Tesis de Licenciatura

λPage

Un bloc de notas para desarrolladores Haskell

Departamento de Computación
Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires

Alumno

Fernando Benavides (LU 470/01)

greenmellon@gmail.com

Directores

Dr. Diego Garbervetsky

Lic. Daniel Gorín

Abstract

El presente documento describe una herramienta para desarrolladores *Haskell* que pretende facilitar la tarea de "debuggear", analizar y entender código, llamada λ***Page***. Con ella el usuario puede manipular "páginas" de texto libre que contengan expresiones *Haskell*, intentar interpretar éstas expresiones independientemente y analizar los resultados obtenidos.

Índice

1. Introducción

1.1. Motivación

> Motivation is what gets you started. Habit is what keeps you going
>
> Jim Rohn
>
> Essstamo mo-ti-va-dos, nene
>
> El "Bambino" Veira

Actualmente estamos presenciando un importante cambio en el desarrollo de sistemas, gracias al éxito de proyectos como CouchDB[1], ejabberd[2] y el chat de Facebook[3], todos ellos desarrollados utilizando lenguajes del paradigma funcional.

Ejemplos de éstos lenguajes de programación, como Haskell[4] o Erlang[5], demuestran ser maduros, confiables y presentan claras ventajas en comparación con los lenguajes tradicionales del paradigma imperativo. Sin embargo, los desarrolladores que deciden realizar el cambio de paradigma se encuentran con el problema de la escasez de ciertas herramientas que les permitan realizar su trabajo más eficientemente. Por el contrario, éstas herramientas abundan en el desarrollo de proyectos utilizando lenguajes orientados a objetos. En particular, nuestro foco de atención se centra sobre aquellas herramientas que permiten realizar *debugging* y *entendimiento* de código a través de *"micro-testing"*[6] .

Los desarrolladores Haskell cuentan actualmente con dos herramientas de este tipo:

GHCi[7] La consola que provee GHC[8] permite a los desarrolladores evaluar expresiones, verificar su tipo o su clase. Cuenta también con un mecanismo de debugging[9] integrado que permite realizar la evaluación de expresiones paso a paso. Pese a ser la herramienta más utilizada por los desarrolladores, *GHCi* tiene varias limitaciones. En particular:

- No permite editar más de una expresión a la vez
- No permite intercalar expresiones con definiciones
- Si bien permite utilizar definiciones, éstas se pierden al recargar módulos
- No es sencillo utilizar en una sesión las definiciones y/o expresiones creadas en sesiones anteriores

[1] http://couchdb.apache.org
[2] http://www.ejabberd.im
[3] http://www.facebook.com
[4] http://www.haskell.org
[5] http://www.erlang.org
[6] Entiéndase "micro-testing" como la tarea de realizar tests eventuales para entender o evaluar algún aspecto de un programa
[8] http://www.haskell.org/ghc
[9] http://www.haskell.org/ghc/docs/6.10-latest/html/users_guide/ghci-debugger.html

Hat[10] Un herramienta para realizar seguimiento a nivel de código fuente. A través de la generación de trazas de ejecución, *Hat* ayuda a localizar errores en los programas y es útil para entender su funcionamiento. Sin embargo, por estar basado en la generación de trazas, requiere la compilación y ejecución de un programa para poder utilizarlo y esto no siempre es cómodo para el desarrollador que puede querer simplemente analizar una expresión particular que incluso quizá no compile aún. Además, su mantenimiento activo parece haber cesado hace más de un año y en su página se observa una importante lista de problemas conocidos[11] y características deseadas[12].

1.2. Trabajos Relacionados

> I like work; it fascinates me. I can sit and look at it for hours
>
> Jerom Klapka

En el mundo de la programación orientada a objetos podemos encontrar herramientas de este tipo, como Java Scrapbook Pages[13] para Java[14] y Workspace[15] para SmallTalk[16]. Utilizando estos aplicativos, los desarrolladores pueden introducir pequeñas porciones de código, ejecutarlas y luego inspeccionar y analizar los resultados obtenidos. Un concepto compartido por ambas herramientas es el de presentar "páginas" de texto en las que varias expresiones pueden intercalarse con partes de texto libre y permitir al desarrollador intentar evaluar sólo una porción de todo lo escrito. Estas páginas pueden ser guardardas y luego recuperadas de modo de poder analizar nuevamente las mismas expresiones. Además permiten crear objetos (lo que para los lenguajes funcionales equivaldría a definir expresiones) locales a la página en uso y utilizarlos en ella.

