, en caso de toparse con comillas dobles (la cadena ha acabado), deja de consumir entrada.

Para terminar, se vuelve a parsear un carácter ¨, que debe estar obligatoriamente. Ahora vemos que nuestra combinación aplicativa sigue una estructura $\mathbf{a} > * \mathbf{b} < * \mathbf{c}$, esto indica que los parsers **a**, **b** y **c** deben tener éxito, pero como sólo se devuelve lo que está apuntado por las flechas, sólo devolveremos lo que haya parseado b, que en este caso corresponde a **(many (noneOf ['"']))**.

De modo que Parsec, como la programación funcional, se basa en hacer sencillas funciones que sean buenas en lo suyo, e irlas combinando para realizar tareas cada vez más complejas. Esta es la base de la filosofía KISS tan popular en los sistemas POSIX.

La siguiente línea da error de tipos:

```
value = bool <|> stringLiteral
```

Solución: crear un tipo algebraico que contenga Bool y String (entre otros). Los tipos de datos algebraicos son una herramienta muy útil para los parsers, ya que permiten saber exactamente a qué tipo pertenece el token que hemos parseado.

```
data JSONValue = B Bool
               | S String
               | N Double --número de JSON
               | A [JSONValue] --array de JSON
               | O [(String, JSONValue)] --objeto de JSON
               | Null String
              deriving Show
```

Como vemos, tenemos un sólo constructor de tipo, **JSONValue**, es decir, nuestros parsers tendrán tipo **Parser JSONValue**. Sin embargo, tenemos 6 constructores de valor, que por simplicidad son simplemente las letras Iniciales de cada tipo de valor a parsear, salvo **Null**, en el cual se usó el nombre completo ya que **N** se usó para el tipo Number de JavaScript.

Veamos ahora el parser principal, es decir, un parser genérico capaz de parsear cualquier valor de JSON:

```
jsonValue :: Parser JSONValue
jsonValue = spaces >> (jsonNull
                    <|> jsonBool
                    <|> jsonStringLiteral
                    <|> jsonArray
                    <|> jsonObject
                    <|> jsonNumber
                    <?> "JSON value")
```

No te preocupes demasiado por el parser **spaces**, lo explicaré más adelante en conjunto con **lexeme**. Pero, ¿qué es esa interrogación? $<?>$ es un combinador que permite dar mensajes de error en caso de parseo fallido. En este caso, se le pasa una String con el mensaje de error que queremos que aparezca. En caso de error, saldrá algo como "expected JSON value", pues ese es el argumento de $<?>$ para cuando falle el parser **jsonValue**.

Lo malo de esto es que seguimos teniendo error de tipos porque:

```
bool :: Parser Bool
stringLiteral :: Parser String
```

Lo bueno es que con $(<\$>) :: Functor f => (a -> b) -> fa -> fb$, que en este caso tendría el tipo: $(<\$>) :: (a -> b) -> Parser a -> Parser b$, podemos solucionarlo.

Recordemos que los constructores de valor son en realidad funciones como otra cualquiera (salvo que empiezan por mayúscula). Por ejemplo, el tipo de **B** es $B :: Bool -> JSONValue$. Por tanto, si hacemos **B ¡$¿bool** tendremos como resultado un **Parser JSONValue**, y eso haremos en todos nuestros parsers anteriormente nombrados.

```
jsonBool'' :: Parser JSONValue
jsonBool'' = B <$> bool

matchNull'' = lexeme matchNull'

jsonStringLiteral :: Parser JSONValue
jsonStringLiteral = lexeme (S <$> stringLiteral)
```

Aquí lo único que nos llama la atención es el parser **lexeme**. **lexeme** está definido en Parsec por defecto, pero nosotros lo programaremos más que nada por razones didácticas.

lexeme es un parser que, recibiendo otro parser, devuelve un parser del mismo tipo, pero que consume todos los espacios (incluyendo tabuladores y newlines) que haya detrás del token parseado.

```
ws :: Parser String -- whitespace
ws = many (oneOf " \t\n")

lexeme :: Parser a -> Parser a
lexeme p = p <* ws
```

De este modo, con aplicar lexeme a cada uno de los parsers que vayamos a usar, tenemos resuelto el problema de los espacios entre tokens.

Bueno, ahora que el problema de los espacios está resuelto...¡sorpresa! no lo está del todo...Como hemos dicho, el combinador **lexeme** se çome"todos los espacios, tabuladores o newlines que encuentre después del token parseado. Pero, ¿y si esos espacios estuvieran antes del primer token que llegamos a parsear? Probablemente se produciría un error.