Dentro del paradigma funcional, con un enfoque similar, aunque un poco más orientado a la presentación y visualización de documentos, Keith Hanna[17] de la Universidad de Kent, ha desarrollado Vital[18]. *Vital* es una implementación de un entorno de visualización de documentos para *Haskell*. Pretende presentar *Haskell* de una manera apropiada para usuarios finales en areas de aplicación como la ingeniería, las matemáticas o las finanzas. Dentro de esta herramienta, los módulos *Haskell* son presentados como documentos en los que pueden visualizarse los valores que en ellos se definen directamente en el lugar en el que aparecen, ya sea de modo textual o gráfico (como "vistas").

Durante el desarrollo de ***λPage*** hemos tenido que enfrentar varios desafíos relacionados principalmente con el desarrollo de interfaces visuales dentro del paradigma funcional. Volcando el conocimiento adquirido durante ese proceso, hemos desarrollado wxhNotepad[19] que es, ante todo, una prueba de concepto sobre cómo desarrollar

[11] http://www.haskell.org/hat/bugs.html
[12] http://www.haskell.org/hat/bugs.html
[13] http://help.eclipse.org/help33/index.jsp?topic=/org.eclipse.jdt.doc.user/reference/ref-34.htm
[14] http://www.java.com
[15] http://wiki.squeak.org/squeak/1934
[16] http://www.smalltalk.org
[17] http://www.cs.kent.ac.uk/people/staff/fkh/
[18] http://www.cs.kent.ac.uk/projects/vital/
[19] http://github.com/elbrujohalcon/wxhnotepad

editores de texto con *wxHaskell*. Gracias a Jeremy O'Donoghue[20], *wxhNotepad* está siendo publicado como un tutorial[21] en sucesivos artículos en su blog

1.3. λ*Page*

> Ancorché lo ingegno umano faccia invenzioni varie, rispondendo con vari strumenti a un medesimo fine, mai esso troverà invenzione più bella, né più facile né più brieve della natura, perché nelle sue invenzioni nulla manca e nulla è superfluo
>
> Leonardo da Vinci
>
> La programación intensiva y el uso prolongado de Tetris sólo lleva a ver estructuras de orden y secuencias en la verdulería y a querer apilar los autos para formar líneas sólidas
>
> Darío Ruellan

λ***Page***[22] se presenta como una herramienta similar al Workspace de *Smalltalk*, que permite a los desarrolladores trabajar con documentos de texto libre que incluyan expresiones y definiciones. λ***Page*** es capaz de identificar las expresiones y definiciones válidas y permite al desarrollador inspeccionarlas, evaluarlas, conocer su tipo y su clase.

En el espíritu de las herramientas provistas por la comunidad de desarrolladores *Haskell*, λ***Page*** se integra con Cabal[23] y Hayoo![24] y se encuentra ya disponible en HackageDB[25].

λ***Page*** presenta una interfaz simple e intuitiva, desarrollada utilizando wxHaskell[26], lo que lo convierte en un sistema multiplataforma.

Por ser una herramienta desarrollada con *Haskell* para *Haskell*, λ***Page*** se diferencia de sus pares del mundo de objetos, al aprovechar conceptos claves como son el tipado fuerte (que permite detectar errores de tipo velozmente, evitando el costo de evaluar expresiones complejas) y la evaluación perezosa (que permite evaluar expresiones infinitas e ir exhibiendo resultados progresivamente).

[20] http://wewantarock.wordpress.com/about/
[21] http://wewantarock.wordpress.com/2010/01/31/building-a-text-editor-part-1/
[22] http://haskell.hpage.com
[23] http://www.haskell.org/cabal
[24] http://holumbus.fh-wedel.de/hayoo
[25] http://hackage.haskell.org/package/hpage
[26] http://haskell.org/haskellwiki/WxHaskell

A diferencia de *GHCi* que es una herramienta "de consola", **λPage** permite visualizar resultados de manera más dinámica, permitiendo que errores intermedios, detectados durante la evaluación de una expresión no impidan continuar con la misma hasta llegar a un resultado más completo.

λPage se encuentra desarrollado utilizando *eprocess*[27], una librería que facilita el manejo de "threads" en un estilo similar al de los procesos *Erlang*. Gracias al uso de esta librería, **λPage** puede realizar tareas en paralelo y por lo tanto permitir al usuario continuar editando los documentos en los que está trabajando mientras espera que se evalúe una expresión e incluso cancelar una evaluación conservando la porción del resultado obtenida hasta ese momento. También gracias al uso de *eprocess*, **λPage** permite detectar cálculos infinitos (o más precisamente, cálculos que demoran demasiado) e informar sobre este hecho al usuario para que ya no siga esperando indefinidamente el resultado de la evaluación solicitada.

[27] http://hackage.haskell.org/package/eprocess

2. Tutorial - Descubriendo λ*Page*

2.1. Instalación

> As a rule, software systems do not work well until they have been used, and have failed repeatedly, in real applications.
>
> Dave Parnas

> The #1 programmer excuse for legitimately slacking off: "My code is compiling"
>
> David Knutz

Para instalar λ***Page*** en *OSX* o *Windows*, se proveen instaladores en el sitio web de λ***Page***, sin embargo,como se ha dicho, λ***Page*** se encuentra en *HackageDB* y por lo tanto el modo oficial de instalarlo es utilizando *Cabal*, con el siguiente comando:

```
$ cabal install hpage
```

Sin embargo, para ello, previamente se deben satisfacer las siguientes dependencias:

wxWidgets 2.8.10+[28] El framework de desarrollo para interfaces de usuario que utiliza *wxHaskell*. Debe ser instalado con los módulos unicode, cmdline, config, log, stl, richtext y clipboard, al menos y con el módulo odbc desactivado.

Haskell Platform[29] Una distribución de *Haskell* que incluye todo lo necesario para compilar e instalar programas desarrollados en este lenguaje (de particular interés para λ***Page***: GHC y happy).

2.1.1. Windows

Para el correcto funcionamiento de λ***Page*** los usuarios de *Windows XP* deben instalar el C++ 2008 SP1[30].

2.1.2. Linux

En algunas distribuciones de Linux es conveniente, además de la instalación de *Haskell Platform* instalar las librerías de *Monad Transformers* ejecutando, por ejemplo:

```
$ sudo aptitude install libghc6-mtl-dev libghc6-mtl-doc
```

[30] http://www.microsoft.com/downloads/details.aspx?familyid=A5C84275-3B97-4AB7-A40D-3802B2AF5FC2

2.2. QuickStart

> We learn by example and by direct experience because there are real limits to the adequacy of verbal instruction
>
> Malcom Gladwell

> How is education supposed to make me feel smarter? Besides, every time I learn something new, it pushes some old stuff out of my brain - remember when I took that home winemaking course, and I forgot how to drive?
>
> Homer Simpons

TODO: Tutorial donde se noten las features de λ***Page***

3. Desarrollo - ¿Cómo se hizo *λPage*?

3.1. Arquitectura General

> If you think good architecture is expensive, try bad architecture
>
> Brian Foote and Joseph Yoder

Las principales decisiones de arquitectura que se tomaron durante el desarrollo de ***λPage*** tuvieron como principales motivaciones los siguientes requerimientos:

Conexión con GHC ***λPage*** debía conectarse con el motor de GHC a través de su API de modo de poder detectar e interpretar expresiones. Para ello se utilizó hint[31]

Paralelismo ***λPage*** debía permitir al usuario editar sus documentos mientras esperaba el resultado de la evaluación de alguna expresión. Para ello se creó *eprocess* y se implementó un modelo de procesos utilizándolo.

Errores Controlados ***λPage*** no debía fallar si la evaluación de una expresión fallaba. Más aún, también debía detectar posibles evaluaciones infinitas e informar estas situaciones al usuario

Teniendo en cuenta estos requerimientos, la arquitectura resultante puede ser descripta con el diagrama de la figura 1. Esta figura presenta el estado del sistema en un instante dado. Cada bloque representan un proceso o "thread" en ejecución. Cada uno de estos procesos se ejecuta dentro del entorno de una mónada, la cual se encuentra identificada en la esquina superior derecha del bloque. En el diagrama podemos identificar los siguientes componentes:

UI Manager Este es el thread que inicia el programa, genera y administra la interfaz del usuario utilizando las herramientas provistas por *wxHaskell*. En este thread se mantiene el estado visual de la aplicación: el estado de los controles, la última búsqueda realizada, etc.