Solución: añadir el parser spaces a nuestro parser principal **jsonValue**. Esto se hizo mediante el operador monádico >>, que en la mónada de Parsec tiene el efecto de ejecutar ese parser, y si tiene éxito no guardar el resultado del parsing, sino pasar al siguiente. Se ha usado >> para

ilustrar el uso de esta función, ya que se había introducido antes: $*>$, también llamado "star arrow".

A continuación creemos un parser que permita parsear números. Para ello usaremos la función **parseFloat**, que permite parsear cualquier tipo de número, incluso con signo, exponente, parte decimal...es decir, el formato de coma flotante.

```
jsonNumber :: Parser JSONValue
jsonNumber = N <$> parseFloat
```

¡Listo! ya tenemos un parser más. Ahora veamos algo un poco más complejo, los arrays de JSON. Un array de JSON tiene el siguiente formato:

```
[
    {"firstName":"John", "lastName":"Doe"},
    {"firstName":"Anna", "lastName":"Smith"},
    {"firstName":"Peter","lastName":"Jones"}
]
```

Como vemos, tenemos:

- 1. Un carácter abrir corchetes [
- 2. Un conjunto de tokens de JSON, separados por comas.
- 3. Un carácter cerrar corchetes]

Sabido esto, lo único nuevo que tenemos que introducir aquí es el parser **sepBy**. **sepBy** recibe dos argumentos, el primero es el parser que se usará para cada token, y el segundo es el parser que se usará para el separador o separadores. Veamos el parser completo.

```
array :: Parser [JSONValue]
array =
```

```
  (lexeme $ char '[')
  *>
  ( jsonValue `sepBy` (lexeme $ char ',') )
  <*
  (lexeme $ char ']')
```

Como los arrays contienen tokens de JSON, lo que hacemos es una llamada recursiva a **jsonValue**. De este modo, vemos que dentro de un array de JSON puede haber "lo que sea"(siempre que esté correctamente escrito y formateado) pero el array debe empezar por el carácter [y terminar con] para garantizar que dicho formato sea correcto. Como vemos, este parser nos devuelve una lista de **JSONValue**, y eso no es un tipo **JSONValue**. Por tanto, debemos aplicar $fmap(< \$ >)$, en este caso de manera infija:

```
jsonArray :: Parser JSONValue
jsonArray = A <$> array
```

Ahora parsearemos algo parecido pero no del todo igual, los objetos de JSON. El formato de los objetos es:

- 1. Un carácter abrir llaves **{**
- 2. Una lista de pares separador por el carácter dos puntos ':'
- 3. Un carácter cerrar llaves **}**

```
jsonObject :: Parser JSONValue
jsonObject = O <$> ((lexeme $ char '{') *>
                    (objectEntry `sepBy` (lexeme $ char ','))
                    <* (lexeme $ char '}'))

objectEntry :: Parser (String, JSONValue)
objectEntry = do
  key <- lexeme stringLiteral
  char ':'
  value <- lexeme jsonValue
  return (key, value)
```

Ahora consigamos que el parser **jsonBool'** sea capaz de lidiar con espacios, tabuladores y nuevas líneas después del token que parsea:

```
jsonBool' = lexeme jsonBool''
```

Ya casi hemos terminado, pero aún falta un pequeño detalle. ¿Y si alguien se equivoca y escribe por ejemplo "falseee", o "nullpointer", o cualquier otra cosa siguiendo a las palabras reservadas **true**, **false** o **null**? Nuestro parser lo aceptaría, cuando eso no debería ser así. Queremos exactamente esas palabras, ni un carácter más ni uno menos, para que nuestro parser sea correcto. Para ello, Parsec nos provee con un parser que falla en caso de que otro esté seguido de ciertos carácteres, es **notFollowedBy**. **notFollowedBy** recibe un parser, y si éste tiene éxito, falla. Un ingenioso truco que nos saca del atolladero de manera muy sencilla y casi autoexplicativa.

```
jsonBool :: Parser JSONValue
jsonBool = jsonBool' <* notFollowedBy alphaNum

jsonNull :: Parser JSONValue
jsonNull = matchNull'' <* notFollowedBy alphaNum
```

Por último, creemos la función **main** que nos permitirá compilar el programa. Para ello seguiremos los siguientes pasos:

- 1. Mostrar por pantalla qué queremos.
- 2. Obtener el nombre del fichero por entrada estándar (teclado) y ligarlo al nombre **filename**.
- 3. Aplicar nuestro parser principal (**jsonValue**) a nuestro fichero **filename** mediante la función **parseFromFile**.

```
main = do
  putStr "Nombre_fichero: "
  filename <- getLine
  parseFromFile jsonValue filename
```

Eso es todo, ya tenemos nuestro parser de JSON funcionando.