HPage Server Este proceso, iniciado por el **UI Manager**, es el que comunica a la interfaz del usuario con la máquina virtual de GHC, a través del **Hint Server**, captura sus errores y lo reinicia en caso de ser necesario. En este proceso se mantiene el estado general de la aplicación: sus páginas, expresiones, paquetes y módulos cargados, etc.

Hint Server Este proceso, iniciado por el **HPage Server**, mantiene una conexión con la máquina virtual de GHC (a la cual se muestra en la figura conectado a través de una línea de puntos)

Char Filler Este proceso, iniciado por el **UI Manager** cumple una muy sencilla función: utilizando los procedimientos de envío y recepción de mensajes provistos por *eprocess*, espera recibir un caracter (o sea, una expresión de tipo Char), para luego evaluarlo y enviar como respuesta su valor en forma normal.

[31] http://projects.haskell.org/hint

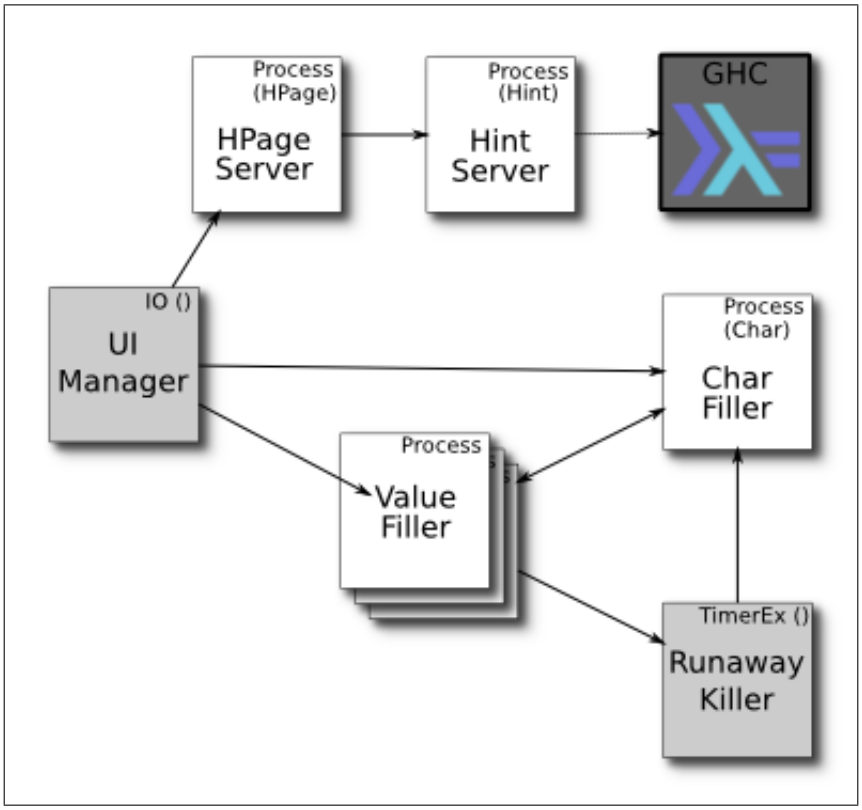

Figura 1: Arquitectura de λ***Page***

Value Filler Estos procesos, iniciados por el **UI Manager** ante cada evaluación solicitada por el usuario son los encargados de procesar el resultado obtenido del **HPage Server**. Cabe recordar aquí que *Haskell* trabaja con evaluación "lazy", por lo cual el resultado obtenido no ha sido aún completamente procesado. Cada **Value Filler** se encarga de evaluar un resultado y mostrarlo por pantalla, para ello envía y recibe mensajes del **Char Filler** a fin de procesar cada caracter a mostrar.