Capítulo 3

El lenguaje Scheme

3.1. Introducción a Scheme

El lenguaje Scheme es un dialecto de Lisp. Es un lenguaje que, como Lisp, es funcional. Es un lenguaje relativamente sencillo de aprender, con un poco de estudio se puede llegar a tener un conocimiento medio en poco tiempo.

Scheme fue elegido para escribir uno de los mejores libros sobre Ciencias de la Computación que se han escrito: Structure and Interpretation of Computer Programs, de la editorial MIT Press. Asimismo, para mis pruebas he usado el intérprete de Scheme del MIT, llamado mit-scheme.

El lenguaje usa la notación polaca para las llamadas a funciones y para cualquier tipo de evaluación. Además, cualquier definición o llamada a una función deberá hacerse entre paréntesis. Asimismo, las listas, contengan lo que contengan, se representan como una lista de ciertos valores de cierto tipo, entre paréntesis y separados por espacios.

Veamos una sesión con el intérprete oficial de Scheme para familiarizarnos un poco con el lenguaje:

```
(+ 3 5)
8
```

```
(sqrt 144)
12
```

Las funciones se definen mediante la palabra reservada **define**, una lista que contiene el nombre de la función y la lista de argumentos (todos separados por espacios), y, por último, el cuerpo de la función. Todo esto debe estar entre paréntesis.

```
(define (factorial n)
  (if (= n 0)
    1
    (* n (factorial (- n 1)))))
```

Las funciones **car** y **cdr** actúan sobre una lista. La función **car** devuelve el primer elemento de la lista (si éste existe), mientras que **cdr** devuelve una lista compuesta por el segundo elemento de la lista y todos los que le siguen.

```
(define (rev lst) (rev-rec lst ()))

(define (rev-rec lst acc)
  (cond ((null? lst) acc)
        (else (rev-rec (cdr lst)
                       (cons (car lst) acc)))
        ))
```

3.2. Segundo apartado de este capítulo

3.3. Tercer apartado de este capítulo

Capítulo 4

Título del Capítulo Cuatro

En el capitulo 1 se describio bla, bla, bla.....

Capítulo 5

Conclusiones y trabajos futuros

Este capítulo es obligatorio. Toda memoria de Trabajo de Fin de Grado debe incluir unas conclusiones y unas líneas de trabajo futuro

Capítulo 6

Summary and Conclusions

This chapter is compulsory. The memory should include an extended summary and conclusions in english.

6.1. First Section

Capítulo 7

Presupuesto

Este capítulo es obligatorio. Toda memoria de Trabajo de Fin de Grado debe incluir un presupuesto.

7.1. Sección Uno

Tipos	**Descripcion**
AAAA	BBBB
CCCC	DDDD
EEEE	FFFF
GGGG	HHHH

Tabla 7.1: Tabla resumen de los Tipos

Apéndice A

Título del Apéndice 1

A.1. Algoritmo XXX

```
*****************************************************************************
*
* Fichero .h
*
*****************************************************************************
*
* AUTORES
*
*
* FECHA
*
*
* DESCRIPCION
*
*
*****************************************************************************/
```

A.2. Algoritmo YYY

```
/*****************************************************************************
 *
 * Fichero .h
 *
```

```
*****************************************************************************
*
* AUTORES
*
* FECHA
*
* DESCRIPCION
*
*
*****************************************************************************/
```

Apéndice B

Título del Apéndice 2

B.1. Otro apendice: Seccion 1

Texto

B.2. Otro apendice: Seccion 2

Texto

Bibliografía

[1] D. H. Bailey and P. Swarztrauber. The fractional Fourier transform and applications. *SIAM Rev.*, 33(3):389–404, 1991.

[2] Miran Lipovaca. *Learn You a Haskell for Great Good!: A Beginner's Guide*. No Starch Press, 2011.

[3] Blas C. Ruiz. *Razonando con Haskell. Un curso sobre programación funcional*. Thomson, 2004.

[4] David D. Spivak. *Category Theory for the Sciences*. MIT Press, 2014.