Runaway Killer Este thread, creado utilizando la clase *TimerEx* provista por *wxHaskell*, es iniciado por cada **Value Filler** al momento de enviar un nuevo caracter al **Char Filler**. El objetivo del **Runaway Killer** es el de detectar prorcesamiento "posiblemente" infinito. Básicamente, pasado un segundo de procesamiento, reinicia el **Char Filler** e informa al **Value Filler** que lo inició que el caracter que se esperaba procesar ha demorado demasiado y podría desencadenar una evaluación infinita.

Para un mayor detalle, la figura 2 nos muestra un diagrama de secuencia correspondiente a un proceso de evaluación. Para poder brindar un ejemplo completo, hemos "fabricado" un tipo que responde al siguiente código:

```
data WithIfiniteChar = WIC

instance Show WithIfiniteChar where
    show WIC = ['c', head . show $ length [1..]]
```

Como puede observarse, al intentar mostrar la expresión `WIC`, *λPage* se encontrará con una cadena cuyo segundo caracter no puede computar pues requiere un cálculo infinito, en la figura representamos este caracter con la letra Ω. Allí es donde entra en acción el **Runaway Killer** para informar esta situación al usuario.

3.2. Diseño

> Design and programming are human activities;
> forget that and all is lost
>
> — Bjarne Stroustrup

3.2.1. eprocess

TODO: Por qué y cómo hicimos eprocess (link a realworldhaskell)

3.2.2. Servers

TODO: Por qué y cómo hicimos hint-server y hpage-server

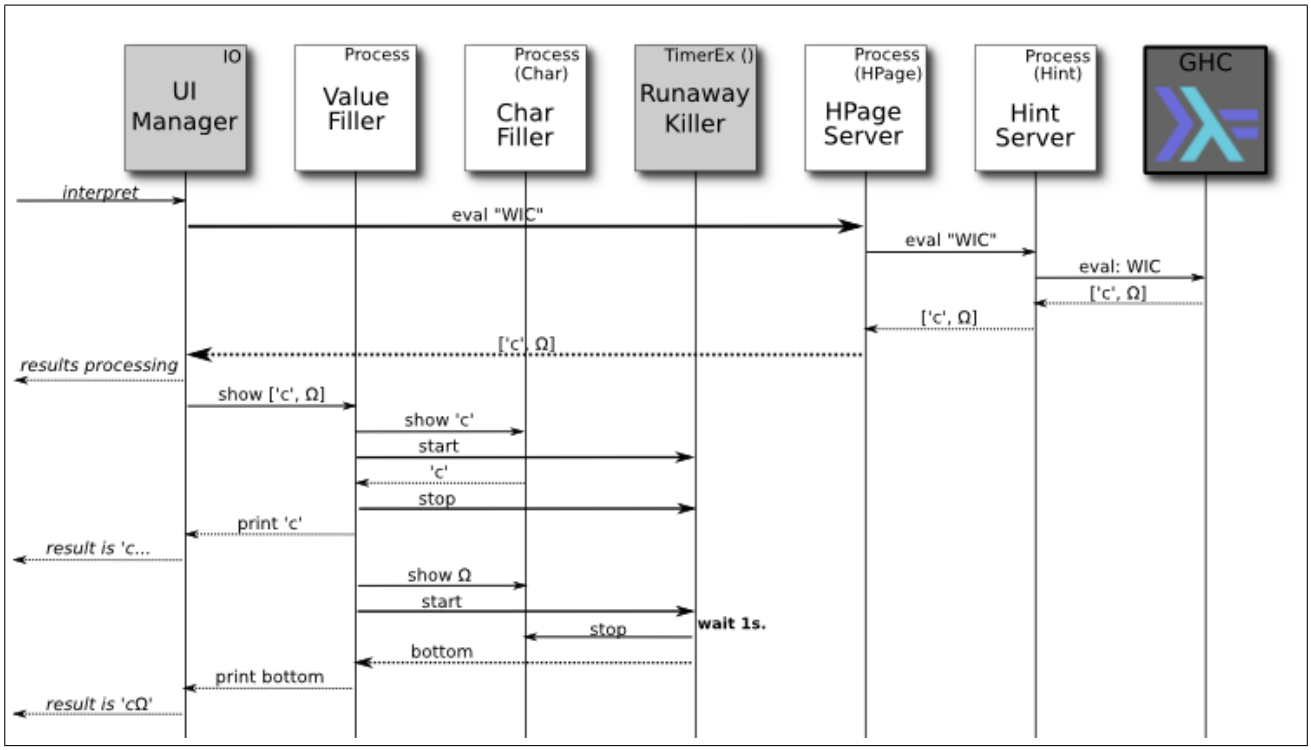

Figura 2: Secuencia de Evaluación de “WIC”

3.2.3. Bottoms

El lenguaje *Haskell* tiene una característica única: *la evaluación perezosa* o *lazy evaluation*. Gracias a esta característica, las expresiones *Haskell* no son completamente evaluadas (reducidas a forma normal) hasta el momento en que realmente se necesita conocer su valor. Dado que **λPage** presenta al usuario un interprete de expresiones, es necesario que esté preparado para soportar esta característica. En particular, dentro de **λPage** las expresiones son reducidas a forma normal al momento de intentar mostrar el resultado de su evaluación al usuario. En ese momento, lo único que se sabe es que la expresión a mostrar es de clase *Show*. **λPage** intenta entonces ejecutar la función `show` sobre la expresión a mostrar y pueden suceder varias cosas:

- Por supuesto, puede suceder que al evaluar la expresión se obtenga una cadena de caracteres, en cuyo caso **λPage** simplemente presentará el resultado al usuario
- Puede suceder que, intentando evaluar la expresión se obtenga una cadena de caracteres de longitud infinita. La política de **λPage** en este caso es permitir al usuario decidir cuándo desea abortar la evaluación y mostrar la porción del resultado obtenida hasta ese momento
- Puede suceder también que, intentando evaluar un caracter, se genere un cálculo "infinito". En este caso, **λPage** aborta el proceso que se encuentra intentando evaluar el caracter en cuestión e informa este hecho al usuario, para luego continuar evaluando los siguientes caracteres
- Es posible también que, al intentar evaluar un caracter, se produzca una excepción. En este caso, **λPage** atrapa la excepción, la informa al usuario y luego continúa evaluando los siguientes caracteres
- Otra posibilidad es que, luego de presentar un caracter, al intentar obtener el resto de la cadena, **λPage** encuentre un cálculo "infinito" o una excepción. En estos casos, se informa la situación al usuario y se aborta la evaluación de la expresión.

3.2.4. Cabal

TODO: Por qué y cómo nos integramos con cabal

3.2.5. Hayoo!

TODO: Por qué y cómo nos integramos con Hayoo!

3.3. Implementación

> Nothing resolves design issues like an implementation
>
> J. D. Horton
>
> A child of five would understand this. Send someone to fetch a child of five
>
> Groucho Marx

3.3.1. UI

La interfaz gráfica de **λPage** está desarrollada utilizando *wxHaskell*, un framework elegido por ser multi-plataforma y, gracias a estar construído sobre *wxWidgets*, presentar un "look&feel" nativo en distintos entornos. *wxHaskell* es un framework sencillo para utilizar y entender y, pese a que aún se encuentra en período de evolución, es suficientemente estable. Sin embargo, tal como puede verse en *wxhNotepad* hemos tenido que superar varios escollos hasta lograr una UI estable e intuitiva. A los lectores interesados en estos detalles técnicos les recomendamos la lectura de los artículos escritos por Jeremy O'Donoghue en su tutorial Building a text editor[32].

3.3.2. TODO: OTROS

TODO: Ver el código... detectar cosas interesantes y armarles secciones

4. Resultados

4.1. Objetivos Alcanzados

> Results! Why, man? I have gotten a lot of results. I know several thousand things that wont work
>
> Thomas A. Edison

> Kids, you tried your best and you failed miserably. The lesson is: "never try"
>
> Homer Simpsons

TODO: ¿Qué se puede hacer ahora que existe **λPage**?

4.2. Trabajo a Realizar

> Inside every large program, there is a small program trying to get out
>
> C.A.R. Hoare

> I'm a man with a one-track mind, so much to do in one life-time
>
> Queen

So much to do, so little done, such things to be TODO: Future Work

[32] http://wewantarock.wordpress.com/2010/01/31/building-a-text-editor-part-1/

5. Agradecimientos

TODO: Acknowledgments. Recordar a Jeremy, al flaco del Undo/Redo, a los de Hayoo!